

ABSTRACT

Title of Document:

ARCHAEAL TRANSCRIPTIONAL
REGULATION OF CATABOLIC CARBON
MONOXIDE DEHYDROGENASE IN
METHANOSARCINA SPECIES

Kimberly Lynn Anderson, Doctor of Philosophy,
2009

Directed By:

Professor Kevin R. Sowers, Marine Estuarine
Environmental Sciences

In Archaea, the basal transcription machinery is eukaryotic-like, but some components, such as activator and repressor proteins, are bacteria-like. To further gain knowledge into cellular processes of Archaea, the genome of *Methanosarcina thermophila* was searched for helicase genes. A homolog of yeast RAD25, a gene with helicase and nucleotide excision repair (NER) abilities, was isolated. *M. thermophila rad25* has the domains for helicase activity, but the C-terminal end is truncated, indicating that this protein mostly likely does not function in NER. After overexpression, helicase activity assays of Rad25 indicated that it might have helicase activity; however, there appeared to be contaminating proteins in the purification, so it was not possible to assign the activity only to Rad25. Additional work is necessary to characterize this protein. To investigate transcription, catabolic gene regulation was studied, specifically regulation of carbon monoxide dehydrogenase/acetyl CoA

synthase (CODH/ACS) from *Methanosarcina* species. The regions upstream of the transcriptional start site, as well as the 5' leader region of *cdhA*, were investigated for *trans* factors and *cis* elements that might be involved in regulation. Experiments revealed that regulation of *cdhABCDE* does not appear to involve *trans* factors upstream of the transcriptional start site. However, deletion analysis indicated that the 5' leader region does have a role in regulation. Comparing the protein levels to the mRNA levels revealed there was no significant difference between the two, indicating that translational regulation was not a factor. Other experiments ruled out differential mRNA stability as a factor in regulation. A region located between +358 and +405 was important in transcriptional regulation, indicating that regulation occurred at the level of transcription elongation. A model for regulation of catabolic CODH/ACS by differential elongation is proposed. Although 5' leader regions identified for other archaeal genes have been postulated to be involved in regulation, this was the first study to demonstrate a regulatory role by an archaeal leader sequence for differential elongation. Identifying regulatory mechanism(s) of catabolic genes such as CODH/ACS is critical for understanding the regulatory strategies employed by the methanoarchaea to efficiently direct carbon and electron flow during biomass conversion to methane.

ARCHAEAL TRANSCRIPTIONAL REGULATION OF CATABOLIC CARBON
MONOXIDE DEHYDROGENASE IN *METHANOSARCINA* SPECIES

By

Kimberly Lynn Anderson

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:
Dr. Kevin R. Sowers, Chair
Dr. Robert Gunsalus
Dr. Zvi Kelman
Dr. Frank Robb
Dr. Harold Schreier
Dr. Daniel Terlizzi

© Copyright by
Kimberly Lynn Anderson
2009

Dedication

This work is dedicated to my husband, Matthew. His unwavering support and faith in me made this dissertation possible.

Acknowledgements

This work was funded in part by the U.S. Department of Energy grants #DE-FG02-93ER20106 and #DE-FG02-07ER64502, the National Science Foundation grant #MCB0110762, and Advanced Research International subcontract for DOE grant #DE-FG02-03ER83596.

I would like to extend my sincerest appreciation to my advisor, Kevin Sowers, for giving me the opportunity to work in his laboratory. His knowledge and support have taught me so much about critical thinking and experimental design. I am a better scientist because of his patience and guidance, and I could not have asked for a better mentor. I would also like to thank the members of my advisory committee: Rob Gunsalus, Zvi Kelman, Frank Robb, and Harold Schreier. Their comments and suggestions have provided me with more insight into my projects, and helped me to see all of the aspects of my research.

There have been many people who helped to move this research forward. I thank Zvi Kelman and Jae-Ho Shin for allowing me to work in their lab and training me in biochemical techniques like EMSA. Thanks to Mark Shirliff at the University of Maryland Baltimore for training me in two-dimensional gel electrophoresis techniques. David Grahame at Uniformed Services University was very helpful in providing the purified Cdh extract for binding studies.

Many thanks go out to all the members of the Sowers lab, both past and present. Ethel Apolinario has been a source of invaluable advice and assistance. She has also been one of my strongest supporters. Thanks go to Birthe Kjellerup for her valuable discussions and for reviewing my dissertation. Sonja Fagervold and

Sheridan MacAuley, two previous graduate students in the lab, provided support and encouragement, as well as participating in many fruitful discussions.

Finally, I would like to thank my family and friends for their constant support and encouragement. Your faith in me gave me the strength and the courage to follow my dreams and achieve my goals. I thank you all from the bottom of my heart.

Statement of Contributions

Karin Jackson performed the isolation of *rad25* from *Methanosarcina thermophila*, primer extension analysis, and the northern analysis.

Ethel Apolinario was responsible for the construction of the deletion constructs and the beta-galactosidase assays for the leader region deletions.

Sheridan MacAuley performed the beta-galactosidase assays on the gross deletion constructs.

Jae-Ho Shin (Center of Advanced Research in Biotechnology) performed the helicase assays on *M. thermophila* Rad25.

Tom Santangelo (Ohio State University) performed the *in vitro* transcription assays.

Table of Contents

Dedication	ii
Acknowledgements	iii
Statement of Contributions	v
Table of Contents	vi
List of Tables	ix
List of Figures	x
List of Abbreviations	xii

Chapter 1: Introduction	1
<u>1.1. General introduction</u>	1
<u>1.2. Methanogenic Archaea and their role in the environment</u>	3
<u>1.3. Substrate utilization by methanogens</u>	5
1.3.1. Acetate utilization by methanogenic Archaea	10
1.3.2. Carbon monoxide dehydrogenase/acetyl Co-A synthase	10
<u>1.4. Transcription in Archaea</u>	12
1.4.1. Archaeal transcriptional machinery and mechanisms	12
1.4.2. Regulation of transcription in Archaea	13
1.4.3. Role of 5' untranslated leader regions in transcriptional regulation	15
<u>1.5. Objectives of Research</u>	18
1.5.1. Objective 1: Identification of putative Rad25 transcription factor from <i>M. thermophila</i> TM1	18
1.5.2. Objective 2: Identification of <i>cis</i> regions and <i>trans</i> factors involved in the regulation of carbon monoxide dehydrogenase/acetyl Co-A synthase from <i>Methanosarcina</i> species	19
1.5.3. Objective 3: Determination of the role of the 5' leader region of carbon monoxide dehydrogenase/acetyl Co-A synthase	20

Chapter 2: Characterization of a <i>rad25</i> helicase homolog in <i>Methanosarcina thermophila</i> TM1	21
<u>2.1. Abstract</u>	21
<u>2.2. Introduction</u>	21
<u>2.3. Materials and Methods</u>	23
2.3.1. Nucleotide sequence accession numbers	23
2.3.2. Strains and plasmids	24
2.3.3. Media and Growth	24
2.3.4. Cloning and sequence analysis of DNA	25
2.3.5. Southern analysis	25
2.3.6. RNA isolation, Northern hybridization, and primer extension analysis	26
2.3.7. Protein purification	27
2.3.8. Site-directed mutagenesis	28
2.3.9. Helicase activity assay	28
2.3.10. Disruption of <i>rad25</i> from <i>M. thermophila</i>	29

2.4. Results.....	30
2.4.1. Isolation of archaeal <i>M. thermophila rad25</i>	30
2.4.2. Phylogenetic analysis.....	34
2.4.3. Expression analysis of <i>M. thermophila rad25</i>	35
2.4.4. Identification of transcription initiation site	36
2.4.5. Detection of helicase activity.....	38
2.4.6. Knockout of <i>M. thermophila rad25</i>	39
2.5. Discussion	41
Chapter 3: Identification of <i>trans</i> factors involved in the regulation of carbon monoxide dehydrogenase in <i>Methanosarcina</i> species	45
3.1. Abstract	45
3.2. Introduction.....	46
3.3. Materials and Methods.....	48
3.3.1. Methanogen strains and media.....	48
3.3.2. Electrophoretic mobility shift assay.....	48
3.3.3. Affinity chromatography of DNA binding proteins	50
3.3.4. Two-dimensional analysis of DNA binding proteins	51
3.3.5. Competition assays	51
3.3.6. Construction of <i>cdh'</i> :: <i>lacZ</i> deletion plasmids	52
3.3.7. Reporter gene expression analysis	53
3.4. Results.....	55
3.4.1. Identification of putative protein binding sites using electrophoretic mobility shift assays.....	55
3.4.2. Isolation and sequence analysis of DNA binding proteins using affinity chromatography	58
3.4.3. Specificity of Cdh binding.....	62
3.4.4. Effect of sequence deletions on the regulation of <i>cdh</i> by different substrates.....	65
3.5. Discussion	67
Chapter 4: A 5' leader region regulates expression of Methanosarcina CO dehydrogenase/acetyl-coenzyme A synthase	70
4.1. Abstract	70
4.2. Introduction.....	71
4.3. Materials and Methods.....	73
4.3.1. Archaeal and bacterial strains	73
4.3.2. Media and cell growth	74
4.3.3. Sequence analysis	74
4.3.4. Construction of <i>cdh'</i> :: <i>lacZ</i> deletion plasmids	75
4.3.5. Identification and deletion of putative mRNA secondary structures	77
4.3.6. Reporter gene expression analysis	78
4.3.7. <i>In vitro</i> transcription assays	78
4.3.8. RNA extraction and quantitative reverse transcriptase PCR	79
4.3.9. Transcript stability studies	80
4.3.10. Construction of TBP::lacZ reporter plasmids	80

4.4. Results.....	82
4.4.1. Identification of conserved sequence in carbon monoxide dehydrogenase/acetyl Co-A synthase (CODH/ACS) orthologs	82
4.4.2. Effect of sequence deletions on the regulation of <i>cdh</i> by different substrates.....	84
4.4.3. Effect of deletions in the sequence within the 5' leader region of <i>cdhA</i> ...	86
4.4.4. Comparison of Cdh::LacZ fusion protein and <i>cdh</i> transcript levels with different substrates	89
4.4.5. <i>In vitro</i> transcription of 5' leader region.....	89
4.4.6. Transcript stability of <i>cdh</i> during growth with different substrates.....	92
4.4.7. Effect of substrates on elongation of <i>cdh</i> transcript.....	94
4.4.8. Regulation of different TATA-binding proteins by different substrates ...	95
4.5. Discussion	97
Chapter 5: Discussion and future perspectives	102
Appendix: Supplemental Information.....	111
Bibliography	113

List of Tables

Table 3.1.	<i>E. coli</i> and <i>M. acetivorans</i> plasmid constructs.....	53
Table 3.2.	Identification of protein spots from two-dimensional gel electrophoresis.....	62
Table 4.1.	<i>E. coli</i> and <i>M. acetivorans</i> plasmid constructs.....	75
Table S2.1.	Primers used in Chapter 2 for characterization of <i>rad25</i> from <i>M.</i> <i>thermophila</i>	111
Table S3.1	Primers used in Chapter 3.....	111
Table S4.1.	Primers used to construct deletion mutants, TBP constructs, and qRTPCR products in Chapter 4.....	112

List of Figures

Figure 1.1.	Coenzymes in the methanogenic pathway.....	7
Figure 1.2.	Four methanogenic pathways found in <i>Methanosarcina</i>	9
Figure 2.1.	Alignment of Rad25 protein domains.....	32
Figure 2.2.	Alignment of helicase protein sequences.....	33
Figure 2.3.	Northern analysis of <i>M. thermophila rad25</i> expression.....	35
Figure 2.4.	Primer extension analysis of the 5' terminus of <i>rad25</i> mRNA from <i>M. thermophila</i>	37
Figure 2.5.	DNA helicase activity of Rad25 from <i>M. thermophila</i>	38
Figure 2.6.	Polyacrylamide gel electrophoresis of purified Rad25.....	40
Figure 3.1.	Alignment of <i>cdhA</i> promoter region from four <i>Methanosarcina</i> species.....	56
Figure 3.2.	Electrophoretic mobility shift assay of DNA fragments from the region upstream of the <i>cdhA</i> promoter.....	57
Figure 3.3.	Polyacrylamide gel electrophoresis of proteins isolated from affinity chromatography with the 5' 371 bp leader sequence DNA of the <i>cdh</i> operon bound to biotin beads.....	59
Figure 3.4.	Two-dimensional gel electrophoresis (2DGE) of DNA-binding proteins from acetate- and TMA-grown cell free extracts.....	61
Figure 3.5.	Electrophoretic mobility shift assay with 16 µg partially purified Cdh complex binding to region upstream of <i>cdhA</i> transcriptional start site.....	64

Figure 3.6.	Map of deletions upstream of the <i>Methanosarcina thermophila</i> TM1 <i>cdhA</i> transcriptional start site and the effects of those deletions on regulation by different substrates.....	66
Figure 4.1.	Alignment of <i>cdhA</i> promoter region from four different <i>Methanosarcina</i> species.....	83
Figure 4.2.	Map of deletions within the 5' leader region of <i>M. thermophila</i> TM1 <i>cdhA</i>	84
Figure 4.3.	Effects of deletions downstream of the transcriptional start site on regulation by different substrates.....	86
Figure 4.4.	Effects of deletions in putative mRNA hairpin loops in the UTR of <i>M. thermophila</i> TM1 <i>cdhA</i>	88
Figure 4.5.	<i>In vitro</i> transcription assay of <i>cdhA</i> promoter and 5' leader region.....	91
Figure 4.6.	Differences in transcript levels after addition of actinomycin D.....	93
Figure 4.7.	Differences in levels of transcripts of varying lengths demonstrated by qRT-PCR.....	91
Figure 5.1.	Proposed model for regulation of <i>cdhABCDE</i>	105

List of Abbreviations

2DGE	two-dimensional gel electrophoresis
aa	amino acid
bp	base pair
cDNA	complementary DNA
CODH/ACS	carbon monoxide dehydrogenase/acetyl Co-A synthase
CSPD	disodium 3-(4-methoxy Spiro(1,2-dioxethane-3,2'-(5'-chloro)tricyclo[3.3.1.1. ^{3,7}]decan)-4-yl)phenyl phosphate
dNTP	dinucleoside triphosphate
DTT	dithiothreitol
EMSA	electrophoretic mobility shift assay
kb	kilobase
LacZ	β -galactosidase
NER	nucleotide excision repair
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
qRT-PCR	quantitative reverse transcriptase polymerase chain reaction
SDS	sodium dodecyl sulfate
TAE	tris acetate EDTA
TBE	tris borate EDTA
TBP	TATA-binding protein
TMA	trimethylamine
UTR	untranslated leader region

Chapter 1: Introduction

1.1. General Introduction – Archaea

The Archaea are the third lineage of the modern tree of life. This group was first proposed by Carl Woese and colleagues based on phylogenetic analysis of 16S rRNA gene sequences (36, 130). Although the Archaea are phylogenetically distinct from Bacteria and Eukaryotes, they have many characteristics in common with both. Features of Archaea that are similar to Eukaryotes include a multi-subunit RNA polymerase, eukaryotic-like basal transcription components for initiation, and employment of an unmodified methionine for translation. However, the morphology of Archaea is similar to Bacteria, as they both are unicellular organisms that lack intracellular compartmentalization and nuclear membranes. In addition, Archaea are also capable of many metabolic processes that are found in Bacteria, such as nitrogen fixation. Archaea have their own unique characteristics as well. These characteristics include unique cofactors and enzymes and the occurrence of phytanyl ether lipids instead of fatty ester lipids in membranes. A majority of these organisms grow in conditions that were once considered the extremes for life, such as extreme temperatures, high salinity, and highly reduced and anoxic environments. Members of the Archaea include non-methanogenic hyperthermophiles and thermophiles, halophiles, and methanogens.

The original division of Archaea was based on physiology and included three groups: methanogens, haloarchaea, and sulfur-dependent thermophiles (36). In the

current classification of the Archaea, which is based on phylogeny, the domain is now divided into four different kingdoms: *Euryarchaeota*, *Crenarchaeota*, *Korarchaeota*, and *Nanoarchaeota* (24, 105). The two kingdoms, *Euryarchaeota* and *Crenarchaeota*, contain the majority of known Archaeal species. These two kingdoms, while defined by certain Archaeal species, are also defined by different biochemical properties, such as differences in DNA replication machinery (35). *Euryarchaeota* have ssDNA-binding protein RPA, a heterotrimer PCNA sliding clamp, and no identified cell division protein. *Crenarchaeota*, on the other hand, have a ssDNA-binding protein SSB, a homotrimeric PCNA, and the FtsZ cell division protein (57). The kingdom *Euryarchaeota* includes all known species of extreme halophiles and methanogens, and some species of non-methanogenic extreme thermophiles. Examples of organisms belonging to this kingdom include members of the genera *Halobacteria*, *Methanosarcina*, and *Pyrococcus*. The *Crenarchaeota* are made up of thermophiles, including members of the genera *Sulfolobus* and *Pyrobaculum*. Some potential psychrophilic Crenarchaeotes have been identified through phylogenetic analysis, such as *Cenarchaeum symbiosium*, which was identified in a deep cold water marine sponge, but these organisms have not yet been cultured in the laboratory (91).

The remaining two kingdoms contain fewer members and have been described more recently. *Nanoarchaeum equitans* is the sole identified member of the kingdom *Nanoarchaeota* (53). This organism is a symbiote of the Crenarchaeote, *Ignicoccus spp.* It is a hyperthermophilic organism with a rapid rate of evolution and is one of the smallest known living organisms. The evidence for whether *Nanoarchaeum*

belongs in a separate kingdom depends on which phylogenetic marker is used. 16S rDNA phylogenies indicate that this organism belongs in a separate kingdom (25). However, using different ribosomal proteins to determine the relationship of *Nanoarchaeum* to other Archaea indicates that this organism may be more closely related to Euryarchaeotes, namely Thermococcales. Members of the fourth kingdom *Korarchaeota* were identified in hot environments (11). This phylogeny is based on comparative sequence analyses of 16S rRNA genes, as these organisms have not been cultured in the laboratory.

1.2. Methanogenic Archaea and their role in the environment

Methanogenic Archaea, or methanogens, were the first members of the Archaea to be described. Biological methane production is an anaerobic microbial process that occurs ubiquitously on Earth. This process was first demonstrated by Alessandro Volta, who collected the gas from disturbed lake sediments and showed that it was flammable. The gas was originally called “combustible air” and later shown to be methane gas. Approximately 80% or more of the methane produced yearly is produced through biological pathways, making methanogens key producers of this greenhouse gas (51).

Methanogens are found in anoxic environments that contain organic material, such as swamps, aquatic sediments, digestive tracts of animals, and near hydrothermal vents (127). These microorganisms also have the ability to grow in

different environmental conditions, such as varying salt concentrations, from fresh to marine water, and at different temperatures, ranging from -2.5 to 110 °C.

In the environment these organisms are often part of a consortium of anaerobic microorganisms that break down organic matter into carbon dioxide and methane. The first group of organisms in this process, the fermentative bacteria, convert complex organic matter into hydrogen and carbon dioxide and partially oxidized intermediates such as fatty acids and primary alcohols. The second physiological group, the acetogenic bacteria, further oxidizes the intermediates to hydrogen, acetate, and carbon dioxide. The third physiological group, the methanogenic Archaea, converts acetate, hydrogen, formate, and carbon dioxide formed by the first two groups into methane. These organisms exist as syntrophs, meaning methanogens require the other groups to produce the substrates for their growth and methanogenesis. The other two groups benefit from the utilization of hydrogen by the methanogens, as the production and accumulation of hydrogen creates a thermodynamically unfavorable environment. Methanogens maintain a low hydrogen partial pressure environment that allows reactions to take place that would normally not be thermodynamically feasible. For fermentors, the ability to reduce protons under low hydrogen partial pressure generated by the methanogens enable them to generate more energy by producing more oxidized fermentation products such as acetate. For the hydrogen producing acetogens that are growth-limited by the accumulation of hydrogen, methanogens use the hydrogen, removing the growth inhibition. This process is called interspecies hydrogen transfer.

1.3. Substrate utilization by methanogens

In the environment, the two most utilized substrates for methanogenesis are acetate and hydrogen. Of these two substrates, the majority of methane produced by these organisms in the environment is derived from acetate, with approximately 75% of methane coming from this source (71). Methanogens are also capable of using a few other substrates for methanogenesis, including formate, methylamines, methylsulfides, and methanol. Recent studies have also reported the ability of some methanogenic Archaea species to use carbon monoxide as a substrate, although this process does not produce methane, but rather acetate and formate as end products. However, not all methanogenic Archaea can use all substrates. Even though the majority of methane produced in the environment is derived from acetate, there are only two genera currently known that are capable of using acetate as a substrate: *Methanosarcina* and *Methanosaeta*. For *Methanosarcina*, which are able to utilize multiple substrates, acetate is not the most energetically favorable, and when two substrates are available, these organisms will exhibit diauxic growth, consuming the other substrate before utilizing acetate (17, 34, 60, 74, 107). These organisms will use the substrate with the highest free energy, such as methanol or trimethylamine, before using acetate as a substrate for methanogenesis.

There are four main catabolic pathways for methanogenesis: carbon dioxide (CO₂) reduction with hydrogen (H₂), methyl reduction using H₂, dismutation of methanol or methylated amines, and fermentation of acetate. These pathways all involve multiple coenzymes and cofactors (Figure 1.1) (94). Methanofuran (MFR) is a one-carbon carrier similar to molybdopterin. Tetrahydromethanopterin (H₄MPT)

is a single carbon carrier which is analogous to tetrahydrofolate in bacterial and eukaryotic systems. Originally, this cofactor was thought to be unique to methanogenic Archaea, but it has since been found in other Archaea. Coenzyme M (HS-CoM) was also thought to be unique to Archaea, but it has been found in methylotrophs and in bacteria capable of epoxy degradation. There are two factors used in methanogenesis, 7-mercaptoheptanoylthreonine phosphate (HS-HTP) and cofactor F₄₃₀, that are unique to methanogenic Archaea.

Each of the four pathways of methanogenesis involves multiple steps (Figure 1.1). Some steps are unique to each pathway, while other steps are shared amongst all the pathways. The carbon dioxide reduction pathway involves the sequential reduction of CO_2 via electrons from H_2 to produce methane (Figure 1.2A). CO_2 is first reduced, along with coenzyme MFR, to generate formyl-MFR. The formyl group is transferred to H_4MPT to yield formyl- H_4MPT , which is cyclized to methenyl- H_4MPT . This substrate is reduced twice to methyl- H_4MPT . The methyl group from this compound is transferred to coenzyme CoM-SH, which is further reduced to CH_4 . The other product of this final step, CoM-SS-HTP, is reduced to regenerate CoM-SH and HPT-SH.

There are two pathways through which methylotrophic substrates, such as methanol and trimethylamine, can be converted to methane. The first pathway involves the transfer of the methyl group from these substrates to a corrinoid protein (Figure 1.2.B). The corrinoid proteins are substrate specific. The methyl group from the corrinoid protein is transferred to coenzyme CoM-HS. The methyl CoM-SH is reduced to CH_4 as described in the final steps of the process above. Some of the methyl groups that are generated from this process are oxidized in reverse through a pathway identical to the reduction of CO_2 , which generates electrons to reduce the CoM-SS-HTP. The second pathway involves the reduction of the methyl groups from the methylotrophic substrates through the oxidation of hydrogen (Figure 1.2.C).

The fourth pathway of methanogenesis, the fermentation of acetate, is shown in Figure 1.2.D and is discussed at length in the next section.

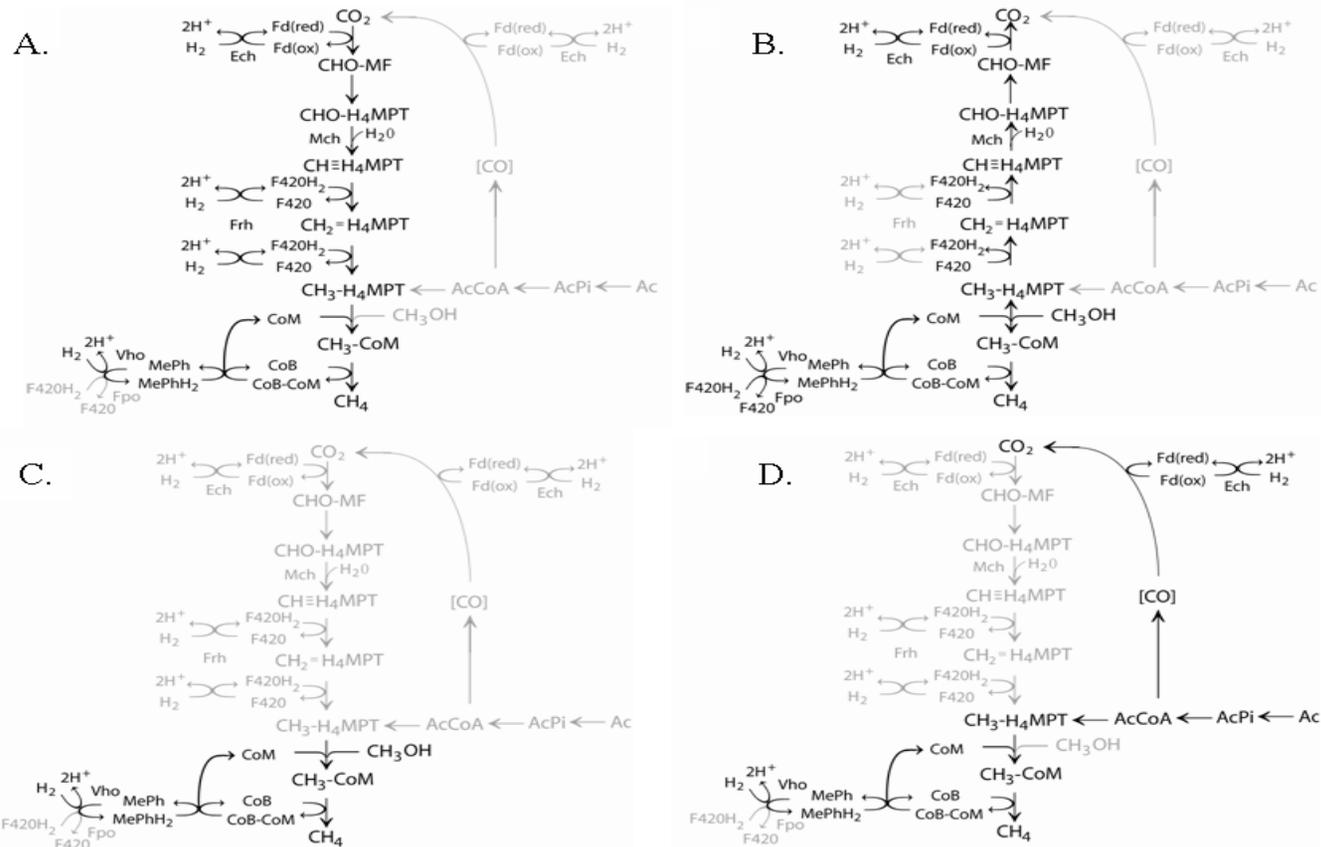


Figure 1.2. Four methanogenic pathways found in *Methanosarcina*. A, the reduction of CO₂ to methane using via oxidation of H₂. B, the methylotrophic pathway. C, the methyl reduction pathway. D the aceticlastic pathway. Steps in black are those specifically involved in each pathway. Ech, ferredoxin-dependent hydrogenase; Frh, F₄₂₀-dependent hydrogenase; Vho, methanophenazine-dependent hydrogenase; Fpo, F₄₂₀ dehydrogenase; CHO-MF, formyl-methanofuran; CHO-H₄ MPT, formyl- tetrahydromethanopterin; CH≡H₄MPT, methenyl tetrahydromethanopterin; CH₂=H₄MPT, methylene-tetrahydromethanopterin; CH₃-H₄MPT, methyl-tetrahydromethanopterin; CH₃-CoM, methyl-coenzyme M; CoM, coenzyme M; CoB, coenzyme B; CoM-CoB, mixed disulphide of CoM and CoB; Mph/MphH₂, oxidized and reduced methanophenazine; F₄₂₀/F₄₂₀H₂, oxidized and reduced Factor 420; Fd(ox)/Fd(red), oxidized and reduced ferredoxin; Ac, acetate; Ac-Pi, acetyl-phosphate; Ac-CoA, acetyl-coenzyme A. [Modified from (44) in (109)].

1.3.1. Acetate utilization by methanogenic Archaea

The acetate utilization pathway in methanogens is a very important pathway, as the majority of the biologically produced methane is derived from acetate. Of the two genera that can utilize acetate as a substrate for methanogenesis, *Methanosaeta* is an obligate acetotroph (58). Members of this genus can only grow using acetate as a substrate. In contrast, members of the genus *Methanosarcina* are able to use most substrates, including methylamines, methanol, carbon dioxide/hydrogen, carbon monoxide, and acetate. Because of this, *Methanosarcina* species are the most metabolically diverse of all the methanogens.

Because of its important role in global methanogenesis, the catabolic acetate utilization has been extensively studied (Figure 1.2D). The first step involves the activation of acetate to form acetyl-CoA. In *Methanosarcina* species, this is accomplished by the enzymes phosphotransacetylase and acetate kinase. The enzyme acetyl-CoA synthase is responsible for this step in *Methanosaeta* species. The next enzyme, carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS), catalyzes the release of CO₂ and transfers a methyl group onto H₄MPT, where the methyl group subsequently proceeds through the reductive pathway discussed above.

1.3.2. Carbon monoxide dehydrogenase/acetyl co-A synthase

In acetate catabolism, one of the limiting steps involves the enzyme carbon monoxide dehydrogenase/acetyl CoA synthase (CODH/ACS). This enzyme contains five subunits, including a nickel/iron-sulfur component, a corrinoid/iron-sulfur

component, and a component of unknown function. In 1986, CODH/ACS was first detected in *Methanosarcina thermophila* (121). It was also demonstrated that this protein was much more abundant when cells were grown under acetate as opposed to methanol. Later, genetic analysis of the gene encoding this protein, *cdhABCDE*, indicated that *cdhABCDE* was differentially expressed based on substrate, and that the regulation of this operon appeared to be at the level of transcription (114). A previous study using a plasmid-mediated *lacZ* fusion system revealed that expression of LacZ while under the control of the *M. thermophila cdhABCDE* promoter was 54-fold lower when cells were grown on methanol versus acetate (5). These same results have been seen in proteomic and microarray studies, confirming that this operon is more highly expressed during growth under acetate conditions than methylophilic conditions (65).

The genomes of *Methanosarcina mazei*, *Methanosarcina acetivorans*, and *Methanosarcina barkeri* have all been sequenced (30, 37, 73). The genomes of all three of these organisms have two full copies of *cdhABCDE*. These two copies have greater than 95% identity, making it hard to distinguish between the two copies. Proteomic studies of *M. acetivorans* have indicated that one copy of the *cdhA* gene (MA3860) is more highly expressed than the other *cdhA* copy (MA1016) when cells are grown on acetate (65-67). In contrast, a study by Grahame (42) found that only one of the carbon monoxide dehydrogenase/acetyl CoA synthase complexes was present, leading the author to conclude that this species had only one copy of the gene, which is the ortholog of the highly expressed MA3860.

1.4. Transcription in Archaea

1.4.1. Archaeal transcriptional machinery and mechanisms

The components involved in transcription in Archaea are very similar to those found in Eukarya. The core machinery of archaeal transcription involves a RNA polymerase that shares many properties with the RNA polymerase II from Eukarya (63). Archaeal RNA polymerase is a multisubunit protein whose very structure is similar to that of the eukaryotic RNA polymerase. The gene sequences of the individual subunits are also highly conserved between Eukarya and Archaea. The archaeal RNA polymerase cannot initiate transcription on its own, but requires the presence of two transcription factors described below.

Transcription initiation in Archaea involves the recognition of an AT-rich promoter region, also known as the TATA-box. This region is recognized by a TATA-binding protein (TBP), which is functionally interchangeable with the TBP from yeast and humans (126). TBP binds to the promoter, and the protein TFB, a functional analog of eukaryotic TFIIB, is recruited to the complex. This transcription factor recognizes the B recognition element (BRE), a sequence located upstream of the TATA box. The new complex, containing both the TBP and TFB, recruits the RNA polymerase. The archaeal RNA polymerase structure resembles that of eukaryotic RNA polymerases.

Archaea also contain a homolog of eukaryal TFIIE, which is called TFE. In Eukaryotes, this transcription factor is used to recruit TFIIH. A homolog for TFIIH has not been found in Archaea. TFE appears to promote transcription *in vitro* from

weak promoters in Archaea. Another transcription factor present in Archaea is TFS, which is homologous to the C-terminal part of TFIIIS in Eukaryotes. This protein induces cleavage of paused RNA polymerase and also gives the polymerase proofreading activity (62).

1.4.2. Regulation of transcription in Archaea

In Archaea, there are many examples of transcriptional activators and repressors that control transcription. Activators of transcription have been described previously such as Ptr2, an activator found in *Methanocaldococcus jannaschii*, which binds upstream of the target gene and recruits TBP to activate transcription (87). This protein, which is a member of the Lrp family of regulators, regulates transcription of ferredoxin A and rubredoxin 2. The Lrp family of regulators are found within the Bacteria and Archaea and are mostly involved in the regulation of amino acid metabolism-related genes (23). Another member of the Lrp family, LysM, activates transcription of genes involved in lysine biosynthesis in the absence of lysine in *Sulfolobus solfataricus* (22). In *Halobacterium*, the transcriptional regulator Bat activates transcription of multiple genes involved in the synthesis of purple membranes, and this regulator is responsive to light intensity and oxygen tension (9).

Repressor proteins that inhibit transcription have also been identified in Archaea. One example of a repressor is LrpA from *Pyrococcus furiosus*. *Pyrococcus furiosus* LrpA represses transcription of its own gene by preventing RNA polymerase from binding to the promoter (28), and it is similar to other bacterial Lrp family transcriptional regulators. Other Archaea, such as *S. solfataricus* and *M. jannaschii*,

also have Lrp family proteins that act as repressors (15, 85). The *Archaeoglobus fulgidus mdr1* (metal dependent repressor) gene codes for a protein that can bind to DNA to prevent polymerase recruitment. This binding depends on the presence of bivalent cations. The protein, Mdr1, is in the DtxR family of bacterial transcriptional regulators (14). Mdr1 controls expression of its own gene, as well as three genes encoding an iron-importing ABC transporter that are cotranscribed with *mdr1*. Another example involves a repressor protein, NrpR, from *Methanococcus maripaludis* (68). This repressor is a helix-turn-helix protein that binds upstream of *nif* and *glnA* promoter regions to prevent transcription depending on nitrogen availability. When levels of ammonia are high, NrpR represses transcription of *nif*. NrpR has two operator sequences to which it can bind, OR₁ and OR₂ (69). Derepression is achieved through binding of 2-oxoglutarate, which is an indicator of low nitrogen availability. Although this protein appears to be unique to Euryarchaeotes, the cooperative binding to two operators is similar to other repression mechanisms found in bacteria.

Archaeal-specific regulators of transcription have also been described. These transcriptional control mechanisms are unique to the Archaea, and are not found in the other domains of life, to date. One example of an activator that appears to be unique to Archaea is GvpE. The GvpE protein from *Haloarchaea* sp. activates the transcription of gas vesicle formation genes (59). This protein has a motif that is similar to bZIP eukaryotic transcriptional regulators, but it has no sequence similarity to eukaryal transcriptional regulators. GvpE is one protein partner in a positive-

negative regulation pair. While GvpE is the activator protein, the repressor protein GvpD inhibits the activity of GvpE (134).

Another example of archaeal-specific transcription regulation involves sugar transport systems. In *Pyrococcus* there are two distinct ABC transporters that control uptake of the sugars maltose/trehalose (*mal* genes) and malodextrin (*mdx* genes) (64). The expression of the genes encoding these transporters is controlled by the global transcriptional regulator TrmB. TrmB binds to the promoter regions of these genes and inhibits transcription. The sequence to which it binds at the promoters is different depending on which sugar transport system is being regulated. When TrmB is bound to the promoter of the maltose/trehalose transporter genes, the inhibition can be reversed with the addition of either maltose or trehalose. TrmB also binds to the promoter of the malodextrin transporter genes. This binding is reversed by addition of malodextrin. However, only the presence of the sugar whose transport is being inhibited can cause TrmB to release from the promoter. For example, the addition of maltose or trehalose causes the depression of *mal* gene, but not of *mdx* genes. This regulator responds in a different way to two different ligands, depending on the DNA sequence. This type of regulatory mechanism appears to be present in only Archaea.

1.4.3. Role of 5' untranslated leader regions in transcriptional regulation

In bacterial systems, long 5' untranslated leader regions (UTRs) are involved in transcriptional regulation via multiple mechanisms. Usually, these mechanisms involve the formation of secondary structures within the RNA. One example of this type of regulation is the attenuation mechanism, which involves the formation of

different RNA secondary structures that either inhibit or promote transcription elongation (133). This type of regulation employs stem loop structures, along with a long stretch of uridines immediately following one of the structures. Another mechanism found in bacteria involves intrinsic terminators such as riboswitches. In this type of regulatory motif, found in many bacterial systems, the RNA forms different secondary structures that either permit or terminate transcriptional elongation (103). There are multiple permutations of this system present in both eukaryotes and prokaryotes. One of the most well described systems involves regulation of tryptophan biosynthesis by a conformational change in mRNA secondary structure that causes early transcription termination of tryptophan-producing genes in the presence of tryptophan (132). Other systems involve a regulatory protein (103) or ligands (78, 80, 117, 118, 129) that bind to the RNA, which creates a conformational change that affects elongation of the transcript. Early termination of transcriptional elongation in sequence distal to the regulatory UTR is consistent with regulatory mechanisms involving changes in secondary structure, such as attenuation or riboswitches.

There have been reports of 5' UTRs in Archaea. For example, 5' UTRs have been found in the gene for monomethylamine methyltransferase in *M. barkeri* (26), methanol methyltransferase genes from *M. barkeri* (98), and carbon monoxide dehydrogenase genes from *M. thermophila* (114). Although not all of these 5' UTRs have been determined to play a role in regulation in gene expression, there have been a few reports of regulatory 5' UTRs in methanogenic Archaea. One report of a potential regulatory 5' UTR in Archaea is the 113 bp UTR identified upstream of the

gene of a DEAD-box RNA helicase in the Antarctic methanogen *Methanococoides burtonii* (70). This 5'UTR of *deaD*, which contains cold-box elements similar to those found in bacteria, is regulated in response to cold shock. There is also a 5' UTR in *Escherichia coli cspA*, which is a cold induced gene. This 5' UTR has been implicated in transcriptional stability at cold temperatures and instability at higher temperatures (79). The mechanisms of regulation from this 5' UTR are unclear. However, an analogous region in *M. burtonii* has not been confirmed to have a role in regulation.

Another example of a 5' UTR in methanogenic Archaea is the methanol transferase genes *mtaCB1*, *mtaCB2*, and *mtaCB3*. These genes all appear to have long 5'UTRs (18). Deletion analysis of these regions showed that they had an effect on gene expression, indicating that they are important in regulation. However, the mechanism of this regulation was not determined.

Carbon monoxide dehydrogenase/acetyl Co-A synthase (*cdh*) from *M. thermophila* has a 371bp 5' leader region (114). An earlier study involving northern analysis showed that this gene is differentially transcribed in response to different substrates, when growth on acetate and methylotrophic substrates was compared (114). These results were confirmed in a more recent study (5) that employed a translational fusion of the *cdh* promoter region to *lacZ*. In this study gene expression was up to 54-fold and 31-fold greater when cells were grown on acetate as opposed to methanol or trimethylamine, respectively. However, the role of the 5' leader region in regulation was not investigated.

Although there is evidence that suggests 5' UTRs are involved in regulation in methanogenic Archaea, the exact mechanism of their involvement has yet to be determined.

1.5. Objectives of research

The overall goal of this work was to further investigate transcriptional regulation in *Methanosarcina* species, focusing specifically on carbon monoxide dehydrogenase/acetyl Co-A synthase as a model to study the mechanisms of catabolic gene regulation in the Archaea. My hypothesis is that catabolic carbon monoxide dehydrogenase/acetyl Co-A synthase is transcriptionally regulated in response to substrate by a mechanism or mechanisms that involve the 5' leader region.

1.5.1. Objective 1: Functional analysis of an archaeal Rad25 homolog in *Methanosarcina thermophila* TM1

The first objective was to identify putative transcription factors in the genome of *Methanosarcina thermophila* TM1. A putative helicase was identified with homology to *rad25* helicase. Genome sequence analysis was done to characterize the gene, including confirming its expression and mapping the transcriptional start site. Further experiments were also conducted to determine if 1) the protein was functional *in vitro* and 2) had any role in transcription. The results of these experiments are presented in Chapter 2.

1.5.2. Objective 2: Identification of *cis* regions and *trans* factors involved in the regulation of carbon monoxide dehydrogenase/acetyl CoA synthase from *Methanosarcina* species

The second objective of this work was to study regulation of gene expression of a gene involved in the catabolic fermentation of acetate to methane. The gene that was studied was *cdhABCDE* (MA3860), which encodes the five subunit carbon monoxide dehydrogenase/acetyl co-A synthase. This enzyme, as was discussed earlier, is responsible for a critical step in the conversion of acetate to methane. Previous work had indicated that this gene was regulated at the level of transcription (114). The region upstream of the start of transcription of *cdhA*, the first gene in the operon, was investigated for the presence of *cis* elements and *trans* factors that might be responsible for the regulation of this operon. Using the sequence located upstream of the transcription start site, electrophoretic gel shift assays were used to determine if differential protein binding was occurring in this region. Two dimensional gel electrophoresis was used to identify any proteins that might differentially bind to this sequence in the presence of different substrates. These proteins were tested for their ability to bind to the sequence upstream of the transcriptional start site. The results of this research are presented in Chapter 3.

1.5.3. Objective 3: The role of the 5' leader region of carbon monoxide dehydrogenase/acetyl CoA synthase

The third objective was to examine the 5' leader region located downstream of the transcriptional start site of *cdhABCDE* (MA3860). This 5' leader region is highly conserved among *Methanosarcina* species. 5' leader regions are associated with transcriptional and translation regulation in Bacteria, and they have been shown to play a role in regulation in Archaea. However, the mechanism of this regulation in Archaea has not yet been determined. The goal of the research was to determine if this 5' leader region did play a role in regulation of gene expression, and if so, by what mechanism the regulation was occurring. Deletion analysis of the 5' leader sequence was used to determine if this region was involved in transcriptional regulation. *In vitro* transcription assays were also conducted to study the affects of the 5' leader region on transcription. Quantitative reverse transcriptase PCR reactions were completed to study transcript abundance and to determine where in the sequence the regulation was occurring downstream of the transcriptional start site. Results of this work are presented in Chapter 4.

Chapter 2: Characterization of a *rad25* helicase homolog in *Methanosarcina thermophila*

2.1. Abstract

Helicases play an essential role in many cellular functions, including DNA replication, recombination, repair, and gene transcription. A homolog of Rad25, a transcriptional helicase protein, was identified in *Methanosarcina thermophila*. Although eukaryotic Rad25 is a component of the TFIIF complex, which functions in both transcription initiation and nucleotide excision repair, there is no evidence to indicate that the archaeal homolog is required for transcription initiation in Archaea. *M. thermophila* Rad25 has 30% amino acid sequence identity with yeast RAD25 and has high similarity in all conserved helicase domains. Analysis of the *M. thermophila* Rad25 protein showed divergent sequence at the C-terminus in a region normally required for eukaryal nucleotide excision repair, suggesting that this archaeal protein functions either in transcription or as part of a modified repair mechanism. Further analysis indicated that *rad25* was expressed in *M. thermophila*. Using *in vitro* helicase assays, it was also shown that this protein may be capable of 3' – 5' helicase activity.

2.2. Introduction

Helicases are essential in DNA replication and transcription. These proteins couple nucleoside triphosphate hydrolysis to the separation and unwinding of two complementary nucleic acid strands. The Rad25 protein is an important member of

the helicase superfamily of proteins. In yeast, RAD25 was shown to be involved in both transcription and DNA repair, coupling with RAD3 to form part of the TFIIH complex (45). RAD3 helicase functions in a 5' to 3' direction. RAD25 functions in a 3' to 5' direction, which is necessary for DNA replication and transcription. The combination of these two proteins make the TFIIH complex a bi-directional helicase, which is necessary for nucleotide excision repair (NER) (88, 119). Defects in either of the helicases lead to severe genetic diseases in humans such as xeroderma pigmentosum and trichothiodystrophy (33).

In eukaryal transcription, TFIIH is recruited to the pre-initiation complex by TFIIE, and in the presence of ATP catalyzes unwinding of the promoter to form a transcription bubble (120). The Rad25 helicase is needed both for unwinding the double stranded DNA and for release of RNA polymerase II from the pre-initiation complex to allow productive elongation (84). The role of TFIIH seems to be to move along the DNA, suppressing abortive transcription until the elongation complex has cleared the transcriptional start site (116).

In addition to its role in transcription, the yeast RAD25 protein also functions in nucleotide excision repair (NER). During eukaryotic NER, TFIIH and several NER specific proteins are recruited to damaged DNA, making two single-strand incisions. The damaged piece is removed as part of an oligonucleotide fragment (12). NER requires both Rad3 and Rad25 for the process to occur (45, 108). In the archaeon *Methanothermobacter thermoautotrophicus* NER activity is more similar to that of the bacterial system (82). However, this may not be the case for all Archaea, as homologues of the eukaryal NER system are detected in other species (6).

There have been reports of functional helicase activity in Archaea, such as the replicative helicase MCM in *M. thermoautotrophicus* and *Sulfolobus solfataricus* (27, 56). However, activity from other types of helicases such as transcriptional or repair helicases has not been reported in the Archaea. A homolog of the eukaryotic transcription factor TFIIH, which contains Rad25, has not been detected in Archaea. However, sequencing of a number of archaeal genomes has revealed proteins with similarity to eukaryal Rad25 (90). In an effort to identify the functional role of these putative archaeal Rad25 homologs the gene encoding *M. thermophila* Rad25 was identified. This protein was overexpressed to determine if it functioned as a 3' to 5' helicase *in vitro*. The similarity of helicase domains between Rad25 from *M. thermophila* and eukaryal Rad25, combined with the activity of the archaeal protein suggest that this enzyme may have a role in archaeal transcription or DNA repair *in vivo*.

2.3. Materials and Methods

2.3.1. Nucleotide sequence accession numbers

The nucleotide sequence of *M. thermophila rad25* (accession number AF265295) has been submitted to the GenBank database. *Methanosarcina acetivorans rad25* (accession number: NP617309), *Methanosarcina mazei* gene MM3000 (NP635024), *Archaeoglobus fulgidus* gene AF0358 (NP069194), *Pyrococcus horikoshii* gene PH0450 (NP142430), *Thermoplasma acidophilum* gene Ta1016 (NP394476), *S. solfataricus* gene SSO0473 (NP342006), *Halobacterium* NRC-1 gene H1047 (NP046038), *Saccharomyces cerevisiae* gene SSL2 (NP012123),

Arabidopsis thaliana gene MYC6 (BAB08508), *Drosophila melanogaster* gene hay (AAF50150) and *Homo sapiens* gene ERCC3 (NP000113), were used for analysis.

2.3.2. Strains and plasmids

M. acetivorans C2A (=DSM 2834) and *M. thermophila* TM-1 (=DSM 1825), were obtained from sources described previously and maintained as frozen stocks (115). *Escherichia coli* strain SURE (Stratagene) was used for recombinant plasmid construction, *E. coli* strain XL1-blue (Stratagene) was used for phage growth experiments and site-specific mutagenesis constructs, and *E. coli* strain BL21 (DE3) was used for overexpression of Rad25. *M. thermophila* TM-1 DNA, prepared as described previously (116), was used to construct a genomic library with Lambda FIX (Stratagene) according to manufacturers directions. Plasmid pCR2.1 (Invitrogen) was used for cloning PCR amplified products, and plasmid pWM315 was used to subclone *M. thermophila* DNA for sequencing (77).

2.3.3. Media and growth

Artificial marine mineral medium was prepared for growth of *M. thermophila* with 100 mM trimethylamine as growth substrate by methods described previously (112). The medium was prepared anaerobically in a N₂/CO₂ (4:1) atmosphere using a modification of the Hungate technique (7). LB liquid and agar-solidified plates were prepared for growth of *E. coli* (96). Ampicillin (100 µg/ml) or kanamycin (100

µg/ml) was added to the media for selection and maintenance of *E. coli* transformants. *M. thermophila* was incubated at 35°C; *E. coli* was incubated at 37°C.

2.3.4. Cloning and sequence analysis of DNA

Degenerate PCR primers 68 and 69 (Table S2.1) were designed for domains I and V, respectively, of conserved helicase sequences. PCR amplification products (1 min at 94°C, 1 min at 48°C and 2 min at 72°C for 30 cycles using 100 ng of each primer, 3 mM MgCl₂, 0.8 mM dNTP's, 1x Taq buffer and 1.5 units Taq polymerase in a volume of 50 µl) from 1 µg total genomic DNA were cloned into the pCR2.1 vector (Invitrogen) and sequenced by dye terminator cycle sequencing on an ABI 373 automated sequencer (PE Applied Biosystems, Foster City, CA). DNA sequence was analyzed using DNAMAN (Version 3.0, Lynnon BioSoft, 1994) and the Wisconsin GCG program (Genetics Computer Group, 1991, Madison, WI).

2.3.5. Southern analysis

Library screening was performed as described in Molecular Cloning: a Laboratory Manual with 100-fold genome coverage (96) to detect *rad25* homolog in *M. thermophila*. Southern analysis was performed as described using 10 µg total genomic DNA digested overnight with *EcoRI* (96). Samples were hybridized overnight at 42°C and washed twice for 15 min each in 2 x SSC/0.1% SDS at 42°C. The cloned PCR products were digoxigenin (DIG) labeled for use as hybridization probes. Hybridization products were visualized by CSPD luminescence following manufacturers' directions (Boehringer Mannheim).

2.3.6. RNA isolation, Northern hybridization and primer extension analysis

M. thermophila cultures were grown at 35°C in anaerobic minimal marine medium supplemented with 100 mM trimethylamine or 100 mM acetate as growth substrate (3). Cells were harvested during exponential growth (O.D.₅₅₀=0.6 for TMA-grown cells, O.D.₅₅₀=0.16 for acetate-grown cells) by rapid chilling in a dry ice/ethanol bath to 0°C. Cultures were transferred to chilled centrifuge bottles and pelleted at 4°C for 5 min at 9,000 x g. Cells were lysed by resuspension in ice-chilled 30 mM NaOAc (pH 5.2) followed by addition of SDS to a final concentration of 2% and immediate boiling for 30 sec. RNA was isolated by acidic phenol extraction and ethanol precipitation (83, 116).

Northern hybridization was performed with 10 µg or 20 µg total RNA as described previously (116). DIG-labeled probe was made from a gel-purified PCR generated 1358 bp fragment of the *M. thermophila rad25* gene. Samples were hybridized overnight at 42°C in ULTRAhyb buffer (Ambion). Blots were washed twice for 5 min each in 2 x SSC/0.1% SDS at room temperature and twice for 15 min in 0.1 x SSC/0.1% SDS at 42°C. Hybridization blots were developed using CSPD luminescence as described above.

Primer extension analysis was performed on freshly prepared total RNA. Primer 92 complementary to nucleotides 53-30 and primer 93 complementary to nucleotides 171-148 of the *rad25* gene (Table S2.1) were end-labeled with [γ -³²P]ATP using T4 polynucleotide kinase (New England Biolabs) and subsequently purified on a NENSORB 20 column (Dupont-NEN) following manufacturer's directions. Labeled primer was annealed to either 5 µg or 20 µg of *M. thermophila*

total RNA. Controls included 5 µg *M. acetivorans* total RNA and no RNA. cDNA was synthesized using the Thermoscript RT-PCR System (Gibco-BRL) following manufacturer directions. Primer extension reactions were compared to sequencing products prepared with primers 92 or 93 and labeled using [α -³²P] dATP. Template pRad326 DNA for sequencing reactions was prepared and sequenced according to Del Sal et al (29).

2.3.7. Protein purification

Rad25 protein from *M. thermophila* was over-expressed using *E. coli* BL21 (DE3) codon plus cells (Stratagene). To overexpress the protein, the *rad25* gene from *M. thermophila* was labeled with N-terminal His₆-tag, and this protein was expressed using the plasmid pJK107. Cells were grown at 37°C in 0.5 liters of LB broth with Kanamycin until reaching an O.D.₆₀₀ of 0.3, induced by addition of 0.5 mM IPTG for 3 hr, and harvested. Bacterial lysates were prepared by sonication in 0.5 M NaCl, 20 mM Tris, pH 8.0 (buffer A), containing 10 mM imidazole and 20 mg lysozyme. Lysates were centrifuged for 15 min at 10,000 x g, supernatant was extracted and centrifuged for 90 min at 100,000 x g. Protein (7.5 mg) was mixed with 2 ml of Ni chelate resin (Qiagen) for 2 hr at 4°C with gentle shaking. The mixture was loaded onto a column, washed with 40 ml buffer A containing 20 mM imidazole, and eluted with 10 ml buffer A containing 500 mM imidazole. Protein fractions purified to near homogeneity were stored in elution buffer with 20% glycerol at -80°C.

2.3.8. Site directed mutagenesis

A single base pair mutation was introduced into the Walker A box in *rad25* from *M. thermophila* using the QuikChange Site-Directed Mutagenesis Kit following the manufacturer's directions (Stratagene). PCR primers 130 and 131 (Table S2.1) were used to introduce the mutation to change Lys 106 to Glu.

2.3.9. Helicase activity assay

The recombinant Rad25 protein fractions were gel filtered prior to using in the assay. The oligonucleotide used in preparation of the helicase substrate was an 18-mer, 5'-CCCGACTGGAAAGCGGGC-3', referred to as Hel-18. After labeling with [γ - 32 P]ATP (GE Healthcare) using T4 polynucleotide kinase, the oligonucleotide was annealed to 100 nucleotides of the longer oligonucleotide 5'-CTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC TGC(TTTG)₁₅-3' in a ratio of 1:2, and the annealed product was gel-purified as described previously (10). DNA helicase activity was measured as previously described (56) in a reaction mixture (15 μ l) containing 20 mM Tris-HCl (pH 8.5), 10 mM MgCl₂, 2mM DTT, 100 μ g/ml BSA, 5 mM ATP or AMP-PNP (Roche Molecular Biochemicals), 5 fmol of 32 P-labeled DNA substrate (3,000/ cpm fmol) and *M. thermophila rad25* protein. After incubation at 37°C for 1 h, the reaction was stopped by adding 5 μ l of 5 \times loading buffer (100 mM EDTA, 1.0 % SDS, 0.1 % xylene cyanol, 0.1 % bromophenol blue, and 50 % glycerol). Aliquots were loaded onto an 8 %

polyacrylamide gel in 0.5× TBE (90 mM Tris, 90 mM Boric acid, 1 mM EDTA) and electrophoresed for 1 h at 200 V at 4°C.

2.3.10. Disruption of *rad25* from *M. thermophila*

To disrupt *M. thermophila rad25*, pJK107 was digested with *EcoRI* to produce a 1.9 kb DNA fragment containing *rad25*. The partial *rad25* fragment was cloned into pCR2.1 (Invitrogen) creating pKA1. The *pac* cassette, which confers resistance to puromycin, was PCR amplified from pEA103 using primers 157 and 158 (Table S2.1), generating *MfeI* sites flanking the product. Both pKA1 and *pac* PCR product were digested with *MfeI* and ligated to create pKA2. All products were confirmed by sequencing. pKA2 was digested with *EcoRI* to generate a linear fragment that was subsequently transformed into *M. thermophila* as described by Metcalf et al (77). To confirm that the transformation procedure worked, pWM313 was transformed into *M. thermophila*. Transformants were selected for puromycin resistance.

2.4. Results

2.4.1. Isolation of archaeal *M. thermophila rad25*

To identify helicase genes in *M. thermophila*, degenerate PCR primers were designed for domain I, the Walker A type nucleotide binding domain, and for domain V of conserved helicase sequences from mammals and yeast (40). PCR amplification products from total genomic DNA were cloned and sequenced. Multiple products were identified and searched against sequences in GenBank (3). The largest PCR product was approximately 2kb, and the sequence of this fragment showed a strong identity (55%) to the nucleotide sequence from *Archaeoglobus fulgidus* AF0358 (NP069194), previously identified as Rad25 through genome sequencing annotation.

Five positive plaques were detected by Southern analysis of the *M. thermophila* λ phage genomic library with the 2 kb PCR product. Restriction analysis followed by Southern hybridization of one of these clones revealed a single 6kb *EcoRI* fragment with homology to the DNA probe. This fragment was cloned into a low copy number shuttle vector pWM315 to create pRad326. The ends of the insert were sequenced. The insert extended 418 nucleotides and approximately 4 kb from the 3' and 5' flanking region of the PCR amplified fragment, respectively. Sequencing upstream from the 3' end, a gene of 1473 nucleotides was identified that encoded a putative open reading frame (ORF) of 489 amino acids (predicted molecular mass 58kDa) with 25% amino acid identity to RAD25 from *S. cerevisiae*.

Alignment of the predicted translated protein to other Rad25 proteins showed averages of 31% identity between domains I and VI to Rad25 genes from Eukarya,

and 52% to annotated genes identified as Rad25 in archaeal genome sequences (Fig. 2.1). Two DNA helicase sequence motifs GxGK(T/S) and DExH (*M. thermophila* amino acids (aa) 103-107 and 185-188, respectively) are present in *M. thermophila* Rad25, and a bipartite motif for nuclear transport KR_X₁₀RKK (aa 410-424) is present at the C-terminus of the protein. The six domains identified in *S. cerevisiae* (40), indicating helicase function, are also present in *M. thermophila*. Domains I, II, III and VI show the strongest conservation between Eukarya and Archaea (Fig. 2.1). Seven imperfect repeats of a 4 amino acid element QSKS occur at the C-terminus. Repetitive elements of this nature are also found in Arg/Ser rich splicing factors. Interestingly, none of the other identified Rad25 genes contain such a repetitive element.

		Domain I			Domain II
Mt	94	KWGVLVLP TGSGKTLVGIRAI	...	172	EILGNRFGLLIFDEVHHLPAVGY
Ph	87	-r-ii-----iiameim	...	163	-l--dk-m-i-----c-----ea-
Af	89	-r-ci-----h-ama--	...	167	-k-----m-----es-
Hb	100	-r-sv-----flavq--	...	182	ney-dq---vv-----pt-
Hs	334	rs--i---c-a--s---vt-a	...	429	wlktqew--m-l----ti--kmf
Sc	380	rs-ii---c-a-----t-a	...	476	flt-rew-fi-l----vv--amf
		Domain III			Domain IV
Mt	204	PCRLGLTATYEREDGLHTELNRLAG	...	353	TFFIPAITYRTPSKERN SIL
Ph	195	-y-----fp--s-n--el-pd-i-	...	344	k-l-----hk-dk---se--
Af	199	-f-----f-----r-ei-kevv-	...	348	v-l-----h--sre--ee--
Hb	214	-y-----t--e-nv-ed-i-	...	363	e-iv-c--hq-qtd--te--
Hs	461	h-k-----lv---dkivd--f-i-	...	574	rlnk-y-ygp-sqq--mq--
Sc	508	hak-----lv---dkigd--f-i-	...	621	kmgk-f-ygs--qq--mn--
		Domain V			Domain VI
Mt	373	EKFRIG.SYRAVVTSKVLDEGIDVPEANVGV.IVSGTGSKRAYVQRLGRILRKKEG			
Ph	364	k---k-.i-k---s-q-----d-s---.i-----p-eli-----sap-			
Af	368	-g--t-.rf--i-s-q-----d-----.m--s--a-e-i-----psk-			
Hb	383	-r--t-.e-sml---q-----a----i.-l--sa---q-a-----pt-d			
Hs	594	qn-khnpkintifi---g-tsf-l-----liq-s-hg--r-qea-----v--a-k-			
Sc	641	qn-qyndqintifl---g-ts--l---tcliq-s-hy--r-qea-----a-r.			

Figure 2.1. Alignment of Rad25 protein domains. An alignment of the six helicase-defining domains identified in *S. cerevisiae* shows strong conservation in Archaea and Eukarya. The *M. thermophila* Rad25 protein sequence from *M. thermophila* (Mt) is aligned with Rad25 protein sequences from *Pyrococcus horikoshii* gene PH0450 (Ph), *Archeoglobus fulgidus* gene AF0358 (Af), *Halobacterium* NRC-1 gene H1047 (Hb), *Homo sapiens* gene ERCC3 (Hs), and *Saccharomyces cerevisiae* gene SSL2 (Sc). Amino acid sequence identical to *M. thermophila* is indicated by a dash, differences are shown in lower case and gaps are shown by periods. Numerical positions within protein sequences are shown prior to each domain.

All archaeal Rad25-like proteins are truncated compared with the eukaryal proteins, primarily at the N-terminus, but they retain the core elements for helicase function (Fig. 2.2). In *M. thermophila* and the other archaeal species, the residues involved in nucleotide binding (domain I) are conserved, while the C-terminus is both truncated and divergent. These characteristics suggest that the archaeal protein is not involved in the typical eukaryal type of NER.

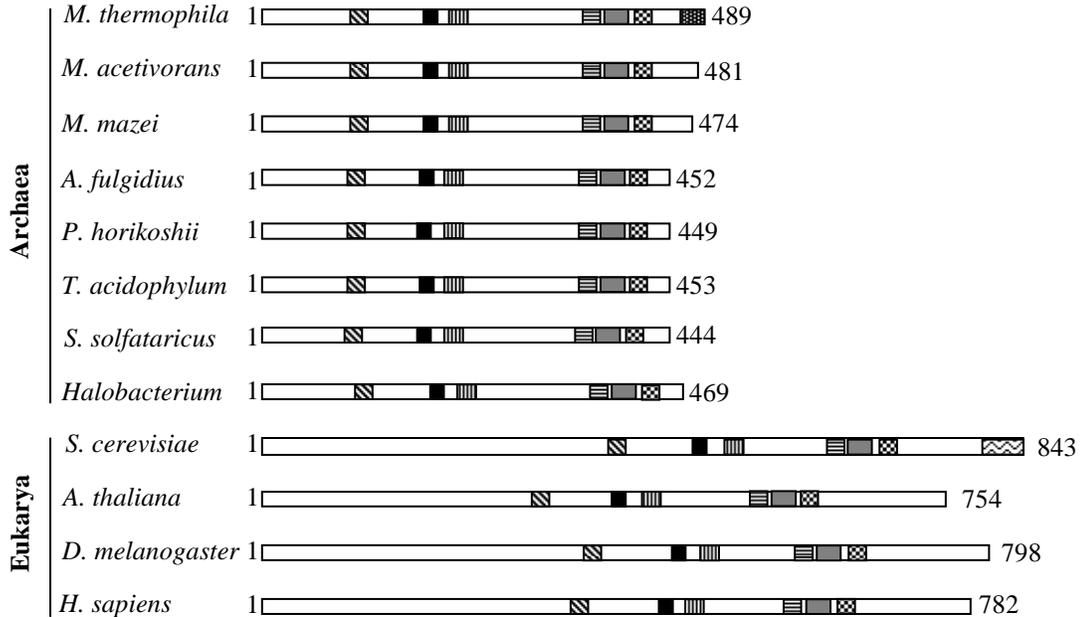


Figure 2.2. Alignment of helicase protein sequences. Sequences shown are *Methanosarcina thermophila*, *Methanosarcina acetivorans*, *Methanosarcina mazei*, *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, *Thermoplasma acidophilum*, *Halobacterium NRC-1*, *Sulfolobus solfataricus*, *Pyrobaculum aerophilum*, *Saccharomyces cerevisiae*, *Arabidopsis thaliana*, *Drosophila melanogaster*, and *Homo sapiens*. Representations are to scale and protein sequence length is indicated. Domains: I, diagonal stripe; II, black; III, vertical stripe; IV, horizontal stripe; V, gray; VI, checkered. Regions: QSKS in *M. thermophila*, dots; region necessary for nucleotide excision repair in *S. cerevisiae*, wavy lines.

2.4.2. Phylogenetic analysis of *M. thermophila rad25*.

The OMNIOME genomic database was searched using BLAST with the predicted *M. thermophila rad25* protein sequence to determine the distribution of this helicase among the Archaea. Protein sequences with significant identities were found in the euryarchaeotes *M. acetivorans*, *M. mazei*, *A. fulgidus*, *Halobacterium* sp NRC-1, *Pyrococcus horikoshii* and *Pyrococcus abyssi*, and in the crenarchaeotes *Pyrobaculum aerophilum*, *Sulfolobus solfataricus* and *Thermoplasma acidophilum*. These species represent mesophilic methanogens, methanogenic and non-methanogenic thermophiles, and extreme halophiles. The widespread presence of this protein among physiologically disparate groups of Archaea suggests that it has a critical function. Closely related genes are also found in non-archaeal species including the Bacteria *Mycobacterium tuberculosis* and *Treponema pallidum*, and the Eukarya *Saccharomyces cerevisiae* and *Homo sapiens*. These genes showed strong similarity in conserved DNA helicase domains, while showing greater divergence at amino and carboxyl termini (Fig. 2.2). The eukaryal sequences show a long amino-terminus extension and a shorter carboxyl-terminus extension. Deletion of the carboxyl-terminal 45 amino acids in the yeast RAD25 protein confers UV sensitivity by disrupting nucleotide excision repair functions of the protein. The archaeal Rad25 helicases do not have this carboxyl-terminus region, suggesting that an alternative repair mechanism exists in the Archaea.

2.4.3. Expression analysis of the *M. thermophila rad25* gene.

Northern analysis was used to determine whether the *rad25* gene from *M. thermophila* was expressed during exponential growth. RNA was hybridized with a DIG labeled 1358 bp probe derived from the coding region of *M. thermophila rad25* gene. Northern hybridization detected two RNA species of approximately 2.7 kb and 1.5 kb in size in *M. thermophila* strain TM-1 (Figure 2.3, lanes 1 and 2). The 2.7kb mRNA is 1.2 kb larger than the expected size for a transcription product consisting of only *M. thermophila rad25* gene, indicating that *M. thermophila rad25* is part of an operon and is cotranscribed with the ORF located directly downstream of *rad25*. The function of this ORF is currently unknown. The 2.7 kb mRNA is expressed when cells are grown on both methylotrophic (Fig. 2.3, lane 1) and on acetate-grown cells (Fig. 2.3, lane 3).

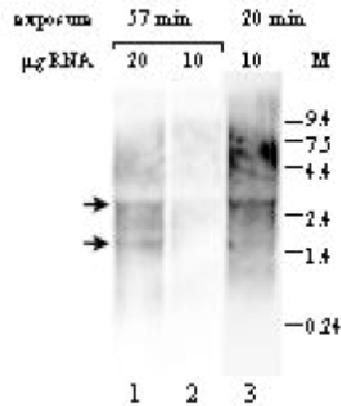
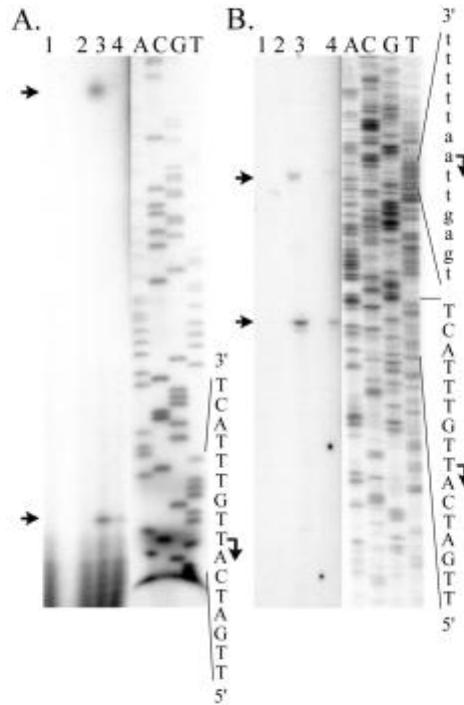


Figure 2.3. Northern analysis of *M. thermophila rad25* expression. A. *M. thermophila rad25* probe, 57 min exposure lanes 1 and 2, and 20 min exposure lane 3. Lane 1, 20 µg and lane 2, 10 µg RNA from TMA-grown cells; lane 3, 10 µg TM1 RNA from acetate-grown cells; M, RNA molecular weight markers, from top to bottom: 9.4 kb, 7.5 kb, 4.4 kb, 2.4 kb, 1.4 kb, 0.24 kb. Figure courtesy of Karin Jackson.

2.4.4. Identification of transcription initiation site.

Primer extension analysis was performed to identify the 5' end of the *M. thermophila rad25* transcript (Fig. 2.4A and B). Labeled primer 92 (Fig. 2.4A) or 93 (Fig. 2.4B) was annealed to either 5 µg (lane 4) or 20 µg (lane 3) of *M. thermophila* total RNA, and controls containing 5 µg *M. acetivorans* control RNA (lane 1) or no RNA (lane 2), and cDNA was synthesized. Primer extension reactions were compared to sequencing products prepared using the same primers 92 or 93. This comparison revealed that the principal start site is the A of an ATG codon at the 5' end of *rad25*. Both primers 92 and 93 demonstrated the same start site in independent experiments. A canonical TATA box element and the B recognition element are both present 23 to 38 nucleotides upstream of the start site of the *rad25* gene. The distance from these elements to the transcription start site conforms to the typical distance in Archaea (Fig. 2.4C) (39, 49, 95). A putative ribosomal binding site is present 78 to 83 nucleotides downstream of the transcription initiation site. A secondary weaker transcriptional start site is 78 nucleotides upstream. The secondary start site does not correlate with any known ORF or tRNA encoding region, nor is there any RNA polymerase recognition sequence upstream of this weaker start site.



C.

```

      -80      -70      -60      -50      -40      -30      -20      -10
      |↓      |      |      |      |      |      |      |
AAAAAAATTAACTCATGCAATAAACTCCAGCCTGAGCCTGCTGATAACTGAATTATTATATATATCATGTCCCGGCTCTTAGTAAACA
                                                                                   ↓↓↓↓
ATGATCAAATCAGCTTCAAGCAGGGAAGTATTCTCATAAAGGAAATGTAAGGGTCCCGAAGTCCATATGGGATGAAAGGAGT
M I K I S F K Q G T I L I K G N V R V P N S I W D E R S

GGAAGCTTCAGGGCTCCTGCAATGTATTACAGGAAATATCAACTACATGAGAGAGTCCAGAATTGATTTTGAAGATAGCGTC
G S F R A P A M Y Y R E I I N Y M R E S R I D F E D S V

CTTGACTTCCTT
L D F L

```

Figure 2.4. Primer extension analysis of the 5' terminus of *rad25* mRNA from *M. thermophila*. Panels A and B: The 5' terminus of *M. thermophila* TM1 M*rad25* mRNA was mapped to the same sites (arrows) using primer 92 (panel A) and primer 93 (panel B) Lanes: **1**, *M. acetivorans* mRNA control; **2**, no RNA control; **3**, 20 µg TM1 mRNA; **4**, 5 µg TM1 mRNA; A,C,G,T, antisense DNA sequencing reactions. Panel C shows the sequence of the 5' end of *rad25*. The putative promoter sequence is in bold. Single arrows indicate a putative ribosomal binding site. Sequences are numbered relative to the transcription initiation sites as determined by primer extension analysis. Figure courtesy of Karin Jackson.

2.4.5. Detection of helicase activity in *M. thermophila rad25*

Recombinant Rad25 was overexpressed in *E. coli*, purified, and assayed to determine whether there was DNA helicase activity. The protein was incubated with a DNA helicase substrate in the presence of ATP or non-hydrolyzable analogue (AMP-PNP) (Fig. 2.5). *M. thermophila* Rad25 unwound the duplex DNA in the presence of ATP (Fig. 2.5, lanes 3-5) but not in the presence of non-hydrolyzable ATP (Fig. 2.5, lane 6). The helicase was shown to be 3' to 5', and no other NTP or dNTP was able to substitute for ATP in the helicase reaction (data not shown).



Figure 2.5. DNA helicase activity of Rad25 from *M. thermophila*. Helicase activity is indicated as primer release in the presence of Mtrrad25 protein. DNA helicase activity assay was carried out with 25ng (lane 3), 75ng (lane 4), and 225ng (lane 5) Rad25 protein in the presence of 5 fmol of substrate and either 5 mM ATP (lane 3 – 5) or AMP-PNP (lane 6). Lane 1 is substrate alone and lane 2 is boiled substrate. Assay done in collaboration with Jae-Ho Shin and Zvi Kelman.

A single base pair change was introduced into the Walker A motif in *rad25*, to determine if this motif was indeed present in the expressed protein, and if this motif was essential, as it is in other helicase proteins. This mutation changed aa 106 from lysine to glutamic acid. If this motif is present and essential for function, the mutation would cause the protein not to function as a helicase. However, it was noted during the purification that there appeared to be other proteins from the recombinant *E. coli* extract within the extract (Fig. 2.6A and B). The purified protein fractions from both pJK107 (Fig. 2.6A), which encoded the recombinant wild type Rad25, and pKH1 (Fig. 2.6B), which encoded the recombinant mutant Rad25, had other proteins in the extract, along with the overexpressed Rad25. These proteins could be responsible for the helicase activity, therefore the results were inconclusive.

2.4.6. Knockout of *M. thermophila rad25*

Gene disruption of *rad25* in *M. thermophila* was attempted using the *pac* cassette as a selectable marker, which confers resistance to puromycin for mutant selection. The *rad25* gene disrupted with *pac* was used in a homologous recombination in an attempt to disrupt genomic *rad25*. Unfortunately, two attempts at disrupting *rad25* yielded only colonies that appeared to be inherently resistant to puromycin, as the *pac* cassette could not be detected in the genome of *M. thermophila* using PCR. The positive control strains, which were transformed with pWM313, were successful both times, indicating that the lack of colonies was not due to procedural error.

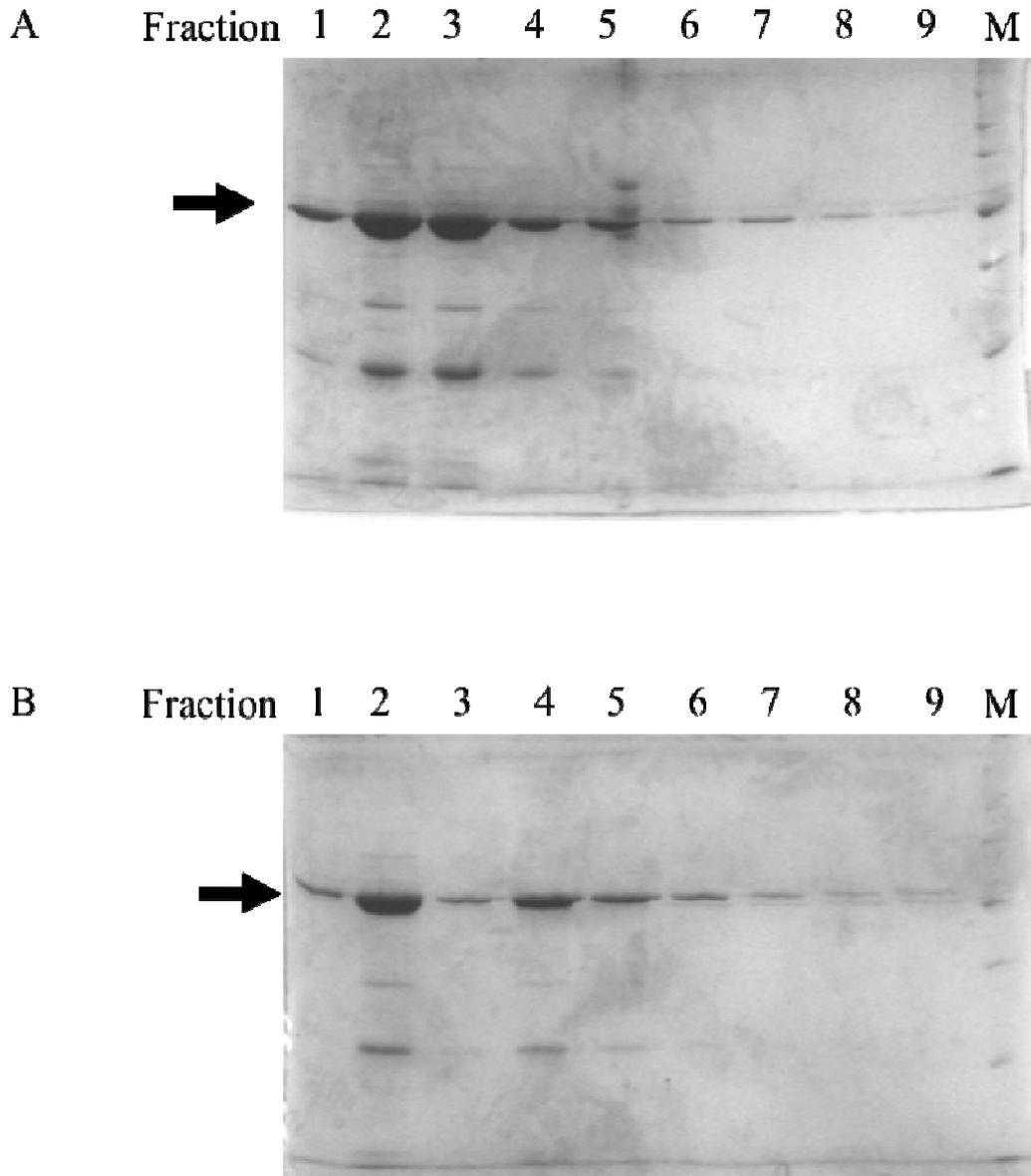


Figure 2.6. Polyacrylamide gel electrophoresis of purified Rad25. Fractions were collected after purification of His-tagged wildtype Rad25 (panel A) and Rad25 with a single basepair change in the 106th amino acid, changing this from lysine to glutamic acid (panel B). 10 μ l of each fraction were run on a polyacrylamide gel. The band marked by the arrow represents Rad25. Numbers at the top of the gel represent the fractions of the purified protein extract. M, protein marker.

2.5. Discussion

The Rad25 protein sequence of *M. thermophila* shows strong similarity to RAD25 protein from yeast, particularly in domains I, II, III and VI. In the yeast protein, mutation of the ATP binding domain results in a lethal phenotype through inhibition of the transcription-related functions of RAD25 (88). This domain is conserved in all archaeal sequences analyzed, indicating that the ATP binding and hydrolysis functions of Rad25 are conserved between the eukaryal and archaeal domains.

In the Eukarya, the Rad25 helicase participates in transcription as a component of eukaryal transcription factor TFIID. Several components of the eukaryal transcription complex such as the TATA binding protein (TBP), TFIIB, and RNA polymerase have homologs in the Archaea (39, 47). Of the human transcription factors, a functional archaeal homolog has been identified for TBP (85, 95), TFIIB (50), TFIIS (48), and TFIIE α (46). The archaeal TFIIE α lacks the C-terminal region, which is associated with the binding of TFIID (46, 83). This observation combined with the lack of evidence for a TFIID complex suggests that either archaeal Rad25 is recruited to the transcription initiation complex by an alternative mechanism or the helicase has an alternative function in the Archaea.

The C-terminal region, whose deletion results in a loss of nucleotide excision repair function in yeast RAD25 (88), is not present in any of the archaeal Rad25 proteins identified by genomic sequence analysis. Deletion and divergence at the C-terminus of *M. thermophila* Rad25 and other Rad25 homologs in Archaea suggest that transcription and nucleotide excision repair, which are linked in eukaryotes, are

not linked in Archaea. In *M. thermophila* a unique QSKS imperfectly hepta-repeated sequence is substituted. This repetitive element was not found in any other Rad25 gene, even those from archaeal species, suggesting a unique function of this protein in *M. thermophila*. Rad25 from *M. thermophila* does not contain the conserved elements for RNA helicase proteins (40, 89, 99), but it does contain the conserved elements involved in DNA helicase activity. Since Archaea lack a homolog of Eukaryotic TFIIF, which contains Rad25, it is likely that *M. thermophila* Rad25 is not involved in transcription, but is possibly involved in DNA replication.

Genes which are similar to those found in bacterial DNA damage repair (*uvrABC*) have been found in Archaea, mostly in the genomes of organisms growing below 55°C (82, 106). There are also uncharacterized proteins in Archaea that contain features suggesting that they comprise a DNA repair system similar to the UmuC-DinB-Rad30-Rev1 superfamily (75). It is unclear from these results whether the Archaea follow a bacterial-like or a eukaryal-like system for DNA repair. However, it is clear that a function for Rad25 in eukaryal-type nucleotide excision repair in Archaea is not supported by analysis of the *M. thermophila rad25* gene. While the promoter element preceding the *M. thermophila rad25* shares considerable nucleotide identity with the canonical archaeal promoter, there appears to be an unidentified open reading frame downstream of *M. thermophila rad25*. Analysis of the expression of *M. thermophila rad25* identified an mRNA fragment larger than the expected translation product. These data indicate that *rad25* from *M. thermophila* is co-transcribed as an operon with the downstream gene, a process common in the

Archaea and Bacteria (14). Future work to elucidate the function of a putative ORF downstream of *rad25* may identify transcriptional components unique to the Archaea.

Partially purified *Methanosarcina* Rad25 protein exhibits DNA helicase activity *in vitro*. Although *rad25* was tested with both 3' to 5' and 5' to 3' specific substrates for DNA unwinding, helicase activity was detected only in the 3' to 5' direction. However, when the partially purified fractions of both the wildtype and single basepair mutant Rad25 were run on polyacrylamide gels, it was noted that there were other products that co-eluted with this His-tagged Rad25. Therefore, it cannot be ruled out that one of those proteins could contain the helicase activity that was detected in the assays. Further experiments need to be done to purify the putative *M. thermophila* Rad25 and to characterize its activity.

Multiple attempts to disrupt gene expression of *rad25* in *M. thermophila* were not successful. In yeast, mutation of the ATP binding domain results in a lethal phenotype through inhibition of transcription-related functions of RAD25 (88). This domain is conserved among the archaeal sequences analyzed in this paper. Because of this, it could be possible that disrupting *rad25* causes a lethal mutation, which is why a disruption mutant has not been obtained. Further experiments need to be done to try and disrupt this gene, including a partial disruption of the carboxyl terminus. Since fully disrupting this gene may result in a lethal phenotype, a partial disruption may not be lethal to the organism, and the effects of the deletion on the physiology of the cells could be studied.

Genome sequence analysis identified helicases in many archaeal species encompassing members of all branches of Euryarchaeota and Crenarchaeota. These

helicases also contain conserved domains that were found in other organisms. It is clear from these analyses that *rad25* helicase genes have a broad presence in Archaea. This observation combined with experimental evidence that the gene is expressed and its deletion is lethal supports the conclusion that this gene has a critical role in cell function. Since this gene appears to be conserved among different members of the Archaea, its function is likely to be important for cellular processes. Future experimentation will be directed at identifying the specific role of this helicase in *M. thermophila*. One approach to determining the function of this protein would be to delete different regions of the protein, such as the carboxyl terminus, to determine what the effect was of the deletion. Another approach is to use quantitative reverse transcriptase PCR (qRT-PCR). Cells would be exposed to ionizing radiation, and the mRNA levels of *rad25* transcript would be compared to wild type cells to see if this gene is upregulated when cells are exposed to radiation.

The current work described in this chapter affirms that this gene is indeed expressed. *M. thermophila* Rad25 does not contain the conserved elements for RNA helicase activity. *M. thermophila* do not contain a homolog of TFIIH, the transcription factor which contains Rad25 in Eukaryotes. These two pieces of evidence indicate that Rad25 is most likely not involved in transcription. It does, however, contain conserved elements for a DNA helicase, indicating that it may have a role in DNA replication. Further work, however, needs to be done in order to confirm the role this protein has in Archaea.

Chapter 3: Identification of *trans* factors involved in the regulation of carbon monoxide dehydrogenase in *Methanosarcina* species

3.1. Abstract

Expression of the archaeal CO dehydrogenase/acetyl-coenzyme A synthase (CODH/ACS) operon (*cdhABCDE*), which catalyzes the dismutation of acetyl CoA in the acetoclastic methanogenic pathway, is highly regulated by *Methanosarcina* spp. in response to growth on acetate versus the methylotrophic substrates methanol and trimethylamine. However, the mechanism of this regulation is unknown. Many of the known mechanisms of regulation within the Archaea involve repressor or activator proteins that bind upstream of the start of transcription to either inhibit or activate transcription. Using the sequence upstream of the transcriptional start site from -256 to -76, electrophoretic mobility shift assays revealed differential protein binding between extracts of *Methanosarcina acetivorans* grown with trimethylamine or acetate. This same sequence was used as a substrate in affinity chromatography, which confirmed the differential binding. Affinity chromatography, coupled with two-dimensional gel electrophoresis, allowed for the identification of differentially binding proteins. CdhA, CdhB, and CdhD were all identified as binding to the sequence upstream of the transcriptional start site. Competition assays to test the binding of a mostly purified extract containing the Cdh complex indicated that this extract contained proteins able to bind to DNA, but the binding was not specific. Using a LacZ expression strain, it was shown that deletion of the region upstream of

the transcriptional start site did not have an effect on expression. Based on the results of these studies, it was concluded that *cdhABCDE* is not regulated via *trans* factors located upstream of the transcriptional start site. In contrast to other archaeal genes that are regulated by binding of either a repressor or activator protein to DNA in proximity of the promoter, the CODH/ACS operon does not appear to be regulated by mechanisms that require activators or repressors.

3.2. Introduction

Approximately 70% of the biogenic methane in the environment is derived from acetate (71) and the regulation of the genes involved in the acetoclastic pathway likely has a significant impact on anaerobic degradation processes and the global biogenesis of methane. There are only two genera reported to date whose members are capable of using acetate as a substrate to produce methane: *Methanosarcina* and *Methanosaeta* (111). While *Methanosaeta* species are obligately acetoclastic, *Methanosarcina* species are able to grow on a variety of substrates, including CO₂ with hydrogen, methylamines, methanol, and acetate. Catabolic genes in *Methanosarcina* spp. are highly regulated in response to substrate. However, when given a choice between two substrates, *Methanosarcina* species will preferentially use the substrate with the higher free energy, like the methylotrophic substrates trimethylamine and methanol, before using acetate (17, 34, 60). The high metabolic diversity, combined with the availability of a tractable genetic system and complete genome sequence from three species, makes members of the genus *Methanosarcina*

spp. ideal models to study the regulation of catabolic gene expression in the methanogenic Archaea.

There are several enzymes involved in acetoclastic methanogenesis, including carbon monoxide dehydrogenase/acetyl CoA synthase (CODH/ACS) (1, 2, 43, 61, 72). This enzyme catalyzes the dismutation of acetyl CoA. The genes encoding this subunit include an operon of 5 genes, with a 5' leader region located downstream of the transcriptional start site. Previous studies have shown that this gene, *cdhABCDE*, is regulated in response to substrate (5, 114, 121). It has also been demonstrated that this gene is regulated at the level of transcription (114). However, the mechanism of this regulation is not understood.

Transcriptional regulation in Archaea has been a topic of many studies. For example, within the literature, there are descriptions of transcriptional activators. Some of these activators include *Methanocaldococcus jannaschii* Ptr2 (87), *Halobacterium* Bat (9), and *Sulfolobus solfataricus* LysM (22). There are also examples of repressor proteins, such as *Pyrococcus furiosus* LrpA (15), *Methanococcus maripaludis* NrpR (68), and *Archaeoglobus fulgidus* Mdr1 (86). All of these transcriptional regulators bind to DNA upstream of the transcriptional start site to affect transcription.

The purpose of this study was to identify *trans* elements that mediate expression of the catabolic CODH/ACS complex in the acetoclastic *Methanosarcina*. The regions both upstream and downstream of the transcriptional start site from *Methanosarcina acetivorans* C2A *cdhABCDE* were amplified with PCR and used as a template to determine if differential protein binding was occurring upstream of the

transcriptional start site. Differential binding was detected and further analyzed to identify and characterize the proteins that appeared to be capable of binding.

3.3. Materials and Methods

3.3.1. Methanogen strains and media

Methanosarcina acetivorans C2A (=DSM 2834), *Methanosarcina thermophila* TM-1 (=DSM 1825), and *Methanosarcina barkeri* Fusaro were obtained from sources described previously and maintained as frozen stocks (115). Artificial marine mineral medium was prepared for growth of methanogenic Archaea by methods described previously (113). Growth substrates used were sodium acetate or trimethylamine-HCl (TMA) at a final concentration of 0.05 M. Methanogen medium was prepared anaerobically in a N₂-CO₂ (4:1) atmosphere by a modified Hungate technique (110). *M. acetivorans* strains were incubated at 35 °C.

3.3.2. Electrophoretic Mobility Shift Assay

Electrophoretic mobility shift assays (EMSA) involved three different 100bp double stranded DNA oligonucleotides (dsDNA) generated by annealing two complimentary oligonucleotides. The synthesized oligos were used as substrates for binding of proteins from cell free extract. The sequence of the first oligonucleotide was 5'-

CTATATAATTATTTGTTTTTAATATGAAAAAAGAAATAATAAATAACAA
ATATTAAAATCGACGAAAAAAATATAGAGGCTGAAGAAGGTCAAATCA

A-3', and the second oligonucleotide sequence was 5'-

GCTGAAGAAGGTCAAATCAAAAAAACGGCAACATAAAGCCTATAGACTA
TTTTTCACGAAATTGATCTTCGAGATTTTTTTTACGAATAGTGTGCCTCTA-

3'. These two oligos were from a region -256 to -156 and -176 to -76 bp upstream of the transcriptional start site. A third oligonucleotide from the coding region of MA1017 was used as a control for nonspecific binding of proteins: 5'-

CGTCTTTGCGACCTTCTCCAGATCGAGCAGAATGAGGCTGCTGACCGGAT
GGGAATATCCCGAAAACTTTCTGGAGCGACCTCCAAAGGGCACGGCAA

A-3'. This oligo was used to look at nonspecific binding of proteins, to rule out any proteins that were able to bind to both this substrate and the specific substrate. DNA labeling with [γ - 32 P]ATP and purification were performed as described previously (92). Cell-free extracts of acetate- and trimethylamine-grown *M. acetivorans* in exponential phase were made by resuspending 2 g of cells in 4 ml 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.0, containing 1 mM phenylmethylsulphonylfluoride. The cells were lysed by passing cell suspension through a French pressure cell at 140 mPa and lysate was centrifuged at 10,000 x g for 20 min to obtain cell-free extract.

Complexes formed between the proteins in the cell-free extract and dsDNA were detected by a gel mobility shift assay in reaction mixtures (15 μ l) containing 20 mM HEPES-NaOH (pH 7.5), 10 mM MgCl₂, 1 mM dithiothreitol, 100 μ g/ml bovine serum albumin, 2 μ L (60-100 fmol) of one of three 32 P-labeled oligonucleotides mentioned above (3000-6000 cpm/fmol), and 1 μ L (20 μ g) of cell-free extract from either acetate-grown or trimethylamine-grown cells. Samples were incubated at either 4°C or 37 °C for 30 min. This was done to inhibit any nuclease activity in the

extract. 5 µl of 5x loading buffer (0.1% xylene cyanol, 0.1% bromphenol blue, 50% glycerol) was added to the reaction. Aliquots of the reaction mixture were electrophoresed for 4 h at 150 V through an 8 % polyacrylamide gel containing 5 % glycerol in 0.5x TBE (45 mM Tris, 45 mM boric acid, 1 mM EDTA). The images were visualized using a Storm phosphoimager (Molecular Dynamics).

3.3.3. Affinity chromatography of DNA-binding proteins.

Both the sequence upstream of the transcriptional start site and the 371 bp leader region were amplified by PCR. The 204 bp sequence upstream of the promoter was amplified from *M. acetivorans* genomic DNA with primers EMSA1-bio and EMSA2 (Table S3.1), which created a 204 bp product labeled with biotin. The 371bp sequence immediately downstream of the promoter was PCR amplified from both *M. acetivorans* and *M. thermophila* with primers UTR1-bio and UTR2 (Table S3.1), which created a biotin labeled product. The biotin labeled-PCR products (1 µg) were bound to streptavidin-coated paramagnetic beads (Promega). The DNA-bead complex was incubated for 30 minutes at room temperature with 400 mg cell free extract from either acetate- or TMA-grown *M. acetivorans*, which were made by resuspending 2 g of cells in 4 ml 10 mM HEPES, pH 7.0, containing 1 mM phenylmethylsulphonylfluoride. The cells were lysed by passing cell suspension through a French pressure cell at 140 mPa and lysate was centrifuged at 10,000 x g for 20 min to obtain cell-free extract. The DNA-bead complexes were washed to remove any unbound protein, and DNA-binding proteins were eluted in 0.1% SDS. The eluted proteins were run on a 10% SDS-PAGE and were silver stained.

3.3.4. Two-Dimensional analysis of DNA binding proteins.

The DNA region used as a substrate for the gel mobility shift assay was also used to purify potential DNA binding proteins. The 204 bp region was amplified by PCR from *M. acetivorans* and labeled with biotin using primers EMSA1-biotin and EMSA2 (Table S3.1). The biotin labeled-PCR product was bound to streptavidin-coated paramagnetic beads (Promega), and this was incubated with 900 mg cell free extract from either acetate- or TMA-grown *M. acetivorans*. The DNA-bead complexes were washed to remove any unbound protein, and DNA-binding proteins were eluted in 0.1% SDS. 9 mg of eluted proteins were run on using 2-dimensional gel electrophoresis, the gels were silver stained and analyzed to determine differential binding. Two dimensional electrophoresis was conducted according to the principles of O'Farrell (81) and as outlined by Gorg et al. (41) and Brady, et al. (19). A total of 14 spots were chosen for identification by the Mass Spectrometry Lab at the University of Texas Medical Branch using peptide mass fingerprinting.

3.3.5. Competition assays

Competition assays were used to determine the specificity of binding of Cdh complex. This method used one 204 bp oligo that was generated by PCR using primers EMSA1 and EMSA2 (Table S3.1) using genomic DNA from *M. acetivorans*. This oligo was located beginning at -251 bp upstream of the start of transcription. Unlabeled specific competitor DNA was generated using the same PCR primers. Unlabeled nonspecific DNA was generated from *M. thermophila* DNA using primers 292 and 293 (Table S3.1), which amplified a 200bp region of *rad25* within the

structural gene. The assay was performed as described above, except partially purified Cdh was substituted for cell-free extract. 16µg of partially purified carbon monoxide dehydrogenase complex (Cdh) from *M. thermophila*, a gift from David A. Grahame, was incubated with 50ng labeled oligo in the presence of increasing amounts of unlabeled specific and nonspecific competitor DNA to confirm the binding ability of this protein. Excesses of 50X, 100X, and 200X of both specific and nonspecific competitor DNA were used. This method was also used to test for binding of the high molecular weight extract to the same 200bp region of *M. thermophila* and *M. barkeri*. The 200 bp templates for these reactions were located at -251 bp upstream of the start of transcription and were generated using primers TM1up1 and TM1up2 for *M. thermophila* and Barkeri-up1 and Barkeri-up2 for *M. barkeri* (Table S3.1).

3.3.6. Construction of *cdh'*::lacZ deletion plasmids

Deletions upstream of the transcriptional start site were generated from the pEA64 (Table 3.1) template using reverse primer 60 (Table S3.1), which hybridized within the plasmid with forward primers that generated *ClaI* restriction sites at different locations upstream of the *cdh* promoter. Forward primer 62 generated a *ClaI* restriction site 262 bp upstream of the transcriptional start site to create Δ62. Forward primer 63 generated a *ClaI* restriction site 71 bp upstream of the transcriptional start site to create Δ63. Forward primer 64 generated *ClaI* restriction site 29 bp upstream of the transcriptional start site to create Δ64, which deleted part of the *cdh* promoter TATA box. PCR generated DNA fragments were cloned directly into pCR2.1

(Invitrogen), then digested with *NsiI/SalI* and cloned into the *NsiI/SalI* restriction sites of plasmid pcdhp::lacZ. The latter plasmid was digested with *XhoI/BamHI* and the fragment containing the cdh-lacZ fusion was ligated into the lacZ reporter shuttle vector pEA103 replacing wildtype *cdh* sequence. Once the plasmids were transformed into *M. acetivorans*, the sequence of each plasmid was verified by extracting the plasmid and sequencing for confirmation.

Table 3.1. *E. coli* and *M. acetivorans* plasmid constructs.

Plasmids	Parent	Characteristics or description of construct	Source
pCR2.1	-----	Cloning vector, Amp ^R , Pur ^R , lacZ α	Invitrogen
pCDH1.3	-----	cdhA', Amp ^R	(114)
pcdhp::lacZ	-----	cdhA'::lacZ, Amp ^R , Ori pBR322	(5)
pEA64	pEA61	Cdh'::CAT, Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pEA103	pWM307	Cdh'::lacZ, Ori pC2A, Amp ^R , Pur ^R , Ori R6K	(5)
pEA108	pEA103	Cdh' Δ 62::lacZ, Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pEA109	pEA103	Cdh' Δ 63::lacZ, Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pEA110	pEA103	Cdh' Δ 64::lacZ, Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study

3.3.7. Reporter gene expression analysis.

M. acetivorans transformed with either the wild-type reporter plasmid pEA103 or deletion plasmids were inoculated (10% v/v) in triplicate into liquid medium containing 0.1 M sodium acetate, TMA, or methanol and incubated to mid-exponential growth phase. Cultures were sampled (1.0 ml) to perform β -galactosidase assays as described previously (5). 1 ml samples were centrifuged at 13,000 rpm for 5 minutes. Cell pellets were resuspended in 1 ml Z Buffer (60 mM Na₂HPO₄⊕7H₂O, 40 mM NaH₂PO₄⊕2H₂O, 10 mM KCl, 1 mM MgSO₄⊕7H₂O, pH 7.0). Cells were sonicated and centrifuged at 13,000 rpm for 15 minutes. The

supernatant was transferred to a new tube. The protein concentration of the supernatant was determined using the Bradford assay. Triplicate samples of the supernatant (100 μ l) were combined with 700 μ l Z Buffer + β -mercaptoethanol (27 μ l/ 10 ml Z Buffer). The mixture was incubated at 30 $^{\circ}$ C. 160 μ l ONPG (4 mg/ml) was added to each mixture, and the time for the color to change to yellow was measured. The reaction was stopped with the addition of 400 μ l 1 M Na_2CO_3 , and the reactions were assayed for absorbance at 420 and 550 nm.

3.4. Results

3.4.1. Identification of putative protein binding sites using electrophoretic mobility shift assays.

In four methanosarcinal species, the region located downstream of the transcriptional start site of *cdhABCDE* is highly conserved, with similarity between 80-90%. In contrast, the overall sequence upstream of the transcriptional start site was less conserved between the four methanosarcinal orthologs, and the formation of secondary structures was not evident. However, short lengths of sequence were conserved among species, which suggested that protein binding might occur in these regions (Fig. 3.1). Electrophoretic mobility shift assays (EMSA) were conducted to identify potential regions of protein binding. Two radiolabelled 100 bp oligonucleotides, oligo #1 and oligo #2, whose sequences were located at -256 to -156 bp and -176 to -76 bp, respectively, were incubated with cell-free extract of *M. acetivorans* grown with acetate and trimethylamine to determine if the catabolic substrates caused differential protein binding to the oligonucleotides (Fig. 3.2). Differential mobility shifts were observed between the acetate-grown cell-free extract (extract A) and the trimethylamine-grown cell-free extract (extract T) with both oligonucleotides. Incubation of the reaction mixture at both 4°C and 35°C did not have a significant effect on the banding patterns. The two different temperatures were used to control for nuclease activity within the cell free extracts.

<i>M. thermophila</i>	TATGAAGAATCTTTAAAATGAGTTGTGGTTAGTATGAAAAATGAAAAATAATAACAA	-208
<i>M. acetivorans</i>	--g--a-----a--t---t-t---tt---a-----a---t--ta-----	-208
<i>M. mazei</i>	--a--a-----gggct-t---cac-----g-----att-g-----	-207
<i>M. barkeri</i>	--ac--at--t-ac-t---a-t---t-g-ac-c--gt---aagcgt--t-gt---t--	-205
<i>M. thermophila</i>	ATATTAAAATCAATGGAAAAACTATTGGGAATCAAGAATACTCATTCAAAAAACCAGT	-148
<i>M. acetivorans</i>	-----g-c-a-----a--a-a-gc-g-----ggtca-a-----ggca	-148
<i>M. mazei</i>	-----ata-----a-----a--tc---t-t-----tcatt-----g-ta	-147
<i>M. barkeri</i>	t--a--gc-ca--g..-gt---aca-ca-tg-t-ta-tggt-tt-a-tg-----cac	-147
<i>M. thermophila</i>	ATCTACAGGCATTGAGGGCTTTTCTTTTCGGAGTAGATCTTCAAGA.TTTTTTATGAAGA	-89
<i>M. acetivorans</i>	-ca--a-.-c-at--acta---t-ca--a-a-t-----g---t-----c---t-	-89
<i>M. mazei</i>	tg--ttgacatagagtat----t-gatca-aa--tct--g-g-tt-----a-ga-ga-	-87
<i>M. barkeri</i>	-----c-a-tct----t-g-accta--at-acaat-a-c--g-aag-c--ct	-87
<i>M. thermophila</i>	AGGTATCACTAAATTCATTTTACCCATTGCGAAATAGGGTGTACAAATTAGAAAATGT	-29
<i>M. acetivorans</i>	gt--gc-t---c-----c-----ca-----a-----	-29
<i>M. mazei</i>	c--gtctc-ac-t-c-at-----tg--cg-..---a-----	-29
<i>M. barkeri</i>	gat-t-g-ac-ttcc-ac--ac--atc-cga--tgc-..-----a-	-29
<i>M. thermophila</i>	TATATAGATTTGGTACATTTAGACTTTTAATTAGTGTATTATTATCCATCGGTAGCGACCT	+32
<i>M. acetivorans</i>	-----t-ac-----	+32
<i>M. mazei</i>	-----g-a-----a-----	+32
<i>M. barkeri</i>	-----g-a-----a-----	+32
<i>M. thermophila</i>	CTGCTCAAGATTAAGGTCCCAATCGTTGGTAAAAACGGTTTTTTTTGAGACATAGGGCGTA	+92
<i>M. acetivorans</i>	-----t-----c-----a--t-----	+92
<i>M. mazei</i>	-----t-----c-----a-a-----	+92
<i>M. barkeri</i>	-----t-g-----t-----c-----a-----	+92
<i>M. thermophila</i>	AAAAGGGATTTGGAAGAGAAAATAAGGCCCTAAAAACGACTTTTACAAG.AATATATTCTC	+151
<i>M. acetivorans</i>	---a-a---a-----t--t-a-----gg---g-a-----c-	+151
<i>M. mazei</i>	---a-----t-----c-t-g-t-----t-----a-a-----c-	+152
<i>M. barkeri</i>	---t--c---.---t-tg-at-a-----t-t-.g-a----tc-	+151
<i>M. thermophila</i>	TATAAAAACGTTCTTTATTTGAAAAAACGTTGGTATAACGTTTAATGTCCAAATTGAGAG	+211
<i>M. acetivorans</i>	-----gct-----t-----g-----g-g--t-----c-----	+210
<i>M. mazei</i>	--c-----c-t-----c-----a--c--c--a-t-----a-----	+211
<i>M. barkeri</i>	-c-----act---gc--g-----cgc-g-ta-a-tat--ggtat-tg---aaa--t-	+211
<i>M. thermophila</i>	CTATGAACGTTTTACCATAAAAACGTCAT...TTAAACGGTTGTAACCTTTGAAGCAGACGG	+268
<i>M. acetivorans</i>	---aa-----g---g---t-c---.-----a-g--g-----	+267
<i>M. mazei</i>	---a-gt---g---g-----..-----a--a-----	+268
<i>M. barkeri</i>	g-gaat-gtg---ta-cag-tg--at-tac---c-----a-.------	+270
<i>M. thermophila</i>	TAAGATATAAAATGAATAATCATTAACAATATGCCGCTT..ATCAGGATTACAATCGAAT	+326
<i>M. acetivorans</i>	---g---t---t-----g-----a-.-a....--ag--gc--cca	+321
<i>M. mazei</i>	c--a-----t---t-----g-----aag-.-a-.-g-----tc-	+325
<i>M. barkeri</i>	c--gag.....-tt-gg-a---ca-----aagaat-a-aat--g-----tc-	+325
<i>M. thermophila</i>	TGGCGAAGTCGACTTTAAACAAATTAAGGAGGTAAAGCTCAC <u>ATG</u> AGCAAACCTAACTACC	+386
<i>M. acetivorans</i>	-----t-----a-----	+361
<i>M. mazei</i>	-----t-----t-----a-----	+365
<i>M. barkeri</i>	--at-----t-----t-----a-----	+365
<i>M. barkeri</i>	-----a-	

Figure 3.1. Alignment of *cdhA* promoter region from four *Methanosarcina* species. Numbering in right column is based on the transcription start site for *cdhABCDE* (114). The start of transcription is marked with an arrow and +1. The start of translation (ATG) is underlined.

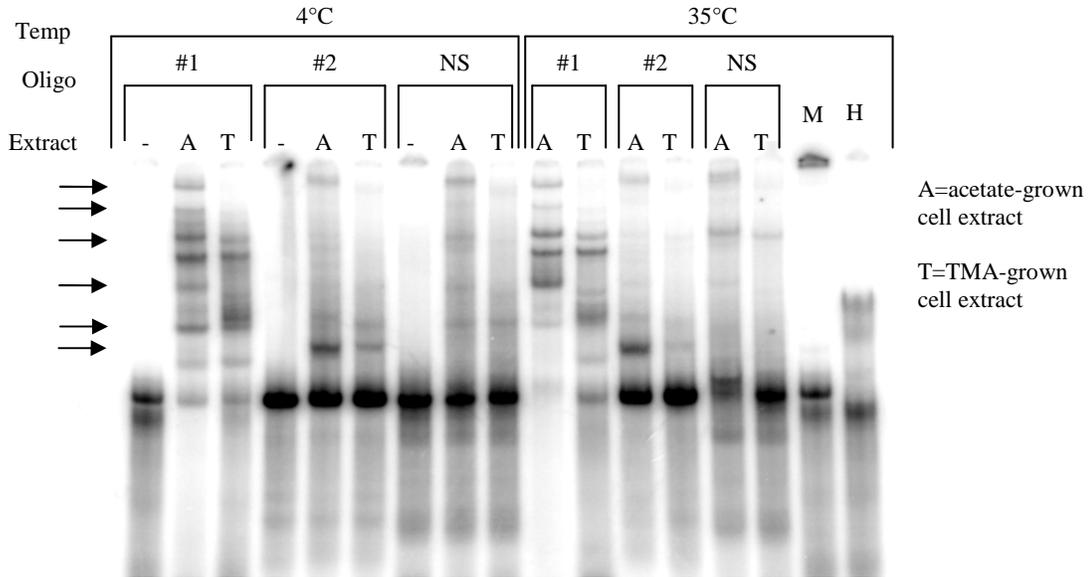


Figure 3.2. Electrophoretic mobility shift assay of DNA fragments from the region upstream of the *cdhA* promoter. Each band represents a dsDNA fragment with bound protein. Minichromosome maintenance (M) protein and HmtB histone (H) from *Methanothermobacter thermoautotrophicum* were used as positive controls. Oligonucleotides 1 and 2 are 100bp fragments beginning 256bp and 176bp upstream of the start of transcription, respectively. Oligonucleotide NS is a 100bp fragment from a structural gene (MA1017), used as a non-specific sequence. 20 μ g of either acetate-grown cell free extract (A) or TMA-grown cell free extract were added to each mixture. Reaction mixtures were incubated at 4°C and 35°C for 30 minutes. Lanes labeled with (-) for extract represent samples in which no extract was added. Arrows indicate some of the differentially shifted bands.

3.4.2. Isolation and sequence analysis of DNA binding proteins using affinity chromatography.

It was investigated whether differential protein binding could be the mechanism of regulation within the 371 bp untranslated region using modified affinity columns. Proteins that were able to bind to this DNA region were separated by running extracts over streptavidin-labeled paramagnetic beads bound to biotin labeled DNA. The DNA used for this experiment was PCR generated fragment of the 5' leader sequence located downstream of the transcriptional start site. The proteins that could bind were eluted and visualized to determine if there was differential protein binding. Differential protein binding patterns were observed between the acetate and trimethylamine grown cell free extract, but the binding patterns for each extract were the same as those present within the negative control columns (lanes 1 and 2) that contained affinity beads with no DNA attached (Figure 3.3). This indicates that the binding detected was nonspecific in nature, and that differential protein binding to DNA in this region is most likely not the cause of the difference in expression.

Extract	A	T	A	T
Lane	1	2	3	4

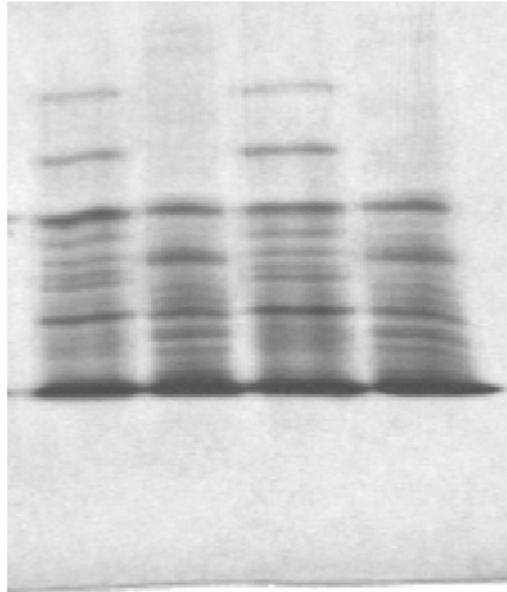


Figure 3.3. Polyacrylamide gel electrophoresis of proteins isolated from affinity chromatography with the 5' 371 bp leader sequence DNA of the *cdh* operon bound to biotin beads (lanes 3 and 4). 400 mg of cell free extracts from acetate-grown cells (A) or TMA-grown cells (T) were passed over the columns. Proteins that bound were eluted and run on the gel. Lanes 1 and 2 were negative control columns, without bound DNA.

To identify which DNA binding proteins that bound upstream of the transcriptional start site might be involved in differential expression of *cdh*, DNA sequence that was used in the EMSA was labeled with biotin and bound to streptavidin-coated paramagnetic beads. Proteins that bound to this DNA region were eluted and analyzed using two-dimensional gel electrophoresis (2DGE) (Fig. 3.4).

There were several proteins that differentially bound to the DNA, both in the acetate and the TMA-grown cell-free extract. Fourteen of these proteins were eluted from the gel for identification by peptide mass fingerprinting (Table 3.2). Some of the eluted proteins could not be identified, either due to low quantity of the protein or because the protein sequence was not in the databases. Protein spots 1, 2, and 3 from acetate-grown extract were identified as three products of the *cdh* operon, CdhA (alpha subunit), CdhB (beta subunit), and CdhD (delta subunit). Protein spot 14 from trimethylamine grown extract was identified as a corrinoid protein associated with trimethylamine methyltransferase.

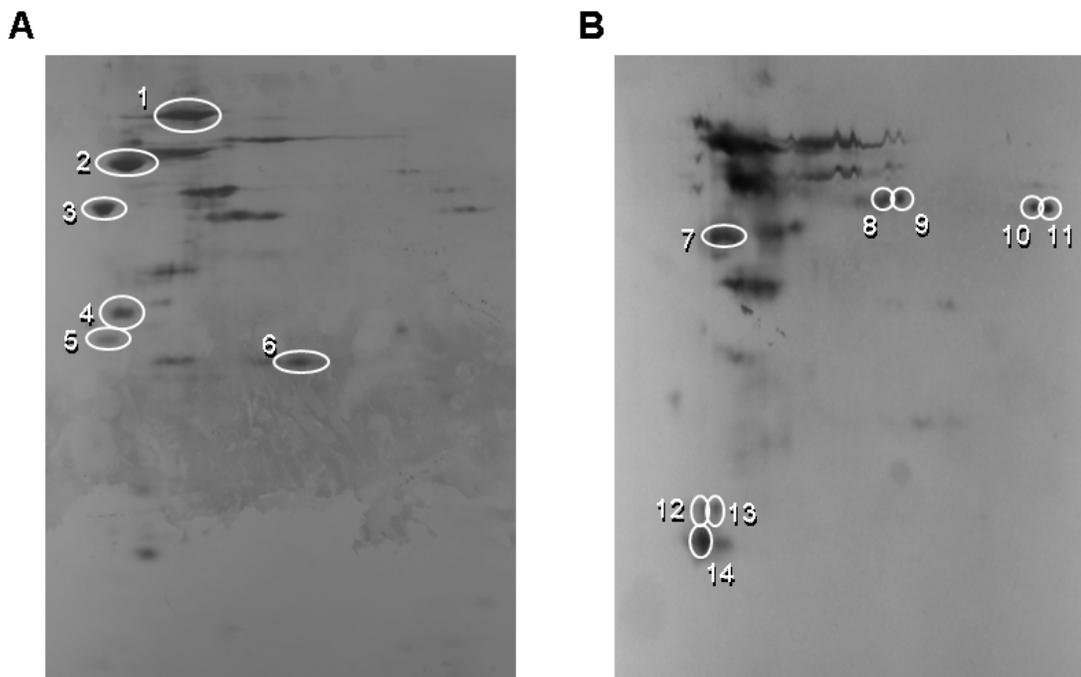


Figure 3.4. Two-dimensional gel electrophoresis (2DGE) of DNA-binding proteins from acetate- and TMA-grown cell free extract. 1 μ g of 200 bp DNA template, which was located -251 bp upstream of the transcriptional start site of *cdhA*, was bound to streptavidin paramagnetic beads and incubated with 900 mg of either acetate-grown or TMA-grown cell free extract. 9 mg of eluted proteins were run on 2DGE. Panel A, cells grown with acetate. Panel B, cells grown with trimethylamine. Circles and numbers refer to spots that were identified through peptide mass fingerprinting.

Table 3.2. Identification of protein spots from two-dimensional gel electrophoresis.

Protein Spot	Identification
Acetate CFE ¹	
1	Acetyl-CoA decarboxylase/synthase complex alpha subunit 1
2	Acetyl-CoA decarboxylase/synthase complex beta subunit 1
3	Acetyl-CoA decarboxylase/synthase complex delta subunit 1
4	ND ³
5	ND
6	Methyl coenzyme-M reductase, gamma subunit (MA4847)
TMA CFE ²	
7	DNA topoisomerase VI subunit A
8	Thiamine biosynthesis protein ThiC (MA4239)
9	Thiamine biosynthesis protein ThiC (MA4239)
10	Elongation factor Tu (MA1256)
11	Elongation factor Tu (MA1256)
12	ND
13	ND
14	Trimethylamine corrinoid protein (MA0529)

¹Acetate CFE = cell free extract from acetate grown *M. acetivorans*

²TMA CFE = cell free extract from trimethylamine grown *M. acetivorans*

³ND = unable to determine identity

3.4.3. Specificity of Cdh binding

To confirm the DNA-binding abilities of Cdh, partially purified Cdh complex from *M. thermophila* was used in a gel mobility shift assay (Fig. 3.5). The substrate for this assay was the DNA sequence from *M. acetivorans* used previously for affinity chromatography. A competition assay was also conducted to determine if the binding of Cdh was specific. An excess of unlabeled specific or nonspecific competitor DNA was added to the reactions. Figure 3.5 shows the results of the competition assay. As the amount of specific competitor increased, the bands representing protein binding

decreased in intensity. When nonspecific competitor was used, the band intensity decreased for only one band; however, the intensity decrease did not seem to be significant enough to indicate the binding was specific. These results confirm the binding ability of a protein in the Cdh extract, but do not confirm the specificity of binding.

Experiments were also conducted to determine if Cdh was able to bind to the same 200bp region in other *Methanosarcina* species (Fig. 3.5). This region was PCR amplified from the genomes of *M. thermophila* and *M. barkeri*. The DNA was then incubated with the Cdh extract to determine if binding occurred with these sequences as well. The Cdh extract was also able to bind to these DNA sequences. As was the case with the DNA from *M. acetivorans*, the binding of this Cdh extract did not appear to be specific.

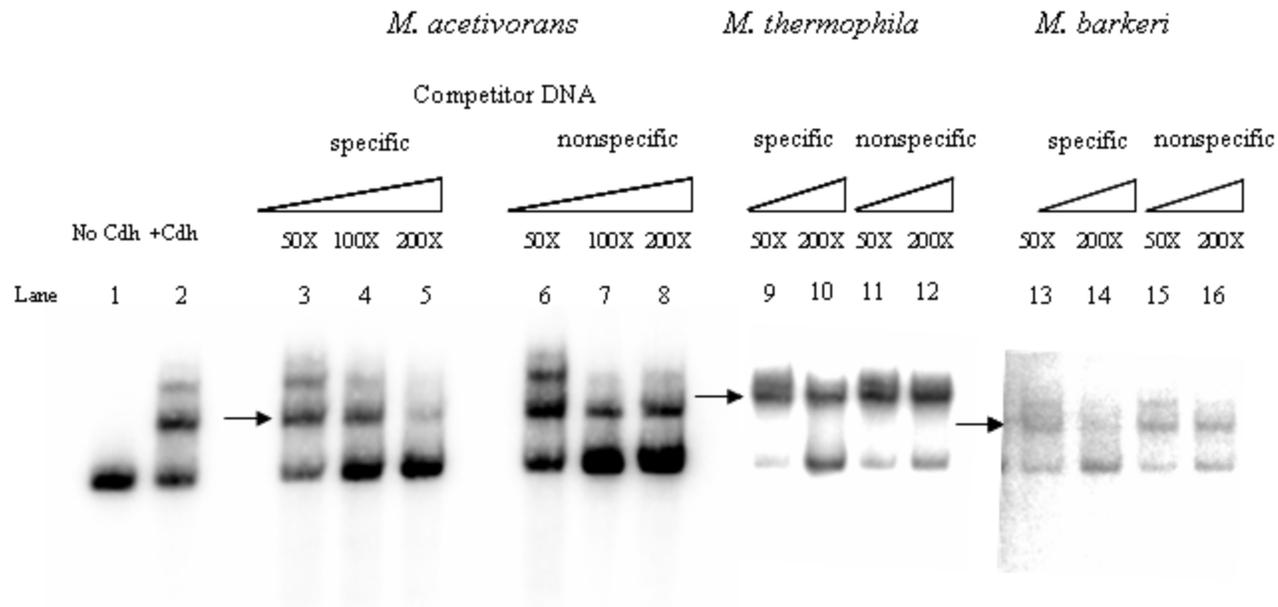


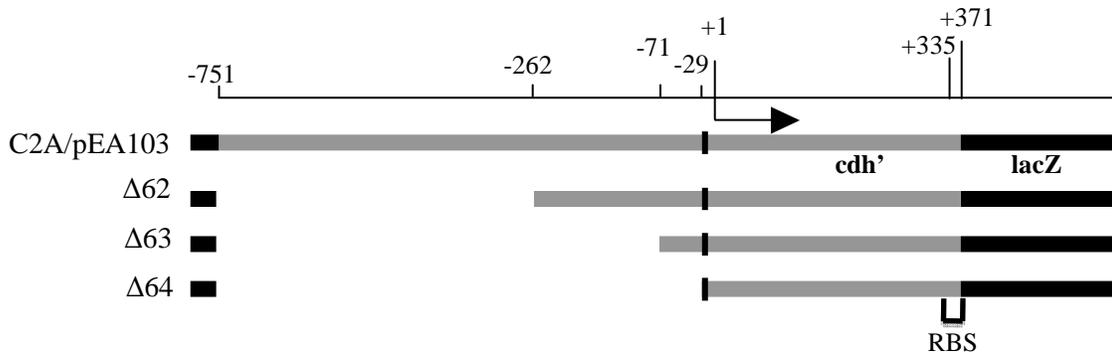
Figure 3.5. Electrophoretic mobility shift assay with 16 μ g partially purified Cdh complex binding to region upstream of *cdh* transcriptional start site. Mobility shift using 50 ng of 200 bp oligonucleotide from flanking region 266 bp upstream of *cdh* promoter. Lanes 1-8, DNA from *M. acetivorans*. Lane 1, labeled DNA probe only. Lane 2, labeled DNA probe and Cdh only. Lanes 3-5, 50X, 100X, and 200X specific competitor DNA, respectively. Lanes 6-8, 50X, 100X, and 200X nonspecific competitor DNA, respectively. Lanes 9-12, DNA from *M. thermophila*. Lanes 9-10, 50X and 200X specific competitor DNA, respectively. Lanes 11-12, 50X and 200X nonspecific competitor DNA, respectively. Lanes 13-16, DNA from *M. barkeri*. Lanes 13-14, 50X and 200X specific competitor DNA, respectively. Lanes 15-16, 50X and 200X nonspecific competitor DNA, respectively. Nonspecific competitor DNA was a 200 bp fragment from within the structural *M. thermophila rad25*.

3.4.4. Effect of sequence deletions on the regulation of *cdh* by different substrates.

In a prior report that employed a translational fusion of the *cdh* promoter to *lacZ*, CODH/ACS from *M. thermophila* TM1 was shown to be up-regulated up to 54-fold during growth on acetate, compared with growth on methanol or trimethylamine (5). To determine whether sequences located upstream of the transcriptional start site might be involved in CODH/ACS regulation, a series of deletions were generated upstream of the promoter (Fig. 3.6A). The *cdh* deletion sequences were ligated into the *lacZ* reporter plasmid as translational fusions and transformed into *M. acetivorans*. The effects of these deletions on the expression of β -galactosidase in response to different substrates are shown in Fig. 3.6B.

Deletion of DNA sequences upstream of the promoter ($\Delta 62$ or $\Delta 63$) did not have a significant effect on expression of β -galactosidase compared with wild type sequence in cells grown with acetate, trimethylamine or methanol. These data indicate that this region is not involved in regulation of expression based on substrate.

A.



B.

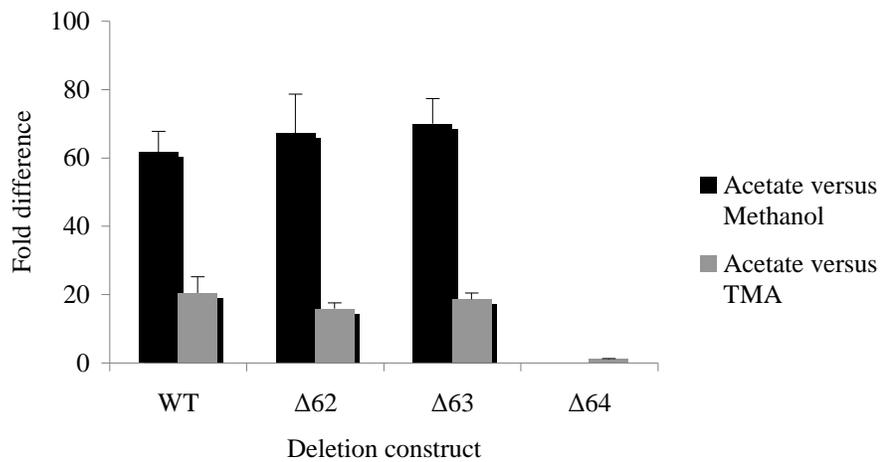


Figure 3.6. Map of deletions upstream of the *Methanosarcina thermophila* TM1 *cdhA* transcriptional start site and the effects of those deletions on regulation by different substrates. Panel A shows the location of the upstream deletions. Position numbers above relative to start of transcription indicated by arrow. RBS is the location of the predicted ribosomal binding site. Panel B shows the fold difference in LacZ expression for either acetate vs. methanol or acetate vs. TMA grown cells.

3.5. Discussion

When grown on acetate, members of the genus *Methanosarcina* generate CODH/ACS, a key enzyme complex that catalyzes the dismutation of acetate and acetylation of coenzyme CoA (43, 61). This gene was previously shown to be regulated at the level of transcription (114). Since the majority of known mechanisms of transcriptional regulation in Archaea involve *trans* factors such as activators or repressors, *cdhABCDE* was studied to determine if *trans* factors might be involved in regulation either upstream or downstream of the transcriptional start site.

Downstream of the transcriptional start site of *cdhABCDE* there is a 371 bp 5' leader region. Long 5' UTRs are involved in transcriptional regulation in bacteria via multiple mechanisms. Usually, these mechanisms involve the formation of secondary structures within the RNA (128, 131). Some of these mechanisms also involve *trans* factors that bind to cause conformational shifts in the RNA. Using affinity chromatography, it was concluded that, although there was differential binding between substrates, the binding was the same as the negative control, indicating that the binding was nonspecific. Therefore, the results indicate that there are no apparent *trans* factors binding to this region downstream of the transcriptional start site under the conditions that were tested. Further experiments would need to be conducted to determine if there are any substrate specific DNA-binding proteins that can bind to this region under different conditions.

The region within 232 bp upstream of the *cdhABCDE* transcriptional start site was also studied to determine if this region contained the binding site for DNA-binding proteins. Gene regulation involving both repressors and activators binding

upstream of transcriptional start sites has been reported in the Archaea (13, 38). Within the methanogenic Archaea nitrogen regulation is controlled by a helix-turn-helix DNA-binding repressor protein, NrpR (68, 69). The binding of this protein to DNA directly interferes with the binding of RNA polymerase, repressing expression of genes involved in nitrogen fixation. Since the 100 bp sequences used in this study were located 256 bp and 176 bp upstream of the transcriptional start site, it is unlikely that regulation of CODH/ACS expression occurred by occlusion of RNA polymerase binding to the promoter, which has been shown to bind to DNA around 29 bp upstream of the start of transcription in DNA footprinting studies in *Methanococcus vannielii* (122). Although homologs to a number of repressor type proteins encoding helix-turn-helix motifs have been identified in the annotated sequences of methanosarcinal genomes, palindromic sequence typically associated with their binding to DNA were not detected upstream of the CODH/ACS promoter. However, analysis of this region revealed sequences that are similar to known eukaryotic activator protein binding sites, including N-Oct-3, which regulates transcription of neuronal cells, NIT2, a global-acting positive regulator of nitrogen-utilizing genes, and GCN4, a yeast transcriptional activator protein (101). These proteins, however, have no known homologs in *Methanosarcina* spp.

Several proteins were shown to bind differentially to this region in response to acetoclastic (CdhA, CdhB, and CdhD) or methylotrophic (MttB) growth. However, further analysis of a partially purified extract of Cdh complex did not prove this binding was specific. The binding of MttB is intriguing as this suggests that binding

of this protein may have a role in suppressing *cdh* expression during methylotrophic growth, but studies with purified MttB are required to confirm a role for this protein.

The results of this study indicate that the region between -256 and -78 is most likely not important in the regulation of *cdhABCDE*. The results of these experiments did not detect any regulatory proteins, either putative or confirmed, that were binding to this region upstream of the *cdhA* transcriptional start site. Analysis of the sequence upstream of the transcriptional start site did not detect any motifs that are usually associated with DNA-binding proteins, such as direct repeats. Using a *lacZ* translational fusion to determine the effects of different substrates on expression, it was shown that the region upstream of the start of transcription was not involved in regulation. There also does not appear to be specific sequences for DNA-binding proteins within the 5' leader region located downstream of the transcriptional start site.

The majority of regulatory mechanisms found in Archaea involve DNA-binding proteins. The results presented in this chapter indicate that *trans* acting elements are not involved in the regulation of expression of CODH under the conditions described in this chapter. Deletions made within the region located upstream of the transcriptional start site indicate that this region is not involved in regulation. However, further experiments are needed to determine if there are any *trans* acting elements capable of regulation the expression of CODH under condition other than those used in this chapter. These experiments should focus on the 5' leader region and should encompass putative *trans* elements that may act on the DNA or the RNA sequence.

Chapter 4: A 5' leader region regulates expression of Methanosarcinal CO dehydrogenase/acetyl-coenzyme A synthase

4.1. Abstract

Expression of the archaeal CO dehydrogenase/acetyl-coenzyme A synthase (CODH/ACS) operon (*cdhABCDE*), which catalyzes the cleavage of the acetyl C–C bond of acetyl CoA in the acetoclastic pathway, is highly regulated by *Methanosarcina* spp. in response to growth on different substrates. However, the mechanism of CODH/ACS regulation is unknown. A highly conserved 371 base pair untranslated leader region (UTR) was identified as a *cis* regulatory element of this gene *in vivo* by deletion analysis using translational fusions. Deletions within the UTR significantly reduced the differential expression of the operon between acetoclastic and methylotrophic growth observed in wild type cells. One apparent regulatory region located within the leader region contains several putative mRNA secondary stem-loop structures, one of which appears to have a poly-U region typical of mechanisms associated with differential transcription termination. Deletions of these putative structures also had an effect on expression under methylotrophic substrates. Comparison of LacZ amounts and mRNA amounts, as calculated by quantitative reverse transcriptase PCR (qRT-PCR) showed that protein and mRNA levels were not significantly different, indicating that translational regulation was most likely not a factor. QRT-PCR along different lengths of the UTR showed a 15 fold difference in transcript levels 358 bp downstream of the 5' end, indicating that

approximately 22% of *cdh* regulation was the result of differential transcription initiation; the remaining difference in *cdh* expression observed was the result of differential transcript levels near the 3' end of the UTR. Transcript stability studies confirmed that differential transcript levels observed near the 3' end resulted from elongation termination rather than differential message degradation. In contrast to other archaeal genes that are regulated by a repressor or activator binding near the promoter, the CODH/ACS operon appears to be attenuated by a novel mechanism located within the conserved leader region that has not been described previously in the Archaea. The results indicate that methanosarcinal CODH/ACS expression is controlled by multiple mechanisms as part of the regulatory strategy employed by these methanogenic Archaea to efficiently direct carbon and electron flow in anaerobic consortia during fermentative processes.

4.2. Introduction

Methanogenic species in the genus *Methanosarcina* are the most metabolically diverse among methanogens, with the ability to grow by CO₂ reduction with H₂, methyl reduction with H₂, acetoclastic fermentation of acetate, or methylotrophic catabolism of methanol, methylated amines, and dimethylsulfide (125). *Methanosarcina acetivorans* is also reported to grow with CO (93). These organisms preferentially use substrates with higher free energy, such as methanol and trimethylamine, before utilizing acetate, exhibiting diauxic growth in the presence of both types of substrates (17, 34, 60, 74, 107). Despite their preference for non-acetoclastic substrates and the relatively lower growth rates and yields of

Methanosarcina spp. with acetate, approximately 70% of the biogenic methane produced globally is generated from the catabolism of acetate (71). Therefore, regulation of genes involved in the aceticlastic pathway is likely to have a significant impact on anaerobic degradation processes and the global biogenesis of methane.

Several enzymes associated with aceticlastic methanogenesis, including acetate kinase, phosphotransacetylase, carbonic anhydrase, and carbon monoxide dehydrogenase/acetyl CoA synthase (CODH/ACS), are regulated in response to substrates (52, 54, 65). Expression of the CODH/ACS operon (*cdhABCDE*), which catalyzes the catabolism of acetyl CoA in the aceticlastic pathway, is highly regulated in response to growth on acetate compared with methanol and trimethylamine (TMA), but the mechanism(s) of this regulation is not known (114, 121). Two CODH/ACS operons are detected in the annotated genome sequences of *M. acetivorans* (37), *Methanosarcina barkeri* (73) and *Methanosarcina mazei* (30). In contrast, only one *cdhABCDE* operon has been detected in *M. thermophila* (42). In *M. acetivorans*, both of these operons appear to be regulated in response to substrate, although the extent of the regulation is greater in one (MA3680) than the other (MA1016) (65). Northern analysis of CODH/ACS mRNA from *Methanosarcina thermophila* indicates that the regulation of this CODH/ACS ortholog in response to substrate occurs, at least in part, at the level of transcription (114). The regulation of this ortholog was corroborated with translational fusions of the *M. thermophila* CODH/ACS promoter to *lacZ* (5) and by peptide fragment analysis of a CODH/ACS component, CdhA, in *M. mazei* (32). This CODH/ACS operon has an unusually long 371 bp 5' leader region (114). In contrast, the 5' leader region of CODH/ACS in the

obligately acetoclastic *Methanosaeta concilii* (formerly “*Methanotherix soehgenii*”) is less than 100 bases long and has no significant sequence similarity to the methanosarcinal 5’ leader regions (31). These observations suggest that CODH/ACS 5’ leader region has a regulatory role, possibly at the transcriptional and/or post-transcriptional level.

The purpose of this study was to localize regulatory regions that mediate expression of the catabolic CODH/ACS complex in the acetoclastic *Methanosarcina* to determine whether the 5’ leader region is involved in expression. The promoter and leader sequence from *M. thermophila* TM1 CODH/ACS operon was fused to *lacZ* as a translational reporter to study differential gene expression on different substrates. Deletion analyses of sequences adjacent to the promoter combined with quantitative reverse transcriptase PCR (qRT-PCR) analyses of the transcript indicate that CODH/ACS expression is regulated in part by a sequence located downstream of the CODH/ACS promoter in the 5’ leader in response to acetate and methylotrophic substrates.

4.3. Materials and Methods

4.3.1. Archaeal and Bacterial strains.

Methanosarcina acetivorans C2A (=DSM 2834) and *Methanosarcina thermophila* TM-1 (=DSM 1825) were obtained from sources described previously

and maintained as frozen stocks (115). *Escherichia coli* SURE was obtained from Stratagene. *Escherichia coli* strain DH5 α / λ pir was obtained from W. Metcalf (77).

4.3.2. Media and cell growth.

Artificial marine mineral medium was prepared for growth of *M. acetivorans* by methods described previously (113). Growth substrates used were sodium acetate, methanol or trimethylamine-HCl (TMA) at a final concentration of 0.05 M.

Methanogen medium was prepared anaerobically in a N₂-CO₂ (4:1) atmosphere by a modified Hungate technique (110). Plating on solidified medium was performed in an anaerobic glove box as described previously by Apolinario et al (4). LB liquid or agar-solidified plates were used for *E. coli* (96). Puromycin (2.0 μ g/mL) and ampicillin (100 μ g/ml) were added to media for selection and maintenance of *M. acetivorans* and *E. coli* transformants, respectively. *M. acetivorans* strains were incubated at 35°C; *E. coli* strains were incubated at 37°C.

4.3.3. Sequence analysis.

Sequences within the *cdh* mRNA leader sequence from the transcriptional start site to the beginning of the *cdhA* structural gene (ATG) from four *Methanosarcina* species were aligned using DNAMAN PC software [DNAMAN version 3.0, Lynnon BioSoft]. *M. thermophila* sequence was obtained from Genbank Accession No. L20952 from 961 to 1390, *M. acetivorans* sequence was obtained from Genbank Accession No. AE010299 from region 4741963 to 4742326, *M.*

barkeri sequence was obtained from Genbank Accession No. NC007355 from region 240857 to 241493, and *M. mazei* sequence was obtained from Genbank Accession No. NC007355 from 819301 to 819937.

4.3.4. Construction of *cdh'*::*lacZ* deletion plasmids.

Primers were designed to delete sequences either downstream or within the promoter upstream of *cdhA* (Table S4.1), which is the first gene in the *cdhABCDE* operon encoding catabolic CODH/ACS. The transcriptional start site was previously determined (76, 114). The reporter plasmid pEA64, which contained the *M. thermophila cdh* promoter from pCDH1.3 (114), was used as a PCR template for deletions in the *lacZ* reporter described in Table 4.1.

Table 4.1. *E. coli* and *M. acetivorans* plasmid constructs.

Plasmids	Parent	Characteristics or description of construct	Source
pCR2.1	-----	Cloning vector, Amp ^R , Pur ^R , <i>lacZ</i> α	Invitrogen
pWM315	-----	Ori R6K, Ori pC2A, Amp ^R , Pur ^R , <i>lacZ</i> α	(77)
pCDH1.3	-----	<i>cdhA'</i> , Amp ^R	(114)
pcdhp:: <i>lacZ</i>	-----	<i>cdhA'</i> :: <i>lacZ</i> , Amp ^R , Ori pBR322	(5)
pEA61	pWM315	<i>Cdh'</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pEA64	pEA61	<i>Cdh'</i> ::CAT, Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pEA103	pWM307	<i>Cdh'</i> :: <i>lacZ</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	(5)
pEA110	pEA103	<i>Cdh'</i> Δ 64:: <i>lacZ</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pEA111	pEA103	<i>Cdh'</i> Δ 65:: <i>lacZ</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pEA112	pEA103	<i>Cdh'</i> Δ 66:: <i>lacZ</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pEA113	pEA103	<i>Cdh'</i> Δ 67:: <i>lacZ</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pKA6	pEA103	TBP-1' <i>::lacZ</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pKA7	pEA103	TBP-2' <i>::lacZ</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study
pKA8	pEA103	TBP-3' <i>::lacZ</i> , Ori pC2A, Amp ^R , Pur ^R , Ori R6K	This study

Deletion 64, located upstream of the transcriptional start site, was generated from the pEA64 template using reverse primer 60 (Table S4.1), which hybridized within the plasmid with forward primers that generated *ClaI* restriction sites at different locations upstream of the *cdh* promoter. Forward primer 64 generated *ClaI* restriction site 29 bp upstream of the transcriptional start site to create $\Delta 64$, which deleted part of the *cdh* promoter TATA box. The PCR generated DNA fragment was cloned directly into pCR2.1 (Invitrogen), then digested with *NsiI/SalI* and cloned into the *NsiI/SalI* restriction sites of plasmid pcdhp::lacZ. The latter plasmid was digested with *XhoI/BamHI* and the fragment containing the *cdh-lacZ* fusion was ligated into the lacZ reporter shuttle vector pEA103 replacing wildtype *cdh* sequence.

Deletions downstream of the transcriptional start site were created by PCR using pCDH1.3 (Table 4.1). The same forward primer 53 was used with reverse primer 65 to generate deletion $\Delta 65$ with a *SalI* restriction site 211 bp downstream of the CODH/ACS transcriptional start site. Likewise, reverse primers 66 and 67 were used to generate a *SalI* restriction site 133 bp ($\Delta 66$) and 28 bp ($\Delta 67$), respectively, downstream of the transcriptional start site of *cdhA*. PCR fragments were ligated directly into pCR2.1 (Invitrogen), then digested with *PstI* and *SalI* and ligated into pcdhp::lacZ. Finally, the latter plasmid was digested again with *XhoI/BamHI* and ligated into pEA103 replacing wildtype *cdh* sequence. After transformation, each deletion plasmid was extracted from *M. acetivorans* and confirmed for the correct sequence.

4.3.5. Identification and deletion of putative mRNA secondary structures.

A PCR-based overlap procedure was used to create deletions in DNA encoding putative mRNA loop structures in the region downstream of the transcriptional start site (102). Primers B1 and C1 were used to create the loop 1 deletion 28 to 50 bp downstream of the transcriptional start site (Table S4.1). Primers B2 and C2 were used to create the loop 2 deletion 56 to 69 downstream of the transcriptional start site. Primers B3 and C3 were used to create the loop 3 deletion 84 to 123 bp downstream of the transcriptional start site. Primers B4 and C4 were used to create the polyU deletion 71 to 77 bp downstream of the transcriptional start site. The 1st PCR product (A1 & B primers) and 2nd PCR product (C & D1 primers) were amplified from pEA103 template in a 50 μ L PCR mixture using the following cycle parameters: 25 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, and elongation at 72°C for 1 min. The final PCR product that included the subdeletion was generated using A1 and D1 primers with the first two PCR products as DNA templates in a 50 μ L PCR mixture using the following cycle parameters: 25 cycles of denaturation at 95°C for 1 min, annealing at 60°C for 2 min, and elongation at 72°C for 3 min. Correct size and yield of PCR products were confirmed by electrophoresis in 0.8% (v/v) agarose (Fisher Biotech) in TAE (Tris base, 40 mM; acetic acid, 20 mM; EDTA, 1 mM; pH 8.5).

PCR fragments were digested with *Xho*I and *Bam*HI restriction endonucleases and ligated into linear pEA103 digested with the same restriction enzymes to remove the wild-type *cdh* promoter fragment. After transformation, each deletion plasmid was extracted from *M. acetivorans* and confirmed for the correct sequence.

4.3.6. Reporter Gene Expression Analysis.

M. acetivorans transformed with either the wild-type reporter plasmid pEA103, deletion plasmids or the control plasmid pWM315 (no *cdh* promoter) were inoculated (10% v/v) in triplicate into liquid medium containing 0.1 M sodium acetate, TMA, or methanol and incubated to mid-exponential growth phase. Cultures were sampled (1.0 ml) to perform β -galactosidase assays as described previously (5). 1 ml samples were centrifuged at 13,000 rpm for 5 minutes. Cell pellets were resuspended in 1 ml Z Buffer (60 mM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 40 mM $\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$, 10 mM KCl, 1 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, pH 7.0). Cells were sonicated and centrifuged at 13,000 rpm for 15 minutes. The supernatant was transferred to a new tube. The protein concentration of the supernatant was determined using the Bradford assay. Triplicate samples of the supernatant (100 μl) were combined with 700 μl Z Buffer + β -mercaptoethanol (27 μl / 10 ml Z Buffer). The mixture was incubated at 30 °C. 160 μl ONPG (4 mg/ml) was added to each mixture, and the time for the color to change to yellow was measured. The reaction was stopped with the addition of 400 μl 1 M Na_2CO_3 , and the reactions were assayed for absorbance at 420 and 550 nm.

4.3.7. *In vitro* transcription assays.

In vitro transcription assays were performed as previously described with some modifications. The templates for the reaction were PCR products amplified from pEA103. The first product was generated with primers Sowers1 and Sowers2 (Table S4.1). This product began 210bp upstream of the transcriptional start site of *cdhA* and terminated 200 bp into the *lacZ* structural gene. The second template was

made using primers Sowers3 and Sowers2. This product contained the *Methanothermobacter thermoautotrophicum hmtB* promoter for initiating transcription of the 5' leader region of *cdhA*. The product also terminated 200 bp into *lacZ*. The *in vitro* transcription assay was performed as described previously (97), with the modification that the limiting nucleotide was CTP, which was added with the other three nucleotides as the chase after elongation had stalled. *In vitro* transcription assays were done at varying temperatures to account for variability.

4.3.8. RNA extraction and quantitative reverse transcriptase PCR.

RNA was extracted from *M. acetivorans* exponential phase cells grown on either acetate or methanol with the RNeasy kit (Qiagen) following the manufacturer's instructions. RNA (50 ng) was used as template for each qRT-PCR using the iScript one-step RT-PCR kit (Bio-Rad) following the manufacturer's instructions. The primers used in the qRT-PCR reactions are listed in Table S4.1. To compare the fold differences in expression, the C(t) values for each reaction were normalized to the C(t) value for the gene product of MA4504 (124). The fold differences were calculated using the formula (100):

$$\text{Fold difference} = 2^{(\Delta\Delta C(t))}$$

where $\Delta\Delta C(t)$ is the difference in the normalized C(t) values of acetate- versus methanol-grown cultures.

4.3.9. Transcript stability studies.

Transcript stability was determined by quantitating *cdh* mRNA in exponentially growing cultures of *M. acetivorans* grown with either acetate or methanol after inhibition of transcription with Actinomycin D. Samples were taken 0, 2, 5, 10, 15, and 30 minutes after addition of the 100 ug/ml Actinomycin D, and RNA was extracted from these samples as described above. The RNA was quantified and used as a template for qRT-PCR using primers 221 and 315 (Table S4.1). These primers were specific for the entire 5' leader region as well as 40 bp of the structural gene, generating a product that was 405 bp in length.

4.3.10. Construction of TBP::*lacZ* reporter plasmids

Plasmid reporter constructs for genes encoding methanosarcinal TATA binding proteins (TBPs) are shown in Table 4.1. *M. acetivorans* C2A genomic DNA was used as a template for PCR to generate the promoter sequences for ligation with *lacZ* using primers specific for the three genes encoding TATA binding proteins (TBPs) in *M. acetivorans* (Table S4.1). For TBP-1 (MA4331), forward primer #161 and reverse primer #162 were used to PCR amplify the region 599bp upstream of the start of translation. These primers introduced an *Xho*I site 595bp upstream and a *Bam*HI site 37bp downstream of the translational start site. TBP-2 (MA0179) was made using forward primer #167 and reverse primer #168, amplifying the region 550bp upstream of the start of translation with an *Xho*I site 543bp upstream and a *Bam*HI site 2bp downstream of the translational start site. TBP-3 (MA0278) was made using forward primer #165 and reverse primer #166, creating a product that

started 590bp upstream of the translational start site with an *XhoI* site 582bp upstream and a *BamHI* site 3bp downstream of the translational start site. Once the PCR-generated promoter fragments were made, they were digested with *XhoI/BamHI* in order to make the fusion.

To construct the fusion plasmid, pEA103 was used as the backbone. This plasmid was digested with *XhoI/BamHI* to remove the *cdh* promoter region, and the remaining backbone of the plasmid was ligated with the PCR-generated TBP promoter fragments to form the reporter plasmids pKA6, pKA7, and pKA8 containing the promoters of *tbp-1*, *tbp-2*, and *tbp-3*, respectively.

4.4. Results

4.4.1. Identification of conserved sequence in carbon monoxide dehydrogenase/acetyl CoA synthase (CODH/ACS) orthologs.

Members of the genus *Methanosarcina* contain multiple copies of several catabolic genes including two copies of the *cdh*ABCDE operon that encodes carbon monoxide dehydrogenase/acetyl CoA synthase (CODH/ACS) (30, 37, 73). However, only one *cdh*ABCDE operon has been detected in *M. thermophila* (42). The current study focused on the CODH/ACS ortholog in *M. thermophila* (Genbank Accession No. L20952, = *M. acetivorans* MA3860, = *M. barkeri* MbarA0204, = *M. mazei* MM0684), which had been reported previously to be highly regulated in response to substrate (5, 32, 114). To identify conserved regions of DNA that might be critical for *cdh* regulation, sequences upstream of the translation start site from CODH/ACS orthologs in *M. thermophila*, *M. acetivorans*, *M. mazei*, and *M. barkeri* were aligned (Fig. 4.1). The alignments revealed that the orthologous *cdh* 5' UTR regions between *Methanosarcina*. spp. were highly conserved ranging from 73% to 87% sequence similarity, which is similar to the sequence conservation of 77% to 86% similarity observed between methanosarcinal *cdhA* encoding orthologs. In contrast, only 45% and 62% similarity is observed within the region 256 bp upstream of the transcription start site. These regions were evaluated further by deletion analysis to determine whether the highly conserved UTR has a role in CODH/ACS regulation.

<i>M. thermophila</i>	TATGAAGAATCTTTAAATGAGTTGTGGTTAGTATGAAAAATGAAAAATAATAACAA	-208
<i>M. acetivorans</i>	--g--a-----a--t---t-t---tt---a-----a---t--ta-----	-208
<i>M. mazei</i>	--a--a-----gggct-t---cac-----g-----att-g-----	-207
<i>M. barkeri</i>	--ac--at--t-ac-t---a-t---t-g-ac-c--gt---aagcgt--t-gt---t--	-205
<i>M. thermophila</i>	ATATTAAATCAATGGAAAAACTATTGGGAATCAAGAATACTCATTCAAAAAACCAGT	-148
<i>M. acetivorans</i>	-----g-c-a-----a--a-a-gc-g-----ggtca-a-----ggca	-148
<i>M. mazei</i>	-----ata-----a-----a--tc---t-t-----tcatt-----g-ta	-147
<i>M. barkeri</i>	t--a--gc-ca--g.-.-gt---aca-ca-tg-t-ta-tggt-tt-a-tg-----cac	-147
<i>M. thermophila</i>	ATCTACAGGCATTGAGGGCTTTTCTTTTCGGAGTAGATCTTCAAGA.TTTTTTATGAAGA	-89
<i>M. acetivorans</i>	-ca--a-.-c-at--acta---t-ca--a-a-t-----g---t-----c---t-	-89
<i>M. mazei</i>	tg--ttgacatagagtat----t-gatca-aa--tct--g-g-tt-----a-ga-ga-	-87
<i>M. barkeri</i>	-----c-a-tct----t-g-accta--at-acaat-a-c--g-aag-c--ct	-87
<i>M. thermophila</i>	AGGTATCACTAAATTCATTTTACCCATTGCGAAATAGGGTGTACAAATTAGAAAATGT	-29
<i>M. acetivorans</i>	gt--gc-t---c-----c-----ca-----	-29
<i>M. mazei</i>	c--gtctc-ac-t-c-at-----tg--cg-. .---a-----	-29
<i>M. barkeri</i>	gat-t-g-ac-ttcc-ac--ac--atc-cga--tgc. .-----a-	-29
<i>M. thermophila</i>	TATATAGATTTGGTACATTTAGACTTTTAATTAGTGTATTATATCCATCGGTAGCGACCT	+32
<i>M. acetivorans</i>	-----	+32
<i>M. mazei</i>	-----t-ac-----	+32
<i>M. barkeri</i>	-----g-a-----a-----	+32
<i>M. thermophila</i>	CTGCTCAAGATTAAGGTCCCAATCGTTGGTAAAAACGGTTTTTTTTGAGACATAGGGCGTA	+92
<i>M. acetivorans</i>	-----t-----c-----a--t-----	+92
<i>M. mazei</i>	-----t-----c-----a-a-----	+92
<i>M. barkeri</i>	-----t-g-----t-----c-----a-----	+92
<i>M. thermophila</i>	AAAAGGGATTTGGAAGAGAAATAAGGCCCTAAAAACGACTTTTACAAG.AATATATTCTC	+151
<i>M. acetivorans</i>	---a-a---a-----t--t-a-----gg---g-a-----c-	+151
<i>M. mazei</i>	---a-----t-----c-t-g-t-----t-----a-a-----c-	+152
<i>M. barkeri</i>	---t--c---.---t-tg-at-a-----t-t-.g-a----tc-	+151
<i>M. thermophila</i>	TATAAAAACGTTCTTTATTTGAAAAACGTTGGTATAACGTTTAATGTCCAAATTGAGAG	+211
<i>M. acetivorans</i>	-----gct-----t-----g-----g-g--t-----c-----	+210
<i>M. mazei</i>	--c-----c-t-----c-----a--c--c--a-t-----a-----	+211
<i>M. barkeri</i>	-c-----act---gc--g-----cgc-g-ta-a-tat--gttat-tg---aaa--t-	+211
<i>M. thermophila</i>	CTATGAACGTTTTACCATAAAAACGTCAT...TTAAACGGTTGTAACCTTTGAAGCAGACGG	+268
<i>M. acetivorans</i>	---aa-----g--g---t-c---. .-----a-g--g-----	+267
<i>M. mazei</i>	---a-gt---g---g-----..-----a--a-----	+268
<i>M. barkeri</i>	g-gaat-gtg---ta-cag-tg--at-tac---c-----a.------	+270
<i>M. thermophila</i>	TAAGATATAAATGAATAATCATTAAACAATATGCCGCTT..ATCAGGATTACAATCGAAT	+326
<i>M. acetivorans</i>	---g---t---t-----g-----a..-a....--ag--gc--cca	+321
<i>M. mazei</i>	c--a-----t---t-----g-----aag..-a-.-g-----tc-	+325
<i>M. barkeri</i>	c--gag.....-tt-gg-a---ca-----aagaat-a-aat--g-----tc-	+325
<i>M. thermophila</i>	TGGCGAAGTCGACTTTAAACAAATTAAGGAGGTAAAGCTCAC <u>ATG</u> AGCAAACCTAACTACC	+386
<i>M. acetivorans</i>	-----t-----a-----	+361
<i>M. mazei</i>	-----t-----t-----a-----	+365
<i>M. barkeri</i>	--at-----t-----t-----a-----	+365

Figure 4.1. Alignment of *cdhA* promoter region from four different *Methanosarcina* species. Numbering in right column is based on the transcription start site for *cdhABCDE* (114). The start of transcription is marked with an arrow and +1. The start of translation (ATG) is underlined.

4.4.2. Effect of sequence deletions on the regulation of *cdh* by different substrates.

In a prior report that employed a translational fusion of the *cdh* promoter to *lacZ*, CODH/ACS from *M. thermophila* TM1 was shown to be up-regulated up to 54-fold during growth on acetate, compared with growth on methanol or trimethylamine (5). To identify sequences involved in CODH/ACS regulation a series of deletions were generated downstream in the 371 bp 5' leader region between the transcriptional start site and the predicted ribosomal binding site of the *cdhA* encoding gene (Fig. 4.2). The *cdh* deletion sequences were ligated into the *lacZ* reporter plasmid as translational fusions and transformed into *M. acetivorans*. The effects of these deletions on the expression of β -galactosidase in response to different substrates are shown in Fig. 4.3.

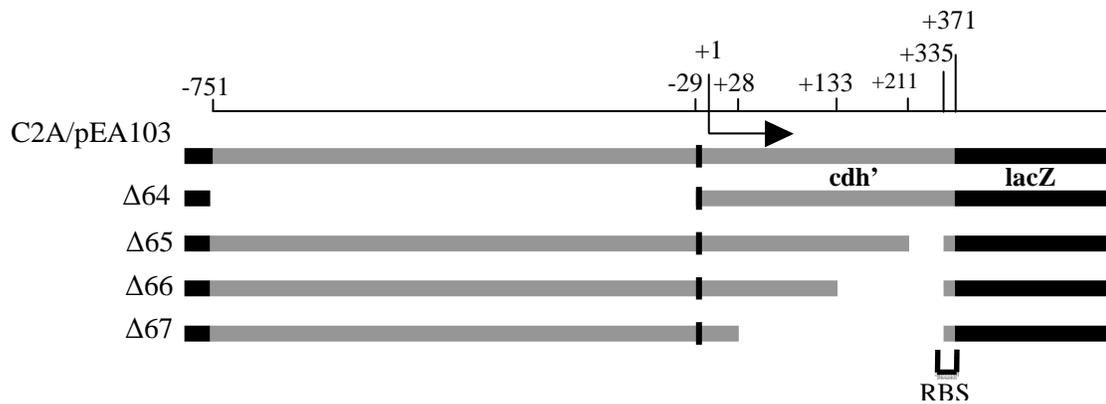


Figure 4.2. Map of deletions within the 5' leader region of *M. thermophila* TM1 *cdhA*. Position numbers above relative to start of transcription indicated by arrow. RBS is the location of the predicted ribosomal binding site.

Some deletions within the 5' UTR did have an effect on expression during methylotrophic growth. Figure 4.3 shows that $\Delta 66$ and $\Delta 67$ had only a 4- and 13-fold difference in expression, respectively, between acetate and methanol grown cells compared to the 61-fold difference seen in wild type constructs. The same was observed for cells grown on TMA, where $\Delta 66$ and $\Delta 67$ were 7- and 5-fold higher, respectively, compared to 20-fold difference for the wild type construct. These data indicate that this region is involved in regulation of expression based on substrate. Differences were also observed between expression of all constructs when grown on methanol and TMA. The methanogenic pathway for these substrates is the same, so it is not clear why expression under methanol and TMA is different. It is possible that there is some difference in the mechanism of regulation between growth on the two substrates that could account for this difference.

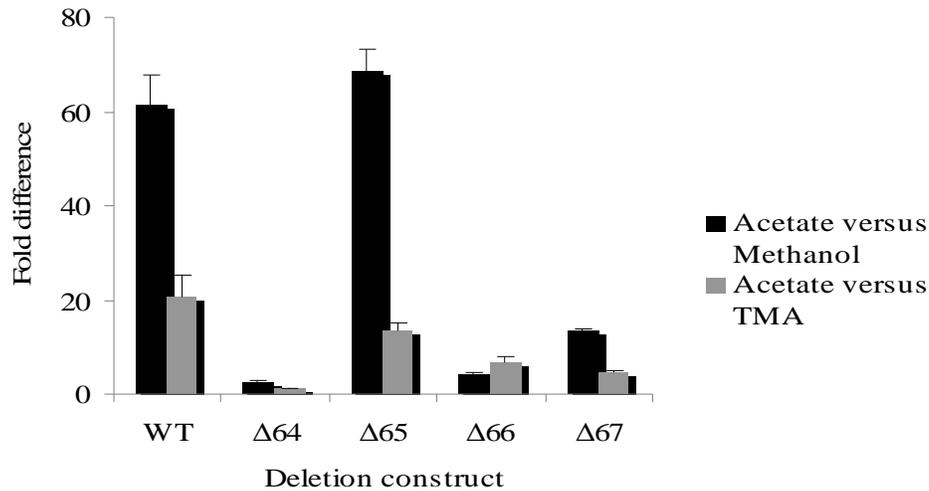


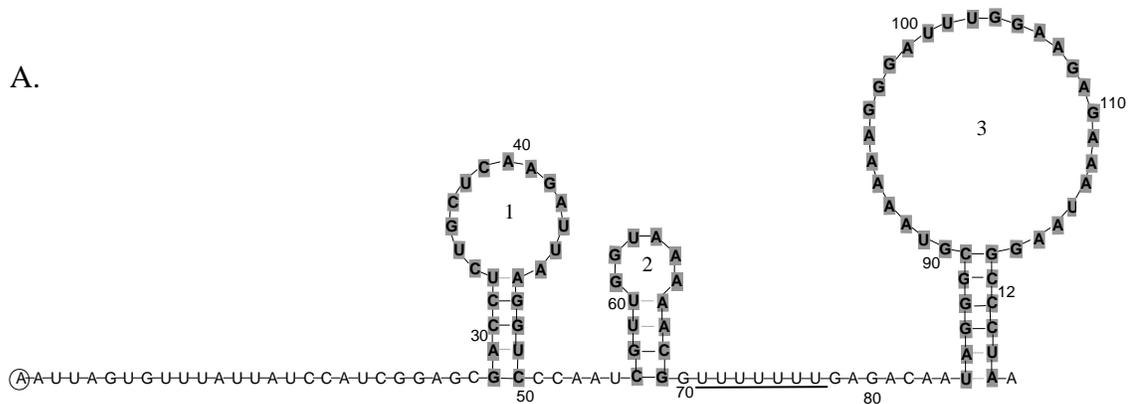
Figure 4.3. Effects of deletions downstream of the transcriptional start site on regulation by different substrates. The graph shows the fold difference in expression for either acetate vs. methanol or acetate vs. TMA grown cells.

4.4.3. Effect of deletions in sequence within the 5' leader region of *cdhA*.

Based on the results mentioned above, the 5' leader region was investigated further. Putative secondary structures within the 5' leader region mRNA were found in all four *Methanosarcina* species that consist of three stem-loop structures (Figure 4.4A). The ΔG values for these structures were calculated to be -4.10 kcal/mol, -1.5 kcal/mol, and -6.90 kcal/mol for loops 1, 2, and 3, respectively. In all four species a poly-U sequence was located directly downstream of the second putative loop structure. To investigate whether these potential secondary structures have any effect on the regulation of *cdh*, promoter sequence with deletions of putative secondary structures was fused to *lacZ* and assayed for β -galactosidase expression.

The deletion of the putative structures, as well as the poly-U sequence, had a significant effect on the expression of β -galactosidase (Figure 4.4B). During methylotrophic growth on trimethylamine, the deletion of each of the putative secondary structures significantly increased expression of β -galactosidase from 3- to 8-fold. Deletion strains grown with methanol showed slightly greater β -galactosidase expression, ranging from 4- to 11-fold. In contrast deletion strains grown acetically showed different effects on expression. When putative loops 1 (Δ loop1) and 2 (Δ loop2) were deleted, the expression of β -galactosidase decreased by half, while the deletions of the poly-U sequence (Δ polyU) and loop 3 (Δ loop3) had no effect on expression during aceticlastic growth.

A.



B.

Growth substrate	β -gal specific activity unit $\mu\text{g/mL}$ protein				
	wildtype	Δ Loop1	Δ Loop2	Δ Poly U	Δ Loop3
Acetate	24.18 \pm 2.83	13.86 \pm 0.18 (0.6 \downarrow)	10.62 \pm 1.64 (0.4 \downarrow)	26.35 \pm 2.17 (1.1 \uparrow)	25.02 \pm 1.86 (1.0)
Methanol	0.45 \pm 0.03	2.56 \pm 0.00 (5.7 \uparrow)	1.59 \pm 0.08 (3.5 \uparrow)	4.48 \pm 0.00 (10 \uparrow)	2.97 \pm 0.12 (6.6 \uparrow)
Trimethylamine	0.51 \pm 0.08	3.93 \pm 0.03 (7.7 \uparrow)	1.70 \pm 0.04 (3.3 \uparrow)	2.08 \pm 0.02 (4.1 \uparrow)	1.30 \pm 0.04 (2.5 \uparrow)

Figure 4.4. Effects of deletions in putative mRNA hairpin loops in UTR of *M. thermophila* TM1 *cdhA*. A) Predicted stem loop structures located within the 5' untranslated leader sequence based on algorithms described in material and methods. B) Effects of deletions of putative stem loop structures and poly-U region. Numbers are reported as β -galactosidase specific activity in $\mu\text{g/mL}$ of protein. Numbers and arrows above each bar represent the fold increase (\uparrow) or decrease (\downarrow) in activity between the wild-type strain and the mutant strains. Values are means and standard deviations for 3 replicate cultures.

4.4.4. Comparison of Cdh::LacZ fusion protein and *cdh* transcript levels with different substrates.

RNA was extracted from *M. acetivorans* wild-type cells to determine the difference between *cdh* transcript levels in acetate- and methanol-grown cells. The fold difference of transcript level was determined by qRT-PCR and compared to the fold difference of β -galactosidase level expressed by the *cdhA'*::*lacZ* reporter in cells grown with acetate and methanol. The fold difference of transcript in cells grown with acetate and methanol determined by qRT-PCR was 68 ± 20 compared with 62 ± 6 fold difference of β -galactosidase level expressed by the *cdhA'*::*lacZ* reporter. The fold difference between transcript and protein levels was not significant, indicating translation is not a factor in the regulation of CODH/ACS expression.

4.4.5. *In vitro* transcription of 5' leader region

The region located upstream of the transcriptional start site as well as the 5' leader region was used as a template (Sowers1/2 template) for *in vitro* transcription assays to determine if there was a definitive sequence element that was stopping transcription. Another template was also used (Sowers3/2 template), this one containing the definitive promoter from *M. thermoautotrophicum htmB*, since it was known where transcription occurred from this promoter. The results of the *in vitro* transcription assay are shown in Figure 4.5. Transcription was initiated from both templates, and temperature was not a factor, as each template had the same pattern of transcription termination.

Interestingly, transcription occurred as expected from the template with the *htmB* promoter. However, transcription from the native *cdh* promoter initiated further downstream than the determined start site. Even taking into account this discrepancy, both templates appeared to have the same pattern. Regardless of the promoter recognized, the pattern of transcriptional elongation, pausing, and termination remained the same. Most of the transcripts that initiated from either template did not reach the end of the template, indicating that transcription was not efficient throughout this 5' leader region. Each band on the gel represents a length of transcript, where the transcription machinery either paused (P) or completely stopped (S) and detached from the template. The transcripts in the pausing lanes were more abundant than those in the complete stoppage lanes. The large numbers of transcripts that were stopped along the entire length of the 5' leader region indicate that transcription of this sequence is difficult with the basal components used in this assay.

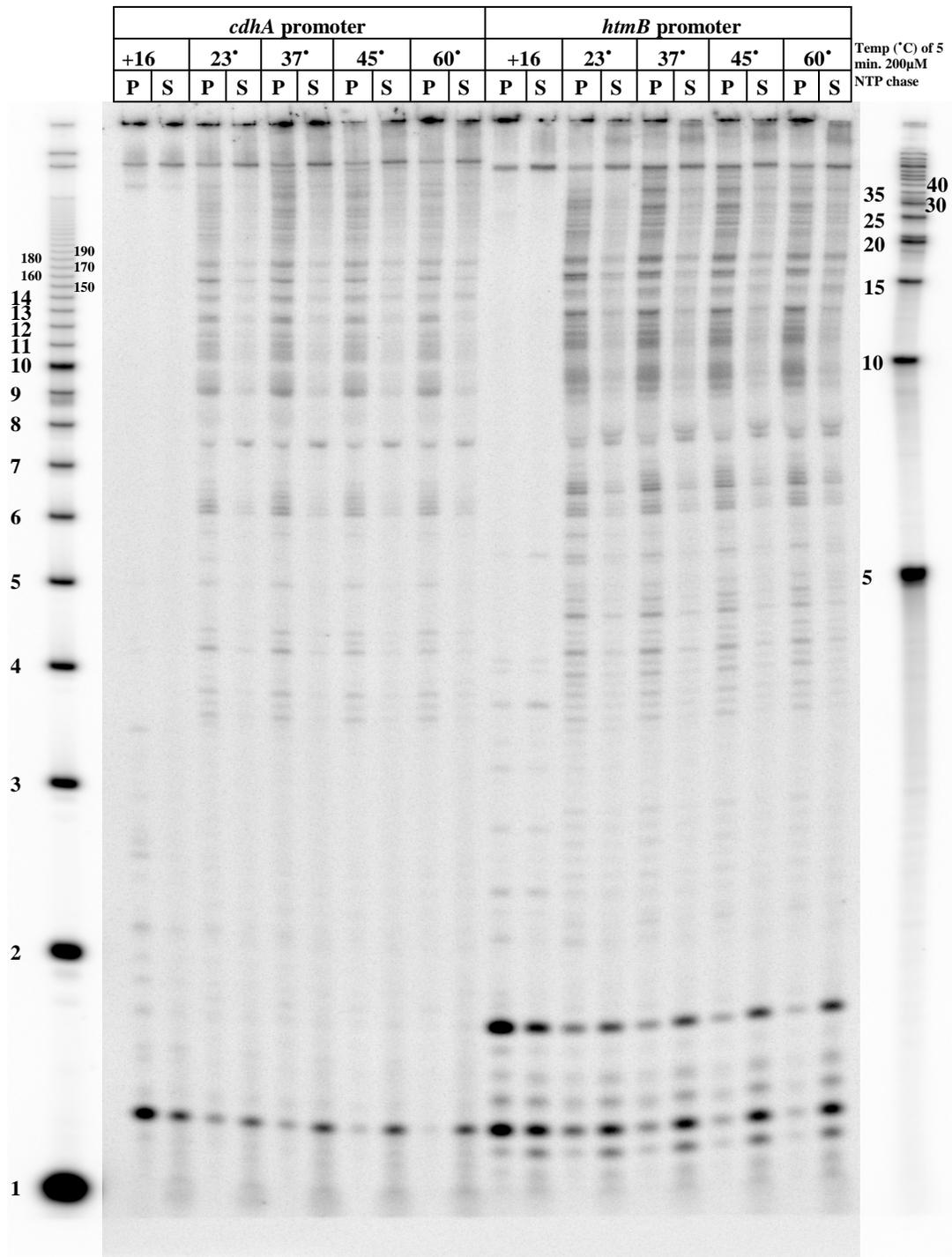


Figure 4.5. *In vitro* transcription assay of *cdhA* promoter and 5' leader region. Each band represents transcripts where elongation has paused (P) or terminated and the RNA has come loose (S). The reactions were done at multiple temperatures, shown at the tops of the lanes. The templates either contained the native *cdhA* promoter or *Methanothermobacter thermoautotrophicum htmB* promoter. Assay done by Tom Santangelo at the Ohio State University.

Interestingly, transcription occurred as expected from the template with the *htmB* promoter. However, transcription from the native *cdh* promoter appeared to be initiated further downstream than the determined start site. Even taking into account this discrepancy, both templates appeared to have the same pattern. This means that, regardless of the promoter recognized, the pattern of transcriptional elongation, pausing, and termination remained the same. Most of the transcripts that initiated from either template did not reach the end of the template, indicating that transcription was not efficient throughout this 5' leader region. Each band on the gel represents a length of transcript, where the transcription machinery either paused (P) or completely stopped (S) and detached from the template. The transcripts in the pausing lanes were more abundant than those in the complete stoppage lanes. The large numbers of transcripts that were stopped along the entire length of the 5' leader region indicate that transcription of this sequence is difficult with the basal components used in this assay.

4.4.6. Transcript stability of *cdh* during growth with different substrates.

Actinomycin D has been shown to inhibit transcription in Archaea (16, 55), and we confirmed that actinomycin D effectively inhibited transcription in *M. acetivorans* by adding actinomycin D to exponentially growing cultures containing ³H-uridine and monitoring the incorporation of the labeled compound over time. The results indicated that actinomycin D inhibited incorporation of the ³H-uridine as a result of transcription inhibition, while cells without actinomycin D continued to show ³H-uridine incorporation. Actinomycin D was added to *M. acetivorans* grown

on either acetate or methanol, and *cdh* mRNA was amplified at multiple timepoints after transcription inhibition using qRT-PCR. Figure 4.6 shows the fold difference of *cdh* transcript between acetate- and methanol- grown cells. No significant difference ($p>0.05$) was observed between the ratio of C(t) values of acetate- versus methanol-grown cells at any of the time points, indicating that the relative stability of the message isolated from acetate- and methanol-grown cells was similar.

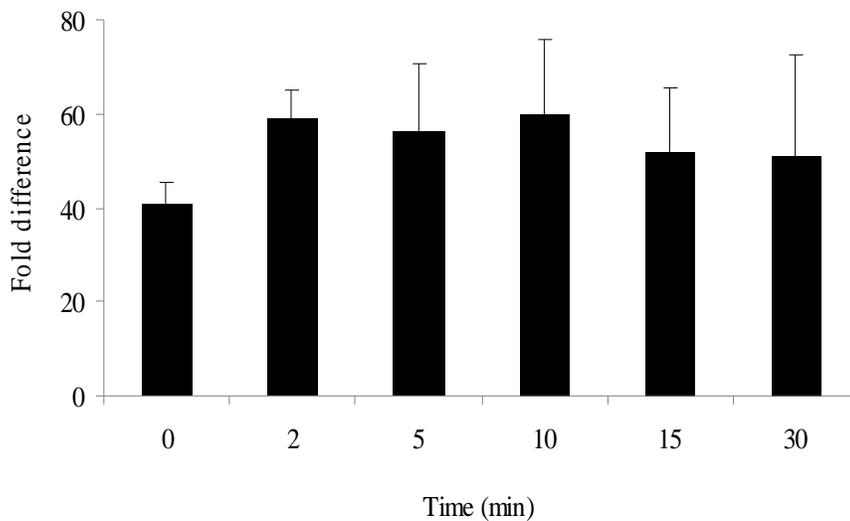


Figure 4.6. Differences in transcript levels after addition of actinomycin D. Fold differences were calculated from the C(t) values for RNA from acetate- versus methanol-grown cells after the addition of actinomycin D (100 μ g/ml) at 0, 2, 5, 10, 15, and 30 minutes. One way analysis of variance (ANOVA) showed no significant difference ($p>0.05$) between the fold differences at the different time points.

4.4.7. Effect of substrates on elongation of *cdh* transcript.

To determine whether differential elongation occurred within the 5' UTR as a possible mechanism of regulation, nine primers that hybridized along the 5' length of the transcript were used with qRTPCR to detect any changes in transcript length between acetate- and methanol-grown cells (Figure 4.7). A 15-fold difference in transcript level was observed within 358 bases of the 5' end in cells grown on acetate and methanol. However, transcript levels 405 bases downstream of the 5' end showed a significantly greater difference, increasing to 68-fold difference between acetate and methanol-grown cells. These results suggested that either early termination of elongation or differential post-transcriptional mRNA processing occurs between 358 and 405 bases downstream of the 5' end of the *cdh* transcript during methylotrophic growth.

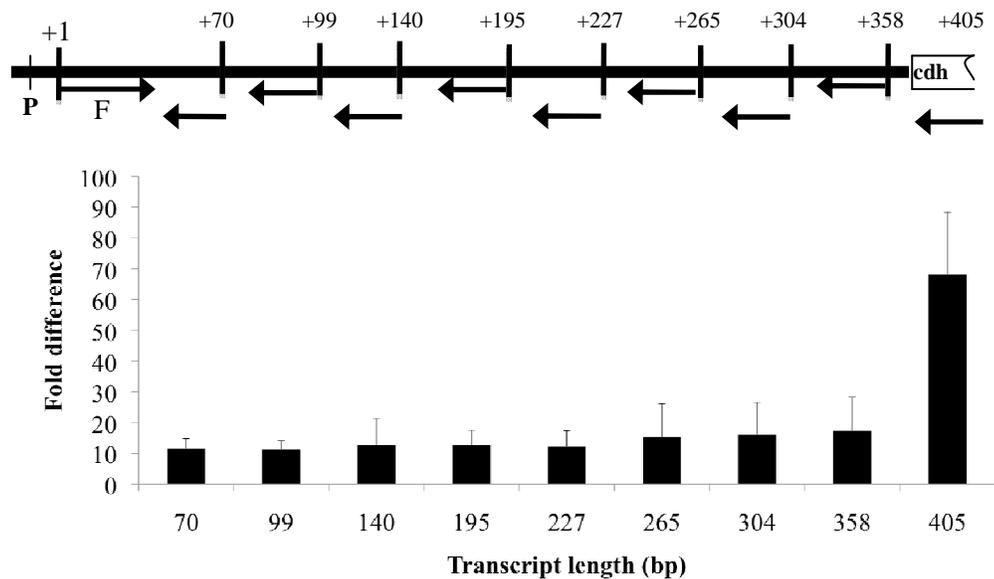


Figure 4.7. Differences in levels of transcripts of varying lengths demonstrated by qRT-PCR. Fold differences were calculated from the $C(t)$ values for RNA from acetate- versus methanol-grown cells. Arrows above the graph represent the position of primers used to generate templates of different lengths.

4.4.8. Regulation of different TATA-binding proteins by different substrates

Using primers to generate transcripts of different lengths, it was determined that the majority of the regulation within the 5' leader region was at the level of transcription elongation. A small portion this regulation, however, appeared to be at the level of transcription initiation. Since *trans* factors acting upstream of the transcriptional start site were not found, components involved in transcription initiation were investigated. It has been theorized that use of different TBPs could be involved in regulation at the level of transcription initiation. To determine whether differential transcription factor pairing was a possible mechanism of regulation for

cdh, *M. acetivorans* with translational fusions of TATA-binding protein (TBP) promoters ligated to *lacZ* assayed for β -galactosidase activities during growth with different substrates. The results of these experiments revealed that neither *tbp-2::lacZ* or *tbp-3::lacZ* were expressed under any substrate conditions, as the specific activities ranged from 0 to 0.04 per $\mu\text{g/ml}$ of protein. *tbp-1::lacZ*, however, was expressed when cells were grown with any of the three substrates. It was also observed that *tbp-1::lacZ* had a higher level of expression when grown in the presence of acetate with a specific activity of 44.4 ± 4.2 per $\mu\text{g/ml}$ of protein. In contrast, the specific activities in cells grown on methanol and TMA were 10.93 ± 0.9 and 11.5 ± 0.4 per $\mu\text{g/ml}$ of protein, respectively.

4.5. Discussion

When grown on acetate, members of the genus *Methanosarcina* express CODH/ACS, a key enzyme complex that catalyzes the cleavage of the acetyl C–C bond of coenzyme CoA (43, 61). As the CODH/ACS operon is highly expressed during acetate growth, the effects of deletions in sequences adjacent to the transcriptional start site were studied to identify *cis* elements critical for regulation. Deleting sequence upstream of the TSS had no significant effect on regulation indicating that upstream *trans* elements were not involved in regulation of transcription initiation. In contrast, downstream deletions in sequence within the highly conserved 371 base pair 5' leader region prior to the ribosomal binding site were shown to have a significant effect on the regulation of CODH/ACS expression. Further deletion of putative secondary structures within the 5' leader sequence were also shown to have a significant effect on regulation.

Differential expression of transcript within the 358 bp downstream of the 5' end accounted for approximately 22% of the difference in *cdh* expression observed between acetate and methanol grown cells. However, when sequence within the 5' leader region 133 bases downstream of the promoter was deleted, gene expression was up-regulated relative to wild type during methylotrophic growth of *M. acetivorans* with TMA or methanol. *Trans*-acting DNA binding factors such as repressors typically bind immediately adjacent to the promoter to cause disruption of RNA polymerase binding. The observations in this study suggest that the observed regulation of transcriptional initiation results from a factor that either binds to the

promoter or interacts with the basal transcription initiation components. Possible mechanisms might include differential expression of a TATA binding protein or TFB. Multiple genes encoding for multiple copies of TBPs and transcription factor B (TFB) have been observed in the genomes of several Archaea. This includes, among the methanogens, three copies of TBP in *M. acetivorans*, *M. barkeri* and *M. mazei*, two copies of TBP in *Methanospirillum hungatei* and two copies of transcription initiation factor B (TFB) in *Methanopyrus kandleri*. We also tested for the possibility that the three TBPs detected in methanosarcinal genomes could form different pairings with TFB in response to different growth substrates to regulate gene expression, but only TBP-1 was expressed at detectable levels. However, some role in *cdh* regulation by TBP-1 cannot be ruled out as it was expressed at greater levels during growth on acetate. Up-regulation of TBP and TFB have been reported for other Archaea in response to UV irradiation (8) and heat shock (104, 123), but a clear association between TBP/TFB expression and regulation of a target promoter has yet to be established.

The results also indicate that the 5' leader region has a role in post-transcriptional regulation of CODH/ACS via early termination of elongation during methylotrophic growth. This conclusion is supported by several observations. First, post-transcriptional regulation by differential translation was ruled out, as the fold difference of the protein levels and the transcript levels were not significantly different. Second, *cdh* transcript stability was similar in cells grown acetoclastically and methylotrophically, ruling out differential mRNA degradation as a possible mechanism. Finally, a significant difference in transcript levels was observed 405

bases downstream of the 5' end of the transcript. Deletion of the leader region encoding sequence between +211 and +335 had no effect on expression in the lacZ reporter fusion and regulation and qRT-PCR analysis of *cdh* transcript showed differential transcript levels occurred between 358 and 405 bp downstream of the 5' end. This observation is consistent with Northern analysis of CODH/ACS transcript that showed a rapid decay of signal after acetate-grown cells were spiked with methanol using a probe that hybridized 34 bases into the 3' leader region (114). Deletions upstream of *cdhA* within UTR sequence between +28 to +211 also had a significant effect on expression, which suggests that this region has a role in the differential termination of elongation downstream. However, the exact role this sequence plays has yet to be determined. The results support the hypothesis that the CODH/ACS operon is regulated by two mechanisms: differential transcription initiation that accounts for approximately a quarter of the differential expression and early termination of elongation that accounts for the balance of the differential expression observed between acetate and methanol-grown cells.

In bacterial systems, long 5' leader regions are involved in transcriptional regulation via multiple mechanisms. Usually, these mechanisms involve the formation of secondary structures within the RNA. One example of this type of regulation is the attenuation mechanism, which involves the formation of different RNA secondary structures that either inhibit or promote transcription elongation (133). This type of regulation involves the formation of stem loop structures, along with a long stretch of uridines immediately following one of the structures. Interestingly, a conserved stretch of predicted uridines occurs in positions +71 to +77

in 5' leader regions of all four methanosarcinal CODH/ACS orthologs. Another mechanism found in bacteria involves intrinsic terminators such as riboswitches. This type of regulation motif has been found in many bacterial systems in which the RNA forms different secondary structures to either allow or terminate transcription (103). There are multiple permutations of this system present in both eukaryotes and prokaryotes. One of the most well known systems involves regulation of tryptophan biosynthesis by a conformational change in mRNA secondary structure that causes early transcription termination in the presence of tryptophan (132). Other systems involve a regulatory protein (103) or ligands (78, 80, 117, 118, 129) that bind to the RNA, which creates a conformational change that either terminates or allows transcription to occur. Early termination of transcriptional elongation in sequence distal to the regulatory UTR is consistent with regulatory mechanisms involving changes in secondary structure, such as attenuation or riboswitches. Although the experimental evidence indicates that the 5' leader region of methanosarcinal CODH/ACS has a role in controlling early termination of elongation, the precise regulatory mechanism is currently unknown.

In the Archaea, there are several examples of transcriptional regulation mechanisms. Most of these examples involve activator or repressor proteins which bind upstream of the transcriptional start site and either induce or inhibit transcription. The difference between these known systems and the one described here is that the previous studies involve transcriptional regulation occurring prior to the start of transcription. 5' UTRs identified within the Archaea include the 113 bp UTR identified upstream of a DEAD-box RNA helicase in the Antarctic methanogen

Methanococoides burtonii (70) and the methyltransferase genes in *Methanosarcina* spp. (18). In both of these examples, the 5' UTR was implicated in regulation, although the role of the 5' UTR in regulation was not confirmed. The system described in this paper appears to involve sequences downstream of the start of transcription, well into the 5' leader region.

To the best of our knowledge this is the first evidence of regulation of transcriptional elongation by a 5' leader region as a mechanism for gene regulation in the Archaea. Further studies to identify putative *trans*-acting elements and secondary structures are necessary to characterize the paradigm for catabolic CODH/ACS regulation in these Archaea.

Chapter 5: Discussion and Future Perspectives

The current knowledge of how transcription occurs in Archaea has indicated that the machinery involved in transcription appears to be homologous to eukaryotic systems. Some transcription factors, such as TBP and TFB, have been identified. *In vitro* transcription systems have demonstrated that transcription in Archaea can be achieved with relatively few transcription factors when compared to eukaryotic systems. However, this is not to say that there are not other proteins that facilitate transcription in Archaea. With the sequencing of multiple Archaeal genomes, multiple putative transcription factors and regulators have been annotated. Further experiments are needed to characterize the function of these putative transcription factors. With the development of *in vitro* transcription systems, gene expression systems, and gene deletion systems, proteins of unknown function can be studied to verify their annotation and to determine their role *in vivo* in cellular processes.

Understanding how catabolic gene regulation occurs in methanogenic Archaea is essential to understanding how these organisms function in fermentative processes, yet little is known about this regulation. Carbon monoxide dehydrogenase/acetyl CoA synthase is a key enzyme in acetoclastic methanogenesis, and this gene is regulated at the transcriptional level in response to substrate (65). The work presented here shows for the first time that the 5' leader region located directly downstream of the transcriptional start site of *cdhA* is involved in transcriptional

regulation (Chapter 4). Other examples of 5' untranslated leader regions involved in regulation are present in the literature (18, 70). There are also studies in haloarchaea that have characterized the abundance of leadered transcripts. Brenneis et al demonstrated that, of 40 genes studied, the majority were leaderless. Other studies on haloarchaea have indicated that these 5' UTRs have a role in translational regulation (20, 21). Overall, it appears as though most archaeal transcripts do not have long 5' UTRs.

As stated above some archaeal 5' UTRs are involved in regulation, but a mechanism(s) for this regulation has not yet been identified. This study is the first to show that transcriptional regulation of archaeal CODH/ACS occurs, for the most part, at the level of transcription elongation from within the 5' leader region. A smaller fraction of the regulation appears to occur at the level of transcription initiation, although this result needs to be investigated further.

Based on the results presented in this dissertation, I propose a model for the regulation of expression of CODH/ACS (Figure 5.1). When *Methanosarcina* species are grown on methylotrophic substrates, such as methanol or trimethylamine, a terminator protein binds to the end of the transcribed RNA and translocates along the length of the RNA. This protein would function similar to Rho-dependent termination in bacteria, involving the helicase protein Rho. Once the protein reached the DNA-RNA complex, it would unwind the RNA, causing termination of transcription. However, when cells are grown in acetate, a *trans* factor would be activated that could bind to the RNA, generating secondary structures in the RNA. These structures would inhibit the progress of the terminator protein, preventing it

from interfering with transcription, and allowing transcription to continue. The *trans* factor protein would be inhibited in the presence of methylotrophic substrates, to the point that it would only be active when there was only acetate present.

In the genome of *M. acetivorans*, there is no annotated homolog of Rho present. However, there are homologs of other factors involved in Rho-dependent termination, such as NusG. There is an annotated homolog of NusG in *M. acetivorans* (MA4273), *M. mazei* (MM_1010), and *M. barkeri* (Mbar_A0616). It is possible that there is a homolog of Rho present, but it has not been correctly annotated. Carefully analysis of the genome, along with functional studies, would help to determine if a homolog of this transcription factor is present. But the model presented in the previous paragraph is supported by the data in this dissertation. I have shown that regulation is at the level of transcription elongation, and that this regulation is occurring within the 5' leader region. Putative stem loop structures and a conserved polyU sequence appear to be present in the 5' leader region and possibly play a role in regulation. Based on the *in vitro* transcription assay, there are numerous site along the 5' leader region where pausing of the RNA polymerase occurs during transcription. The work in Chapter 3 failed to identify any *trans* factors that bind to the 5' leader region, but it is still possible that the conditions tested were not ideal for isolation of the protein. Further research is needed to determine if there indeed are any *trans* factors that can possibly bind to this region and are involved in transcriptional regulation.

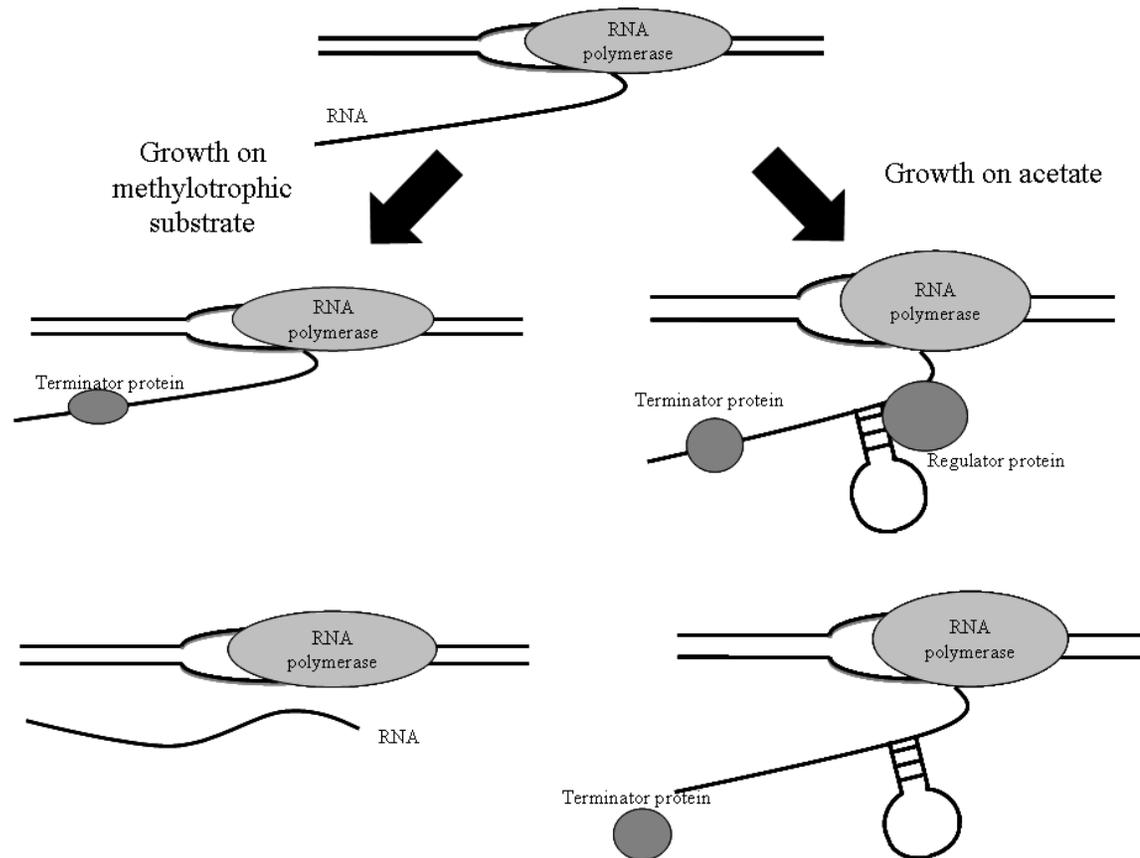


Figure 5.1. Proposed model for the regulation of *cdhABCDE*. The steps on the left illustrate what occurs during methylotrophic growth. Under these conditions, a terminator protein binds to the newly synthesized RNA, interfering with the RNA polymerase and terminating transcription. The steps on the right illustrate what occurs during acetate growth. Under this condition, the terminator protein still binds, but secondary structures which form within the RNA, due to a regulator protein, prevent termination of transcription.

The model presented in the previous paragraph is not the only possible mechanism for regulation. There are multiple mechanisms present in other organisms that involve 5' leader regions. Usually, these mechanisms involve the formation of secondary structures within the RNA. One example of this type of regulation is the attenuation mechanism, which involves the formation of different RNA secondary structures that either inhibit or promote transcription elongation (133). These types of regulatory mechanisms are usually characterized by a region of poly-U's located downstream of one of the putative loop structures. Interestingly, one of the predicted loop structures found in this dissertation also had a stretch of poly-U's downstream. This type of regulatory mechanism can involve the intrinsic formation of secondary structure, or the formation of the structures can be mediated by a regulatory protein. This type of mechanism could also be involved here, as the poly-U region would allow the RNA polymerase to pause during transcription. Under growth on methylotrophic substrates, the terminator protein could catch up to the elongation complex and cause the termination of transcription. For growth under acetate, a regulatory protein could interfere with the terminator protein, preventing transcription termination.

Another mechanism found in bacteria involves intrinsic terminators such as riboswitches. This type of mechanism involves ligands (78, 80, 117, 118, 129) that bind to the RNA, which creates a conformational change that either terminates or allows transcription to occur. While this mechanism could be feasible for regulation

of *cdhABCDE*, further research would be necessary to determine what ligand could be controlling the regulation.

The 5' leader region of *cdhABCDE* needs to be further characterized to determine its specific role in regulation of the gene. As stated previously, 5' leader regions have been implicated in transcription and translation regulation in bacteria. It is feasible that these regions in Archaea function in a similar way, involving putative secondary structures. This type of mechanism would allow the cells to respond quickly to changes in the environment. Since acetate is not the favored substrate for these organisms, this type of mechanism would allow the organisms to quickly shut off expression of CODH/ACS, which when expressed accounts for up to 20% of the total cell proteins, when in the presence of preferred substrates, such as methylotrophic substrates. Future work should focus on investigating this region for specific mechanisms, such as riboswitches and attenuators. Using *in vitro* transcription assays would be essential for these studies, in order to determine the exact sequence that is important for the regulation. Unfortunately, there is no *in vitro* transcription assay available at the moment that uses the components from *M. acetivorans*, although this system will hopefully be developed in the near future. Currently, the only system available for use with methanogenic Archaea involves components from *Methanothermobacter thermoautotrophicum*. A *M. acetivorans in vitro* transcription system will allow for the use of native transcriptional components to study transcription, which combined with site-directed mutagenesis is essential for identifying mechanisms at both the DNA and RNA levels.

Further work is needed to characterize the exact mechanism by which the regulation is occurring, particularly at the level of transcription initiation. More investigation into the putative promoter region is necessary to identify factors involved in differential transcription initiation. *Methanosarcina* species often contain multiple copies of some transcription factors, such as TATA-binding protein (TBP) and transcription factor B (TFB). Other Archaea have also been predicted to contain multiple copies of TBPs and TFBs. It has been speculated that differential transcription factor pairing could play a role in transcriptional regulation, although this theory has not been proven. Previous work in our lab in which the putative promoter regions of the three TBPs present in the genome of *Methanosarcina acetivorans* were fused to *lacZ* to make translational fusions demonstrated that TBP-1 appears to be regulated in response to substrate. Further analysis is needed to evaluate if this gene could be playing a role in the transcriptional initiation regulation that is seen with *cdhABCDE*. Overexpression of the putative TBP proteins would allow for other experiments, such as DNA-binding assays, to determine if different TBPs have varying affinities for promoters under different conditions. More characterization of transcription initiation is also necessary to further determine which transcription factors are present, and to identify the role of these factors in transcription. Development of an *in vitro* transcription system that uses *Methanosarcina* components is essential for identifying other factors. Since genome annotation has discovered multiple putative transcription factors, these factors could be overexpressed and purified and used in *in vitro* transcription assays to determine their effects on transcription. The availability of gene disruption systems also allows

investigators to disrupt the genes of interest to determine their effect on transcription *in vivo*.

This is the first study to elucidate the mechanism of regulation involving a 5' leader sequence in Archaea. Understanding the mechanism of regulation from this 5' leader will allow greater understanding of catabolic gene regulation, as well as increasing the knowledge of how the acetoclastic pathway is regulated. Since the majority of the biologically produced methane in methanogenic consortia of microorganisms is produced from acetate, which is also a rate-limiting step, understanding how this process is regulated is essential for understanding how carbon flow is effectively regulated throughout the consortia for optimal conversion of biomass to methane and carbon dioxide. Biomass conversion is directly dependent on the interaction of three groups of microorganisms, and one of the rate limiting steps involves acetoclastic methanogenesis. Because methanogenic Archaea are important for many industries, including waste treatment and processing, understanding more about how catabolic pathways function and are regulated can lead to more efficient processes. This is also important, because methane is both a greenhouse gas and an energy source, thus being able to control methane production and substrate utilization would help increase production of biogas as an alternative form of energy.

Thoroughly understanding how transcriptional regulation of *cdhABCDE* occurs will also help to give insight to other possible catabolic gene regulation mechanisms within the methanogenic Archaea. Since the only known regulation mechanisms in Archaea involve DNA-binding proteins that bind near the start of

transcription, this mechanism would be the first of this type to be described for Archaea. It would present a new mechanism by which Archaea are regulating gene expression, and it would open the field to the study of other 5' leader regions, which may be acting in the same fashion. This understanding of the regulation of methanogenic processes, in particular aceticlastic methanogenesis, would have global importance. The conversion of biomass by a methanogenic consortia provides a renewable energy resource in the form of methane, and is also functional as waste treatment processes.

Although 5' leader regions identified for other archaeal genes have been postulated to be involved in regulation, this was the first study to demonstrate a regulatory role by an archaeal leader sequence for differential transcription elongation. Based upon the results described in this dissertation, a model can be proposed for the regulation of catabolic CODH/ACS. This model involves a Rho-dependent like mechanism, similar to that found in bacteria. When grown on methylotrophic substrates, transcription elongation would be terminated by a *trans* element that binds to the RNA. When grown under acetate, the terminator protein would be inhibited by another *trans* element, which would allow transcription to continue. Further testing is necessary to confirm the proposed model. Identifying the regulatory mechanism of catabolic genes such as CODH/ACS is critical for understanding the regulatory strategies employed by methanogenic consortia for efficient conversion of biomass to methane.

Appendix A: Supplemental Information

Table S2.1. Primers used in Chapter 2 for characterization of *rad25* from *M. thermophila*.

<u>Primer</u>	<u>Primer sequence</u>	<u>Purpose</u>
#68	5' GANATGCCNTCNGGNACNGGNAA-3'	Cloning helicase sequences
#69	5'-CAAANTCNATNCCTTCNGANACN-3'	Cloning helicase sequences
#92	5'-CTTACATTTCCCTTTATGAGAATA-3'	Primer extension
#93	5'-AAGGACGCTATCTTCAAATCAAT-3'	Primer extension
#130	5'-CAGGAAGTGGAGAGACCCTTGTTGG-3'	Site-directed mutagenesis
#131	5'-CCAACAAGGGTCTCTCCACTTCCTG-3'	Site-directed mutagenesis
#157	5'-CGCCCAATTGGTAATACTATTC-3'	PCR for <i>pac</i> cassette
#158	5'-CACAGGAACAATTGACGGCTG-3'	PCR for <i>pac</i> cassette

Table S3.1. Primers used in Chapter 3.

<u>Primer</u>	<u>Sequence</u>
UTR1-bio	5'-bio-AATTAGTGTTTATTATCCATCGGTAGCG-3'
UTR2	5'-CATGTGAGCTTTACCTCCTTAATTTG-3'
EMSA1-bio	5'-bio-GGAAAAATCTATATAATTATTTGTTTTAATATGAA-3'
EMSA2	5'-GGGTAAAGTGGAATGTAGAGGCACACTATTC-3'
EMSA1	5'-GGAAAAATCTATATAATTATTTGTTTTAATATGAA-3'
#292	5'-GGCTTCAAAAACCGCTTCTCTGGC-3'
#293	5'-GAGCTTGCAGGGCTCCTTATTGAAACC-3'
TM1up1	5'-TGAAGAATCTTTAAAATGAGTTGTGGTTAGTATGAA-3'
TM1up2	5'-GGGTAAAATGGAATTTAGTGATACCTTCTTC-3'
Barkeri-up1	5'-ACAAATATTTACATAATAATTTGTTGTGAACACGAC-3'
Barkeri-up2	5'-TGGGTAAAGTGGAATGTTCAAATC-3'
#60	5'-GCCATTGGGATATATCAACGG-3'
#62	5'-AGAATCGATAAAAATGAGTTGT-3'
#63	5'-ATTTTATCGATTGCGAAATAG-3'
#64	5'-TTATATCGATTTGGTACATTT-3'

Table S4.1. Primers used to construct deletion mutants, TBP constructs and qRTPCR products in Chapter 4.

Primers	Name	Location in reference to Transcriptional start site
Deletion constructs		
TGAAGAAGTATCGATAATCAA	#53	—
GCCATTGGGATATATCAACGG	# 60	—
TTATATCGATTTGGTACATTT	# 64	- 29 bp
CTCAATGTCGACATTAACGT	# 65	+ 211 bp
AAGTCGACTTTAGGGCCTTAT	# 66	+ 133 bp
CGTACCGGTCGACAATAAAC	# 67	+ 28 bp
AGGAAACAGCTATGACCATGATTACGGTA	A1	—
AACCGTTTTTACCAACGATTGG/GCTACCGATGGATAATAAAC	B1	+ 51 bp / + 8 bp
CCAATCGTTGGTAAAAACGGTT	C1	+ 51 bp
TCATCACCGGATCCAGAGC	D1	+ 365 bp
TACGCCCTATGTCTCAAAAAACATTGGGACCT	B2	+ 46 bp
GTTTTTTTGAGACATAGGGCGTA	C2	+ 70 bp
CTTGTAAGAGTCGTTTTTGTCTCAAAAAAC	B3	+ 70 bp
GAGACAAAAACGACTTTTACAAG	C3	+ 78 bp
CCTTTTTACGCCCTATGTCTCCCGTTTTTACCAACG	B4	+ 56 bp
GAGACATAGGGCGTAAAAAGG	C4	+ 78 bp
TBP constructs		
CGAAATCCTCGAGGTGCGGTC	#161	—
GGTGGATCCAACCACGTTTTTC	#162	—
GGGGAGACTCGAGAAAACCTCAGG	#165	—
CTCTATAGTTATTGTGGATCCCATGCC	#166	—
CCCCGTCCTCGAGATACAGTCTC	#167	—
GTTCTCTATAGTTATTGTGGGATCCATAACAAC	#168	—
qRTPCR products		
AATTAGTGTTTATTATCCCATCGGTAGCG	#221	+ 1 bp
CCTCCTAAATTTGTTTAAAGTCGACTTCG	#222	+ 358 bp
CCGTTTTTGCCAACGATTGAG	#228	+ 70 bp
CTTCTAAATTCCTTTTTTACGCCCTATATCTT	#229	+ 99 bp
CTTCCAAAAGTCGTTTTTLAGGGC	#306	+ 140 bp
CAAACGTCATACCAACATTTTTTCAAAT	#307A	+ 195 bp
GGCAAAACGTTTTTLAGCTCTCG	#307	+ 227 bp
CCGCCTTAAAGTTACAACCGTTTAAATG	#307B	+ 265 bp
AGCGGCATATTCTTAATGATTATACATT	#308	+ 304 bp
ATCTGAACGGATTCCAGATCTTC	#315	+ 405 bp
CGCGCTCCGTTTCAGCCAC	#334	N/A ¹
GTTACCTCCAACATCACAAATCCTGG	#335	N/A ¹

Bibliography

1. **Aceti, D. J., and J. G. Ferry.** 1988. Purification and characterization of acetate kinase from acetate-grown *Methanosarcina thermophila*: Evidence for regulation of synthesis. *J Biol Chem* **263**:15444.
2. **Alber, B. E., and J. G. Ferry.** 1994. A carbonic anhydrase from the archaeon *Methanosarcina thermophila*. *Proc Natl Acad Sci USA* **91**:6909-6913.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Meyers, and D. J. Lipman.** 1990. Basic local alignment tool. *J Mol Biol* **215**:403-410.
4. **Apolinario, E. A., and K. R. Sowers.** 1996. Plate colonization of *Methanococcus maripaludis* and *Methanosarcina thermophila* in a modified canning jar. *FEMS Microbiol Lett* **145**:131-137.
5. **Apolinario, E. E., K. M. Jackson, and K. R. Sowers.** 2005. Development of a plasmid-mediated reporter system for in vivo monitoring of gene expression in the archaeon *Methanosarcina acetivorans*. *Appl Environ Microbiol* **71**:4914-4918.
6. **Aravind, L., D. R. Walker, and E. V. Koonin.** 1999. Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res* **27**:1223-1242.
7. **Balch, W. E., G. E. Fox, L. J. Magrum, C. R. Woese, and R. S. Wolfe.** 1979. Methanogens: reevaluation of a unique biological group. *Microbiol Rev* **43**:260-296.

8. **Baliga, N. S., S. J. Bjork, R. Bonneau, M. Pan, C. Iloanusi, M. C. Kottemann, L. Hood, and J. DiRuggiero.** 2004. Systems level insights into the stress response to UV radiation in the halophilic archaeon *Halobacterium* NRC-1. *Genome Res* **14**:1025-1035.
9. **Baliga, N. S., S. P. Kennedy, W. V. Ng, L. Hood, and S. DasSarma.** 2001. Genomic and genetic dissection of an archaeal regulon. *Proc Natl Acad Sci USA* **98**:2521-2525.
10. **Barcelona, M. J., H. M. Liljestrand, and J. J. Morgan.** 1980. Determination of low molecular weight volatile fatty acids in aqueous samples. *Anal Chem* **52**:321-325.
11. **Barns, S. M., C. F. Delwiche, J. D. Palmer, and N. R. Pace.** 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci USA* **93**:9188-9193.
12. **Batty, D. P., and R. D. Wood.** 2000. Damage recognition in nucleotide excision repair of DNA. *Gene* **241**:193-204.
13. **Bell, S. D.** 2005. Archaeal transcriptional regulation--variation on a bacterial theme? *Trends Microbiol* **13**:262-265.
14. **Bell, S. D., S. S. Cairns, R. L. Robson, and S. P. Jackson.** 1999. Transcriptional regulation of an archaeal operon *in vivo* and *in vitro*. *Mol Cell* **4**:971-982.
15. **Bell, S. D., and S. P. Jackson.** 2000. Mechanism of autoregulation by an archaeal transcriptional repressor. *J Biol Chem* **275**:31624-31629.

16. **Bini, E., V. Dikshit, K. Dirksen, M. Drozda, and P. Blum.** 2002. Stability of mRNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. RNA **8**:1129-1136.
17. **Blaut, M., and G. Gottschalk.** 1982. Effect of trimethylamine on acetate utilization by *Methanosarcina barkeri*. Arch Microbiol **133**:230-235.
18. **Bose, A., M. A. Pritchett, and W. W. Metcalf.** 2008. Genetic analysis of the methanol- and methylamine-specific methyltransferase 2 genes of *Methanosarcina acetivorans* C2A. J Bacteriol **190**:4017-4026.
19. **Brady, R. A., J. G. Leid, A. K. Camper, J. W. Costerton, and M. E. Shirliff.** 2006. Identification of *Staphylococcus aureus* proteins recognized by the antibody-mediated immune response to a biofilm infection. Infect Immun **74**:3415-3426.
20. **Brenneis, M., O. Hering, C. Lange, and J. Soppa.** 2007. Experimental characterization of Cis-acting elements important for translation and transcription in halophilic Archaea. PLoS Genet **3**:e229.
21. **Brenneis, M., and J. Soppa.** 2009. Regulation of translation in haloarchaea: 5'- and 3'-UTRs are essential and have to functionally interact in vivo. PLoS ONE **4**:e4484.
22. **Brinkman, A. B., S. D. Bell, R. J. Lebbink, W. M. de Vos, and J. van der Oost.** 2002. The *Sulfolobus solfataricus* Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability. J Bio Chem **277**:29537-29545.

23. **Brinkman, A. B., T. J. Ettema, W. M. de Vos, and J. van der Oost.** 2003. The Lrp family of transcriptional regulators. *Mol Microbiol* **48**:287-294.
24. **Brochier, C., P. Forterre, and S. Gribaldo.** 2004. Archaeal phylogeny based on proteins of the transcription and translation machineries: tackling the *Methanopyrus kandleri* paradox. *Genome Biol* **5**:R17.
25. **Brochier, C., S. Gribaldo, Y. Zivanovic, F. Confalonieri, and P. Forterre.** 2005. Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol* **6**:R42.
26. **Burke, S. A., S. L. Lo, and J. A. Krzycki.** 1998. Clustered genes encoding the methyltransferases of methanogenesis from monomethylamine. *J Bacteriol* **180**:3432-3440.
27. **Carpentieri, F., M. De Felice, M. De Falco, M. Rossi, and F. M. Pisani.** 2002. Physical and functional interaction between the mini-chromosome maintenance-like DNA helicase and the single-stranded DNA binding protein from the crenarchaeon *Sulfolobus solfataricus*. *J Biol Chem* **277**:12118-12127.
28. **Dahlke, I., and M. Thomm.** 2002. A *Pyrococcus* homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment. *Nucleic Acids Res* **30**:701-710.
29. **Del Sal, G., G. Manfioletti, and C. Schneider.** 1989. The CTAB-DNA precipitation method: a common mini-scale preparation of template DNA

- from phagemids, phages or plasmids suitable for sequencing. *Biotechniques* **7**:514-520.
30. **Deppenmeier, U., A. Johann, T. Hartsch, R. Merkl, R. A. Schmitz, R. Martinez-Arias, A. Henne, A. Wiezer, S. Bäumer, C. Jacobi, H. Brüggemann, T. Lienard, A. Christmann, M. Bömeke, S. Steckel, A. Bhattacharyya, A. Lykidis, R. Overbeek, H.-P. Klenk, R. P. Gunsalus, H. J. Fritz, and G. Gottschalk.** 2002. The genome of *Methanosarcina mazei*: Evidence for lateral gene transfer between Bacteria and Archaea. *J Mol Microbiol Biotechnol* **4**:453-461.
31. **Eggen, R. I., A. C. Geerling, M. S. Jetten, and W. M. de Vos.** 1991. Cloning, expression, and sequence analysis of the genes for carbon monoxide dehydrogenase of *Methanotherix soehngenii*. *J Biol Chem* **266**:6883-6887.
32. **Eggen, R. I. L., R. Vankranenburg, A. J. M. Vriesema, A. C. M. Geerling, M. F. J. M. Verhagen, W. R. Hagen, and W. M. Devos.** 1996. Carbon monoxide dehydrogenase from *Methanosarcina frisia* Go1 - characterization of the enzyme and the regulated expression of two operon-like *cdh* gene clusters. *J Biol Chem* **271**:14256-14263.
33. **Ellis, N. A.** 1997. DNA helicases in inherited human disorders. *Curr Opin Genet Dev* **7**:354-363.
34. **Ferguson, T. J., and R. A. Mah.** 1983. Effect of H₂-CO₂ on methanogenesis from acetate or methanol in *Methanosarcina* spp. *Appl Environ Microbiol* **46**:348-355.

35. **Forterre, P., C. Brochier, and H. Philippe.** 2002. Evolution of the archaea. *Theor Pop Biol* **61**:409-422.
36. **Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, D. T. A., R. S. Wolfe, W. E. Balch, R. S. Tanner, L. J. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonner, B. Lewis, J., D. Stahl, A., K. Luehrsen, R., K. Chen, N., and C. Woese, R.** 1980. The phylogeny of prokaryotes. *Science* **209**:457-463.
37. **Galagan, J. E., C. Nusbaum, A. Roy, M. G. Endrizzi, P. Macdonald, W. FitzHugh, S. Calvo, R. Engels, S. Smirnov, D. Atnoor, A. Brown, N. Allen, J. Naylor, N. Stange-Thomann, K. DeArellano, R. Johnson, L. Linton, P. McEwan, K. McKernan, J. Talamas, A. Tirrell, W. J. Ye, A. Zimmer, R. D. Barber, I. Cann, D. E. Graham, D. A. Grahame, A. M. Guss, R. Hedderich, C. Ingram-Smith, H. C. Kuettner, J. A. Krzycki, J. A. Leigh, W. X. Li, J. F. Liu, B. Mukhopadhyay, J. N. Reeve, K. Smith, T. A. Springer, L. A. Umayam, O. White, R. H. White, E. C. de Macario, J. G. Ferry, K. F. Jarrell, H. Jing, A. J. L. Macario, I. Paulsen, M. Pritchett, K. R. Sowers, R. V. Swanson, S. H. Zinder, E. Lander, W. W. Metcalf, and B. Birren.** 2002. The genome of *Methanosarcina acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res* **12**:532-542.
38. **Geiduschek, E. P., and M. Ouhammouch.** 2005. Archaeal transcription and its regulators. *Mol Microbiol* **56**:1397-1407.
39. **Gohl, H. P., B. Grondahl, and M. Thomm.** 1995. Promoter recognition in archaea is mediated by transcription factors: identification of transcription

- factor aTFB from *Methanococcus thermolithotrophicus* as archaeal TATA-binding protein. *Nucleic Acids Res* **23**:3837-3841.
40. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res* **17**:4713-4730.
41. **Gorg, A., C. Obermaier, G. Boguth, A. Harder, B. Scheibe, R. Wildgruber, and W. Weiss.** 2000. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **21**:1037-1053.
42. **Grahame, D. A., S. Gencic, and E. DeMoll.** 2005. A single operon-encoded form of the acetyl-CoA decarboxylase/synthase multienzyme complex responsible for synthesis and cleavage of acetyl-CoA in *Methanosarcina thermophila*. *Arch Microbiol* **184**:32-40.
43. **Grahame, D. A., and T. C. Stadtman.** 1987. Carbon monoxide dehydrogenase from *Methanosarcina barkeri*. Disaggregation, purification, and physicochemical properties of the enzyme. *J Biol Chem* **262**:3706-3712.
44. **Guss, A. M., B. Mukhopadhyay, J. K. Zhang, and W. W. Metcalf.** 2005. Genetic analysis of mch mutants in two *Methanosarcina* species demonstrates multiple roles for the methanopterin-dependent C-1 oxidation/reduction pathway and differences in H-2 metabolism between closely related species. *Mol Microbiol* **55**:1671-1680.

45. **Guzder, S. N., P. Sung, V. Bailly, L. Prakash, and S. Prakash.** 1994. RAD25 is a helicase required for DNA repair and RNA polymerase II transcription. *Nature* **369**:578-581.
46. **Hanzelka, B. L., T. J. Darcy, and J. N. Reeve.** 2001. TFE, an archaeal transcription factor in *Methanobacterium thermoautotrophicum* related to eukaryal transcription factor TFIIE. *J Bacteriol* **183**:1813-1818.
47. **Hausner, W., and G. T. Frey, M.** 1991. Control regions of an archaeal gene. A TATA box and an initiator element promote cell-free transcription of the tRNA^{Val} gene of *Methanococcus vannielii*. *J Mol Biol* **222**:495-508.
48. **Hausner, W., U. Lange, and M. Musfeldt.** 2000. Transcription factor S, a cleavage inducitor factor of the archaeal RNA polymerase. *J Bio Chem* **275**:12393-12399.
49. **Hausner, W., and M. Thomm.** 1995. The translation product of the presumptive *Thermococcus celer* TATA-binding protein sequence is a transcription factor related in structure and function to *Methanococcus* transcription factor B. *J Biol Chem* **270**:17649-17651.
50. **Hausner, W., J. Wettach, C. Hethke, and M. Thomm.** 1996. Two transcription factors related with the eucaryal transcription factors TATA-binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase. *J Biol Chem* **271**:30144-30148.
51. **Heddrich, R. C., and W. B. Whitman.** 2005. Physiology and Biochemistry of the Methane-Producing Archaea. *In* M. Dworkin (ed.), *The Prokaryotes*:

An Evolving Electronic Resource for the Microbiological Community, 3rd ed.
Springer-Verlag, New York, N.Y.

52. **Hovey, R., S. Lentes, A. Ehrenreich, K. Salmon, K. Saba, G. Gottschalk, R. P. Gunsalus, and U. Deppenmeier.** 2005. DNA microarray analysis of *Methanosarcina mazei* Go1 reveals adaptation to different methanogenic substrates. *Mol Genet Genomics* **273**:225-239.
53. **Huber, H., M. J. Hohn, R. Rachel, T. Fuchs, V. C. Wimmer, and K. O. Stetter.** 2002. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**:63-67.
54. **Jablonski, P. E., A. A. Dimarco, T. A. Bobik, M. C. Cabell, and J. G. Ferry.** 1990. Protein content and enzyme activities in methanol-grown and acetate-grown *Methanosarcina thermophila*. *J Bacteriol* **172**:1271-1275.
55. **Jager, A., R. Samorski, F. Pfeifer, and G. Klug.** 2002. Individual *gvp* transcript segments in *Haloferax mediterranei* exhibit varying half-lives, which are differentially affected by salt concentration and growth phase. *Nucleic Acids Res* **30**:5436-5443.
56. **Kelman, Z., J. K. Lee, and J. Hurwitz.** 1999. The single minichromosome maintenance protein of *Methanobacterium thermoautotrophicum* Delta H contains DNA helicase activity. *Proc Nat Acad Sci USA* **96**:14783-14788.
57. **Kelman, Z., and M. F. White.** 2005. Archaeal DNA replication and repair. *Curr Opin Microbiol* **8**:669-676.

58. **Kendall, E. M., and D. R. Boone.** 2004. The Order Methanosarcinales. *In* M. Dworkin (ed.), *The Prokaryotes: An evolving Electronic Resource for the Microbiological Community*, 3rd ed. Springer-Verlag, New York, N.Y.
59. **Kruger, K., T. Hermann, V. Armbruster, and F. Pfeifer.** 1998. The transcriptional activator GvpE for the halobacterial gas vesicle genes resembles a basic region leucine-zipper regulatory protein. *J Mol Biol* **279**:761-771.
60. **Krzycki, J. A., R. H. Wolkin, and J. G. Zeikus.** 1982. Comparison of unitrophic and mixotrophic substrate metabolism by an acetate-adapted strain of *Methanosarcina barkeri*. *J Bacteriol* **149**:247-254.
61. **Krzycki, J. A., and J. G. Zeikus.** 1984. Characterization and purification of carbon monoxide dehydrogenase from *Methanosarcina barkeri*. *J Bacteriol* **158**:231-237.
62. **Lange, U., and W. Hausner.** 2004. Transcriptional fidelity and proofreading in Archaea and implications for the mechanism of TFS-induced RNA cleavage. *Mol Microbiol* **52**:1133-1143.
63. **Langer, D., J. Hain, P. Thuriaux, and W. Zillig.** 1995. Transcription in Archaea: similarity to that in Eucarya. *Proc Natl Acad Sci USA* **92**:5768-5772.
64. **Lee, S. J., C. Moulakakis, S. M. Koning, W. Hausner, M. Thomm, and W. Boos.** 2005. TrmB, a sugar sensing regulator of ABC transporter genes in *Pyrococcus furiosus* exhibits dual promoter specificity and is controlled by different inducers. *Mol Microbiol* **57**:1797-1807.

65. **Li, L., Q. Li, L. Rohlin, U. Kim, K. Salmon, T. Rejtar, R. P. Gunsalus, B. L. Karger, and J. G. Ferry.** 2007. Quantitative proteomic and microarray analysis of the archaeon *Methanosarcina acetivorans* grown with acetate versus methanol. *J Proteome Res* **6**:759-771.
66. **Li, Q., L. Li, T. Rejtar, B. L. Karger, and J. G. Ferry.** 2005. Proteome of *Methanosarcina acetivorans* Part I: an expanded view of the biology of the cell. *J Proteome Res* **4**:112-128.
67. **Li, Q., L. Li, T. Rejtar, B. L. Karger, and J. G. Ferry.** 2005. Proteome of *Methanosarcina acetivorans* Part II: comparison of protein levels in acetate- and methanol-grown cells. *J Proteome Res* **4**:129-135.
68. **Lie, T., and J. Leigh.** 2003. A novel repressor of *nif* and *glnA* expression in the methanogenic archaeon *Methanococcus maripaludis*. *Mol Microbiol.* **47**:235-246.
69. **Lie, T. J., G. E. Wood, and J. A. Leigh.** 2005. Regulation of *nif* expression in *Methanococcus maripaludis* - Roles of the euryarchaeal repressor NrpR, 2-oxoglutarate, and two operators. *J Biol Chem* **280**:5236-5241.
70. **Lim, J., T. Thomas, and R. Cavicchioli.** 2000. Low temperature regulated DEAD-box RNA helicase from the Antarctic archaeon, *Methanococcoides burtonii*. *J Mol Biol* **297**:553-567.
71. **Lovley, D. R., and M. J. Klug.** 1982. Intermediary metabolism of organic matter in the sediments of a eutrophic lake. *Appl Environ Microbiol* **43**:552-560.

72. **Lundie, L. L., and J. G. Ferry.** 1989. Activation of acetate by *Methanosarcina thermophila* - purification and characterization of phosphotransacetylase. J Biol Chem **264**:18392-18396.
73. **Maeder, D. L., I. Anderson, T. S. Brettin, D. C. Bruce, P. Gilna, C. S. Han, A. Lapidus, W. W. Metcalf, E. Saunders, R. Tapia, and K. R. Sowers.** 2006. The *Methanosarcina barkeri* genome: comparative analysis with *Methanosarcina acetivorans* and *Methanosarcina mazei* reveals extensive rearrangement within methanosarcinal genomes. J Bacteriol **188**:7922-7931.
74. **Mah, R. A., M. R. Smith, T. Ferguson, and S. Zinder.** 1981. Methanogenesis from H₂-CO₂, methanol, and acetate by *Methanosarcina.*, p. 131-142. In H. Dalton (ed.), Microbial growth on C-1 compounds. Heyden and Son, Ltd., London.
75. **Makarova, K. S., L. Aravind, N. V. Grishin, I. B. Rogozin, and E. V. Koonin.** 2002. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. Nucleic Acids Res **30**:482-496.
76. **MaupinFurlow, J. A., and J. G. Ferry.** 1996. Analysis of the CO dehydrogenase/acetyl-coenzyme A synthase operon of *Methanosarcina thermophila*. J Bacteriol **178**:6849-6856.
77. **Metcalf, W. W., J. K. Zhang, E. Apolinario, K. R. Sowers, and R. S. Wolfe.** 1997. A genetic system for Archaea of the genus *Methanosarcina*:

- Liposome-mediated transformation and construction of shuttle vectors. Proc Natl Acad Sci USA **94**:2626-2631.
78. **Mironov, A. S., I. Gusarov, R. Rafikov, L. Errais Lopez, K. Shatalin, R. A. Kreneva, D. A. Perumov, and E. Nudler.** 2002. Sensing Small Molecules by Nascent RNA: A Mechanism to Control Transcription in Bacteria. Cell **111**:747-756.
79. **Mitta, M., L. Fang, and M. Inouye.** 1997. Deletion analysis of *cspA* of *Escherichia coli*: requirement of the AT-rich UP element for *cspA* transcription and the downstream box in the coding region for its cold shock induction. Mol Microbiol **26**:321-335.
80. **Nahvi, A., N. Sudarsan, M. S. Ebert, X. Zou, K. L. Brown, and R. R. Breaker.** 2002. Genetic control by a metabolite binding mRNA. Chem Biol **9**:1043.
81. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J Biol Chem **250**:4007-4021.
82. **Ögrünc, M., D. F. Becker, S. W. Ragsdale, and A. Sancar.** 1998. Nucleotide excision repair in the third kingdom. J Bacteriol **180**:5796-5798.
83. **Ohkuma, Y., and R. G. Roeder.** 1994. Regulation of TFIIF ATPase and kinase activities by TFIIE during active initiation complex formation. Nature **368**:160-163.
84. **Ostapenko, D., and O. Gileadi.** 2000. Rad25p, a DNA helicase subunit of yeast transcription factor TFIIF, is required for promoter escape *in vivo*. Gene **245**:109-117.

85. **Ouhammouch, M., R. E. Dewhurst, W. Hausner, M. Thomm, and E. P. Geiduschek.** 2003. Activation of archaeal transcription by recruitment of the TATA-binding protein. *Proc Natl Acad Sci USA* **100**:5097-5102.
86. **Ouhammouch, M., and E. P. Geiduschek.** 2005. An expanding family of archaeal transcriptional activators. *Proc Natl Acad Sci USA* **102**:15423-15428.
87. **Ouhammouch, M., and E. P. Geiduschek.** 2001. A thermostable platform for transcriptional regulation: the DNA-binding properties of two Lrp homologs from the hyperthermophilic archaeon *Methanococcus jannaschii*. *Embo Journal* **20**:146-156.
88. **Park, E., S. N. Gudzer, M. H. Koken, J. I. Jaspers, G. Weeda, J. H. Hoeijmakers, S. Prakash, and L. Prakash.** 1992. RAD25 (SSL2), the yeast homolog of the human xeroderma pigmentosum group B DNA repair gene, is essential for viability. *Proc Natl Acad Sci* **89**:11416-11420.
89. **Pause, A., N. Methot, and N. Sonenberg.** 1993. The HRIGRXXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. *Mol Cell Biol* **13**:6789-6798.
90. **Peterson, J. D., L. A. Umayam, T. Dickinson, E. K. Hickey, and O. White.** 2001. The Comprehensive Microbial Resource. *Nucleic Acids Res* **29**:123-125.
91. **Preston, C. M., K. Y. Wu, T. F. Molinski, and E. F. DeLong.** 1996. A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc Natl Acad Sci USA* **93**:6241-6246.

92. **Rigaut, G., A Shevchenko, B Rutz, M Wilm, M Mann and B Séraphin.** 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnol* **17**:1030-1032.
93. **Rother, M., and W. W. Metcalf.** 2004. Anaerobic growth of *Methanosarcina acetivorans* C2A on carbon monoxide: An unusual way of life for a methanogenic archaeon. *Proc Natl Acad of Sci USA* **101**:16929-16934.
94. **Rouviere, P. E. W., R.S.** 1988. Novel biochemistry of methanogenesis. **263**:7913-7916.
95. **Rowlands, T., P. Baumann, and S. P. Jackson.** 1994. The TATA-binding protein: a general transcription factor in eukaryotes and archaebacteria. *Science* **264**:1326-1329.
96. **Sambrook, J., and E. F. M. Fritsch, T.** 1989. p. 2.60 - 2.81. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
97. **Santangelo, T. J., and J. N. Reeve.** 2006. Archaeal RNA Polymerase is Sensitive to Intrinsic Termination Directed by Transcribed and Remote Sequences. *J Mol Biol* **355**:196-210.
98. **Sauer, K., U. Harms, and R. K. Thauer.** 1997. Methanol:coenzyme M methyltransferase from *Methanosarcina barkeri* - Purification, properties and encoding genes of the corrinoid protein MT1. *Eur J Biochem* **243**:670-677.
99. **Schmid, S. R., and P. Linder.** 1992. D-E-A-D protein family of putative RNA helicases. *Mol Microbiol* **6**:283-291.

100. **Schmittgen, T. D., and K. J. Livak.** 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**:1101-1108.
101. **Schug, J.** 2003. Using TESS to Predict Transcription Factor Binding Sites in DNA Sequence. *In* A. D. Baxevanis (ed.), *Current Protocols in Bioinformatics*. J. Wiley and Sons.
102. **Senanayake, S. D., and D. A. Brian.** 1995. Precise large deletions by the PCR-based overlap extension method. *Mol Biotechnol* **4**:13-15.
103. **Shimotsu, H., M. I. Kuroda, C. Yanofsky, and D. J. Henner.** 1986. Novel form of transcription attenuation regulates expression the *Bacillus subtilis* tryptophan operon. *J Bacteriol* **166**:461-471.
104. **Shockley, K. R., D. E. Ward, S. R. Chhabra, S. B. Connors, C. I. Montero, and R. M. Kelly.** 2003. Heat shock response by the hyperthermophilic archaeon *Pyrococcus furiosus*. *Appl Environ Microbiol* **69**:2365-2371.
105. **Simonson, A. B., J. A. Servin, R. G. Skophammer, C. W. Herbold, M. C. Rivera, and J. A. Lake.** 2005. Decoding the genomic tree of life. *Proc Natl Acad Sci USA* **102 Suppl 1**:6608-6613.
106. **Smith, D. R., L. A. DoucetteStamm, C. Deloughery, H. M. Lee, J. Dubois, T. Aldredge, R. Bashirzadeh, D. Blakely, R. Cook, K. Gilbert, D. Harrison, L. Hoang, P. Keagle, W. Lumm, B. Pothier, D. Y. Qiu, R. Spadafora, R. Vicaire, Y. Wang, J. Wierzbowski, R. Gibson, N. Jiwani, A. Caruso, D. Bush, H. Safer, D. Patwell, S. Prabhakar, S. McDougall, G. Shimer, A. Goyal, S. Pietrokovski, G. M. Church, C. J. Daniels, J. I. Mao,**

- P. Rice, J. Nolling, and J. N. Reeve.** 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: Functional analysis and comparative genomics. *J Bacteriol* **179**:7135-7155.
107. **Smith, M. R., and R. A. Mah.** 1978. Growth and methanogenesis by *Methanosarcina* strain 227 on acetate and methanol. *Appl Environ Microbiol* **36**:870-879.
108. **Song, J. M., B. A. Montelone, W. Siede, and E. C. Friedberg.** 1990. Effects of multiple yeast *rad3* mutant alleles on UV sensitivity, mutability, and mitotic recombination. *J Bacteriol* **172**:6620-6630.
109. **Sowers, K.** 2009. Methanogenesis, p. 265-286. *In* M. Schaechter (ed.), *Encyclopedia of Microbiology*, 3rd edition. Elsevier, Inc.
110. **Sowers, K. R.** 1995. Growth of *Methanosarcina* spp. as single cells, p. 61-62. *In* F. T. Robb, K. R. Sowers, S. DasSharma, A. R. Place, H. J. Schreier, and E. M. Fleischmann (ed.), *Archaea: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
111. **Sowers, K. R.** 2004. Methanogenesis, p. 659-679. *In* M. Schaechter (ed.), *The Desk Encyclopedia of Microbiology*. Elsevier Academic Press, San Diego.
112. **Sowers, K. R., J. E. Boone, and R. P. Gunsalus.** 1993. Disaggregation of *Methanosarcina* spp. and growth as single cells at elevated osmolarity. *Appl Environ Microbiol* **59**:3832-3839.
113. **Sowers, K. R., and R. P. Gunsalus.** 1988. Adaptation for growth at various saline concentrations by the archaeobacterium *Methanosarcina thermophila*. *J Bacteriol* **170**:998-1002.

114. **Sowers, K. R., and R. P. Gunsalus.** 1993. Transcriptional regulation of the carbon monoxide dehydrogenase gene (*cdhA*) in *Methanosarcina thermophila*. *J Biol Chem* **268**:23172-23178.
115. **Sowers, K. R., J. L. Johnson, and J. G. Ferry.** 1984. Phylogenetic relationships among the methylotrophic methane-producing bacteria and emendation of the family Methanosarcinaceae. *Int J Syst Bacteriol* **34**:444-450.
116. **Spangler, L., X. Wang, J. W. Conaway, R. C. Conaway, and A. Dvir.** 2001. TFIID action in transcription initiation and promoter escape requires distinct regions of downstream promoter DNA. *Proc Natl Acad Sci USA* **98**:5544-5549.
117. **Sudarsan, N., J. E. Barrick, and R. R. Breaker.** 2003. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* **9**:644-647.
118. **Sudarsan, N., J. K. Wickiser, S. Nakamura, M. S. Ebert, and R. R. Breaker.** 2003. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev* **17**:2688-2697.
119. **Sung, P., S. N. Guzder, L. Prakash, and S. Prakash.** 1996. Reconstitution of TFIID and requirement of its DNA helicase subunits, Rad3 and Rad25, in the incision step of nucleotide excision repair. *J Biol Chem* **271**:10821-10826.
120. **Svejstrup, J. Q., P. Vichi, and J. M. Egly.** 1996. The multiple roles of transcription/repair factor TFIID. *Trends Biochem Sci* **21**:346-350.
121. **Terlesky, K. C., M. J. K. Nelson, and J. G. Ferry.** 1986. Isolation of an enzyme complex with carbon monoxide dehydrogenase activity containing a

- corrinoid and nickel from acetate-grown *Methanosarcina thermophila*. *J Bacteriol* **168**:1053-1058.
122. **Thomm, M., and B. A. R. Sherf, J. N.** 1988. RNA polymerase-binding and transcription initiation sites upstream of the methyl reductase operon of *Methanococcus vannielii*. *J Bacteriol* **170**:1958-1961.
123. **Thompson, D., J. Palmer, and C. Daniels.** 1999. Expression and heat-responsive regulation of a TFIIB homologue from the archaeon *Haloferax volcanii*. *Mol Microbiol* **33**:1081-1092.
124. **Veit, K., C. Ehlers, A. Ehrenreich, K. Salmon, R. Hovey, R. P. Gunsalus, U. Deppenmeier, and R. A. Schmitz.** 2006. Global transcriptional analysis of *Methanosarcina mazei* strain Go1 under different nitrogen availabilities. *Mol Genet Genomics* **276**:41-55.
125. **Welander, P. V., and W. W. Metcalf.** 2005. Loss of the *mtr* operon in *Methanosarcina* blocks growth on methanol, but not methanogenesis, and reveals an unknown methanogenic pathway. *Proc Natl Acad Sci USA* **102**:10664-10669.
126. **Wettach, J., H. P. Gohl, H. Tschochner, and M. Thomm.** 1995. Functional interaction of yeast and human TATA-binding proteins with an archaeal RNA polymerase and promoter. *Proc Natl Acad Sci USA* **92**:472-476.
127. **Whitman, W. B., D. C. Coleman, and W. J. Wiebe.** 1998. Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* **95**:6578-6583.

128. **Winkler, W., A. Nahvi, and R. R. Breaker.** 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**:952-956.
129. **Winkler, W. C., A. Nahvi, N. Sudarsan, J. E. Barrick, and R. R. Breaker.** 2003. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat Struct Biol* **10**:701-707.
130. **Woese, C. R., and G. E. Fox.** 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci USA* **74**:5088-5090.
131. **Yanofsky, C.** 1981. Attenuation in the control of expression of bacterial operons. *Nature* **289**:751-758.
132. **Yanofsky, C.** 1988. Transcription attenuation. *J Biol Chem* **263**:609-612.
133. **Yarnell, W. S., and J. W. Roberts.** 1999. Mechanisms of intrinsic transcription termination and antitermination. *Sci* **284**:611-615.
134. **Zimmermann, P., and F. Pfeifer.** 2003. Regulation of the expression of gas vesicle genes in *Haloflex mediterranei*: interaction of the two regulatory proteins GvpD and GvpE. *Mol Microbiol* **49**:783-794.