INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Identification and characterization of mycoplasma promoters

Knudtson, Kevin Lee, Ph.D.

Iowa State University, 1993
Identification and characterization of mycoplasma promoters

by

Kevin Lee Knudtson

A Thesis Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Microbiology, Immunology and Preventive Medicine
Major: Immunobiology

Approved:
Signature was redacted for privacy.
In Charge of Major Work
Signature was redacted for privacy.
For the Major Program
Signature was redacted for privacy.
For the Major Department
Signature was redacted for privacy.
For the Graduate College

Members of the Committee
Signature was redacted for privacy.

Iowa State University
Ames, Iowa
1993
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>vii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ix</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>LITERATURE REVIEW</td>
<td>4</td>
</tr>
<tr>
<td>General Background and Classification of Mycoplasmas</td>
<td>4</td>
</tr>
<tr>
<td>Gene Transfer in Mycoplasmas</td>
<td>13</td>
</tr>
<tr>
<td>Development of Cloning Systems for Mycoplasmas</td>
<td>17</td>
</tr>
<tr>
<td>Regulation of Gene Expression in Mycoplasmas</td>
<td>22</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>34</td>
</tr>
<tr>
<td>Bacterial Strains and Culture Media</td>
<td>34</td>
</tr>
<tr>
<td>Reagents and Buffers</td>
<td>34</td>
</tr>
<tr>
<td>Transformation of <em>Acholeplasma</em> spp., <em>M. gallisepticum</em> and <em>E. coli</em></td>
<td>39</td>
</tr>
<tr>
<td>Preparation of <em>Acholeplasma</em> ISM1499 Chromosomal DNA</td>
<td>40</td>
</tr>
<tr>
<td>Plasmid Constructions</td>
<td>41</td>
</tr>
<tr>
<td>Preparation of RNA and Measurement of mRNA Levels</td>
<td>43</td>
</tr>
<tr>
<td>β-galactosidase Assays</td>
<td>43</td>
</tr>
<tr>
<td>Immunoblot Analysis</td>
<td>44</td>
</tr>
<tr>
<td>DNA Sequencing</td>
<td>45</td>
</tr>
<tr>
<td>Primer Extension Analysis</td>
<td>45</td>
</tr>
<tr>
<td>Exo/Mung Deletion Analysis</td>
<td>46</td>
</tr>
<tr>
<td>RESULTS</td>
<td>47</td>
</tr>
<tr>
<td><em>M. capricolum</em> Promoter trp'-lacZYA Fusion Studies</td>
<td>47</td>
</tr>
<tr>
<td>Tn4001lac Promoter Probe Studies</td>
<td>50</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Construction, message levels, and β-galactosidase activities of the
trp'-lacZYA fusion plasmids

Figure 2. Placement of unique BamHI and SmaI RE sites in an IS256 arm
of transposon Tn4001

Figure 3. Removal of the translational stop codons from the inverted repeat of an
IS256 element and construction of Tn4001lac

Figure 4. Construction of promoter probe transposons designed to rescue adjacent
chromosomal sequences

Figure 5. Construction of the promoter probe vector, pISM2050

Figure 6. Acholeplasma ISM1520 transformants

Figure 7. Quantitation of lacZ transcript levels of the lacZ fusion vectors in
Acholeplasma ISM1520

Figure 8. Immunoblot analysis of the lacZ fusions in Acholeplasma

Figure 9. Primer extension analysis of the transcriptional site of the
lacZ fusion transcript of ISM2050.8

Figure 10. Exo/Mung deletion analysis of the ISM1499 chromosomal
DNA region of pISM2050.2

Figure A1. Sequence upstream of the lacZ gene in pISM2050.1

Figure A2. Sequence upstream of the lacZ gene in pISM2050.2

Figure A3. Sequence upstream of the lacZ gene in pISM2050.8

Figure A4. Sequence upstream of the lacZ gene in pISM2050.18

Figure A5. Sequence upstream of the lacZ gene in pISM2050.19

Figure A6. Sequence upstream of the lacZ gene in pISM2050.25

Figure A7. Sequence upstream of the lacZ gene in pISM2050.39

Figure A8. Sequence upstream of the lacZ gene in pISM2050.40

Figure A9. Sequence upstream of the lacZ gene in pISM2050.66

Figure A10. Sequence upstream of the lacZ gene in pISM2050.69
Figure A11. Sequence upstream of the lacZ gene in pISM2050.70
Figure A12. Sequence upstream of the lacZ gene in pISM2050.86
Figure A13. Sequence upstream of the lacZ gene in pISM2050.89
LIST OF TABLES

Table 1. Bacterial strains and plasmids 35
Table 2. Transformation frequencies of Tn4001 derivatives in Acholeplasma 54
Table 3. Insert size and β-gal activity of lacZ fusion constructs in Acholeplasma and E. coli 64
Table 4. Alignment and determination of the putative -10 and -35 promoter regions of the promoters driving the expression of lacZ in the plasmid pISM2050 derivatives 75
Table B1. Experimental design of the Acholeplasma challenge study 132
Table B2. Absorbance values of an ELISA examining the antibody response to recombinant Acholeplasmas expressing the E. coli lacZ 134
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>adenosine</td>
</tr>
<tr>
<td>AIDS</td>
<td>acquired immunodeficiency syndrome</td>
</tr>
<tr>
<td>Ap</td>
<td>ampicillin</td>
</tr>
<tr>
<td>AMV</td>
<td>avian myeloblastosis virus</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>β-gal</td>
<td>β-galactosidase</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>CAP</td>
<td>catabolite activator protein</td>
</tr>
<tr>
<td>CFU</td>
<td>colony forming unit</td>
</tr>
<tr>
<td>cpm</td>
<td>counts per minute</td>
</tr>
<tr>
<td>CRP</td>
<td>cyclic AMP receptor protein</td>
</tr>
<tr>
<td>cAMP</td>
<td>cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>cat</td>
<td>chloramphenicol acetyl transferase</td>
</tr>
<tr>
<td>Da</td>
<td>daltons</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetraacetic acid</td>
</tr>
<tr>
<td>G</td>
<td>guanine</td>
</tr>
<tr>
<td>Gm</td>
<td>gentamicin</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropylthio-β-galactoside</td>
</tr>
<tr>
<td>IS</td>
<td>insertion sequence</td>
</tr>
<tr>
<td>Km</td>
<td>kanamycin</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pair</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodaltons</td>
</tr>
<tr>
<td>MCS</td>
<td>multiple cloning site</td>
</tr>
<tr>
<td>-mer</td>
<td>used to denote an oligonucleotide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>MLO</td>
<td>mycoplasma like organism</td>
</tr>
<tr>
<td>NET</td>
<td>Sodium chloride-EDTA-Tris buffer</td>
</tr>
<tr>
<td>p</td>
<td>plasmid</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PPLO</td>
<td>pleuropneumonia like organism</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>rbs</td>
<td>ribosomal binding site</td>
</tr>
<tr>
<td>RE</td>
<td>restriction enzyme</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>T</td>
<td>thymine</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>TE</td>
<td>Tris [hydroxymethyl aminomethane]-EDTA [ethylene diamine tetraacetic acid] buffer</td>
</tr>
<tr>
<td>Th</td>
<td>transposon</td>
</tr>
<tr>
<td>U</td>
<td>units</td>
</tr>
<tr>
<td>X-gal</td>
<td>5-bromo-4-chloro-3-indolyl-β-D galactoside</td>
</tr>
</tbody>
</table>
I would like to thank Drs. A. Myers, P. A. Pattee, P. S. Paul, and M. J. Wannemuehler for serving on my examining committee. I appreciate their helpful suggestions and encouragement. I would like to extend a very special thanks to Dr. F. Chris Minion. His patience, suggestions, and encouragement have made it possible for me reach this level in my career. Our scientific discussions were stimulating as well as enlightening. I consider him to be more than my major advisor, I consider him to be a good friend.

It is impossible to spend as much time in one area as I have and not make some very good friends. I would like to thank Greg Mahairas, Laura Hruska, Tim Freier, Paul Kapke, Eli Tigges, Jian Cao, Brenda Smiley, Karalee Taylor, Debra Moore, Anna Yeung, and Tina VanDyk. Their support inside the lab as well as outside the lab setting were equally as important in keeping me going.

I would like extend my deepest thanks and love to the most important person I have met here and in my life, my wife Linda Knudtson. Her love and support have kept me going and have played an important role in making this possible.

Finally, I would like to extend a very special thanks to my parents, Vernon and Suzanne Knudtson, for patiently waiting and offering support when needed.
INTRODUCTION

The term "mycoplasmas" is used as the trivial name for all members of the class Mollicutes. The class Mollicutes is composed of six genera which include: Mycoplasma, Ureaplasma, Spiroplasma, Acholeplasma, Anaeroplasma, and Asteroleplasma (186). The members of this class are cell wall-less eubacteria that possess the smallest genome size thought necessary for autonomous existence (184). As a result of their small genome size, mycoplasmas have limited biosynthetic pathways. Generally, the macromolecules required for growth must be scavenged from the environment because mycoplasmas possess only catabolic and conversion pathways to obtain amino acids and nucleotides (reviewed in 62, 63). Thus, complex undefined media are required for their growth. Despite an apparent fragile existence, mycoplasmas have been shown to colonize humans, animals, plants, and insects (reviewed in 185).

Mycoplasmas have become recognized as important human and agricultural pathogens. They are generally mucosal pathogens in animals, colonizing the urogenital and respiratory tracts (222). The human pathogen Mycoplasma pneumoniae, which causes tracheobronchitis and primary atypical pneumonia, is the leading cause of pneumonia in older children and young adults (111). Ureaplasma urealyticum is a cause of nongonococcal urethritis in men and is often associated with pelvic inflammatory disease in women (111). Recently, a strain of Mycoplasma fermentans, Mycoplasma incognitus, has been associated with AIDS (reviewed in 125). In addition, mycoplasmas have had a great impact on both the agronomy and livestock industries (reviewed in 120, 222).

Mycoplasmas are thought to be a product of degenerative evolution from Gram positive bacteria (135). Like many Gram positive bacterial species, mycoplasmas are A + T-rich (60 - 77 %) organisms (186). Because of the phylogenetic relationship between mycoplasmas and Gram positive bacteria and their lack of a cell wall, they have been compared to Gram positive
bacterial protoplasts. Consequently, the basis for the development of gene transfer methods and tools (i.e., transposons and plasmids) in mycoplasmas has been derived from Gram positive bacterial systems. Gene transfer via conjugation, transfection, and transformation have been demonstrated in mycoplasmas (reviewed in 51, 52).

The use of *Escherichia coli* cloning hosts has generated misleading results in studies of the regulation of both mycoplasma transcription and translation. Gafny et al. showed that both *E. coli* and *Mycoplasma capricolum* RNA polymerase recognized the *M. capricolum* rRNA gene promoter, but during amino acid starvation, transcription was activated in *E. coli* instead of repressed (67). Notarnicola et al. have shown that *E. coli* initiated translation at internal sites in a *Mycoplasma hyorhinis* lipoprotein gene (170). In addition, most mycoplasmas, except for *Acholeplasma*, have an alternative codon usage in which the UGA codon encodes tryptophan (94, 265). The alternative codon usage results in the expression of a truncated mycoplasma gene product that may or may not be detectable (101). Therefore, it is important to be able to examine mycoplasma genes and their regulatory elements in their normal genetic background, the mycoplasma.

The lack of genetic tools has made the development of mycoplasma cloning systems difficult. Only two transposons, Tn916 and Tn4001, have been shown to be useful for studying mycoplasma genetics (53, 54, 133). Broad host-range Gram positive plasmids have been examined as possible cloning vectors, but they have proven to be unstable. Naturally occurring mycoplasma plasmids have been examined as possible cloning vectors, but they have not been shown to maintain and express a cloned gene (reviewed in 52). A cloning system has been developed in spiroplasmas that uses a spiroplasma virus as a cloning vector, but, the virus has a limited host range (231, 232). Mahairas and Minion have developed a cloning system based on the ability to integrate cloned genes into mycoplasma chromosomes via homologous recombination (132, 134). The stability and versatility of the integrated plasmid cloning system
make it possible to incorporate any DNA sequence into the host including cloning vectors designed to study gene regulation.

Therefore, the goals of this study were to expand the mycoplasma cloning system developed by Mahairas and Minion to include a reporter gene for the detection of mycoplasma upstream gene regulatory elements. The approach involved the development of both plasmid- and transposon-based reporter gene vectors. Once mycoplasma promoter sequences were identified, the sequences were analyzed, and a consensus promoter sequence was be derived.
LITERATURE REVIEW

General Background and Classification of Mycoplasmas

History of mycoplasmas

Mycoplasmas were originally isolated in 1898 by Nocard and Roux (169) as an agent of bovine pleuropneumonia. As a result, during the first half of this century, similar organisms were known as "pleuropneumonia-like organisms" or "PPLOs" (85). Despite nearly one hundred years of study with these organisms, their growth requirements and genetics remain an enigma. In addition, their taxonomic classification has been one of the most confusing and controversial battles encountered for any bacterial group.

Prior to the 1930's, mycoplasmas were considered to be viruses because they could pass through filters that blocked the passage of other bacteria. As the true nature of viruses became apparent, however, it was obvious that mycoplasmas could not be considered to be viruses. Interestingly, the mycoplasma species recently identified in AIDS patients was originally thought to be a novel virus (126, 127).

The most difficult battle for independent taxonomic status for mycoplasmas was the distinction between L-forms of other bacteria and mycoplasmas. Following the observation that bacteria could be induced to grow as stable L-forms (118), the notion that mycoplasmas were merely L-forms of other bacteria predominated. The isolation of walled bacteria in mycoplasma cultures enhanced this notion (181). Following nearly 30 years of heated debate, the controversy was resolved based on genomic analysis of G + C content (181). The A + T-rich mycoplasmas could be definitively differentiated from the G + C-rich bacterial contaminants.

Just as the establishment of taxonomic status for mycoplasmas was highly controversial, the appropriate placement of mycoplasmas in the bacterial phylogenetic tree also was a subject for heated debate. One model proposed by Neimark suggested that the different mycoplasmas
arose from different branches of the bacterial phylogenetic tree through degenerative evolution (167, 168). For example, the model predicted that the genus *Mycoplasma* arose from one species of bacteria while the genus *Acholeplasma* arose from another. The other model, proposed by Morowitz and Wallace (155, 255), predicted that mycoplasmas were the first bacteria. Neither of these models, however, are accepted today. The present model suggests that mycoplasmas arose through degenerative evolution from Gram positive bacteria (probably *Lactobacillus* spp.) and that the entire class of *Mollicutes* have a common origin (135, 136).

This model was supported by 16S rRNA gene sequence homology studies (257). The *Mollicutes* phylogenetic order, from earliest to latest branch, was predicted to be *Asteroleplasma, Anaeroplasma, Acholeplasma, Spiroplasma, Mycoplasma, and Ureaplasma* (257). Thus, the association of mycoplasmas with L-forms is not entirely unjustified because they are thought to be a product of degenerative evolution from Gram positive bacteria (135).

Dujardin-Beaumetz was the first to cultivate a mycoplasma on solid medium (cited in ref. 85). He described the distinctive "fried-egg" colony morphology that has become the trademark of mycoplasmas. The fried-egg colony appearance is achieved by the production of a dark center that penetrates into the medium and is surrounded by a thinning periphery.

The first isolation of a mycoplasma from a human source was reported in 1937 by Dienes and Edsall (44). It was not until 1962, however, that etiological proof for the involvement of a mycoplasma in human disease was provided (26). This organism is presently known as *Mycoplasma pneumoniae*, the etiologic agent of primary atypical pneumonia. In addition to being human pathogens, mycoplasmas have become known as important agricultural pathogens, severely affecting the agronomy and livestock industries.

In 1936, another mycoplasma was isolated from sewage and was thought to be a saprophyte (117). A similar organism was later isolated from the bovine genital tract (57). These two organisms, unlike the pleuropneumonia causing agents, did not require sterols for
growth. Because they lacked a cell wall, they were later assigned a separate genus within the class Mollicutes called Acholeplasma (59).

Organisms which produced tiny colonies (7 to 15 \( \mu \text{m} \) in diameter) were first isolated from men with primary and recurrent nongonococcal urethritis by Shepard in 1954 (218). As a result of their tiny colony morphology, they were originally referred to as T-strains and were later named Ureaplasma urealyticum (219). To definitively prove that Ureaplasma urealyticum was an etiological agent in nongonococcal urethritis, Taylor-Robinson infected himself (244).

Spiroplasmas were first observed in 1961 when they were mistaken for spirochetes (177). Although they were observed throughout the decade as infectious agents of insects and plants, they were not defined as a Mollicute until 1973 (211, 259).

The first strictly anaerobic mycoplasma was reported in 1966 by Hungate (91). This organism was characterized by Robinson (199) and was originally assigned the name Acholeplasma bactoclasticum. Because this organism required sterols for its growth, it was assigned a separate genus for strictly anaerobic mycoplasmas, Anaeroplasma (200, 201).

Taxonomy of mycoplasmas

The current classification system of mycoplasmas is an expansion of the classification system originally presented by Edward and Freundt in 1956 (58). The name "mycoplasma" was derived from the Greek word mykes meaning fungus and plasma for something molded or formed to describe the fungus-like morphology of M. mycoides (185). Mollicutes is presently the only class in the division Tenericutes which is composed of wall-less bacteria (186). The name "mycoplasmas" is loosely used to describe any organism within the class Mollicutes despite the introduction of 5 new genera since mycoplasmas were first classified.

The class Mollicutes consists of 3 orders, Mycoplasmatales, Acholeplasmatales, and Aneroplasmatales. Because this work did not involve species from the order Aneroplasmatales, this discussion will focus primarily on the other two orders. For the sake of completeness,
there is another group of organisms within the class *Mollicutes*, the mycoplasma-like organisms (MLOs), that have tentatively been placed there because they lack a cell wall. Because they have not been cultivated, their metabolic or other distinctive properties have not been defined (185).

*Mycoplasmatales.* The order *Mycoplasmatales* is composed of facultative anaerobic sterol-requiring mollicutes that may or may not ferment glucose and hydrolyze arginine. The order has two families, *Mycoplasmataceae* and *Spiroplasmataceae* (186).

The family *Mycoplasmataceae* is composed of two genera, *Mycoplasma* and *Ureaplasma* (186). Ninety-two of the 125 currently recognized species are members of the genus *Mycoplasma* (185). Members of this genus have a wide habitat including humans, animals, plants, and insects. The other genus within this family, *Ureaplasma*, possesses species that are pathogens of humans and animals, and are characterized by having the ability to hydrolyze urea to CO₂ and NH₃ (220).

The family *Spiroplasmataceae* contains one genus, *Spiroplasma*. Arthropods and plants are the primary habitat for the eleven species within this genus. These organisms are characterized by helical filaments and their genome sizes are among the largest (1,350 to 2,575 kb) of the *Mollicutes* (14, 121).

*Acholeplasmatales.* The order *Acholeplasmatales* is composed of mollicutes that are facultative anaerobes that do not require sterols for growth (186). The members of this order are the only mollicutes that are able to modulate membrane fluidity by carotenoid synthesis, selective incorporation and elongation of fatty acids, and *de novo* synthesis of fatty acids (183). In addition, these organisms can metabolize glucose, but they cannot hydrolyze arginine or urea (250). This order has one family, *Acholeplasmataceae*, which has one genus, *Acholeplasma* (186). Members of this genus are primarily found in animals, plants and insects (251).
Growth requirements of mycoplasmas

The Mollicutes are among the most fastidious groups of bacteria, and their close association with their plant or animal host have led to dependence on their host for many of their macromolecule precursors (144). The ability to determine the specific nutrient requirements for a microorganism usually requires the use of defined growth media. Unfortunately, fully defined or semi-defined growth media have only been developed for a few Mollicute species including *M. mycoides*, *M. gallisepticum*, *M. capricolum*, *A. laidlawii*, and Spiroplasma spp. (24, 74, 79, 119, 202, 204, 246). In addition, the ability to define the organic growth requirements is confounded by the interaction between metabolic pathways (144). For example, alanine was essential for growth of *M. mycoides* only in the absence of pyridoxal (202), and tyrosine and phenylalanine were not required by *A. laidlawii* when alanine was added to the medium (246). Also, nutritional antagonism has been demonstrated between alanine and glycine in *M. mycoides* (202). This type of antagonism is common among eubacteria (e.g., abundant glucose levels inhibiting the utilization of other sugars in the medium). Thus, the precise nutritional requirements for many of the mycoplasmas are not known at this time because of the lack of well-defined growth media.

Generally, the organic elements required for mycoplasma growth must be scavenged from the environment. Mycoplasmas must acquire many of their amino acids because they lack the pathways to synthesize them. For example, *M. arthritidis* was shown to catabolize arginine, glutamine, glutamic acid, aspartic acid, histidine, leucine, and threonine under aerobic conditions and tyrosine and tryptophan under anaerobic conditions (226). Fisher et al. has compiled a list of amino acid requirements of organisms known to grow on a defined or semi-defined media (63).

Mycoplasmas require nucleic acid precursors for growth because they lack the de novo pathways for their synthesis (150). Mycoplasmas are capable, however, of converting one
nucleotide to another. For example, *M. mycoides* subsp. *mycoides* was reported to be able to convert guanine to adenine (150, 166, reviewed in 62).

Sterols and fatty acids are essential for growth for all mollicutes except *Acholeplasma* and *Asteroleplasma* species. However, sterols were shown to be incorporated into the membrane by *A. laidlawii* if provided in the medium (12). Sterols account for 20% of the mass of the total membrane lipids in mycoplasmas. Interestingly, a cholesterol-rich growth medium was shown to reduce the growth rate of *M. capricolum* (122). Fatty acids were reported to be required for phospholipid synthesis and their uptake by *M. capricolum* required ATP and a transport protein (35). *M. mycoides* apparently lacked the ability to alter the chain length of saturated or unsaturated fatty acids, thus both fatty acids were required in a defined medium (204). Ureaplasmas were shown to synthesize both saturated and unsaturated fatty acids (207). Exogenous phospholipids such as sphingomyelin and phosphatidylcholine could be incorporated directly into the membranes of mycoplasmas and spiroplasmas (39, 209). For further background of lipid incorporation, biosynthesis, and metabolism see the reviews of McElhaney (142) and Smith (227).

Glucose was originally used as an energy source in the preparation of defined media (202). A study examining ten different mycoplasmas and acholeplasmas suggested that mollicutes do not use the TCA pathway, but rather they appear to ferment glucose via the Embden-Meyerhof-Parnas pathway (31, 43, 139). Glucose can be replaced by other sugars or pyruvate, but growth rates may be affected (145, 146). Studies suggested that all glucose-fermenting mycoplasmas also ferment maltose, starch, and glycogen (186, 205). Arginine hydrolysis via the arginine dihydrolase pathway has been shown to be utilized by some mycoplasmas as an energy source (61). There are no reports of mycoplasmas using lactose as an energy source. Thus, it does not appear that mycoplasmas possess the ability to breakdown lactose to glucose and galactose. This implies that *lacZ* could be an effective reporter gene in
mycoplasmas. For a review of carbohydrate metabolism and energy utilization in mollicutes see Pollack (176).

Mycoplasmas appear to require a large number of vitamins for their growth (202, 204, 205). The vast vitamin requirement was based on the large number of vitamins and coenzymes (6 to 8) used in preparation of the defined media for *M. mycoides* Y and *A. laidlawii* B (204). Studies performed with spiroplasmas showed that only nicotinic acid or riboflavin were required for growth (22). The best growth rates were achieved when both vitamins were present, but high concentrations of the vitamins appeared to reduce the growth rate (23). Thus, the vitamin requirements for *Mollicute* growth has not been precisely determined. It is apparent that macromolecules must be scavenged from the environment because they do not possess complete anabolic pathways to synthesize necessary macromolecules. Moreover, the metabolic pathways they do possess appear to be catabolic or are involved in interconversion (182).

Little work has been done on the inorganic requirements of the *Mollicutes*. The preparation of the defined media and semi-defined media for *M. mycoides* and *A. laidlawii* implied the requirement for PO$_4^{3-}$, Mg$^{2+}$, K$^+$, and Fe$^{2+}$(204). The author suggested that other elements, namely metals, would be required but adequate amounts could be obtained as contaminants on glassware (204). Tryon and Baseman have shown that *M. pneumoniae* is able to bind human lactoferrin as a potential iron acquisition mechanism for mycoplasmas (249). It should be noted that high concentrations of metal ions (Zn$^{2+}$, Mn$^{2+}$, Co$^{2+}$, and Fe$^{2+}$) inhibited mycoplasmal growth (203).

The temperature range for mycoplasma growth generally is from 7$^\circ$ to 40$^\circ$C (144). The optimal temperature is usually indicative of the natural habitat in which the mycoplasma is found. Mycoplasmas will grow and survive over a narrow pH range (144). The optimal pH for most species of vertebrate origin is 7.4 with a range of 6.5 to 8.0 (205). The optimal temperature and atmospheric conditions for mycoplasma growth are discussed by Gardella and Del Giudice (69).
Genomic composition of mycoplasmas

The genomes of mycoplasmas are small and have a low G + C content as compared to other bacteria (184). For nearly twenty years, the genome size served as one of the major taxonomic criterion for distinguishing the genera of the Class Mollicutes. Recent data from analysis of over 40 different mollicutes shows that the genomic sizes of the Mycoplasma and Ureaplasma species range from 500- to 1,300-kb, while the Spiroplasma and Acholeplasma species range from 1,350- to 2,600-kb (5, 8, 32, 115, 121, 124, 179, 198, 236, 258). Therefore, some of the genome sizes of Mycoplasmas and Ureaplasmases are similar to that of Acholeplasma and Spiroplasma. This makes distinguishing these species by genomic size a less reliable tool for taxonomic classification.

The size of the genome is an indication of the number of genes it encodes. A genome of 5 x 10^8 Da (approx. 750-kb) has the potential to code for approximately 650 proteins of an average molecular mass of 4 x 10^4 Da if no intergenic spacer regions are present (153, 154). Twenty to thirty percent of the bacterial genome is usually composed of spacer regions. Thus, it would be predicted that a mycoplasma genome would encode about 500 proteins (100). Two-dimensional gel analysis of the M. capricolum genome revealed about 350 spots suggesting that this mycoplasma encodes at least 350 proteins (100). As a comparison, Escherichia coli is known to encode for more than 1,100 proteins with a genome size of 4,000-kb (123).

The mycoplasma genomes examined to date have one or two sets of rRNA operons encoding the 23S, 16S, and 5S rRNAs (1, 73, 215, 216). This contrasts to the 10 rRNA operons in each of their closest relatives Clostridium spp. (70) and Bacillus spp. (98) and the 7 in E. coli (105). The M. capricolum genome possesses 30 genes that encode 29 tRNA species (14, 161), whereas E. coli possesses 78 genes for 45 tRNA species (109) and B. subtilis possesses at least 51 genes for 31 tRNA species (254). Therefore, it appears that the number of mycoplasma tRNA and rRNA species is considerably lower than in other eubacteria.
The G + C content for mycoplasmas range from 25 to 40 mol% (186). The G + C content of *M. capricolum* spacer regions are about 20 mol%, the protein genes are about 30 mol%, and the rRNA and tRNA genes are 46 to 54 mol% (163). The G + C content of the *M. capricolum* rRNA and tRNA is lower than the corresponding regions in other eubacteria (163). The low G + C content of mycoplasmas is hypothesized to be a product of A·T-biased directional mutation pressure that was exerted on the entire genome (163, 238). As the mycoplasma evolved, selective pressure must have occurred on the entire genome to replace the G-C pairs with A-T pairs. Muto et al. (163) suggest that the spacers are functionally the least important regions of the genome, and mutations in those regions are selectively neutral resulting in a mutation rate higher than any other regions. The A·T shift in the rRNA and tRNA genes is least pronounced because most of the transcripts serve a biological function. Protein genes are more variable than the RNA genes because of the redundancy of the genetic code.

Based on two-dimensional protein analysis of the *M. capricolum* genome and the relatively low number of rRNA and tRNA genes, it appears as though the *M. capricolum* genome codes for about 400 genes (184). Morowitz (153) suggested that mycoplasmas have a limited number of genes that are required for growth and has discarded many unnecessary genes as it evolved.

The basic elements for DNA replication (64), transcription of DNA to RNA (66), and translation of RNA into proteins (257) have been conserved in mycoplasmas. In addition, a number of DNA-related systems have also been maintained including DNA recombination and repair (56, 114, 131, 134, 261), DNA restriction/modification (137), and heat shock proteins (37), implying the necessity of these functions. It has not been determined whether proof-reading capability during DNA replication has been retained in mycoplasmas (136). Thus, it appears as though mycoplasmas possess the minimal genetic information necessary for life.
Gene Transfer in Mycoplasmas

The ability to genetically manipulate mycoplasmas through gene transfer is fundamental if an understanding of the genetic regulation of mycoplasmas is to be obtained. Gene transfer via conjugation, transfection, and transformation have been demonstrated in mycoplasmas (for reviews see 51, 52). Transduction has not yet been described for mycoplasmas. Because of the phylogenetic relationship between mycoplasmas and Gram positive bacteria and their lack of a cell wall, they have been compared to Gram positive bacterial protoplasts. Thus, many of the methodologies and tools (i.e., transposons and plasmids) developed to transfer genes in mycoplasmas were originally derived from Gram positive bacterial systems.

Mycoplasma conjugation

The acquisition of the tetM gene in clinical isolates of M. hominis and U. urealyticum was thought to occur through a conjugation event involving Enterococcus faecalis as a donor in vivo (194, 195, 197). Further investigations suggested that the mechanism of conjugal transfer might involve the Gram positive transposon Tn9/6 based on several observations. First, cell to cell contact was required for DNA transfer between Enterococcus faecalis and Mycoplasma hominis (196). Second, it was shown that the Gram positive conjugative transposon, Tn9/6, was transferred from E. faecalis to M. hominis and M. pulmonis at a frequency of $10^{-6}$ to $10^{-7}$ transconjugants per recipient colony forming unit (CFU) (50, 196). Finally, tetM is the antibiotic resistance marker found on Tn9/6 (71). Conjugal transfer of Tn9/6 using mycoplasmas as a donor has not yet been reported, however. Naglich and Andrews described comobilization of plasmids pC194 and pUB110 in the conjugal transfer of Tn9/6 from Bacillus subtilis to Bacillus thuringiensis subsp. israelensis (165). Thus, conjugative transposons may provide a means of introducing various antibiotic resistances or other genetic markers into mycoplasmas.
Another phenomenon of gene transfer in \textit{Mollicutes} requiring cell-to-cell contact has been demonstrated in \textit{Acholeplasma} and \textit{Spiroplasma} (6, 7, 131). Gene transfer between \textit{Acholeplasma} was demonstrated using cells that possessed either the gentamicin resistance marker from Tn\textit{4001} or the tetracycline resistance marker from Tn\textit{916}. Broth or agar surface mating mixtures yielded colonies possessing both of the antibiotic resistance markers. Moreover, the antibiotic resistance markers were located at the same place on the chromosome in the transconjugants as in the parents (131). The mating phenomenon was DNase I insensitive (6, 7, 131) ruling out a role for natural transformation. The matings occurred in the absence of polyethylene glycol (PEG) which has been shown to be effective in promoting the fusion of \textit{Bacillus subtilis} protoplasts (25) as well as in the transformation and transfection of mycoplasmas (54, 224). Triparental matings in mycoplasmas have not been observed (6).

The mechanism of DNA transfer in the \textit{Acholeplasma} and \textit{Spiroplasma} mating systems is not clear. The mechanism resembles conjugation in that the transfer was not DNase-sensitive and required cell-to-cell contact similar to the clumping observed in pheromone-induced enterococcal mating (47, 48). On the other hand, gene transfer resembled a protoplast fusion event because transposon markers in the progeny were in the same chromosomal locations as the parental strains and prolonged mating in non-selective media enhanced the frequency of colonies possessing both markers. Hopwood reported a similar phenomenon with \textit{Bacillus} protoplast fusions in which single colonies harvested from non-selective regeneration media gave rise to multiple phenotypes (89). The generation of multiple phenotypes was suggested to be a result of interruption of recombination at different chromosomal locations during fusion. Thus, the long incubation time in non-selective media for the spiroplasma matings might allow for complete acquisition of both markers.

Mahairas and Minion reported that mating did not occur with \textit{M. pulmonis}, \textit{M. gallisepticum}, and the \textit{A. laidlawii} strain ATCC 14192, and that mating appeared to require the presence of a trypsin-sensitive membrane protein (131). Thus, it appears that this mating
phenomenon only occurs in a few members of the *Mollicutes, S. citri* and an *Acholeplasma* species. Because it is advantageous for an organism to be able to acquire useful genes from the environment, the limited evidence presented here might suggest that these two strains continue to possess a DNA acquisition mechanism that other mollicutes with smaller genomes have lost. Evidence to support this notion comes from the observation that low-frequency transformation of spiroplasmas occurs in the absence of PEG (33, 212) suggesting the possible existence of a natural transformation system in some mycoplasma species.

**Transformation and transfection of mycoplasmas**

Initial attempts to transform mycoplasmas was based on the treatment of the cells with divalent cations (21, 65) similar to the methods used in preparing *E. coli* competent cells (82, 83). These attempts were not repeatable by another laboratory, however (196). The ability to effectively and repeatably transform mycoplasmas was based on the observation that PEG promoted the fusion of Gram positive protoplasts (25, 89). Because mycoplasmas resemble protoplasts, Sladek and Maniloff first utilized PEG-mediated transformation of *Acholeplasma* in transfection studies using a mycoplasma virus L2 (225).

Transformation protocols have since been developed for introducing DNA into a number of *Acholeplasma, Mycoplasma,* and *Spiroplasma* species (reviewed in 52). The DNA introduced into mollicutes has been in the form of a virus, transposon, or plasmid with a frequency ranging from $10^{-4}$ to $10^{-8}$ transformants per colony forming unit. The DNA size also affected transformation frequency as Mahairas and Minion showed an inverse relationship with plasmid size (134). A similar observation was reported with the 39-kb dsDNA mycoplasma virus L3. While attempts to transfect *A. laidlawii* with L3 using PEG were not successful, electroporation was effective (130). Thus, the inefficient transformation frequencies and plasmid size have become important considerations when conducting genetic studies in mycoplasmas.
The PEG concentration, stage of growth of the cell, and amount of DNA must be considered when introducing DNA into mycoplasmas. King and Dybvig have shown that transformation frequencies of *Mycoplasma mycoides* subsp. *mycoides* increased with higher PEG concentrations (53 - 62% w/v) (103). On the other hand, Mahairas and Minion reported that transformation of *Acholeplasma* ISM1499 was successful only at PEG concentrations of 34 to 37% (w/v) with an optimal concentration of 35% (w/v) (134). King and Dybvig reported that optimal transformation occurred when *M. mycoides* subspecies *mycoides* was in the mid-late logarithmic phase of growth (103), but Mahairas and Minion failed to observe a similar relationship with *Acholeplasma* ISM1499 (134). Transformation frequency significantly decreased after mid to late log growth phase suggesting that once mycoplasma cells have reached a certain stage of growth they are less likely to be transformed. Generally, 5 to 10 μg of homogeneous DNA were required per transformation (103, 134). The addition of yeast tRNA was shown to enhance transformation frequencies in acholeplasmas (54, 134) by acting as substrate for a potent surface nuclease (149).

Electroporation has also been used to successfully introduce DNA into mycoplasmas. The efficiency of transformation using electroporation versus PEG varies between the species of mycoplasmas and the size of the DNA. Lorenz et al. showed a 10-fold decrease in the efficiency of transfection by electroporation of the 4.5-kb mycoplasma virus L1 DNA compared to the PEG method. However, electroporation was the only means whereby the 39.5-kb mycoplasma virus L3 DNA could be introduced into *A. laidlawii* (130).

Transformation frequency of *M. mycoides* subsp. *mycoides* with plasmid pAM120 was 2 logs more efficient using PEG (103) as compared to electroporation (260). It should be noted that the electroporation conditions used (260) were markedly different than the conditions that have been successfully used with other mycoplasmas. Therefore, the transformation frequency could increase if other conditions were used. In addition to the mycoplasma species already
mentioned, electroporation has been used to transform *M. pneumoniae* (D. Krause, personal communication), *M. gallisepticum* (Minion, unpublished data), and *S. citri*. (231).

**Development of Cloning Systems for Mycoplasmas**

The development of a cloning system in mycoplasmas is essential if studies of their gene regulation are to be conducted. Although some mycoplasma genes have been cloned and expressed in *E. coli* (13, 110, 157), the majority of mycoplasma genes may not be expressed because of alternative codon usage (265). Moreover, gene regulatory signals differ between mycoplasmas and *E. coli* (67). Therefore, it is necessary to be able to study gene regulation in its natural setting, the mycoplasma cell.

The lack of genetic tools has made the development of cloning systems difficult in mycoplasmas. Only two transposons, Tn916 and Tn4001, have been shown to be useful for studying mycoplasma genetics (53, 54, 133). A number of broad host-range Gram positive plasmids have been examined as possible cloning vectors but they have proven to be unstable (49). Mycoplasma viruses have also been examined as possible cloning vectors (231), and preliminary studies show promise. Problems with low transformation frequencies and the requirement of high amounts of DNA which are inherent with the PEG transformation protocol have hindered development of an effective mycoplasma cloning system.

**Alternative codon usage in mycoplasmas**

Mycoplasma DNA is A + T-rich compared to the DNA of other bacteria. For example, the coding regions of *M. capricolum* have an A + T content of 75 mol% whereas the coding regions for *E. coli* are 51 mol% (162). The A + T content for the coding regions for *M. pneumoniae*, on the other hand, are only 60 mol% (93). Many of the tRNAs of *M. capricolum* have been purified and sequenced (3), and examination of the anticodon regions show an extremely high level (90%) of A or U in the third position (162, 172). Muto and Osawa have
examined the nucleotide usage in all three positions in the codons of *M. capricolum* as well as their use in other bacteria (163). They found that mycoplasmas were A + U biased in all three positions of the codon. The bias was highest in the third position, followed by the first position. The second position was the most invariable because changes at that position would cause a change in the amino acid. *M. capricolum* possesses only 29 different tRNA genes (2, 162, 264) for the 62 possible codon combinations. It has been predicted that there are limited number of tRNA genes found in *M. mycoides* (214). Samuelsson et al. (214) hypothesized that mycoplasmas have eliminated a number of the tRNA genes by taking advantage of the "wobble hypothesis" proposed by Crick (34). Thus, the tRNA in some instances needs only to recognize the first two nucleotides in the codon, and tRNAs that would recognize the other three possible codons would not be necessary. Examination of the tRNA species produced in *M. capricolum* support this notion in that one tRNA species appears to be available for the codons that have the same nucleotides in the first two positions and code for the same amino acid (160).

The most notable feature of the codon usage of most mycoplasmas is the deviation from the universal genetic code from a stop codon to a tryptophan encoded codon (i.e., UGA) (265). Based on a survey of 6,814 codons used in genes in *M. capricolum*, the UGA codon for tryptophan was used 42 times, the UGG codon for tryptophan was used 7 times, and the CGG codon for tryptophan was not used (160). Because the UGA codon was preferentially used, a protein possessing a tryptophan would probably include the UGA codon. In addition to *M. capricolum*, other *Mycoplasma* species have been shown to use the UGA codon for tryptophan including *M. arginini* (110), *M. gallisepticum* (94, 96), *M. genitalium* (94), *M. hyopneumoniae* (107), *M. hyorhinis* (46), and *M. pneumoniae* (93, 94, 129). Other *Mollicutes* that have been reported to use the UGA codon for tryptophan include *Ureaplasmas* (13) and *Spiroplasmas* (14, 28). *Acholeplasma laidlawii* does not use the UGA codon for tryptophan, but continues to use UGA as a stop codon (241).
Translational termination in *E. coli* is mediated by two codon-specific protein factors, release factor 1 and release factor 2. Release factor 1 recognizes the stop codons UAA and UAG, and release factor 2 recognizes the stop codons UGA and UAA (20). Because the UGA codon is used for tryptophan in most mycoplasmas, it remains to be determined whether release factor 2 has been deleted or if release factor 2 has become specific for UAA. Moreover, it is not known if mycoplasmas use release factors because they have not been identified.

The alternate codon usage of mycoplasmas creates two major problems in using *E. coli* to express mycoplasma genes. First, the UGA codon would be read as a translational stop in *E. coli* instead of a tryptophan resulting in truncated protein products (101). Unfortunately, some of the truncated protein products were not detectable with specific or convalescent antisera and thus escaped screening (13, 36, 237). Second, different codon usage preferences and tRNA concentrations may interfere with translation (3). In *E. coli*, there is a positive correlation between codon frequency and the relative concentrations of corresponding tRNAs (160). Thus, mycoplasma genes may not be expressed or are expressed at low levels in *E. coli* because the necessary tRNAs may be limiting.

**Development of cloning systems in mycoplasmas**

The first cloning system for mycoplasmas was developed by Mahairas and Minion, and was based on the ability to integrate cloned genes into mycoplasmal chromosomes via homologous recombination (132, 134). A cloning system involving two plasmids that share a region of homology such as an origin of replication and/or an antibiotic resistance marker was developed to simplify cloning steps (132). The cloning vectors are not able to replicate extrachromosomally in mycoplasmas. Thus, the plasmids must integrate into the chromosome in order to get expression and maintenance of the resistance marker. A recipient strain was created by transforming with one of the plasmids containing a mycoplasmal chromosomal DNA fragment. The plasmid integrated into the chromosome via homologous recombination.
conferring antibiotic resistance (gentamicin or tetracycline) on the transformed cell. A second plasmid, containing the cloned gene of interest, was then used to transform the recipient strain. The second plasmid integrated into the previously integrated plasmid via homologous recombination. The presence of the second plasmid is selected for and maintained by a different antibiotic resistance marker than the one used to construct the recipient strain. Mahairas and Minion used the cloning vectors pSP64 (Promega Corp.) and pKS (Stratagene) (131, 134), and attempted to express the P1 gene and operon as well as a gene encoding a 31-kDa protein from *Brucella abortus* in *Acholeplasma* (132). Although integration into the chromosome of these genes was confirmed by DNA-DNA hybridization, the protein products were not detected by screening with antisera. More recent studies have confirmed the versatility of this system by introducing recombinant plasmids into *M. gallisepticum* (Cao and Minion, unpublished results).

The primary advantage of this system is its versatility because any plasmid, whether it is of Gram positive or Gram negative origin, can be inserted into the chromosome. Thus, the cloning vectors can be shuttled between mycoplasmas and other cloning hosts. In addition, this system will accommodate DNA inserts of at least 15-kb and the inserted DNA is stable, even in the absence of selection (134). Because the cloned gene is integrated into the chromosome, it has a copy number of one or two. Thus, the copy number of the cloned gene would be similar to the way it is normally represented in the bacterial cell. If high levels of expression of the cloned gene were desired, higher gene copy numbers might be more advantageous, but high copy number plasmid cloning vectors are not presently available. The disadvantage of this system is the requirement that the cloning and amplification steps must occur in another cloning host due to the inherent limitations of the PEG transformation procedure. In addition, because the cloned insert can only integrate into one site in the chromosome, regulation in that region may also influence levels of gene expression.
Another cloning system has been developed in *S. citri* using the *Spiroplasma* spp. virus SpV1 as a cloning vehicle (231). SpV1 has a single-stranded DNA genome with a double-stranded replicative form that can be used as a cloning vehicle (188). Initially, Stamburski et al. cloned and expressed the *E. coli* chloramphenicol acetyl transferase (CAT) gene in *S. citri* (231). They were able to introduce *cat* at a frequency of $6 \times 10^5$ transfectants/μg of DNA via electroporation. Moreover, the vector provided the promoter, translational start region, and termination sequence for the *cat* gene. In order to show that genes possessing a UGA codon would still be expressed, one TGG tryptophan codon of the *cat* gene was mutated to the stop codon TGA (232). The altered *cat* gene was expressed in *S. citri*, but not in *E. coli*. The system was then used to clone and express a 1.4-kb fragment of the P1 gene of *M. pneumoniae* containing 7 UGA codons (140). Thus, this cloning and expression system has demonstrated the ability to express mycoplasmal genes. The advantages of this system include the proven ability to clone and express mycoplasmal genes, high transformation frequencies, high copy number of the recombinant in spiroplasmas, and an effective reporter gene, *cat* (187). The system's major disadvantage is instability. As the insert size increased, the ability to maintain and express the cloned gene decreased. In addition, there is a limited host range because the virus will only infect spiroplasmas.

A third cloning system has utilized naturally occurring mycoplasma plasmids as cloning vectors (55, 102). Unfortunately, there have been only three plasmids isolated from the genus *Mycoplasma*, and only two have been characterized (11, 104, 253). The mycoplasma plasmids pKMK1 and pADB201 were both isolated from *M. mycoides* subsp. *mycoides* and belong to the family of Gram positive plasmids that replicate by a single-stranded intermediate (77, 104). King and Dybvig have cloned the *tetM* gene from Tn916 into plasmid pKMK1 to generate plasmid pJI3, and they were able to introduce this plasmid into *M. mycoides* subsp. *mycoides*. This plasmid, however, was unstable and tended to delete portions of the plasmid (King and Dybvig, personal communication, 55). The potential advantages of this system remain to be
proven. While this system shows promise, there are precedents of plasmid instability using plasmids from Gram positive bacteria as cloning vectors (60, 76, 77).

It should be noted that broad host range Gram positive plasmids pVA868, pVA920, and pNZ18 have also been examined as potential cloning hosts in mycoplasmas (49, 239). Plasmids pVA868 and pVA920 were highly unstable and spontaneously deleted cloned fragments, whereas, pNZ18 was stable in A. laidlawii. A. laidlawii, however, uses UGA as a stop codon, offering no advantage over E. coli as a cloning host. These plasmids have not been shown to transform any of the other Mollicutes. Opal suppressor strains have also been examined as a means of obtaining full-length expression of mycoplasmal proteins in E. coli (Smiley and Minion, submitted; 189).

Regulation of Gene Expression in Mycoplasmas

The regulation of gene expression in mycoplasmas remains an enigma because of the lack of mycoplasma-based genetic systems to perform valid studies. Other cloning hosts such as E. coli, have been shown to be inappropriate for examining gene regulation (67). The present knowledge of gene regulation in mycoplasmas is based on models developed for other bacterial genes; it remains to be determined whether they will apply to mycoplasmas. Because mycoplasmas possess the smallest genomes necessary for autonomous existence, it is possible that they may also possess novel means of regulating their genes.

Transcription in mycoplasmas

Transcription, the synthesis of mRNA from a DNA template, occurs through the action of a DNA-dependent RNA polymerase following the specific recognition of a promoter sequence upstream of a gene or operon. The E. coli RNA polymerase holoenzyme is an 480-kDa protein that is composed of 4 subunits and can be divided into two components, the core enzyme (α2, β, β') and a sigma factor (σ) (123). The α subunit is 40-kDa and is involved in promoter
binding. The 155-kDa $\beta$ subunit is involved in nucleotide binding and the 160-kDa $\beta'$ subunit is involved in template binding during elongation. The 85-kDa sigma subunit is involved in specific promoter recognition and is released once chain elongation has initiated (123). Mycoplasma RNA polymerase has been partially characterized from two genera. Gadeau et al. isolated spiroplasma RNA polymerase by running a purified cell lysate over a heparin-agarose column (66). The product was run on a SDS polyacrylamide gel and three major bands of 140, 130-, and 38-kDa were present. The authors suggest that they correspond to the $\beta$, $\beta'$, and $\alpha$ subunits, respectively, of the core enzyme. A partially purified *M. capricolum* fraction also revealed two large proteins of 140- and 130-kDa (160).

A notable feature of the mycoplasma RNA polymerase is its insensitivity to the antibiotic rifampicin (66, 219). Rifampicin binds to the $\beta$ subunits of eubacterial RNA polymerase and prevents transcription initiation (123). *Spiroplasma* spp., *A. laidlawii*, and *M. mycoides* was shown to be 10 to 1,000 times more resistant to rifampicin than *E. coli* (66). Therefore, the structure of the mycoplasma RNA polymerase is similar to other eubacteria. The size or protein conformation of the individual subunits may differ because the mycoplasma RNA polymerase appeared to possess the same subunits as the *E. coli* RNA polymerase yet the mycoplasma RNA polymerase was resistant to rifampicin.

The sigma factor of mycoplasma RNA polymerase has not been identified. The sigma factor is important because it confers promoter specificity on the transcriptional process (86). A number of bacteria, including *E. coli* and *B. subtilis*, have been shown to possess more than one sigma factor each recognizing a different set of promoters. By altering the sigma factor component in RNA polymerase, global changes in transcription can be initiated such as during a heat shock response or during induction of sporulation (92, 164, 235). High temperature induced proteins that cross-react with the DnaK and GroEL heat shock proteins of *E. coli* have been identified in mycoplasmas (37). The presence of multiple mycoplasma sigma factors have
not been addressed, and it remains to be determined whether mycoplasmas respond to environmental changes.

**Promoters.** The analysis of mycoplasma promoters has been based on the ability of upstream regions of mycoplasma genes to generate a product in *E. coli* and on transcriptional start site mapping studies (67). Defining a mycoplasma promoter by its ability to drive the expression of a gene in *E. coli* could be misleading because mycoplasma DNA, like *E. coli* promoters, is A + T-rich. It has been shown with another A + T-rich organism, *Streptococcus pneumoniae*, that strong promoter-acting sequences were commonly observed in randomly cloned DNA fragments in *E. coli*, and they occurred more frequently than in randomly cloned *E. coli* sequences (27). Thus, mycoplasma DNA may generate pseudo promoter activity in *E. coli* (159, 184). The mapping of transcriptional start sites by S1 nuclease or primer extension have merely shown where transcription begins and does not necessarily imply that the upstream regions at -10 and -35 are the mycoplasma RNA polymerase binding or recognition sites.

Hawley and McClure analyzed over 110 *E. coli* promoter sequences and derived a consensus promoter sequence at regions -35 and -10 nucleotides upstream of the transcriptional start site (84). The -10 region or the consensus TATAAT is referred to as the Pribnow box (123). Unfortunately, relatively few mycoplasma genes have been sequenced, and most of the genes analyzed so far are rRNA, tRNA, and ribosomal protein genes (29, 30, 160). Comparison of mycoplasma regions upstream of the transcriptional start sites with the *E. coli* consensus promoter sequence showed that the -10 region was conserved, but the -35 region was more variable showing significant heterogeneity (29, 30). It should be noted that for the *E. coli* promoters examined by Hawley and McClure (84) 3.9 of 6 bases of the -35 consensus and 4.2 of 6 bases of the -10 consensus promoter sequence matched. The observation of *E. coli*-like promoters has been noted in other A + T-rich organisms, including *B. subtilis* and *S. pneumoniae* (151, 156). Because the likelihood of randomly encountering an *E. coli*-like
promoter sequence is high in these A + T-rich organisms (159), their RNA polymerases are expected to have a more stringent sequence requirement than does *E. coli* RNA polymerase (151, 156). There have been a number of attempts to develop a promoter search algorithm which would efficiently detect promoter sites within a large DNA sequence (158, 159, 171, 173, 234). These programs have been based on the ability to detect *E. coli* promoters.

Two models for the association of RNA polymerase with the -35 and -10 regions of the promoter have been offered. The bipartite model suggests that the -10 region is involved in the melting-in of the RNA polymerase near the transcriptional start site and the -35 region is involved in the specific recognition of the promoter by RNA polymerase (72). The other model, proposed by Stefano and Gralla, suggests that there is simultaneous recognition by the RNA polymerase at the -10 region, the -35 region, and the spacer elements between the two regions (233). Reznikoff and McClure, however, offer a more general model that a preferred promoter-RNA polymerase interaction exists for a given promoter, but that other influences such as ancillary proteins or supercoiling affect the ideal interaction between the promoter and RNA polymerase (192).

The transcription initiation sites have been mapped for several of mycoplasma genes and they appear 6 to 8 bp downstream of the putative -10 region (67, 242, 243). This observation agreed with the distance reported for *E. coli* (180). The apparent conservation of the Pribnow box sequence and its distance from the start site in mycoplasmas has led investigators to suggest that the *M. capricolum* RNA polymerase recognizes promoter sequences resembling those in *E. coli* (67, 160). They supported their hypothesis with the observation that promoters of the rRNA and tRNA genes were recognized by *E. coli* RNA polymerase both in vivo and in vitro (67). This experiment does not prove their hypothesis, however. Their evidence, at best, merely indicates that the melting-in region (Pribnow box) and its distance from the start site have been conserved in mycoplasmas. Because of the variability at the -35 regions of mycoplasma promoters, it would be unwarranted to assume that this region is recognized by
the mycoplasma RNA polymerase until fine structure analysis of this region has been performed in mycoplasmas.

**Transcription initiation.** Regulation of transcription initiation generally falls into two forms, negative and positive regulation. Moreover, several genetic regulatory systems have exhibited dual control effected by the same protein, with positive or negative regulation determined by the protein binding site relative to the promoter (reviewed in 192). Other means of regulating transcription initiation include DNA supercoiling (45, 178), DNA methylation (106), and altering the specificity of the RNA polymerase by substituting sigma factors (86).

Positive and negative regulation are defined by the response of the operon when no regulatory protein is present. Genes under negative control are constitutively expressed unless they are switched off by a repressor protein. One model system of negative regulation is the lac operon (reviewed in 192). Negative regulation of the lac operon occurs via the lac repressor protein which binds to the operator region (lacO) located between the promoter (lacP) for lacZ and the translational start site of the structural gene. This prevents the RNA polymerase from forming an open complex to initiate transcription. The lac repressor is a tetramer of identical 38 kDa subunits produced by lacI, located upstream of lacP. Mutations in lacI result in the constitutive phenotype irrespective of whether the inducer is present or not (148). The other component of this negative regulatory system is the inducer which binds the repressor causing an allostERIC change in the configuration of the repressor, and reduces its affinity for the operator region. The inducer is a β-galactoside (e. g., lactose, allolactose, and isopropylthio-β-galactoside or IPTG) which is the substrate for the lacZ product.

Positive regulation at transcription initiation refers to the requirement for a protein factor that is called an activator or inducer (193). This protein may be an ancillary protein that allows RNA polymerase to initiate transcription at specific promoters, a factor that replaces one of the subunits of RNA polymerase thus altering its promoter recognition specificity (i. e., sigma factor), or a new RNA polymerase (193).
Perhaps the best known example of an activator is the catabolite activator protein (CAP) or cyclic AMP receptor protein (CRP) which is activated in response to nutrient changes in *E. coli* (40) as well as other bacteria (193). *E. coli* preferentially utilizes glucose over other sugars as an energy source. For example, if both glucose and lactose were amply supplied in the medium, glucose would be preferentially metabolized and the use of lactose would be repressed in a response termed catabolite repression (40). Catabolite repression is initiated by the ability of glucose to reduce the level of cyclic AMP (cAMP) in the cell by an unknown mechanism. But if cAMP levels are high, it can bind to the CAP rendering it capable of assisting RNA polymerase in transcription initiation. The CAP-cAMP complex has been shown to activate the *pap* operon which codes for type P pilus in *E. coli*, and consequently, is involved in regulation of virulence factors (reviewed in 245).

For the *lac* operon, DNase protection assays demonstrated that the complex of cAMP-CAP bound just upstream of lacP and protected nucleotides -87 to -49, whereas, the RNA polymerase protected nucleotides -48 to +5 (191). Because the presence of CAP stimulated the initiation of *lac* mRNA synthesis by 50-fold and because the CAP protein was adjacent to the RNA polymerase binding region, it is believed that the CAP provides an additional binding site for RNA polymerase (192). The protein-protein interaction acts to either stabilize the RNA polymerase-DNA complex or induces a favorable conformational change in the RNA polymerase. Another hypothesis is that CAP induces a change in the DNA that enhances the ability of RNA polymerase to bind (190). It should be noted that the *lac* repressor has been shown to bind to nucleotides -3 to +21 which overlapped the RNA polymerase binding site. Thus, this finding supports the notion that the repressor prevents the binding of RNA polymerase (191, 192).

Regulation of transcription initiation has not been demonstrated in mycoplasmas. Muto et al. suggested that most of the genes encoded within the mycoplasma genome are constitutively expressed because the genome only possesses genes that are necessary for growth (160).
Moreover, they asserted that the negative control elements used by other eubacteria in response to environmental changes may not be required in mycoplasmas. They believed that gene regulation in mycoplasmas is a quantitative mechanism of gene regulation rather than an on or off switch for gene expression. They proposed that the quantitative regulation is similar to the stringent control response that regulates the levels of many housekeeping genes in *E. coli* in response to poor growth conditions (reviewed in 19).

Stringent control is an adaptive response that is part of a global regulatory network activated in response to amino acid starvation (19). The stringent response results in a 10- to 20-fold reduction in rRNA and tRNA synthesis, and a 3-fold reduction in mRNA synthesis. The response is triggered by the presence of an uncharged tRNA in the A site of the ribosome during translation. The major regulatory signals in this response are guanosine tetraphosphate (ppGpp) and guanosine pentaphosphate (pppGpp) which act as negative regulators of transcription initiation and/or elongation. Stringently controlled promoters in *E. coli* contain a conserved G + C-rich discriminator region just 3' to the -10 region of the promoter at -5 to +2 (248). Similar regions have been described in other bacteria including *Salmonella typhimurium* and *B. subtilis* (248). A similar G + C-rich discriminator region does not appear to be present in mycoplasmas because *M. capricolum* *rrnA* was activated rather than repressed in the presence of high levels of ppGpp in *E. coli* (67). This suggests that the control mechanism is different from that in *E. coli*.

In addition to the discriminator region in *E. coli*, A + T regions centered at -70 and -51 were found to be conserved in rRNA upstream regions (248). It has been shown in both *E. coli* and mycoplasmas that the region upstream of the -35 region affects levels of gene expression (R. Herrmann, University of Heidelberg, personal communication; 99).

**Transcription termination and attenuation.** Regulation of gene expression at the level of transcription termination also occurs in bacteria. The termination signals generally fall
into three categories: simple (rho-independent) terminators, attenuators, and rho-dependent terminators (reviewed in 175, 262).

The simple terminators consist of a series of uridine residues at the 3' end of the transcript which are preceded by a G + C-rich region of dyad symmetry capable of forming a hairpin loop (262). It has been shown that base analog substitutions or point mutations that strengthen or weaken the hairpin structure increase or decrease, respectively, the termination signal (175). The polyuridine region facilitates dissociation of the transcript from the template. Rho-independent terminator-like structures have been reported for the rRNA, tRNA, and ribosomal protein genes of *M. capricolum* that possess A + T-rich hairpin loops instead of the G + C-rich sequences in *E. coli* (161, 172, 215). In contrast, it has been shown that the transcription terminator of spiroplasmas possess G + C-rich hairpin structures (28, 206, 230). Thus, there appears to be a difference in the termination signals used by mollicutes, and the transcriptional terminators should be a consideration when developing a cloning system to express mycoplasma genes.

An attenuator is a transcription termination site located within an operon that acts to regulate the expression of the operon's gene product (266). This type of regulation is associated with amino acid biosynthesis such as the histidine, isoleucine, leucine, phenylalanine, threonine, and tryptophan operons (175). The *trp* operon has been the model system in studying attenuation in *E. coli* (266). Similar attenuator-like sequences have been described for tRNA operon coding for tRNAs that recognize the UGA and UGG codons for tryptophan (264, 265). The attenuator-like sequence is located between the two tRNA genes and results in a 5-fold decrease in the level of the tRNA<sub>UGG</sub>. This correlates with the 5- to 10-fold difference in which these codons are used by the cell (263, 264).

Rho-dependent termination, unlike simple termination, requires the protein rho (ρ) factor for transcription termination (175). Rho factor's role in transcription termination is unclear because a consensus rho-binding site has not been determined. The hairpin structure of a rho-
dependent terminator is not necessarily G + C-rich, nor is there a polyuridine region downstream. There is usually an 80-bp region upstream of the hairpin that contains pyrimidines (usually C) in the third position of every fourth codon. It is suggested that the rho protein winds itself up in the transcript via attachment to these pyrimidine residues in an ATP-dependent manner resulting in transcription termination (175). Rho-dependent termination has not been reported for mycoplasmas, but J. M. Bove (University of Bordeaux) has indicated that spiroplasmas may exhibit rho-dependent termination (9th International Organization for Mycoplasmology).

**Promoter probe vectors.** Promoter probe vectors provide a means of identifying or locating DNA sequences containing promoters in vivo. Generally, they are designed with unique cloning sites located upstream of a promoterless gene encoding an assayable and/or selectable function such as lacZ (β-galactosidase) (221), galK (galactokinase) (210), cat (chloramphenicol acetyltransferase) (78), gusA (β-glucuronidase; GUS) (68), lux (luciferase) (228), phoA (alkaline phosphatase) (138), or an antibiotic resistance marker (240). Introduction of the promoter probe vector into a cell should result in a null phenotype because no promoter is present. If a fragment containing a promoter in the proper orientation is cloned upstream of the reporter gene, the reporter gene's phenotype should be expressed.

*lac* fusions have been shown to be useful in a variety of applications (reviewed in 221). These applications include: 1) the study of the regulation of a gene or operon, 2) the detection of genes that respond to a particular regulatory signal, 3) the study of the localization of a protein within the cell, and 4) the detection of a protein or gene and the ability to assay for the protein or gene when no assay for the gene product exists. In addition, *lac* fusion vectors have been widely used because the biochemical and genetic aspects of the *lac* operon have been well studied and there are a number of indicator media available for detection of β-galactosidase activity (221). Gronenborn and Messing (75) removed the *lacZ* gene regulatory elements by
introducing an EcoRI site at the fifth codon, and Casadaban et al. (17) introduced a BamHI site at the eighth codon. Expression of the truncated lacZ gene requires that cloned upstream sequences possess a promoter, ribosomal binding site, and translational start codon. Operon fusion vectors have also been constructed using lacZ and a truncated trp promoter (16). The truncated trp promoter possesses only the translational elements, so that the cloned upstream fragment need only provide the promoter in any reading frame. Operon fusion vectors, however, have been difficult to use because of high background expression when placed on a high copy number plasmid. To avoid the high background, Simons and Kleckner introduced a transcription termination signal upstream of lacZ to avoid read-through from adjacent genes (223).

In addition to plasmid based promoter probe vectors, transposable elements carrying lacZ have been used to detect promoters directly on the chromosome. These vectors include derivatives of Mu phage, Mu d lac 1 and 2 (18) and λ plac Mu (108), and many transposons such as Tn3 (229), Tn10 (256), Tn5 (113), and Tn917 (268). These vectors are able to locate promoters in the chromosome by inserting into regions of active transcription and/or translation which drive the expression of the reporter gene carried within the transposon.

Translation in mycoplasmas

The mechanisms of translational regulation in mycoplasmas are merely speculation at this point because of a lack of study. Genetic systems are not available to study translation in mycoplasmas. Like other aspects of gene regulation previously discussed, the precise regions involved are hypothetical based on observations made in other bacterial systems. The probable ribosomal binding sites, as well as translational start sites, for over twenty M. capricolum ribosomal genes have been reported (172). In addition, the probable translational start regions for a number of spiroplasmal proteins (28), the arginine deiminase gene of M. arginini (110), and the urease gene of U. urealyticum (13) have been reported. All of these genes appeared to
use AUG as a start codon except for 1 gene of unknown function in \textit{S. citri} (28). Moreover, there appeared to be ribosomal binding sites at the proper distance (5 to 9 nucleotides) upstream of the start site.

The 3' ends of the 16S rRNA sequence of \textit{M. capricolum} and \textit{S. citri} have been described and were shown to be identical (28, 97). Moreover, the last 15 nucleotides of the 3' ends of the 16S rRNA genes of these mycoplasma species are identical to the last 15 nucleotides of the 16S rRNA gene of \textit{B. subtilis} and have a 12-nucleotide homology with \textit{E. coli} 16S rRNA (28). Thus, at the sequence level, the translational initiation regions of mycoplasmas appear to be similar to those of \textit{E. coli} and \textit{B. subtilis}. A report by Notarnicola et al. (170) has even made the issue involving mycoplasma translational start regions more confusing. They observed translational initiation inside a mycoplasma coding sequence in \textit{E. coli}.

The ability to clone and express Gram positive genes in Gram negative bacteria has been reported to occur efficiently, but the ability to clone and express Gram negative genes in Gram positive bacteria has been inefficient (81). One possible reason for the difference in the ability to express genes from Gram negative bacteria in Gram positive is the differences in binding stringencies of their respective ribosomes and mRNA (81, 143). The base pairing of the 16S rRNA with the ribosomal binding sites of \textit{E. coli} had a generally lower free energy of binding than that of ribosomes and mRNA of Gram positive bacteria.

Differences in the stringency of binding of 16S rRNA and the ribosomal binding sites between Gram positive and negative bacteria cannot account for all the differences in the ability to express in the heterologous host. Hager and Rabinowitz showed that T7 genes, which possess a long ribosomal binding site, were not translated as well by \textit{B. subtilis} ribosomes as genes from the \textit{B. subtilis} phage \textphi 29 that has the same ribosomal binding site as T7 genes (80, 81). Because the spacing between the ribosomal binding sites and the start sites was nearly identical, the authors suggested that the difference in translation efficiency was due to secondary structure or interaction with the ribosome at another region in the transcript.
Secondary structure of the 5' untranslated region of the \textit{gltBDF} operon of \textit{E. coli} has been shown to regulate expression (252). Also, enhancing sequences recognized by the ribosome located within the 5 to 9 nucleotide spacer, 5' of the ribosomal binding site, or 3' of the translational start site have been reported (reviewed in 141).

Loechel et al. have reported that they have identified a novel translational initiation region from \textit{M. genitalium} that functioned in \textit{E. coli} (128). The efficiency of this translation initiation was affected by the secondary structure around the site as well as sequences upstream and downstream of the translational start site. Therefore, it appears that secondary structure of the translation initiation region, as well as enhancer-like sequences, may also affect the ability to translate genes in mycoplasmas.
MATERIALS AND METHODS

Bacterial Strains and Culture Media

The strains and plasmids used in this study are listed in Table 1. Acholeplasma strain designations ISM2004-2010, or ISM2050.X and ISM1499.2050.X represent mycoplasma recombinants containing plasmids pISM2004-pISM2010 or pISM2050.X, respectively. The X suffix designates a derivative of the parent lacZ fusion vector, pISM2050, containing a cloned Acholeplasma chromosomal DNA fragment. Fragment sizes of these cloned fragments are given in Table 1. Strain ISM1499 is a laboratory isolate that has been serotyped by J. G. Tully (NIAID) to be related to Acholeplasma oculi and Acholeplasma laidlawii. All references to M. pulmonis ISM1499 should be regarded as Acholeplasma ISM1499.

Acholeplasma strains were grown in PPLO broth (Difco Laboratories, Detroit, Mich.) supplemented with 15% gamma globulin-free horse serum (GIBCO Laboratories, Grand Island, N.Y.), 2.5% yeast extract, 0.5% glucose, 2.5 μg/ml of Cefobid (Pfizer, Inc., New York, N.Y.), 0.02% herring sperm DNA (Sigma Chemical Co., St. Louis, Mo.), 0.002% phenol red, and (when required) 1% Noble agar (Difco). Selective levels of antibiotics for Acholeplasma ISM1499 were 2 μg/ml tetracycline (Sigma) and 15 μg/ml gentamicin (Sigma), and resistant cultures were maintained at 10 μg/ml tetracycline and 15 to 20 μg/ml gentamicin. E. coli cultures were grown in Luria broth containing 100 μg/ml ampicillin (Sigma) or 12.5 μg/ml tetracycline when appropriate.

Reagents and Buffers

Restriction enzymes (RE), T4 DNA ligase, and Klenow enzyme were purchased from Promega Corporation (Madison, Wis.) or GIBCO BRL (Gaithersburg, Md.) and used according to manufacturer's instructions. 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) (GIBCO BRL or Gold Biotechnology, Inc., St. Louis, Mo.) was used at a concentration of
Table 1. Bacterial strains and plasmids

<table>
<thead>
<tr>
<th>Bacterial strain or plasmid</th>
<th>Relevant genotype and phenotypea</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>φ80dlacZΔM15 endA1 recA1 hsdR17 supE44 thi-1 gyrA relA1 F- Δ(lacZYA argF)U169</td>
<td>BRL</td>
</tr>
<tr>
<td>CSH50</td>
<td>ara D(lac-pro) strA thi F-</td>
<td>(147)</td>
</tr>
<tr>
<td>CSH50.Xb</td>
<td>CSH50(pISM2050.X)</td>
<td>This study</td>
</tr>
<tr>
<td>χ289</td>
<td>F- supE42 λ- T3F</td>
<td>R. Curtiss III</td>
</tr>
<tr>
<td>XL1-Blue</td>
<td>recA1, lac-, endA1, gyrA96, thi, hsdR17, supE44, relA1 [F' proAB, lacI, lacZAM15, Tn10] Tc$^r$</td>
<td>Stratagene</td>
</tr>
</tbody>
</table>

| **Acholeplasma sp.**        |                                  |        |
| ISM1499                     | Laboratory isolate              | This study |
| ISM1499.X                   | ISM1499 with integrated plasmid X Gm$^r$ | This study |
| ISM1520                     | ISM1499 with integrated pISM1026 Tc$^r$ | (132)  |
| ISMX                        | ISM1520 with integrated plasmid X Gm$^r$ Tc$^r$ | This study |

| **Plasmids**                |                                  |        |
| pDIA15                      | trp-lacZYA Km$^r$                | (41)   |
| pKS$^-$                     | Multipurpose cloning vector Ap$^r$ | Stratagene |
| pMC1871                     | Prokaryotic fusion vector Tc$^r$ | Pharmacia |
| pSP64                       | Multipurpose cloning vector Ap$^r$ | Promega |
| pSP71                       | Multipurpose cloning vector Ap$^r$ | Promega |
| pISM205                     | pKS$^-$ containing the 3.1-kb lacZ fragment from pMC1871 in which the downstream BamHI RE site has been altered Ap$^r$ | This study |

a Ap$^r$, Tc$^r$, Gm$^r$, and Km$^r$ denote resistance to ampicillin, tetracycline, gentamicin, and kanamycin, respectively. RE denotes restriction endonuclease.

bX denotes the plasmid designation which was then used to denote the strain designation (e.g., Strain ISM2050.1 = ISM1520 with pISM2050.1, Strain ISM1499.2050.1 = ISM1499 with pISM2050.1, and Strain CSH50.2050.1 = CSH50(pISM2050.1).
<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pISM1001</td>
<td>13.5-kb plasmid containing Tn4001 Ap&lt;sup&gt;f&lt;/sup&gt; Gm&lt;sup&gt;f&lt;/sup&gt;</td>
<td>(133)</td>
</tr>
<tr>
<td>pISM1003</td>
<td>5.75-kb plasmid containing the gentamicin resistance marker from Tn4001 Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>(134)</td>
</tr>
<tr>
<td>pISM2002</td>
<td>pISM1003 with <em>M. capricolum</em> rrnA P2 promoter Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2003</td>
<td>pISM1003 with <em>E. coli</em> rrnB P1 promoter Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2004</td>
<td>pISM1003 with <em>trp-lacZYA</em> Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2005</td>
<td>pISM2002 with <em>trp-lacZYA</em> Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2006</td>
<td>pISM2003 with <em>trp-lacZYA</em> Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2009</td>
<td>pISM2004 with <em>E. coli</em> rrnB P1 promoter containing 700 bp of the upstream region of the <em>M. capricolum</em> rrnB P2 promoter Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2010</td>
<td>pISM2004 with <em>M. capricolum</em> rrnA P2 promoter containing 700 bp of the upstream region of <em>tetM. capricolum</em> rrnB P2 promoter Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2011</td>
<td>pISM2004 with <em>E. coli</em> rrnB P1 promoter containing 700 bp of the upstream region of the <em>M. capricolum</em> rrnB P2 promoter in the reverse orientation Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2012</td>
<td>pISM2004 with <em>M. capricolum</em> rrnA P2 promoter containing the -35 region of the <em>E. coli</em> rrnB P1 promoter and the -10 region of <em>tetM. capricolum</em> rrnB P2 promoter Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2013</td>
<td>pISM2012 containing the <em>trp-lacZYA</em> Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2025</td>
<td>10.3-kb rescue plasmid derivative of pSP64 containing the <em>tetM</em> from Tn916 and the gentamicin resistance marker from Tn4001 Ap&lt;sup&gt;f&lt;/sup&gt; Tc&lt;sup&gt;f&lt;/sup&gt; Gm&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2030</td>
<td>pISM2002 containing the B52 adhesion gene from <em>M. pneumoniae</em> Ap&lt;sup&gt;f&lt;/sup&gt; Gm&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2031</td>
<td>pISM2002 containing the B52 adhesion gene from <em>M. pneumoniae</em> in the reverse orientation Ap&lt;sup&gt;f&lt;/sup&gt; Gm&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2032</td>
<td>pISM2008 containing the B52 adhesion gene from <em>M. pneumoniae</em></td>
<td>This study</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>pISM2033</td>
<td>pISM2008 containing the B52 adhesion gene from <em>M. pneumoniae</em> in the reverse orientation</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2034</td>
<td>pISM2002 containing the PI adhesion gene from <em>M. pneumoniae</em></td>
<td>This study</td>
</tr>
<tr>
<td>pISM2035</td>
<td>pISM2002 containing the PI adhesion gene from <em>M. pneumoniae</em> in the reverse orientation</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2036</td>
<td>pISM2008 containing the PI adhesion gene from <em>M. pneumoniae</em></td>
<td>This study</td>
</tr>
<tr>
<td>pISM2037</td>
<td>pISM2008 containing the PI adhesion gene from <em>M. pneumoniae</em> in the reverse orientation</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2046</td>
<td>pKS* with <em>BamHI</em> and <em>Smal</em> RE sites removed</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2048</td>
<td>pISM2046 containing a 1.3-kb fragment containing the ca. 1-kb distal end of an IS256 arm of Tn4001</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2048.5</td>
<td>pISM2048 with the insertion of <em>BamHI</em> and <em>Smal</em> RE sites ca. 30 bp from the distal end of an IS256 arm of Tn4001</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050</td>
<td>Promoterless <em>lacZ</em> fusion vector</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.1</td>
<td>pISM2050 with 2.8-kb of ISM1499 chromosomal DNA</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.2</td>
<td>pISM2050 with 0.9-kb of ISM1499 chromosomal DNA</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.8</td>
<td>pISM2050 with 1.0-kb of ISM1499 chromosomal DNA</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.18</td>
<td>pISM2050 with 0.9-kb of ISM1499 chromosomal DNA</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.19</td>
<td>pISM2050 with 1.5-kb of ISM1499 chromosomal DNA</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.25</td>
<td>pISM2050 with 0.8-kb of ISM1499 chromosomal DNA</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.39</td>
<td>pISM2050 with 1.0-kb of ISM1499 chromosomal DNA Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>pISM2050.40</td>
<td>pISM2050 with 1.7-kb of ISM1499 chromosomal DNA Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.66</td>
<td>pISM2050 with 2.0-kb of ISM1499 chromosomal DNA Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.69</td>
<td>pISM2050 with 2.5-kb of ISM1499 chromosomal DNA Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.70</td>
<td>pISM2050 with 2.3-kb of ISM1499 chromosomal DNA Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.86</td>
<td>pISM2050 with 2.2-kb of ISM1499 chromosomal DNA Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2050.89</td>
<td>pISM2050 with 1.0-kb of ISM1499 chromosomal DNA Gm&lt;sup&gt;f&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2061</td>
<td>pKS&lt;sup&gt;-&lt;/sup&gt; containing a 5.8-kb PstI-SacI fragment from pISM1001 harboring Tn4001 Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2062</td>
<td>pISM2061 containing the BamHI and Smal RE sites in an IS256 arm of Tn4001 Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2062&lt;sub&gt;lac&lt;/sub&gt;</td>
<td>pISM2062 containing the 3.1-kb lacZ fragment from pMC1871 Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2062.2</td>
<td>pISM2062 with two translational stop codons altered Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2062.3</td>
<td>pISM2062 with three translational stop codons altered Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2062.&lt;sub&gt;2lac&lt;/sub&gt;</td>
<td>pISM2062.2 containing the 3.1-kb lacZ fragment from pMC1871 Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2062.&lt;sub&gt;2lac(rev)&lt;/sub&gt;</td>
<td>pISM2062.2 containing the 3.1-kb lacZ fragment from pMC1871 in the reverse orientation Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2062.&lt;sub&gt;3lac&lt;/sub&gt;</td>
<td>pISM2062.3 containing the 3.1-kb lacZ fragment from pMC1871 Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
</tbody>
</table>
Table 1. (continued)

<table>
<thead>
<tr>
<th>Plasmid Code</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pISM2062.3lac(rev)</td>
<td>pISM2062.3 containing the 3.1-kb lacZ fragment from pMC1871 in the reverse orientation Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2064</td>
<td>pISM2062.2 containing pISM205 Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2065</td>
<td>pISM2062.3 containing pISM205 Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2075</td>
<td>Exo/Mung deletion clone of pISM2050.2 Lac&lt;sup&gt;-&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2076</td>
<td>Exo/Mung deletion clone of pISM2050.2 Lac&lt;sup&gt;+&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2077</td>
<td>Exo/Mung deletion clone of pISM2050.2 Lac&lt;sup&gt;+&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2078</td>
<td>Exo/Mung deletion clone of pISM2050.2 Lac&lt;sup&gt;+&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2079</td>
<td>Exo/Mung deletion clone of pISM2050.2 Lac&lt;sup&gt;+&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2080</td>
<td>Exo/Mung deletion clone of pISM2050.2 Lac&lt;sup&gt;+&lt;/sup&gt; Ap&lt;sup&gt;r&lt;/sup&gt; Gm&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2081</td>
<td>pKS containing a 1.4-kb fragment harboring the 16S rRNA gene of ISM1499 Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>pISM2083</td>
<td>pSP71 containing the promoter containing insert from pISM2050.2 Ap&lt;sup&gt;r&lt;/sup&gt;</td>
<td>This study</td>
</tr>
</tbody>
</table>

80 μg/ml when scoring for the Lac<sup>+</sup> phenotype in Acholeplasma or 25 μg/ml in E. coli. All other chemicals and reagents were obtained from Sigma except where noted. Isopropylthio-β-D-galactoside (GIBCO BRL or Gold Biotechnology) was used at a final concentration of 10 mM. Oligonucleotides were synthesized by the Iowa State University Nucleic Acid Facility.

Transformation of Acholeplasma spp., M. gallisepticum, and E. coli

The transformation of Acholeplasma was performed as described previously (134). Mycoplasma gallisepticum was transformed via electroporation using the BTX model 600...
electroporator (Biotechnologies & Experimental Research, Inc., San Diego, Cal.), using 2 mm gap cuvettes, with settings of 2.5 kV, 50 μF, and 129 ohm, to generate a pulse length of 5 msec. Mid-log phase cells were washed 3 times with prechilled electroporation buffer (272 mM sucrose, 8 mM HEPES, pH 7.4). After the final wash the cells were resuspended in one-fourth to one-tenth the original culture volume with electroporation buffer and placed on ice. One hundred microliters of cells were mixed with 20 μl of electroporation buffer containing 7 to 10 μg of DNA and placed in the cuvette on ice. After a 10 min incubation the mixture was pulsed. The cells were immediately resuspended in 1 ml of complete PPLO media, transferred to a prechilled microfuge tube, and incubated on ice for 10 min. The cells were incubated at 37°C for 0 to 2 hr prior to plating. X-gal was spread on the agar surface 1 to 2 h prior to plating the mycoplasma transformation mixture. Blue colonies were picked and analyzed for β-galactosidase (β-gal) activity. Transformation of *E. coli* was achieved by the method of Hanahan (82) or by electroporation using cuvettes with a 1 mm gap, settings of 0.81 kV, 50 μF, and 129 ohm, to generate a pulse length of 5 msec.

**Preparation of *Acholeplasma ISM1499* Chromosomal DNA**

Chromosomal DNA was prepared by washing the cells isolated from a 500 ml culture by centrifugation (10,000 x g, 15 min) with phosphate buffered saline (PBS, 10 mM sodium phosphate, 150 mM NaCl, pH 7.4) and resuspending the pellet in 5 ml of NET buffer (0.15 M NaCl, 0.08 M ethylene diamine tetraacetic acid (EDTA), 0.1 M Tris, pH 7.5) containing 10 mg/ml Proteinase K. After a 2 h incubation at 37°C, the cells were lysed with 1 ml of a detergent solution containing 1% sodium dodecyl sulfate (SDS), 1% Nonidet P-40, and 1% deoxycholate. The DNA was extracted with phenol until the interface between the organic and aqueous phases was clear. The DNA was then extracted once with phenol-chloroform. The DNA was purified on a cesium chloride-ethidium bromide buoyant density gradient (213).
DNA bands were collected, ethidium bromide extracted, and salt removed as described previously (213).

The chromosomal DNA was partially digested with Sau3AI and fragments were separated by a sucrose gradient as described previously (213). One-to six-kilobase fragments were isolated and cloned into pISM2050 using standard recombinant DNA techniques.

**Plasmid Constructions**

**Construction of *M. capricolum* promoter-*trp'-*lacZYA translational fusions**

The -35 and -10 regions of the *M. capricolum* rrnA P2 promoter or the -35 and -10 regions of the *E. coli* rrnB P1 promoter (87) were cloned into the EcoRI and BamHI sites of plasmid pISM1003 (133) to generate plasmids pISM2002 or pISM2003, respectively. Plasmids pISM2005 and pISM2006 were constructed by inserting the 7.1-kb BamHI fragment containing *trp'-*lacZYA from plasmid pDIA15 (41) downstream of the promoters in plasmids pISM2002 and pISM2003. Plasmids pISM2005 and pISM2006 are identical to plasmids pISM2005 and pISM2006, respectively, except that an additional 700-bp region upstream of the *M. capricolum* rrnB P2 promoter has been placed upstream of the promoters. Plasmid pISM2011 is identical to plasmid pISM2010 except that the *M. capricolum* promoter and rrnB P2 upstream sequences are in the reverse orientation.

**Construction of pISM2062 by using site-directed mutagenesis via the polymerase chain reaction (PCR)**

The Smal and BamHI RE sites were placed into an IS256 arm of Tn4001 by using oligonucleotide-directed mutagenesis via PCR, generating plasmid pISM2048.5. The mutant oligonucleotide primer, 5′ TTGTGTTAAAAAGTCCCCGGGATCCAAAAAGGCCCCATAT-3′, and the T3 primer provided with the plasmid pKS- were used to synthesize a 1-kb fragment
flanked with *Hind*III and *BamHI* RE sites. The complementary mutant oligonucleotide primer, 5'-ATATGGCTTTTTGGATCCCCGGGACTTTTACACAA-3', and the T7 primer were used to synthesize a 300-bp fragment flanked with *BamHI* and *PstI* RE sites. The DNA fragments were generated with a Coy Tempcycler. One hundred nanograms of each primer was added to 50 ng of template (pISM2048) and the PCR was run using the following program: four cycles of 7 min denaturation at 94°C, 10 min cooling, 5 min annealing at 25°C and 4 min of polymerization at 72°C; twenty cycles of 90 s denaturation at 94°C, 90 s annealing at 48°C, and 2 min of polymerization at 72°C; and one cycle of 90 s denaturation at 94°C, 90 s annealing at 48°C and 7 min of polymerization at 72°C.

**Construction of the promoter probe vector, pISM2050**

Plasmid pISM2050 is a derivate of plasmid pISM1003 (133). Plasmid pISM1003 was digested with *BamHI* and the 3.1-kb *BamHI* fragment of plasmid pMC1871 (16) containing a promoterless *lacZ* gene was ligated into the site. In order to generate a unique *BamHI* RE site for cloning, the downstream *BamHI* site was removed by partial digestion with *BamHI* followed by a fill-in reaction with Klenow fragment of DNA polymerase I, ligation to circularize, and transformation into *E. coli*. Transformants were selected on ampicillin-containing media, and the proper construct was confirmed by restriction enzyme analysis.

Plasmid pISM2050 was constructed with a gentamicin marker for selection in mycoplasmas, a promoterless *lacZ* reporter gene with unique upstream *BamHI* and *SmaI* RE sites for cloning *Sau3AI* and blunt end fragments, respectively, and an ampicillin marker and origin of replication for amplification of DNA in *E. coli* prior to transformation into *Acholeplasma* . The ampicillin gene and origin of replication also provided a region of homology to insert the plasmid into the *Acholeplasma* recipient strain ISM1520 via homologous recombination (134).
Preparation of RNA and Measurement of mRNA Levels

Total RNA from a 5 to 10 ml overnight culture was prepared using RNA STAT-60 isolation reagent (guanidinium thiocyanate-phenol mixture) (Tel-Test "B", Inc., Friendswood, Tex.) according to manufacturer's instructions. Messenger RNA levels were measured by slot blot analysis using a Minifold II apparatus (Schleicher & Schuell, Keene, N.H.) and the procedure described previously (213). Two micrograms of total RNA were placed into each slot and transferred to nitrocellulose (Schleicher & Schuell) and probed with a 1.4-kb fragment containing the 16S rRNA gene from ISM1499 or a 3-kb fragment containing the lacZ gene. The blot was exposed to X-ray film or was examined with a PhosphorImager and analyzed with the ImageQuant program (Molecular Dynamics, Sunnyvale, Calif.). Levels of the lacZ mRNA were adjusted by normalizing to the 16S rRNA levels to account for differences in amounts of RNA that may have been loaded in each slot for each strain.

β-galactosidase Assays

β-galactosidase assays were performed as described by Schleif and Wensink (217). Briefly, 5 ml of an Acholeplasma culture was washed once with PBS and the pellet was resuspended in 2 ml of Z buffer (0.1 M sodium phosphate, pH 7.0; 0.001 M magnesium sulfate; 0.1 M 2-mercaptoethanol). Fifty microliters of 0.1% SDS were added to 1 ml of the cell suspension and vortexed for 15 sec. One hundred microliters of the lysed cell suspension were added to 1 ml of Z buffer and equilibrated to 37°C. In order to examine the β-gal activity of an unlysed cell preparation, 100 μl of the remaining cells suspended in Z buffer were added to another 1 ml of Z buffer and similarly equilibrated to 37°C. Two hundred microliters of a 4 mg/ml o-nitrophenyl-β-D-galactopyranoside (Sigma) solution were added. The reaction was stopped by adding 0.5 ml of 1 M sodium carbonate. The absorbance of the reaction was measured with a Spectronic 20 (Bausch & Lomb) at an absorbance of 420 nm. E. coli cultures were assayed similarly except that 100 μl of a 1 ml culture was used, and the cells were
permeabilized with 0.1% SDS and a couple of drops of chloroform. Fully induced \textit{E. coli} strain \chi289 was used as a positive control.

Protein concentration was measured by using the Bio-Rad protein assay (Bio-Rad, Richmond, Calif.) according to the manufacturer's directions. The number of colony forming units (CFU) per milligram of protein was determined by dividing the number of CFU/ml by mg/ml from the average of three cultures each measured in triplicate. The amount of \(\beta\)-gal enzyme required to give an OD = 1.0 at an absorbance of 420 nm in one minute is \(4.45 \times 10^{12}\) monomers (217). The levels of \(\beta\)-gal activity, therefore, was expressed as the number of monomers per CFU as shown below. Cultures were assayed in triplicate.

\[
\text{Units of } \beta\text{-gal activity} = \frac{A_{420} \times 4.45 \times 10^{12} \text{ monomers}}{\text{time (min)} \times \text{mg protein} \times \text{CFU/mg}}
\]

**Immunoblot Analysis**

Protein samples containing 25 \(\mu\)g protein from washed mycoplasmas were resuspended in water and lysed with 0.01% (final concentration) SDS, SDS-PAGE sample buffer (116) was added, the samples were boiled for 5 min and then separated on a 7.5% polyacrylamide resolving gel. Following electrophoresis for 4 h at 25 mAmp constant current, proteins were transferred to nitrocellulose following the procedure of Towbin (247). The blots were analyzed using a 1:3,000 dilution of a monoclonal antibody to \(\beta\)-gal (Promega) followed by goat anti-mouse antibody conjugated to alkaline phosphatase (Kirkgaard & Perry Laboratories, Inc., Gaithersburg, Md.) (1:1,000 dilution). The blot was developed by using the BCIP/NBT one component alkaline phosphatase substrate system (Kirkgaard & Perry Laboratories, Inc.).
DNA Sequencing

The mycoplasma chromosomal DNA driving the expression of the lacZ gene was sequenced using the Sequenase Kit Version 2.0 (United States Biochemical) or was sequenced by the Nucleic Acid Facility at Iowa State University. The sequencing primer was an oligonucleotide that was 30 bases long (30-mer) (5'-GCTGGCGAAAGGGGAATGCATGCTGC AAGGC-3') that is reverse and complimentary to the lacZ gene at approximately 50 nucleotides downstream of the lacZ-chromosomal DNA fusion point. The mycoplasma inserts were sequenced by either isolating the insert from a gel or subcloning the fragment into pSP71, because pKS also possessed lacZ sequences. The inserts isolated from a 0.6% Tris acetate agarose gel were purified by GeneClean (San Diego, Calif.). The fragments were subcloned into either the Clal or Clal/HindIII RE sites of pSP71. The plasmid pSP71 derivatives were purified using Qiagen (Qiagen Inc., Chatsworth, Calif.) plasmid purification columns. Double-stranded DNA was denatured in 0.2 M NaOH, 0.2 mM EDTA at 37°C for 30 min. Following ethanol precipitation, the DNA was resuspended in 7 μl H2O. The sequencing reaction was conducted as described by the manufacturer. Approximately 20 ng of the primer was used for each sequencing reaction.

Primer Extension Analysis

The transcriptional start sites of the lacZ fusion constructs were determined by using primer extension following the protocol described elsewhere (4). Briefly, 10 μl total RNA (containing 10 to 100 μg RNA) were mixed with 1.5 μl 10X hybridization buffer (1.5 M KCl, 0.1 M Tris, 0.01M EDTA, pH 8.3) and 3.5 μl of primer end-labeled with [γ-32P]ATP. After 1 min at 90°C, the mixture was placed at 65°C for 90 min and then cooled slowly to room temperature. To each primer extension reaction a mixture containing 0.9 μl 1 M Tris (pH 8.0), 0.9 μl 0.5 MgCl2, 0.25 μl 1 M dithiothreitol, 6.75 μl of 1 mg/ml actinomycin D, 1.33 μl of 5 mM deoxynucleotide mix (all 4 deoxynucleotides), 20 μl H2O, and 0.2 μl of 25 U/μl avian
myeloblastosis virus (AMV) reverse transcriptase was added. After an 1 h incubation at 42°C, RNase A was added to a final concentration of 15 μg/ml and incubated for an additional 15 min at 37°C. The DNA was resuspended in 3 to 5 μl of TE (0.01 M Tris, 0.001 M EDTA, pH 8.0) buffer following ethanol precipitation. Three microliters of the primer extension product were electrophoresed on an 8% acrylamide sequencing gel. As a reference, a sequencing reaction of the upstream lacZ fusion region with the sequencing primer was loaded beside the primer extension product. The primer extension product was also compared to a sequencing ladder generated from pISM2083 using the sequencing primer.

**Exo/Mung Deletion Analysis of the pISM2050.2 Insert**

Plasmid pISM2050.2 was digested with PstI and Smal restriction enzymes, and then subjected to Exonuclease III and Mung Bean deletion analysis according to manufacturer's (Stratagene) instructions. Briefly, 20 μg of plasmid pISM2050.2 were first digested with PstI then Smal. Following a 0 (no enzyme), 1, 2 or 3 min digestion with 20 U Exonuclease III/μg DNA, the mixture was heated to 68°C and placed on dry ice. Fifteen units of Mung Bean nuclease were added to the mixture and incubated for 30 min at 30°C. Following a phenol-chloroform extraction, the deletion products were ligated and used to transform *E. coli* DH5α.
RESULTS

*M. capricolum* Promoter *trp'-lacZYA* Fusion Studies

Construction of *lacZ* translational fusions with the *M. capricolum* *rrnA* P2 promoter and β-gal activities

Initial studies were designed to explore the possibility of using the *E. coli* *lacZ* as a reporter gene for studying genetic regulation in mycoplasmas. The study examined the ability of a defined mycoplasma promoter to generate a translational fusion with the *trp'-lacZYA* operon. Plasmids pISM2004 - 2006 and pISM2009 - 2011 were constructed and transformed into *Acholeplasma* ISM1520 as described in the Materials and Methods. The resulting *Acholeplasma* recombinants were analyzed for β-gal activity, and the results are shown in Figure 1A. The overall levels of β-gal activity in *Acholeplasma* were very low in comparison to *E. coli*. The *M. capricolum* promoter was able to drive the expression of *lacZ* in both *E. coli* and *Acholeplasma* ISM1520. The *M. capricolum* promoter in the reverse orientation with respect to *lacZ* (pISM2011), however, was not able to generate β-gal activity in either *Acholeplasma* or *E. coli*. Slightly higher levels of β-gal activity were generated in *Acholeplasma* with the additional 700 bp region from the *M. capricolum* *rrnB* P1 promoter region in plasmid pISM2010 supporting an earlier observation that the region upstream of the promoter influences transcriptional levels (87, 99).

Analysis of transcript levels

To examine the relative strengths of each promoter, transcript levels of *Acholeplasma* strains ISM2004, ISM2005, ISM2006, ISM2009, and ISM2010 were measured. Figure 1B shows that the *M. capricolum* *rrnA* P2 promoter (ISM2005 and ISM2010) generated the highest transcript levels of the *lacZ* gene in *Acholeplasma*. The transcript levels using
Figure 1. Construction, message levels, and β-galactosidase activities of the trp'-lacZYA fusion plasmids. (A) Construction of the trp'-lacZYA fusion plasmids and definition of units of β-galactosidase activity are described in the Materials and Methods. Single headed arrows show the direction of transcription. Double headed arrow and hatched bar denotes the 700 base pair region upstream of the M. capricolum rnb P2 promoter. Shaded boxes represent the trp'-lacZYA operon (B). Slot blot analysis of the RNA levels from Acholeplasma ISM1499, ISM1520, ISM2004, ISM2005, ISM2006, ISM2009, and ISM2010 probed with either 16S rRNA (left) or lacZ-specific (right) gene probe. The values shown on the right hand side of the figure represent the counts in a defined volume after subtracting background and adjusting for the relative amounts of total RNA loaded into each slot. Levels of RNA used were adjusted by first normalizing the relative amounts of RNA used in the rRNA slots then similarly adjusting the RNA levels in the lacZ slots. Abbreviations: M. capricolum, M. cap; Ampicillin, Ap; Gentamicin, Gm; and, origin of replication, ori.
A

<table>
<thead>
<tr>
<th>Year</th>
<th>Description</th>
<th>Strain</th>
<th>ISM1499</th>
<th>ISM1520</th>
<th>ISM2004</th>
<th>ISM2005</th>
<th>ISM2006</th>
<th>ISM2009</th>
<th>ISM2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>M. cap rrnA P2</td>
<td>E. coli</td>
<td>0</td>
<td>0</td>
<td>40,652</td>
<td>543,651</td>
<td>0</td>
<td>32,574</td>
<td>329,117</td>
</tr>
<tr>
<td>2005</td>
<td>E. coli rrnB P1</td>
<td>E. coli</td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
<td>32,574</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>E. coli rrnB P1</td>
<td>E. coli</td>
<td></td>
<td></td>
<td>3,911</td>
<td>0</td>
<td></td>
<td>32,574</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>M. cap rrnA P2</td>
<td>E. coli</td>
<td></td>
<td></td>
<td>4,163</td>
<td>4</td>
<td></td>
<td>32,574</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>M. cap rrnA P2</td>
<td>E. coli</td>
<td></td>
<td></td>
<td>4,922</td>
<td>1</td>
<td></td>
<td>32,574</td>
<td></td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>Strain</th>
<th>rRNA</th>
<th>lac</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISM1499</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>ISM1520</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>ISM2004</td>
<td>40,652</td>
<td></td>
</tr>
<tr>
<td>ISM2005</td>
<td>543,651</td>
<td></td>
</tr>
<tr>
<td>ISM2006</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>ISM2009</td>
<td>32,574</td>
<td></td>
</tr>
<tr>
<td>ISM2010</td>
<td>329,117</td>
<td></td>
</tr>
</tbody>
</table>
the *E. coli* promoter only (pISM2006) was negligible whereas the addition of the 700 bp *M. capricolum* *rrnB* P1 upstream sequences (ISM2009) did yield detectable, but low levels of the *lacZ* transcript. Surprisingly, the *E. coli* consensus promoter was unable to drive expression of *lacZ* in *Acholeplasma* despite close homology in the -10 and -35 sequences (67). Transcript levels of ISM2004 were detectable and were probably due to read through from adjacent genes, a common problem when using *lacZ* translational fusion vectors (221, 223).

**Tn4001lac Promoter Probe Studies**

**Placement of unique *SmaI* and *BamHI* RE sites in a IS256 arm of Tn4001**

In order to construct the Tn4001lac promoter probe transposon, unique RE sites were first placed into one of the IS256 arms of Tn4001. The unique RE sites were then used as insertion sites the *lacZ* gene. This was accomplished in several steps outlined in Figure 2. First, because pKS+ contains *BamHI* and *SmaI* RE sites, these sites were removed, generating plasmid pISM2046 Figure 2A). The *SmaI* and *BamHI* RE sites were inserted 28 bases from the distal end of the IS256 arm by first cloning one IS256 arm into plasmid pISM2046 (Figure 2B) and then using PCR to create the desired mutation (Figure 2C). Because the inverted repeat region only spans the first 26 bp of the IS256 element (15), it was predicted that the transposase would recognize and bind to the modified inverted repeat normally (10, 42).

In order to reduce the size of the Tn4001 containing plasmid, and to have the mutated Tn4001 eventually reside on a vector that does not possess the *BamHI* and *SmaI* RE sites, the 5.7-kb *PstI-SalI* fragment containing Tn4001 was ligated into plasmid pISM2046, generating plasmid pISM2061 (Fig 2D). Subsequent analyses of transformation frequencies using the various Tn4001 derivatives generated in this study were compared to the frequency obtained using plasmid pISM2061. The plasmids carrying the transposon derivatives cannot replicate
Figure 2. Placement of unique BamHI and Smal RE sites in an IS256 arm of transposon Tn4001 (A) Removal of BamHI and Smal RE sites from plasmid pKS+ by digestion with XbaI and Smal followed by filling in with Klenow fragment of DNA polymerase I and ligation to circularize, generating plasmid pISM2046. (B) Insertion of the 1.3-kb PstI-HindIII fragment from plasmid pISM1001 containing IS256 into plasmid pISM2046. A 1.3 kb fragment from a PstI and HindIII digest of plasmid pISM1001 was isolated from a 0.6% agarose gel, purified using GeneClean, and ligated into plasmid pISM2046, generating plasmid pISM2048. (*) indicates that only the inserts of plasmids pISM2048 and pISM2048.5 are shown. (C) The Smal and BamHI RE sites were placed into IS256 using oligonucleotide-directed mutagenesis via PCR, generating plasmid pISM2048.5. (D) The 5.7 kb PstI-SalI fragment containing Tn4001 was ligated into plasmid pISM2046, generating plasmid pISM2061. (E) The mutant IS256 arm was used to replace the wild type arm. The 1.3-kb PstI-HindIII fragment from plasmid pISM2048.5 was inserted into pISM2061, generating plasmid pISM2062 (Tn4001mod). The double headed arrow spans the region encompassing Tn4001.
in mycoplasmas. Thus, the selectable phenotype was generated by transposition of the transposon from the plasmid to the chromosome following transformation.

The mutated IS256 arm was used to replace the wild type arm to generate Tn4001mod on plasmid pISM2062 (Figure 2E). The mutation was confirmed by sequence analysis. The transformation frequency of plasmid pISM2062 was similar to that of plasmid pISM2061 in Acholeplasma ISM1499 (Table 2). This suggested that the inverted repeat functions have not been altered by the mutation.

**Removal of the translational stop codons and the construction of Tn4001lac**

The 3.1-kb promoterless lacZ gene from pMC1871 (16) was cloned into the BamHI site of plasmid pISM2062 to generate transposon Tn4001.2062lac (Figure 3). Although this construct transformed Acholeplasma ISM1499, blue colonies did not appear on X-gal containing media. The lack of β-gal activity was expected because the IS256 sequence (15) shows that there are translational stop codons in all three reading frames contained within the inverted repeat (shown in bold letters in Figure 3). Thus, translation originating from adjacent mycoplasma chromosomal sequences would be terminated prior to lacZ.

To overcome premature translation termination, the translational stop codons were altered by inserting a pair of synthesized complimentary oligonucleotides that replaced the wild type IS256 inverted repeat region with an inverted repeat region possessing altered stop codons (Figure 3). The mutant adapter was designed to alter two or three of the stop codons. The mutant adapter was ligated into plasmid pISM2062 (Figure 3A) following PstI and BamHI digestions to generate transposons Tn4001.2062.2 (two stop codons altered) and Tn4001.2062.3 (all three stop codons altered). This resulted in the loss of 300 bp of unnecessary plasmid DNA between the PstI site of the cloning vector and the end of the IS256 arm. Alteration of the stop codons was confirmed by sequence analysis (data not shown).
Table 2. Transformation frequencies of Tn4001 derivatives in Acholeplasma

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Tn4001 Derivative</th>
<th>$\text{Tf}^a$</th>
<th>% Blue$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pISM2061</td>
<td>wild type</td>
<td>$2.0 \times 10^{-6}$ [1,3]</td>
<td>0 (0/2,796)</td>
</tr>
<tr>
<td>pISM2062</td>
<td>Tn4001 mod</td>
<td>$2.2 \times 10^{-6}$ [1,3]</td>
<td>0 (0/2,963)</td>
</tr>
<tr>
<td>pISM2062 lac</td>
<td>2062 lac</td>
<td>$4.0 \times 10^{-7}$ [2,3]</td>
<td>0 (0/599)</td>
</tr>
<tr>
<td>pISM2062.2</td>
<td>2062.2</td>
<td>$2.4 \times 10^{-6}$ [1,3]</td>
<td>0 (0/573)</td>
</tr>
<tr>
<td>pISM2062.3</td>
<td>2062.3</td>
<td>$6.0 \times 10^{-7}$ [1,4]</td>
<td>0 (0/573)</td>
</tr>
<tr>
<td>pISM2062.2 lac</td>
<td>2062.2 lac</td>
<td>$7.1 \times 10^{-7}$ [2,3]</td>
<td>7.7 (575/7,510)</td>
</tr>
<tr>
<td>pISM2062.2 lac(rev)</td>
<td>2062.2 lac(rev)</td>
<td>$9.6 \times 10^{-7}$ [2,3]</td>
<td>0 (0/1,572)</td>
</tr>
<tr>
<td>pISM2062.3 lac</td>
<td>2062.3 lac</td>
<td>$1.0 \times 10^{-7}$ [2,4]</td>
<td>4.1 (6/146)</td>
</tr>
<tr>
<td>pISM2062.3 lac(rev)</td>
<td>2062.3 lac(rev)</td>
<td>$1.2 \times 10^{-8}$ [2,4]</td>
<td>0 (0/10)</td>
</tr>
<tr>
<td>pISM2065</td>
<td>2065</td>
<td>$2.5 \times 10^{-8}$ [2,4]</td>
<td>8.8 (27/307)</td>
</tr>
</tbody>
</table>

$^a$ $\text{Tf} =$ transformation frequency. The values represent the average of at least three experiments. Transformation frequencies were defined as the number of Gm-resistant transformants divided by the number of viable cells used in the transformation mixture. The data were analyzed by the Student $t$ test to ascertain the effects of plasmid size, transposon size, and alteration of the inverted repeat sequence (stop codons) on transformation frequency. The standard error of the means = $1.89 \times 10^{-7}$. Significance group differences are indicated by [].

Plasmid size had a significant effect on transformation frequency; group 1 is significantly different from group 2 ($P<0.002$). Alteration of the third stop codon also had a significant effect on transformation frequency; group 3 is significantly different from group 4 ($P<0.001$). There were no significant differences for the effect of transposon size or alteration of the first two stop codons.

$^b$ Numbers in parentheses indicate the total number of blue colonies divided by the total number of white colonies.
Transposons Tn4001.2062.2 and Tn4001.2062.3 were introduced into Acholeplasma ISM1499, and the frequencies of transformation (Table 2) indicated that both functioned in Acholeplasma. The frequency of transformation of plasmid pISM2062.3 carrying transposon Tn4001.2062.3 was reduced by 4-fold.

The Tn4001lac promoter probe transposon was constructed by cloning the 3.1-kb promoterless lacZ gene from plasmid pMC1871 into transposons Tn4001.2062.2 and Tn4001.2062.3 to generate the corresponding Tn4001.2062.2lac and Tn4001.2062.3lac transposon derivatives. In addition, the lacZ was cloned in the reverse orientation to generate transposons Tn4001.2062.2lac(rev) and Tn4001.2062.3lac(rev). All of these Tn4001lac derivatives were able to transform Acholeplasma (Table 2). The transformation frequencies of plasmids pISM2062.2lac and pISM2062.2lac(rev) were 2- to 3-fold lower than that of the control plasmid, pISM2061. The transformation frequencies of plasmids pISM2062.3lac and pISM2062.3lac(rev) were 20- to 150-fold lower than pISM2061. Analysis of variance using the Student's t test showed a significant (P<0.002) effect of increasing the plasmid size (by the addition of the 3.1-kb lacZ fragment) on transformation frequency. Alteration of the codon nearest to the BamHI site in the IS256 inverted repeat also significantly (P<0.001) affected transformation frequency. There was no a significant difference for the effect of alteration of the first two stop codons on transformation frequency. It should be noted that pMB1-based plasmids do not replicate in mycoplasmas and only the transposition event can be observed. These experiments were unable to distinguish between the uptake of plasmid DNA (transformation) and transposition of Tn4001 into the chromosome. Thus, the combination of these two events are reported as transformation frequency in Table 2.
Figure 3. Removal of the translational stop codons from the inverted repeat of an IS256 element and construction of Tn4001lac. (A) Plasmid pISM2062 was digested with *PstI* and *BamHI* to liberate the inverted repeat of IS256 and 300 bp of intervening plasmid sequences. Relevant sequences are shown adjacent to the corresponding plasmid map. Bold letters indicate translational stop codons in IS256. Base changes are indicated by outlined letters. (B) Complimentary oligonucleotides were annealed by mixing at equal molar ratios, heating to 70°C for 3 min, and slowly cooling to room temperature. The mutant oligonucleotide adapter was then ligated into the *PstI* and *BamHI* sites of plasmid pISM2062, generating plasmids pISM2062.2 (two stop codons altered) and pISM2062.3 (three stop codons altered). (*) indicates that a 50/50 mixture of that base was used during synthesis of the oligonucleotides. (C) The 3.1-kb *BamHI* fragment containing the promoterless *lacZ* from plasmid pMC1871 was ligated into the *BamHI* sites of plasmids pISM2062.2 and pISM2062.3 generating plasmids pISM2062.2 lac and pISM2062.3 lac, respectively.
Analysis of mycoplasmal Lac fusions

The testing of the Tn4001 constructs and the analysis of β-gal activity was performed initially with Acholeplasma ISM1499. Expression of β-gal activity was observed with transposons Tn4001.2062.2lac and Tn4001.2062.3lac at a frequency of 8 and 4 percent of the Acholeplasma transformants, respectively (Table 2). Intensity of blue color on X-gal containing media (data not shown) and β-gal assays with this species showed that the transposon had inserted downstream of promoters of varying strengths. Levels of β-gal activity of three independent Tn4001.2062.2lac Acholeplasma transformants were 129, 276, and 1,020 units and levels of β-gal activity of three Tn4001.2062.3lac transformants were 89, 297, and 790 units. The level of β-gal activity measured with each transformant was very reproducible varying less than 5% between cultures and experiments. Expression of β-gal activity was not observed in constructs in which lacZ was placed in the reverse orientation (Table 2).

To examine the versatility of the Tn400Ilac derivative Tn4001.2062.2lac, M. gallisepticum was transformed with plasmid pISM2062.2lac and the transformants plated on X-gal-containing media. Approximately 3% of the transformants appeared to be Lac+ based on blue colony formation on X-gal-containing media. Levels of β-gal activity of four independent transformants giving blue colonies were 37, 44, 52, and 102 units. The level of β-gal activity of the white colonies was less than 1 unit.

Construction and use of the Tn400Ilac rescue transposons

Transposons Tn4001.2064 and Tn4001.2065 are derivatives of Tn4001.2062.2 and Tn4001.2062.3, respectively, containing plasmid pISM205 within the IS256 arm of Tn4001. Plasmid pISM205 is a 6-kb plasmid that was generated by cloning the 3.1-kb lacZ fragment from plasmid pMC1871 into plasmid pKS (Figure 4A). The BamHI site downstream of lacZ
Figure 4. Construction of promoter probe transposons designed to rescue adjacent chromosomal sequences. (A) Plasmid pISM2065 is a derivative of pISM2062.3, containing plasmid pISM205 within the IS256 arm of Tn4001. Plasmid pISM2065 was constructed by digesting plasmid pISM2062.3 with PvuII, isolation of the 5.7-kb fragment, circularization with T4 DNA ligase, digestion with BamHI, and ligating with BamHI digested plasmid pISM205. (B) The rescue system permits the recovery of adjacent chromosomal DNA by first digesting the chromosomal DNA with a unique enzyme in the multiple cloning site (e.g., PstI), followed by ligation and transformation into E. coli.
(A) 

- **ori**: Orientation of replication initiation
- **Ap**: Ampicillin resistance
- **p- lacZ**: Lacz promoter
- **Gm**: Gm resistance
- **HindIII**: Restriction enzyme
- **PstI**: Restriction enzyme
- **XhoI**: Restriction enzyme
- **SmaI**: Restriction enzyme
- **BamHI**: Restriction enzyme

(B) 

- **2065**: Genetic element
- **chr.**: Chromosome
- **Ap**: Ampicillin resistance
- **ori**: Orientation of replication initiation
- **Gm**: Gm resistance
- **p- lacZ**: Lacz promoter
- **PstI**: Restriction enzyme
- **SmaI**: Restriction enzyme
- **BamHI**: Restriction enzyme
- **HindIII**: Restriction enzyme
transcription was removed by partial digestion, fill-in, and ligation. Plasmids pISM2064 and pISM2065 were constructed by digesting plasmids pISM2062.2 and pISM2062.3, respectively, with PvuII, isolation of the 5.7-kb fragment, circularization with T4 DNA ligase, digestion with BamHI and ligating with the BamHI-digested plasmid pISM205. Orientation of the fragments in the resulting recombinant plasmid was confirmed by RE analysis, and the function of these transposons was confirmed by transformation into Acholeplasma. Plasmid pISM2065 transformed ISM1499 equally as well as pISM2062.3lac and pISM2062.3lac(rev). An analysis of variance for the effect of transposon size on transformation frequency showed the effect was not significant. Plasmid pISM2064 failed to transform ISM1499 despite confirmation by RE analysis that it was constructed properly.

A transposon rescue system was developed, similar to the Tn917lac rescue system (267), as a means of recovering mycoplasma chromosomal sequences containing promoters (Figure 4). The resulting Tn4001lac derivative, Tn4001.2065, possesses within one of its IS256 arms a plasmid capable of replication and selection in E. coli as well as the promoterless lacZ. Thus, mycoplasma chromosomal sequences that were able to drive β-gal expression can be recovered and cloned in E. coli as illustrated in Figure 4B. The rescued fragment can be sequenced directly by using the T3 or reverse primer that recognizes corresponding sequences within pKS. To test the functionality of this system, chromosomal DNA was obtained from different pISM2065 transformants, digested with PstI, diluted, ligated, and then transformed into E. coli. Recombinant plasmids obtained from these experiments contained chromosomal DNA sequences that drive β-gal expression in both Acholeplasma and E. coli. The transformation frequency by using pISM2065, however, was 2 logs lower than that of using the plasmid containing the wild type transposon.
lacZ fusion library construction and analysis

Because of the relatively low levels of β-gal activity resulting from the translational fusion system described above, a transcriptional lacZ fusion system (16) was examined. This system required that the DNA upstream of the truncated lacZ gene possess the translational start elements as well as the promoter to generate a β-gal fusion protein. Acholeplasma ISM1499 chromosomal DNA and vector pISM2050 (Figure 5) were digested with Sau3Al or BamHI, respectively, ligated together, and the ligation mixtures were transformed into E. coli DH5α. Approximately 20% of the E. coli transformants demonstrated the blue phenotype. Analysis of plasmid preparations from both blue and white colonies showed that all of the blue colonies contained mycoplasmal chromosomal DNA inserts while only 20-30% of the white colonies had plasmids with inserts. As a result, only plasmids from blue colonies were used for further study. Pools of plasmid DNA from five independent colonies were transformed into Acholeplasma ISM1520 where integration of the plasmid occurred within plasmid sequences previously introduced into the chromosome (132). If mycoplasma transformants from these plasmid pools demonstrated β-gal activity based on blue colony formation on X-gal-containing media, individual plasmids were used to transform ISM1520. Only 13 of 140 plasmid preparations giving promoter activity in E. coli were also positive for β-gal activity in Acholeplasma.

The thirteen lacZ fusion constructs that demonstrated β-gal activity in ISM1520 were also used to transform ISM1499 (Table 3). In this experiment, the region of homology for integration of the plasmid into the chromosome was the cloned chromosomal fragment. Only 9 of the 13 fusions were able to transform ISM1499. Constructs were confirmed by DNA-DNA hybridization analysis (data not shown). The lac fusion constructs were transformed into E. coli CSH50, a Δ(lac-pro) genotype, to simplify the interpretation of the β-gal assay data.
Figure 5. Construction of the promoter probe vector, pISM2050. The vector was constructed by ligating a 3.1-kb BamHI fragment containing the promoterless lacZ from plasmid pMC1871 into the BamHI site of plasmid pISM1003. The arrow indicates the direction of transcription of lacZ. The asterisk denotes the BamHI site that was inactivated. Abbreviations: Promoterless lacZ, p-lacZ; Ampicillin, Ap; Gentamicin, Gm; Tetracycline, Tc; and origin of replication, ori.
Table 3. Insert size\(^a\) and β-gal activity\(^b\) of lacZ fusion constructs in *Acholeplasma* and *E. coli*

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Insert Size (kb)</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ISM1520</td>
<td>ISM1499</td>
</tr>
<tr>
<td>pISM2050</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>pISM2050.1</td>
<td>2.8</td>
<td>17</td>
</tr>
<tr>
<td>pISM2050.2</td>
<td>0.9</td>
<td>1,428</td>
</tr>
<tr>
<td>pISM2050.8</td>
<td>1.0</td>
<td>48</td>
</tr>
<tr>
<td>pISM2050.18</td>
<td>0.9</td>
<td>48</td>
</tr>
<tr>
<td>pISM2050.19</td>
<td>1.5</td>
<td>375</td>
</tr>
<tr>
<td>pISM2050.25</td>
<td>0.8</td>
<td>211</td>
</tr>
<tr>
<td>pISM2050.39</td>
<td>1.0</td>
<td>172</td>
</tr>
<tr>
<td>pISM2050.40</td>
<td>1.7</td>
<td>37</td>
</tr>
<tr>
<td>pISM2050.66</td>
<td>2.0</td>
<td>35</td>
</tr>
<tr>
<td>pISM2050.69</td>
<td>2.5</td>
<td>131</td>
</tr>
<tr>
<td>pISM2050.70</td>
<td>2.3</td>
<td>186</td>
</tr>
<tr>
<td>pISM2050.86</td>
<td>2.2</td>
<td>39</td>
</tr>
<tr>
<td>pISM2050.89</td>
<td>1.0</td>
<td>73</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strain c</th>
<th>Insert Size (kb)</th>
<th>ISM1520</th>
<th>ISM1499</th>
<th>CSH50</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSH50</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>(\chi^{289})</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>5,382</td>
</tr>
</tbody>
</table>

\(^a\) NA = Not applicable

\(^b\) ND = Not done

\(^c\) *E. coli* strains CSH50 and \(\chi^{289}\) were fully induced with isopropyl β-D-thiogalactopyranoside.
Quantitation of Lac activity

*Acholeplasma* transformants containing *lacZ* fusions gave colonies of varying blue intensity on X-gal-containing media. Figure 6 shows a mixture of strains ISM2050.1 (very weak β-gal activity), ISM2050.2 (strong β-gal activity), and ISM2050 (vector only, no β-gal activity). β-gal assays were performed on the *lacZ* fusion constructs that were introduced into ISM1520, ISM1499, and CSH50 (Table 3). The results showed that levels of β-gal production varied by 100-fold between the various *lacZ* fusion constructs in *Acholeplasma*. Levels of β-gal activity generated by integration of the *lacZ* fusions in either ISM1520 and ISM1499 were similar except with pISM2050.66 and pISM2050.69. Recombinants demonstrating high levels of β-gal activity in *Acholeplasma* did not always generate high levels of activity in *E. coli*. In addition, pISM2050.89 generated more β-gal activity in *Acholeplasma* than in *E. coli*. Levels of β-gal activity of whole cells when the SDS lysis step was omitted were also examined with *Acholeplasma* strains ISM2050, ISM2050.2, ISM2050.25, and ISM2050.70. Their levels of β-gal activity were 3, 68, 19, and 27 units, respectively. The cell-free, ISM1499 cell, and ISM1520 cell controls in all experiments failed to generate any measurable β-gal activity.

Levels of the *lacZ* fusion transcripts were also measured for the ISM1520 strains harboring the *lacZ* fusion plasmids (Figure 7). These results correlated with the results of the β-gal assays in that strains demonstrating higher transcript levels generally had higher levels of β-gal activity. There were inconsistencies between β-gal and *lacZ* fusion transcript levels observed in *Acholeplasma* strains ISM2050.18, ISM2050.66, and ISM2050.69. These strains demonstrated higher transcript levels than the levels of β-gal assay would have predicted.

Immunoblot analysis was performed on several recombinant *Acholeplasma* strains to demonstrate the production of a β-gal fusion protein (Figure 8). *Acholeplasma* ISM1520 strains with high levels of β-gal activity reacted more strongly with the anti-β-gal monoclonal antibody supporting the results of the β-gal assay.
Figure 6. Acholeplasma ISM1520 transformants. Appearance of Acholeplasma ISM1520 transformant colonies grown on mycoplasma media containing 80 μg per ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). Colonies with large dark centers (i.e., blue color) are ISM2050.2, smaller dark centers, ISM2050.1, and lightest center are ISM2050 (vector only).
Figure 7. Quantitation of lacZ transcript levels of the lacZ fusion vectors in Acholeplasma ISM1520. Slot blot analysis of the mRNA levels from Acholeplasma ISM1499, ISM1520, ISM2050, ISM2050.1, ISM2050.2, ISM2050.8, ISM2050.18, ISM2050.19, ISM2050.25, ISM2050.39, ISM2050.40, ISM2050.66, ISM2050.69, ISM2050.70, ISM2050.86, and ISM2050.89 probed with either 16S rRNA (left) or lacZ-specific (right) gene probe. Plasmid pSP71 was included as a negative control. The 3.1-kb DNA fragment containing lacZ or the 1.4-kb fragment containing the 16S rRNA gene were used as positive controls. The values shown represent the counts in a defined volume after adjusting for the relative amounts of total RNA loaded into each slot as described in Figure 1.
<table>
<thead>
<tr>
<th>Strain</th>
<th>16S</th>
<th>lacZ</th>
<th>Counts/vol.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1499</td>
<td></td>
<td></td>
<td>66,517</td>
</tr>
<tr>
<td>1520</td>
<td></td>
<td></td>
<td>112,580</td>
</tr>
<tr>
<td>2050</td>
<td></td>
<td></td>
<td>77,006</td>
</tr>
<tr>
<td>2050.1</td>
<td></td>
<td></td>
<td>387,036</td>
</tr>
<tr>
<td>2050.2</td>
<td></td>
<td></td>
<td>1,941,175</td>
</tr>
<tr>
<td>2050.8</td>
<td></td>
<td></td>
<td>246,890</td>
</tr>
<tr>
<td>2050.18</td>
<td></td>
<td></td>
<td>595,954</td>
</tr>
<tr>
<td>2050.19</td>
<td></td>
<td></td>
<td>52,873</td>
</tr>
<tr>
<td>2050.25</td>
<td></td>
<td></td>
<td>661,166</td>
</tr>
<tr>
<td>2050.39</td>
<td></td>
<td></td>
<td>947,472</td>
</tr>
<tr>
<td>2050.40</td>
<td></td>
<td></td>
<td>66,904</td>
</tr>
<tr>
<td>2050.66</td>
<td></td>
<td></td>
<td>540,979</td>
</tr>
<tr>
<td>2050.69</td>
<td></td>
<td></td>
<td>803,629</td>
</tr>
<tr>
<td>2050.70</td>
<td></td>
<td></td>
<td>1,054,501</td>
</tr>
<tr>
<td>2050.86</td>
<td></td>
<td></td>
<td>312,903</td>
</tr>
<tr>
<td>2050.89</td>
<td></td>
<td></td>
<td>476,371</td>
</tr>
<tr>
<td>pSP71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16S/lacZ</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 8. Immunoblot analysis of the lacZ fusions in Acholeplasma. Proteins were detected with a monoclonal antibody to β-galactosidase. Lanes: A, ISM2050; B, ISM2050.70; C, ISM2050.69; D, ISM2050.40; E, ISM2050.39; F, ISM2050.25; G, ISM2050.19; H, ISM2050.2, I, ISM2050.1; and J, χ289, fully induced with isopropylthio-β-D-galactoside.
Mapping of *Acholeplasma ISM1499* promoters by primer extension

The transcriptional start sites for the promoters driving the expression of the *lacZ* in the fusion plasmids pISM2050.1, pISM2050.2, pISM2050.8, pISM2050.18, pISM2050.25, pISM2050.69, and pISM2050.70 were mapped by primer extension. The quality of the RNA was examined by first running 5 to 10 μg of RNA in a 1.2% agarose gel containing formaldehyde and then staining the gel with acridine orange. If the bands corresponding to the 23S rRNA, 16S rRNA and tRNA were sharp the RNA was considered to be free from degradation by nucleases and was used for the primer extension reaction. RNA-DNA hybridization analysis, using the 16S rRNA as a probe, was also performed to confirm that the 16S rRNA band was not smeared.

Initial experiments were conducted with a primer that was 21 bp in length that was reverse and complimentary to a region approximately 45 bp upstream of the *lacZ* fusion point. The $T_m$ of the oligonucleotide was determined to be 38.5°C. As a result, initial reactions gave the appearance of false-priming due to the appearance of many primer extension products. An oligonucleotide of 30 bp (30-mer) was then constructed that has a $T_m$ of 85°C to reduce the amount of primer extension products. The use of this 30-mer did reduce the number of background bands detected in the reaction. It should be noted that the appearance of false-priming could be observed if the unbound $^{32}$P was not first removed by precipitating the primer 3 times with ethanol.

Preliminary studies showed that the primer extension product signals were weak for all reactions excepts for those involving ISM2050.2. Because ISM2050.2 generated the highest level of *lacZ* fusion transcript (Figure 7), it was reasoned that the primer extension signal could be increased by increasing the amount of target transcript. Thus, the effects of increasing the total RNA up to 100 μg per reaction was examined. Reactions that used more than 50 μg of
total RNA either did not generated any bands or yielded high levels of background as demonstrated by the appearance of many distinct bands. The reactions that failed to generate any primer extension products were probably due to the inability to completely resuspend 100 μg of RNA in 10 μl of H2O. Thirty to fifty micrograms of total RNA were found to be optimal for the reaction.

The transcriptional start site for the promoter driving the expression of lacZ in pISM2050.8 is shown in Figure 9. Each primer extension product was run adjacent to a sequencing ladder generated by using the sequencing primer with plasmid pISM2083 (containing the insert of pISM2050.2 and the adjacent 800 bp of the lacZ). The transcriptional start site was determined by first counting the number of bases between the lacZ fusion point on the ladder to the primer extension product. The number of bases was then compared to the sequence of the pISM2050.8 insert-lacZ fusion region. All promoters were mapped in this manner except for the promoter contained within the pISM2050.2 insert. This promoter was mapped by directly comparing the primer extension product with the pISM2050.2 insert-lacZ fusion sequence contained in pISM2083.

**Exo/Mung deletion analysis of plasmid pISM2050.2**

Plasmid pISM2050.2, which has a 900-bp chromosomal DNA fragment from ISM1499, was shown to generate high levels of β-gal activity when introduced into Acholeplasma ISM1520 (Table 3). Exo/Mung deletion analysis was performed on plasmid pISM2050.2 to determine the minimum upstream region required to drive the expression of lacZ. Following exonuclease III and Mung Bean nuclease digestion, the products were ligated and used to transform E. coli DH5α. Both blue and white colonies were obtained from the 3 min Exonuclease III reaction mixture. Five blue and five white colonies were selected and plasmid DNAs prepared. Restriction enzyme analysis showed that the blue colonies possessed plasmids that were larger than those from the white colonies. Plasmid DNA was prepared with five of
Figure 9. Primer extension analysis of the transcriptional start site of the *lacZ* fusion transcript of ISM2050.8. The primer extension product (indicated by the arrow) was run against a sequencing ladder generated with pISM2083. Lanes G = guanine, A = adenine, T = thymine, C = cytosine, and 1 = the primer extension product from ISM2050.8.
these clones and the DNA was used to transform ISM1520. *Acholeplasma* ISM1520 transformed with plasmid pISM2075 demonstrated no β-gal activity by either a blue phenotype on X-gal-containing media or by β-gal assay. Transformants possessing the other plasmids demonstrated blue colony formation on X-gal containing media and β-gal assay. The levels of β-gal activity of these transformants ranged from 861 to 1,080 units. All of the plasmids in the transformants generating the Lac⁺ phenotype mapped to within 100 bp of each other (Figure 10). The deletion of the Lac⁻ transformant included the entire 900-bp insert as well as part of the lacZ. The mutation endpoints were mapped by sequence analysis. The sequence downstream of the mutation endpoints represent the insert DNA remaining following the Exonuclease III/Mung Bean digestions. The transformants generated by using the deletion plasmids from the 0, 1, and 2 min Exonuclease III digestions all gave the Lac⁺ phenotype on X-gal-containing media.

**Preliminary determination of an *Acholeplasma* ISM1499 consensus promoter sequence**

The transcriptional start sites and upstream regions the lacZ fusion transcripts for seven of the ISM2050 derivatives are shown in Table 4. The upstream regions were aligned at the putative sequences encompassing the -10 and -35 promoter regions. The regions were defined based on their similarity to the *E. coli* consensus promoter (84). A consensus promoter sequence was derived based upon the frequency with which a particular base appeared in the same position in the -10 and -35 regions. Also shown in Table 4 are the putative -10 and -35 regions reported for the *M. pneumoniae* P1 operon (95), *M. capricolum* rRNA P2, rRNBl (160) and the *E. coli* consensus promoters. The average number of bases that align between the ISM1499 promoter regions and the *E. coli* consensus promoter regions are 4.4 bp for the -10 region and 3 bp for the -35 region. Moreover, the spatial relationships between the
Figure 10. Exo/Mung deletion analysis of the ISM1499 chromosomal DNA region of ISM2050.2. Arrows indicate the point where deletion ends. The sequence to the right of the arrows represent the downstream region remaining. The bases to the left of the arrows have been deleted for each plasmid. The putative -35 and -10 regions of the promoter, ribosomal binding site, translational start site, chromosomal DNA -lacZ fusion point (Sau3AI site) are underlined. An asterisk indicates the translational start site. The lacZ region is denoted in bold.
Table 4. Alignment and determination of the putative -10 and -35 promoter regions of the promoters driving the expression of \textit{lacZ} in the plasmid pISM2050 derivatives

<table>
<thead>
<tr>
<th>Source of Promoter</th>
<th>-35</th>
<th>-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>pISM2050.1\textsuperscript{a}</td>
<td>TCACTGCACC--TGAAGAAACTACATAAATATAAAACTTTGAAGAA</td>
<td></td>
</tr>
<tr>
<td>pISM2050.2</td>
<td>TATCTTCCCTT--TAATTACTATTATAGTATAATATAAAAAGT</td>
<td></td>
</tr>
<tr>
<td>pISM2050.8</td>
<td>ACTTACGCTACAATCTAAAACCCAAATGGTCAAGTACATGCTGA</td>
<td></td>
</tr>
<tr>
<td>pISM2050.18</td>
<td>AGAATTCTGCTTTAATCCTTTTATATTAAAATAATCCATCA</td>
<td></td>
</tr>
<tr>
<td>pISM2050.25</td>
<td>ATGTTGATGTATTACATACATGATATAATATAGGG</td>
<td></td>
</tr>
<tr>
<td>pISM2050.69</td>
<td>TTAGATAATAGATCCCGTTTACACTCATGATATAATATAGGTAGG</td>
<td></td>
</tr>
<tr>
<td>pISM2050.70</td>
<td>TGGGTTATGATAATGAATTTACATGATATAATATAGGGACATAAC</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{b}Consensus sequence of the promoters driving the expression of \textit{lacZ} in the pISM2050 fusion derivatives. Upper case letters in the consensus sequence denote that at least 5 of the 7 sequences contained that base in that position. Lower case letters indicate that at least 3 of the 7 sequences contained that base in that position. \textit{W} = \textit{A} or \textit{T}, \textit{n} = any base.

\textsuperscript{a}The sequences of the upstream regions of \textit{lacZ} fusion constructs are aligned to the putative -10 and -35 regions based on their similarity to the consensus \textit{E. coli} promoter sequence. Bold letters indicate transcriptional start sites. \textit{A} = adenine, \textit{C} = cytosine, \textit{G} = guanine, and \textit{T} = thymine.

\textsuperscript{c}The \textit{M. pneumoniae} \textit{P1} operon and \textit{M. capricolum} \textit{rrn} \textit{A} and \textit{rrnB1} are conserved regions in the promoter regions.

\textsuperscript{d}Consensus sequence of the promoters driving the expression of \textit{lacZ} in the pISM2050 fusion derivatives. Upper case letters in the consensus sequence denote that at least 5 of the 7 sequences contained that base in that position. Lower case letters indicate that at least 3 of the 7 sequences contained that base in that position. \textit{W} = \textit{A} or \textit{T}, \textit{n} = any base.
The sequence for the *M. pneumoniae* PI operon promoter was reported by Inamine et al. (95).

The sequences for the *M. capricolum* *rrnA* P2 and *rrnB1* promoters were reported by Muto et al. (160).

The sequence for the *E. coli* consensus promoter was reported by Hawley and McClure (84).

Transcriptional start and the -10 region as well as between the -10 region and the -35 region were similar to the *E. coli* consensus promoter.

The acholeplasma consensus -10 and -35 promoter regions were also similar to the *M. pneumoniae* PI, *M. capricolum* *rrnA* P2 and *rrnB1* promoters. The distance between the -35 and -10 regions of the *M. pneumoniae* PI promoter was least 6 bp less than the acholeplasma promoters.
DISCUSSION

The ability to examine gene regulation of mycoplasmas has been hampered by the lack of a suitable genetic system to adequately assess the influence of their gene regulatory elements (67, 170). Because of the alternative codon usage (94, 265) and the potential differences between regulatory proteins of mycoplasmas and other bacterial cloning hosts, it is important that mycoplasma gene regulation be studied in its natural host. This laboratory has shown that the transposon Tn4001 will transpose from suicide delivery vectors to the chromosome of mycoplasmas (133). In addition, a plasmid-based cloning system which relies on the ability to introduce foreign DNA into the chromosome via homologous recombination with plasmid sequences previously introduced into the chromosome was developed in this laboratory (132, 134). Thus, the aims of the studies reported in this dissertation were to expand the mycoplasmal cloning systems developed in this laboratory to include a useful reporter gene for the examination of mycoplasma gene regulatory elements in vivo. Once the promoter probe vectors were developed, their ability to locate mycoplasma promoters was examined.

*M. capricolum* Promoter *trp*-lacZYA Fusion Studies

Because the *E. coli lacZ* gene has been shown to be an effective reporter gene in other heterologous bacterial backgrounds (9, 113, 229, 268), initial studies were designed to examine whether lacZ could also be used as a reporter gene in mycoplasmas using defined promoters to drive the expression of a *trp*-lacZ operon fusion. The generation of high levels of β-gal in *E. coli* using the *M. capricolum rRNA* P2 promoter showed that the *E. coli* RNA polymerase recognized *M. capricolum* promoter sequences. In contrast, the *E. coli rRNA* P1 promoter was not recognized by the *Acholeplasma* RNA polymerase as evidenced by the lack of mRNA produced (Figure 1B). Although the -35 and -10 regions of these promoters were
similar (67, 87), the *Acholeplasma* RNA polymerase was more stringent supporting previous studies using in vitro transcription assays (87).

Low levels of β-gal activity in *Acholeplasma* were observed using the *M. capricolum* promoter in a trp'-lacZ operon fusion (Figure 1A). This could be accounted for, in part, by the single copy insertion into the chromosome or by poor recognition of the *M. capricolum* promoter by the *Acholeplasma* RNA polymerase. A more likely explanation, however, is that the translation initiation region of the *E. coli* trp gene promoter was not efficiently recognized in *Acholeplasma*. The trp'-lacZ mRNA transcript levels in ISM2005 and ISM2010 are relatively high (Figure 1B) despite the low levels of β-gal being produced (Figure 1A). 16S rRNA sequence homology suggests that mollicutes are more closely related to Gram-positive organisms than Gram-negative organisms (257). Thus, the differences in translational specificity that has been demonstrated between the Gram-negative and Gram-positive bacteria (81) appears to include mycoplasmas as well. It was clear from these initial studies, however, that lacZ could be used as a reporter gene in mycoplasmas.

**Tn4001lac Promoter Probe Studies**

The primary goal of these studies was to develop a transposon-based promoter probe vector for use in mycoplasmas. Because our laboratory had previously shown that Tn4001 could be introduced into mycoplasmas from a plasmid delivery vector (133), a study was undertaken to develop a Tn4001lac promoter probe transposon. In addition, a means of rescuing the promoter-containing sequences in the mycoplasma chromosome that drive the expression of the lacZ in the transposon was pursued.

The construction of the Tn4001lac promoter probe vector was initiated by inserting unique BamHI and Smal RE sites into an IS256 arm of Tn4001, and then inserting the lacZ gene into one of these RE sites. The location of these RE sites in the IS256 element of Tn4001 was an
important consideration. The right and left IS256 elements were shown to possess identical 1,324-bp sequences (15). Therefore, the probability that only one of the arms encodes the functional transposase, as shown with Tn5 (208), was unlikely. The IS256 element has 26-bp terminal inverted repeats in which 17 bp match. In order to avoid interrupting the terminal inverted repeats, the BamHI and SmaI RE sites were placed 28 bp from the end of the IS256 element to generate plasmid pISM2062 (Figure 2).

The insertion of the BamHI and SmaI RE sites, however, could have affected the transposition activity of Tn4001 as a consequence of interrupting the promoter for the transposase. The mutation resulted in the insertion of an additional 11 bp between the -10 and -35 region of the predicted transposase promoter (15). There are two possible reasons that transformation frequency was not affected by the mutation. First, the change in the spatial relationship between the -10 and -35 region was tolerated and transcription still proceeded. This is unlikely, however, because it has been shown that changing the distance between the -10 and -35 region affects promoter strength (84). The second explanation is that the transposase produced from the unaltered IS256 element was able to provide the necessary transposition capabilities. It has been shown in Tn5 and TnJ0 that the transposase was able to complement in cis when the inverted repeat regions were relatively close to the transposase gene (152, 174).

The 3.1-kb lacZ fragment was inserted into pISM2062 to generate plasmid pISM2062lac. Although this construct could transform Acholeplasma ISM1499, no blue colonies appeared on X-gal-containing media. This was not unexpected because examination of the sequence between the end of the IS256 element and the lacZ fragment showed translational stop codons in all three reading frames (Figure 3) (15). Alteration of the stop codon nearest to the BamHI RE site resulted in a significant ($P<0.001$) decrease in transformation frequency as compared to the wild type (Table 2). There was not a significant effect of altering the other two codons as compared to the wild type.
The decrease in transformation frequency as a result of altering the codon nearest to the BamHI RE site may be explained in two ways. First, the site around this stop codon has the greatest base homology between the inside and outside terminal inverted repeat regions suggesting that this site may be involved in transposase binding or recognition. Byrne et al. (15) have postulated that the transposase may recognize sequences common to both of the terminal inverted repeat regions. Second, this stop codon also lies within the predicted -35 region of the transposase (15). Thus, the reduction in transposition as measured by its effect on transformation frequency could either be a result of altering the transposase binding site or decreasing the promoter strength of the transposase.

The addition of the 3.1-kb lacZ gene to plasmids pISM2062, pISM2062.2, or pISM2062.3 (generating plasmids pISM2062lac, pISM2062.2lac and pISM2062.3lac) also resulted in a significant (P<0.002) decrease (3- to 7-fold) in transformation frequency. This could have been due to an increase in plasmid size, or the placement of the lacZ gene in the IS256 arm could have affected stability of the transposon. In a previous study with an integrative plasmid system in mycoplasmas (134), it was shown that larger plasmids have lower transformation frequencies than smaller plasmids.

In addition, Byrne et al. (15) have described a region of dyad symmetry encompassing the transposase promoter and ribosomal binding site able to form a stem-loop structure. Placement of the lacZ gene into the IS256 arm would have either eliminated the formation of the stem-loop structure or would have reduced its stability. The formation of the stem-loop structure is thought to be involved in protecting the transposon from the influence of adjacent active promoter sequences (15, 38, 112). The decreased transformation frequencies resulting from the placement of the lacZ fragment into the IS256 element may suggest that the stem-loop structure has a role in stabilizing the transposon in the chromosome.

The effects of increased plasmid size and altering the codon nearest to the BamHI site appear to be independent. The additive effect of the modification of the third stop codon and the
increase in plasmid size is exemplified by the 50- to 100-fold difference in transformation frequencies between plasmids pISM2061, pISM2062, or pISM2062.2 and pISM2062.3lac, pISM2062.3lac(rev), or pISM2065 (Table 2). The insertion of lacZ or the altering of the third stop codon only resulted in a 5- to 10-fold decrease in transformation frequency.

A Student's t test was conducted to examine for significant effects of transposon size on transformation frequency. This was done by comparing the transformation frequencies between pISM2065 and pISM2062.3lac and pISM2062.3lac(rev). All three of these plasmids possess an IS256 element with all three stop codons altered and are similar in size. However, pISM2065 has the 3-kb plasmid inserted into the IS256 arm. The results showed that transposon size did not significantly affect transformation frequency.

β-galactosidase activity was observed in 7.7, 4.1, and 8.8 percent of the Acholeplasma ISM1499 transformants with transposons Tn4001.2062.2lac and Tn4001.2062.3lac, and Tn4001.2065, respectively (Table 2). Blue color on X-gal-containing media (data not shown) and β-gal assays indicated that the transposons had inserted in regions of active transcription in the chromosome. Varying levels of β-gal activity were measured, however. This indicated that either the transposon had inserted downstream of promoters of varying strengths, the stability of the lacZ fusion transcript differed, or stability of the protein was altered. β-galactosidase activity was not observed in constructs in which lacZ was placed in the reverse orientation showing that the promoters driving the expression of lacZ lie outside the transposon. The versatility of the Tn4001lac was tested by using plasmid pISM2062.2lac to transform M. gallisepticum. The recovery of Lac+ transformants demonstrate that this transposon would be effective in locating promoters of other species of mycoplasmas.

Another aim of these studies was to develop a means of rescuing sequences adjacent to the transposon that are driving the expression of lacZ contained in the transposon. A transposon rescue system modeled after the Tn917lac rescue system developed by Youngman et al. (267) was developed with Tn4001 (Figure 4). The resulting Tn4001lac derivative, Tn4001.2065,
possesses within one of its IS256 arms a plasmid capable of replication and selection in *E. coli* as well as the promoterless *lacZ*. Thus, mycoplasma chromosomal sequences that are able to drive β-gal expression can be recovered and cloned in *E. coli* as illustrated in Figure 4B. Acholeplasma chromosomal sequences have been successfully rescued using this system.

Plasmid pISM2065 has a major drawback in that it transforms *Acholeplasma* ISM1499 at a very low frequency (10⁻⁸ transformants/CFU) probably due to the altered third stop codon. Therefore, only 10 to 50 transformants can be expected to be recovered per transformation. Unfortunately, plasmid pISM2064 did not transform ISM1499. Based upon the transformation frequency results with its parent plasmid pISM2062.2, a 10-fold increase in transformation frequency would have been expected with pISM2064. Plasmids pISM2064 and pISM2065 have identical RE patterns, yet pISM2064 was not functional. The inability to transform ISM1499 with plasmid pISM2064 cannot be explained at this time.

In summary, a Tn4001lac promoter probe vector for use in mycoplasmas has been developed. Placement of the BamHI and *SmaI* RE sites into the IS256 arm of Tn4001 did not effect transformation frequency. Removal of all three stop codons and increasing the plasmid size appeared to reduce the transformation frequency. Nevertheless, these Tn4001lac derivatives are suitable for the detection and analysis of mycoplasma promoters and/or translational start elements in vivo. Moreover, the ability to rescue adjacent chromosomal sequences driving the expression of *lacZ* using transposon Tn4001.2065 should facilitate the rapid analysis of mycoplasma regulatory sequences.

**Plasmid pISM2050 Promoter Probe Studies**

**Construction and use of the plasmid promoter probe vector pISM2050**

A transcriptional fusion vector was constructed (Figure 5) in hopes of increasing β-gal activity. This vector required both the transcriptional and translational sequences cloned
upstream of the promoterless lacZ for the expression of β-gal. Thus, unlike the M. capricolum studies described above, both the transcriptional and translational start regions were from Acholeplasma. A library of cloned fragments was produced and screened in E. coli prior to introduction into Acholeplasma because of the large quantities of plasmid DNA (7 to 10 μg) required for transformation into mycoplasmas (134). Interestingly, only 10% of the recombinants that demonstrated promoter activity in E. coli also showed activity in ISM1520. This may be explained in two ways. First, pooled plasmid preparations were used to initially transform ISM1520, and it is possible that in some of the pooled preparations, promoter-containing constructs were under represented, did not transform, and thus, all transformed colonies would appear to have the Lac− phenotype. In an attempt to prevent this, it was assumed that each plasmid of the pool would be represented in the transformed colonies if more than 50 transformants arose from the plasmid mixture. Second, the A + T-rich mycoplasma chromosome may have generated pseudo-promoter activity in E. coli. This is supported by similar studies in another A + T-rich organism, S. pneumoniae. Chen and Morrison observed more strong promoter-acting sequences with randomly cloned S. pneumoniae DNA fragments in E. coli than with randomly cloned E. coli DNA fragments (27). It is clear from these studies that random mycoplasma sequences can result in false promoter activity in E. coli thereby demonstrating the importance of examining gene regulation in the original mycoplasma host.

The lacZ fusion constructs were introduced into two different Acholeplasma chromosomal locations to examine whether location would affect β-gal activity levels. Integration was targeted to a single chromosomal site in ISM1520 or to the original location of the cloned fragment in ISM1499. The levels of β-gal activity were similar whether the plasmid constructs were introduced into ISM1520 or into ISM1499 except for plasmids pISM2050.66 and pISM2050.69 (Table 3). The 3- to 5-fold change in levels of β-gal activity in these constructs
suggests that chromosomal location may influence gene expression. The inability to transform ISM1499 with plasmids pISM2050.2, pISM2050.8, pISM2050.86, and pISM2050.89 despite repeated attempts, suggests that the integration of these plasmids into the chromosome at their normal location may be lethal.

The difference in levels of β-gal between *Acholeplasma* and *E. coli* may be due to the lacZ fusion construct copy number. In *E. coli*, the fusion construct exists as a high copy number plasmid while in *Acholeplasma* the plasmid is integrated into the chromosome. The higher levels of β-gal activity generated by the integration of plasmid pISM2050.89 in *Acholeplasma* than in *E. coli* was unexpected. This may suggest that there is more than one class of promoters in mycoplasmas. Hudson and Stewart (90) have isolated 2 classes of *S. aureus* promoters; one that directs chloramphenicol acetyltransferase expression in both *E. coli* and *S. aureus*, and another that only directs expression in Gram positive hosts. Their results suggest that there is something unique about this second class of Gram positive bacterial promoters such as the need for a regulatory factor or recognition by a specific class of sigma factors of the RNA polymerase.

Transcript levels of the lacZ fusion products from the *Acholeplasma* ISM1520 strains transformed with the pISM2050 derivatives were measured and compared to levels of β-gal activity. Generally, strains demonstrating higher levels of β-gal activity (Table 3) also had higher transcript levels (Figure 7). There were some exceptions, however. Strains ISM2050.18, ISM2050.66, and ISM2050.69 had relatively low levels of β-gal activity compared to the relative level of transcript measured. This indicated that either the lacZ message in these strains was unstable or the message was not efficiently translated.

**Analysis of Acholeplasma ISM1499** promoters

The primer extension studies (Figure 9) showed that indeed the *Acholeplasma* ISM1499 chromosomal sequences adjacent to lacZ in the plasmid derivatives of pISM2050 contained the
promoters that were driving the expression of lacZ. Unfortunately, 6 of the 13 promoters could not be mapped. The promoters driving the expression of lacZ contained in ISM2050.19 and ISM2050.40 probably could not be mapped because of the low amount of lacZ fusion transcript produced in these strains (Figure 7). The reason(s) that the other transcriptional start sites could not be determined remains an enigma at this time. One possible reason is that the transcriptional start is too far upstream from the point in which the primer binds the message. Reverse transcriptase has a tendency to pause or stop in regions of high secondary structure and the enzyme is the most efficient when the 5' end of the mRNA is within 100 bp of the primer target region (213).

Exonuclease III/Mung Bean nuclease deletion analysis of the promoter driving lacZ expression in ISM2050.2 (Figure 10) supported the primer extension mapping studies in that the translational start mapped within the sequences deleted in pISM2075 (Lac̅) and present in pISM2076 (Lac⁺). Unfortunately, there were not any deletion mutations that ended between these two constructs to precisely confirm the minimal upstream distance required for transcriptional activity.

The sequences upstream of the transcriptional start sites were aligned (Table 4). The -10 and -35 regions were assigned based upon their similarity to the E. coli consensus sequence as reported by Hawley and McClure (84). Defining a mycoplasma promoter by its similarity to the E. coli consensus promoter could be misleading because mycoplasma DNA, like E. coli promoters, is A + T-rich. Therefore, it is especially important with mycoplasma promoters to correlate the sequence data with promoter mapping studies.

Alignment of the sequences upstream of the mapped transcriptional start sites suggested that the -10 regions were similar to the -10 region of the E. coli consensus promoter, but the -35 regions were more variable. The average of 4.4 bp conserved at the -10 region is similar to the 4.2 bp conservation observed with the E. coli and phage promoters examined by Hawley and McClure (84). There was only a 3 bp conservation at the -35 region of the acholeplasma
promoters which is about 1 bp less than the 3.9 bp observed with the *E. coli* and phage promoters. The observation that the -10 region was more like the *E. coli* consensus promoter than the -35 region is consistent with previous reports examining the putative mycoplasma promoter regions (29, 30, 95, 160).

The result that the *Acholeplasma* promoters were similar to the *E. coli* consensus promoter is not unexpected because the *lacZ* fusion constructs were initially screened in *E. coli* for their ability to give the Lac+ phenotype. Thus, only the *Acholeplasma* upstream regions that are recognized by *E. coli* RNA polymerase were considered for future analysis in *Acholeplasma*.

The putative -10 and -35 regions were also similar to the *E. coli* consensus promoter region with respect to distance from the transcriptional start site (6 to 8 bp) and the intervening distance between the regions (16 to 18 bp). The *Acholeplasma* promoter sequence was different from the putative *M. pneumoniae* PI promoter (95) with respect to the distance between the -10 and -35 regions (Table 4). The shorter intervening distance suggests that the PI gene promoter may be regulated differently; perhaps the promoter is recognized by a different sigma factor. In *E. coli*, the genes that respond to nitrogen starvation possess promoters in which the intervening distance is only 6 bp and are recognized by the rpoN gene (formally ntrA) product, σ^60^ (88).

Should mycoplasma promoter regions be defined by their similarity to the *E. coli* consensus sequence? Just because all mycoplasma promoters have been defined in this manner (160), does not validate this approach. It could be argued that if the parameters for locating a promoter region are limited to their similarity to the *E. coli* consensus sequence, then naturally only *E. coli* consensus-like sequences will be assigned as promoter regions.

The studies described above demonstrate that *Acholeplasma* promoter-containing DNA have *E. coli*-like promoter regions that may or may not function as RNA polymerase recognition sequences in *Acholeplasma*. Additional studies involving fine structure mapping, deletion analysis, and site-directed mutagenesis will be required to answer this question. The
ideal method in which to define a mycoplasma consensus promoter would be to align many upstream regions by their confirmed translational start sites and correlate that information with the fine structure mapping studies. Unfortunately, only a few mycoplasma transcriptional start sites have been actually mapped. Thus, as more mycoplasma upstream regions are confirmed, an analysis for a "true" consensus mycoplasma promoter sequence should be undertaken.

Summary of the plasmid pISM2050 promoter probe vector studies

In summary, these data demonstrate the need to be able to examine mycoplasma gene regulation in a homologous system. Clearly, *E. coli* was less stringent in its promoter recognition sequences than *Acholeplasma*, and therefore the use of *E. coli* in studying mycoplasmal promoter structure and function should be limited to an initial screening. These studies also showed that the promoter probe vector, pISM2050, can identify cloned mycoplasmal chromosomal fragments containing transcriptional and translational control sequences. The levels of gene expression from cloned fragments vary by 100-fold indicating that gene regulation is occurring at the transcriptional and/or translational level in *Acholeplasma*. Immunoblot analysis (Figure 8), measurement of message levels (Figure 7), and the intensity of blue color on X-gal-containing media (Figure 6) confirm this conclusion. Primer extension studies showed that the promoter driving the expression of *lacZ* was from acholeplasma. Moreover, the putative promoter regions resemble the *E. coli* consensus promoter at the -10 region but were more variable at the -35 region.

Summary

The integrative plasmid system lends itself to analysis of gene regulatory elements in mycoplasmas by offering the ability to identify gene regulatory elements, to introduce defined mutations in promoter sequences, and has the potential to identify and study transcriptional
regulators. These studies showed that *lacZ* could be used as a reporter gene in mycoplasmas. These studies also demonstrated the need to be able to examine mycoplasma gene regulation in its natural host, the mycoplasma, because of the pseudo-promoter activity in *E. coli*. In addition, mycoplasmas regulate levels of gene expression because differing levels of β-gal activity were generated between the various *lacZ* fusion constructs. The transposon studies showed that Tn4001lac could integrate and detect sites of active transcription in both *Acholeplasma* and *Mycoplasma*. Thus, this vector should be useful for the detection of promoters for a variety of different mycoplasmas. Finally, the regions upstream of 7 transcriptional start sites were sequenced, and the putative promoters were determined. The -10 regions of the *Acholeplasma* promoters were similar to the -10 region of the *E. coli* consensus promoter. The putative -35 regions of the *Acholeplasma* promoters were more variable.
REFERENCES


28. Chevalier, C., C. Saillard, and J. M. Bove. 1990. Organization and nucleotide sequences of the *Spiroplasma citri* genes for ribosomal protein S2, elongation factor Ts, spiralin,


R. N. McElhaney, L. R. Finch and J. B. Baseman (ed.), Mycoplasmas: molecular
biology and pathogenesis. American Society for Microbiology, Washington, D.C.

63. Fischer, R. S., B. E. Fischer, and R. A. Jensen. 1992. Sources of amino acids, p. 201-
209. In J. Maniloff, R. N. McElhaney, L. R. Finch and J. B. Baseman (ed.),
Mycoplasmas: molecular biology and pathogenesis. American Society for Microbiology,
Washington, D.C.

64. Fujita, M. Q., H. Yoshikawa, and N. Ogasawara. 1992. Structure of the dnaA and DnaA-
box region in Mycoplasma capricolum chromosome: conservation and variations in the

of Mycoplasma hominis tet(s) and Mycoplasma salivarium tet(s) for genetic transformation
to tetracycline resistance by DNA extracted from Mycoplasma hominis tet(r). J. Infect.
Dis. 139:444-451.

rifampin and characterization of spiroplasmal DNA-dependent RNA polymerase. J.

capricolum ribosomal RNA operons: identical activities but different regulation in
homologous and heterologous cells. Nucleic Acids Res. 16:61-76.


Mycoplasmas: molecular biology and pathogenesis. American Society for Microbiology, Washington, D.C.


Mycoplasma pneumoniae. 9th International Congress of the International Organization for Mycoplasmology 2:66.


Williams and Wilkins, Baltimore.


APPENDIX A: SEQUENCES UPSTREAM OF LACZ IN THE pISM2050 FUSION PLASMIDS
Figure A1. Sequence upstream of the \lacZ\ gene in pISM2050.1. The \lacZ- chromosomal DNA fusion is shown in bold (GATC). The putative -35 and -10 promoter regions are underlined and the translational start site is indicated by an (*). The potential ribosomal binding site (rbs) and translational start site (start) is also underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of *A. laidlawii*. 
Figure A2. Sequence upstream of the lacZ gene in pISM2050.2. The lacZ- chromosomal DNA fusion is shown in bold (GATC). The putative -35 and -10 promoter regions are underlined and the translational start site is indicated by an (*). The potential ribosomal binding site (rbs) and translational start site (start) is also underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of *A. laidlawii*. 
Figure A3. Sequence upstream of the lacZ gene in pISM2050.8. The lacZ-chromosomal DNA fusion is shown in bold (GATC). The putative -35 and -10 promoter regions are underlined and the translational start site is indicated by an (*). The potential ribosomal binding site (rbs) and translational start site (start) is also underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of *A. laidlawii*.
<table>
<thead>
<tr>
<th>Frame 2</th>
<th>Ser UKN Ile Thr Gln Pro Ile UKN UKN Glu Val Ser UKN UKN Ser Phe Glu</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AGT AGN ATC ACT CAA CCA ATA ATN NTT GAA GTC TCT NAA NCC AGT TTT GAA</td>
</tr>
<tr>
<td></td>
<td>10 19 28 37 46</td>
</tr>
</tbody>
</table>

| Tyr Val Asn Thr UKN UKN Ile UKN UKN Asp Ala Asn Ile Pro Arg Gly Ser Phe Tyr |
|-------------------------------|---------------------------------------------------------------|
| TAT GTG AAT ACN CNA NAC ATT ATN NAA GAT GCT AAT ATT CCA AGA GGG TCT TTT TAC |
| 61 70 79 88 97 106                                                |

| Gln Tyr Phe Glu Asp Lys UKN Asp Met Tyr Glu Tyr Ile Met Asp Tyr Ile Ser Ser |
|-------------------------------|---------------------------------------------------------------|
| CAG TAC TTT GAA GAT AAG NCG GAT ATG TAT GAA TAT ATG ATG TAT ATT AGT TCA |
| 118 127 136 145 154 163                                                |

| Ile Lys Arg Tyr Phe Lys Ser Ile Phe Glu Ala Val Asn Leu Asn Phe Ile Glu |
|-------------------------------|---------------------------------------------------------------|
| ATA AAA AGA TAT TAT TTT AAA AGT ATA TTT GAA GCA GTG AAT CTG AAT TTT ATA GAG |
| 175 184 193 202 211 220                                                |

| Arg Ile Glu Ala Ile Tyr Leu Ala Gly Val Lys Phe Lys Ser Glu Asn Pro Asp Phe |
|-------------------------------|---------------------------------------------------------------|
| CGA ATA GAG GCA ATT TAT TTA GCG GGT GTA AAA TTT AAG TCC GAG AAC CCT GAT TTT |
| 232 241 250 259 268 277                                                |

<table>
<thead>
<tr>
<th>-35 -10 *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val Arg Ala Gly Glu Phe Met Leu Lys Ser Pro Leu Tyr Leu Asn Asn Pro Ser Val</td>
</tr>
<tr>
<td>GTA AGA GCA GGA GAA TTC ATG CTT AAA TCA CCT TTA ATT ATT AAT AAT CCT GCA GTA</td>
</tr>
<tr>
<td>289 298 307 316 325 334</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>rbs</th>
<th>start</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKN Lys Gly Leu Glu Gln Met Ile Ser Ile Tyr Glu Ser Trp Ile Ile Asn Asp Pro</td>
<td></td>
</tr>
<tr>
<td>NCC AAA GGT TTA GAA CAA ATG ATT TCA ATC TAC GAG TCT TGG ATT ATC AAT GAT CCC</td>
<td></td>
</tr>
<tr>
<td>346 355 364 373 382 391</td>
<td></td>
</tr>
</tbody>
</table>

Figure A4. Sequence upstream of the *lacZ* gene in pISM2050.18. The *lacZ*- chromosomal DNA fusion is shown in bold (GATC). The putative -35 and -10 promoter regions are underlined and the translational start site is indicated by an (*). The potential ribosomal binding site (rbs) and translational start site (start) is also underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of *A. laidlawii*. 
Figure A5. Sequence upstream of the lacZ gene in plSM2050.19. The lacZ- chromosomal DNA fusion is shown in bold (GATC). The potential ribosomal binding site (rbs) and translational start site (start) is underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of A. laidlawii.
Figure A6. Sequence upstream of the lacZ gene in pISM2050.25. The lacZ-chromosomal DNA fusion is shown in bold (GATC). The putative -35 and -10 promoter regions are underlined and the translational start site is indicated by an (*). The potential ribosomal binding site (rbs) and translational start site (start) is also underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of A. laidlawii.
Figure A7. Sequence upstream of the lacZ gene in pISM2050.39. The lacZ-chromosomal DNA fusion is shown in bold (GATC). The potential ribosomal binding site (rbs) and translational start site (start) is underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of A. laidlawii.
Figure A8. Sequence upstream of the lacZ gene in pISM2050.40. The lacZ-chromosomal DNA fusion is shown in bold (GATC). The potential ribosomal binding site (rbs) and translational start site (start) is underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of *A. laidlawii*. 
Figure A9. Sequence upstream of the lacZ gene in pISM2050.66. The lacZ-chromosomal DNA fusion is shown in bold (GATC). The potential ribosomal binding site (rbs) and translational start site (start) is underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of A. laidlawii.
Figure A10. Sequence upstream of the lacZ gene in pISM2050.69. The lacZ-chromosomal DNA fusion is shown in bold (GATC). The putative -35 and -10 promoter regions are underlined and the translational start site is indicated by an (*). The potential ribosomal binding site (rbs) and translational start site (start) is also underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of A. laidlawii.
Figure A11. Sequence upstream of the lacZ gene in pISM2050.70. The lacZ-chromosomal DNA fusion is shown in bold (GATC). The putative -35 and -10 promoter regions are underlined and the translational start site is indicated by an (*). A potential ribosomal binding site was not found.
Figure A12. Sequence upstream of the lacZ gene in pISM2050.86. The lacZ-chromosomal DNA fusion is shown in bold (GATC). The potential ribosomal binding site (rbs) and translational start site (start) is underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of A. laidlawii.
Figure A13. Sequence upstream of the *lacZ* gene in pISM2050.89. The *lacZ*-chromosomal DNA fusion is shown in bold (GATC). The potential ribosomal binding site (rbs) and translational start site (start) is underlined. The potential ribosomal binding site was determined by aligning the sequence with the last 15 bp at the 3' end of the 16S rRNA gene of *A. laidlawii*. 
APPENDIX B: IMMUNE RESPONSE OF A FOREIGN GENE EXPRESSED IN ACHOLEPLASMA ISM1499

Rationale and Experimental Design

The goal of this experiment was to determine whether a foreign gene expressed in mycoplasmas could induce a specific immune response. The overall goal of my research has been to develop an expression cloning system for use in mycoplasmas. A consensus promoter sequence was to be derived from the promoter probe studies and an expression vector for use in mycoplasmas was to be generated. Finally, the expression cloning system was to be tested by placing a foreign gene downstream of the consensus promoter and determine whether an immune response could be induced to that gene upon challenge. Because the promoter probe studies showed that the E. coli lacZ was expressed in Acholeplasma, an experiment to determine whether mice would mount an immune response to β-galactosidase was conducted.

Twenty specific pathogen free HE Ouj mice that were 4 months of age were used in the study. They were grouped as shown in Table B1. Following a prebleed, each mouse was injected i.p. with 100 μg of whole cells (killed by repeated freezing and thawing) resuspended in incomplete Freund's adjuvant. After 7 days, blood was obtained via an orbital puncture. The mice were boosted with another 100 μg of whole cells resuspended in incomplete Freund's adjuvant 9 days following the primary challenge. Five days later blood was obtained via an orbital puncture. Serum was separated from the RBCs by centrifugation in a microfuge and the samples were frozen at -20°C until they were assayed for anti-β-galactosidase activity by ELISA.

The ELISA was performed by first coating a 96-well flat bottom plate with 100 μl of 10 μg/ml β-galactosidase suspended in 0.1M Na2CO3, pH 9.6. Following a blocking step overnight with PBS containing 5% BSA, 100 μl of the test serum diluted 1:250 was added to
each well. After a 5 h incubation at room temperature, the wells were washed with 0.9% NaCl containing 0.05% Tween-20. One hundred microliters of the conjugate, horse radish peroxidase labeled goat anti-mouse Ig, diluted 1:1,000 were added to each well. Following a 4 h incubation at room temperature, 100 µl of the TMB substrate was added. The reaction was stopped by adding 100 µl of 1 M phosphoric acid, and the plate was read with a microtiter plate reader at an absorbance of 415 nm. The positive control was a monoclonal antibody to β-galactosidase diluted 1:100. The negative controls for the ELISA assay included wells in which either no Ag (β-galactosidase), test antisera, or conjugate was added. All samples and controls were performed in triplicate.

Table B1. Experimental design of the *Acholeplasma* challenge study

<table>
<thead>
<tr>
<th>Group</th>
<th>Na</th>
<th>Sex</th>
<th>β-gal activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>3</td>
<td>F</td>
<td>n.a.</td>
</tr>
<tr>
<td>ISM2050.1</td>
<td>5</td>
<td>M</td>
<td>19</td>
</tr>
<tr>
<td>ISM2050.2</td>
<td>4</td>
<td>M</td>
<td>1,193</td>
</tr>
<tr>
<td>ISM2050.70</td>
<td>4</td>
<td>M</td>
<td>201</td>
</tr>
<tr>
<td>ISM2050</td>
<td>4</td>
<td>F</td>
<td>1</td>
</tr>
</tbody>
</table>

aN = number of mice.

bUnits of β-gal activity defined as the number of monomers of the enzyme per CFU generated in one min. (See Materials and Methods of dissertation.)
Results and Discussion

The results of the ELISA assay are shown in Table B2. The values shown in Table B2 represent the average absorbance values of the triplicate measurements for each mouse as well as the average for all mice within a group. The results show that a low level antibody response to β-galactosidase as indicated by absorbance levels was already present in these mice and introduction of more antigen in the form of a recombinant bacteria did not induce a greater antibody response. The ELISA controls worked as expected. The absorbance values of the positive control (Mab to β-galactosidase) were not remarkably high but were nevertheless positive. There were not any differences in antibody levels between the negative control groups (PBS and ISM2050) and the groups (ISM2050.2 and ISM2050.70) that would have been predicted to induce the highest levels of antibody response. Moreover, there was not an increase in antibody response as a result of the primary and secondary injections of the antigen. Because there already appeared to be a low level antibody response to β-galactosidase, the primary and secondary injections should have induced an amnestic response if the β-galactosidase protein was being recognized by the mice. β-galactosidase levels (Table 1) were determined for these Acholeplasma strains prior to the repeated freezing and thawing process to kill the bacteria. Therefore, strains ISM2050.2 and ISM2050.70 were producing relatively high levels of β-galactosidase that could have been recognized by the mouse.

The inability to induce an antibody response was probably due the choice of antigen used in this experiment. An existing antibody response to β-galactosidase was not unexpected because other bacteria (e.g., E. coli) that the mouse was exposed to produce this enzyme. The levels of β-galactosidase introduced with these recombinant strains was probably not high enough to induce a higher antibody response. Thus, this experiment suggested that a killed Acholeplasma recombinant vaccine was not effective in inducing an immune response to the foreign antigen, β-galactosidase.
Table B2. Absorbance values of an ELISA examining the antibody response to recombinant *Acholeplasmas* expressing the *E. coli lacZ*.

<table>
<thead>
<tr>
<th>Group</th>
<th>Day 0</th>
<th>Day 7</th>
<th>Day 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>0.122</td>
<td>0.119</td>
<td>0.100</td>
</tr>
<tr>
<td>ISM2050.1</td>
<td>0.129</td>
<td>0.134</td>
<td>0.121</td>
</tr>
<tr>
<td>ISM2050.2</td>
<td>0.138</td>
<td>0.133</td>
<td>0.121</td>
</tr>
<tr>
<td>ISM2050.70</td>
<td>0.118</td>
<td>0.106</td>
<td>0.136</td>
</tr>
<tr>
<td>ISM2050</td>
<td>0.140</td>
<td>0.156</td>
<td>0.143</td>
</tr>
<tr>
<td>Mab</td>
<td>0.479</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Ag</td>
<td>0.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Ab</td>
<td>0.032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No conjugate</td>
<td>0.013</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C: DEVELOPMENT OF A VACCINE TO *MYCOPLASMA HYOPNEUMONIAE*

**Introduction**

Mycoplasmal pneumonia of swine, as a result of *Mycoplasma hyopneumoniae* infection, is considered to be one of the most economically important causes of disease-associated loss in swine. Although *M. hyopneumoniae* infections rarely result in mortality, depressed growth rate, depressed feed consumption, and predisposition to secondary infection can result. Efforts to combat this disease have not proven to be entirely effective. Antibiotics have been shown to improve mycoplasma pneumonia; however, they do prevent the establishment of infection and subsequent lesions. Live, attenuated, bacterin, subunit, and *M. hyopneumoniae* supernatant vaccines have been tested, but have met with limited success.

The goal of this study was to employ a recombinant DNA approach to construct an effective vaccine to *M. hyopneumoniae*. Monoclonal antibodies (Mab) have been produced in the laboratory of Dr. R. Ross to a 65-kDa major surface antigen of *M. hyopneumoniae*. These Mabs inhibit hemagglutination by *M. hyopneumoniae*. Therefore, the approach was to locate the gene encoding this 65 kDa protein by screening a *M. hyopneumoniae* genomic library with these Mabs. Once the gene was mapped by restriction analysis, it would be placed into an expression vector and transformed into a vaccine strain of *Salmonella*. The recombinant *Salmonella* vaccine strain would be given orally to swine to induce a mucosal immune response to *M. hyopneumoniae*. It was hoped that the recombinant approach would enhance the immune response to mycoplasma antigens on mucosal surfaces and thus provide protection.
Preparation of the M. hyopneumoniae genomic library

M. hyopneumoniae chromosomal DNA was partially digested with Sau3AI and separated by size via sucrose density gradient centrifugation. Fractions possessing 9- to 22-kb fragments were pooled and ligated into λFIX cloning vector according to manufacturer's (Stratagene) instructions. Once the inserts were ligated into λFIX and packaged in vitro, the recombinant phage was plated on E. coli strain Q359, a P2 lysogen. The spi(-) selection prevented the growth of nonrecombinant phage. Approximately 1,200 independent recombinant phage were obtained and amplified in E. coli strain JM105.

Monoclonal antibody production

Two hybridoma cell lines producing Mabs to the 65-kDa protein of M. hyopneumoniae were a gift from Dr. R. Ross. The lines designated 80.1 and 27.10 were grown in RPMI medium containing 15% fetal bovine serum. Ten million cells were injected into pristane-primed mice and ascites fluid was collected 4 to 7 days later. The antibody-containing ascites fluid was aliquoted and stored at -70°C.

Screening of the M. hyopneumoniae genomic libraries

One hundred microliters of 4 x 10² pfu/ml were mixed with 100 µl of approximately 10¹⁰ JM105 cells/ml. After 15 to 20 min incubation at 37°C, the mixture was placed into 3 ml of soft agar and poured onto phage agar plates. Following incubation for 16 h at 37°C, the plaques were transferred to nitrocellulose filters. The filters were blocked with 3% BSA in Tris-saline for 30 min and then incubated for 1.5 h at room temperature with a 1:100 or 1:1,000 dilution of ascites fluid (80.1 or 27.10). The filters were washed (Tris-saline, 10 min; Tris-saline w/ 0.05% NP-40, 15 min, 2 times; Tris-saline, 10 min). The second antibody, ¹²⁵I-labeled goat
anti-mouse immunoglobulin was added for 1 h at room temperature. The filters were washed as previously described, dried, and exposed to film.

**Immunoblot blot analysis**

Approximately 50 μg of *E. coli* 81918, *M. hyopneumoniae*, and recombinants clones in *E. coli* lysates were loaded onto 10% SDS polyacrylamide gels. Following electrophoresis, the protein was transferred to nitrocellulose with a Hoeffer electroblotting transfer unit. Antigen detection on nitrocellulose was the same as the procedure used to screen the libraries.

**Results and Discussion**

**Construction of the *M. hyopneumoniae* genomic library**

Ten- to twenty-kilobase *M. hyopneumoniae* DNA fragments were ligated into λFIX, packaged in vitro, and plated on *E. coli* Q359. Approximately 1,200 independent recombinant phage were obtained. Because of the small size of the *M. hyopneumoniae* genome (ca. 800 kb), it was determined that only 300 independent recombinant phage containing 15-kb inserts would be required to obtain a complete genomic library at the 99% confidence level. Thus, the 1,200 independent clones in this library should represent the entire genome.

**Monoclonal antibody characterization**

The Mabs 80.1 and 27.10 both recognize a 65-kDa band when *M. hyopneumoniae* whole cell lysates are used in immunoblot analysis. The 80.1 Mab appeared to bind with greater affinity than did the 27.10 Mab. ELISA analysis showed that the 80.1 Mab was over 10 times more effective in recognizing whole cells than the 27.10 Mab. As a result, the 80.1 Mab was used to screen the *M. hyopneumoniae* genomic library.
Screening the *M. hyopneumoniae* genomic library

The phage library was plated on JM105 at a density of 500 plaques per plate and screened with Mab 80.1 at a dilution of 1:100. Initially there was high background, but it was eliminated by using the Mab 80.1 at a dilution of 1:1,000. Potentially positive clones were picked and stored at 4°C. Three microliters of the positive phage lysate was spotted onto soft agar containing JM105 cells and screened with both Mabs 80.1 and 27.10. Differing amounts of the ascites fluid of each Mab (1:1,000 for 80.1 and 1:50 for 27.10) were required to obtain a positive reaction, however. Eleven of the potentially positive clones continued to be positive upon rescreening.

Two of the positive clones were used by immunoblot analysis. The Mab 80.1 recognized a 39.5-kDa protein, but the Mab 27.10 failed to recognize any protein using this approach. The observation that Mab 27.10 weakly recognized the same positive plaques as Mab 80.1, but did not recognize the 39.5-kDa recombinant protein upon immunoblot analysis may indicate that this epitope is not very stable in *E. coli*. Thus, the epitope recognized by this Mab may have been lost during the immunoblot analysis procedure. It should be noted that the Mab 27.10 binds strongly to the 65-kDa antigen in immunoblot analysis of mycoplasmal whole cells.

The observation that the Mab 80.1 recognized a 39.5-kDa band indicated that either the protein was expressed in a truncated form in *E. coli* or only part of the gene was cloned. Expression of truncated gene products in *E. coli* is a possibility because of the alternate codon usage between *E. coli* and mycoplasmas. Mycoplasmas use the UGA codon for a tryptophan while it is a stop codon in *E. coli*. Thus, mycoplasma proteins containing a tryptophan encoded by a UGA would be prematurely truncated in *E. coli* during translation.

The truncation of the 65-kDa protein in *E. coli* may also be a result of improper post-translational modification such as the linkage of the protein antigen to a lipid component. A protein of identical molecular weight was characterized as a lipoprotein in *M. hyopneumoniae* by another laboratory (K S. Wise, Univ. of Missouri).
APPENDIX D. DEVELOPMENT OF A RECOMBINANT VACCINE TO
SERPULINA HYODYSENTERIAE

Description and Rationale

The etiologic agent of swine dysentery, Serpulina hyodysenteriae (formerly Treponema hyodysenteriae), causes a mucohemorrhagic diarrhea in swine. Although the disease has been recognized since 1921, the mechanism(s) of pathogenesis of this disease has not been determined. The endotoxin and hemolysin, however, have been implicated as major virulence factors. The goal of the project, therefore, was to produce an effective vaccine to S. hyodysenteriae based on inducing an immune response to the lipopolysaccharide O antigen. This would be accomplished by using a recombinant DNA approach in which the genes producing the S. hyodysenteriae O antigen would be cloned and expressed using an alternative vaccine delivery system developed in Salmonella spp. Previous work in the laboratory of Dr. Wannemuehler with murine and porcine challenge studies showed that the antibody response in both convalescent and immunized animals appeared to be against the serotype antigen. In addition, it was shown that mice could be protected from S. hyodysenteriae infection by using an a monoclonal antibody to the S. hyodysenteriae LPS. Unfortunately, there are at least seven S. hyodysenteriae serotypes. Thus, the possibility of using a multivalent approach was considered.

Initial experiments in the laboratory of Dr. Roy Curtiss III (Washington University, St. Louis) indicated that they had developed an avirulent Salmonella vaccine delivery system for swine. The Salmonella could colonize in the intestine of the pig for at least ten days without producing any obvious signs of disease. Moreover, a significant antibody response was detected one week after challenge.

Thus, in order to develop an effective vaccine against S. hyodysenteriae, the following approach was taken. First, the gene(s) encoding the S. hyodysenteriae LPS O antigen would
be cloned into *Escherichia coli*. Contiguous and noncontiguous *S. hyodysenteriae* genomic libraries were constructed in λZAP or in cosmid pISM24. A contiguous library consisted of large genomic inserts while the noncontiguous library was composed of small randomly associated genomic DNA fragments. The noncontiguous approach was taken because it has been shown that the LPS genes for other bacteria are located at a number of sites in the genome. Therefore, it was hoped that the noncontiguous library could generate a cassette in which the *S. hyodysenteriae* LPS would be produced. The libraries were screened with hyperimmune rabbit antisera or a monoclonal antibody to the LPS of *S. hyodysenteriae*. Second, once a positive clone was detected the genes related to the O antigen would be mapped. The O antigen coding sequence would then be cloned into the *Salmonella* expression vector, pYA804. Third, the recombinant *Salmonella* would be tested to determine whether an O antigen-specific immune response could be induced in mice and swine. And fourth, the vaccine would be tested to determine whether oral immunization could give protection against *S. hyodysenteriae*.

Materials and Methods

**Preparation of *S. hyodysenteriae* chromosomal DNA**

Two liters of *S. hyodysenteriae* B78 grown in a modified trypticase soy broth were washed once with PBS and resuspended in 1/200th of the original volume in NET buffer (0.1M Tris, 0.15M NaCl, and 0.08M EDTA, pH 7.5). Lysozyme was added to a final concentration of 100 mg/ml and incubated for 2 h at 37°C. Proteinase K (100 mg/ml) was then added and incubated at 65°C for an additional 2 h. The cells were then lysed with a lysis buffer (1% Triton X-100, 1% NP-40, and 1% deoxycholate) and proteinase K was immediately added and the lysate was incubated at 65°C for 2 h. The DNA was extracted with phenol and
phenol-chloroform, ethanol-precipitated and dialyzed against TE. Some of the DNA preparation was placed on a CsCl gradient for further purification.

The chromosomal DNA was partially digested with \textit{Sau3AI} and the fragments were separated over a sucrose density gradient. The fragments were pooled into 1- to 5-kb, 5- to 10-kb, 10- to 20-kb, and 20- to 50-kb pools for use in the appropriate cloning vector.

\textbf{Preparation of \textit{S. hyodysenteriae} genomic libraries}

\textbf{Preparation of genomic libraries in lambda.} The \textit{S. hyodysenteriae} libraries were constructed in the lambda vectors \textit{\lambda}ZAP (Stratagene) and \textit{\lambda}GEM-11 (Promega) according to manufacturer's instructions. To prepare the genomic library in \textit{\lambda}ZAP, 5- to 10-kb fragments were ligated into the \textit{BamH}I site, packaged using Gigapack (Stratagene), and allowed to infect \textit{E. coli} BB4 or XL1-blue cells. The library was amplified one time. To prepare the genomic library in \textit{\lambda}GEM-11, 10- to 20-kb fragments were partially filled in with Klenow and ligated into the \textit{XhoI} site, packaged using Packagene, and used to infect \textit{E. coli} LE392 cells.

\textbf{Preparation of genomic libraries in cosmid pISM24.} The 20- to 50-kb fragments were ligated into the unique \textit{BamH}I site of cosmid pISM24 to generate the contiguous genomic library. The noncontiguous library was generated by using the 5- to 10-kb fragments. The ligation mixes were packaged using Gigapack and used to infect \textit{E. coli} strains HB101, BB4, LE392, and XL1-blue. The colonies were picked and grown up in 200 µl volumes in a 96-well plate. One half of the culture was placed in 40% glycerol and stored at -20°C and to other half of the culture was placed in 15% glycerol and stored at -70°C in 96-well plates. Forty-eight or 96-pin replicators were used to inoculate plates for subsequent screening.
Screening the *S. hyodysenteriae* genomic libraries

The genomic libraries were screened with either pooled rabbit antisera to *S. hyodysenteriae* B78 whole cells or a monoclonal antibody (B78.18.2) to the *S. hyodysenteriae* B78 O antigen. Lambda libraries were screened from plaque lifts from plates containing about 1000 plaques. The membranes were blocked overnight with TSB (0.05M Tris, 0.15M NaCl, and 3% BSA, pH 7.5) at either 4°C, room temperature, or 37°C. Rabbit hyperimmune antisera that was diluted 1:250 in TSB and absorbed with lysed *E. coli* cells was added to the filters and gently shaken for 2 h at room temperature. The filters were washed (1 time with TS, 10 min; 2 times with TSN (TS with 0.05% NP-40), 20 min; and 1 time with TS, 10 min) then secondary antibody (^125^I-Protein G, approx. 10^6^ cpm/ml) diluted in TSB was added and incubated at room temperature for 2 h. Autoradiography was used to detect reactive plaques.

The cosmid libraries were screened by first inoculating LB containing 100 μg/ml ampicillin with a 48-pin replicator from the 96-well plate stocks. Colony lifts were performed and the cells were either lysed with chloroform and placed in TSB containing 100 μg/ml lysozyme and 1 μg/ml DNase or were not lysed at all. The rest of the screening protocol was the same as above.

Results and Discussion

Construction of *S. hyodysenteriae* genomic libraries

Two λZAP libraries were constructed from 2 different ligation mixes. It was predicted that 1,500 to 3,000 independent plaques would be necessary to be 99% confident that the entire genome (estimated to be 3,300 kb) would be cloned using 5- to 10-kb inserts. The 8,000 and 11,000 independent clones obtained with the two libraries should be sufficient to contain the entire *S. hyodysenteriae* genome. Scoring for the blue/white phenotype on plates containing X-gal and IPTG showed that less than 1% of the plaques were blue indicating that most of the
clones possessed insert DNA. Fifteen plaques from each library were picked and DNA was prepared. Restriction enzyme analysis showed that 15 of 15 clones in one library and 14 of 15 clones in the other library appeared to possess inserts. Each library was amplified one time to give libraries with titers of $1 \times 10^8$ and $1.4 \times 10^8$ pfu/ml. No plaques were generated using the λGEM-11 vector probably due to the inability to effectively partially fill-in the insert DNA prior to ligation with the λGEM-11 vector. Thus, only the λZAP library was used to screen for the O antigen gene(s).

The cosmid libraries were generated by ligating *S. hyodysenteriae* genomic fragments into the cosmid pISM24. Cosmid pISM24 is a derivative of cosmid pJB8 in which the multiple cloning site from pBluescript was inserted into the EcoRI site of pJB8. The contiguous library was constructed with inserts that were greater than 20-kb whereas the noncontiguous library was constructed with 5-to 10-kb fragments. Eight colonies from each library were picked and plasmids were prepared. Restriction enzyme analysis showed that the library was composed of large inserts of ca. 40-kb and the differing restriction patterns indicated that the inserts were different. DNA-DNA hybridization analysis using *S. hyodysenteriae* genomic DNA as a probe showed that the insert DNA was *S. hyodysenteriae*. These results indicated that a *S. hyodysenteriae* genomic library had been constructed using cosmid pISM24. Therefore, approximately 1,750 contiguous and 2,600 noncontiguous independent recombinant clones were picked and stored in 96-well plates for future analysis.

**Screening the *S. hyodysenteriae* libraries generated using λZAP**

The amplified *S. hyodysenteriae* libraries were initially screened with rabbit hyperimmune antisera. High background was encountered in the initial attempts. The high background was alleviated by first absorbing both the primary and secondary antibodies with lysed *E. coli* HB101 cells prior to the screening procedure. The positive control in these experiments was
lysed *S. hyodysenteriae* B78 cells and the negative control was lysed *E. coli* HB101 cells. No positive clones were generated but both the positive and negative controls reacted as predicted. The screening of these libraries was repeated with the monoclonal B78.18.2 as the primary antibody. Again, no positive clones were detected despite both controls working as expected.

It would have been predicted that some positive clones would have been detected if *S. hyodysenteriae* genes were being expressed. The amplified library was rechecked and the clones did not appear to have inserts. Therefore, the inability to detect positive clones at this point was due to the lack of recombinants containing *S. hyodysenteriae* DNA.

The packaging mixes were then reexamined to ensure that *S. hyodysenteriae* DNA fragments were indeed inserted into λZAP. Restriction enzyme analysis as well as DNA-DNA hybridization analysis using *S. hyodysenteriae* genomic DNA as a probe confirmed previous observations that the unamplified library contained in the packaging mix had *S. hyodysenteriae* inserts. The probe did not bind to the λZAP DNA negative control or to the bands that corresponded to the vector in the recombinant lanes. The unamplified library was then screened with both the rabbit hyperimmune antisera and the monoclonal antibody B78.18.2. Again, no positive clones were detected and the controls worked as expected.

The inability to detect positive clones, especially with the rabbit hyperimmune antisera, is difficult to explain. One possibility is that the gene products were lethal to the *E. coli* cell. This is unlikely because it could be reasoned that some *S. hyodysenteriae* gene products would be stably maintained and expressed in *E. coli*, yet the antisera failed to detect anything.

**Screening the *S. hyodysenteriae* libraries generated using pISM24**

Because efforts to detect the *S. hyodysenteriae* O antigen were unsuccessful using lambda vectors, cosmid vectors were examined. In *E. coli*, multiple genes from different regions of the genome are required for the expression of the O antigen. Thus, it was predicted that the genes
required for the expression of the *S. hyodysenteriae* LPS would be similarly spread around the genome.

One of the reasons for trying to express the O antigen in *E. coli* K-12 is that *E. coli* K-12 does not express the O antigen in its LPS. Thus, it was hoped that the gene product in the cosmid clones would complement the O antigen mutation and permit expression on the surface of the *E. coli* cell. This was demonstrated previously with *Neisseria gonorrhoeae* in that the O antigen from *N. gonorrhoeae* was expressed on the surface of HB101.

Therefore, one screening approach was to determine whether the *S. hyodysenteriae* O antigen was being expressed on the surface of the *E. coli* HB101 by using unlysed cells. Twenty-six potentially positive clones were detected with both the rabbit hyperimmune rabbit antisera and the monoclonal antibody. The positive reactions were weak, however. None of the 26 clones continued to give a positive reaction upon reexamination. Immunoblot analysis was tried in an attempt to increase the amount of the gene product so it could be detected by the antisera. In addition, ten of the potentially positive clones were treated with Proteinase K prior to the lysis step to try to enhance the detection process and to examine whether Proteinase K treatment alters the antigen recognized by the antisera to *S. hyodysenteriae*. All attempts to detect the O antigen by immunoblot analysis were unsuccessful. ELISA analysis was also performed and no positive reaction was detected. It should be noted that in the colony lifts, immunoblot analyses, and ELISA analyses that both the positive (lysed *S. hyodysenteriae* cells) and negative (lysed HB101 cells) controls worked as expected.

To determine why these clones were no longer reacting positively with the antisera, DNA was prepared by the Triton lysis procedure. Restriction enzyme analysis of these clones showed that most of the inserts were now less than 12-kb. DNA-DNA hybridization analysis with *S. hyodysenteriae* genomic DNA confirmed that about 10-kb of each clone continued to be present. Apparently the inserts in these cosmids were very unstable.
It should be noted that the cosmids could only be grown in HB101 and not in BB4, XL1-blue, or LE392 *E. coli* cells. This suggested a possible restriction/modification barrier to the cloned inserts. None of the clones from the contiguous library gave a positive reaction with either the hyperimmune rabbit antisera or the monoclonal antibody. A total of 17 attempts were made to locate positive clone from both the contiguous and noncontiguous genomic libraries.

In summary, the attempts to clone the gene(s) responsible for *S. hyodysenteriae* O antigen production were unsuccessful. Apparently, the cosmid clones were very unstable because after several rounds of amplification and growth the insert size was greatly reduced. Instability of cosmid inserts has been noted in other systems. As an effort to reduce cosmid instability the possibility of using a low copy number cosmid, pPR691, was considered. The inability to detect any *S. hyodysenteriae* gene products in *E. coli* with the rabbit hyperimmune antisera despite confirmation of cloned inserts remains an enigma. Perhaps the *S. hyodysenteriae* gene regulatory regions (i.e., promoters and translation initiation regions) are not recognized in *E. coli*. 