INFORMATION TO USERS

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

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

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

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

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Studies on the *cad* operon of *Escherichia coli* K-12: A pH regulatory system in bacteria

Meng, Shi-Yuan, Ph.D.

Rice University, 1992
RICE UNIVERSITY

STUDIES ON THE CAD OPERON OF ESCHERICHIA COLI K-12: A pH REGULATORY SYSTEM IN BACTERIA

by

SHI-YUAN MENG

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE:

George N. Bennett
Associate Professor of Biochemistry and Cell Biology

Kathleen M. Beckingham
Associate Professor of Biochemistry and Cell Biology

John S. Olson
Professor of Biochemistry and Cell Biology

Micheal Stern
Assistant Professor of Biochemistry and Cell Biology

Ka-Yiu San
Associate Professor of Chemical Engineering

Houston, Texas
November, 1991
Studies on the cad Operon of Escherichia coli K-12: A pH Regulatory System in Bacteria

Shi-Yuan Meng

ABSTRACT

The expression of biodegradative lysine decarboxylase of E. coli is induced in cells grown at low pH, and the maximum induction was observed in the presence of excess lysine under anaerobic conditions. The cadA gene encoding the lysine decarboxylase resides in the cadBA operon. The sequence of about 5 kb of the cadBA operon (3.7 kb) and its flanking regions has been determined. The cadB gene encodes a protein very similar to the ArcD protein, the arginine:ornithine antiporter of Pseudomonas aeruginosa. Both cadA and cadB genes have been subcloned into different expression vectors and the protein products have been expressed and observed in maxicell experiments. The cadB gene has been found to facilitate the excretion of cadaverine, one of the end products of the enzymatic reaction of CadA (lysine decarboxylase), and therefore CadB probably encodes a lysine: cadaverine antiporter. A model for detoxification of extracellular high H⁺ concentration based on the functions of CadA and CadB is proposed. Studies on regulation of the cad operon were performed using lac fusion techniques. A high affinity pH-responsive site(s) was found to be located over 100 bp upstream from the cad promoter by a series of deletion and mutation experiments, and the involvement of a positive pH regulator was confirmed in protein titration experiments. Random mutagenesis and in vivo footprinting were conducted to determine the importance of specific residues in the cad upstream region. The residues detected in these studies of the cad promoter region suggest a complex regulatory mechanism. The possible routes of regulation by the availability of lysine and oxygen were also explored. Hypothetical models account for experimental results are presented.
ACKNOWLEDGMENTS

I would like to thank Dr. George Bennett for his advice and guidance throughout this work. I am grateful to my parents for their encouragement and support. I cherish the life shared with my friends at Rice. This research was supported by a grant to Dr. G. N. Bennett from the National Institute of Health (G.M. 39420).
Table of Contents

CHAPTER 1. INTRODUCTION

pH Responses of Bacteria ......................................................................................... 1
  pH limits for the growth of Escherichia coli ......................................................... 1
  Intracellular pH variations .................................................................................... 2
  Membrane energetics ............................................................................................. 3
  Possible components of pH homeostatic system .................................................... 4
  Possible mechanisms for pH homeostasis of neutrophilic bacteria ....................... 5
    Alkalization of cytoplasm .................................................................................... 5
    Acidification of cytoplasm .................................................................................. 6
  pH responses observed in other systems ............................................................... 6
    Anaerobiosis ......................................................................................................... 6
    Chemotaxis ......................................................................................................... 7
    Heat shock response ............................................................................................ 7
    Osmolarity .......................................................................................................... 7
    SOS response ...................................................................................................... 8
    Spore germination ............................................................................................... 8
    Starvation ............................................................................................................ 8
  Genes involved in pH regulation in bacteria ......................................................... 9
    Induction by acid pH ........................................................................................... 9
    Induction by alkaline pH .................................................................................... 12
  Bacterial virulence and pH ................................................................................... 12

Global Regulation in Bacteria ................................................................................ 14
  Definitions .............................................................................................................. 14
  Signal transduction in bacteria ......................................................................... 15
Sensor-regulator two-component regulatory systems...........................15
Signal transduction by phosphorylation...........................................15

Significance and Application of pH Regulation..................................22

CHAPTER 2. MATERIALS AND METHODS

Materials..........................................................................................24
  Bacterial strains.................................................................24
  Plasmid DNA.................................................................24
  Bacterial phages...............................................................24
  Bacterial media.................................................................24
  Oligonucleotide primers.......................................................25
  Antibiotics and chemicals used for selection of strains...............25
  Enzymes.................................................................................25
  Reaction kits...........................................................................26
  Radioactive labels...............................................................26

Methods.........................................................................................38
  Plasmid isolation...............................................................38
    Birnboim procedure (mini-prep)..........................................38
    QIAGEN plasmid purification (midi-prep)..........................39
    Isolation of methylated plasmid DNA.................................39
  Isolation of chromosomal DNA...............................................40
  Gel electrophoresis............................................................40
  Recombinant DNA techniques...............................................40
  Southern hybridization.........................................................41
  Purification of synthetic oligonucleotides...............................41
CHAPTER 3. STRUCTURE AND PHYSIOLOGICAL ROLE OF THE cad OPERON

Identification of the Inducible Lysine Decarboxylase Coding Gene cadA

INTRODUCTION

Amino acid decarboxylases

Biodegradative and biosynthetic amino acid decarboxylases

Lysine decarboxylase coding gene cadA
RESULTS........................................................................................................51
  Localization of cadA gene........................................................................51
  Subcloning and expression of of cadA gene..............................................64
  Amino acid sequence of biodegradative lysine decarboxylase...............71
  Comparison of the amino acid sequences of the inducible lysine decarboxylase with other bacterial amino acid decarboxylases..........................71
  Lysine decarboxylases of H. alvei and E. coli...........................................71
  Comparison of biodegradative LDC and biosynthetic ODC of E. coli.....79

DISCUSSION..................................................................................................79

Characterization of the cadB Gene.................................................................85

INTRODUCTION...............................................................................................85
  Membrane associated proteins...............................................................85
  Membrane protein topology.................................................................85
  Alkaline phosphatase (PhoA) fusions.......................................................87
  arc operon of Pseudomonas aeruginosa....................................................87

RESULTS AND DISCUSSION.........................................................................90
  Identification of cadB gene .................................................................90
  Subcloning of cadB gene.........................................................................90
  Expression of cadB with maxicells........................................................95
  Confirmation of CadB sequence...............................................................98
  Homology between CadB of E. coli and ArcD of P. aeruginosa..............98
  Secondary structure of CadB....................................................................98
Physiological Role of \textit{cad} Operon .................................................. 108

INTRODUCTION .................................................................................. 108
  Basic amino acid transport systems ........................................... 108
  Arginine and ornithine transport ............................................. 108
  Lysine transport ........................................................................ 109
  Basic amino acid transport in other bacteria ............................. 111
  Polyamine content of \textit{E. coli} .................................................. 112
  Arginine: ornithine antiport in \textit{Streptococcus lactis} .................. 115

RESULTS ......................................................................................... 116
  Cadaverine excretion ............................................................... 116

DISCUSSION .................................................................................... 117

Organization of the \textit{cad} Operon and its Neighboring Genes ......... 129

INTRODUCTION ................................................................................ 129
  Genomic mapping of \textit{E. coli} .................................................. 129
  Genes located in the 93.2 to 94 min on \textit{E. coli} genome ............ 129

RESULTS AND DISCUSSION ......................................................... 130

CHAPTER 4. REGULATION OF THE \textit{cad} OPERON

INTRODUCTION ................................................................................ 133
  Selected topics of bacterial virulence ....................................... 134
    \textit{Salmonella typhimurium} .................................................. 134
    \textit{Vibrio cholerae} ............................................................... 135
    \textit{Agrobacterium tumefaciens} ............................................ 137
  Regulation of \textit{cad} operon of \textit{E. coli} .................................... 143
RESULTS.............................................................................................................143
Construction of operon and protein lac fusions.................................144
β-Galactosidase activities of lac fusion constructs..............................144
Differential expression of cadB and cadA genes.................................151
PCR random mutagenesis of fragment (primer+)-(R-).........................151
The DNA segment between the SspI and EcoRV sites is essential for pH induction of the cad operon..................................................151
Deletion and mutation analysis of the octonucleotide palindrome sequence between SspI and EcoRV..................................................153
Protein titration studies.................................................................158
Titration of pH-responsive regulator(s)..............................................158
Titration of the pH-responsive activator(s) with deleted and mutated fragments.................................................................158
PCR mutagenesis.............................................................................161
Protein titration with PCR-mutagenized fragments..........................165
Lysine effect.................................................................................165
In vivo footprinting.........................................................................175
Oxygen effect on cad promoter activated by CadC............................184

DISCUSSION...............................................................................................184

CHAPTER 5. POSSIBLE WORKING MODELS FOR REGULATION OF THE cad OPERON

INTRODUCTION..............................................................................................197
Transcription initiation........................................................................197
Homology of sigma factors...............................................................198
Sigma factors of E. coli.......................................................................200
Differences between sigma 70 and sigma 54 promoters....................202
Nucleoproteins ................................................................. 204
DNA bending ........................................................................ 206
DNA supercoiling for the regulation of gene expression .......... 207
Post-transcriptional control and regulation ......................... 208
  Translation initiation determined by mRNA sequence and structure ... 208
  Translational regulatory systems ........................................ 209
  mRNA stability ............................................................... 210

SPECULATIONS .................................................................. 211
  Sigma 70 or sigma 54 promoter ........................................ 212
  Looping, wrapping, or domino effect ................................. 212
  The role of cadC, cadR, and osmZ in regulation of cad operon .... 213

REFERENCES ..................................................................... 226
List of Figures

1. Model for signal transduction in two-component regulatory systems..................17
2. Two families of homologous signal transduction proteins in bacteria...............19
3. Chromosomal regions cloned in pLC4-5 and pKER65........................................52
4. Southern hybridization mapping of cad:: lac fusion strains.............................54
5. Nucleotide and the deduced amino acid sequences of cadBA operon..................56
6. Construction of plasmid pCADA.......................................................................65
7. Protein expression of lysine decarboxylase from pCADA.................................67
8. Protein expression of cadA from pCADA with maxicells.................................69
9. Hydropathy profiles of lysine decarboxylases of E. coli and H. alvei................73
10. Garnier and Robson secondary structure analysis of lysine decarboxylase and ornithine decarboxylase.................................................................75
11. Sequence alignment of lysine decarboxylase of H. alvei and biosynthetic ornithine decarboxylase of E. coli to biodegradative lysine decarboxylase of E. coli.................................................................77
12. The arc operon and the arginine deiminase pathway of p. aeruginosa..............88
13. Amino acid composition of CadB.......................................................................91
14. Construction of pCADB..................................................................................93
15. Protein expression of cadB from pCADB with maxicells.................................96
16. Hydrophobicity profiles of CadB of E. coli and ArcD of P. aeruginosa.............99
17. Amino acid sequence homology between CadB of E. coli and ArcD of P. aeruginosa.................................................................101
18. Garnier and Robson secondary structure analyses of CadB and ArcD.............103
19. Fusion positions and the corresponding β-galactosidase activities of cad::lacZ protein fusions.................................................................106
20. Schematic representation of a lysine: arginine shuttle in L. lactis subsp. lactis which leads to accumulation of arginine.........................113
21. Model for detoxification of extracellular high H+ concentration by the CadB-CadA system.................................................................125
22. Chromosomal localization of *cad* operon and its neighboring genes..................131
23. Model for the interaction between ToxR and ToxS .............................................138
24. Similarity between *virG* promoters and *E. coli* promoters induced by phosphate starvation (A) or by heat shock stresses (B)..............................141
26. Construction of cloning vector pRS552-1 and pXA-1..........................................147
27. Operon and protein fusion constructs of the *cad* upstream regulatory region...147
28. Deletion and mutation of the *SspI* to *EcoRV* region by PCR.................................154
29. Base pair changes in primers SsRV2 to SsRV6.....................................................156
30. Effect of multicy copy plasmid-borne *cad* upstream regions on expression from the chromosomal pH-responsive *cad* promoter...........................................159
31. Base pair changes in PCR mutagenized fragments cloned in pTMU plasmids..162
32. Effect of single base pair changes in the -161 to -96 region.................................166
33. Lysine effect on pH induction of *cad* promoter..................................................168
34. Effect of multicy copy plasmid-borne *cad* upstream regions on the chromosomal lysine-responsive *cad* promoter activity in GNB8385K.......171
35. Effect of mutated *cad* upstream regions on pTMU plasmids on the chromosomal lysine-responsive *cad* promoter activity in GNB8385K.......173
37. Effect of *cadC* gene copy number on the pH induction of the *cad* ::*lac* fusions..........................................................178
38. Effect of multicy copy *cadC* gene on the single copy chromosomal *cad* promoter......................................................................................................................180
39. *In vivo* methylation of pH-responsive regulatory region of *cad* operon........182
40. Important DNA residues involved in pH induction of *cad* operon....................185
41. Oxygen effect on *cad* promoter activated by CadC.............................................187
42. Structure and sequence homology between *cad* and *adi* upstream regions.....194
43. pH effect on the expression of *cad* operon.......................................................218
43.2. Lysine effect on the expression of the cad operon..........................220
43.3. Oxygen effect on the expression of the cad operon at low pH............222
43.4. OsmZ effect on the expression of the cad operon..........................224
List of Tables

1. Bacterial strains.................................................................27
2. Cloning vectors.......................................................................28
3. Plasmid constructs.................................................................29
4. Ingredients of CadR medium...................................................32
5. Oligonucleotide primers..........................................................33
6. Amino acid composition of lysine decarboxylases of *E. coli* and *H. alvei*......72
7. Comparison of the physical and chemical properties of arginine-, lysine-, and ornithine decarboxylases of *E. coli*...........................................................81
8. Comparison of amino acid sequences of the pyridoxal 5'-phosphate binding site of amino acid decarboxylases in bacteria.........................................................84
9. Cadaverine excretion assay I....................................................118
10. Cadaverine excretion assay II..................................................119
11. The *cadR* effect on cadaverine excretion assay I....................120
12. The *cadR* effect on cadaverine excretion assay II...................121
13. pH responses of *lac* fusions to mutagenized *cad* upstream regulatory region...152
14. Concensus promoter sequences recognized by sigma factors............199
15. Properties of polII and σ^{54} pol, not shared by σ^{70} pol..................203
CHAPTER 1. INTRODUCTION

pH Responses of Bacteria

Cytoplasmic pH is an important parameter in bacterial cell physiology. Most enzymatic reactions are active only within a narrow pH range. Cells have to maintain an appropriate internal pH for essential functions and therefore tight regulation of pH homeostasis is present universally in bacteria. Bacterial cells are in direct contact with the outer environment where the pH condition can suddenly go through a drastic change. How bacterial cells sense the pH shift of the ever-changing environment and conduct proper responses in order to survive and prosper are the principle concerns of this research.

In nature bacteria are distributed over a wide range of habitats. They have been found in acidic sulfur springs with proton concentrations of pH 1 to 2 and in soda lakes of pH 11 (Lanworthy, 1978). The majority of bacteria fall into the category of neutrophiles which grow over the range from pH 5 to 9, whereas the acidophiles and alkalophiles can grow under more extreme pH conditions. Although a wide range of extracellular proton concentrations are encountered, the intracellular proton concentrations of bacteria in each category are kept almost constant with small fluctuations. Acidophiles exhibit pH\textsubscript{i} values in the range of 6.5 to 7.0, neutrophiles have pH\textsubscript{i} (internal pH) values of 7.5 to 8.0, and alkalophiles have pH\textsubscript{i} values of 8.4 to 9.0 (Booth, 1985).

pH limits for the growth of Escherichia coli

*E. coli* is a representative of the neutrophiles; it grows optimally between pH\textsubscript{o} (external pH) 6 and 8 and can grow more slowly at pH\textsubscript{o} values of a unit or so beyond these limits. The pH\textsubscript{i} of *E. coli* remains constant at pH 7.6 to pH 7.8 (Padan et al., 1981). When the pH\textsubscript{i} value of *E. coli* was artificially titrated with weak acid to 7.2,
growth rate was decreased by a factor of two; growth ceased completely when pH_{i} dropped to below 6.6 to 6.8 (Booth, 1985). *E. coli* is highly intolerant to perturbations that causes pH_{i} to become more alkaline than pH 8 (Zilberstein et al., 1981,1984). The effect of pH_{i} on bacterial growth has been applied in preservation of food from microbial spoilage. The inhibition of bacterial growth by weakly acidic food preservatives appears to be a consequence of the inhibition of metabolism by lowering pH_{i} (Salmond et al., 1984). *E. coli cya* and *crp* mutants in glucose-minimal medium grew more slowly than their parents, and this difference was intensified as the external pH increased from pH 6 to pH 7.8 while the parents were comparatively unaffected by external pH. This pH sensitivity could be overcome by adding cAMP to the medium (Ahmad and Newman, 1988)

**Intracellular pH variations**

Although different organisms exhibit different capacities to regulate their cytoplasmic pH, in many organisms the pH_{i} values vary by only 0.1 units per pH unit change in external pH (Booth, 1985). The perturbations of pH_{i} caused by large shift of pH_{o} are generally small due to the intrinsic impermeability of the lipid bilayer (Raven and Beardall, 1980) and very specific control of ion flux through the membrane. However, the addition of weak acids or bases to cell suspensions can lead to permanent changes in pH_{i} (Slonczewski et al., 1982), and cells depleted of potassium are sensitive to extracellular pH perturbation (Kroll and Booth, 1983). On the other hand, proton consumption and acid metabolite production by essential metabolic processes are the major sources for perturbations of internal pH. This problem may be so acute that some organisms (e.g. *Clostridium acetobutylicum*) evolved mechanisms to convert acidic products to neutral compounds (Bahl et al., 1982). In *E. coli*, some decarboxylases and deaminases are induced respectively by low and high pH media (Gale and Epps, 1942), the enzymatic reactions and end products can cause pH_{i} problems if they are not properly
balanced. The production of polyamines and polycarboxylates by the reaction of amino acid decarboxylases and deaminases was suggested as being part of an effort to shift the pH of the medium back to neutrality (Gale and Epps, 1942)

**Membrane energetics**

Energy transduction of bacterial cells largely depends on the circulation of protons across the membrane (Harold, 1986; Maloney, 1987). By electrogenic transport of protons outward, cells are able to maintain an electrochemical gradient of protons, Δp, across the membrane. Through various molecular devices, protons are allowed to return to the cytoplasm while performing useful work (e.g. active transport) at the expense of the proton motive force (PMF). Cells need to maintain a transmembrane PMF which is constituted of a membrane electrochemical potential (Δ φ) and a pH gradient (ΔpH) for various functions. \[ \text{PMF} = \Delta p = \Delta \phi - K \Delta \text{pH} \]. \[ K = 2.3 \frac{RT}{F} \] (R, gas constant; T, absolute temperature; F, Faraday constant). Change in one of the components is often accompanied by a compensatory change in the other, so that the PMF supply remains constant. For the prokaryotic cell, the magnitude of the pH gradient is a function of the external pH, \[ \Delta \text{pH} = \text{pH}_1 - \text{pH}_0 \], and that is reflected by a relatively constant internal pH (Meury and Kepes, 1982). A mechanism for pH homeostasis based on the modulation of the ΔpH and Δφ created by the primary proton pumps of the cell, that is, electron transport-linked pumps and/or proton translocating ATPase, appears to be common to many bacterial species (Kobayashi, 1985; Kobayashi et al., 1984). The central role of H⁺ in energized membrane is suggested as an intermediate between the energy-producing and energy-consumption reactions, and H⁺ movements probably dominate bacterial membrane transport (Maloney, 1987). In spite of the intricate relationship of H⁺ potential with many known or unknown systems, it is possible to study its influence on a particular reaction by dissipating the proton gradient across the membrane. This can be achieved by short circuiting the proton current with H⁺-ionophores (e.g. gramicidin). In
a series of elegantly designed experiments, several roles of the proton gradient were
determined including cellulose synthesis in *Acetobacter xylinum*, penetration of DNA into
bacterial cells, and regulation of cytoplasmic pH (Padan and Schuldiner, 1987). If ATP
is not supplied by glycosylation or other substrate level phosphorylation, it is essential to
drive synthesis of ATP by energized membrane. And it is also necessary for the uptake
of nutrients, establishment of K\(^+\) and Na\(^+\) gradients across the membrane, and
maintaining intracellular pH homeostasis. However, there are specific reactions in
different bacteria somehow modulated by the proton gradient or membrane potential in a
manner unrelated to their energetic role. The uses of pH\(_i\) and membrane potential as
regulators in prokaryotic cells are reviewed by Padan and Schuldiner (1987). The
methods and concepts on energy coupling and the role of alkaline cation transport in
neutrophilic and acidophilic bacteria are assessed by Bakker (1990).

**Possible components of pH homeostatic system**

The potential candidates for pH homeostatic system in bacteria include:

**Cytoplasmic buffers.** The principle components of the cytoplasmic buffer are the
amino acid side chains of proteins (Sanders and Slayman, 1982). The cytoplasmic
buffering capacity of several bacterial organisms pointed to values of around 50 to 100
nmol of H\(^+\) per pH unit per mg of cell protein at around pH\(_i\) 7, and the buffering capacity
increased at both acid and alkaline extremes of pH\(_i\) (Booth, 1985). However,
cytoplasmic buffers can offset only a limited amount of acidification or alkalization,
which can be eventually outstripped by continuous pH stress. **Production of acids
and bases.** Production of acids by fermentation processes, and the production of
polyamines and polycarboxylates by decarboxylases and deaminases reactions can pose a
serious problems to cells if they are not properly dealt with. The synthesis of some
decarboxylases and deaminases is known to be affected by pH of the culture medium
(Gale and Epps, 1942), but this has not been shown to be directly involved in pH\(_i\)
control. **Active transport of H⁺ (or OH⁻).** The controlled transport of H⁺ is considered to be the dominant mechanism of control of pHi. However, the excess membrane potential generated by the extrusion of large quantities of protons by the primary ATP-driven proton pump has to be dissipated in order to continue the proton pumping process for alkalization of pHi. This is done by coupling either cation influx or anion efflux. For acidification of the cytoplasm, similar cation-proton antiports have been suggested to play a dominant role (Rosen, 1986). In such an antiport, the internal cations (Na⁺ or K⁺) are exchanged for external protons resulting in acidification of pHi. However, the internal cation concentration may become too low for this type of exchange, therefore cycling of cations has to be considered. A role for anion-hydroxyl ion exchange in pH homeostasis has not been found in bacteria, but it is possible that such an antiporter exists in bacteria.

**Possible mechanisms for pH homeostasis of neutrophilic bacteria**

**Alkalization of cytoplasm.** Since potassium is the major cellular cation, uptake of this ion through a potassium: H⁺ antiport may well be a major factor acting to increase pHi (Koyama and Nosoh, 1985). Within certain limits there was a direct relationship between the potassium uptake and the pH gradient generated in *E. coli*. When the alkalization of the internal pH reached to pH 8.1, the relationship broke down. At this point a mechanism for acidification of the cytoplasm overcame the tendency for pHi to rise and the pHi fell back to 7.6 (Kroll and Booth, 1983). However, the potassium transport in generating pH gradient is not directly under the control of pHi. Other factors affecting potassium transport, e.g., an increase in osmotic pressure, might transiently perturb pHi. Higher membrane proton-translocating ATPase activity was found in *Streptococcus faecalis* cells grown at acidic pH₀ than cells grown at alkaline pH₀, and this enhancement was indicated at the level of protein synthesis (Abrams and Jensen, 1984).
**Acidification of cytoplasm.** The mechanism of the acidification of the cytoplasm is unclear, but it is sensitive to both pH$_i$ and pH$_o$ and might be mediated via the magnitude of ΔpH (Kroll and Booth, 1983). Proton uptake could be performed through $K^+$: $H^+$ and $Na^+$: $H^+$ antiports. At alkaline pH$_o$, the generation of pH gradient, inside acid (normally cells maintain alkaline inside at acid to neutral pH$_o$), may depend on the presence of a potassium gradient, directed outward (Nakamura et al., 1984). In membrane vesicles the antiport has been reported to be electroneutral at acidic pH and electrogenic (2$H^+$ per $Na^+$) at alkaline pH. The flux of $Na^+$ through the $Na^+$: $H^+$ antiport could also account for the decrease in the pH gradient as the external pH was increased (Padan, et al., 1976; Ishikawa et al., 1987).

**pH responses observed in other systems**

**Anaerobiosis.** In view of the external acidification which accompanies anaerobic growth in bacteria and the increased reliance on the ΔpH as a source of proton motive force, it would not be surprising to find some overlap between genes induced by low pH and by anaerobiosis. In the absence of oxygen, the energy has to be supplied either by fermentation or by the anaerobic respiration of alternative electron acceptors such as nitrate, fumarate, trimethylamine oxide, or dimethyl sulfoxide (Winkelman and Clark, 1986). Anaerobic respiratory genes are repressed by oxygen and require the appropriate electron acceptor as an inducer. In many but not all cases, the *fnr* gene which encodes an activator is necessary for the expression of anaerobic respiratory genes (Winkelman and Clark, 1986). So far, pH-induced genes do not appear to be under the control of Fnr protein. The implication of anaerobiosis in pH regulation is not clear, but the *aniG* locus of *S. typhimurium* as well as the *E. coli* *adi* and *cadA* genes have pointed out to have some interplay between anaerobiosis and acid regulation (Aliabadi et al., 1988; Auger et al. 1989).
**Chemotaxis.** Acid conditions elicit a repellent chemotactic response (Kihara and Macnab, 1981), whereas mildly basic conditions elicit an attractant response (Repaske and Adler, 1981). The change of swimming behavior is achieved through the control of the methylation of the transmembrane methylation-dependent protein MCPI (Tsr). This protein can sense both the external and internal pH. Demethylation occurs in response to acidification of the cytoplasm or the medium, and become remethylated when the pH homeostasis is restored (Slonczewski et al., 1982; Nettleton and Ordal, 1989).

**Heat shock response.** The induction of the synthesis of heat shock proteins by alkaline pH shift has been observed in *E. coli* (Taglicht et al., 1987). A rapid increase in membrane potential was detected immediately upon an alkaline pH$_0$ shift, the same effect induced by heat shock (Zilberstein et al., 1984). Overlap of heat shock proteins with acid tolerance proteins were also observed (Hyde and Portalier, 1990; Foster, 1992). Some heat shock proteins are known to be able to degrade or help refold denatured proteins (Gottesman, 1989). It is not clear whether a heat shock response to pH variations is induced by damaged proteins under adverse conditions.

**Osmolarity.** The relative amount of the two major outer membrane porin proteins, OmpF and OmpC, are influenced by many environmental factors such as osmolarity, ionic strength, composition of the culture medium, presence of local anaesthetics such as procaine, carbon source and growth temperature (Lugtenberg et al, 1976; Kawaji et al., 1979). Among them, osmolarity is the major factor. It has been shown that the fluctuation in porin ratios can also be induced by varying the pH of the growth medium and that these effects are at least partly controlled by the EnvZ protein (Heyde and Portalier, 1987). Growth at acid pH resulted in a slight increase in protein OmpC and a more prominent decrease in OmpF and LamB; the effect of acid pH on the amounts of OmpF and LamB was both observed when cells were grown in LB or Tris
minimal medium. The pH-dependent expression of *ompC* and *ompF* is regulated at a transcriptional level and also at a post-transcriptional level for *ompF*, and both levels of regulation are *envZ*-dependent. In *envZ*" strain, *ompF* gene expression was stimulated, but not repressed, by lowering the pH, suggesting that EnvZ mediated a negative pH control at low pH and the existence of an *envZ*-independent positive regulatory mechanism operating at acid pH on *ompF* gene expression (Heyde and Portalier, 1987).

**SOS response.** Failure of pH homeostasis is accompanied by cessation of cell division and start of cell filamentation. Inhibition of cell division leading to filamentation is part of the pleiotropic effect of the SOS response to DNA damage (Walker, 1984). Perturbation of pH$_i$ at alkaline pH$_o$ may trigger the SOS system in *E. coli* (Schuldiner et al., 1986). The effect may be mediated secondarily by changes such as damage to the DNA caused by alkaline pH$_i$, or it can also be more direct via a highly pH-sensitive protein such as *lexA* product. This protein exhibits autodigestion activity upon shift to alkaline pH (Little, 1984).

**Spore Germination.** It has been suggested that the low pH$_i$ of the dormant spore contributes to its metabolic dormancy and resistance. Upon the germination, an increase of pH$_i$ was observed (Moir, 1990).

**Starvation.** Some acid tolerant mutants are also auxotrophs suggesting starvation could render protection from external acid conditions (Foster and Hall, 1991). This can be explained by a partial cross protection to acid as a result of starvation for a specific nutrient. It is known that starvation for carbon source can lead to general protection against other stresses such as heat shock (Matin, 1991).
Genes involved in pH regulation in bacteria

**Induction by acid pH.** Two classes of pH-stimulons have been isolated from *E. coli* by *lac* fusion technique: *exa*, inducible by external acidification but not by membrane-permeable weak acids, and *ina*, inducible by internal acidification (mediated by weak acids) but not by external acidification (Slonczewski et al., 1987). The two *exa* strains behaved similarly with respect to pH range and maximum, time course, and increased induction under anaerobic conditions. Significant degree of induction was also observed to result from the addition of NaCl instead of low-pH buffer. However, this osmo-induction may be only transient since NaCl concentration did not cause significant difference of the steady-state induction of *exa-1* and *exa-2*. These two locus differ in that *exa-1* was more strongly induced than *exa-2* at low pH and more strongly repressed at neutral to high pH. The two *ina* strains, *ina-1* and *ina-2*, appeared identical with respect to pH range of induction, concentration range of weak acid response, and neither anaerobiosis nor osmolarity had any significant effect.

In *Salmonella typhimurium*, two anaerobiosis-induced genes, *hyd* and *aniG* (maps to 64 min), are found to be also regulated by alterations in external pH (Aliabadi et al., 1988). They differed in that *hyd* only responded to a pH stimulus if it was placed in an anaerobic situation, whereas *aniG* responded under both aerobic and anaerobic environments. The *earA* (external acidification regulator) mutation (maps to 86 min) not only relinquished *aniG* from pH control but also eliminated the complex medium requirement; studies on *aniG* mutant suggested that acid regulation was the primary function of EarA. However EarA did not affect oxygen regulation. The cloning of *earA* has revealed a second gene *earB*, immediately upstream from *earA*. EarA was detected in maxicells as a 33 Kd membrane-associated protein while the *earB* product was not observed in the same experiment. Multicopy of *earB* gene product conferred a partial
earA⁻ phenotype, indicating that EarB either titrated EarA repressor or acted directly as a positive regulator (Foster and Aliabadi, 1989). The other aniA (hyd) locus was unaffected by earA. The aniG locus is independent of both oxA and oxA (two positive regulatory genes in anaerobic regulation equivalent to fnr in E. coli); its anaerobic expression requires complex medium and is glucose repressed. Later, the aniG locus was determined to be exogenously coinduced by acid and mannose. It is also shown that the expression of aniG requires cAMP and CRP but only in the presence of EarA. It indicated that EarA is a repressor protein that binds to the operator region of aniG under alkaline growth conditions. Full derepression requires anaerobiosis as well as an acid environment, mannose and CRP/cAMP. earA itself is not regulated by pH or anaerobiosis. Two proteins were visualized on two-dimensional PAGE in aniG⁺ear mutants but not in ΔaniG ear mutant. The more acidic one was in the periplasmic fractions, while the truncated protein is found in the cytoplasmic fractions (Foster and Aliabadi, 1989).

The adaptive acidification tolerance response (ATR) was established in S. typhimurium (Foster and Hall, 1990). Logarithmically grown cells (pH 7.6) shifted to mild acid (pH 5.8) for one doubling time as an adaptive procedure were 100 to 1000 times more resistant to subsequent strong acid challenge (pH 3.3) than were unadapted cells shifted directly from pH 7.6 to 3.3. The ATR requires protein synthesis. Cells adapted in glucose minimal medium buffered with significant amount of citrate in the lower pH ranges were analyzed on two-dimensional PAGE. The ATR polypeptides (12 proteins were observed to increase while 6 decreases during the ATR protocol) had only minor overlap with other stress-induced proteins by two-dimensional PAGE analysis. No cross protection was noted for hydrogen peroxide, SOS, or heat shock.
The defensin-sensitive \textit{phoP} mutants of \textit{S. typhimurium} were also sensitive to external acidification, implying a correlation between acid tolerance and virulence (Fields et al., 1989). The \textit{unc (atp)} (H$^+$-translocating ATPase) mutants were also extremely acid sensitive, suggesting that the H$^+$-translocating ATPase is essential for the development of ATR (Kobayashi, 1985). ATPase activity from other organisms has also been shown to increase in response to a decrease in pH$_i$ (Kobayashi et al., 1984). Recently the \textit{fur} gene product (ferric uptake regulator) has been implicated as a major regulator (Foster and Hall, 1991). Mutations in \textit{fur} confer Atr$^-$ and acid-sensitive phenotypes as well as prevent the expression of several acid-regulated genes. The role of Fur in this system appears to be independent of iron, since ATR is unaffected by iron availability.

A two-phase working model has been envisioned to explain how cells respond to acid stress. The first phase, preshock, happens as the environmental pH approaches 5.8, when the cell will induce the ATR-associated pH homeostasis system. This system will operate at pH$_o$ values below 4. During severe acid stress, this mechanism will maintain the pH$_i$ near 5.5 and minimize acid denaturation of internal proteins. The second phase occurs when the pH$_o$ drops to between pH 5 and 3. A completely different set of proteins are induced during this type of acid shock, including several of the heat shock proteins classified as chaperonins (Heyde and Portalier, 1990; Foster et al., 1991).

Acid shock proteins described by Heyde and Portalier (1990) are induced by an external acid shift from pH 6.9 to pH 4.3. Cells were grown in rich MOPS medium. At least sixteen proteins are induced, including four heat shock proteins GroE, DnaK, HtpG and HtpM. Their pH induction was RpoH-dependent. Three other pH-induced proteins were identified as stress proteins induced either by osmolarity, aerobiosis, or low temperature. Seven proteins specifically induced by acid shift were called \textit{acid shock}
proteins (ASP). There are both RpoH-dependent and independent induction, and these ASP are not membrane but soluble proteins.

**Induction by alkaline pH.** Much less is known about alkaline pH induced genes than acid pH induced gene. A sodium-proton antiporter may be required for pH homeostasis in alkaline media (Ishikawa et al., 1987). It has been known that an alkaline pH shift induces the SOS response, via alkalinization of the cytoplasm (Schuldiner et al., 1986). Increase in pH\textsubscript{i} because of failure of pH\textsubscript{i} homeostasis induces the resistance to killing by UV irradiation and the expression of uvrA promoter, which are part of the pleiotropic effect of the SOS response to DNA damage. Some heat shock genes showed transient induction at a high external pH (Taglicht et al., 1987). The kinetics of the alkaline pH induction resembled that induced by heat shock, and a dependence on the rpoH gene was observed. The first gene in *E. coli* showing induction by alkaline pH\textsubscript{0} shift (pH 6.5 to 8.5) is called \textit{alx} which maps to 67.5 min (Bingham et al, 1990).

**Bacterial virulence and pH**

In order to infect host cells and cause diseases, microbial pathogens have to pass several mammalian host defense systems to secure their establishment and subsequent transmission. There are multilayers of barriers to overcome especially for pathogens following the fecal-oral route. Not only the organism must be able to tolerate life outside the host for variable periods of time, it has to survive the harsh environment encountered in the digestive track once ingested. These include higher temperatures, extremes of pH, different available nutrients, digestive enzymes, high concentrations of bile salts, peristalsis, etc (Finlay, and Falkow, 1989a). After attachment to the epithelial cells of the intestine, colonization and finally invasion and penetration through deeper tissues, the pathogenic bacteria face the even more formidable defense mechanism led by phagocytosis and subsequent acidification by fusion with lysosomes to form
phagolysosomes in macrophages (Finlay, and Falkow, 1989a). Most bacteria can hardly sustain such harsh treatment but pathogenic bacteria can. Apparently, the pH response plays an important role in bacterial pathogenicity. Studies on what kinds of survival mechanisms evolved in bacterial pathogens have proceeded to the genetic and molecular level. Recent progress in *Shigella spp* was reviewed by Hale (1991), and an earlier review on genetic regulation of bacterial virulence was published (DiRita and Mekalanos, 1989).
Global Regulation in Bacteria

In order to survive and prosper, bacterial cells are constantly adjusting themselves to the ever-changing environment in the most energy efficient way, that is, through various regulatory mechanisms. Many of these adjustments are carried out by changing the patterns of gene expression, and the pleiotropic effects of certain mutations on bacterial growth, metabolism and behavior are commonly observed, attesting to the existence of global regulation in bacteria.

Definitions

In a global regulatory network, sets of operons, scattered physically throughout the bacterial genome and sometimes representing disparate functions, are coordinately controlled (Gottesman, 1984). A global regulon is defined as a network of operons under the control of a common and regulatory protein in a global regulatory network (Gottesman, 1984). However, most commonly, even a simple environmental change can induce several regulons. Therefore a stimulon is used to refer to the entire sets of regulons responding to a single environmental stimulus (Smith and Neidhardt, 1983b). Sometimes, in order to emphasize the intricacy and overlapping nature of a global regulatory network, modulon was introduced with the definition of a group of operons and/or regulones under different specific controls but sharing a pleiotropic regulatory protein (Iuchi and Lin, 1988; Iuchi et al., 1990). For example, the crp modulon and the fnr modulon. An overall description of global regulation was reviewed by Neidhardt (1987a), and the advancement of some individual multigene systems has been compiled in detail (Neidhardt, 1987b). The confirmed regulatory genes of different stimuli include those involved in utilization of nitrogen, carbon and phosphate, as well as in response to stresses from heat shock, osmolarity, oxidative stress, starvation (stringent control), DNA damage (SOS response). Regulation by specific regulatory proteins in chemotaxis
and anaerobic respiration are also known. Since then, more and more reports have identified global regulation in other multigene regulatory systems, for example, capsule synthesis (Stout and Gottesman, 1990), sporulation, bacterial virulence, flagellar biogenesis, secretion of degradative enzymes, oxygen regulation, symbiotic nitrogen fixation, di- and tri-carboxylate transport, hydrogenase, exoprotein synthesis, alginate production, phosphoglycerated transport, motility and development, uptake of hexose phosphates, to name a few (Stock et al., 1989). Obviously, global regulation is a common scenario rather than a rare isolated case in bacteria.

Signal transduction in bacteria

Sensor-regulator two-component regulatory systems. Genetic and biochemical analyses of the responses of bacteria to environmental stimuli have revealed that signal transduction in response to certain environmental stimuli is usually carried out through the communication between pairs of two classes of proteins. Each pair is composed of a sensor protein which can detect the change in the surroundings and send a signal to its partner, a regulator protein. In turn, the sensitized regulator would affect the transcription initiation of certain sets of operons either positively or negatively, and thus constitutes a global regulatory network. Significant homology has been observed among proteins of each class (Deretic et al., 1989; Kofoed and Parkinson, 1988; Stock et al., 1990; Ronson et al., 1987). In most cases, the conserved domain of the sensor protein extends over about 250 amino acids in the C-terminus, and a conserved domain of about 120 amino acids in the N-terminus of the regulator protein (Albright et al., 1989).

Signal transduction by phosphorylation. In a typical sensor-regulator two-component regulatory system, a transmembrane domain separates the sensor into the periplasmic domain which somehow senses a specific environmental stimulus and causes autophosphorylation of its C-terminus, the phosphate group is then to the N-terminus of
the corresponding regulator by the protein kinase activity of the sensor protein. Phosphorylation of the regulator alters the activity of its C-domain, which ultimately carries out the appropriate response. The C-domain of regulator protein is believed to possess DNA binding activity, however no helix-turn-helix motif which recognizes DNA dyad symmetry sequence has been identified. They may constitute a different type of DNA binding protein. The simplified model for signal transduction is depicted in Fig. 1. Such families of homologous signal transduction proteins are involved in a variety of responses in various bacteria (Fig. 2). **Protein phosphorylation** is one of the posttranslational covalent modifications, which is important to the functioning of proteins in cellular metabolism and responses. Protein phosphorylation in prokaryotes (Cozzone, 1988) and its roles in signal transduction (Stock et al., 1989 and 1990) have been reviewed. In a signal transduction relay, the sensor is a kinase that uses ATP to phosphorylate itself at the histidine residue. The phosphoryl group is then transferred to an aspartic acid side chain within the conserved domain of the regulator. The level of regulator phosphorylation is further controlled by associated phosphatase activities. The chemistry and function of phosphorylation at different substrate amino acids are compared (Stock et al., 1990). The thermostability of phosphorylation at histidine and aspartate residues is quite distinct from the more commonly observed phosphorylations at serine, threonine and tyrosine side chains. The intrinsic stability of phosphoserine is more suitable for the induction of energetically unfavorable conformational changes involved in critical enzymatic activities. For phosphohistidine the "high-energy" molecule is just what is needed for a protein functioning as a catalytic phosphate group donor. Phosphorylation of this amino acid is almost independent of the protein it is a part. On the contrary, the stability of the phosphoaspartate residue in the protein is dependent on the rest of the protein, which is consistent with its role in triggering a conformational change. The signal detection
Figure 1. Model for signal transduction in two-component regulatory systems. A. depicted as a stimulus and response pathway (excerpted from Neidhardt, 1987). B. Signal transduction between sensor and regulator. Diagonal shading indicates conserved domains (excerpted from Albright et al., 1989).
A

Stimulus
↓
Sensor
↓
Signal
↓
Regulator

Operon_a  Operon_b  Operon_n
↓  ↓  ↓
Regulon proteins
↓
Response

Return

B

ENVIRONMENTAL STIMULUS
↓

N  C  SENSOR

REGULATOR
N  C

TARGET RESPONSE
Figure 2. Two families of homologous signal transduction proteins in bacteria. Histidine protein kinases and response regulators have been identified in a wide range of Gram-positive and Gram-negative species. The kinases share a homologous domain of about 100 amino acids generally located near the C terminus (open ovals). Members of this family have homologous sequences surrounding a conserved histidine residue (inverted triangles), and most members of the family have stretches of hydrophobic residues characteristic of membrane-spanning sequences (filled boxes). The response regulators all share a homologous domain of about 100 residues generally located at the N terminus (filled ovals). Many response regulator proteins fall into one of three subfamilies defined by homologies between their C-terminal domains (hatched and striped boxes). The FrzE and VirA proteins contain both histidine kinase and response regulator domains. Abbreviations: At, Agrobacterium tumefaciens; Bp, Bradyrhizobium parasponiae; Bs, Bacillus subtilis; Ea, Enterobacter aerogenes; Ec, Escherichia coli; Ka, Klebsiella pneumoniae; Mx, Myxococcus xanthus; Pa, Pseudomonas aeruginosa; Rc, Rhodobacter capsulatus; Rj, Rhizobium japonicum; Rl, Rhizobium leguminosarum; Rm, Rhizobium meliloti; Rt, Rhizobium trifolii; Sa, Staphylococcus aureus; St, Salmonella typhimurium. Figure 2 is excerpted from Stock, et al., 1990.
<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>SPECIES</th>
<th>HISTIDINE PROTEIN KINASE</th>
<th>RESPONSE REGULATOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemotaxis</td>
<td>Ea, Ec, St</td>
<td>CheA</td>
<td>CheY, CheB</td>
</tr>
<tr>
<td>Motility and development</td>
<td>Mx</td>
<td>FtsE</td>
<td>SpeoF</td>
</tr>
<tr>
<td>Sporulation</td>
<td>Bs</td>
<td>SpoIIJ</td>
<td>SpoOA</td>
</tr>
<tr>
<td>Exoprotein synthesis</td>
<td>Sa</td>
<td></td>
<td>AgrA</td>
</tr>
<tr>
<td>Alginate production</td>
<td>Pa</td>
<td></td>
<td>AlgR</td>
</tr>
<tr>
<td>Nitrogen regulation</td>
<td>Bp, Ec, Kd, Kp, Rs, Rm, St</td>
<td>NR$_{II}$</td>
<td>NR$_{I}$</td>
</tr>
<tr>
<td>Dicarboxylic acid transport</td>
<td>Rl, Rl, Rm, Rt</td>
<td>DctB</td>
<td>DctD</td>
</tr>
<tr>
<td>Phosphoglycerate transport</td>
<td>St</td>
<td>PglB</td>
<td>PglA</td>
</tr>
<tr>
<td>Outer membrane protein</td>
<td>Ec, St</td>
<td>EnvZ</td>
<td>OmpR</td>
</tr>
<tr>
<td>Phosphorous assimilation</td>
<td>Bs, Ec</td>
<td>PhoR</td>
<td>PhoB</td>
</tr>
<tr>
<td>ND</td>
<td>Ec</td>
<td>PhoM</td>
<td>ORF2</td>
</tr>
<tr>
<td>Virulence</td>
<td>At</td>
<td>VirA</td>
<td>VirG</td>
</tr>
<tr>
<td>F plasmid expression and</td>
<td>Ec</td>
<td>CpxA</td>
<td>ArcA, Flj</td>
</tr>
<tr>
<td>aerobic metabolism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrogen fixation</td>
<td>Pm</td>
<td>FliL</td>
<td></td>
</tr>
<tr>
<td>Degradative exoenzyme</td>
<td>Bs</td>
<td>DegS</td>
<td>DegU</td>
</tr>
<tr>
<td>expression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uptake of hexose phosphates</td>
<td>Ec</td>
<td>UhpB</td>
<td>UhpA, ComA</td>
</tr>
<tr>
<td>Competence</td>
<td>Bs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reductase expression</td>
<td>Ec</td>
<td>NcrX</td>
<td>NcrL, UvrC-ORF2</td>
</tr>
<tr>
<td></td>
<td>Ec</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
appears to use two routes: either the sensor receives an environmental stimulus directly or there is another primary environmental sensor (e.g. the MCP proteins in chemotaxis) that receives the stimulus and transduces it into an intracellular (or intramembrane) signal that is then detected by the sensor of the two-component regulatory system. A good illustration and summary of the adaptive responses of chemotaxis, nitrogen and phosphate regulation, osmoregulation, as well as other systems in bacteria are presented in the review of Stock et al. (1989). **Cross-talk** by phosphorylation and dephosphorylation has been observed in both directions for some members of the two component regulatory families. That is, a particular sensor can communicate (cross-talk) with regulators from other systems using the conserved mode of signal transduction. This *in vivo* could provide a means for integrating many types of environmental stimuli.

A variation of two-component signal transduction system is the **one-component system**. In the toxR system of *Vibrio cholerae*, both the transcriptional activation and stimuli sensing seems to be executed through one protein ToxR, which regulates cholera toxin expression. The conserved domain of ToxR to other two-component regulators is in the cytoplasmic N-terminus whereas the C-terminus of ToxR is located in the periplasmic space presumably functioning as a sensor (Miller et al., 1987). Another example is the **arcB gene** product of *E. coli* which is responsible for anaerobic repression of operons and regulons of aerobic function. A homology has been found between the central region of ArcB and the C-terminal region of several sensor proteins and another conserved domain in the C-terminal region of ArcB shows homology with the regulator proteins of two-component systems. Thus ArcB seems to be a sensor protein which has incorporated a regulator domain (Iuchi et al., 1990).
Significance and Application of pH Regulation

Understanding of bacterial cell physiology could provide some insight of otherwise difficult-to-dissect fundamental rules of life in higher organisms. Even for the comparatively simpler, most studied E. coli cells, endless questions still remain unanswered. However, thanks to advanced genetic and molecular tools and techniques, it is now possible to study bacterial cell physiology in various aspects at the genetic and molecular level. One such field is pH responses. Study of bacterial pH regulation not only satiates scientific curiosity but also elicits pragmatic implication. External pH influences bacterial cell function primarily through the transmembrane proton motive force and the intracellular pH. The productivity of protein is affected by the intracellular metabolism and cell membrane transport of nutrients and metabolic products, which in turn are influenced by external pH.

Large-scale production of some precious biological materials has been made possible with the advent of modern genetic engineering. A suitable expression vector and an accommodating host cell are two major considerations for maximizing gene expression. An ideal expression vector might have the following features: small size, high genetic stability in the absence of selection, low copy number under standard conditions, inducible transcription control and efficient translation signals, and inducible copy number control (Polisky, 1986). The regulation of cloned gene expression can have an important influence on the productivity of the recombinant microorganisms. Overproduction of exogenous protein often imposes a burden to the growth of the cells by competing for the energy reserves and various components required for the synthesis of the host own proteins. The results could be more serious than reduced growth rate and plasmid instability during multiplying cycles, sometimes minute production of the foreign protein can be deleterious to inhibit or kill the cells before they can reach a proper
concentration for an economical harvest. Therefore, inducible plasmid promoters are often utilized to control the time and extent of plasmid gene expression. Major strong promoters now in use include λ \( P_L \), \( trp \), \( tac \), \( phoA \), and \( lpp \)-derived promoters. Common ways to initiate transcription are temperature shift or chemical induction, however they all have substantial limitations (Brosius, 1988).

In large bioreactors, on the industrial scale, pH homeostasis of a culture can be ensured by external control with a pH-stat which will adjust the pH to a desired value by injecting acid or base. The parameters of pH and oxygen saturation in a bioreactor can be easily monitored, and the control is feasible and comparatively more economical. Therefore, pH regulation of \( cad \) promoter sort appears to be an attractive system worth further exploitation. The \( cad \) promoter has shown some promising features of an expression promoter, for example, tight and reversible control by external pH has been observed, and an initial 200-fold induction has been achieved (Tolentino, 1991). Higher yield is expected by improving expression vectors and hosting cells, as well as by adjusting production parameters after more knowledge about this regulatory system is available.
CHAPTER 2. MATERIALS AND METHODS

Materials

Bacterial strains

_E. coli_ strains used in and constructed during this study are described in Table 1.

Plasmid DNA

Cloning vectors are described in Table 2. Clones of _cadA_ and/or _ cadB_ and _lac_ fusion clones are listed in Table 3.

Bacterial phages

Bacterial phages IR1, R408 and M13K07 are from Enea and Zinder (1982), Bullock (1987), and De Lorenzo (1988) respectively.

Bacterial media

_Falkow decarboxylase medium_ was described (Falkow, 1958). **pH-buffered modified Falkow lysine decarboxylase medium** contained 5 g of Bactopeptone, 3 g of yeast extract, 5 g of L-lysine (monohydrochloride) per liter and was buffered with 20 g (100 mM) of MES for pH 5.5 or 23 g (100 mM) of HEPES for pH 8. **CadR medium** was essentially the same as the lysine-free medium used for the selection of SAEC (S-2-aminoethyl-L-cysteine) resistant strains (Popkin and Maas, 1980; Tolentino, 1991), except SAEC was not added and the medium was buffered with 100 mM MES for pH 5.5 or with 100 mM HEPES for pH 8. A detailed recipe for CadR medium is described in Table 4. CadR medium contains all the ingredients necessary for cell growth and all the amino acids except for lysine. In the plus lysine control experiments, 0.5% L-lysine hydrochloride was added to the CadR medium. **LB medium.** One liter of LB contains 10 g of NaCl, 10 g of Bacto-tryptone, 5 g of Bacto-
yeast and 30 mg cysteine. **pH-buffered Modified Falkow lysine decarboxylase plates with X-Gal.** X-Gal (40 µg/L) was added to the medium for the selection of pH-inducible promoter. If the *lacZ* gene is fused to a fragment that contains a pH responsive promoter, induced expression of β-galactosidase will cause the color change of the colony to blue by utilizing 5-bromo-4-chloro-3-indolyl-β-D-galactosidase (X-gal). **MacConkey lactose plates** (40 g of MacConkey agar base and 10 g of lactose per liter) were also used for detection of β-galactosidase activity. Antibiotics were included in the media at all times to retain plasmids during growth.

**Oligonucleotide primers**

Universal primer and -40 primer were provided by the manufacturer; other oligonucleotides for DNA sequencing and PCR reactions were synthesized on a Biosearch model synthesizer by K. Muthukrishnan in this department. The sequences of primers used in this study are listed in Table 5.

**Antibiotics and chemicals used for selection of strains**

Ampicillin (Ap, 30 µg/ml), chloramphenicol (Cm, 32 µg/ml), kanamycin (Km, 25 µg/ml), spectinomycin (Sp, 100 µg/ml), streptomycin (Str, 50 µg/ml), tetracycline (Tc, 15 µg/ml), isopropyl-β-D-thiogalactopyranoside (IPTG, 40 µg/ml), and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal, 40 µg/ml).

**Enzymes**

Restriction enzymes, modifying enzymes and T4 DNA ligase were purchased from Promega Corp., New England Biolabs, Inc., and Bethesda Research Laboratories, Inc., and were used according to manufacturers recommendations.
Reaction kits

The Sequenase Version 2.0 sequencing kit from United States Biochemical, Corp., and the GeneAmp\textsuperscript{TM} DNA amplification reagent kit was purchased from Perkin-Elmer Cetus.

Radioactive materials

Tran\textsuperscript{35}S-label (Sp. Act. >1000Ci/mmol), [\alpha-\textsuperscript{32}P]-dATP (Sp. Act. >3000 Ci/mmol) and [\gamma-\textsuperscript{32}P]-dATP (Sp. Act. >4000Ci/mmol) were purchased from ICN Biomedicals, Inc. [\textsuperscript{35}S]-dATP (500 Ci/mmol) was from Dupont-NEN Research Products.
### Table 1. Bacterial strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>71-18</td>
<td>(lac-proAB) { F' proAB lacI^{\varnothing} Z \Delta M15 }</td>
<td>Dente et al., 1983</td>
</tr>
<tr>
<td>CSR603</td>
<td>F(^{-}) phr-1 recA1 uvrA6 supE44 rpsL31</td>
<td>Sancar et al., 1981</td>
</tr>
<tr>
<td>CSR603F(^{+})</td>
<td>CSR603 { F' Tn10 proAB lacI^{\varnothing} Z \Delta M15 }</td>
<td>This work</td>
</tr>
<tr>
<td>GJT001</td>
<td>MC4100 cadR(^{-})</td>
<td>Tolentino, 1991</td>
</tr>
<tr>
<td>GJT001F(^{+})</td>
<td>GJT001 { F' Tn10 proAB lacI^{\varnothing} Z \Delta m15 }</td>
<td>This work</td>
</tr>
<tr>
<td>GJT002</td>
<td>GNB8385K cadR(^{-})</td>
<td>Tolentino, 1991</td>
</tr>
<tr>
<td>GJT002F(^{+})</td>
<td>GJT002 { F' Tn10 proAB lacI^{\varnothing} Z \Delta M15 }</td>
<td>This work</td>
</tr>
<tr>
<td>GNB581</td>
<td>MC4100 cadA :: Mu dX (Ap(^{+}), lac)</td>
<td>Auger et al., 1989</td>
</tr>
<tr>
<td>GNB7145K</td>
<td>MC4100 adi :: Mu dI 1734 (Km(^{+}), lac)</td>
<td>Auger et al., 1989</td>
</tr>
<tr>
<td>GNB8385</td>
<td>MC4100 cadB :: Mu dX (Ap(^{+}), lac)</td>
<td>Auger et al., 1989</td>
</tr>
<tr>
<td>GNB8385K</td>
<td>MC4100 cadB :: Mu dI 1734 (Km(^{+}), lac)</td>
<td>Auger et al., 1989</td>
</tr>
<tr>
<td>MC4100</td>
<td>(arg-lac)U169 rpsL150 relA1</td>
<td>Casadaban et al., 1979</td>
</tr>
<tr>
<td>XL1-blue</td>
<td>lac { F' Tn10 proAB lacI^{\varnothing} Z \Delta M15 }</td>
<td>Bullock, 1987</td>
</tr>
</tbody>
</table>
### Table 2. Cloning Vectors

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant features</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloning vectors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBGS18⁺</td>
<td>Phagemid Km⁺</td>
<td>Spratt et al., 1986</td>
</tr>
<tr>
<td>pEMBL8⁺</td>
<td>Phagemid Ap⁺</td>
<td>Dente et al., 1983</td>
</tr>
<tr>
<td>pEMBL9⁺</td>
<td>Phagemid Ap⁺</td>
<td>Dente et al., 1983</td>
</tr>
<tr>
<td>pPV9⁺+700</td>
<td>Protein expression vector Ap⁺</td>
<td>Vermersch, 1988</td>
</tr>
<tr>
<td>pRS550</td>
<td>{} lacZ operon fusion</td>
<td>Simons, et al., 1987</td>
</tr>
<tr>
<td>pRS551</td>
<td>{} lacZ operon fusion</td>
<td>Simons, et al., 1987</td>
</tr>
<tr>
<td>pRS552</td>
<td>{} lacZ Protein fusion</td>
<td>Simons, et al., 1987</td>
</tr>
<tr>
<td>pRS577</td>
<td>{} lacZ Protein fusion</td>
<td>Simons, et al., 1987</td>
</tr>
<tr>
<td>pRS552-1</td>
<td>pRS552 with cloning sites replaced by polylinker BE</td>
<td>This work</td>
</tr>
<tr>
<td>pXA</td>
<td>Derived from pCON4 Str⁺ and Sp⁺</td>
<td>De Lorenzo, 1988</td>
</tr>
<tr>
<td>pXA-1</td>
<td>pXA with polylinker BE</td>
<td>This work</td>
</tr>
</tbody>
</table>
Table 3. Plasmid constructs

<table>
<thead>
<tr>
<th>Clone</th>
<th>Descriptions</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCADA</td>
<td><em>cadA</em> cloned in pBGS18+</td>
<td>This work</td>
</tr>
<tr>
<td>pCADB</td>
<td><em>cadB</em> cloned in pPV9+700</td>
<td>This work</td>
</tr>
<tr>
<td>pKER65</td>
<td>Mu d5005 <em>cad</em>+ <em>mel</em>+ <em>adi</em>+</td>
<td>Auger et al., 1989</td>
</tr>
<tr>
<td>pLC4-5</td>
<td><em>lysU</em>+ <em>cadA</em>+</td>
<td>Clarke and Carbon, 1979</td>
</tr>
</tbody>
</table>

**lac** fusion plasmids

<table>
<thead>
<tr>
<th>Operon fusion</th>
<th>Protein fusion</th>
<th>Titrating plasmid</th>
<th>Fragment description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSM10</td>
<td>pSM552-1-10</td>
<td>-</td>
<td><em>BamHI</em>-EcoRI2747 [pKER65]</td>
</tr>
<tr>
<td>pSM170</td>
<td>pSM552-1-170</td>
<td>-</td>
<td><em>PvuII</em>-PvuII1115 [pSM10]</td>
</tr>
<tr>
<td>pSM180</td>
<td>-</td>
<td>-</td>
<td>EcoRV303-EcoRI2747 [pSM10]</td>
</tr>
<tr>
<td>pSM210</td>
<td>-</td>
<td>-</td>
<td><em>BamHI</em>-EcoRV303 [pSM10]</td>
</tr>
<tr>
<td>pSM225</td>
<td>-</td>
<td>pTM220</td>
<td><em>PvuII</em>-EcoRV303 [pSM10]</td>
</tr>
<tr>
<td>pSM230</td>
<td>-</td>
<td>-</td>
<td><em>BamHI</em>-PvuII1 [pSM10]</td>
</tr>
<tr>
<td>pSM265</td>
<td>-</td>
<td>-</td>
<td>SspI162-HaeII703 [pSM10]</td>
</tr>
<tr>
<td>pSM270</td>
<td>-</td>
<td>-</td>
<td><em>PvuII</em>-HaeII703 [pSM10]</td>
</tr>
<tr>
<td>pSM290</td>
<td>pSM577-290</td>
<td>-</td>
<td><em>BamHI</em>-HpaI1633 [pSM10]</td>
</tr>
<tr>
<td>pSM303</td>
<td>pSM552-303</td>
<td>-</td>
<td>(Primer+)-(R-)</td>
</tr>
<tr>
<td>pSM315</td>
<td>pSM552-315</td>
<td>-</td>
<td>SspI162-HaeII703 [pSM10]</td>
</tr>
<tr>
<td>pSM325</td>
<td>-</td>
<td>-</td>
<td>SspI162-PvuII1115 [pSM10]</td>
</tr>
<tr>
<td>pSM335</td>
<td>-</td>
<td>-</td>
<td><em>BamHI</em>-(R-)</td>
</tr>
<tr>
<td>pSM345</td>
<td>pSM552-345</td>
<td>-</td>
<td>SspI162-(R-)</td>
</tr>
<tr>
<td>pSM355</td>
<td>pSM552-355</td>
<td>-</td>
<td>RsaI99-(R-)</td>
</tr>
<tr>
<td>pSM360</td>
<td>-</td>
<td>-</td>
<td><em>BamHI</em>-ScaI [pSM10]</td>
</tr>
</tbody>
</table>
### lac fusion plasmids (cont.)

<table>
<thead>
<tr>
<th>Operon fusion</th>
<th>Protein fusion</th>
<th>Titrating plasmid</th>
<th>Fragment description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSM370</td>
<td>pSM552-1-370</td>
<td>-</td>
<td>BamHI*-SalI_{1435}[pSM10]</td>
</tr>
<tr>
<td>pSM380</td>
<td>pSM552-1-380</td>
<td>-</td>
<td>BamHI*-XhoI_{2102} [pSM10]</td>
</tr>
<tr>
<td>pSM390</td>
<td>pSM577-390</td>
<td>-</td>
<td>BamHI*-(cadB-)</td>
</tr>
<tr>
<td>pSM405</td>
<td>pSM552-405</td>
<td>pTM400</td>
<td>EcoRV_{303}-RsaI_{533} [pSM10]</td>
</tr>
<tr>
<td>pSM415</td>
<td>pSM552-415</td>
<td>pTM410</td>
<td>SspI_{162}-RsaI_{533} [pSM10]</td>
</tr>
<tr>
<td>pSM425</td>
<td>pSM552-425</td>
<td>pTM420</td>
<td>(SsRV1)-(R-)</td>
</tr>
<tr>
<td>pSM435</td>
<td>pSM552-435</td>
<td>pTM430</td>
<td>(SsRV2)-(R-)</td>
</tr>
<tr>
<td>pSM445</td>
<td>pSM552-445</td>
<td>pTM440</td>
<td>(SsRV3)-(R-)</td>
</tr>
<tr>
<td>pSM455</td>
<td>pSM552-455</td>
<td>pTM450</td>
<td>(SsRV4)-(R-)</td>
</tr>
<tr>
<td>pSM475</td>
<td>-</td>
<td>-</td>
<td>(Primer+)-(T-)</td>
</tr>
<tr>
<td>pSM485</td>
<td>pSM552-485</td>
<td>-</td>
<td>(SsRV1)-(T-)</td>
</tr>
<tr>
<td>pSM495</td>
<td>pSM552-495</td>
<td>pTM490</td>
<td>(SsRV2)-(T-)</td>
</tr>
<tr>
<td>pSM505</td>
<td>pSM552-505</td>
<td>-</td>
<td>(SsRV3)-(T-)</td>
</tr>
<tr>
<td>pSM515</td>
<td>pSM552-515</td>
<td>-</td>
<td>(SsRV4)-(T-)</td>
</tr>
<tr>
<td>pSM535</td>
<td>pSM552-535</td>
<td>-</td>
<td>(Primer+)-(A-)</td>
</tr>
<tr>
<td>pSM545</td>
<td>pSM552-545</td>
<td>-</td>
<td>(SsRV1)-(A-)</td>
</tr>
<tr>
<td>pSM555</td>
<td>pSM552-555</td>
<td>-</td>
<td>(SsRV2)-(A-)</td>
</tr>
<tr>
<td>-</td>
<td>pSM552-565</td>
<td>-</td>
<td>(SsRV3)-(A-)</td>
</tr>
<tr>
<td>pSM575</td>
<td>pSM552-575</td>
<td>-</td>
<td>(SsRV4)-(A-)</td>
</tr>
<tr>
<td>pSM585</td>
<td>pSM552-585</td>
<td>-</td>
<td>(SsRV5)-(A-)</td>
</tr>
<tr>
<td>pSM595</td>
<td>pSM552-595</td>
<td>-</td>
<td>(SsRV6)-(A-)</td>
</tr>
<tr>
<td>pSM605</td>
<td>pSM552-605</td>
<td>pTM600</td>
<td>(SsRV5)-(R-)</td>
</tr>
</tbody>
</table>
### lac fusion plasmids (cont.)

<table>
<thead>
<tr>
<th>Operon fusion</th>
<th>Protein fusion</th>
<th>Titrating plasmid</th>
<th>Fragment description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSM615</td>
<td>pSM552-615</td>
<td>pTM610</td>
<td>(SsRV6)-(R-)</td>
</tr>
<tr>
<td>pSM625</td>
<td>pSM552-625</td>
<td>-</td>
<td>(SsRV5)-(T-)</td>
</tr>
<tr>
<td>pSM635</td>
<td>pSM552-635</td>
<td>-</td>
<td>(SsRV6)-(T-)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>pTM640</td>
<td>(SsRV2)-(SsRV7-)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>pTM650</td>
<td>(SsRV9)-(SsRV7-)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>pTM660</td>
<td>(SsRV9)-(SsRV10-)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>pTM680</td>
<td>(SsRV2)-(SsRV10-)</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>pTM690</td>
<td>(SsRV2)-(SsRV11-)</td>
</tr>
<tr>
<td></td>
<td>pXA-1-10</td>
<td>-</td>
<td><em>BamHI</em>-EcoRI2747 [pKER65]</td>
</tr>
<tr>
<td></td>
<td>pXA-1-170</td>
<td>-</td>
<td>PvuII1-PvuII1115 [pSM10]</td>
</tr>
<tr>
<td></td>
<td>pXA303</td>
<td>-</td>
<td>(Primer+)-(R-)</td>
</tr>
</tbody>
</table>

*For operon fusions, pRS550 was used as vector for those plasmid numbers ending with a zero, and pRS551 was used for those ending with a 5.  
*For protein fusions, pRS552, pRS577, or pRS552-1 were used as indicated by the first section of the names.  
*All pTM plasmids used pEMBL8+ as the cloning vector.  
*Insert endpoints separated by hyphen are numbered according to Fig.3 (Chapter 3). The sequences of the oligonucleotide primers are given in Table 5. The plasmid used as the source of fragment is bracketed; endpoints defined by oligonucleotide primers are in parentheses.  
*Site originated from the mini-Mu vector.*
Table 4. Ingredients of CadR medium

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>K$_2$HPO$_4$</td>
<td>7 g/L</td>
<td>DL-Alanine</td>
<td>400 mg/L</td>
</tr>
<tr>
<td>KH$_2$PO$_4$</td>
<td>3 g/L</td>
<td>L-arginine</td>
<td>100 mg/L</td>
</tr>
<tr>
<td>Na$_3$ citrate. 2H$_2$O</td>
<td>0.5 g/L</td>
<td>L-Aspartate</td>
<td>150 mg/L</td>
</tr>
<tr>
<td>MgSO$_4$.7H$_2$O</td>
<td>0.1 g/L</td>
<td>L-cysteine</td>
<td>50 mg/L</td>
</tr>
<tr>
<td>(NH$_4$)$_2$SO$_4$</td>
<td>1 g/L</td>
<td>L-Glutamate</td>
<td>200 mg/L</td>
</tr>
<tr>
<td>Glucose@</td>
<td>2 g/L</td>
<td>DL-Histidine*</td>
<td>75 mg/L</td>
</tr>
<tr>
<td>Thiamine*</td>
<td>2 μg/ml</td>
<td>DL-Isoleusine</td>
<td>100 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Leucine</td>
<td>50 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DL-Phenylalanine*</td>
<td>100 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Proline*</td>
<td>100 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Serine</td>
<td>50 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Threonine</td>
<td>50 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Tryptophan*</td>
<td>75 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Tyrosine*</td>
<td>50 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Valine</td>
<td>50 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Glycine</td>
<td>50 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Methionine</td>
<td>100 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Asparagine</td>
<td>100 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L-Glutamine</td>
<td>100 mg/L</td>
</tr>
<tr>
<td>HEPES (pH 8) or MES (pH 5.5)</td>
<td></td>
<td></td>
<td>100 mM</td>
</tr>
</tbody>
</table>

After all the ingredients (except for * and @) were dissolved, the medium was pH to 8 or 5.5. *: Filter sterilization. @: Autoclave separately.
Table 5. Oligonucleotide primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>-40</td>
<td>GTTTTCCCAGTCACGAC@</td>
</tr>
<tr>
<td>Universal</td>
<td>GTAAAAACGACGGCCAGT@</td>
</tr>
<tr>
<td>Primer 1-</td>
<td>CGCAAGAGTAGAAATACGGC1795#</td>
</tr>
<tr>
<td>Primer 2</td>
<td>GCCGTATTTCTACTCTTGCG1815#</td>
</tr>
<tr>
<td>Primer 3</td>
<td>TGTTCTACGCTCGCAAAAATG1985#</td>
</tr>
<tr>
<td>Primer 4</td>
<td>TTTTTTACCTGAGATATGACTATGAA2123#</td>
</tr>
<tr>
<td>Primer 5-</td>
<td>AGCGCATATTCAAAGAAGC2420#</td>
</tr>
<tr>
<td>Primer 6-</td>
<td>AAAGCACCTGAGTGCAG1414#</td>
</tr>
<tr>
<td>Primer 7</td>
<td>CGTCTGAACCTCCAGATT2217#</td>
</tr>
<tr>
<td>Primer 8-</td>
<td>CGATGCTGGCCGGAATAC3641#</td>
</tr>
<tr>
<td>Primer 9</td>
<td>ATGTGGTGATGGGTATTTG4599#</td>
</tr>
<tr>
<td>Primer 10-</td>
<td>TAAGGCACGCATGCGTA4398#</td>
</tr>
<tr>
<td>Primer 11-</td>
<td>GGTATTCCGACCAAAGA2626#</td>
</tr>
<tr>
<td>Primer 12-</td>
<td>GTGGACGCGACCAGATACCG3327#</td>
</tr>
<tr>
<td>Primer 13-</td>
<td>CATGATGCACACCACCAC3314#</td>
</tr>
</tbody>
</table>
Primer 14  CACATGTATCTTGACCCG3582#
Primer 15  CCCGCCGGGAGTTCTCT4091#
Primer 16- GCGGTAATAACCAAATCGC998#
Primer 17  ACCTGCGAACCTAGCAAG805#
Primer 18  GGGTATCGCCTGTATTG1084#
Primer 19- CCATAAACTTTCCGGAAAG1619#
Primer 20- CATAGTCATATCTCCAGG2104#
Primer C- GCGAACCCTACAAAACCGCAG2351#
Primer D  CTGATGATGATGAGCGATGT2887#
Primer E- ACAGCAGACCATCATCGGTAG3035#
Primer F- TGCCAGGTGCTGTGACAGA3520#
Primer G  CACTGAGCCTGCTGCG3770#
Primer H- GGATAGTGAGCGCCGATT4161#
Primer I- CAATACCCAGCACCACAT4582#
Primer J  CCCCAATCGCCTGCGGCGC4792#
Primer K  AGAAAGCCGTCGGTTCTGG4138#
Primer N- ATGAAGTTCACGGATGG2170#
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pX</td>
<td>CTGGCACAGCGCTGGACGCG*</td>
</tr>
<tr>
<td>N-</td>
<td>ATGAAGTTTCACGGATG GG2170#</td>
</tr>
<tr>
<td>PCR primers</td>
<td></td>
</tr>
<tr>
<td>cadA-</td>
<td>CCCATGGGATC CAAATATTGC2131#</td>
</tr>
<tr>
<td>cadB+</td>
<td>GGAGAAGAGCCATGGGTTCCTGCC116#</td>
</tr>
<tr>
<td>cadB-</td>
<td>CGCTCGAGATGAAAGGAGGAGGCCTCG2054#</td>
</tr>
<tr>
<td>cadC+</td>
<td>GCCGCGAATTCAATAAGTAAGTGCTCCGGAC*</td>
</tr>
<tr>
<td>cadC-</td>
<td>CGCCCGGAATTGC CCGGCAAATAAATTGCGCGA AC412#</td>
</tr>
<tr>
<td>cadC2</td>
<td>ATTAATATTAGATCTTT CTTCCCTTTCCAATGAGTTTC*</td>
</tr>
<tr>
<td>cadBR-</td>
<td>ATGTGAAATTCCACCGCGGTGGTTATCC209#</td>
</tr>
<tr>
<td>Primer+</td>
<td>GGGGAATTC GGGGATGAAACCGGAAGCAG209#</td>
</tr>
<tr>
<td>Primer-</td>
<td>GGGGATCCCTCTTTGGCAGAACTCATGCTC101#</td>
</tr>
<tr>
<td>PrimerA-</td>
<td>CCCATGGGATC CAAATATTGC1495</td>
</tr>
<tr>
<td>pX</td>
<td>CTGGCACAGCGCTGGACGCG@</td>
</tr>
<tr>
<td>PrimerR-</td>
<td>GGGGATCCCTCTTTGGCAGAACTCATGCTC82</td>
</tr>
<tr>
<td>PrimerT-</td>
<td>GAATTTGGATCCCGGT CATGATG29</td>
</tr>
<tr>
<td>SsRV1</td>
<td>GTTTGGAGA ATTC GTATG TTCCTTT GG-227</td>
</tr>
<tr>
<td>SsRV2</td>
<td>CCAGAGAATTC CACGCAAATCC-152</td>
</tr>
</tbody>
</table>
SsRV3  GTCACGCAAT\textbf{GAATTC}CTAAACATTTAATG_{135}
SsRV4  CCATT\textbf{GAAATTC}CTAAATGTTTTATCTTTTCATG_{121}
SsRV5  CATTAGA\textbf{ATTTC}TATCTTTTCATG_{121}
SsRV6  ACG\textbf{GAATTC}ATTGTAAAACATTTAGA\textbf{ATTGTAT}TTTTTCA_{123}
RV7-   GGTTTT\textbf{GGATCC}ACACATCAGG_{105}
SsRV9  GTCAC\textbf{GAAATTC}ATTGTAAAC_{143}
RV10-  CACAT\textbf{CAGGATCC}CAAGTTG_{116}
RV11-  CCTAAG\textbf{GGATCC}GGTGAAATAAACACAAAAAG_{62}

Polylinker

BE  \text{5'AA\textbf{TT}CGCGAGATCTGCGAG\textbf{AATTC}CCTCGAGGGCCCA3'}
       \text{3'AGCGCGCTTTAGGACG\textbf{TCTTTAAGGGACGTCGGCGGTTTAG5'}}

Cloning sites (EcoRI or BamHI) used to insert the PCR product into the vector are indicated in bold type. Bases that were changed from the original sequence are underlined. Unmarked or "+" following primer name indicates the primer corresponds to the top strand or left end of the PCR product; the designation "." indicates the sequence shown corresponding to the bottom strand or right end of the PCR product. The number at the 3' end of the oligonucleotide without "#" refers to the position of the 3' terminal base in the regulatory region with the transcriptional start point defined as +1. The number followed by "#" refers to the numbering of the "cad new" nucleotide sequence file built with DNA Strider sequence analysis program. (Cont.)
(Cont.)*. Sequences of primers cadC+ and cadC2 can be found in "cadCcadBA" DNA strider file. @. Primer -40 and universal primer were described by the manufacturer (USB). Primer pX anneals to sequence upstream from the EcoRI-Smal-BamHI cloning sites of plasmid vector pXA.
Methods

Plasmid Isolation

Birnboim Procedure (Mini-Prep). Cells were pelleted from 5 ml overnight LB culture and resuspended in 110 µl freshly-made lysozyme solution (2 mg/ml lysozyme, 5 mM glucose, 10 mM EDTA, 25 mM Tris-Cl, pH 7.9). The suspension was immediately transferred to an 1.5 ml Eppendorf tube and incubated on ice for 30 min. Then 220 µl of alkaline-SDS solution (0.2 M NaOH, 1% SDS) was added, and mixed by inversion until the mixture became clear and viscous. After 5 min on ice, 165 µl of 3 M sodium acetate (pH 4.6) was added and mixed by inversion; a clot of white precipitate formed at this time. The incubation was continued for another 1 hour. The cell debris and chromosomal DNA were removed by centrifugation for 5 min in a microcentrifuge at maximal speed. Plasmid DNA was precipitated from 400 µl of supernatant by mixing with 1 ml of ethanol followed by centrifugation as above for 5 min. The precipitate was rinsed once with ethanol and dried briefly in a SpeedVac concentrator (RNA is harder to dissolve when completely dried), and then redissolved in 100 µl of 1× TE (10 mM Tris-Cl, 0.1 mM EDTA, pH 7.9), treated with RNase (10 µg/ml) at 37 °C for 10 min. The reaction was stopped by TE-equilibrated phenol/chloroform (1:1) extraction and 90 µl of supernatant was carefully removed after centrifugation for 10 min. Addition of 10 µl of 3 M sodium acetate allowed the precipitation of nucleic acid with 2 volumes of 95 % ethanol by centrifugation for 8 min. The dried pellet was finally dissolved in 50 - 100 µl of sterile deionized H₂O. Preparation of RNase stock solution and the equilibration of phenol has been described elsewhere (Maniatis et al., 1982).

Comments: This protocol gives a good yield of plasmid DNA for restriction analysis and transformation, but a large quantity of RNA is still present thus it is not recommended for ligation experiments.
**QIAGEN Plasmid Purification (Midi-prep).** A cell pellet from 150 ml overnight culture was resuspended in 4 ml buffer P1 (100 μg/ml RNase in 50 mM Tris-Cl, 10 mM EDTA, pH 8). Then 4 ml of buffer P2 (0.2 M NaOH, 1% SDS) was added and the mixture was immediately mixed by inversion and centrifuged in the Beckman J2-21 centrifuge at 15 rpm for 30 min. The supernatant (free of particles) was added to a QIAGEN-tip 100 column pre-equilibrated with 3 ml of buffer QBT (750 mM NaCl, 50 mM MOPS, 15% ethanol, pH 7) and allowed to pass through the column by gravity flow. The column was then washed with 10 ml of buffer QC (1 M NaCl, 50 mM MOPS, 15% ethanol, pH 7) and the plasmid was with 5 ml of buffer QF (1.25 M NaCl, 50 mM MOPS, 15% ethanol, pH 8.2). The plasmid DNA was precipitated with 0.7 volume of isopropanol and centrifuged in the Beckman J2-21 centrifuge at 12 rpm for 30 min. The pellet was air-dried for 10 min and redissolved in sterile water.

Comments: A good yield of high quality DNA, almost free of RNA. This was generally suitable for ligation.

**Isolation of Methylated Plasmid DNA.** The cell pellet from 500 ml of dimethylsulfate (DMS) treated culture was washed once with 30 ml WL buffer (25 mM Tris-Cl, 10 mM EDTA, pH 8) and resuspended in 8 ml WL buffer. Cells were ruptured by quick freeze-thaw 3 times and lysozyme treatment (2.5 ml of lysozyme 2 mg/ml) on ice for 10 min. Then 13 ml of Brij lysis mix (1% Brij 58, 0.4% sodium deoxycholate, 62.5 mM EDTA, 50 mM Tris-Cl, pH 8) and 1 ml 5 M NaOH were added. The mixture was mixed by inverting and centrifuged immediately in the Beckman J2-21 centrifuge at 17 k rpm for 50 min. The supernatant was mixed with 14 ml of 2.5 M NaCl and 25% PEG (Sigma 8000) by vortexing. After 1 hour incubation on ice, the mixture was centrifuged at 8 k rpm for 10 min in the Backman centrifuge; the pellet was then resuspended in 5 ml TE, extracted once with phenol/chloroform (1:1), and precipitated
with 1/10 volumes sodium acetate (3 M, pH 7) and 2 volumes of ethanol (95 %). The final pellet was dissolved in 12 ml of QIAGEN plasmid midi-prep P1 + P2 + P3 buffer mix (potassium dodecylsulfate has been removed by centrifugation), and incubated on ice for 30 min to allow RNase digestion before the solution was applied to a QIAGEN-tip 100 plasmid preparation column. Equilibration, washing, elution and precipitation were performed according to the manufacturer.

Comments: Methylated DNA is both sensitive to strong acid and alkaline conditions, therefore methylated plasmid DNA was purified by mild Brij lysis procedure.

Isolation of chromosomal DNA

Chromosomal DNAs were isolated and purified using standard procedures (Rodriguez and Tait, 1983)

Gel electrophoresis

Polyarylamide gel electrophoresis of DNA was conducted as described by Maniatis et al. (1982). Denaturing sequencing gels were prepared according to Sanger and Coulson (1978). Agarose gel electrophoresis was described by McDowell et al (1977). Ethidium bromide staining of DNA and visualization under ultraviolet illumination were described by Swenstrom and Shank (1978). SDS-polyacrylamide gel electrophoresis of protein samples and staining by coomassie brilliant blue were according to the method of Laemmli (1970).

Recombinant DNA techniques

All cloning experiments were conducted according to standard procedures (Maniatis 1982). Restriction endonuclease digestions were performed as recommended by the supplier of the enzyme. DNA fragment isolation. Following separation by electrophoresis, desired DNA fragments were eluted overnight from polyacrylamide gels
with buffer X (0.5 M ammonium acetate, 0.01 M magnesium acetate, 1% SDS, 0.1 mM EDTA), and precipitated with 2 volumes of ethanol as described by Maxam and Gilbert (1977). **Ligation and transformation.** Competent cells were either prepared by calcium chloride treatment or by Hanaham procedure (Hanaham, 1983). Shotgun cloning as well as cloning of specific fragments eluted from polyacrylamide gels with buffer X were used to prepare suitable subclones. Plasmid DNA was isolated and subjected to restriction enzyme analysis to identify the insert and to determine the orientation.

**Southern hybridization**

DNA samples were digested with appropriate enzymes and separated on agarose gels. Prior to transfer, the DNA was irradiated with UV light for 5 min, denatured in 0.5 M NaOH and 1.5 M NaCl for 30 min, and neutralized in 1 M Tris-Cl and 1.5 M NaCl for 30 min. The DNA was then transferred to nitrocellulose filters (0.45 μM) by the technique of Southern (1975). Blotted filters were immobilized by baking at 80 °C for 30 min. Hybridization with ³²P-labeled probe was performed according to the procedures provided by the manufacturer of nitrocellulose filters (Schleicher and Schuell).

**Purification of synthetic oligonucleotides**

Synthetic oligonucleotides in a screw-cap vial was incubated in a 55 °C water bath for 5 hours to allow the deprotection reaction to complete. The vial was then put in a hood and allowed to cool for a few minutes before the cap was removed. Ammonia was evaporated during overnight ventilation in the hood. The oligonucleotide solution was transferred to Eppendorf tubes and dried in a Speed Vac concentrator. The dried oligonucleotides were resuspended in 300 μl H₂O and the concentration of this crude oligonucleotide was determined by measuring the absorbance of the diluted (about 200
fold) solution at 260 nm. An aliquot of 60 to 100 µg of oligonucleotides (in a volume of 150 µl) was loaded to an one ml syringe filled with TE-washed, H₂O-equilibrated Sephadex 25 (G25-150) and centrifuged for about 7 sec (before the centrifuge reached the maximum speed) in a physicians compact centrifuge (Clay Adams). The rest of the crude oligonucleotide solution was dried and stored at -20 oC. The was collected and the column was washed twice with 100 µl H₂O. Both fractions were saved. The concentration of the first elute and the two washes were measured and the fractions with significant absorbance (usually the first elute and the first wash) were dried and resuspended in a proper amount of H₂O to make a final concentration of 1-2 µg.

**DNA sequencing**

**Sanger dideoxy chain termination sequencing.** All fragments subjected to sequencing were cloned into pEMBL vectors (Dente et al., 1983). One ml overnight LB culture of 71-18 bearing the plasmid was re-inoculated to 50 ml LB in a 250 ml Klett flask. The cells were allowed to grow aerobically at 37 oC to an absorbance of 70 Klett units (with H₂O as a blank). Then, synthesis of the single-stranded DNA was induced by infecting cells with 15 µl of phage IR1 (1.2 × 10¹² PFU). Cells were pelleted after another 4-6 hours of incubation by centrifugation at 10 k rpm for 10 min in a Beckman J2-21 centrifuge. The pellet was discarded and 28 ml of the supernatant was mixed with 7 ml of 20 % PEG (Sigma 8000)/ 2.5 M NaCl. The mixture was vortexed well and centrifuged at 10 k rpm for 10 min after incubated at room temperature for 15 min. The pellet was dissolved in 0.5 ml 1 × TE by vortexing and transferred to an Eppendorf tube. The suspension was extracted with 1 volume of phenol/chloroform twice (save upper layer) followed by two times of ether extraction (layer separation could be enhanced by brief spin). The upper ether layer and sometimes milky staff at the interphase were removed during this process. The final aqueous layer should be clear. The residual ether in the solution was evaporated in a hood for 3-5 min. One tenth volumes of 3M sodium
acetate and 2 volumes of ethanol (95 %) were then added to the aqueous layer and vortexed well. Single-stranded DNA was pelleted by centrifugation at maximal speed for 10 min and dried in a Speed Vac concentrator. Finally, the dried pellet was dissolved in 50 - 100 μl of sterile water. Usually 5 μl of DNA solution is enough for DNA sequencing experiments. Sequencing was carried out by modified dideoxy chain termination method as described by U.S. Biochemicals Sequenase 2.0 sequencing protocol. Synthesized oligonucleotide primers based on the known sequences were used to bridge the gap regions.

**Maxam and Gilbert chemical sequencing.** End-labeled double-stranded DNA was cleaved by a second enzyme and separated on a polyacrylamide gel. The band positions of appropriated fragments were detected on the gel by aligning the gel with its autoradiogram. Desired DNA fragments were cut out of the gel and eluted with buffer X as described above. Modified chemical sequencing procedure was described (Maxam and Gilbert, 1980).

**DNA and protein sequence analyses**

DNA sequences from autoradiograms were read manually, and were analyzed with DNA Strider (National Institute of Health) and MacVector (International Biotechnologies, Inc.) sequence analysis softwares. The Eugene data bank system provided by the Molecular Biology Information Resources at Baylor College of Medicine was used to perform homology searches and secondary structure prediction.

**Growth conditions**

Temperature: all cultures were grown at 37 °C. Aerobic growth: 10 ml of culture in 250 ml flask was agitated vigorously (200 rpm) and let grow to an OD₆₀₀nm of 0.3; anaerobic growth: 10 ml of culture in 16 ml screw-cap test tube was let grow to an OD₆₀₀nm of over 4.5 without agitation. pH effect was assayed by comparing growth in
modified Falkow medium with lysine at pH 5.5 or pH 8. Lysine effect was assayed by comparing growth in CadR medium with or without lysine.

**β-Galactosidase assay**

β-galactosidase assays were performed basically according to Miller (1972) and Auger (1986) except that cells were grown under different conditions (described as above) to detect a pH or lysine effect on the induction of cad operon.

**lysine decarboxylase assay**

Determination of lysine decarboxylase activity manometrically by capturing $^{14}$C-labeled carbon dioxide was described by Boeker and Fischer (1983). Another spectrophotometric method was described by Phan et al (1982). All reagents and procedures used for the lysine decarboxylase assay were as described by Phan et al. (1982), except the reaction temperature was changed from 40 °C to 43 °C. 1.02 × 10⁻² M TNBS (2',4',6'-trinitrobenzylsulfonic acid) was added to react with cadaverine, lysine and trace amounts of other amines and amino acids at 43 °C for 5 min. The intensely colored product N,N'-bistrinitrophenyllysine (TNP-lysine) was separated from N,N'-bistrinitrophenylcadaverine (TNP-cadaverine) by extraction of the water layer with 2 ml of toluene. The extraction was done by vortexing the mixture at top speed for 20 sec. The absorbance of the final extract was read at 340 nm and diluted if necessary.

**Determination of cadaverine content**

The procedure is exactly the same as measuring lysine decarboxylase colorimetrically, except that at the final step the OD$_{340nm}$ was measured for the organic layer (TNP-cadaverine). **Cadaverine excretion assay.** Stains were grown in modified Falkow lysine decarboxylase media buffered with 100mM MES for pH5.5 or with 100 mM HEPES for pH8. Anaerobic cultures were grown in 16 × 125 mm screw-cap test tubes containing 10 ml of modified Falkow lysine decarboxylase medium and
were incubated at 37 °C without shaking. IPTG to a concentration of 80 µg/ml was added to the culture when the cell density reached an OD$_{600}$ of 0.3 to induce the expression of $cadA$ or $cadB$ carried by plasmid pCADA or pCADB respectively. Cells were pelleted after the absorbance reached 0.65 - 0.7, and one ml of each supernatant was tested for the cadaverine content spectrophotometrically. In order to verify that other trace amount of amines and amino acids present in the supernatant did not affect the validity of cadaverine excretion assay using the spectrophotometric method, appropriate media blanks were used. Also in a separate specificity experiment, [U$^{14}$C]-lysine (1.5 µCi/ml) was added to the culture 10 min before the cells were pelleted. Subsequent reaction of the supernatant with TNBS and extraction with toluene were conducted as described above. An aliquot of the extract was mixed with scintillation fluid and counted using Beckman LS3801 liquid scintillation system.

**PCR methods**

Primers used for PCR amplification are listed in Table 5. Standard PCR procedure was performed according to the protocol provided by Perkin Elmer Cetus except that the initial melting time and temperature were set at 92 °C for 2 min, followed by 25 cycles of 1 min melting at 92 °C, 1 min annealing temperature at 60-65 °C, and 2 min of primer extension at 72 °C. The final extension was allowed to continue for 8 min. The heating and cooling cycles were performed with the Techne PHC-2 PCR system. Random mutagenesis of a defined DNA segment was achieved using a modified PCR reaction which incorporated 0.5 mM MnCl$_2$ into the standard reaction buffer and adjusted the dGTP to dATP ratio to 5 to increase the mutation frequency (Leung et al., 1989). The temperature and time settings were the same as described for standard PCR procedure except the reaction cycle was increased to 30 cycles. The amplified DNA fragments were separated on a polyacrylamide gel and the fragments was cut out of the gel and eluted with buffer X as described previously.
In vivo footprinting

DNA methylation. The method used for in vivo methylation of plasmid DNA was a modification of Martin et al. (1986). A ten ml overnight culture of the strain harboring a suitable plasmid was inoculated into 500 ml pH-buffered modified Falkow lysine decarboxylase medium, and grown at 37 °C to an A600nm of 0.5 for anaerobic growth at pH 5.5, and 0.4 for aerobic (with agitation) growth at pH 8. Rifampicin was then added to a final concentration of 200 µg/ml. After 8 min, 1 ml of dimethylsulfate (99%) was added followed by vigorous shaking for 30 sec. The culture was then quick chilled by pouring onto 250 g crushed ice. Cells were harvested by centrifugation, washed once with WL buffer (25 mM Tris-Cl, pH 8, 10 mM EDTA), and finally resuspended in 8 ml WL buffer. Isolation of methylated DNA has been described previously. No conspicuous difference in the extent of methylation was observed between cells grown at pH 5.5 and pH 8. The supernatants and washes containing DMS were treated with NaOH to about 2M for a few hours before proper disposal.

Detection of methylated residues in a DNA fragment. Purified methylated plasmid (about 4 µg) was digested with the appropriate enzyme and end labeled with [α-32P]-dATP and DNA polymerase-Klenow fragment according to protocol of Maniatis et al. (1982). Unincorporated label and nucleotides were removed using a Sephadex G-50 spun column (maniatis, 1982). The elute was ethanol precipitated with sodium acetate (pH 7) and washed once with ethanol and dried. Linearized plasmid DNA was then cleaved by a second restriction enzyme and separated on a 5% polyacrylamide gel. The desired end-labeled fragment was identified by comparing the signal on its autoradiogram and was cut out of the gel, eluted overnight with buffer X and precipitated with ethanol. The ethanol-washed and dried end-labeled fragment was dissolved in 63 µl of sterile H2O in an 1.5 ml screw-cap microcentrifuge tube, and 7 µl of piperidine
(100%) was then added. The strand scission at modified G residues was catalyzed by incubating the 10% piperidine mixture at 90 °C for 30 min. Then the piperidine was removed by drying under vacuum in a Speed Vac concentrator. The drying process was repeated two more times after dissolved in 30 μl and 20 μl of H₂O at each time. The pellet was finally dissolved in 10 μl of H₂O plus 6 μl of sequencing stop buffer, and an aliquot of the sample was counted for its radioactivity. A proper amount of sample (ideally less than 5 μl) was loaded to a 6% sequencing gel with thickness of 1.2 mm. The sequence of the same fragment without previously methylated in vivo was determined using modified Maxam and Gilbert method (Current Protocols in Molecular Biology, F.M. Ausubel et al., 1987), which provided the sequence ladder for identification of in vivo methylated G residues. The gel was covered with a layer of plastic wrap and directly exposed to a XAR-5 imaging film without prior fixing and drying procedure.

Construction of operon fusions and protein fusions

There are two general ways to monitor promoter activity with a reporter gene: operon fusion and protein fusion. Operon fusion vectors used in this study have an intact lacZ gene with its own Shine-Dalgarno (SD) sequence, whereas protein fusion vectors have functional lacZ gene with its first 8 codons truncated. Therefore, the lacZ gene in the protein fusion vector has to be fused in phase to a protein coding sequence and translated from the SD sequence of the target sequence.

Bacterial mating

Overnight cultures of F' and F− strains were reinoculated into LB (without drug) and grown at 37 °C to mid-log phase. Then cells were mixed at a ratio of 1F' to 20F− and allowed to grow for another two hours with gentle agitation. Then a series of
dilution of the culture with LB medium were plated on LB plates with appropriate drugs to select successful conjugates.

**Protein expression**

**Protein expression vectors.** Plasmids pBS18+, and pPV9+700 carrying a strong promoter (*lacUV5* or *tac* promoter) in front of a polylinker were used as protein expression vectors. Theoretically, a structure gene can be expressed from an exogenous promoter, if an SD sequence is located at a proper position before the coding sequence and the sequence between the promoter and the SD sequence does not interfere with the recognition of SD by RNA polymerase. If the intrinsic SD is not an efficient one, it can be replaced by an exogenous SD sequence as was designed for pPV9+700.

**Expression of gene product with maxicells.** Transformed maxicells were grown in special medium and UV-irradiated as described by Sancar et al. (1981). The *tac* promoter was induced with IPTG (80 μg/l) 30 min prior to labeling with Tran-35S. Protein samples were analyzed by SDS-PAGE and autoradiography.
CHAPTER 3. STRUCTURE AND PHYSIOLOGICAL ROLE OF THE cad OPERON

Identification of the Inducible Lysine Decarboxylase Coding Gene cadA

INTRODUCTION

Amino acid decarboxylases

The general description of bacterial, mammalian and plant amino acid decarboxylases has been reviewed by Boeker and Snell (1972). Most of them use PLP (pyridoxal 5'-phosphate) as coenzyme to decarboxylate the amino acid to its corresponding amines, however the involvement of PQQ (pyrroloquinoline quinone) in the catalytic reaction is also indicated in E. coli glutamate decarboxylase (Van der Meer et al., 1989) and in some other cases (Duine and Jongejan, 1989). The conserved sequence for PLP-binding site in E. coli was found to be Ser-X-His-Lys with the lysine residue as the PLP attachment site (Vaalter et al., 1986). The recent progress in the studies of PLP enzymes has be reviewed (Hayashi, 1990). As for the chemical and physical properties of the biodegradative and biosynthetic amino acid decarboxylases of E. coli as well as their similarities and differences, a detailed review was included in the thesis by Auger (1987).

Biodegradative and biosynthetic amino acid decarboxylases

Two types of amino acid decarboxylases exist in E. coli. Biosynthetic enzymes are involved in the synthesis of polyamines and are expressed at low levels regardless of pH variation (Tabor and Tabor, 1985); biodegradative enzymes are strongly induced by acid media and hence have a possible role in maintaining pH homeostasis (Gale, 1946) and in control of the intracellular CO₂ concentration (Guirard and Snell, 1964). Unlike arginine
and ornithine decarboxylases of which both types of enzymes have been well characterized (Nowark and Borker, 1981; Applebaum et al., 1975; Wu and Morris, 1973; Blethen et al., 1968; Applebaum et al., 1977), only the inducible form of lysine decarboxylase has been analyzed (Sabo et al., 1974; Sabo and Fischer, 1974). Some evidence has indicated the presence of a pH-constitutive form of lysine decarboxylase in E. coli (Wertheimer and Leifer, 1983; Goldemberg, 1980). Other inducible amino acid decarboxylases include histidine and glutamate decarboxylases (Gale, 1940).

**Lysine decarboxylase coding gene cadA**

The production of large quantities of amino acid decarboxylases by many species of bacteria under conditions of low pH has been known for many years (Gale 1946). In the presence of lysine, the expression of biodegradative lysine decarboxylase encoded by the cadA gene of *E. coli* is induced by extracellular low pH and the magnitude of induction is further augmented under anaerobic condition (Auger et al., 1989; Slonczewski, 1987). The level of inducible lysine decarboxylase of *E. coli* B under optimal inducing conditions can reach up to two percent of total cell protein (Tabor and Tabor, 1985). Similar high level induction by low pH and oxygen-limiting conditions also happens to biodegradative arginine and ornithine decarboxylases of *E. coli* (Gale, 1946). The physical and chemical properties of the inducible lysine decarboxylase have been well characterized (Sabo, 1974; Sabo et al., 1974) and reviewed by Auger (1987). The active form of inducible lysine decarboxylase is a decamer of identical subunits of about 80 Kd, and the decamer is the pentameric association of . Each monomer contains one pyridoxal-5'-phosphate covalently bound to the ε-amino group of a lysine residue. The N-terminal 7 amino acid sequence and the 15 amino acid sequence at the active site of the inducible lysine decarboxylase have been determined in chemical degradation experiments.
Mutants of \textit{cadA} have been isolated and the gene was first mapped to around 92 minutes on the \textit{E. coli} genome (Tabor et al., 1980). Subsequent detailed mapping of in \textit{vivo} Mu dX operon fusion strains deficient in lysine decarboxylase activity and mini-Mu plasmids bearing \textit{cadA} gene revealed that \textit{cadA} was at 93.7 minutes (Auger et al., 1989). In studies of protein expression from the plasmid pLC4-5, it was found that \textit{cadA} and \textit{lysU} were both expressed (Van Bogelen et al., 1983), and correlations of the expression of these two genes have been observed (Hershield et al., 1984). However the orientation and order relative to the surrounding genes were not established.

\textbf{RESULTS}

\textbf{Localization of \textit{cadA} gene}

Initial restriction mapping, cassette insertion and deletion experiments on pLC4-5 derivatives indicated that \textit{cadA} was close to the end of the insert that contains an \textit{XhoI} site (Fig. 3). From previous Southern hybridization mapping experiments on GNB581 and GNB8385, two Mu dX fusion strains deficient in lysine decarboxylase activity, the insertion positions of Mu dX were found to be in the 0.6 kb \textit{PstI-XhoI} fragment and in the 0.9 kb \textit{EcoRI} fragment respectively (Fig. 3 and Fig. 4).

According to the estimated molecular weight of inducible lysine decarboxylase (Sabo et al., 1974), \textit{cadA} requires about 2 kb for its coding region; therefore, fragments covering a region of about 3 kb were subcloned and subjected to sequencing. The match of the nucleotide sequences with the known peptide sequences of the N-terminus and the active site (Sabo and Fischer, 1974) provided direct evidence for the identification of the gene and its orientation (Fig. 5). Only one glutamine residue at the active site did not match the originally stated glutamic acid residue, which can be explained by loss of the amino group during chemical degradation of the peptide (Auger et al., 1989).
Figure 3. Chromosomal regions cloned in pLC4-5 and pKER65. The coding regions of cadB, cadA, and lysU are covered by black rectangles and the direction of transcription is indicated by arrow. The insertion positions of Mu dX in Mu lac fusion strains GNB581 and GNB8385 are indicated by dark triangles. Dotted boxes represent vector sequences of ColEI and Mu d5005. Regions beyond each side of $\mathcal{J}$ are not shown in this figure. Relevant restriction sites: B, BamHI; E, EcoRI; K, KpnI; P, PstI; R, EcoRV; Sc, ScaI; V, PvuII; X, XhoI.
Figure 4. Southern hybridization mapping of cad::lac fusion strains.
Chromosomal DNA isolated from GNB581 (lane 1), GNB8385 (lane 2), MC4100 (lane 3), and plasmid DNA pLC4-5 (lane 4), were digested with PstI (panel A) or EcoRI (panel B), separated by gel electrophoresis, transferred to a nitrocellulose filter, and hybridized with [32P]-labeled pLC4-5. The band present in MC4100 but absent in GNB581 (0.9 kb EcoRI) or GNB8385 (3.4 kb PstI) indicates the position of the insertion of the Mu lac phage. The positions of markers are as shown.
Figure 5. Nucleotide and the deduced amino acid sequences of cadBA operon. The total sequence shown starts from a PvuII site and ends with a PstI site. The right side numbers indicate the last base pair at the end of each row of the double-stranded DNA sequence. Base pair 1 to 131 is part of the C-terminal coding region of cadC (Watson et al., 1991). The transcriptional start point (+1) and the -10 and -35 sequences of cad operon were identified by Watson et al. (1991). CadB and CadA amino acid sequences are translated from the top strand and are positioned below each codon. Shine-Dalgarno sequence (SD) of each ORF is indicated. The N-terminal and active site sequences of lysine decarboxylase used to identify the cadA gene are underlined. The only discrepancy with the 15 amino acid residues at the active site sequence determined by Sabo and Fischer (1974) resides at the first nucleotide of the glutamine codon, as indicated by an asterisk. This residue would need to be G for a glutamate residue. The stem-loop structure 35 bp downstream from the cadB start codon is indicated by a pair of arrows. Also indicated is the GC-rich hair-pin structure downstream from the cadA coding sequence, and this is followed by a T region as shown. The nucleotide sequence of the cad operon has been deposited in GenBank, accession number M76411.
CAGCTGATGCTATATCTCACCCTTTTAATTGACCGCACCGGACACCCCTTTAGGGA
GTCGACTAGTATAGAGTGCGGAAATAATGCGGCTCCCCTTTTGTGGAAATGAGCCT

TTGAAAATGTTATATTCCAGACTCTCGGTCTTTATATGTTACCTTCTCTGCAAAATTTC
AATCTTTACATATTAGGTCTGAGGACAAGAATACAATGGAATAGACTGGTTTAAG

TGTCTTCAAGATAAGTAATCCGGGTTGGATTTAGTCTCGAAATATTTGTTGGTGGATT
AACGAAAGTCTTTACCATGAGGCCCCAACTAATAAGAGCCCTTTTATAAACGAAACTCAA

TGTATGCTCGTGTTGATTATAATATGTGTCGCAATGAAACTTATTATTGGGCTTAATTGGA
AACATAAGCAGAACCATAATTATACAGGCGGTAAAATAACGCGTATAAAAAATT

CATATAATTTAACCACAGAATGCTCAGCAGATCCATTGGAACATTAATATGTGTATCTTTTC
GTATATTTAAATTTGCTCTCTTACAGTGGTATTGTAACATTTGTAATTTTCAAAATAGAAAG

ATGATATCAACTTGGCATCTCTGATGTGAATAAAAACCTCAAGTCTCTCATTACCAGAA
TACATATGTTAGGCTAGCAGACTACACAAATATTATTATTTGAGATTCAAGATGAGCTTT

+1
GCCAGTTTTTGCTGGTCTCGAGGAAATAGTTATACTACTCATGACCCGAATCCCAAATCTCAAAA
CGGTCACACCAAGACAGCTCTCTTATTATCAATATGTAATCTGACTGGGCTCGGATTAAAATTTT

-SD
TGAANCTGGGAAGAGGAGATGACTCTGCAAGAGAATTTGGCTCGGCTTATGCTTGACGTGCTT
ACTTACTGCCCTCTCTCGACTCAAGACGGTTCTCTGACCCGATAAACGACATGCGCCA

cadB MetSerSerAlaLysLysIleGlyLeuPheAlaCysThrGlyGly>

GTTGTTGGGCCGTTATGATGCGGGCTGTTACATTACCTGCGAAATCTAGCAAGT
CAAAGCCGCGGCATATTATACTACCCCTGGCCGATAACGTAATAATGAGGCCTTGATGCTCA
ValValAlaGlyAsnMetMetGlySerGlyIleAlaLeuLeuProAlaAsnLeuAlaSer>

ATGGCGGTTGTTAATACGATGTTGATTATTCTCTTTATGTTGCAATTGGCTGCGC
TAGCCCGGATGTTACGAGCCCCAATCTAATAGATATATTCACACAGTTACCGGCACGCGC
IleGlyGlyIleAlaIleTrpGlyTrpIleIleSerIleGlyAlaMetSerLeuAla>

TATGTTATATGCCGACCTGGCAAAACAAACGCCAAAGAGTGGGCCCGCATTGGCTTATGCC
ATACATATGAGGCTTTGTTTTGGGCGTTGTTACCCGCTTGTTAAGGATACGG
TyrValTyrAlaArgLeuAlaThrLysAsnProGlnGlnGlyGlyProIleAlaTyrAla>

GGAGAAATTTCCCTCGATTGGTTTTACAGAGGCTTTATTACCATGTCTAAGCTG
CCCTTTAAAAGGGCGTGAAACCAAGAGCGTCCAGCAGAAATAATGCGTATCAGGACC
GlyGluIleSerProAlaPheGlyPheGlyThrGlyValLeuTyrTyrHisAlaAsnTrp>
GGCTTGCCACTCTCCTTTTTTGGTCAATAGGAGAACACATGAAAACACCCTCACACGCTG
CGAACGGTGAGGGAAAAACGAGTTTTCTTCTTTTGACTTTTTTGAGGATGTCCGAC

CGCGGATATATATATATCCTGCGATCCCAAATCTTGAGATATTTTACGCGATG
GCAGCCTATATATAGATAGCAAGCTAGTGGTAGACCCCTCTAGTAAAGGAAAAATGCCGATC

CGTCCTTACATTCTACTATCTACACCACAGCTTGTTGATGATAAACATTGCCATA
GCACCGAATGAGTAGAGAGATAGTGGTAGTCAACAAACTACTATTGATACGTAT

GCCGGTTGAGCCGAATGCTGCTCGCTGAGAGAGAAATGCAATGGCGATAAGCCGGCACCAGAC

CCGGACCCCGCTGGCTGGCAACCCGACTGCAG
GGCTGGCGACGGCCGGTTGGCGACGTC
Because the insertion position of Mu dX (lac Ap') in GNB8385 and the start codon of cadA were close to the end of the chromosomal insert in pLC4-5, there was a possibility that the plasmid might not contain a complete upstream regulatory region. Therefore, pKER65 (Auger et al., 1989), an in vivo mini-Mu constructed plasmid bearing cadA with a longer upstream region, was also used for sequencing experiments and further studies. There was no variation in the overlapping sequenced region due to strain differences between pLC4-5 originally constructed from E. coli CS520 (Clarke and Carbon. 1976) and pKER65 constructed from E. coli MC1040-2 (Casadaban and Cohen, 1979). The nucleotide sequence reveals an open reading frame of 2145 bp, which encodes a protein of 715 amino acid residues (Fig. 5). The single termination codon TAA is followed by a hair-pin structure typical of a rho-independent transcription terminator (Platt, 1986). A Shine-Dalgarno sequence GGAGA is 7 bp upstream of the translation start codon.

Subcloning and expression of cadA gene

A 2.6 kb ScaI fragment from pLC4-5 containing the Shine-Dalgarno sequence, the structural gene and the transcription terminator was put under the control of the lacUV5 promoter by ligation into the SmaI site on the plasmid pBG518+ (Spratt et al., 1986). The resulting plasmid, pCADA bears the entire cadA gene plus its own SD sequence and transcription terminator (Fig. 6). Because the translation of the α fragment of lacZ is terminated prematurely owing to stop codons present upstream from the SD sequence of cadA in the 2.6 ScaI fragment, the translation of cadA transcript is not hampered. A protein corresponding to the size of inducible lysine decarboxylase was induced by IPTG at pH 8 and identified on an SDS-PAGE gel of total cell extracts (Fig. 7). A Maxi-cell experiment also demonstrated a single band corresponding to an 81 kd protein (Fig. 8). Plasmid pCADA also expressed a functional lysine decarboxylase upon induction by
Figure 6. Construction of plasmid pCADA. A 2.6 kb Scal fragment was ligated to the SmaI site of pBGS18+ in an orientation such that cadA would be under the control of the lac promoter provided by the vector. The Scal fragment contains the SD sequence, cadA structural gene, translation stop codon, and transcription terminator. The sequence upstream from the SD sequence provides stop codons for translation from lacZ α fragment. Abbreviations: B, BamHI; E, EcoRI; P, PstI; H, HindIII; S, Smal; Sc, Scal; Km, Kanamycin; lacPOZ*: lac promoter, operator, and α fragment.
Figure 7. Protein expression of lysine decarboxylase from pCADA. Strain MC4100 harboring plasmid pCADA (lane 2) or vector pBGS18+ (lane 3) was grown aerobically in modified Falkow lysine decarboxylase medium at pH 8 in the presence of IPTG (a non-inducing condition for pH regulated expression of the chromosomal *cadA* gene). The IPTG-induced expression of *cadA* from pCADA is indicated by arrow (lane 2). This band comigrated with the purified *E. coli* lysine decarboxylase purchased from Sigma Chemical (lane 1). No expression of *cadA* was observed from pBGS18+ (lane 3). The positions of size markers are as shown.
Figure 8. Protein expression of *cadA* from pCADA with maxicells. $[^{35}S]$-labeled protein extracted from UV-irradiated maxicell CSR603 bearing pBGS18+ (panel A, lane 1) or pCADA (panel A, lane 2) were separated on an SDS-polyacrylamide gel as described in Methods. The gene product of *cadA* is indicated by arrow and the positions of protein markers are shown in Kd.
IPTG as determined by lysine decarboxylase indicator media (Falkow, 1958) and cadaverine excretion assay (see later section).

**Amino acid sequence of biodegradative lysine decarboxylase**

The amino acid composition and molecular weight of inducible lysine decarboxylase deduced from the DNA sequence are quite similar to the results reported from chemical degradation (Table 6, Sabo and Fischer, 1974). The result of Kyte and Doolittle (Fig. 9, Kyte and Doolittle, 1982) hydrophobicity analysis is consistent with the fact that the inducible lysine decarboxylase is a soluble cytosolic protein. The predicted profile of the Garnier and Robson (Garnier et al., 1978) secondary structure (Fig. 10) does not provide enough clues to predict the folding pattern during the formation of, decamer, or long aggregate under different pH and ionic conditions (Sabo et al., 1974).

**Comparison of the amino acid sequence of the inducible lysine decarboxylase with other bacterial amino acid decarboxylases**

Two sequences in the GenBank were found to be significantly similar to the biodegradative lysine decarboxylase of *E. coli*: biosynthetic ornithine decarboxylase (*speC*) of *E. coli* (S. M. Boyle, GenBank accession no M33766) and lysine decarboxylase of *Hafnia alvei* (Fecker et al., 1986, GenBank accession no X03774).

**Lysine decarboxylases of *H. alvei* and *E. coli***. Although *H. alvei* and *E. coli* are distantly related among *Enterobacter* (Steigerwalt et al., 1976), their lysine decarboxylase nucleotide sequences share 73% homology [number of the same nucleotide residues/total number of nucleotide residues of the *E. coli* lysine decarboxylase gene] and the amino acid sequences share an 80% homology [number of identical amino acid residues/total number of amino acid residues of *E. coli* lysine decarboxylase] (Fig. 11); however, only 41% of the amino acids use the same codon. The amino acids and
Table 6. Amino acid compositions of lysine decarboxylases of *Escherichia coli* and *Hafnia alvei*.

<table>
<thead>
<tr>
<th></th>
<th>E. coli B n/80000D</th>
<th>E. coli K-12 n/81207D</th>
<th>H. Alvei n/83074D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>46.3</td>
<td>46</td>
<td>50</td>
</tr>
<tr>
<td>Cys</td>
<td>6</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Phe</td>
<td>32.9</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>Gly</td>
<td>43.1</td>
<td>44</td>
<td>47</td>
</tr>
<tr>
<td>His</td>
<td>21.5</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>Ile</td>
<td>46.6</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>Lys</td>
<td>41.3</td>
<td>43</td>
<td>46</td>
</tr>
<tr>
<td>Leu</td>
<td>54.5</td>
<td>64</td>
<td>67</td>
</tr>
<tr>
<td>Met</td>
<td>29.3</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>Asp</td>
<td>74.5</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td>Asn</td>
<td>----</td>
<td>37</td>
<td>40</td>
</tr>
<tr>
<td>Pro</td>
<td>39.5</td>
<td>36</td>
<td>44</td>
</tr>
<tr>
<td>Glu</td>
<td>65.2</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Gln</td>
<td>----</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Arg</td>
<td>28.7</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>Ser</td>
<td>37.9</td>
<td>40</td>
<td>48</td>
</tr>
<tr>
<td>Thr</td>
<td>46.9</td>
<td>49</td>
<td>44</td>
</tr>
<tr>
<td>Val</td>
<td>37.5</td>
<td>38</td>
<td>43</td>
</tr>
<tr>
<td>Trp</td>
<td>9.6</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Tyr</td>
<td>31.9</td>
<td>34</td>
<td>33</td>
</tr>
</tbody>
</table>

The deduced amino acid composition of lysine decarboxylases of *E. coli* K-12 is close to the results collected from analysis of lysine decarboxylase of *E. coli* B by chemical degradation (Sabo and Fischer, 1974). The deduced amino acid composition of a pH-constitutive lysine decarboxylase of *H. alvei* (Fecker et al., 1986) is also shown. The number of each amino acids present in each molecule (n / molecular weight) is displayed.
Figure 9. Hydropathy profiles of lysine decarboxylases of *E. coli* and *H. alvei*. The hydrophobicity profiles of lysine decarboxylases of *E. coli* and *H. alvei* are shown. Each residue (abscissa) is scored for its hydropathy according to Kyte and Doolittle (1982). The higher the value (ordinate), the more hydrophobic the position is. A score below 0 indicates a hydrophilic position. Very similar patterns are shown for these two lysine decarboxylases.
E. coli

H. alvei
Figure 10. Garnier and Robson secondary structure analysis of lysine decarboxylase and ornithine decarboxylase. The top figure is the secondary structure profile of the biodegradative lysine decarboxylase of *E. coli* as compared to the profile of the lysine decarboxylase of *H. alvei* shown in the middle. The bottom figure displayed is the profile of the biosynthetic ornithine decarboxylase (*speC*) of *E. coli*. Abbreviations: C, random coil; T, reverse turn; S, beta sheet; H, alpha helix. The region around 320 to 415 comprising the active site region of lysine decarboxylase shows similar pattern. Calculation of plot was described (Garnier et al., 1978).
Figure 11. Sequence alignment of lysine decarboxylase of *H. alvei* and biosynthetic ornithine decarboxylase of *E. coli* to biodegradative lysine decarboxylase of *E. coli*. The sequence alignment of three amino acid decarboxylases is shown. Amino acid residues of the lysine decarboxylase of *H. alvei* (middle sequence) and the biosynthetic ornithine decarboxylase of *E. coli* (bottom sequence) are replaced by dashes if they are the same as the biodegradative lysine decarboxylase of *E. coli* (top sequence). Gaps between sequences are filled with dots. Extra residues appearing in the gaps are displayed with lower-case letters. The lysine residue for binding of pyridoxal 5'-phosphate is indicated by arrow.
spacing exhibit strong regions of homology except for a small portion at the carboxyl end. The two enzymes catalyze the same decarboxylation reaction of lysine to cadaverine. The difference is that, in *E. coli*, the basal level of lysine decarboxylase is very low, but upon induction by acid pH and anaerobic conditions the expression rises sharply (Gale, 1946); on the other hand, in *H. alvei*, the production of the enzyme is independent of pH changes (Sakazaki, 1981). This may suggest an explanation for the observation that although these two lysine decarboxylase genes are very similar, no similarity was found in the flanking regions.

**Comparison of biodegradative LDC and biosynthetic ODC of *E. coli.***

Biosynthetic ornithine decarboxylase catalyzes the decarboxylation of ornithine to putrescine in the polyamine biosynthesis pathway. Although the overall amino acid sequences of biodegradative lysine decarboxylase and biosynthetic ornithine decarboxylase only share 29% homology, there are several more conserved domains including the pyridoxal 5'-phosphate binding site. The conserved domains are aligned to biodegradative lysine decarboxylase throughout the sequence except the first 120 N-terminal amino acid residues (Fig. 11).

**DISCUSSION**

In the study of amino acid decarboxylases of *E. coli*, one should distinguish biodegradative amino acid decarboxylases from those biosynthetic enzymes involved in polyamine biosynthesis: arginine decarboxylase (*speA*), ornithine decarboxylase (*speC*), S-adenosyl-L-methionine decarboxylase (*speD*), and diaminopimelic acid decarboxylase (*lysA*). The biosynthetic decarboxylases are produced in low amounts and are unaffected by pH variation (Tabor and Tabor, 1985). In contrast, the expression of the biodegradative decarboxylases: arginine decarboxylase (*adi*), glutamic acid decarboxylase (*gadS*), histidine decarboxylase, lysine decarboxylase (*cadA*) and ornithine
decarboxylase are all induced by low pH and special nutritional requirements (Gale, 1946). Among them the inducible arginine, lysine and ornithine decarboxylases exhibit similarities in overall structure though they display individual characteristics (Table 7, and Tabor and Tabor, 1985). The biodegradative arginine and lysine decarboxylases have much higher turnover rate than other amino acid decarboxylases (Applebaum et al., 1975). This characteristic coupled with the high level of production under inducing conditions indicates their capacity to generate large amounts of carbon dioxide and polyamines to cope with an acidic environment. Although inducible arginine and lysine decarboxylases exist in almost all E. coli strains, inducible ornithine decarboxylase is absent in E. coli K-12 (Applebaum et al., 1977). The evolutionary relationship of these three amino acid decarboxylases has been discussed based on similarities and differences in physical and chemical properties and in distribution (Applebaum et al., 1975), and on available sequence information (Moore and Boyle, 1990).

Biosynthetic arginine decarboxylase (Moore and Boyle, 1990) shares no homology with biosynthetic ornithine decarboxylase (S.M. Boyle, GenBank accession no. M33766) or biodegradative lysine decarboxylase, and sufficient physical differences are known even between the two forms of arginine decarboxylases. On the other hand, the amino acid sequence of biosynthetic ornithine decarboxylase bears a strong resemblance to biodegradative lysine decarboxylase, and provides further evidence that biosynthetic ornithine decarboxylase and biodegradative arginine, lysine and ornithine decarboxylases share a common origin (Applebaum et al., 1975). Intriguingly, biodegradative arginine decarboxylase has a subunit structure more similar to biodegradative lysine decarboxylase than to biosynthetic arginine decarboxylase, which might indicate the functional similarity between the biodegradative enzymes. The high degree of similarity between the pH inducible lysine decarboxylase of E. coli and the pH non-inducible lysine decarboxylase of H. alvei might be explained by a loss of specific controls in the H. alvei case. The
Table 7. Comparison of the physical and chemical properties of arginine-, lysine-, and ornithine decarboxylases of *E. coli*.

<table>
<thead>
<tr>
<th></th>
<th>Optimum pH</th>
<th>Active Form at pHopt</th>
<th>Subunit Mol. Wt. (kd)</th>
<th>Km (mM)</th>
<th>Coenzyme / Subunit</th>
<th>Turnover No. (nmol/min/subunit)</th>
<th>Coding Gene (min)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biosynthetic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADC</td>
<td>8.4</td>
<td>Tetramer</td>
<td>74</td>
<td>0.03</td>
<td>1PLP</td>
<td>1.12</td>
<td>speA 63.5</td>
<td>Wu and Morris, 1973.</td>
</tr>
<tr>
<td>ODC</td>
<td>8.1</td>
<td>Dimmer</td>
<td>81.6</td>
<td>5.6</td>
<td>1PLP</td>
<td>9.6</td>
<td>speC 64</td>
<td>Applebaum <em>et al.</em>, 1977</td>
</tr>
<tr>
<td><strong>Biodegradative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADC</td>
<td>5.2</td>
<td>Decamer</td>
<td>82</td>
<td>0.65</td>
<td>1PLP</td>
<td>32.2</td>
<td>adi 93.4</td>
<td>Nowak and Boeker, 1981.</td>
</tr>
<tr>
<td>LDC</td>
<td>5.7</td>
<td>Decamer</td>
<td>81.2</td>
<td>1.5</td>
<td>1PLP</td>
<td>80</td>
<td>cadA 93.7</td>
<td>Sabo <em>et al.</em>, 1974.</td>
</tr>
<tr>
<td>ODC&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.9</td>
<td>Dimer</td>
<td>80</td>
<td>3.6</td>
<td>1PLP</td>
<td>10.4</td>
<td></td>
<td>Applebaum <em>et al.</em>, 1975</td>
</tr>
<tr>
<td>GDC</td>
<td>3.8</td>
<td>Hexamer</td>
<td>50</td>
<td>0.54</td>
<td>1PLP</td>
<td>5.5</td>
<td>gadS 82</td>
<td>Straubach and Fischer, 1970</td>
</tr>
</tbody>
</table>

<sup>a</sup> Biodegradative ODC does not exist in *E. coli* K-12.
non-inducible lysine decarboxylase may substitute for arginine decarboxylase as an component in polyamine synthesis in *H. alvei* as the role ascribed to the putative biosynthetic lysine decarboxylase in *E. coli* (Goldemberg, 1980). One possible reason why lysine decarboxylase is constitutively expressed in *H. alvei* may be the absence of a biosynthetic arginine decarboxylase (Sakazaki, 1981). Recent sequence data from a plasmid bearing the lysine decarboxylase gene of *Bacillus subtilis* indicates good homology to that of *E. coli* (H. Hemila, personal communication).

The existence of a second lysine decarboxylase has been observed in *E. coli* and mutants with reduced lysine decarboxylase activity have been isolated from an *E. coli* polyamine auxotroph (Goldemberg, 1980). The expression of the putative lysine decarboxylase was not induced by low pH and anaerobic conditions, and the inhibitory effect of putrescine and spermidine on the enzyme argued against the possibility that the low activity might represent a basal level of the inducible lysine decarboxylase (Wertheimer and Leifer, 1983). It was suggested that the cadaverine produced by the biosynthetic lysine decarboxylase might partially substitute for some of the functions of putrescine or spermidine (Goldemberg, 1980). Neither has the locus of the second gene been mapped nor has the purified protein been reported. Our preliminary Southern hybridization experiment using *cadA* to probe *E. coli* chromosomal DNA failed to identify a second region homologous to *cadA*. One simple explanation is that the genes encoding the two lysine decarboxylases have little similarity.

Concerning the enzymatic function of lysine decarboxylase, it was early recognized that it required a cofactor, pyridoxial 5'-phosphate (Gale, 1946; Sabo and Fischer, 1974). The pyridoxial 5'-phosphate enzymes have been recently reviewed (Hayashi et al., 1990). Besides lysine (the attachment binding site of pyridoxal 5'-phosphate) and the adjacent histidine residues, histidine decarboxylase and inducible arginine, lysine and
glutamic acid decarboxylases all have the conserved sequence Ser-X-His-Lys, while biosynthetic arginine decarboxylase is an exception as it does not have a serine residue (Table 8). The recent observation of pyrroloquinoline quinone in glutamic acid decarboxylase of *E. coli* and pig kidney aromatic-L-amino acid decarboxylase (Van der Meer, et al., 1989) raises the possibility that other decarboxylases may also contain this molecule.
Table 8. Comparison of amino acid sequences of the pyridoxal 5'-phosphate binding site of amino acid decarboxylases in bacteria.

<table>
<thead>
<tr>
<th>Peptide Sequence of Pyridoxal 5'-Phosphate Binding Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDC&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Biosynthetic</td>
</tr>
<tr>
<td>ADC</td>
</tr>
<tr>
<td>ODC</td>
</tr>
<tr>
<td>Biodegradative</td>
</tr>
<tr>
<td>ADC</td>
</tr>
<tr>
<td>LDC</td>
</tr>
<tr>
<td>ODC</td>
</tr>
<tr>
<td>GDC</td>
</tr>
</tbody>
</table>

<sup>a</sup> Sequence is from *Morganella morganii*. 
Characterization of the *cadB* Gene

INTRODUCTION

Membrane associated proteins

The bacterial cell membrane is a dynamic structure exercising many functions critical to a living organism. Although membrane proteins (aside from the structural proteins) only constitute a small fraction of the total cellular protein, the importance of membrane proteins is certainly no less than cytosolic proteins. The outer membrane serves to protect the cytoplasmic membrane and mediate diffusion processes. Major outer membrane proteins include *murcin lipoprotein* whose main function is maintaining membrane structure, *porins* (encoded by *ompF, ompC, and phoE*) producing relatively nonspecific pores or channels that allow passage of small hydrophilic molecules, and proteins involved in specific diffusion processes (e.g. *LamB*). Many of cytoplasmic membrane proteins are involved in *bioenergetic* and *biosynthetic* reactions, whereas other proteins are involved in *transport* of specific solutes.

Membrane protein topology

The relation between the amino acid sequence and the three-dimensional structure of proteins is a central problem in molecular biology. The elucidation of the mechanism of a functional protein often depends on good understanding of the structure of the protein. From the studies of a small number of proteins with known folding pattern, it implies that folding patterns arise from the intrinsic general properties of the polypeptide chains and of the secondary structures. The general principles that govern protein folding patterns has been recently reviewed by Chothia (1990). However, the folding patterns of proteins are influenced not only by their chemical structure but also by the environment. The geometry of inter and intra-molecular bonds which differs between proteins in hydrophilic and hydrophobic environments and dictates the secondary structure of the
molecule. Therefore the current empirical secondary structure prediction methods based on aqueous-soluble protein folding motifs (Chou and Fasman, 1974; Garnier et al., 1978) may not provide good prediction for membrane proteins. In fact, Wallace (1988) found a poor correlation between the predicted and actual secondary structures of membrane proteins. Most models proposed for membrane transport proteins involve multiple membrane-spanning α-helices that pack together to form a hydrophilic pore which is lined by polar faces of several of the membrane-spanning helices. In this model, a segment of 22 amino acid is the minimum to form an α-helix that spans a lipid bilayer. The helix could be largely or entirely composed of hydrophobic residues or it could be amphipathic with the nonpolar side interacting with the hydrophobic fatty acid side chains. But Lodish (1988) pointed out that for a membrane protein arranged in multiple layers, the segment of the inner core which is not in contact with lipid bilayer does not need to be a α-helix and may be shorter than 22 amino acids. The outer helices would serve simply to anchor the protein in place with the membrane, while the inner shorter domains actually govern the transport process. It was reported in a recent review (Nilsson and von Heijne, 1990) that positively charged residues can act as topological determinants both in signal peptides and in integral membrane proteins. In contrast, negatively charged residues are less potent topological determinants. The "positive inside" rule suggests that positively charged residues could lock the end of a hydrophobic domain to the cytoplasmic side of the membrane. An explanation is that all charged residues provide a barrier to translocation, with positively charged ones being more difficult to translocate than negatively charged ones. Positive charge at the N-terminal side of an apolar segment promotes translocation of the downstream region, whereas positive charge at the C-terminal side tends to prevent translocation.
Alkaline phosphatase (PhoA) fusions

The use of alkaline phosphatase fusions to determine membrane protein subcellular location has been recently reviewed (Manoil et al., 1990). The phoA gene which encodes E. coli periplasmic alkaline phosphatase is enzymatically active only when it is exported across the cytoplasmic membrane into the periplasmic space and it usually exists as a dimer. The inactivity of cytoplasmic alkaline phosphatase appears to be due to the absence of essential intrachain disulfide bonds. Experimental results indicate that the sequence N-terminal to a site in a membrane protein are often sufficient enough to decide its cytoplasmic or periplasmic location. By measuring the enzymatic activity of phoA fusions to different positions, the cytoplasmic and the periplasmic domains of a transmembrane protein can be decided. On the other hand, β-galactosidase is most active in the cytoplasm, therefore the combined use of alkaline phosphatase and β-galactosidase fusions provides positive enzymatic signals for both periplasmic and cytoplasmic domains in the analysis of membrane protein topology (Manoil, 1990).

arc operon of Pseudomonas aeruginosa

The arginine deiminase pathway (ADI) (Fig. 12), one of several pathways of arginine catabolism in Pseudomonas spp., uses arginine as an energy source for mobility and slow growth under anaerobic, nitrate-free conditions (Vander Wauven et al., 1984). This pathway is induced under conditions of limited oxygen, depletion of the carbon energy source, and without an alternative electron acceptor for anaerobic respiration. The three enzymes involved in the ADI pathway, arginine deiminase, catabolic ornithine carbamoyltransferase and carbamate kinase are encoded by the arcABC genes. A transmembrane protein is encoded by the arcD gene which is located 5’ to arcABC genes and cotranscribed in the same unit (Luthi et al., 1990). The transcription initiation site
Figure 12. The *arc* operon and the arginine deiminase pathway of *P. aeruginosa*. P, promoter of the *arc* operon; CM, cytoplasmic membrane; cOTC, catabolic carbamoyltransferase; CK, carbamate kinase. H, *Hind*III; B, *Bam*HI; Bg, *Bgl*II; S, *Sca*I; Sm, *Sma*I (Gamper et al., 1991).
has been determined, and an ANR (a regulatory protein required for anaerobic growth on nitrate medium) consensus sequence (TTGACNNNNATCAG) similar to FNR recognition site (TTGATNNNNATCAA) in *E. coli* was found about 40 bp upstream from the transcription start site (Gamper et al., 1991). Anaerobic induction of the *arc* operon occurs at the level of transcription and requires the ANR box in *cis* and the ANR protein in *trans*. In *Streptococcus lactis*, the arginine uptake and ornithine release mediated by an arginine: ornithine antiporter has been well studied (Poolman et al., 1987). Since arginine is also taken in with concomitant release of an equal amount of ornithine in *P. aeruginosa* and *arcD* mutants lose the ability to use extracellular arginine as an energy source, ArcD may well be an arginine: ornithine antiporter or at least part of it (Luthi et al., 1990).

**RESULTS AND DISCUSSION**

**Identification of *cadB* gene**

The finding that the pH regulated promoter of *cadA* could not be detected within a 600 bp upstream from *cadA* by earlier primer extension studies and by promoter cloning experiments led to the suspicion that *cadA* might reside in a multigene operon. Indeed, sequencing of the upstream region revealed another ORF designated as *cadB*. This ORF encodes a polypeptide of 444 amino acid residues (Fig. 5) mostly consisting of neutral amino acid residues (Fig. 13).

**Subcloning of *cadB* gene**

To identify its gene product and understand its physiological role, a PCR fragment containing the *cadB* gene was amplified, and cloned into a multicopy expression plasmid pPV9+ 700 (Vermersch, 1988); the resulting plasmid was designated as pCADB (Fig. 14). Plasmid pPV9+700 contains an *NcoI* cloning site downstream of the *tac* promoter
Figure 13. Amino acid composition of CadB. The distribution of acid (A) and basic (B) residues in CadB is shown in the figure at top. The number and percentage of each amino acid in CadB are also listed.
444 Amino Acids  

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>n(%)</th>
<th>MW</th>
<th>MW(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>59</td>
<td>13.3</td>
<td>4191</td>
<td>9.0</td>
</tr>
<tr>
<td>Cys</td>
<td>8</td>
<td>1.8</td>
<td>824</td>
<td>1.8</td>
</tr>
<tr>
<td>Asp</td>
<td>9</td>
<td>2.0</td>
<td>1035</td>
<td>2.2</td>
</tr>
<tr>
<td>Glu</td>
<td>7</td>
<td>1.6</td>
<td>903</td>
<td>1.9</td>
</tr>
<tr>
<td>Phe</td>
<td>22</td>
<td>5.0</td>
<td>3235</td>
<td>6.9</td>
</tr>
<tr>
<td>Gly</td>
<td>45</td>
<td>10.1</td>
<td>2565</td>
<td>5.5</td>
</tr>
<tr>
<td>His</td>
<td>7</td>
<td>1.6</td>
<td>959</td>
<td>2.1</td>
</tr>
<tr>
<td>Ile</td>
<td>38</td>
<td>8.6</td>
<td>4297</td>
<td>9.2</td>
</tr>
<tr>
<td>Lys</td>
<td>12</td>
<td>2.7</td>
<td>1537</td>
<td>3.3</td>
</tr>
<tr>
<td>Leu</td>
<td>45</td>
<td>10.1</td>
<td>5088</td>
<td>10.9</td>
</tr>
<tr>
<td>Met</td>
<td>20</td>
<td>4.5</td>
<td>2620</td>
<td>5.6</td>
</tr>
<tr>
<td>Asn</td>
<td>19</td>
<td>4.3</td>
<td>2166</td>
<td>4.6</td>
</tr>
<tr>
<td>Pro</td>
<td>16</td>
<td>3.6</td>
<td>1552</td>
<td>3.3</td>
</tr>
<tr>
<td>Gln</td>
<td>6</td>
<td>1.4</td>
<td>768</td>
<td>1.6</td>
</tr>
<tr>
<td>Arg</td>
<td>8</td>
<td>1.8</td>
<td>1248</td>
<td>2.7</td>
</tr>
<tr>
<td>Ser</td>
<td>34</td>
<td>7.7</td>
<td>2959</td>
<td>6.3</td>
</tr>
<tr>
<td>Thr</td>
<td>28</td>
<td>6.3</td>
<td>2829</td>
<td>6.1</td>
</tr>
<tr>
<td>Val</td>
<td>37</td>
<td>8.3</td>
<td>3665</td>
<td>7.9</td>
</tr>
<tr>
<td>Trp</td>
<td>11</td>
<td>2.5</td>
<td>2046</td>
<td>4.4</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tyr</td>
<td>13</td>
<td>2.9</td>
<td>2119</td>
<td>4.5</td>
</tr>
<tr>
<td>Z</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 14. Construction of pCADB. Two primers (primers cadB+ and cadB-) flanking the cadB coding sequence were designed to contain unique restriction site (NcoI or XhoI) for facile insertion into the protein expression vector pPV9+700. Since the NcoI site (CCATGG) in the forward primer gives a cohesive end CATGG right in front of the ATG start codon of cadB (CATGA), the amplified cadB gene is now under control of the tac promoter without influence of its own upstream regulatory sequence. The single mismatch in the second codon did not seem to affect the function of cadB. The replication origin is from pBR322. Ap, Ampicillin resistance; DHFR, dihydrofolate reductase.
and a proper spaced SD sequence, therefore it is possible to clone a structural gene right from its ATG start codon. The amplified fragment flanked by primer cadB+ and primer cadB-, with only one base pair change in the second codon, was digested by restriction enzymes and ligated to the NcoI and XhoI sites of pPV9+700.

Expression of cadB with maxicells

Transformation, isolation and identification of the pCADB construct were successfully accomplished in strain 71-18. However, pCADB caused extremely slow growth (transformant showed up only after 72 hours of incubation) when transformed into CSR603 for maxicell experiments. Slow growth of MC4100 harboring pCADB was also observed. The different effects of pCADB on different strains could be explained by the presence of a lacIQ gene in 71-18, which strongly repressed the expression of cadB from the tac promoter. This observation indicates that even in the absence of exogenous inducers (IPTG or lactose), the amount of CadB membrane protein expressed from a multicopy plasmid in a rich medium is sufficient enough to cause growth problems. Therefore, a lacIQ gene present on the F factor of strain XL-1 was transferred into CSR603 through conjugation by a mating experiment. Because CSR603 is rpsL and the F factor from XL-1 carries Tn10, it is possible to select CSR603F' on an LB plate with streptomycin and tetracycline. The lacIQ gene indeed rescued CSR603F' from growth inhibition by pCADB. The expression of cadB in maxicells has been described in Methods. Two bands appeared on an SDS-PAGE gel by maxicell analysis of pCADB despite high background labeling (Fig. 15). The unusual migration rate relative to the deduced molecular weight of CadB (~47 kd) suggests that this protein may have an atypical structure and this might correlate with the proposed membrane role of the protein as certain membrane proteins have been observed to migrate as significantly smaller proteins in similar gel systems. The much faster migration rate was also observed for the ArcD protein, however, only a single band was detected in a maxicell experiment (Luthi
Figure 15. Protein expression of cadB from pCADB by maxicells. $[^{35}S]$-labeled protein extracted from UV-irradiated maxicell CSR603F' bearing pCADB (lane 1) and pPV9+700 (lane 2) were separated on an SDS-polyacrylamide gel as described in Methods. The gene products of cadB are indicated by arrows and the positions of protein markers are shown. Two polypeptides expressed from pCADB but not present in pPV9+700 migrated significantly faster than the predicted molecular weight of CadB (~47 kd). The protein of the faster migrating band may be a processed or degraded form of the slower migrating protein. The positions of size markers are shown on the side.
et al., 1990). The faster migrating bend could be a processed or a degraded form of the slower migrating bend.

**Confirmation of CadB sequence**

In order to check for proper retention of the coding triplet phase of the *cadB* sequence, several in phase protein fusions were made at various points within *cadB* from very early in the gene to one fusion in the 3rd to last codon. These *cadB::lac* protein fusions were assayed for their pH responses and similar levels of induction were observed among the *cadB::lac* protein fusion constructs (see later section, Fig. 19). This supported the idea that the entire ORF was being translated in the phase shown and that the true size of the protein is ~47kDa. Its highly hydrophobic profile (Fig. 16, Kyte and Doolittle, 1982) suggested that CadB was a membrane associated protein.

**Homology between CadB of *Escherichia coli* and ArcD of *Pseudomonas aeruginosa***. In addition to some similarities between short stretches of CadB sequence with a few peptide sequences from a GenBank search, the most significant finding was the homology with ArcD of *P. aeruginosa*. Not only do they have comparable sizes (444 aa vs. 482 aa), but the similarity is observed throughout their sequences (Fig. 16 and Fig. 17). Although the homology was not reflected by using identical amino acid residues, a very close relatedness between CadB and ArcD was apparent. The arrangement of CadB and ArcD membrane spanning domains as revealed by Kyte and Doolittle hydropathy graphs strongly suggested a similar natural folding pattern for both CadB and ArcD (Fig. 16), and therefore they might perform a similar function or at least adopt a similar structure in their physiological roles.

**Secondary structure of CadB**. Although CadB and ArcD resemble each other dramatically in terms of hydrophathy profiles and sequence relatedness, no similarity could be detected in the Garnier and Robson secondary structure analysis (Fig. 18).
Figure 16. Hydrophobicity profiles of CadB of E. coli and ArcD of P. aeruginosa. The hydrophobicity of CadB and ArcD sequences is displayed as Kyte and Doolittle plots (1982). The higher the positive values, the more hydrophobic the region is. The distribution of acid (A) and basic (B) amino acids is compared to the respective hydropathy plot. With the help of phoA fusion technique, on which side of the membrane located each hydrophilic region can be determined.
Figure 17. Amino acid sequence homology between CadB of *E. coli* and ArcD of *P. aeruginosa*. The extensive similarity between CadB and ArcD is indicated by the straight line drawn with Lawrence homology domains dot matrix. Dots represent the same or conservatively related amino acid residues between CadB and ArcD.
Sequence Homology Between CadB and ArcD
Figure 18. Garnier and Robson secondary structure analyses of CadB and ArcD. The structure prediction of CadB (upper panel) and ArcD (lower panel) is shown. Abbreviations: C, random coil; T, reverse turn; S, β sheet; H, α helix.
CadB tends to have more extended β-sheet regions whereas ArcD has more α-helix structure. The difference may not be significant because of the intrinsic deficiency of current programs for prediction of membrane protein structure. The hydrophobic domains with enough length are usually considered as the transmembrane domains, however, the prediction of the topology of an integral membrane protein still relies on phoA protein fusion experiments and alkaline phosphatase assays. Whether certain regions are in the cytoplasmic or periplasmic side of the membrane determined with phoA fusions should match the results of lacZ fusions in a complementary way. Construction of phoA fusions to cadB have not be performed, however some cadB::lacZ fusions have been constructed, and the fusion positions with their corresponding β-galactosidase activities are shown in Fig. 19. Although lacZ fusion is not as definitive as phoA fusion in terms of localization of membrane protein domains, the activity of β-galactosidase in the cytoplasmic site is supposed to be higher than in the periplasmic site. Fusion to positions 12 and 441 seems to have higher β-galactosidase activities than others, however it is premature to suggest the orientation of CadB in membrane due to limited fusion points. The only indication here is that the C-terminal hydrophilic domain is probably retained in the cytoplasm.
Figure 19. Fusion positions and the corresponding β-galactosidase activities of cadB::lacZ protein fusions. The fusion positions are indicated by arrows. The fusion point either indicated by restriction site or synthesized primer (parenthesized) and the number of the last cadB codon in the fusion are bracketed. The expression level of lacZ gene at both inducing (pH 5.5) and non-inducing (pH 8) conditions are listed as β-galactosidase units.
<table>
<thead>
<tr>
<th>Protein fusion clone</th>
<th>Fusion position</th>
<th>β-Galactosidase</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSM552-345</td>
<td>6</td>
<td>540</td>
</tr>
<tr>
<td>pSM552-355</td>
<td>6</td>
<td>810</td>
</tr>
<tr>
<td>pSM552-303</td>
<td>6</td>
<td>790</td>
</tr>
<tr>
<td>pSM552-245</td>
<td>12</td>
<td>2000</td>
</tr>
<tr>
<td>pSM552-415</td>
<td>12</td>
<td>2230</td>
</tr>
<tr>
<td>pSM552-1-170</td>
<td>207</td>
<td>920</td>
</tr>
<tr>
<td>pSM552-1-370</td>
<td>313</td>
<td>650</td>
</tr>
<tr>
<td>pSM577-290</td>
<td>379</td>
<td>610</td>
</tr>
<tr>
<td>pSM552-315</td>
<td>379</td>
<td>580</td>
</tr>
<tr>
<td>pSM577-390</td>
<td>441</td>
<td>1170</td>
</tr>
</tbody>
</table>
Physiological Role of cad Operon

INTRODUCTION

Basic amino acid transport systems

The uptake of basic amino acids lysine, ornithine, and arginine is mediated by at least four transport systems in *E. coli* based on biphasic kinetics of entry and on competitive inhibition at low, but not high, substrate concentrations. Each basic amino acid is known to be transported by a specific and a general transport systems (Rosen, 1971).

Lysine specific system (LS). LS is a low affinity (Km= 10 \( \mu \)m) lysine specific (not inhibited by arginine or ornithine) transport system. It is composed of an osmotic shock resistant integral membrane protein using proton motive force as driving force for lysine uptake.

Arginine specific system (AS). AS is a low affinity osmotic shock sensitive, periplasmic binding protein dependent transport system, using ATP as energy source.

Arginine-ornithine specific system (AO). It was determined later that a common transport system existed for arginine and ornithine (Km = 10 \( \mu \)M for arginine) in *E. coli*. (Celis, 1981).

General lysine-arginine-ornithine system (LAO). LAO is a high affinity (Km for lysine = 0.5 \( \mu \)M, Km for ornithine= 1.4 \( \mu \)M), osmotic shock sensitive, periplasmic binding protein dependent system. The low affinity system of AO is sensitive to repression by either arginine or ornithine, whereas the high affinity of LAO system is only repressed by lysine (Celis et al., 1973). The binding protein of AS and
LAO systems have been purified. The arginine specific binding protein (Rosen, 1973b) and the lysine-ornithine specific permease have been purified (Rosen, 1971).

**Arginine and ornithine transport.** Although arginine is a substrate of the LAO binding protein, the transport of arginine is mainly accomplished by the AS system and not by the LAO system. Affinities of LAO binding protein for basic amino acids are in the order of ornithine > lysine > arginine (Rosen, 1971). Mutants resistant to canavanine (a toxic arginine analogue) include alterations in the regulation of arginine biosynthesis, altered arginyl-tRNA synthetase, and loss of transport ability for arginine, lysine and ornithine (Maas, 1972). Although canavanine is transported by only the AS system, the mutant resulting in resistance to cadavanine appears to affect two transport systems. Therefore, it was proposed that AS and LAO transport systems share a common energy coupling protein(s) past the level of the recognition and binding site of the two systems (Rosen, 1973a). Mutation in the \( \text{argP} \) gene (56 min) (Maas, 1972) reduces transport by all three systems (Celis et al., 1973) and a mutant that affects transport of arginine and ornithine but not other related transport systems maps to 62 min close to \( \text{argP} \) on the recalibrated linkage map (Celis, 1977). The arginine biosynthetic derepressed mutants \( \text{argR}^- \) (repressor of arginine biosynthesis) exert no effect on arginine transport, indicating regulation of arginine uptake and biosynthesis are not regulated in a concerted manner, however, exogenous arginine can repress both transport and biosynthesis. It might be an indication that some elements may be shared by the two regulatory systems (Celis, 1977b). The arginine-ornithine binding protein of AO system is encoded by \( \text{abpS} \) (Celis, 1982). Both phosphorylated and non-phosphorylated forms of the arginine-ornithine binding protein have been isolated from periplasmic fluid (Celis, 1984). The dissociation constant of the phosphorylated form for arginine is 5 \( \mu \text{M} \) compared to 0.1 \( \mu \text{M} \) of the unmodified form, suggesting the phosphorylated protein is
the active form which releases the ligand for translocation across the membrane. This also indicates that ATP is the energy source of the AO system (Celis, 1984)

**Lysine transport.** The genetic study of lysine transport started with isolation of lysine analogue thiosine (thiolysine, S-aminoethylcysteine) resistant mutants. The rate of decarboxylation of thiosine by lysine decarboxylase is 15 to 20 % of lysine. Thiosine inhibits growth of microorganisms in the absence of exogenous lysine due to replacement of lysine in protein synthesis and thus interference of enzyme activity. Thiosine resistant mutants have been found to consist of several types including those with an altered level of lysyl-tRNA synthetase, altered activity of aspartokinase III and aspartokinase-homoserine dehydrogenase activities (enzymes in lysine biosynthesis pathway), and inhibited lysine uptake by both LS and LAO systems (Popkin and Maas, 1980). Mutant strains that are resistant to thiosine and excrete lysine into the medium were isolated, and an impaired ability to take up lysine from the medium due to overproduction of lysine was observed (Halsall, 1975). The dual activity of the LS system is described in an unpublished paper by Rosen (1976) and indicates that in cells grown in minimal medium LS is not affected by osmotic shock but becomes sensitive in cells grown in enriched medium. An analysis of the energetics of the LS system indicates that the energy source of LS partially comes from an ATP-derived energy and partially from a membrane proton motive force in contrast to other transport systems which appear to be coupled to only one of the two sources of energy. Thiosine resistant mutants *lysP (cadR)* occur spontaneously in minimal medium cultures at high frequency, and some of them not only show lowered level of lysine uptake by both LS and LAO systems but also show derepressed level of lysine decarboxylase but not other basic amino acid decarboxylases. Besides the above phenotypes, the *lysP* lesion also has lowered sensitivity of lysine transport to repression by lysine and more rapid excretion of endogenously synthesized lysine (Popkin and Maas, 1980). Despite the derepressed level of lysine decarboxylase
in exponential cultures of the mutant, extracts of these cultures had no detectable cadaverine pool, although cadaverine was present in cells grown to stationary phase (Popkin and Maas, 1980). Because of the mutants spontaneous nature and high frequency of occurrence, the pleiotropic effects of the lysP lesion are thought to be caused by a single mutation. However, it is possible this is a deletion mutation or a polar nonsense mutation that inactivates several adjacent genes. A model of a single regulatory protein encoded by lysP which controls lysine transport and lysine decarboxylase was proposed based on the observation that lysine transport is repressible by lysine whereas lysine decarboxylase is inducible by lysine (Popkin and Maas, 1980). In this model, the regulatory protein, when bound to lysine, serves as an active repressor for lysine transport and at the same time an inactivated repressor for the expression of lysine decarboxylase. A mutation that changes the affinity to lysine can result in altered regulation by lysine. A mutant with high affinity to lysine might explain the phenotype of lysP mutant discussed above. During the isolation of mutants deficient in cadaverine biosynthesis, two types of mutants were found: one is the lysine decarboxylase gene cadA mutant, the other is a regulatory gene for lysine decarboxylase cadR (46 min, maps close to lysP) (Tabor, et al., 1980). The characteristics of cadR mutants are similar to lysP mutants, and thus they probably carry a mutation in the same locus or region.

**Basic amino acid transport in other bacteria**

Many lactic acid bacteria contain an arginine catabolism specific transport system, the arginine:ornithine antiporter that mediates electroneutral exchange between exogenous arginine and intracellular ornithine produced from degradation of arginine by arginine deiminase pathway (Poolman et al., 1987). High rates of arginine metabolism have been reported, suggesting that arginine is translocated into cells via a transport system with a large capacity. One mole of ATP is synthesized by converting 1 mole of arginine into ornithine. The stoichiometric exchange is tightly regulated. In the absence of arginine,
Streptococcus lactis cells maintain a high intracellular concentration of ornithine (Poolman et al., 1987). In studies of membrane vesicles without arginine antiport activity, of the three basic amino acids lysine, arginine, and ornithine, only lysine uptake in Lactococcus lactis is accumulated to a significant level mediated by a specific PMF-dependent system, but a separate PMF-driven arginine uptake system is absent (Driessen et al., 1989). This lysine uptake seems to be a H⁺: lysine symport. A putative lysine-arginine shuttle consisting of the arginine ornithine antiport, the H⁺: lysine symport and H⁺/ATP pump in L. lactis was proposed (Fig. 20, Driessen et al., 1989). A variety of transport systems have been reported for lysine uptake in different bacteria. These include LS system and ATP-dependent LAO system in E. coli (Rosen, 1971 and 1976), proton: lysine symport in Lactococcus lactis (Driessen et al., 1989), and a counterflow process mediated by arginine: ornithine antiport in L. lactis (Driessen et al., 1989). A H⁺symport was also suggested for lysine uptake in Corynebacterium glutamicum (Luntz et al., 1986). And an antiport system for lysine was demonstrated in C. glutamicum which could be both electroneutral (homologous lysine-lysine exchange) and electrogenic (Broer and Kramer, 1990). The homologous lysine-lysine exchange observed in resting cells is insensitive to uncouplers and ionophores. This is an obligate antiport system and cannot be explained by exchange diffusion. On the other hand the antiport of lysine against some neutral amino acids inevitably causes charge movement and can change or respond to the membrane potential.

Polyamine content of Escherichia coli

The physiological roles of polyamines have been implicated in a wide variety of biological functions including biosynthesis of nucleic acid and protein (Tabor and Tabor, 1984). The biosynthetic pathways of putrescine (1,4-diaminobutane), spermidine (1-aminopropyl-1,4-diaminobutane) and cadaverine (1,5-diaminopantane) are well known in E. coli. Putrescine is the most abundant polyamine in E. coli, with spermidine as the
Figure 20. Schematic representation of a lysine:arginine shuttle in *L. lactis* subsp. *lactis* which leads to accumulation of arginine. Arginine can enter the cell in exchange for ornithine (a) or lysine (c) via the arginine:ornithine or the arginine:lysine antiporter. Ornithine is formed during arginine degradation by arginine deiminase pathway, which yields ATP. ATP can be used for the generation of a PMF via the F$_1$F$_0$ ATPase (d). The PMF functions as a driving force for the accumulation of lysine via the H$^+$: lysine symport (b) (Driessen et al., 1989).
second. There are 3 mg (34.1 μmol) of putrescine and 1 mg (7 μmol) of spermidine per gram of E. coli B dried cells (Morris and Jorstad, 1970). The intracellular concentration of putrescine but not spermidine is a function of the osmotic strength of the medium. As the osmotic strength of the medium increases, the intracellular concentration of putrescine drops. High osmolarity could cause the excretion of putrescine to the medium. It was suggested that excretion might occur via a K⁺: putrescine antiport. Since K⁺ is the primary intracellular balancer of extracellular osmotic strength, exchange of two positive charge for one positive charge of K⁺ could maintain the internal ionic strength when external osmotic strength increases (Munro et al., 1972). An energy dependent polyamine uptake system has been shown in E. coli (Tabor and Tabor, 1966), and polyamine transport appeared to be unidirectional, and dependent on proton motive force (Kashiiwagi et al., 1986). The amount of polyamine content in E. coli was studied by using wild type or strains producing excessive amounts of enzymes for polyamine synthesis; the amount of intracellular putrescine was adjusted by expression of excessive amount of putrescine to the medium (Kashiwagi and Igarashi, 1988). Under conditions of putrescine starvation, mutants unable to synthesize this compound produce detectable cadaverine (Hafner et al., 1979), and exogenous cadaverine stimulates the growth of a putrescine deficient mutant (Dion and Cohen, 1972b), suggesting that cadaverine could act as a substitute for other diamines. Polyamine auxotrophs blocked in the synthesis of putrescine accumulate cadaverine in the absence of exogenous lysine (Dion and Cohen, 1972a). Although it has been stated that the cadaverine generated from decarboxylation of L-lysine is not further metabolized for energy production but is excreted out of the cell as part of a countermeasure to acid environment (Sabo et al, 1974), no direct detailed analysis of this aspect has been reported.
Arginine: ornithine antiport in *Streptococcus lactis*

Galactose-arginine-grown cells of *S. lactis* maintained a high intracellular concentration of ornithine (10 to 15 mM) for several hours when suspended in buffer without an exogenous energy source. In the presence of arginine, a simultaneous uptake of arginine and ornithine excretion was observed in *S. lactis*. This concurrence could be disrupted by removing extracellular arginine from the medium. The addition of arginine to the resting cell suspension resulted in a rapid alkalinization of the medium by ammonia for about 15 sec, followed by a much slower rate of alkalinization owing to steady-state activity of arginine deiminase pathway activity, until all the arginine was consumed (Poolman et al., 1987). The arginine: ornithine exchange occurs in an 1 to 1 ratio and is independent of the proton motive force. The driving force of arginine uptake in intact cells is provided by the concentration gradients of arginine and ornithine formed during arginine catabolism in the cell (Poolman et al., 1987). Similar phenomena were also observed in membrane transport studies with membrane vesicles (Driessen et al., 1987).

**RESULTS**

**Cadaverine excretion**

Upon full induction of the *cad* operon, large quantities of lysine decarboxylase are produced and if high amounts of lysine are present in the media, one would expect that significant amounts of cadaverine should be formed under the induced condition since no other regular pathways to further metabolize cadaverine are present in *E coli*, and excessive levels of polyamines can be toxic (Tabor and Tabor, 1985). Cadaverine was speculated to be excreted from the cell based on some indirect evidence (Hanka and Koessler, 1924), and cadaverine has been isolated from the media (Gale, 1940). The discovery of the *cadB* gene ignited our interest to test for release of cadaverine using a colorimetric assay originally designed for lysine decarboxylase activity assay (Phan et al.,
1982). Two Mu lac fusion strains GNB581 (cadA::lac Ap\(^{r}\)) and GNB8385 (cadB::lac Ap\(^{r}\)) have phenotypes of CadB\(^{+}\)CadA\(^{-}\) and CadB\(^{-}\)CadA\(^{+}\) respectively. Together with MC4100 (CadB\(^{+}\)CadA\(^{+}\)) and in combination with plasmids pCADA and pCADB, various phenotypes of CadB and CadA could be manipulated. For the convenience of drug selection, GNB8385K (cadB::lac Km\(^{r}\)) was used to harbor plasmid pCADB (Ap\(^{r}\)). The strategy to overcome growth inhibition by induced pCADB on CSR603 in maxicell experiments was implemented again by introducing the lacI\(^Q\) gene on the F factor into respective strains through mating. The summary of the results of cadaverine measurements are depicted in Table 9. Cells were grown at different pHs to an OD\(_{600nm}\) of 0.3, then IPTG was added to induced expression of plasmid-borne cadA and cadB genes; the cadaverine contents of the media were measured after the cell density reached an OD\(_{600nm}\) of 0.65 to 0.7. A large amount of cadaverine was detected in the medium at pH 5.5 but not at pH 8 nor in the absence of lysine decarboxylase. The IPTG induced expression gave comparable levels of cadaverine as with pH induction (GNB581F\(^{r}\)/pCADA vs. MC4100F\(^{r}\)). The expression of cadA was a prerequisite for the source of cadaverine. However, in the absence of CadB (GNB8385F\(^{r}\)/pCADA vs. MC4100F\(^{r}\) or GNB581F\(^{r}\)/pCADA), the amount of cadaverine detected in the medium was significantly reduced. To ensure the specificity of the experiment, [U-\(^{14}\)C]-L-lysine was used as the substrate and radioactive cadaverine was measured after an incubation time of 10 min (Table 10). The results also showed a higher excretion of labeled cadaverine in the presence of CadB. CadB appeared to facilitate the excretion of cadaverine at pH 5.5, but the effect of CadB was less conspicuous at pH 8. Since CadR\(^{-}\) mutants exhibit a large pH induction (with or without added lysine) which is the same as the pH induction of CadR\(^{+}\) strains only in the presence of excess lysine, a test of CadR\(^{-}\) on cadaverine excretion was performed. CadR\(^{-}\) strain GJT001F\(^{r}\) showed the same cadaverine excretion pattern as its CadR\(^{+}\) counterpart MC4100F\(^{r}\) when cells were
Table 9. Cadaverine excretion assay I.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Plasmid</th>
<th>A.</th>
<th>B.</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH5.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC4100F'</td>
<td></td>
<td>10.0</td>
<td>100</td>
<td>CadB⁺CadA⁺</td>
</tr>
<tr>
<td>GNB581F'</td>
<td>PCADA</td>
<td>8.2</td>
<td>82</td>
<td>CadB⁺CadA⁺</td>
</tr>
<tr>
<td>GNB581F'</td>
<td></td>
<td>0.1</td>
<td>1</td>
<td>CadB⁺CadA⁻</td>
</tr>
<tr>
<td>GNB8385KF'</td>
<td>pCADB</td>
<td>0.3</td>
<td>3</td>
<td>CadB⁺CadA⁻</td>
</tr>
<tr>
<td>GNB8385F'</td>
<td>pCADA</td>
<td>2.0</td>
<td>20</td>
<td>CadB⁻CadA⁺</td>
</tr>
<tr>
<td>GNB8385F'</td>
<td></td>
<td>0.04</td>
<td>0.4</td>
<td>CadB⁻CadA⁻</td>
</tr>
<tr>
<td>pH8.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC4100F'</td>
<td>pCADB, pCADA</td>
<td>1.3</td>
<td>13</td>
<td>CadB⁺CadA⁺</td>
</tr>
<tr>
<td>MC4100F'</td>
<td>pCADB</td>
<td>0.1</td>
<td>1</td>
<td>CadB⁺CadA⁻</td>
</tr>
<tr>
<td>MC4100F'</td>
<td>pCADA</td>
<td>0.8</td>
<td>8</td>
<td>CadB⁻CadA⁺</td>
</tr>
<tr>
<td>MC4100F'</td>
<td></td>
<td>0.08</td>
<td>0.8</td>
<td>CadB⁻CadA⁻</td>
</tr>
</tbody>
</table>

The growth of cells and the cadaverine excretion assay were described in Methods. Cells with or without plasmids were grown anaerobically in modified Falkow lysine decarboxylase medium at pH 5.5 or pH 8. Expression of plasmid-borne cadB or cadA genes was induced by IPTG and the amount of cadaverine excreted into the medium was determined by measuring the OD₃₄₀nm of TNP-cadaverine (column A). This value is presented in column B as the percentage of the cadaverine content in the growth medium for the CadB⁺CadA⁺strain listed unit as 100 %. The underlined phenotype indicates their phenotype was induced by IPTG as compared to the pH-induced phenotype.
Table 10. Cadaverine excretion assay II.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Plasmid</th>
<th>A.</th>
<th>B.</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH5.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC4100F'</td>
<td></td>
<td>3900</td>
<td>100</td>
<td>CadB⁺CadA⁺</td>
</tr>
<tr>
<td>GNB581F'</td>
<td></td>
<td>920</td>
<td>24</td>
<td>CadB⁺CadA⁻</td>
</tr>
<tr>
<td>GNB8385KF'</td>
<td>pCADB</td>
<td>570</td>
<td>15</td>
<td>CadB⁺CadA⁻</td>
</tr>
<tr>
<td>GNB8385F'</td>
<td>pCADA</td>
<td>1400</td>
<td>37</td>
<td>CadB⁻CadA⁺</td>
</tr>
<tr>
<td>GNB8385F'</td>
<td></td>
<td>800</td>
<td>21</td>
<td>CadB⁻CadA⁻</td>
</tr>
</tbody>
</table>

The growth of cells and the cadaverine excretion assay were described in Methods. Cells with or without plasmids were grown anaerobically in modified Falkow lysine decarboxylase medium at pH 5.5. Expression of plasmid-borne cadB or cadA genes was induced by IPTG and the amount of cadaverine excreted into the medium was determined by measuring the radioactivity of TNP-[¹⁴C]-cadaverine (column A). This value is presented in column B as the percentage of the [¹⁴C]-cadaverine content in the growth medium for the CadB⁺CadA⁺ strain listed unit as 100%. The underlined phenotype indicates their phenotype was induced by IPTG as compared to the pH-induced phenotype.
grown in modified Falkow medium at pH 5.5 (Table 11), and CadR phenotypes did not affect the amount of cadaverine excreted with or without the presence of lysine (Table 12), indicating cadR locus had no effect on cadaverine excretion

**DISCUSSION**

The control of intracellular pH in bacteria is complex and involves coordination of the activities of a number of ion transport systems (Booth, 1985). *E. coli* can grow with little difference in generation time over a pHo range of 5.5-8.5 (Gale and Epps, 1942). The intracellular pH of *E. coli* cells is tightly controlled and the internal pH is nearly constant at pH 7.4 to pH 7.8 (Zilberstein et al., 1984; Slonczewski et al. 1981). A rapid shift to low pH transiently lowers the internal pH but it quickly recovers. Major factors required in maintaining pH homeostasis appear to be a source of energy, the H+/ATP pump and the ability to transport potassium ion, however the complicated interrelationships among transport processes should not be underestimated (Booth, 1985). After the discovery of pH induction of amino acid decarboxylases, their role in a pH response was considered. In support of their importance were the observation of reduced ability of arginine decarboxylase mutants to grow at low pH (Becker, 1967) and the decreased growth of histidine decarboxylase mutant of *Lactobacillus* at low pH (Rescei and Snell, 1972). Proposed models of how the action of amino acid decarboxylases can allow increased growth in acidic growth conditions have considered the production of CO₂ by the reaction, in light of requirement for CO₂ for growth (Neidhardt et al., 1974) and as a compensatory mechanism for the loss of CO₂ at low pH by chemical equilibrium (Buirad and Snell, 1964). Also considered has been the effect of the decarboxylase reaction on the intracellular pH in preventing over acidification (Gale, 1946). At low pH the increased permeability of H⁺, especially in the presence of the weak organic acids produced during fermentation presents a problem in maintaining the
Table 11. *cadR* effect on cadaverine excretion assay I.

<table>
<thead>
<tr>
<th>Strain</th>
<th>A</th>
<th>B</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH5.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC4100F'</td>
<td>10.3</td>
<td>100</td>
<td>CadB⁺CadA⁺CadR⁺</td>
</tr>
<tr>
<td>GJT001F'</td>
<td>11.5</td>
<td>110</td>
<td>CadB⁺CadA⁺CadR⁻</td>
</tr>
<tr>
<td>GJT002F'/pCADB</td>
<td>0.4</td>
<td>4</td>
<td>CadB⁺CadA⁻CadR⁻</td>
</tr>
<tr>
<td>GJT002F'</td>
<td>0.4</td>
<td>4</td>
<td>CadB⁻CadA⁻CadR⁻</td>
</tr>
</tbody>
</table>

Bacterial strains were grown anaerobically in modified Falkow lysine decarboxylase medium at pH 5.5. The measure of cadaverine content in the culture has been described in the previous experiments. The amount of cadaverine excreted into the medium was determined by measuring the OD₃₄₀nm of TNP-cadaverine (column A). This value is presented in column B as the percentage of the cadaverine content in the growth medium for the CadB⁺CadA⁺strain listed as 100 %. No strains with phenotypes of CadB⁻CadA⁺CadR⁺ and CadB⁻CadA⁺cadR⁻ were available for assays.
Table 12. \textit{cadR} effect on cadaverine excretion assay II.

<table>
<thead>
<tr>
<th>Strain</th>
<th>$OD_{340nm}$ + Lysine</th>
<th>$OD_{340nm}$ - Lysine</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH5.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC4100</td>
<td>14.0</td>
<td>4.7</td>
<td>CadB$^+$CadA$^+$CadR$^+$</td>
</tr>
<tr>
<td>GJT001</td>
<td>15.3</td>
<td>4.7</td>
<td>CadB$^+$CadA$^+$CadR$^-$</td>
</tr>
</tbody>
</table>

Bacterial strains were grown anaerobically in CadR medium with or without lysine at pH 5.5. The measure of cadaverine content in the culture has been described in the previous experiments.
appropriate electrochemical gradients. This conjecture for the role of decarboxylases is based on the fact that during the decarboxylase reaction one $H^+$ is removed per molecule of lysine processed. Thus the reaction could eliminate a large quantity of $H^+$ ions from the cell. Another aspect of the action of decarboxylases is the effect on extracellular pH. The subsequent increase in the extracellular pH value due to the large scale removal of lysine and $H^+$, and the production of a polyamine will also act to reduce the problems associated with growth under too acidic conditions. Indeed, specific acid-base indicating media have been developed which allow the screening of bacterial strains for the presence or absence of inducible arginine, lysine or ornithine decarboxylases (Falkow, 1958). The rationale of the method is exactly based on the fact that decarboxylase positive strains could prevent the color change caused by excessive acidification from glucose fermentation. It has been confirmed that the cad operon is induced by extracellular low pH but is not affected by intracellular pH fluctuation (Auger et al., 1989; Slonczewski et al., 1987).

In the proposed model, lysine is taken in vigorously and turned over quickly to cadaverine and CO$_2$. The driving force of this movement would be further aided if the end products can be rapidly removed from the system. Carbon dioxide can freely escape through the cell membrane to the atmosphere and this property was used in screening and identification of cadR and cadA mutants (Tabor et al., 1980). The other product of this reaction, cadaverine, is not normally catabolized since there is no such pathway existing in E. coli. If the cell does not possess some kind of secretion device(s), the accumulation of cadaverine not only will disrupt the electrochemical equilibrium and eventually become toxic but also will slow the lysine decarboxylase reaction, and therefore prevent further processing of $H^+$ and lysine. The existence of membrane protein CadB helps solve this basic problem. While the excretion of polyamines has been studied with respect to $K^+$ and osmolarity (Munro and Sauerbrier, 1973), and polyamine
transport systems have been investigated (Kashiwagi and Igarashi, 1988; Munro et al., 1974), the relation of these studies to the pH response has not been thoroughly considered. An ability of CadB to not only facilitate the excretion of cadaverine but to act in such a way as to coordinate cadaverine excretion with lysine uptake makes an attractive model. In order to resolve this role for cadB, more detailed transport assays need to be undertaken. However, because of the similarity between CadB and ArcD of Pseudomonas aeruginosa, it is tempting to speculate that CadB might constitute a lysine: cadaverine antiporter or at least part of it.

A hypothetical working model for maintaining pH homeostasis by the CadB-CadA system is presented in Fig.21. In this case, amino acid decarboxylases would serve an ideal role for the efficient reduction of extracellular H⁺. The source of the H⁺ used in the decarboxylation could come from several routes: H⁺ accumulated inside the cell by proton permeability, brought in by weak acids or via other transporters. An alternative possibility is that H⁺ is brought in at the same time as lysine by the cadB system. Examples of concomitant uptake of H⁺ together with a specific amino acid are known (Anraku, 1978). The plausible explanation is that symport systems allow the accumulation of some amino acids and carbohydrates from the medium beyond the limitation of electrochemical equilibrium in uniport systems (Munro and Sauerbrier, 1973). The same strategy can also be adopted to raise the external pH, if the significance of H⁺ uptake is considered. It may seem fruitless to bring in extra proton ions while the cell is suffering from acid stress, but if the external low pH can be alleviated, the culture will benefit by extended growth. The H⁺ entry to the cell and the presence of an antiporter which couples uptake of substrate with release of the polyamine product, would comprise a complete system for feeding substrate to the decarboxylase while removing product.
Figure 21. Model for detoxification of extracellular high H\textsuperscript{+} concentration by the CadB-CadA system. Upon encountering low pH environment in a rich media, the cadBA operon is highly induced in E. coli cells. Membrane protein CadB functions as a lysine: cadaverine antiporter, bringing in lysine and excreting cadaverine, the end product of the CadA (lysine decarboxylase) enzymatic reaction. In order to balance the electrochemical charge inside the cell, CadB may take in H\textsuperscript{+} from the medium at the same time it imports lysine into the cell, leaving other mechanisms to cope with the strain on intracellular pH. Alternatively the action of lysine decarboxylase could remove H\textsuperscript{+} from the intracellular environment that enters by other routes or leakage under conditions of low pH\textsubscript{0}. This model would propose that CadB functions best when the medium contains both a high concentration of lysine and H\textsuperscript{+}, this condition together with the enzymatic reaction of CadA, would create an inward concentration gradient for lysine and H\textsuperscript{+} as well as an outward concentration gradient for cadaverine and CO\textsubscript{2}. The driving force for the lysine: cadaverine antiport can solely depend on the concentration gradients described above without other energy expense. In this way, the extracellular H\textsuperscript{+} entering the cell is neutralized inside the cell with minimum interference to cell metabolism.
Cytoplasm (Near Neutral pH)

\[ \text{Lysine} \quad \text{In Excess} \]

\[ +\text{NH}_3 \quad l \quad \text{C-COO}^- \quad l \quad (\text{CH}_2)_4 \quad l \quad +\text{NH}_3 \]

\[ \text{(LDC)} \quad \text{CadA} \quad \rightarrow \quad +\text{NH}_3 \quad l \quad (\text{CH}_2)_5 \quad + \text{CO}_2 \]

Cytoplasmic Membrane

CadB

\[ +\text{NH}_3 \quad l \quad \text{C-COO}^- \quad l \quad (\text{CH}_2)_4 \quad l \quad +\text{NH}_3 \]

Lysine
In Excess

Air

Cadaverine

Medium (Acid pH)
*P. aeruginosa* uses the arginine deiminase (ADI) pathway to convert arginine to ornithine and carbamoylphosphate which serves to generate ATP under anaerobic, nitrate-free conditions (Vander Wauwen et al., 1984). A similar pathway also exists in *Streptococcus* spp., in which the arginine: ornithine exchange has been well studied (Poolman et al., 1987a). An equivalent of arginine uptake and ornithine release should also be present in *P. aeruginosa*. The *arcD* gene which encodes an integral membrane protein is located in the same operon as other genes specifying the enzymes in the ADI pathway (Luthi et al., 1990). Since mutations in *arcD* rendered the cells unable to utilize extracellular arginine as an energy source, ArcD was proposed to be involved in an arginine: ornithine antiport (Luthi et al., 1990). Recent work with *Corynebacterium glutamicum* has led to the observation of a lysine-alanine exchange by an apparent antiport that can alter its exchange properties depending on changing metabolic conditions (Broer and Kramer, 1990). The speculation of CadB as an lysine: cadaverine antiport is based on the observation that a CadB- strain showed a significant decrease in cadaverine excretion even in the presence of CadA (lysine decarboxylase). The energy cost-free arginine: ornithine antiport has been shown to exist in bacteria, and the driving force solely depends on the opposite concentration gradients of arginine and ornithine across the membrane (Poolman et al., 1987). Similarly, the high lysine decarboxylase activity can also create a concentration gradient that would assist in providing the driving force to favor the antiport model. The attempt to detect cadaverine secretion from cells grown in minimal medium supplemented with 10 uM lysine (as opposed to over 43 mM in the modified Falkow medium) was unsuccessful even with both IPTG induced expression of *cadA* and *cadB* (data now shown). This might be explained as the necessity of a high concentration gradient at low pH for proper functioning of *cadB*. If the transporter behaves such that high external H+ concentration increases the uptake of lysine and
concomitant excretion of cadaverine, this would be very supportive of the coordinate role of the CadA CadB system in neutralizing H\textsuperscript{+} from the media.

Is there a specific lysine transport system only functioning for the purpose of raising the external pH or does the expression of cadB modify the usual transport mechanism to a more efficient and energy cost-free antiport? The relation of cadB to previously studied lysine transport systems of E. coli (Anraku, 1978; Popkin and Mass, 1980) is unclear. The only known locus affecting the expression of cadA (and cadB) is the cadR gene which has been mapped to 46 min (Tabor et al., 1980). The spontaneous CadR mutant strains were selected by resistance to SAEC (thiosine, S-aminoethylcysteine), and showed a lowered level of lysine transport and relieved a requirement for the presence of exogenous lysine for the maximal induction of cadA (Popkin and Mass, 1980). SAEC resistant strains have also been shown to excrete large amount of lysine into the medium (Halsall, 1974). However, cadR is not responsible for the pH and air regulation of cadA (Auger and Bennett, 1989). One possible explanation is that cadR is the pH-constitutive lysine transport system or regulates these systems which mainly provide lysine for protein synthesis, whereas cadB dominates the massive lysine uptake at low pH for the purpose of neutralizing H\textsuperscript{+}. This model is supported by the observation that normal lysine transport systems are repressible by lysine (Celis et al., 1973), contradicting the fact that large amount of lysine uptake occurs at low pH. The validity of this hypothesis awaits detailed analysis of the transport behavior of cadB and the elucidation of the sequence and function of the cadR locus.
Organization of the \textit{cad} Operon and its Neighboring Genes

INTRODUCTION

Genomic mapping of \textit{E. coli}

\textit{E. coli} is probably the most studied organism in genetics and molecular biology due to its relative simplicity of genomic organization. The \textit{E. coli} chromosome is divided into 100 min as determined by time of entry of markers in interrupted conjugation experiments. Each unit consists of approximately 47 kb of DNA (Bachmann, 1990). Various genetic methods and tools are available to decide the location of a locus in the linkage map; for a summary see Singer et al. (1989). A collection of strains with antibiotic selectable transposons located at defined positions in the \textit{E. coli} genome are available for Hfr mating and P1 cotransduction mapping (Singer et al., 1989). The latest linkage map of \textit{E. coli} K-12 was compiled by Bachmann (1990). The first physical map of \textit{E. coli} genome was decided by computer analysis of seven restriction sites of a \lambda phage library (Kohara, 1987). Subsequently a cosmid-derived physical map of \textit{E. coli} chromosome (Birkenbihl and Vielmetter, 1989) was compared to the map of Kohara. A computer software program for alignment of a restriction map to the genomic physical map is now available for quick physical mapping (Rudd et al., 1989 and 1991).

Genes located in the 93.2 to 94 min on \textit{E. coli} genome

Ten genes have been located in the 93.2 to 94 min on \textit{E. coli} genome. All of them have been sequenced. \textit{adi} (Stim, unpublished), \textit{melR} (Webster, et al., 1987), \textit{melAB} (Hanatani et al., 1884), \textit{fumB} (Guest et al., 1985; Bell et al, 1989), \textit{lysU} (Clark and Neidhardt, 1990; Le'Ve'que, et al., 1990), \textit{cadBA} (this work), \textit{cadC} (Watson et al., 1991), and \textit{pheU} (Prażkier et al., 1990).
RESULTS AND DISCUSSION

Chromosomal location of cad operon and its neighboring genes. Fig. 22 shows the relative positions of these genes by comparing their restriction enzyme maps with Kohara's E. coli genomic map (Kohara et al., 1987; Rudd et al., 1990). The gene lysU, also present on pLC4-5, has been sequenced (Clark and Neidhardt, 1990; Leveque et al., 1990). It is in the same orientation as cadA based on comparison of restriction maps. According to the restriction map, there is a space of about 2 kb between cadA and lysU. Immediately upstream from cadB, another ORF cadC was recently identified as a positive regulator of the cadBA operon (Watson et al., 1991). The gene lysU which encodes the thermostable lysyl-tRNA synthetase and cadA are both present on pLC4-5 and are transcribed in the same orientation. That lysU can also be induced under the same conditions as cadA (Hickey and Hirshfield, 1990) has prompted consideration of the hypothesis that they may belong to the same operon (Clark and Neidhardt, 1990). However, according to available information, these two genes are about 2 kb apart, and there is a sequence resembling a strong transcription terminator [ΔG of -26.6 kcal calculated with the model of Tinoco et al. (Tinico et al., 1973)] immediately downstream of cadA. The free energy of most hair-pin loop transcription terminators are in the range of 20-30 kcal (Von Hippel et al., 1984); therefore, it is unlikely that there is significant cotranscription of genes downstream of this terminator. However, because of the magnitude of expression of cadA, the possibility of some carry-over transcription to lysU region from cadA promoter can not be completely excluded.
Figure 22. Chromosomal localization of *cad* operon and its neighboring genes. The coordinates on the top line are positioned according to Kohara's map (1987). Minute calibration units follow that of Bachmann (1990). The restriction map is that provided by Kohara. Restriction enzyme site designation: B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; K, *Kpn*I; P, *Pst*I, R, *Eco*RV.
CHAPTER 4. REGULATION OF THE \textit{cad} OPERON

INTRODUCTION

Bacterial cells possess elegant regulatory mechanisms to alleviate stresses in harsh conditions and to take advantage of favorable environments. Many responses have been analyzed at the level of gene expression, and the hierarchy of global regulation has been established in many instances (Neidhardt, 1987). The situation with regard to regulation in response to extracellular low pH is just beginning to be investigated at the molecular level, although the phenomenon of induction of amino acid decarboxylases by low pH has been known for many years (Gale, 1946). Recent work has reported a number of proteins induced by pH shift in \textit{E. coli} (Heyde and Portalier, 1990; Hickey and Hirshfield, 1990) and \textit{S. typhimurium} (Foster and Aliabadi, 1989; Hickey and Hirshfield, 1990) and have compared the pattern to those of other stresses (Heyde and Portalier, 1990). Mu-\textit{lac} fusions (Silhavy and Beckwith, 1985) have been employed in analyzing many complicated regulatory systems and this approach has been used in pH regulatory studies with the observation of internal and external pH induced genes (Slonczewski et al., 1987), pH induction of fusions to the lysine decarboxylase and arginine decarboxylase genes (Auger et al., 1989) and the analysis of pH induced genes of \textit{S. typhimurium} (Aliabadi et al., 1988; Bingham et al., 1990; Foster and Aliabadi, 1989). The relation of pH homeostasis and the acid tolerance response of \textit{Salmonella typhimurium} has also been recently reported (Foster and Hall, 1990). More detailed information of bacterial pH responses are described in Chapter 1. In other systems pH regulation has been related to bacterial virulence. Examples are the response of \textit{Agrobacterium tumefaciens} in plant wounds (Winans, 1990) and \textit{Vibrio cholera} through the \textit{toxR} system (Miller et al., 1987).
Selected Topics of Bacterial virulence

Pathogenic bacteria have evolved regulatory systems capable of responding to various environmental signals that may be encountered during the infectious cycle. Signal transduction and coordinate regulation of the expression of genes and operons encoding virulence factors are a main theme in bacterial pathogenicity (Miller et al., 1989b).

*Salmonella typhimurium*. Virulent species of *Salmonella* can invade epithelial cells and survive within macrophages. A group of genes (*invABC* and *invD*) have been cloned that allow *S. typhimurium* to penetrate tissue culture cells (Galan and Curtiss, 1989). The osmoinducibility of an *invA-phoA* fusion strain is independent of *ompR* but is affected by gyrase inhibitors and in a *topA* mutant, indicating the expression of invasion genes is affected by changes in DNA supercoiling (Galan and Curtiss, 1990). The *phoP* locus mutants are defective in macrophage intracellular survival and show increased sensitivity to small cationic peptides from neutrophil granules known as defensins (Ganz et al., 1985). The *phoP* locus essential for virulence encodes two ORFs termed *phoP* and *phoQ*, that show similarity to other known sensor-regulator prototypes (Miller et al., 1989a). The *phoQ* gene product has a periplasmic domain by which signals within the macrophage may be sensed. Ten out of twenty amino acids in the predicted periplasmic domain of PhoQ are negatively charged. This ionic region may be involved in sensing intracellular signals (defensins, cations, low pH, etc.) unique to phagolysosome. Transcriptional and translational gene fusions generated by transposon insertions that require *phoP* and *phoQ* for expression have been mapped to different chromosomal locations. One of the fusion strains with a virulence defect has the insertion mutation in the *pagC* locus (*phoP*-activated genes), whereas the insertion in the *phoN* gene encoding an acid phosphatase retains the wild-type virulence. The gene
expression of the \textit{phoP} regulon (\textit{pagA}, \textit{pagB}, \textit{pagC}, \textit{psiD} and \textit{phoN}) has shown to be induced 2- to 3-fold by growth in media of low pH and \textit{phoN} expression is induced by limitation of carbon, nitrogen, and phosphorous (Miller and Mekalanos, 1990). These data suggest that starvation and low pH might be the signals recognized by PhoQ in the phagolysosome. A strain with a mutation in the \textit{phoP} locus yields constitutive expression of \textit{pag} loci in rich medium (phenotype PhoPc) and at the same time represses the synthesis of other proteins (\textit{prg} loci), suggesting that PhoP is able to function as a repressor as well as an activator (Miller and Mekalanos, 1990). It is also possible that a \textit{pag} locus encodes a transcriptional repressor of \textit{prg} loci and thus functions as a cascade or the \textit{pag} loci may encode proteins that are capable of causing degradation or modification of proteins and therefore change the protein expression pattern observed. Both PhoP$^-$ and PhoPc strains have attenuated virulence, suggesting that the ability to turn off the expression of certain genes encoding virulence factors is as important as turning on their expression.

\textit{Vibrio cholerae}. Expression of several genes whose products are associated with virulence is under coordinate control of the \textit{toxR} gene product. The most important virulence factors are cholera toxin production and the ability to adhere to and colonize in the small intestine of the host. Cholera toxin is a multimeric protein composed of \textit{A} and \textit{B} two kinds of subunits which are encoded by \textit{ctxAB} operon, and the colonization properties are associated with the production of fimbriae encoded by the \textit{tcp} locus. Several other ToxR regulated genes include \textit{ompT} and \textit{ompU} encoding two outermembrane proteins, toxin-coregulated pilus (\textit{tcp}), and accessory colonization factor (\textit{acf}) (Peterson and Mekalanos, 1988). Recently the \textit{aldA} gene encoding a cytoplasmic protein, aldehyde dehydrogenase, was found to belong to the \textit{toxR} regulon (Parsot and Mekalanos, 1991). ToxR is a 38 kd transmembrane protein that specifically binds to a repeated sequence \textit{TGGTAT} that is 3 to 8 times upstream of the \textit{ctx} transcriptional start
site (Miller et al. 1987). OmpR (Mizuno and Mizushima, 1986) and PhoB (Makino et al., 1986) of E. coli might also recognize similar directly repeated 6-10 bp sequences. Recognition of such multiple repeats may involve the cooperative binding of multiple ToxR monomers that allow a productive contact with the RNA polymerase. ToxR-PhoA fusion proteins indicate that a transmembrane domain exists within the protein separating the protein into a cytoplasmic and a periplasmic domain. The N-terminus of the ToxR polypeptide is cytoplasmic and a homology to other prokaryotic transcription activators is found in this region. The periplasmic domain is probably involved in sensing environmental signals. ToxR apparently has both sensor and activator functions and it has been suggested as a one-component system analogous to the two-component regulatory systems (Ronson, et al., 1987). ToxR is required for the response of V. cholerae to in vitro environmental conditions (NaCl concentration, the presence of certain amino acids, pH, and temperature) which are known to affect the expression of the virulence genes (Miller et al. 1989). Recent data suggest that the expression of the toxR gene is modulated by the heat-shock response and controlled by the level of sigma-32 in the cell (Parsot and Mekalanos, 1990). A heat-shock protein coding gene htpG is located upstream from and in the opposite orientation to toxR. The htpG heat-shock promoter is in close proximity to toxR promoter. The high temperature suppressive effect on toxR expression might be explained as a result of a competition between the σ^70 and the σ^32 RNA polymerases in the short intergenic region. PhoA usually exists as a dimer in periplasm, and the ToxR-PhoA fusion protein is 80-fold more active in activation of ctxAB operon than native ToxR when cells are grown under high salt concentrations. This could be explained as a result of stable dimerization through the PhoA protein in the case of the ToxR-PhoA fusion. Therefore osmoinduction of toxin production may be exerted through the dimerization of the C-terminus of ToxR. The toxS gene located directly downstream from toxR in an apparent operon is required in trans for the
activator function of ToxR, however ToxS does not directly activate the toxR promoter, suggesting that its effect occurs at posttranscriptional level (Miller et al., 1989). Overproduction of ToxR relieves its requirement for ToxS for full activity, which is analogous to the transcriptional activator OmpR in E. coli: in that overexpression of OmpR obviates its requirement for EnvZ (Csonca, 1989). ToxS is a 19 kd membrane-associated protein that resides almost completely within the periplasm and ToxS probably interacts with the C-terminal periplasmic domain of ToxR. A model for the stabilization of functional ToxR dimers by ToxS binding is therefore proposed (Fig. 23, DiRita and Mekalanos, 1991). In this model, dimerization of the C-terminal periplasmic domain of ToxR is necessary for its N-terminal functions of DNA binding and transcriptional activation. ToxS binds to and stabilize spontaneously forming ToxR dimers, particularly under conditions of low ToxR expression. ToxS-blind mutants of ToxR protein which can not be recognized by ToxS were found to have lesions in a of 18 amino acids located near the transmembrane domain but on the cytoplasmic side. This region is within but located on the end of the homologous domain ToxR shared with other transcriptional activators (DiRita and Mekalanos, 1991).

Seventeen Tox-activated genes (tag genes) have been identified by screening TnphoA fusions. Unlike ctxAB which is under the control of toxRS, the expression of some of the tag gene needs other transcriptional activators. A toxT gene is identified that accounts for the expression of the tag genes (DiRita et al., 1991). The expression of toxT is dependent on ToxR, thus a regulation cascade is established in V. cholerae virulence determinants.

Agrobacterium tumefaciens. The molecular biology of tumorigenesis caused by A. tumefaciens in plants has been reviewed (Zambryski, 1988; Winans, et al., 1987).
Figure 23. Model for the interaction between ToxR and ToxS. In this model, ToxR dimers are assembled or stabilized by interaction of monomers with ToxS. Dimers are active for binding to DNA (symbolized by a double helix) and transcriptional activation. Amino acid numbers for ToxR are as described in Miller et al. (1989). The black shaded portions of both ToxR and ToxS represent transmembrane domains. Various postulated domains of ToxR are as follows: a, DNA binding domain; b, linker domain; c, spacer domain; d, ToxS interaction region defined by deletion and TnphoA fusion analysis. ToxS-blind mutations are located in the spacer domain (c) close to the transmembrane segment (DiRita and Mekalanos, 1991).
The environmental control of the expression of the virulence genes was briefly described (DiRita and Mekalanos, 1989). In the presence of wounded plant cells the transposition of T-DNA carried by the Ti plasmid of *A. tumefaciens* into the plant genome requires the expression of the *vir* genes which are also encoded by the Ti plasmid. Two of these genes, *virA* and *virG*, mediate the induction of all other *vir* genes in response to plant exudates containing phenolic compounds. VirA and VirG have been identified as members of a large group of sensor-regulator two-component regulatory systems. Higher gene dosage of *virG* increases the efficiency both of *vir* gene induction and of infection (Jin et al., 1988). VirA is a transmembrane protein kinase that phosphorylates itself and VirG (Jin et al., 1990b). The *virG* gene is induced by acetylsyringone (phenolic compound) in a *virA-virG*-dependent manner, thus allowing positive autoregulation of the gene (Stachel, and Zambryski, 1986), whereas the induction of *virG* by phosphate starvation and acidic media is not *virA* dependent (Winans et al., 1988). A consensus DNA sequence, TNCAATTGAAAPy, *vir* box, has been identified in the 5'-noncoding region of all *vir* genes, and it is a *cis*-acting regulatory sequence that serves as a VirG protein binding site (Jin et al., 1990b). Analysis of the different *vir* gene promoters suggests that not only the presence of the *vir* box but also its relative position are important in the regulation of transcription. Although some *vir* genes have more than one *vir* box, it seems that only one *vir* box at positions -69 to -58 relative to the transcriptional start site is sufficient for induction (Jin et al., 1990b). The transcription of *virG* is enhanced by plant wound released phenolic compounds, phosphate starvation, and acidic media (Winan et al., 1988). Two promoters (P1 and P2) and three *vir* boxes (I, II and III) are found in the 5'-noncoding region of *virG* gene. Induction by phenolic compound requires promoter P1 and *vir* boxes I and III, whereas induction by phosphate starvation requires P1 and III. Promoter P2 is responsible for the induction by acidic media (Winans, 1990). P1 contains a sequence strongly similar to the *pho* box found in the -42
to -22 region of a group of *E. coli* promoters that are inducible by phosphate starvation (Fig. 24), and VirG protein is homologous to the PhoB protein which regulates the phosphate starvation genes through binding to the *pho* box. The sequence of P2 has a strong similarity to *E. coli* heat shock-inducible promoters (Fig. 24). P2 is also induced by high temperature, ethanol, alkaline pH, and heavy metals (Winans, 1990) which are common inducing signals for *E. coli* genes under the sigma-32 control (VanBogelen et al., 1987), suggesting P2 is a member of the heat shock regulon.

**Regulation of *cad* operon of *E. coli***

Bacterial amino acid decarboxylases are dramatically induced by acid pH (Gale, 1946), and thus provide good model systems to study pH regulated promoters. The expression of inducible lysine decarboxylase is affected by pH, as well as substrate and oxygen availability. Maximum induction was observed when cells were grown anaerobically at low pH in a rich medium with an excess of lysine (Sabo et al., 1974). The inducible lysine decarboxylase has been found to be encoded by the distal *cadA* gene in the *cad* operon (Chapter 3), and the proximal *cadB* gene encodes a protein exhibiting dramatic similarity with the proposed arginine: ornithine antiporter ArcD of *Pseudomonas aeruginosa* (Gamper et al., 1991). It is hypothesized that *cadB* facilitates the excretion of cadaverine by serving as a lysine: cadaverine antiporter in the high H⁺ concentration detoxification system (Chapter 3).

The chromosomal Mu *lac* fusion strains of *cadA* and *cadB* have been previously constructed (Auger et al., 1989). It was noted that *cadR* mutants which exempt the *cad* operon from lysine control and at the same time exhibit a lowered level of lysine transport (Popkin and Maas, 1980; Tabor et al., 1980), do not lose the pH regulation of the *cad* operon (Auger and Bennett, 1989). The *cad* operon is not regulated by some known transcriptional regulatory genes: *fnr, cya, crp, pgI* (Auger and Bennett, 1989).
Figure 24. Similarities between virG promoters and E. coli promoters induced by phosphate starvation (A) or by heat shock stresses (B). Line 1 represents sequence alterations in vir box III of pSW288 that abolish induction by phosphate starvation. Line 4 represents a 10-bp deletion in pSW303 and pSW305 that causes an attenuation in induction by acidic media (Winans, 1990).
A

pSW288
P1 (wild type)

consensus phosphate box

TTGCATAAAAATGAAATACTTGGTTCGCATTTTTGTCA

CTGCATAAAAACGTTCAT

22bp +1

B

pSW303, pSW305
P2 (wild type)

consensus heat shock promoter

GCGGTCAAGCGCAATTCTGACGAAAAGTCCCAGACCTGCCCAGAAATTTAGCTGGAGA

TNTCNCCCTTGAA 13-15bp CCCCATTTA +1

-35 -10
Recently, an ORF immediately upstream from *cad* operon was identified which was required for the activation of the *cad* operon (Watson et al., 1991). The amino acid sequence of this new gene, *cadC*, indicates that it has characteristics of the ToxR one-component sensor-regulator family (Watson et al., 1991). We confirm that *cadC* is essential for the pH induction of *cad* operon and characterize the activator binding site(s) by a series of deletion and mutation experiments, as well as by *in vivo* footprinting experiments. The possible routes through which pH and lysine exert their effects are also considered.

**RESULTS**

**Construction of operon and protein *lac* fusions**

Multicopy operon fusion cloning vectors pRS550 and pRS551, and protein fusion vectors pRS552, and pRS577 were used in this work (Fig. 25). A polylinker BE was inserted between the *EcoRI* and *BamHI* cloning sites of pRS552 and pXA to provide the desired orientation and protein coding phase, and the resulting vector was named pRS552-1 and pXA-1 (Fig. 26). Initially, a 3 kb *BamHI*-*EcoRI* fragment from the mini-Mu construct pKER65 was moved into pRS550, and designated as pSM10. This was the first *cad* plasmid construct that demonstrated strong pH induction. The same fragment was also put into pRS552-1, generating pSM552-1-10 as the *cadA:: lacZ* protein fusion counterpart of pSM10.

**β-Galactosidase activities of *lac* fusion constructs**

Because the pH induction of *cad* operon is not strictly dependent upon the oxygen concentration (or Fnr protein) and the difference of pH induction between aerobic and anaerobic conditions is only about two fold (Auger and Bennett, 1989), a less specific regulation (e.g. local DNA superhelicity) is thought to take place. In this research, the
Figure 25. Plasmid maps of pRS550, pRS551, pRS552, and pRS577. Plasmid pRS551 is the operon fusion version of pRS552 and plasmid pRS577 is the protein fusion counterpart of pRS550. Detailed sequence and restriction map of each plasmid are available from Simons et al. (1987). The replication origin (ori) is from pBR322. Genes coding for ampicillin (Ap$^\mathrm{T}$) and kanamycin (Km$^\mathrm{T}$) resistance are as indicated. The direction of transcription of each gene is indicated by arrow. Four tandem copies of the strong transcriptional terminator (T14) from the E. coli rrnB operon are located upstream from the unique cloning sites to block transcription from upstream plasmid promoters. Complete lac operon coding genes (lacZYA) are present in these plasmids except that for protein fusion plasmids, the first 8 codons of lacZ gene is truncated. B, BamHI; E, EcoRI; H, HindIII; Sm, Smal; RV, EcoRV.
Figure 26. Construction of cloning vector pSM552-1 and pXA-1. A custom designed poly linker-1, which has multiple cloning sites and cohesive ends of *EcoRI* (AATT) and *BamHI* (GATC) on each end was ligated to the *EcoRI* and *BamHI* sites on pRS552 and pXA. The resulting plasmids pRS552-1 and pXA-1 contain reversed *BamHI* and *EcoRI* sites from the original plasmids and other new cloning sites. The top strand sequence of the poly linker-1 fused to the ninth codon of *lacZ* gene is as shown. All sites present in poly linker-1 are unique in pXA-1 whereas only *SacII*, *BamHI*, *EcoRI*, and *ApaI* are unique in pRS552-1. A GC-rich region (GCₜ) called GC clamp, located between the transcriptional terminators and the poly linker site in pXA-1, is useful for enrichment of mutants of cloned fragments in saturation mutagenesis experiments (Myers et al., 1985). pXA-1 is both streptomycin (Str⁺) and spectinomycin (Sp⁺) resistant. Km, kanamycin; Ap, ampicillin.
aspect of air regulation has not been pursued in depth. All promoter activity assays were
done under anaerobic conditions unless otherwise specified.

**Differential expression of cadB and cadA genes.** Subsequent lac fusions
constructed with promoter cloning vectors and their corresponding β-galactosidase
activities in MC4100 at different pHs are depicted in Fig. 27. Significant variations were
observed between operon fusions and protein fusions. For protein fusions, cadA fusions
generally exhibited a 2-3 fold higher β-galactosidase activates at pH 5.5 than the cadB
fusions. The differential expression of cadB and cadA suggests that cadA is translated
more efficiently.

**PCR random mutagenesis of fragment (primer+)-(R-).** In an early
incomplete experiment, a fragment of about 500 bp flanked by primers (primer+) and
(R-) was subjected to random mutagenesis. The mutagenized fragments were then fused
to lacZ gene on plasmid pXA to determine the promoter activities and the nature of
mutations. The β-galactosidase activities of the parent plasmid pXA303 and 10 pMu
derivatives are listed in Table 13. Because of the length of the region subjected to
random mutagenesis, most of the fragments sequenced with primer pX have multiple
mutations, compounding the effort to interpret the results. Another problem caused
desertion of this project was the poor quality of the synthesized primer 21, which made it
impossible to sequence the other half of the fragment.

**The DNA segment between the SspI and EcoRV sites is essential for
pH induction of the cad operon.** The shortest restriction fragment that still showed
a full pH response is the SspI-RsaI fragment cloned in pSM552-415. In contrast, the
EcoRV-RsaI fragment in pSM552-405 did not display any pH induction, indicating that
at least part of the essential pH responsive element is located beyond the EcoRV site.
However, the fragment upstream from the EcoRV site did not respond to pH induction
Figure 27. Operon and protein fusion constructs of the *cad* upstream regulatory region. Open reading frames of *cadC*, *cadB*, and *cadA* are represented by open boxes, and arrows point in the direction of transcription. Regions beyond the double lines in *cadC* and *cadA* are truncated. The dotted region is from the vector sequence in pKER65. Intercepted lines below the map indicate fragments cloned. The resulting plasmids and their corresponding β-galactosidase activities in strain MC4100 under inducing and non-inducing conditions are listed adjacent to the diagram. Relevant restriction sites are shown. B: *BamHI*; E: *EcoRI*; Ha: *HaeIII*; Hp: *HpaI*; Pv: *PvuII*; Rs: *RsaI*; RV: *EcoRV*; Ss: *SspI*; X: *XhoI*. * means not available. All β-galactosidase units are average values from at least 3 assays.
<table>
<thead>
<tr>
<th>Operon Fusion</th>
<th>pH 5.5</th>
<th>pH 8</th>
<th>Protein Fusion</th>
<th>pH 5.5</th>
<th>pH 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSM10</td>
<td>2800</td>
<td>67</td>
<td>pSM552-1-10</td>
<td>2500</td>
<td>33</td>
</tr>
<tr>
<td>pSM380</td>
<td>3400</td>
<td>190</td>
<td>pSM552-1-380</td>
<td>2000</td>
<td>52</td>
</tr>
<tr>
<td>pSM290</td>
<td>3300</td>
<td>320</td>
<td>pSM557-290</td>
<td>610</td>
<td>42</td>
</tr>
<tr>
<td>pSM210</td>
<td>420</td>
<td>740</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSM170</td>
<td>1700</td>
<td>150</td>
<td>pSM552-1-170</td>
<td>920</td>
<td>30</td>
</tr>
<tr>
<td>pSM270</td>
<td>6400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSM225</td>
<td>1100</td>
<td>2300</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSM315</td>
<td>4400</td>
<td>930</td>
<td>pSM552-315</td>
<td>580</td>
<td>65</td>
</tr>
<tr>
<td>pSM325</td>
<td>1600</td>
<td>250</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSM265</td>
<td>3800</td>
<td>1400</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSM415</td>
<td>6100</td>
<td>650</td>
<td>pSM552-415</td>
<td>2200</td>
<td>210</td>
</tr>
<tr>
<td>pSM160</td>
<td>52</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSM405</td>
<td>500</td>
<td>610</td>
<td>pSM552-405</td>
<td>210</td>
<td>270</td>
</tr>
</tbody>
</table>
Table 13. pH responses of lac fusions to mutagenized cad upstream regulatory region.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>β-Galactosidase</th>
<th>Induction (pH5.5/pH 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pH 8</td>
<td>pH 5.5</td>
</tr>
<tr>
<td>pXA-1-10</td>
<td>35</td>
<td>380</td>
</tr>
<tr>
<td>pXA-1-170</td>
<td>15</td>
<td>160</td>
</tr>
<tr>
<td>pXA303</td>
<td>48</td>
<td>250</td>
</tr>
<tr>
<td>pMU1</td>
<td>200</td>
<td>400</td>
</tr>
<tr>
<td>pMU2</td>
<td>24</td>
<td>110</td>
</tr>
<tr>
<td>pMU3</td>
<td>24</td>
<td>200</td>
</tr>
<tr>
<td>pMU6</td>
<td>27</td>
<td>120</td>
</tr>
<tr>
<td>pMU8</td>
<td>60</td>
<td>660</td>
</tr>
<tr>
<td>pMU9</td>
<td>35</td>
<td>400</td>
</tr>
<tr>
<td>pMU11</td>
<td>224</td>
<td>273</td>
</tr>
<tr>
<td>pMU13</td>
<td>57</td>
<td>615</td>
</tr>
<tr>
<td>pMU14</td>
<td>61</td>
<td>154</td>
</tr>
<tr>
<td>pMU15</td>
<td>30</td>
<td>64</td>
</tr>
</tbody>
</table>

This batch of assays has lower than usual β-galactosidase activities, but the induction was still observed. The pH 8 activities of pMU1 and pMU11 are increased and the pH 5.5 activity of pMU15 is decreased, while pMU8 and pMU13 seem to have higher pH5.5 activities.
(pSM210) indicating that other essential elements (e.g. promoter sequences) are down stream of the EcoRV site. Two sequence features stand out between the SspI and EcoRV sites: a possible transcription terminator stem-loop structure (a, Fig. 28) and an octonucleotide palindrome (b, Fig. 28).

**Deletion and mutation analysis of the octonucleotide palindrome sequence between SspI and EcoRV.** The participation of the stem-loop structure (a) in pH regulation was ruled out since the two PCR amplified fragments (1)-(R-) vs (2)-(R-) (Fig. 28) cloned in pSM552-425 and pSM552-435 respectively gave essentially the same pH induction results. Further deletions or mutations extending near (b) resulted in either partial reduction of the maximal level of expression as in pSM552-445 and pSM552-445, or almost complete loss of the pH response as in pSM552-605 and pSM552-615. The sequence of the octonucleotide palindrome plus its flanking region and the actual base pair changes incorporated in the primers SsRV2 to SsRV6 of Fig. 28 are shown in Fig. 29. Preservation of the palindrome (b) alone did not allow for full pH induction (primer SsRV3). The right side region of the palindrome seemed more important than the left side sequence (SsRV6 vs SsRV4), although nonidentical changes were made in the two halves of the palindrome. A series of plasmids were made using the PCR fragments and the primers SsRV1-6 to define the left end with the right end defined by primer T at position +41 in the region just before the translational start sequence of cadB (pSM485, pSM495, pSM505, pSM515, pSM625 and pSM635), or at a position within cadA by using primer A- (pSM552-545, pSM552-555, pSM552-565, pSM552-575, pSM552-585 and pSM552-595). Examination of these plasmids for their promoter activities revealed results consistent with the analogous series of cadB fusion constructs. In the case of those fragment constructs having a right end point before the
Figure 28. Deletion and mutation of the SspI to EcoRV region by PCR. PCR fragments were amplified by primers (parenthesized) flanking each fragment. Primers 1 to 6 are the same primers SsRV1 to SsRV6 (Table 2). Short vertical lines indicate mutated residues in each fragment. The pH responses of these fragments are reflected by their β-galactosidase activities from both operon and protein fusions in different vectors. The region of cadA beyond the wiggled line is not shown in this map. Stem-loop structure (a) and octonucleotide palindrome (b) are indicated by a pair of inverted arrows. Relevant positions are numbered from the transcriptional start site as +1 defined by Watson et al. (1991).
Figure 29. Base pair changes in primers SsRV2 to SsRV6. The sequence of the octonucleotide palindrome is presented in bold type and the residues changed in each primer are underlined. The left end EcoRI cloning site is as labeled and the right side primer-extended sequence is indicated by dotted line. The sequence 3' to the EcoRV site is omitted in this figure.
CCAGAGAATGTCCACGCAATCCATTGTA A A C AT TAA ATGT TTA TCTTTTCATGATATC

Eco RV

SsRV2
CCAGAGAATTCACGCAATC

Eco RI

SsRV3
GTCACGCAATGAAATTGTAA A A C AT TAA ATG

Eco RI

SsRV4
CCATTGGA A TTGTAA ATGTTTA TCTTTTCATG

Eco RI

SsRV5
CATTA GAATTGTA TCTTTTCATG

Eco RI

SsRV6
ACGGAATTGTA A A C AT TAGAATTGTA TCTTTTC

Eco RI
*cadB* coding sequence, a relatively high basal level of activity was found at pH 8. The overall results suggested a transcriptional control mechanism.

**Protein titration studies**

Because most regulatory proteins exist in relatively small amount in the cell, the regulator of the *cad* operon, whether it is either a positive or negative acting factor, should be titrated out by the presence of an excess number of binding sites. Fragments containing different upstream regulatory segments were inserted into the multicopy plasmid pEMBL8+ and are designated the pTM series.

**Titration of pH-responsive regulator(s).** The titration effects of the resulting pTM series were tested on the chromosomal *cadB::*lacZ operon fusion strain, GNB8385K, by monitoring the level of β-galactosidase activity compared to that of the parent strain at pH 5.5 and pH 8 (Fig. 30). Unexpectedly, the initial results showed that pTM220 did not demonstrate strong titrating ability nor did pTM400, while pTM240 and pTM410 almost completely abolished the chromosomal *cad* promoter activity. Clearly, the pH responsive site(s) covers sequences on both sides of the *EcoRV* site. The promoter activities of three GNB8385K mutants (GNB8385K-37, -24mc, and -8mc) which exhibit higher promoter activities at both pH 5.5 and pH 8 (X. Shi and B. Waasdorp, unpublished) were also reduced to almost background values at both pHs in the presence of multicopy pTM410 (data not shown). Because the titration experiment only reduced high promoter activity but did not increase promoter activity under conditions where there is little expression, the results indicated the involvement of a pH-responsive activator.

**Titration of the pH-responsive activator(s) with deleted and mutated fragments.** The same series of deleted and mutated fragments that were tested for
Figure 30. Effect of multicopy plasmid-borne cad upstream regions on expression from the chromosomal pH-responsive cad promoter. Fragments cloned in the pTM plasmid series are indicated by intercepted lines. Both ends of the fragments are either defined by existing restriction enzyme sites (as shown in the top map) or by oligonucleotide primers (parenthesized). Coding regions of cadC and cadB beyond double lines are not shown. Primers 1 to 6 and 9 are the same as primers SsRV1 to SsRV6 and SsRV9 respectively. Primers 7-, 10- and 11- are the same as primers RV7-, RV10- and RV11- respectively. The β-galactosidase activities of GNB8385K without or with different pTM plasmids are listed accordingly. #: not tested. Restriction sites: Pv, PvuII; Rs, RsaI; RV, EcoRI; Ss, SspI.
promoter activity (pSM552-425, pSM552-435, pSM552-445, pSM552-455, pSM552-605 and pSM552-615) were also scored for their titrating ability. The results indicated that the titrating ability of the fragment correlated to its pH responsive promoter activity (Fig. 28). Because the titration effect of a fragment was reflected by the promoter activity of chromosomal cad promoter, it was possible to define the regulatory site without having the entire promoter sequence present within the fragment being tested. pTM640 displayed satisfactory titrating ability, however, not as strong as pTM490 and pTM690 (Fig. 30). The more extensively deleted fragments present in plasmids pTM650, pTM660 and pTM 680 only possessed partial titrating ability. Therefore, the 66 bp fragment from position -97 to -161 cloned in pTM640 appeared to contain the most active portion of the regulatory sequence responsive to pH regulation.

**PCR mutagenesis.** The involvement of a long segment of sequence in the pH regulation of cad promoter was indicated by failed attempts to localize the regulatory site(s) within a region shorter than the 66 bp fragment cloned in pTM640 by protein titration. As an alternative, the importance of individual base pairs was explored by localized PCR random mutagenesis in the 66 bp region. Primer SsRV2 and primer RV7- were used to anneal to pSM10 for extension. Primer extension under mutagenic conditions by PCR was conducted (see Methods). The mutagenized PCR fragments were then ligated into pEMBL8+ and 187 clones were obtained from a single ligation reaction. The resulting series of pTMU clones were transformed into GNB8385K and the color change was compared to pTM640/GNB8385K (white to light blue) on modified Falkow, lysine, X-Gal, pH 5.5 plates. GNB8385K transformants that showed darker blue color than pTM640/GNB8385K were assayed for β-galactosidase activity, and the pTMU clones were sequenced to determine the base pair changes. Later, clones that did not show an altered titration effect were also sequenced to locate unimportant base pairs in the DNA-protein interaction. Of those 106 sequenced clones, 50 single mutations, 23
Figure 31. Base pair changes in PCR mutagenized fragments cloned in pTMU plasmids. Double stranded nucleotide sequence of the -161 to -96 region is shown at the top. The 3' ends of the primer sequences are indicated by arrows. The changes of the bottom strand revealed by DNA sequencing are indicated with the new residues detected. For single bp change mutations, the corresponding pTMU plasmids in each category (lower case alphabet) are listed. The number of the name of each plasmid is given. For example, pTMU43 in the category i has a TA to CG mutation at base pair -136; for position -127, both TA to AT (n-T) and TA to GC (n-C) changes are present in pTMU133 and pTMU91 respectively. For double, triple, and multiple mutations, the corresponding pTMU plasmids are listed on the side. Signal "*" means deletion.
Single mutation

-161  -150  -130  -110  -96

CAACCAATCTGATGTAACATTTGCATTATGTCATTTCTGATATCAAACCTTTGAGATCTGTGATG
GTCCGTTAGTACATTTGTAATTACAAAATAGAAGATCTATAGTTTAGGGCTAGGACTACACA

| | | | | | | | | | | | | | | | | | | | | |
| CGG G C TG T G G G CG TGG CG G CC CG G |
| T C G T |

abc d e fgh i jk lm nop qr s tu vw x

\[ a: 164 \]
\[ b: 1 \]
\[ c: 11 \]
\[ d: 36, 37 \]
\[ e: 66, 163 \]
\[ f: 18 \]
\[ g: 25, 148, 174, 185 \]
\[ h: 4 \]
\[ i: 43 \]
\[ j: 112 \]
\[ k-G: 56 \]
\[ k-T: 104, 171, 172 \]
\[ l: 15, 86, 153, 154 \]
\[ m: 88 \]
\[ n-T: 133 \]
\[ n-C: 91 \]
\[ o: 109 \]
\[ p: 13 \]
\[ q: 75 \]
\[ r: 119, 149 \]
\[ s: 158 \]
\[ t-C: 30, 41, 136, 138, 156, 160, 161 \]
\[ t-G: 81, 85, 101 \]
\[ u: 58 \]
\[ v: 111 \]
\[ w-G: 23, 82, 184, 186 \]
\[ w-T: 165 \]
\[ x: 127 \]

(Cont.)
### Double mutation

<table>
<thead>
<tr>
<th>-161</th>
<th>-150</th>
<th>-130</th>
<th>-110</th>
<th>-96</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACGCATCCATTTGTAACATTTAATGTGTTATTTTTCAATGATATAGATCAACTTGCGATGCTGATGTTGCTGACGTAAGAAGAGTAGACATAGTTGAAACGCTAGGACTACACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>159</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>C</td>
<td>.</td>
<td>95</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>C</td>
<td>.</td>
<td>G</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>C</td>
<td>G</td>
<td>183</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>C</td>
<td>G</td>
<td>135</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>C</td>
<td>G</td>
<td>26</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>8,9</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>G</td>
<td>C</td>
<td>2</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>G</td>
<td>G</td>
<td>39,87,113,114</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>G</td>
<td>G</td>
<td>38</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>G</td>
<td>G</td>
<td>27</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>G</td>
<td>G</td>
<td>5</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>G</td>
<td>G</td>
<td>50</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>G</td>
<td>C</td>
<td>57,67,72,79,83</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>C</td>
<td>G</td>
<td>55</td>
</tr>
</tbody>
</table>

### Triple mutation

<table>
<thead>
<tr>
<th>-161</th>
<th>-150</th>
<th>-130</th>
<th>-110</th>
<th>-96</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACGCATCCATTTGTAACATTTAATGTGTTATTTTTCAATGATATAGATCAACTTGCGATGCTGATGTTGCTGACGTAAGAAGAGTAGACATAGTTGAAACGCTAGGACTACACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>G</td>
<td>C</td>
<td>.</td>
<td>125</td>
</tr>
<tr>
<td>.</td>
<td>C</td>
<td>G</td>
<td>.</td>
<td>60</td>
</tr>
</tbody>
</table>

### Multiple mutation

<table>
<thead>
<tr>
<th>-161</th>
<th>-150</th>
<th>-130</th>
<th>-110</th>
<th>-96</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACGCATCCATTTGTAACATTTAATGTGTTATTTTTCAATGATATAGATCAACTTGCGATGCTGATGTTGCTGACGTAAGAAGAGTAGACATAGTTGAAACGCTAGGACTACACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>A</td>
<td>*</td>
<td>.</td>
<td>TG</td>
</tr>
<tr>
<td>.</td>
<td>C</td>
<td>G</td>
<td>.</td>
<td>34,40</td>
</tr>
<tr>
<td>.</td>
<td>C</td>
<td>G</td>
<td>GT</td>
<td>152,18</td>
</tr>
</tbody>
</table>
double mutations, 2 triple mutations, 2 deletions, and 2 multiple mutations were found (Fig. 31).

**Protein titration with PCR mutagenized fragments.** After eliminating those with identical sequence changes, 27 different single mutations (including 3 mutants that were at the same position as another mutation but bore a different nucleotide change) and 15 different double mutants were further assayed for their titration effect. The results are shown in Fig. 32. The importance of each individual base pair position was reflected by the loss of protein binding ability when altered. The higher the bar in Fig. 32, the more crucial is the position. A single base pair change at 9 positions resulted in a loss of more than 80% of the titrating ability; changes in another 9 positions exhibited a 50-80% loss of activator binding ability; the remaining 6 positions had a less than 50% influence on the DNA-protein interaction assayed by titration. All nucleotide changes were found to possess some effect but mutants exhibiting a change of less than 40% are interpreted as minor effects, if any.

**Lysine effect**

It was known that an excess amount of lysine was necessary for the full pH induction of a CadR+ strain; however, no quantitative data regarding lysine regulation of the cad operon were available. To analyze this further, the lysine effect on the plasmid borne lac-fusions (pSM series) was tested by comparing their β-galactosidase activities grown in CadR medium with or without lysine at both pH 5.5 and pH 8. The overall data suggested an approximate 7 fold increase of pH induction by the presence of an excess amount of lysine in the medium at pH 5.5; representative data are shown in Fig. 33. However, the lysine effect is dependent on the pH effect, that is, the 7 fold increase of induction could only be observed at acid pH and when the promoter still responded to the pH induction. Constructs that no longer responded to pH induction because of an
Figure 32. Effect of single base pair changes in the -161 to -96 region. The effect of each individual base pair change (abscissa) in pTMU plasmids is presented on the ordinate as the percentage of the loss of titrating ability in GNB8385K. The higher bar thus indicates a more complete loss of titrating ability. The loss of titrating ability of the mutant was calculated as the $\beta$-galactosidase value of GNB8385K bearing the pTMU plasmid divided by the $\beta$-galactosidase value of GNB8385K. A solid circle indicates a loss of over 80% of the titrating ability. An open circle indicates a loss of 50-80% of titrating ability. Without any mutations, region -161 to -96 as carried by pTM640 has a 85% titrating effect (15% loss of titrating ability) on the $\beta$-galactosidase activity of GNB8385K (Fig. 4). Most of the mutations are AT to GC changes with exception of two AT to TA and AT to CG changes. In the two positions that have different base pair changes in the same position, the same amount of titrating ability was observed.
Figure 33. Lysine effect on pH induction of *cad* promoter. The *cad*
promoter activities of different *lac* fusion plasmids harbored by MC4100 were tested.
The strains were grown in CadR medium with (0.5 %) or without lysine. pSM552-545
carries a protein fusion to *cadA* and pSM552-425 carries a protein fusion to *cadB*.
pSM552-405 also carries a protein fusion to *cadB* but does not have the essential *EcoRV*
upstream sequence for pH induction (Fig.27).
GNB8385K
β-Galactosidase

pH5.5

<table>
<thead>
<tr>
<th></th>
<th>+Lys</th>
<th>-Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td>w/o</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1530</td>
<td>75</td>
<td>33</td>
</tr>
<tr>
<td>5100</td>
<td>340</td>
<td>170</td>
</tr>
<tr>
<td>560</td>
<td>560</td>
<td>250</td>
</tr>
<tr>
<td>1100</td>
<td>910</td>
<td>340</td>
</tr>
<tr>
<td>1200</td>
<td>1200</td>
<td>550</td>
</tr>
</tbody>
</table>

(2) (T-)
(3) (T-)
(4) (T-)
(5) (T-)
(6) (T-)
insufficient pH regulatory region did not respond to lysine; for instance, pSM552-405. One hypothesis suggested that CadR might act as a repressor of the cad operon that could be inactivated upon binding to lysine (Popkin and Maas, 1980). If a repressor site could be found in the regulatory region, it would be strong evidence to support the hypothesis. The same strategy as used for locating the pH regulator binding site based on protein titration experiments was used to investigate the lysine responsive site. Theoretically, if there is a lysine responsive repressor site, it should be able to titrate out the repressor and relieve the lysine requirement for maximal induction. Again, the pTM plasmids were tested for their lysine regulator titration effects on GNB8385K grown at pH 5.5 in CadR medium with or without lysine (Fig. 34). The results showed a lysine effect in all sets of experiments; no evidence supported that there was an independent lysine responsive repressor site. A possible explanation might be that the hypothetical repressor could be only titrated out by massive amounts of lysine but not by multicoys of a binding sequence. No elevation in the β-galactosidase levels is observed (-lysine) with any of the regulatory region containing plasmids, suggesting that they do not compete effectively for an independent repressor site. To further explore the possibility of a closely spaced pH responsive site and a lysine responsive site, pTMU plasmids with single mutations that exhibited altered pH regulator titration ability were tested for their lysine repressor titration ability. If any of these weakened pH activator binding mutants could titrate out a lysine repressor more effectively than they titrated out the pH activator, a separation of the two functions might be observed. However, a lysine effect from 2.5 to 13 fold was observed in all pTMU/GNB8385K transformants (Fig. 35). These data suggest that there might not be an independent lysine repressor binding site in the upstream region of cad operon or that the repressor is not effectively titrated under these conditions. It is also possible that the entire required lysine repressor binding site is not present on these plasmids.
Figure 34. Effect of multicopy plasmid-borne cad upstream regions on the chromosomal lysine-responsive cad promoter activity in GNB8385K. Fragments cloned in the pTM series are shown under the corresponding restriction map of the cad upstream region. Restriction fragments cloned in pEMBL8+ are indicated by intercepted lines. The end of fragments obtained by PCR are indicated by the primers used (in parenthesis), and their sequences are shown in Table 2. Cultures of GNB8385K bearing each plasmid were grown in CadR medium with or without lysine anaerobically at pH 5.5. β-galactosidase was assayed as described in methods.
Effect of Lysine on pH Induction

<table>
<thead>
<tr>
<th>Lysine</th>
<th>pH 5.5</th>
<th>pH 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+Lys</td>
<td>-Lys</td>
</tr>
<tr>
<td></td>
<td>+Lys</td>
<td>-Lys</td>
</tr>
<tr>
<td>pSM552-545</td>
<td>7100</td>
<td>910</td>
</tr>
<tr>
<td>pSM552-425</td>
<td>730</td>
<td>100</td>
</tr>
<tr>
<td>pSM552-405</td>
<td>250</td>
<td>280</td>
</tr>
</tbody>
</table>
Figure 35. Effect of mutated cad upstream regions on pTMU plasmids on the chromosomal lysine-responsive cad promoter activity in GNB8385K. Random mutagenized fragment in pTMU series were tested for their abilities to titrate hypothetical lysine-responsive regulator in the presence or absence of lysine in CadR medium. Theoretically if a lysine-responsive promoter can be titrated without affecting the affinity for pH-responsive activator, differential expression caused by lysine effect should not be seen, which is however not the case implicated in this experiment.
<table>
<thead>
<tr>
<th></th>
<th>+Lys</th>
<th>-Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td>w/o</td>
<td>4800</td>
<td>900</td>
</tr>
<tr>
<td>pTMU1</td>
<td>290</td>
<td>50</td>
</tr>
<tr>
<td>pTMU4</td>
<td>145</td>
<td>17</td>
</tr>
<tr>
<td>pTMU11</td>
<td>2600</td>
<td>800</td>
</tr>
<tr>
<td>pTMU13</td>
<td>1200</td>
<td>200</td>
</tr>
<tr>
<td>pTMU15</td>
<td>2000</td>
<td>630</td>
</tr>
<tr>
<td>pTMU18</td>
<td>320</td>
<td>55</td>
</tr>
<tr>
<td>pTMU23</td>
<td>540</td>
<td>37</td>
</tr>
<tr>
<td>pTMU25</td>
<td>510</td>
<td>150</td>
</tr>
<tr>
<td>pTMU30</td>
<td>1600</td>
<td>180</td>
</tr>
<tr>
<td>pTMU36</td>
<td>1850</td>
<td>230</td>
</tr>
<tr>
<td>pTMU43</td>
<td>600</td>
<td>220</td>
</tr>
<tr>
<td>pTMU56</td>
<td>1600</td>
<td>450</td>
</tr>
<tr>
<td>pTMU58</td>
<td>2400</td>
<td>500</td>
</tr>
<tr>
<td>pTMU66</td>
<td>2600</td>
<td>640</td>
</tr>
<tr>
<td>pTMU75</td>
<td>340</td>
<td>60</td>
</tr>
<tr>
<td>pTMU81</td>
<td>910</td>
<td>72</td>
</tr>
<tr>
<td>pTMU88</td>
<td>310</td>
<td>35</td>
</tr>
<tr>
<td>pTMU91</td>
<td>1700</td>
<td>200</td>
</tr>
<tr>
<td>pTMU104</td>
<td>1600</td>
<td>290</td>
</tr>
<tr>
<td>pTMU109</td>
<td>1700</td>
<td>250</td>
</tr>
<tr>
<td>pTMU112</td>
<td>2100</td>
<td>520</td>
</tr>
<tr>
<td>pTMU113</td>
<td>1300</td>
<td>320</td>
</tr>
<tr>
<td>pTMU119</td>
<td>3100</td>
<td>370</td>
</tr>
<tr>
<td>pTMU127</td>
<td>310</td>
<td>30</td>
</tr>
<tr>
<td>pTMU133</td>
<td>2000</td>
<td>190</td>
</tr>
<tr>
<td>pTMU158</td>
<td>680</td>
<td>78</td>
</tr>
<tr>
<td>pTMU164</td>
<td>2800</td>
<td>570</td>
</tr>
</tbody>
</table>
In vivo footprinting

Attempts to detect the interaction between the activator and the chromosomal binding site by PCR primer extension of in vivo methylated chromosomal DNA (Sasse-Dwight and Gralla, 1990) were not successful. Instead, in vivo footprinting on plasmid DNA was attempted as a technically less difficult route. To avoid ambiguous results due to unsaturated activator sites present on a multicopy plasmid, a PCR fragment containing the cadC gene with its own expression system (Watson et al., 1991; GenBank accession number M67452) was amplified from MC4100 chromosomal DNA and inserted into plasmid pSM485, as well as pSM552-425, pSM552-545 and pEMBL8+. Plasmid pSM485C⁻ therefore has the same copy number of the cadC gene and the activator binding site (Fig. 36). To ascertain β-galactosidase expression levels from the cad promoter prior to in vivo methylation studies, β-galactosidase levels were measured in cadC-plasmid bearing strains. The presence of multicopy cadC on pSM485C⁻ boosted the β-galactosidase levels at both pHs to a higher level, however a good induction ratio (pH 5.5/pH 8) was observed in some constructs (pSM552-545C⁻, Fig.37). Similar effects of multicopy cadC gene carried in pCADC⁺ and pCADC⁻ also were observed on the induction of the single copy chromosomal cad::lac fusion (GNB8285K); however, the activity at pH 5.5 did not go up probably because the single cad promoter was already saturated (Fig. 38). Thus too much expression of cadC could activate the cad promoter at a normally non-inducing pH. Therefore, instead of pSM485C⁻, pSM485 was used as a suitable negative control at pH 8 to reduce background. Rifampicin was added before dimethylsulfate methylation to halt RNA polymerase, in an effort to enhance the stability of activator binding and at the same time provide an internal control for detecting the cad promoter. The protection patterns of the opposite strands differ (Fig. 39). In the top strand, only two G residues were protected. The importance of the two protected G positions of -121 and -135 were consistent with the importance of this region found from

The PCR fragment amplified from E. coli MC4100 chromosomal DNA with primers cadC+ and cadC− was cloned into the EcoRI site of pSM485, pSM552-425, pSM552-545 and pEMBL8+ respectively. The superscript '+' indicates that cadC gene is inserted in the same orientation as the cad promoter in plasmid pSM485, pSM552-425, and pSM552-545, and in the same orientation as the lac promoter in pEMBL8+. The superscript "-" means the opposite direction. Relevant restriction sites and primer locations (parenthesized) are shown in the top map. Dotted Boxes indicate the region corresponding to the top map cloned into the lacZ fusion vector pRS551 or pRS552, and the cloning vector pEMBL8+. 
Figure 37. Effect of cadC gene copy number on the pH induction of multicopy cad::lac fusions. The expression levels of multicopy cad::lac fusion constructs with a single chromosomal cadC gene and multicopy plasmid-borne cadC gene are compared at both pH 5.5 and pH 8.
Effect of cadC Copy No. on Multicopy cad :: lac Fusions

![Bar chart showing β-Galactosidase activity at pH 5.5 and pH 8 for different constructs.]

<table>
<thead>
<tr>
<th></th>
<th>pH 5.5</th>
<th>pH 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSM485</td>
<td>2800</td>
<td>370</td>
</tr>
<tr>
<td>pSM485C-</td>
<td>34000</td>
<td>12000</td>
</tr>
<tr>
<td>pSM552-425</td>
<td>800</td>
<td>67</td>
</tr>
<tr>
<td>pSM552-425C+</td>
<td>11000</td>
<td>4800</td>
</tr>
<tr>
<td>pSM552-425C-</td>
<td>9500</td>
<td>4600</td>
</tr>
<tr>
<td>pSM552-545</td>
<td>5400</td>
<td>150</td>
</tr>
<tr>
<td>pSM552-545C+</td>
<td>30000</td>
<td>30000</td>
</tr>
<tr>
<td>pSM552-545C-</td>
<td>32000</td>
<td>1100</td>
</tr>
</tbody>
</table>
Figure 38. Effect of multicopy cadC gene on the single copy chromosomal cad promoter. The β-galactosidase activity of GNB8385K was measured at pH 5.5 and pH 8 both in the presence or absence of multicopy cadC gene carried on pCADC+ and pCADC−.
Effect of cadC Copy No. on Single Copy cad :: lac Fusion

<table>
<thead>
<tr>
<th>GNB8385K</th>
<th>w/o Plasmid</th>
<th>pCADC+</th>
<th>pCADC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH 8</td>
<td>3500</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>pH 5.5</td>
<td>2600</td>
<td>1800</td>
<td></td>
</tr>
<tr>
<td>w/o</td>
<td>3000</td>
<td>2000</td>
<td></td>
</tr>
</tbody>
</table>
Figure 39. *In vivo* methylation of pH-responsive regulatory region of *cad* operon. *In vivo* methylated plasmid DNA was isolated from MC4100 grown with modified Falkow lysine decarboxylase medium at inducing or non-inducing pH. The methylation patterns of the region between -241 to +41 were revealed by chemical sequencing of DNA fragments uniquely labeled at either end. The methylation patterns of the top strand and the bottom strand (Fig. 31) are displayed in panel A and panel B respectively. Plasmid pSM485C- was methylated at pH 5.5 (lanes 1 and 3) and plasmid pSM485, the control, was methylated at pH 8 (lanes 2 and 4). Differently methylated G residues between pH 5.5 (lanes 1 and 3) and pH 8 (lanes 2 and 4) are indicated with arrows.
the random mutagenesis results. However, an unexpected methylation pattern was seen in the bottom strand. Both protected (positions -48, -68, -74, and -82) and enhanced (position -72) signals were detected and an indication of a doublet appeared at position -40. These changes were located outside of the 66 bp (-97 to -161) critical region as found from titration experiments. Within the 66 bp region, positions -109, -113, -116 and -124 showed a reduced degree of methylation while position -129 showed an enhanced signal; however, these five positions are less obvious (Fig. 40).

**Oxygen effect on cad promoter activated by CadC**

Multicopy cadC constructs pSM552-425C-, pSM552-425C+, pSM552-545C-, and pSM552-545C+ were tested for the response to oxygen concentration at both pH 5.5 and pH 8. The results suggest that the activation of the cad promoter by excess CadC at pH 8 could be suppressed in the presence of proper amount of oxygen in some cases (Fig. 41).

**DISCUSSION**

Information regarding the global response of bacterial cells to environmental stimuli is accumulating at a rapid rate. Studies of global systems have led to the general theme of two-component sensor-regulator networks and the role of phosphorylation (Stock et al., 1989; Stock et al., 1990). The most investigated cases with regard to the role of specific sequences in pH regulation include activation of the virulence genes of *Agrobacterium tumefaciens* and *Vibrio cholerae*. In *A. tumefaciens*, the acid pH inducible promoter of *virG* gene was found to resemble the heat shock promoter in *E. coli* (Winans, 1990) and a 12 bp *vir* box (TNCAATTGAAAPy) has been identified as the VirG protein binding core sequence (Jin et al., 1990). In *V. cholerae*, cholera toxin and other virulent determinants are under the control of the ToxR protein which recognizes a tandem array of a heptanucleotide sequence (TTTTGAT) (Miller et al., 1987). Recently, an ORF denoted *cadC* immediately upstream from the *cad* operon of *E. coli* was identified as
Figure 40. **Important DNA residues involved in pH induction of cad operon.** The significance of DNA residues involved in pH induction of the cad operon is summarized by combining the results from both protein titration with altered regulatory sequences and *in vivo* footprinting experiments. The transcription start point as well as the -10 and -35 regions were determined by Watson et al. (1991); DNA base pairs are numbered accordingly. Base pairs that showed a significant effect on activator competing ability are indicated by solid or open circles as described in Fig. 32. Protected G residues in *in vivo* methylation experiments (Fig. 40) are indicated by solid triangles, and those G residues exhibiting enhanced methylation are indicated by open triangles. Palindrome sequences are indicated by pairs of arrows.
Figure 41. Oxygen effect on *cad* promoter activated by CadC. Strain E2218 is a CadC- strain in which no pH induction of *cad* promoter was observed (E. Olson). *cadB::lac* fusion plasmids (pSM552-425C- and pSM552-425C+) and *cadA::lac* fusion plasmids (pSM552-545C- and pSM552-545C+) were grown aerobically or anaerobically in modified Falkow lysine decarboxylase media at pH 5.5 and pH 8. β-Galactosidase activity of each culture was measured.
### β-Galactosidase Units

<table>
<thead>
<tr>
<th>Strain/Plasmid</th>
<th>pH 5.5</th>
<th></th>
<th>pH 8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>anaerobic</td>
<td>aerobic</td>
<td>anaerobic</td>
<td>aerobic</td>
</tr>
<tr>
<td>GNB8385K</td>
<td>4300</td>
<td>2000</td>
<td>80</td>
<td>110</td>
</tr>
<tr>
<td>E2218/pSM552-425C⁻</td>
<td>9500</td>
<td>8400</td>
<td>4600</td>
<td>400</td>
</tr>
<tr>
<td>E2218/pSM552-425C⁺</td>
<td>10700</td>
<td>7800</td>
<td>4800</td>
<td>350</td>
</tr>
<tr>
<td>E2218/pSM552-545C⁻</td>
<td>32000</td>
<td>24000</td>
<td>1100</td>
<td>370</td>
</tr>
<tr>
<td>E2218/pSM552-545C⁺</td>
<td>30000</td>
<td>20000</td>
<td>30000</td>
<td>8300</td>
</tr>
</tbody>
</table>
encoding a positive regulator of the cad promoter (Watson et al., 1991). The resemblance of the CadC protein to ToxR leads to an interesting question: do V. cholerae and E. coli adopt a similar signal transduction mechanism in response to acid pH?

In efforts to define the cad regulatory region, both lac operon and protein fusions were constructed in order to separate and analyze regulation at both transcriptional and translational levels. Considerably lower β-galactosidase activities were found in cadB::lacZ protein fusions compared to cadA::lacZ protein fusions, although the induced β-galactosidase activities of their respective operon fusions were in the same range. Clearly, certain controls exist at the translational level. These could be due to less efficient translation or lower mRNA stability of the cadB gene mRNA. Effects of this sort have been observed in other operons which encode one membrane protein and another enzymatic protein (McCarthy, 1990; Owolabi and Rosen, 1990). The unusual Shine-Dalgarno sequence of cadB, 5'AGGAGAAGAG3', only one bp from the ATG start codon might not be optimal for the ribosomal recognition and its high G content near the ATG is usually unfavorable for efficient translation (Jacques and Drefus, 1990; McCarthy and Gualerzi, 1990). Thirty-five bp downstream of the ATG of cadB a stem-loop structure [ΔG = -14 kcal/mol as calculated by Tinoco et al.(1973)] appears in the transcript. This could cause a pause site for the progression of RNA polymerase.

Unlike the toxR regulated promoter which possesses a tandemly repeated heptanucleotide sequence upstream from the promoter region, two palindromes are the most obvious sequences located in the region implicated in the pH regulated cad operon. One is an octanucleotide palindrome close to the Eco RV site. Deletion and mutation analysis indicated that although base pairs within this palindrome were important for the pH response, the palindrome structure itself did not seem to be a crucial element. A partial response was still observed when the left-side sequence of the palindrome was
mutated or deleted; in contrast, the response to pH induction was greatly reduced when mutations extended into the right side of the palindrome even with an intact left-side sequence. Furthermore, the significance of the palindrome structure was also decreased by the fact that single mutations outside of the palindrome could result in partial to complete loss of function. Sequences within the palindrome may serve as a recognition signal but certainly it is not the sole determinant of a stable protein-DNA interaction. Other notable features are a hexanucleotide and a pentanucleotide inverted repeats around the -35 region. The possibility that this segment defines a binding site as an operator for a lysine responsive repressor was tested in titration experiments. However, this hypothesis was not supported by the protein titration experiments as no repressor activity could be titrated out by plasmids bearing this segment when grown under the no-lysine condition. Whether these two palindromes have a specific function is not clear.

The region -96 to -161 was sufficient for a stable protein-DNA interaction as indicated by protein titration, however a slight difference in effectiveness was observed when compared to region +41 to -161. One would speculate that the stronger titrating ability of the longer fragment might be attributed to the stabilization effect by communicating with the bound RNA polymerase at the promoter. A fragment in pTM690 (-45 to -161) without the promoter sequence was as effective as the one in pTM490 (+41 to -161) with the promoter sequence (Fig. 30), thus ruling out a direct effect of bound RNA polymerase on the stability of the activator-DNA complex. However, this means that sequence from -45 to -96 still has a minor influence on the activator binding event. This long stretch of important sequence resembles that found for the repeating binding sites of varying affinity upstream of the gln operon (Ninfa et al., 1987).
Although an identical repeating sequence is not found within this region, variations of the sequence ATGTT are found at -137, -123, -99 (or -101), -81 and -57. This could implicate an array of sites of varying affinities which could be filled cooperatively and then activate RNA polymerase. Alternatively, the site could act as an enhancer since the most powerful titrating (activator binding) sites are the most distal from the promoter as is the case with the activator NR1 of gltA p2 (Ninfa et al., 1987). However, enhancers have only been found to act with the σ54 RNA polymerase. According to a recent review, E. coli promoters fall into two categories based on the type of sigma factor required for transcription: σ70 and σ70-like promoters recognize the basal elements known as the -10 and -35 sequences while σ54 promoters have different basal elements located at -12 and -24 (Gralla, 1991). The σ70 promoters subject to activation contain activator sites located where they communicate directly with the polymerase, by contrast, for σ54 promoters, the activators are centered near position -110 and can not touch the polymerase without looping out the intervening DNA. Such enhancer like sequences can be moved farther away in cis and still retain function. The activator site(s) and promoter of cad possess some characteristics of both types of activation.

Although the PCR mutagenesis procedure provided valuable information, the following points should be kept in mind when trying to interpret these results. First, the 3’ ends of the two primers used for PCR extension covered sequences into the 66 bp (-96 to -161) region, therefore the actual region subjected to mutation was only 46 bp (-106 to -151); sequences outside this region were undetected by this particular experiment and should not be considered irrelevant to protein-DNA interaction. Second, all of the 25 mutated positions were at A or T positions in spite of the previously stated randomness of this procedure (Leung et al., 1989). The insufficiency of PCR mutagenesis was partially compensated by in vivo DNA footprinting experiments which most readily detected G
residues in the +41 to -242 region. The positions of the two most protected G residues in the top strand were consistent with the mutation experiments. However, the methylation pattern of the bottom strand was more complex. Both protection and enhancement were observed, and some changes were spotted outside of the 66 bp region, suggesting that although an area downstream from position -96 could not compete independently with the upstream region for binding to the activator, it was protected during the initiation process. One reasonable explanation is that there are high affinity and low affinity binding sites along this stretch of about 115 bp sequence (-46 to -161); binding of the high affinity site(s) initiates the cooperative binding of the low affinity site(s) with the higher affinity sites most distal to the promoter and being those defined by the initial deletion experiments as well as by the mutagenesis experiments. This hypothesis can also account for the involvement of the long stretch of sequence detected by various experiments (Fig. 16). Or certain sequences may serve as an essential structural role rather than being engaged in protein binding (Gralla, 1991). Surprisingly, protection of the promoter region by RNA polymerase was not well detected by in vivo DNA footprinting experiments. This may be due to only a transient occupation of the site by RNA polymerase and not the stable closed complex found with activated σ54 promoters (Sasse-Dwright and Gralla, 1990). A possible problem encountered with in vivo footprinting on plasmid is the ratio of the protein to DNA binding site. If the multicopy binding sites on the plasmid outnumber the amount of the regulatory protein, underprotection is expected. It was true when we compared the methylation patterns of pSM485 at pH 5.5 and pH 8, no difference was detected (data not shown). This problem was overcome by putting a copy of cadC together with its own promoter in the same plasmid that carried the regulatory region.

However, such a construct pSM485C- not only exhibited 10 times more activity at pH 5.5 but also was quite active at pH 8; loss of pH control was also observed with other
similar constructs. Obviously, overproduction of CadC overcame the pH control mechanism, reminiscent of the observation that overproduction of ToxR exempted the requirement for ToxS for pH (and other environmental stimuli) induction. Intriguingly, the high promoter activity at pH 8 caused by the presence of excess CadC could be partially suppressed by increased oxygen concentration, reinstating the unique pH regulatable property of cad promoter in the presence of excess CadC. By shifting cell cultures from alkaline aerobic stage to acidic anaerobic stage, gene expression could be increased by about 100 fold. It has been reported that a 200 fold induction of cad promoter was achieved by optimizing medium composition and growth conditions in the absence of an extra copy of cadC gene (Tolentino, 1991). Now, with cadC dosage as an parameter, it is expected to reach an even higher induction level.

Although cadC has been proved to be required for positive regulation of cadBA and a difference of methylation patterns could be observed between pSM485 and pSM485C-, this does not prove that CadC is in direct contact with the cadBA upstream regulatory region. The pH-responsive activator of the cad operon seems to be specific; the upstream regulatory region of cad operon did not affect the promoter activity of another low-pH induced locus (adi encoding biodegradative arginine decarboxylase) in protein titration of GNB7145K with pTM plasmids. However, some structural similarity might exist between cad and adi upstream regulatory regions. An 14 bp palindrome with 2 miss matches is centered at position 92 bp upstream from the ATG start codon of adi, compared to the octonucleotide palindrome centered 214 bp from the cadB start codon (-139), and there is a good sequence homology in these regions (Fig. 42). This similarity suggest that even though cad and adi do not share the same activator, there might be a common mechanism regarding transcription initiation. Some mutant strains have been isolated that affect the expression of both decarboxylase systems (X Shi and B Waasdrop, unpublished).
Figure 42. Structure and sequence homology between *cad* and *adi* upstream regions. Conserved sequences are depicted with bold type and the palindrome sequences are underlined. The distance between the last residue of the sequence shown and the ATG start codon is indicated. Cap region is filled with dots
cad  CAATCCATTGTAACATT...AATGTTTATCTTTTCAT---197 bp-ATG
adi  TTTATAAACTAAAATTTTCAAAAAAAGTTTTGTTTTTCACG---73 bp-ATG
Experimental results showed that the presence of excess lysine increased induction by about 7 fold at pH 5.5, but this enhancing effect was not observed at pH 8, nor was it efficiently observed in the absence of either CadC or the activator binding site(s). No tested small fragment demonstrated a lysine-inactivated repressor binding site by titration studies in the absence of lysine. The data would be consistent with the lysine effect being exerted through the pH activation system. This finding is unexpected as previous experiments have shown that cadR mutants still exhibit a normal pH response and this result has been exploited in the culturing of pH regulated vectors bearing the cad promoter. There could be a more complicated role for CadR in the response. It could act to compete in the formation of the activated promoter complex through protein-protein interactions or competitively at a location on the DNA that interferes with activator action. As a potential membrane protein involved in lysine transport it could interact with CadC in some way. Recently, a model was presented for the interaction between ToxR and ToxS regarding transcriptional activation of the cholera toxin genes (ctxAB) of V. cholerae (DiRita and Mekalanos, 1991). In this model, ToxS assembles or stabilizes active ToxR dimers through protein-protein interaction. This raises the possibility of an analog of ToxS in this system and a speculation that CadR might act upon CadC. Further definition of the system awaits the analysis of the cadR locus and further dissection of the mechanism of activation of the cad promoter. A number of transposon generated mutants (other than cadC) with altered pH regulation have been isolated and these are being analyzed to determine their role in pH regulation, in what appears to be a complicated and fascinating example of signal response.
CHAPTER 5. POSSIBLE WORKING MODELS FOR REGULATION OF THE cad OPERON

INTRODUCTION

Experimental results of this work have characterized the structure and physiological role of the cad operon, and located the site(s) of pH regulation. The promoter sequence and the upstream positive regulator gene cadC required for regulation of the cad operon by extracellular pH have also been determined (Watson et al., 1991). Analysis of the mechanism of pH sensing and signal transduction can now be approached. The nature of this pH regulatory system is particularly interesting because of the similarities between CadC in regulation of the pH response of E. coli and ToxR in regulation of the virulence determinants of V. cholerae (Watson et al., 1991). Although the oxygen and lysine effects on the regulation of the cad operon have not been investigated extensively, preliminary experimental data suggest that the oxygen and lysine effects are secondary to and dependent on the pH effect. The actual regulatory mechanism adopted in this system still remains unclear; however, several possible considerations have been implicated from this work and information collected from related researchers. Speculative models and possible elements involved in the regulation of cad operon are discussed in this chapter.

Regulation of gene expression in bacteria is most often executed through the control of transcription initiation. However more and more evidence shows that post-transcriptional control also assumes a significant role in regulation of bacterial gene expression especially in the differential gene expression of a polycistronic operon.

Transcription initiation

The general mechanisms of bacterial transcription initiation have been reviewed (Reznikoff et al., 1985). In the first step of transcription initiation, bacterial sigma
factors are required to form a holoenzyme with the core enzyme of RNA polymerase to recognize specific promoter sequences, and they probably assist the melting of the DNA to form an open complex in the -10 region, leading to productive transcription initiation. The structure and function of bacterial sigma factors involved in specific transcription initiation have been reviewed (Doi and Wang, 1986; Helmann and Chamberlin, 1988). In this section, the main focus is on the E. coli promoters, and the sigma factors of other bacteria may be mentioned if relevant. In E. coli, the majority of cellular transcription is initiated by $\sigma^{70}$-RNA polymerase. Other minor RNA polymerase holoenzymes appear to control genes that respond to stressful situations. Four known minor sigma factors are $\sigma^{54}$, $\sigma^{32}$, $\sigma^F$, and $\sigma^S$, which recognize the nitrogen-regulated genes, heat-shock genes, flagella and chemotaxis genes, and growth phase-regulated genes and expression of stationary-phase phenotypes respectively (Table 14). Although only limited minor sigma factors have been found to be involved in cellular response to environmental stimuli, it is certainly a good regulatory mechanism worth being implemented in other systems.

**Homology of sigma factors.** Conserved domains among different sigma factors of different bacteria and phages have been found by amino acid sequence analysis, however this similarity does not always exists in all cases (Helmann and Chamberlin, 1988). The sigma factors $\sigma^{70}$, $\sigma^{32}$ of E. coli and $\sigma^{28}$ of B. subtilis show good homology with most other known sigma factors whereas $\sigma^{54}$ does not reveal similarity by alignment. The sequence of the $\sigma^F$ of E. coli is unknown, however $\sigma^F$ resembles the $\sigma^{28}$ of B. subtilis in many ways. The promoter sequences recognized by these sigma factors are listed in Table 14. All of the known sigma factors are acidic with an excess of negative charge at pH 7. It is especially pronounced for E. coli $\sigma^{70}$. Some functions include protein: protein interaction with core RNA polymerase, promoter recognition, and subsequent promoter binding have been assigned to the conserved domains (Helmann and Chamberlin, 1988).
Table 14. Consensus promoter sequences recognized by sigma factors.

<table>
<thead>
<tr>
<th>Promoter</th>
<th>-35</th>
<th>Distance (bp)</th>
<th>-10</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ⁷₀</td>
<td>tcTIGACa</td>
<td>16-18</td>
<td>TATAT</td>
<td>a</td>
</tr>
<tr>
<td>σ³₂</td>
<td>TntCNCoCTTGAA</td>
<td>13-15</td>
<td>CCCATT</td>
<td>b</td>
</tr>
<tr>
<td>σ⁸</td>
<td>TAAA</td>
<td>15</td>
<td>GCGATAA</td>
<td>c</td>
</tr>
<tr>
<td>σ⁸⁰</td>
<td>GNGTTAAGC</td>
<td>16-20</td>
<td>CGTCC</td>
<td>d</td>
</tr>
<tr>
<td>tRNA*</td>
<td>gTTGac</td>
<td>16-18</td>
<td>TATATCGGCCoC</td>
<td>e</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Promoter</th>
<th>-24</th>
<th>Distance (bp)</th>
<th>-12</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>σ⁵₄</td>
<td>CTGGYAYR</td>
<td>4</td>
<td>TIGCA</td>
<td>f</td>
</tr>
</tbody>
</table>

Sigma factors of *E. coli*. By analyzing known *E. coli* promoter sequences, Gralla (1991) was able to categorize them into two families and summarize the general features of each family. The majority of promoters are recognized by $\sigma^{70}$ protein or its homologs, and have consensus sequences at the -10 and -35 regions. The minority use $\sigma^{54}$ factor and have conserved sequences at the -12 and -24 regions.

A. Promoters recognized by sigma 70 or its homologs. The mechanism of transcription initiation from the $\sigma^{70}$ promoters has been well studied (Collado-Vides et al., 1991). The heat shock proteins under the control of $\sigma^{32}$ promoters have also been reviewed (Lindquist and Craig, 1988), and the $\sigma^F$ was found to be similar to the $\sigma^{28}$ of *B. subtilis* in flagella synthesis and chemotactic responses (Arnosti and Chamberlin, 1989). For $\sigma^{70}$ and $\sigma^{70}$-like promoters, the positions of the effector binding sites are critical to determine whether the promoter is regulated by a positive or a negative mechanism, and how well it is regulated. Operators positioned in the spacer sequence between the -35 and -10 regions are rationized to be the most effective repressor sites, however operators located either upstream or downstream from the promoter are also present, and duplication of operators are common in many promoters. Duplication of operators provides flexibility in complex promoters which respond to multiple signals, and enhances tighter repression in simple promoters through cooperative binding of the repressors. The same activator functioning by binding to an upstream region from the promoter can act as a repressor when the binding site is moved further downstream so as to interfere with the promoter-RNA polymerase interaction (e.g. CRP and FNR proteins). Usually a weaker $\sigma^{70}$ promoter receives stronger activation in the presence of an activator that helps the binding of RNA polymerase to the target promoter sequence, whereas in some cases, a strong promoter which already has a stable interaction with RNA polymerase is actually repressed by the occupation of an similarly positioned activator binding site (Collado-Vides et al., 1991). The alternate favored location for an
activator suggests that stereospecific contacts with polymerase must be preserved. However the stereospecific contacts between a repressor and the polymerase does not seem to be required. The activator sites of σ70 promoters are usually proximal (near -35 region) and located in a much less variable position than the operator sites. The major function of an activator may be assisting promoter recognition of the -35 region by RNA polymerase (Collado-Vides et al., 1991). Distant activator sites do exist for σ70 promoters, but a DNA wrapping process to form a large nucleoprotein complex is necessary to allow close approach of the activator from a distant site to contact the RNA polymerase without a looping-out mechanism (Raibaud, 1989).

**B. Promoters recognized by sigma 54.** The common mechanism of using an activator protein and ATP to catalyze formation of a transcriptionally productive open complex for σ54-dependent gene expression has been described previously (Kustu et al., 1989; Popham et al., 1989). Regulation of nitrogen fixation genes has also been reviewed (Gussin et al., 1986). The σ54 promoters are predominantly regulated by activation alone and this is accomplished primarily through a mechanism employing a remote site. Control of transcription at σ54-dependent promoters is primarily accomplished by modulation of the alternative state of an activator that in its active form binds to a distant site from the transcriptional start point. The amount of σ54 itself does not vary much under inducing or noninducing conditions. Sigma 54 confers on core RNA polymerase the ability to bind specifically to a promoter but it does not confer the ability to form open complexes. Transcription initiation from the poised σ54 holoenzyme needs a push-start by the activator bound at a distant site (Collado-Vides et al., 1991; Gralla, 1991). Some of σ54 promoters contain a proximal IHF(integration host factor)-binding site, generally near -40 or -50. Binding of IHF alone does not stimulate transcription, but enhances the effectiveness of a required activator. The distance between the IHF site and the upstream activator site is about 80 bp in all cases,
suggesting that this distance is important (Collado-Vides et al., 1991). A specific example is the involvement of IHF in the regulation of the \textit{glnHp2} promoter of \textit{E. coli} (Claverie-Martin and Magasanik, 1991). Function of the activator protein NtrC is controled by phosphorylation by its partner NtrB (NRII). Phosphorylated NtrC is able to activate the \(\sigma^{54}\) holoenzyme to initiate transcription. The glutamine rich region at the N-terminus of \(\sigma^{54}\) is required for NtrC-dependent isomerization of closed to open complex at the \textit{glnA} promoter (Kustu et al., 1989).

**Differences between sigma 70 and sigma 54 promoters.** Major differences between these two types of promoters are listed in Table 15 (Collado-Vides et al., 1991; Gralla, 1991). The \(\sigma^{54}\) promoters are activated by enhancer-like elements typically centered at -110 from the transcriptional start site, while the activator sites for virtually all \(\sigma^{70}\) promoters are located in a distance that allows communication by direct contact without looping out intervening DNA sequence. No negative control by repression exist for \(\sigma^{54}\) promoters. The energy from ATP hydrolysis is required for DNA melting of \(\sigma^{54}\) promoter to form open complex and to allow productive transcription. Multiple activator binding sites and duplication of activator sites are generally used for \(\sigma^{54}\) promoters, may be because the probability of touching the polymerase during looping is higher for multiple activator sites. Coactivators, e.g. IHF, are often used for stabilization of a looped complex (Hoover et al., 1990). Transcription initiated by sigma 54 RNA polymerase resembles eukaryotic transcription initiation by polymerase II. The looping mechanism provides greater flexibility in evolutionary terms, that is why it seems to dominate the regulatory apparatus in eukaryotes (Collado-Vides et al., 1991; Gralla, 1991).
Table 15. Properties of pol II and $\sigma^{54}$ pol, not shared by $\sigma^{70}$ pol (Gralla, 1991).

<table>
<thead>
<tr>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activated from a distance</td>
</tr>
<tr>
<td>Requires ATP hydrolysis for initiation</td>
</tr>
<tr>
<td>Has acidic, Gln-rich, and leucine zipper domains in activated</td>
</tr>
<tr>
<td>transcription complexes</td>
</tr>
<tr>
<td>Low basal level expression--is regulated predominantly by</td>
</tr>
<tr>
<td>activation rather by repression</td>
</tr>
<tr>
<td>Recognition of basal elements commonly occurs prior to</td>
</tr>
<tr>
<td>activator intervention</td>
</tr>
<tr>
<td>Coactivators are used commonly</td>
</tr>
<tr>
<td>Multiple activator binding sites are generally used</td>
</tr>
</tbody>
</table>
Nucleoproteins. Bacteria contain small, usually basic, abundant histone-like DNA-binding proteins capable of condensing DNA by coiling and wrapping. Much of the following introduction is based on reviews of histonelike proteins of bacteria (Drlica and Rouviere-Yaniv, 1987; Pettijohn, 1988). Considerable amino acid sequence homology exist among bacterial species. The structure and function of two most abundant histone-like proteins, HU and H1, and a less abundant nucleoprotein IHF are discussed below.

A. HU encoded by hupa (90 min) and hupB (10 min) binds to both single-stranded and double-stranded DNA as well as RNA. The major function of HU is probably to play a structural role in wrapping DNA into nucleosome-like particles. HU can stabilize double-stranded DNA against thermal denaturation. Its native form is a heterodimer. HU can decrease the linking number of bound DNA sequences, and DNA becomes more negatively supercoiled when bound to HU. The HU-DNA interaction is dynamic and undergoing rapid exchange. One dimer of HU can cover one turn of the DNA double helix. HU-DNA interactions show little nucleotide sequence specificity, however it is involved in the coiling and wrapping of specific DNA sequences in different regions of the chromosome. HU has been implicated in conferring conformational DNA changes that promote specific protein-DNA interactions (Flashner and Gralla, 1988). HU may bind cooperatively with the other regulatory proteins that have sequence specificity. HU may influence transcription through its DNA wrapping activity and the effects of this on promoting DNA loop formation. HU is also involved in site-specific recombination and DNA replication. It is proposed that HU is a general facilitator protein of protein-DNA interactions that require certain kinds of distortion of the DNA binding sequence.
B. **H1 (H-NS).** A recent review on the histone-like protein H1 has been published (Higgins et al., 1990). The H1 protein is encoded by the *osmZ* gene in a single gene operon at 27 min on *E. coli* genome. The effects of *osmZ* mutants are highly pleiotropic but specific on the expression of many apparently unrelated plasmid and chromosomal genes. Genes could be either repressed or derepressed by H1 (Goodrich et al., 1990; Hulton et al., 1990). For example, mutations in *osmZ* increase *proU* (transport system for osmoprotectant glycine betaine) transcription both at low and at high osmolarity, and alter the osmolarity controlled expression of *ompC* and *ompF* porin genes (Higgins et al., 1988). Other examples include derepression of some temperature-regulated genes (Dorman et al., 1990; Goransson et al., 1990), and the *bgl* operon (Higgins et al., 1988). Besides their effects on transcription, *osmZ* mutations also increase the frequency of spontaneous deletion formation and site-specific recombination events (Lejeune and Danchin, 1990). The influence of H1 on DNA structure is important to the genetic stability of *E. coli* genome. DNA supercoiling is altered in *osmZ* mutants *in vivo* and binding of purified H1 to double-stranded DNA increases the thermal stability of the DNA *in vitro*. It is thought that the effects of *osmZ* on gene expression are exerted through changes of DNA supercoiling in response to environmental signals such as osmolarity, temperature, and oxygen availability (Higgins et al., 1988 and 1990; Hulton et al., 1990). The other different mechanism of transcription affected by H1 is proposed as transcriptional silencing (Goransson et al., 1990) in which H1 may act as a general 'silencer' of transcription. However, the silencing model can not fully account for various observations (Higgins et al., 1990). The H1 protein of *E. coli* has been purified and sequenced (Falconi et al., 1988). H1 differs from other histone-like proteins in that it is neutral rather than basic. Although H1 binds relatively nonspecifically to double stranded DNA, recent evidences have shown cooperative binding of many molecules of H1 to DNA based on a loose consensus sequence TNTNAN (Rimsky and
Spassky, 1990). The DNA-binding domain of H1 appears to be at the N-terminus (Higgins et al., 1990).

C. **IHF** is a small heterodimeric protein encoded by *himA* (38 min) and *hip* (20 min), that is known to induce DNA bending. In *E. coli*, IHF plays an important role in a variety of processes including site-specific recombination, transposition, plasmid replication and gene regulation (Friedman, 1988). The binding of IHF is thought to promote DNA bending so as to increase the probability of loop formation. For a sigma 54 promoter, IHF binds between the activator site and the promoter in a sequence specific manner. The consensus sequence recognized by IHF has be determined as TNPyAANNPNpTTGNT (Craig and Nash, 1984), (A or T)ATCAANNNNTTPu (Leong et al., 1985), PyAANNNNTTGAT(A or T), or AATCAANNANTTPu (Goodrich et al., 1990) in different reports.

**DNA bending.** DNA distortion can be intrinsic to the DNA itself, happening between conformational changes or DNA supercoiling, and DNA bending is often the result of protein binding. DNA bending can either occur as a curve wrapping around a nucleoprotein complex or as a free loop between proteins bound at separate sites. The local DNA conformational changes resulted from DNA bending and their energetic role in regulating many cellular processes are discussed by Travers (1990). Intrinsic bending of DNA in regions of homopolymeric dA-dT sequence which are at least 4 bp long has been observed (Koo et al., 1986). **DNA looping.** Research results for both eukaryotes and prokaryotes have shown that DNA regulatory elements are multipartite in nature and are located at sites distant from the promoter. A DNA looping model has since been used to interpret protein-protein communication in a variety of regulatory systems. Examples of negative control by DNA looping in *E. coli* are reviewed by Adhya (1989) and Gralla (1989), and cases involved in positive control are discussed by Ptashne (1989). In order
to form a loop, protein bound to two DNA sites must face each other in a given geometry. Thermodynamic considerations need to be brought in when judging the feasibility of DNA looping. A DNA segment of 200 bp or less may need energy to overcome the strain induced by the curving and twisting motion during formation of a DNA loop connected by direct contact of two DNA-bound proteins. However, if the two protein binding sites are too far away, the probability of the contact of two sites becomes a limiting factor. DNA periodicity (integral number of turns) exists in DNA loop formation suggesting that protein contact is in a stereospecific manner. Alternatively, proteins bound to separate DNA sites can also be brought into contact through a DNA wrapping and coiling mechanism without looping out intervening sequence. Examples could be found in the case of polymerase-CRP complex formation during transcription initiation (Zinkel and Crothers, 1991).

**DNA supercoiling for the regulation of gene expression**

DNA supercoiling varies in response to environmental stimuli. For example, an increase in the mean negative superhelicity density is observed for DNA isolated from cells grown at high osmolarity (Higgins et al., 1988) and from cells grown anaerobically (Dorman et al., 1988). Supercoiling may also vary with growth phase, and during growth transitions between carbon-rich and carbon-poor media. DNA becomes more relaxed as growth proceeds (Dorman et al., 1988) and when the cells experience a sudden nutrient shift or starvation (Balke and Gralla, 1987). DNA supercoiling has been implicated in the regulation of a number of genes responsive to the environmental stress such as osmolarity, anaerobiosis, pH, starvation, and temperature (Higgins, 1990b). Mutations in *topA* (topoisomerase I) and *osmZ* (H1) increase the negative supercoiling of cellular DNA and permit *proU* (transport system for the osmoprotectant glycine betaine) expression at low osmolarity, when the gene is normally repressed (Higgins, 1988). In contrast, the regulation of *tonB* (a component of iron siderophore uptake system) by
DNA supercoiling during aerobic and anaerobic shift is in the opposite direction to proU; that is, the expression of tonB is repressed by mutations increasing superhelicity, topA and osmZ mutations, and derepressed by mutations in gyr (gyrase) which result in decreased superhelicity (Dorman et al., 1988). How the specificity of regulation by DNA supercoiling is achieved is not clear. The intrinsic difference of each promoter is probably one of the reasons, and different parts of the chromosome may not respond in the same manner to supercoiling changes caused by different stimuli. Furthermore, the binding of histonelike proteins and other DNA-binding proteins may confer local effects of gross topological changes.

**Post-transcriptional control and regulation**

There are a number of post-transcriptional events that can be responsible for the rate at which a protein is synthesized. The structure of mRNA plays a crucial role not only in the control of the constitutive rates of translation but it also participates in a range of translational regulation (McCarthy and Gualerzi, 1990). From a quantitative viewpoint, initiation rather than elongation is usually the rate-limiting step in translation, and the amount of the transcript to be translated also accounts for the level of translation. Therefore we will only consider questions regarding translation initiation and mRNA stability. However, in some cases slowed elongation may depress the overall translation efficiency.

**Translation initiation determined by mRNA sequence and structure.**

The various events of translation initiation have been reviewed (Gold, 1988). And a review concerning specifically translation initiation in *E. coli* is available (Jacques and Dreyfus, 1990). The nature of most influencing factors is conferred by the intrinsic sequence and structure of mRNA. The information extends beyond the mere presence of a Shine-Dalgarno (SD) element followed by a suitable start codon, the importance of
sequence bracketing a few tens of nucleotides from the start codon should not be ignored. A summary of elements known or proposed to control initiation site selection and initiation efficiency are listed by McCarthy and Gualerzi (1990).

A. Initiation site. The initiation site usually consists of the following elements: a start codon of AUG with GUG, UUG, AUU as possible alternatives; and a SD sequence of 3 or more of UAAGGAGGUGA polypurine track located 5-12 nucleotides upstream from the start codon, which is complementary to a section of of the 3' end of the 16S rRNA. However merely a start codon and a properly positioned SD sequence can not explain statistically the strict specificity of translation initiation. A region of 30-40 bp approximating the ribosome-binding site is required to ensure correct translation initiation (Dreyfus, 1988). The nature of this local discriminative information is not clear. Apart from SD element and start codon, preferable sequences that could make favorable contacts with the translational machinery are thought to increase initiation efficiency.

B. mRNA structure. Since initiation regions need to be freely accessible for efficient translation, intramolecular base pairing and thus a higher order of mRNA structure can certainly interfere with initiation. But little is known about the ways that higher order structure can influence the reaction. It has been observed that translation initiation regions (TIR) of differently expressed genes in a polycistronic operon have different affinities for 30 S ribosomal units. A more stable secondary structure with a low affinity for 30 S ribosomal units was found in the TIR of poorly translated genes.

Translational regulatory systems. Both RNA-RNA and RNA-protein interactions are implicated in regulation of translation. Direct blocking of the ribosomal binding site or stabilization of an inhibitory mRNA secondary structure can contribute to inhibition of translation. Anti-sense RNA is also one of the possible inhibitory routes
(Simons, 1988). Positive action can be achieved by destabilization of inhibitory structures in mRNA. RNase III seems to be able to achieve this purpose not only via cleavage but also via binding alone (Altuvia et al., 1987).

**mRNA stability.** The concentration of a certain mRNA often determines the synthesis rate of protein(s) it encodes, therefore the rate of protein synthesis depends on both the rate of transcription and the stability of the transcript. Based on known mechanisms of mRNA decay, the stability of a given mRNA reflects its susceptibility to digestion by enzymes, and the message for such degradation is usually imbeded in its own sequence and structure. A generalization of the mechanisms of mRNA decay in bacteria was summarized (Belasco and Higgins, 1988). Two types of enzymes are known to be responsible for mRNA decay: exonuclease and endonuclease. Two major exonucleases, polynucleotide phosphorylase and RNaseII degrade single-stranded RNA from the 3' end in a progressive manner. No 5'-exonucleases have been reported in bacteria. Two endonucleases, RNaseIII and RNaseE have shown to participate in mRNA degradation. RNaseIII recognizes certain large stem-loop structures, in some cases cleavage by RNaseIII initiates mRNA degradation. RNaseE seems to recognize a consensus sequence lying upstream from a stem-loop structure. But the actual digestion mechanisms are not clear. The resistance or susceptibility to RNase attack can be determined intrinsically by the sequence and the structure of the message, and by some extrinsic factors such as protection by ribosomes and antisense RNA. There is evidence to show that a stem-loop structure can protect and delay upstream RNA from 3' to 5' exonuclease digestion in a processed polycistronic transcript with different stability. It is also true for many 3' end stem-loop transcription terminators. This type of protection does not appear to be sequence specific or need an auxiliary protein. The REP (repetitive extragenic palindromic) sequences have been found to be involved in protection of upstream mRNA from degradation (Newbury et al., 1987). It is not clear how many
endonucleases participate in mRNA decay. But some degrees of specificity (may be broad) should exist for cleavage sites. Ribosomes may sterically protect potential cleavage sites from endonucleases. For some genes, an anti-sense RNA can disrupt translation by binding to the translational start site of a target transcript, such an RNA duplex might also alter mRNA stability by depriving downstream mRNA from ribosomal protection, masking or creating an endonuclease site, or generating a double-stranded barrier for exonuclease (Simons, 1988). Some evidence indicates that gene expression in response to environmental stimuli can be modulated at post-transcriptional level by selective changes in mRNA stability (Melefors and von Gabain, 1988).

**SPECULATIONS**

Studies of pH regulation at the molecular level only started recently; with limited information collected from few examples, a general scheme is far from being established. However, a broad spectrum of pH effects is expected and the global cellular response has been observed in the changes of protein synthesis patterns revealed from two-dimensional gels. So far, the involvement of specific DNA sequences in pH regulation was only revealed by two reported examples, the virG gene of *A. tumefaciens* and the toxR regulon of *V. cholerae*, neither of them exhibit conspicuous similarity to the adi and cad operons of *E. coli* studied in this laboratory. The novelty of the cad operon resides in the lack of specific targeted sequences that could be clearly defined as certain protein binding sites. The only known two loci that affect cad expression, cadC and cadR, have not been proved to have direct contact with the cad operon, although evidence has shown that CadC is a trans activator required for pH induction and it might interact with a region more than 100 bp upstream from the transcription start site. The results of random mutagenesis and *in vivo* DNA footprinting experiments suggest that sequences of the region from the distal high affinity activator binding site to the RNA polymerase
recognition site are all in some way participate in the initiation process. Even with limited experimental results, we could still speculate on possible mechanisms based on knowledge of other regulatory systems with similar features.

**Sigma 70 or sigma 54 promoter**

Almost all *E. coli* promoters analyzed can be assigned to two categories depending on which type of sigma factor it uses, sigma 70 or sigma 54. The induction of transcription from the *cad* promoter by a distal activator site is a strong suggestion for using σ54 RNA polymerase, however the position and sequence of the *cad* promoter more closely resembles a σ70 promoter. One important difference between σ70 and σ54 promoters is the location of the upstream regulatory site. For σ70 promoters, the regulatory protein must somehow contact the promoter without DNA looping, therefore the position of the regulatory site is crucial for its function; on the other hand, the σ54 promoter can be contacted by the activator bound to a distal site through DNA looping mechanism, therefore the length of the intervening sequence is more flexible and the activator can still be functional even if its binding site has been shifted a considerable distance from its original position. However, the phasing of the upstream regulatory site relative to the σ54 promoter should be considered in this kind of experiment. It is more likely that the *cad* operon uses a weak σ70 promoter because the -10 and -35 sequences have been confirmed by mutation experiments. It would be interesting to test if another weak σ70 promoter can be activated by the same regulatory sequence.

**Looping, wrapping, or domino effect**

How is the *cad* promoter activated by a regulator bound to a distal site? A looping mechanism has been detected for repression but not activation of σ70 promoters. The known distal activation of a σ70 promoter involves DNA wrapping mechanism, in which at least one proximal protein binding site either for the same protein or a different one still
exists to serve as direct contact with RNA polymerase. No proximal binding site has been detected in the *cad* upstream sequence, however protection from *in vivo* methylation does occur at position as proximal as -48. The detection of methylation-protected and -enhanced residues in the intervening sequence between the distal high affinity site and the promoter suggests that the looping mechanism is less likely. Another possible mechanism could be presented as cooperative binding of multiple proteins aligned along the sequence in such a way so that the signal initiated from the distal protein could be finally conveyed to the RNA polymerase in a protein-to-protein manner as in a domino effect. However, the *in vivo* methylation pattern is not homogeneous and therefore it is uncertain if the whole region is filled with homologous or heterologous proteins. No difference was detected in the attempts to identify a protein-bounded DNA fragment in gel retardation experiments with total cellular protein isolated from cells grown at pH 5.5 of at pH 8. It could be because the activator CadC was in the membrane fraction.

**The roles of *cadC*, *cadR*, and *osmZ* in regulation of *cad* operon**

Although CadC has not been proved to directly bind to the *cad* upstream region, and the *cadR* locus might consist of more than one gene, it is convenient to assume that CadC is the binding protein and *cadR* encodes one protein CadR in the simplified models described below. It was first thought that CadC and CadR act independently as a pH-responsive activator and a lysine-responsive repressor respectively. However, experimental data suggest that the signal sensed by CadR needs to go through the CadC route to be effective on the *cad* operon. Inspired by the proposed ToxR working model, a model for the action of CadC and CadR was forged. The DNA binding domain resides in the cytoplasmic N-terminal segment of the transmembrane protein CadC and the periplasmic C-terminal segment is responsible for sensing the pH gradient or membrane potential either directly or indirectly. It has been observed that overproduction of CadC from a multicopy plasmid activates the *cad* promoter at pH 8 and therefore loses the pH
control, but the expression of \textit{cadC} is not pH-inducible, therefore it is illegitimate to conclude that pH regulation is executed at the level of CadC concentration. However, it is reasonable to speculate that high concentration of CadC surmount the regular pH-sensing mechanism. Similar situations happen in the ToxR system, in which dimerization of ToxR is essential for binding to the target DNA. Under normal condition (at low ToxR concentration) another membrane-associated protein ToxS is required for stable dimerization of ToxR, but at high ToxR concentration, ToxS is dispensable for stable formation of ToxR dimers. Signal transduction by phosphorylation common to the two-component regulatory systems has not been observed with the ToxR protein, so it may well be that other forms of signal transduction (such as ToxR-ToxS interaction) exist. Whether dimerization also happens to CadC could be determined by CadC-PhoA fusion experiments. It is not known if there is another protein involved in pH sensing and regulation of \textit{cad} operon, but it is clear that extracellular lysine concentration is sensed by CadR which has been suggested as a lysine transport protein and probably a membrane associated protein. CadR behaves negatively in terms of regulation of \textit{cad} promoter by CadC. The strains with mutations in CadR or in the presence of high external lysine concentration, the dampening effect by CadR on CadC disappears. A reasonable speculation is illustrated in Fig. 43 at the end of this Chapter. At low pH, CadC somehow senses the signal and becomes active through a conformational change. This probably is engaged in dimerization or other forms of aggregate that can be more easily achieved at high concentration. In the absence of exogenous lysine, CadR hampers the activation of CadC to certain degree but not completely, because acidic environment alone is enough to elicit the expression of \textit{cad} operon. This interference could be caused by blocking the dimerization of CadC by competitive binding of CadR to the periplasmic domain of CadC. The dynamic competition between two proteins could explain why the lysine effect is not an all-or-none case. When a high external lysine
concentration is available, CadR changes its conformation and loses its ability to bind to CadC and thus unleashes CadC to achieve full induction. The direct interaction of CadR and CadC in the periplasm explains why no DNA binding site for CadR was found and suggests that lysine is sensed by CadR before it is transported into the cell. CadR may be a lysine transporter, but it is probably not as important in transport of lysine per se as in sensing of lysine at low external pH, because no difference of cadaverine excretion is observed in CadR⁺ or CadR⁻ strains and therefore there is probably no difference in lysine uptake. The major route for lysine uptake at low pH is proposed as the function of CadB. The solidification of the above hypothesis still awaits detailed sequence and structure analyses of CadC and CadR.

Recently, osmZ mutants were found to increase the pH 8 activity of the cad promoter substantial and also raise the pH 5.5 activity slightly but only in the CadC⁺ strain. Apparently, histonelike protein H1 works as a repressor to clamp the cad promoter activity especially at pH 8. This leads to a question about the pH regulation by CadC. The original hypothesis it that the pH signal is sensed by CadC which is not activated at pH 8. Now the high pH 8 activity of osmZ mutants suggests that CadC is active though not to a full extent at pH 8. The cad promoter activity can not be observed at pH 8 is because of the suppression by H1. Therefore, it is reasonable to include H1 in the pH sensing machinery. However, H1 is an abundant cytosolic DNA binding protein, it is difficult to imaging it as a specific pH sensor or regulator. The effects of osmZ mutants are pleiotropic, many of them are attributed to changes in DNA structure and stability. The silencing effect of H1 to cad promoter may be due to increased thermostability or unfavorable superhelicity. Therefore the pH signal is probably sensed by CadC but not by H1. H1 just helps to intensify the signal by clamping down the leaky activation from CadC at pH 8. But at pH 5.5, H1 can no longer suppress fully activated CadC and transcription initiates.
Anaerobiosis is known to increase the *cad* promoter activity by about two fold at both acid and alkaline pHs. It is speculated to be a change in DNA supercoiling under different conditions. The *cad* promoter activity was only slightly increased (may not be significant) in cells grown in high osmolarity medium in an early experiment. But the increased levels of supercoiling induced by high osmolarity and anaerobiosis may be different. Direct testing of the correlation between the topA or gyr activities and the air regulation of *cad* operon should provide valuable information. The involvement of *osmZ* in regulation of the *cad* promoter strongly suggests that DNA supercoiling plays an role in this regulatory system. The loss of pH regulation of *cad* operon was observed at high CadC concentration under anaerobic condition, however good pH control was retained at high CadC concentration under aerobic condition. Apparently, clamping by OsmZ is sensitive to oxygen concentration. OsmZ has a stronger effect under aerobic condition. Aerobiosis is supposed to cause more relaxed DNA supercoiling, similar effect caused by binding of OsmZ, which is not prefered for the expression of many genes. Whether more relaxed DNA leads to tighter binding of OsmZ under aerobic condition or *vice versa* is not clear but they apparently work synergistically to suppress the activation of *cad* promoter by CadC. The speculation on the overall function and regulation of the *cad* operon is illustrated in Fig. 43. The models for the regulation of the *cad* operon under different conditions are based on the following assumptions. The transcription initiation of the *cad* promoter requires the activation by the CadC dimer. The positive regulator CadC forms a stable active dimer at low pH and less stable dimer at high pH. The negative regulator CadR can mask CadC dimer formation competitively in the absence of lysine but becomes inactive upon binding to external lysine. Anaerobiosis and *osmZ* mutations increase the expression of *cad* operon because of the higher level of DNA superhelicity.
This research has not been well investigated at the translational level, but some general guidelines of post-transcriptional control and regulation described in the introduction are helpful for future study. So far, differential expression of the *cad* operon is known and some kind of post-transcriptional control does exist and has been discussed in Chapter 3 and 4.
Figure 43.1. pH effect on the expression of *cad* operon. The expression of *cadC* is not affected by pH conditions. The pH signal is somehow sensed by CadC (C) and the active stable dimer of CadC is formed at pH 5.5. In the presence of lysine, the negative regulator CadR (R) can not interfere with the dimer formation of CadC, and under anaerobic condition, higher supercoiling and less binding of OsmZ (Z) all favor the expression of the *cad* operon. The highly expressed lysine decarboxylase (A) and antiporter CadB (B) then cooperate together to reduce the stress of high extracellular H⁺ concentration through the lysine (Lys)-cadaverine (Cad) antiport system. However, at pH 8, a weaker CadC dimer is formed, and the residual binding of OsmZ is enough to block the low level activation from the CadC dimer. Transcription of *cadC* and *cadBA* is indicated by the zigzag-line with arrow. Thicker line means stronger transcription. An × on top of the zigzag-line means the transcription is blocked.
pH 5.5, +Lys, Anaerobiosis

Cytoplasmic membrane

\[ \text{cadC} \rightarrow \text{cadBA} \]

\[ \text{Lys} \rightarrow \text{A} \rightarrow \text{Cad} \rightarrow \text{CO}_2 \]

\[ \text{Lys} + \text{H}^+ \rightarrow \text{B} \rightarrow \text{Cad} \rightarrow \text{CO}_2 \]

pH 8, + Lys, Anaerobiosis

\[ \text{cadC} \rightarrow \text{cadBA} \]

\[ \text{Lys} \rightarrow \text{R} \rightarrow \text{Cad} \rightarrow \text{CO}_2 \]
Figure 43.2. Lysine effect on the expression of the cad operon. In the absence of lysine, CadR is able to bind to CadC competitively, and thus prevent active CadC dimer formation. However, because of the competitiveness of CadC-CadR binding, CadC still has a chance to form CadC dimers, that explains why the cad promoter is still induced at low pH even if no lysine is present in the medium.
pH 5.5, + Lys, Anaerobiosis

Lys

R

Cytoplasmic membrane

cadC cadBA

Z

Z

Lys + H⁺ A Cad + CO₂

Lys H⁺ Cad CO₂

pH 5.5, - Lys, Anaerobiosis

cadC cadBA

Z

Z

X
Figure 43.3. Oxygen effect on the expression of the *cad* operon at low pH. The more relaxed DNA with more bound OsMZ under aerobic condition hinders transcription initiation from the *cad* promoter. However, at pH 5.5, the CadC dimer is stable enough to activate the *cad* promoter substantially. But the weaker CadC dimer at pH 8 cannot overcome the blocking by OsMZ.
pH 5.5, +Lys, anaerobiosis

Lys → R

Cytoplasmic membrane

cadC ≥ ≥ ≥ cadBA

Lys + H⁺ → Cad + CO₂

Lys  H⁺  Cad  CO₂

pH 5.5, + Lys, Aerobiosis

Lys → R

Cytoplasmic membrane

cadC ≥ ≥ ≥ cadBA

pH 8, + Lys, Aerobiosis

Lys → R

cadC ≥ ≥ ≥ cadBA
Figure 43.4. OsmZ effect on the expression of the cad operon. Without the presence of OsmZ, the cad operon can be expressed even at pH 8. Because the activity of cad promoter at pH 5.5 is already saturated, the cad promoter activity in an osmZ- strain is only slightly increased at pH 5.5. The oxygen concentration still affects the promoter activity of the cad promoter in an osmZ- strain: the promoter is more active under anaerobic condition than aerobic condition. Therefore, the air effect and the OsmZ effect may be additive.
pH 5.5, + Lys, Anaerobiosis, OsmZ^-

\[ \text{Lys} \xrightarrow{\text{C, C}} \text{R} \]

\[ \text{cadC} \rightarrow \text{cadBA} \]

\[ \text{Lys} + \text{H}^+ \rightarrow \text{A} \rightarrow \text{Cad} + \text{CO}_2 \]

\[ \text{Lys} \quad \text{H}^+ \quad \text{Cad} \quad \text{CO}_2 \]

---

pH 8, + Lys, Anaerobiosis, OsmZ^-

\[ \text{Lys} \xrightarrow{\text{C, C}} \text{R} \]

\[ \text{Cytoplasmic membrane} \]

\[ \text{cadC} \rightarrow \text{cadBA} \]

---

pH 8, + Lys, Anaerobiosis

\[ \text{Lys} \xrightarrow{\text{C, C}} \text{R} \]

\[ \text{cadC} \rightarrow \text{cadBA} \]

\( \times \)
REFERENCES


Rudd, K.E., W. Miller, J. Ostell, and D.A. Benson. 1989. Alignment of
Escherichia coli K12 DNA sequences to a genomic restriction map. Nucleic Acids
Res. 18: 313-321.

Rudd, K.E., W. Miller, C. Werner, J. Ostell, C. Tolstoshev, and S.G.
Satterfield. 1991. Mapping sequenced E. coli genes by computer: software,

phage for generating single-stranded plasmid DNA. Gene. 45: 333-338.

Purification and physical properties of inducible Escherichia coli lysine

decarboxylase including a segment of it pyridoxal 5'-phosphate binding site.


preservatives on pH homeostasis in Escherichia coli. J. Gen. Microbiol. 130: 2845-
2850.

Sancar A., R.P. Wharton, S. Seltzer, B.M. Kacinski, N.B. Clark, and
148:45-62

Sanders, D. and C.L.Slayman. 1982. Control of intracellular pH-predominant
role of oxidative metabolism, not proton transport in eukaryotic microorganism

Sanger, F., and A.R. Coulson. 1978. The use of thin acrylamide gels for DNA

Sasse-Dwight, S., and J.D. Gralla. 1990. Role of eukaryotic type functional
domains found in the prokaryotic enhancer receptor factor σ54. Cell 62: 945-954.

Schuldiner, S., S. Agmon, J. Brandsma, A. Chohen, E. Friedman, and
E. Padan. 1986. Induction of SOS functions by alkaline intracellular pH in

Shiota, S., and H. Nakayama. 1989. Micrococcus luteus homology of the
Escherichia coli uvrA gene: identification of a mutation in the UV-sensitive mutant

Silhavy, T. J., Beckwith, J. R. 1985. Uses of lac fusions for the study of


