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.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

REGULATION OF THE EARLY STAGE OF THE TETRAPYRROLE BIOSYNTHETIC PATHWAY IN *ESCHERICHIA COLI*

by

WEI HUA

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1997

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 9807942

Copyright 1997 by
Hua, Wei

All rights reserved.

UMI Microform 9807942
Copyright 1997, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized
copying under Title 17, United States Code.

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

©1997

WEI HUA

All Rights Reserved

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This manuscript has been read and accepted for the Graduate Faculty in Biology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

8/27/97

Date

Chair of Examining Committee

Dr. Sharon Cosloy

9/2/97

Date

Executive Officer

Dr. Charlotte Russell

Dr. Dani McBeth

Dr. Simon Simms

Dr. Denis Liveris

Supervisory Committee

The City University of New York

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT

REGULATION OF THE EARLY STAGE OF THE TETRAPYRROLE BIOSYNTHETIC PATHWAY IN ESCHERICHIA COLI

by

WEI HUA

Mentors: Professor Sharon Cosloy and Professor Charlotte Russell

Heme is the prosthetic group for a number of redox enzymes such as cytochromes, catalases, and peroxidases. 5-Aminolevulinic acid (ALA) is the first committed precursor in the tetrapyrrole biosynthetic pathway. There are two completely different pathways to synthesize ALA: C_4 and C_5. *E. coli* uses the C_5 pathway in which ALA is produced by three consecutive enzymatic reactions from glutamate. Glutamate and $tRNA^{Glu}$ are converted by glutamyl-tRNA synthetase (GTS), the gene product of *gltX*, to glutamyl-tRNA. Glutamyl-tRNA is reduced to glutamate semialdehyde (GSA) by glutamyl-tRNA reductase (GTR), the gene product of *hemA*. GSA is converted to ALA by GSA aminotransferase, the gene product of *hemL*. Because glutamyl-tRNA is also used in protein synthesis and GSA can be converted to ALA non-enzymatically, it is highly probable that GTR is the major site of regulation of ALA synthesis in *E. coli*. A study was undertaken to understand the regulation of this gene.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Analysis of the 5' upstream regulatory region of the *hemA* reveals many interesting features. There are three putative σ^{70} promoters (P1, P2 and P3), and one σ^{32} promoter (P32). There are binding sites for FNR, CRP, NarL and ArcA regulatory proteins, and two overlapping stem loops that resemble the $tRNA^{Glu}$ stem loop. This suggests that *hemA* regulation is complex. We constructed six fragments of the *hemA* 5' upstream region using PCR, and cloned these fragments into the vector, pKK232-8, in front of the promoterless gene for chloramphenicol acetyltransferase (CAT). The promoter strength of these fragments was measured by CAT ELISA assay. The largest fragment (pWH515), which contains 496 bp upstream from the transcription initiation site had the highest level of expression. This fragment contains the putative P1, P2, P3 and P32 promoters. The second largest fragment (pWH359), which contains 340 bp upstream from the transcription initiation site, showed the next highest level of expression. This fragment also has all putative P1, P2, P3 and P32 promoters, but lacks 156 bp upstream region of P3. The other fragment (pWH239), which contains 220 bp upstream from the transcription initiation site, showed the next highest level of expression. This fragment harbors only P1 and P2. The promoter strength dropped with the size of the fragment in the order: pWH515, pWH359, pWH239. A fragment (pWH239*) which is identical to the one in pWH239 except for a 1 bp mutation in the -10 sequence of P1, and 2 bp changes away from P1 and P2, showed very little expression. The fragment (pWH101), which contains 82 bp upstream from the transcription initiation site, showed very little expression. This fragment has P1 only. Another fragment (pWH415) containing only P2 and P3 in a 415 bp upstream region from -81 to -496 bp showed very little expression. These results indicate that P1 is necessary but not sufficient for *hemA* expression; that 138 bp

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

upstream of P1 are essential for high level of expression, and that P2 and P3 are weak promoters.

The role of global regulators FNR, CRP, NarL and ArcA in *hemA* regulation was examined by using the plasmids mentioned above. In studies with pairs of isogenic strains (contaning plasmid) which differ with respect to their production of a functional ArcA, expression of CAT was 2-fold higher in strains with ArcA under aerobic conditions. ArcA appears to be an aerobic activator. *hemA* has a FNR binding site overlapping P1. Expression of CAT was about 2-fold higher in an *fnr*$^-$ than in its isogenic wild type strain under anaerobic conditions. FNR acts like an anerobic repressor in *hemA* expression. These results correlate with the finding that CAT expression with all of these plasmids was higher in cells grown aerobically as compared to cells grown anaerobically.

Aerobic growth on glucose caused a 2-fold decrease in CAT expression in stationary phase but not in log phase. For *crp*$^+$ strains, the expression of CAT decreased about 2-fold with glucose. For *crp*$^-$ strains, the expression of CAT showed no difference between cells which were grown with or without glucose. This suggests that the glucose effect is mediated by CRP.

NarL is the regulator of a two-component regulatory system, involved in nitrate reduction. When pairs of isogenic strains (*narL*$^+$/*narL*$^-$) which harbor pWH515, pWH359 or pWH239 were analyzed, there was a small increase in CAT expression in all strains that were grown with nitrate (1.3-fold). The putative NarL binding site is located on the insert of pWH515, the same site which is mutated on the insert of pWH359, and which is not present on the insert of pWH239. CAT expression from all of these plasmids showed the same increase in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

narL$^+$ and *narL*$^-$ strains. Therefore there is no evidence, at this time, that NarL is involved in the regulation of the *hemA* gene.

Growth of strains harboring the plasmids on rich medium gave higher CAT expression than growth on minimal medium.

Starvation for ALA or heme increased CAT expression more than 2-fold. Thus, ALA or heme starvation seems to stimulate *hemA* expression. The order of promoter strength was a function of the size of the upstream region regardless of the strains and growth conditions used.

The levels of inhibition or enhancement of CAT expression driven by *hemA* promoters are not dramatic (about 2-fold). Thus control of *hemA* expression by regulation of the initiation of transcription may slightly modulate the ability of an *E. coli* cell to synthesize heme in different environments, but the major control or additional control may be at one or more other steps, such as the post-transcriptional level.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGMENTS

This dissertation is dedicated to my parents and Tony for their love, support and belief in me.

I would like to thank the members of my supervisory committee, Dr. Dani Mcbeth, Dr. Denis Liveris and Dr. Simon Simms for their valuable advice and cooperation. I appreciate their time and effort.

I would like to thank my colleagues and friends: Susan Pido, Wei Chen, Sheryl Chen, Yun Peng, Jie Rao, Xianglong Li, Libo Zhou, Yan Yan, Chengyi Zhou, George Ballesteros, Janice Riley, Gary Pestano and Weiyi Chen for their helpful discussions, generous contributions, and warm friendship. They made my everyday in the laboratory enjoyable.

In particular, I want to express my deep gratitude to my mentors, Dr. Sharon Cosloy and Dr. Charlotte Russell. Their guidance, encouragement and support not only helped me complete my training in Ph. D. program; but also will benefit me later in my life. They set good examples for me to be my best.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

	Page
Introduction	1
Materials and Methods	38
Results	52
Discussion	74
Bibliography	87

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

	Page
Table 1. The ArcA-P box, or primary sequence DNA binding site consensus for ArcA-P	21
Table 2. Primers used in this study	39
Table 3. Strains used in this study	40
Table 4. Plasmids used in this study	42
Table 5. Promoter strengths of various fragments of the upstream region of *hemA*	60

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES

	Page
Figure 1. The structure of the heme molecule.	2
Figure 2. Schematic structure of (a) hemes, (b) chlorophylls and bacteriochlorophylls c, d, and e, (c) bacteriochlorophylls a, b, and g, (d) siroheme, (e) corrinoids, and (f) coenzyme F_{430}.	2
Figure 3. The tetrapyrrole biosynthesis pathway in *E. coli*.	3
Figure 4. Location of *E. coli* heme genes on the genetic map.	5
Figure 5. 5' upstream region of *hemA*.	10
Figure 6. Energy-generating metabolic pathways of *E. coli* during (a) aerobic growth and (b) anaerobic growth in the presence or absence of alternative electron acceptors.	16
Figure 7. Steps involved in the positive control of transcription.	18
Figure 8. Mechanism for the anaerobic formation of active FNR, and target functions of FNR as an activator(+) or a repressor (-).	23
Figure 9. The structure of the FNR monomer.	25
Figure 10. Location of FNR152 recognition sites on the pFNRCS consensus site plasmid at the FNR-activated *narK*, *narG* and *dmsA* promoters and at the *hemA* promoter that is repressed by FNR.	26
Figure 11. NarL, NarP, NarQ and NarX dependent control for response to nitrate and molybdate availability.	28
Figure 12. Potential NarL binding sites in different NarL sensitive promoters.	30
Figure. 13. Proposed role of IHF in NarL activation of *narGHJI* expression.	31
Figure 14. Interaction of CRP and RNA polymerase.	33
Figure 15. Sequences of CRP binding sites and -10 sequences and the distance between them are shown for 11 promoters that are activated by cAMP-CRP.	36

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Page

Figure 16. The structure of the CRP monomer. 37

Figure 17. Strategy for subcloning the 5' upstream region of *hemA*. 46

Figure 18. Various fragments of the 5' upstream region of *hemA*. 53

Figure 19. Sequence of the insert in pWH239*. 54

Figure 20. The position of *hemA* primers. 55

Figure 21. Promoter strength of *hemA*. 59

Figure 22. Oxygen effect on promoter strength. 62

Figure 23. ArcA effect on promoter strength. 63

Figure 24. FNR effect on promoter strength. 65

Figure 25. Glucose effect on promoter strength. 67

Figure 26. CRP effect on promoter strength. 68

Figure 27. Nitrate and NarL effect on promoter strength. 70

Figure 28. Medium effect on promoter strength. 71

Figure 29. ALA and hemin starvation effect on promoter strength. 73

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

INTRODUCTION

1. Tetrapyrrole Biosynthetic Pathway

The basic structure of biologically relevant cyclic tetrapyrroles is often the porphyrin ring. Heme is protoporphyrin IX containing ferrous iron (Fe^{++}) (Figure 1). Heme, the most common form of iron in biological systems, is the prosthetic group for a number of redox enzymes such as cytochromes, catalases, nitrate and nitrite reductases and peroxidases. Siroheme is the prosthetic group of sulfite reductase and a nitrite reductase. Modified cyclic tetrapyrroles are the active moieties of the chlorophylls and vitamin B_{12} (Figure 2). The enzymes involved in the biosynthesis of these substances are highly conserved from bacteria to humans.

5-Aminolevulinic acid (ALA) is the first universal committed precursor in this pathway. There are two completely different pathways to synthesize ALA: C_4 and C_5 (Beale and Castelfranco, 1974). In the C_4 pathway, ALA is synthesized from glycine and succinyl-CoA by ALA synthase. In the C_5 pathway, ALA is produced in three consecutive enzymatic reactions from glutamate: glutamate and $tRNA^{Glu}$ are converted by glutamyl-tRNA synthetase (GTS), the gene product of *gltX*, to glutamyl-$tRNA^{Glu}$ which is converted by glutamyl-tRNA reductase (GTR), the gene product of *hemA*, to form glutamic semiadehyde (GSA), which is then converted by GSA aminotransferase, the gene product of *hemL*, to form ALA (Figure 3) (Beale, 1996). The C_5 pathway is widely distributed in plants, most bacteria such as *E. coli* (Li *et al.*, 1988; 1989a), *Salmonella typhimurium* (Elliott, 1989), *Bacillus subtilis* (O'Neill *et al.*, 1989), and archae, such as *Methanobacterium thermoautotrophicum* (Friedmann *et al.*, 1987).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

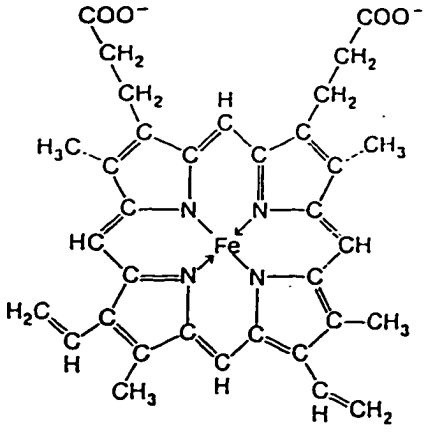

Figure 1. The structure of the heme molecule.

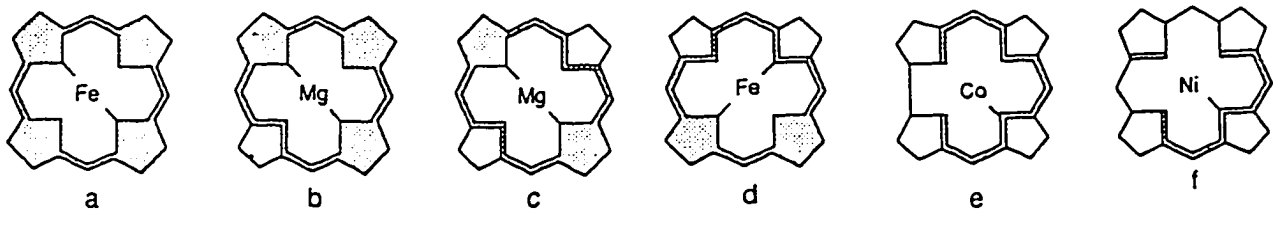

Figure 2. Schematic structure of (a) hemes, (b) chlorophylls and bacteriochlorophylls c, d, and e, (c) bacteriochlorophylls a, b, and g, (d) siroheme, (e) corrinoids, and (f) coenzyme F_{430}, showing the various ligated metals and emphasizing the relative state of oxidation/reduction. Double bonds and conjugated double bond systems are shaded gray (Friedmann and Thauer, 1992).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

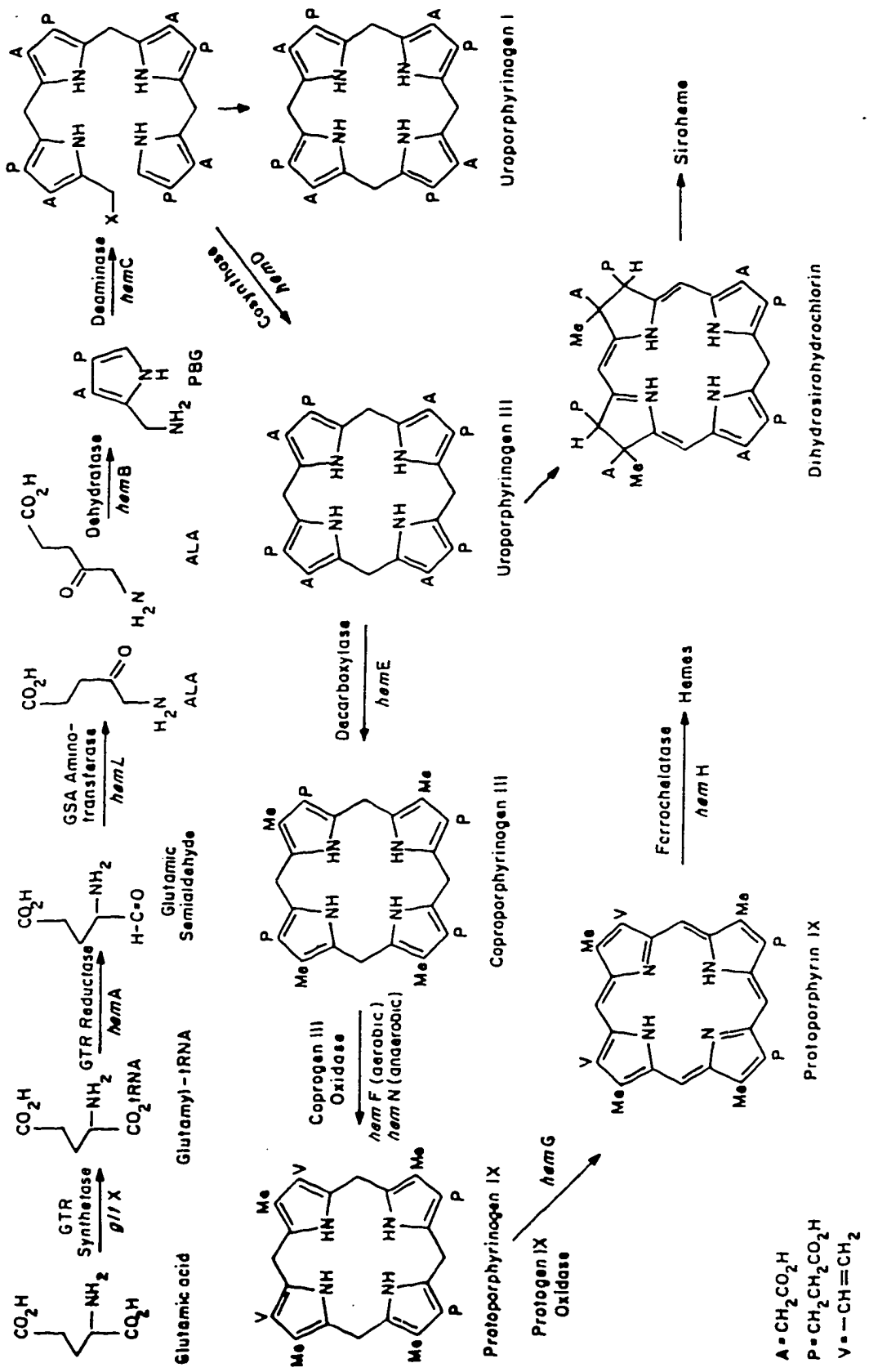

Figure 3. The tetrapyrrole biosynthesis pathway in *E. coli*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The C_4 pathway is present in humans (Jordan, 1991), animals (May *et al.*, 1986), yeast (Urban-Grimal *et al.*, 1984) and some bacteria such as *Bradyhizobium japonicum* (McClung *et al.*, 1987).

Glutamyl-tRNA participates in protein synthesis as well as in the synthesis of aminolevulinic acid (ALA). There are other branch points in the pathway: at uroporphyrinogen III to form cobalamin, coenzyme F_{430} and siroheme; and at protoporphyrin IX to form chlorophylls, bacteriochlorophylls and hemes. *E. coli* does not make vitamin B_{12}, F_{430}, or chlorophylls, but does make heme and siroheme.

Tetrapyrroles are involved in important metabolic processes, such as respiration and cell detoxification. The inner cytoplasmic membrane of bacteria is the site of oxidative phosphorylation. The flavoproteins, quinones, iron-sulfur "FeS" proteins, and cytochromes are inserted into it. All the cytochromes, the components of the electron transport chain in cell respiration system, utilize heme.

Catalases and peroxidases are important for cell detoxification. Catalase catalyzes the decomposition of hydrogen peroxide formed by aerobic oxidation of reduced flavoproteins and oxygen. Peroxidase can use molecular oxygen to remove hydrogen atoms from specific organic substrates in an oxidative reaction that produces hydrogen peroxide. Siroheme is the prosthetic group of sulfite reductase, which catalyzes the reduction of ferric citrate; and of a nitrite reductase, which is involved in anaerobic nitrite respiration (Siegel *et al.*, 1973).

The tetrapyrrole biosynthetic pathway in *E. coli* has been genetically defined by the isolation and characterization of respiratory mutants which are unable to synthesize heme. The genes are widely-scattered on the genome except for *hemC* and *hemD*, which are linked and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

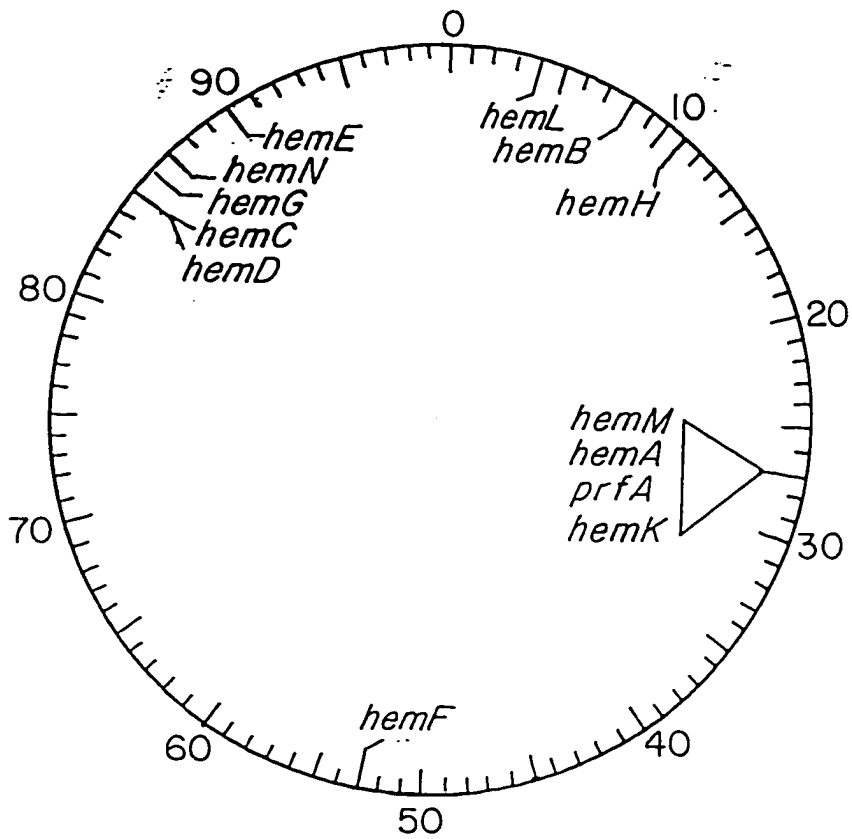

Figure 4. Location of *E. coli* heme genes on the genetic map

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

appear to form an operon (Figure 4). The structural genes which encode the enzymes in the pathway in *E. coli* have been cloned. They are: *hemA* (Li *et al.*, 1989b; Verkamp and Chelm, 1989; Drolet *et al.*, 1989), *hemB* (Li *et al.*, 1988), *hemC* (Thomas and Jordan, 1986), *hemD* (Jordan *et al.*, 1987; Sasarman *et al.*, 1987), *hemE* (Nishimara *et al.*, 1993), *hemF* (Troup *et al.*, 1994), *hemN* (Troup *et al.*, 1995), *hemG* (Jordan *et al.*, 1988), *hemH* (Miyamoto *et al.*, 1991), and *hemL* (Elliott *et al.*, 1987; Grimm *et al.*, 1991) (Figure 3). Another gene which is involved in this pathway has been cloned and designated *hemK* (Nakayashiki *et al.*, 1995). It is located at 27 min where it might be part of the *hemA-prfA-hemK* operon. Its function is still unknown. It might be involved in the oxidation of protoporphyringen to protoporphyrin IX. There is another interesting gene at 27 min (Li *et al.*, 1989b). It was designated *hemM* (Ikemi *et al.*, 1992). *hemA and hemM* are 213 bp apart and divergently transcribed. The role of *hemM* in ALA synthesis is still unknown (Chen *et al.*, 1994).

2. *hemA* in Bacteria Which Use C_5 Pathway

The C_5 pathway is found in plants, algae and most bacteria. They all share the three steps to synthesize ALA. The *hemA* gene in these organisms is highly conserved. In bacteria, the genes which are involved in tetrapyrrole synthesis have been studied most extensively in *E. coli*, *Salmonella typhimurium* and *Bacillus subtilis.*

hemA in *E. coli* is followed immediately by *prfA* which encodes RF1, polypeptide chain release factor 1. There are only 41 nucleotides between the stop codon of *hemA* and the starting codon of *prfA*. *hemK* is located at 27 min on the linkage map of the *E. coli* chromosome immediately after *prfA*. The starting codon of *hemK* overlaps the stop codon of *prfA*. Thus, these

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

three genes are probably in one operon: *hemA-prfA-hemK* in *E. coli*. (Nakayashiki, 1995). Elliott (1992) showed that part of the *hemK* gene from *S. typhimurium* could be cotranscribed from the promoter of *hemA*.

Salmonella typhimurium is a Gram-negative, facultative enteric bacterium like *E. coli*. The *hemA* gene in *Salmonella typhimurium* has been cloned and sequenced and it is highly homologous to the *E. coli hemA* gene (Elliott, 1989). The main difference is that the *hemA* gene in *Salmonella typhimurium* has two insertions, each is about 100 bp, one in the *hemA* promoter region and the other in the far upstream region. The *prfA* gene in *Salmonella typhimurium* has almost the identical sequence as the *prfA* of *E. coli*. The *hemA* and *prfA* belong to the same operon in *Salmonella typhimurium*.

Bacillus subtilis is a Gram positive, aerobic bacterium. The *hemA* gene has been cloned and sequenced (Petricek *et al.*, 1990). The *B. subtilis* glutamyl-tRNA reductase, GTR, showed 34% identity with *E. coli* GTR. In *Bacillus subtilis, hemA* is part of an operon, *hemAXCDBL*. This operon encodes enzymes for synthesis from ALA to uroporphyrinogen III. The interesting *hemX* gene has not been shown to exist in *E. coli*. The gene product of *hemX*, HemX, is a membrane protein. It is not directly involved in ALA or tetrapyrrole synthesis. However, HemX negatively affects the steady-state cellular concentration of GTR (Schroder *et al.*, 1994). The reason that gene organizations are so different for *E. coli* and *Bacillus subtilis* is still unknown. It is possible that the separation of the genes in *E. coli* allows a greater degree of flexibility in the regulation under a wider range of growth conditions, such as aerobic or anaerobic conditions. On

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the other hand, *Bacillus subtilis* is an aerobic bacterium that may have simpler regulatory demands (Beale, 1996).

3. Regulation of *hemA* in *E. coli*

E. coli hemA has been cloned and sequenced (Li *et al.*,1989b; Verkamp and Chelm, 1989). The DNA sequence-derived amino acid sequence is for a protein of 46 kDa minimum MW (Li *et al.*, 1989). GTR from *E. coli* has been difficult to isolate, purify and characterize. Its mRNA is apparently very unstable and the enzyme also appears to have a short half-life (unpublished observations). There is substantial evidence that *hemA* is the structural gene for GTR in *E. coli* (Avissar *et al.*, 1989; Ilag *et al.*, 1991). The size of GTR varies among different organisms. GTR from *Synechocystis* is a 350 kDa protein with 47.5 kDa subunits (Verkamp *etal.*, 1992; Rieble and Beale, 1991). GTR from *Chlamydomonas* is a 130 kDa protein (Chen *et al.*, 1990). There was a report that GTR and GTS could form a 200 kDa complex in the presence of $tRNA^{Glu}$, glutamate and ATP in *Chlamydomonas reinhardtii* (Jahn, 1992). Chen et al (1996) showed that there were two types of GTR complex: 175 kDa and 117-126 kDa in *E. coli*. The 175 kDa protein may represent different complexes of GTR, GTS and $tRNA^{Glu}$; the 117-126 kDa protein may be an dimer of GTR associated with $tRNA^{Glu}$ or a complex of GTR, GTS and $tRNA^{Glu}$.

Glutamyl-$tRNA^{Glu}$ participates in protein synthesis as well as in ALA synthesis, therefore the enzyme, glutamyl-tRNA synthetase (GTS) is probably not the major site for regulation of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ALA synthesis. GSA aminotransferase would be a poor choice for regulation because the substrate, GSA, can form ALA non-enzymatically. *hemA*, the structural gene for GTR, is a good candidate for the major site of regulation of ALA synthesis in *E. coli*. This regulation may be at the transcriptional, translational or enzymatic level; it also may be a combination of these.

Phillipp-Dormston and Doss (1975) found that *E. coli* could accumulate all the porphyrin intermediates in the tetrapyrrole biosynthetic pathway when the growth medium was supplemented with ALA. This suggests that regulation of the pathway is at ALA synthesis.

Over-expression of *hemA* accumulates ALA and porphyrins (Chen *et al.*, 1994). This observation suggests that GTR may be the rate-controlling step of ALA synthesis. Rivera and Walker (1995) observed that over-expression of a rat cytochrome in *E. coli* which creates a demand for more heme in the cell resulted in an increase in cellular heme without accumulation of pathway intermediates. Woodard and Dailey (1995) made a similar observation and concluded that the heme biosynthesis pathway is regulated at ALA levels. They also suggested that regulation of ALA levels might be mediated by a heme sensor.

Heme seems to be a feedback regulator for tetrapyrrole biosynthesis. In a *hemB* mutant which is hemin-permeable, ALA accumulation was inhibited by adding more hemin to the medium and this inhibition was concentration-dependent (Chen, unpublished data).

Analysis of the 5' upstream region of the *hemA* gene shows some very interesting features which appear to be relevant to the regulation of *hemA* expression (Figure 5):

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-513 TAACGCTGACCTTTATTGTCGACTAACTGCACGTTACGCGGTTGAGCATT

-463 CAGCTCCAGTTCCGTGCTGCCCAATGGGTTAGTGAGCAGCAGACGGTAGC

-413 GATCCTGGCCGGTTTGCTGCCAGAAAAAGCGGGCGTACACTTTTTGTTGGT

NarL

-362 CAGAAATATAAGCGAACGCGCCGCGAGTCTGA**TACTGAT**TAAGATTGCG

P3 P32

-313 CACGTCTTGCTGATGCTGACGCCATTGTGGCGAATCCGGGCTTTTGCCAGG

P3 P32 SL1

-262 ACCTTTGGGCGTGGTAACGGAACAGGCAGTGAGCACAAGAGCAGCCAGC

SL2

-212 GT*AGCAGGCGGATAAGACGA*AAATCGGGCAGGGGCATAGTGATGACAAG

cAMP-CRP

-163 TCCT**TGAGA**TACGTT**GCAGT**TATAACCCTTAATGCTAGCGTTACCGTCC

P2

ArcA

*IT2

-114 GCTATCGTCTATGTTCAAGTTGTCTTAATTGCCAGAATCTAACGGCTTTCG

P2

FNR

-63 GCAATTACTCAAAAGGGGGCGCTCTCTTTTATTGATCTTACGCATCCTG

P1

*IT1

-13 TATGATGCAAGCAGACTAACCCTATCAACGTT**GGT**ATTATTTCCCGCAG

P1 +1 SD

+37 **ATG**ACCCTTTTAGCACTCGGTATCAACCATAAAACGGCACCTGTATCGCTG

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Legend:
Promoters: underline (P)
Initiation of transcription: * (IT) and underline
cAMP-CRP binding site: bold and underline
NarL binding site: bold and underline
ArcA binding site: bold and underline
FNR binding site: bold and underline
Stem-loop: boxed (SL)
Inverted repeat: double underline
Shine-Dalgarno sequence: bold and double underline
Initiation of translation (ATG): bold and underline

Figure 5. 5' upstream region of *hemA*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a) There are three putative σ^{70} promoters (Verkamp and Chelm, 1989; Choi *et al.*, 1996). P1 has the -10 sequence: TATGAT at nt -8 to -13; the -35 sequence: TTTATTGATC at nt -26 to -35. P2 has the -10 sequence: TATCGT at nt -107 to -112 ; the -35 sequence: TTAATG at nt -130 to -135. P3 has the -10 sequence: GATGCT at nt -297 to -302; the -35 sequence: TTAAGA at nt -319 to -324. They are poorly homologous to the -10 sequence: TATAAT and -35 sequence: TTGACA, consensus sequences recognized by the σ^{70}-RNA polymerase of *E. coli*. This may indicate these are positively regulated promoters. There are two initiation of transcription sites corresponding to P1 and P2. The first one is at -1 bp and the second one is at -94 bp. Another possible promoter (P32) has the -10 sequence: CGCCAT at nt -289 to -294; the -35 sequence: ATTGCG at nt -314 to -319. This is a putative σ^{32} promoter. It is poorly homologous to the consensus sequence, the -10 sequence: CCCCAT and -35 sequence: CTTGAA.

(b) There is a putative cAMP-CRP binding site with the sequence: TGAGAnnnnnnGCAGT at nt -144 to -159. ALA synthesis may be under catabolic control.

(c) There is a putative FNR binding site located near the -35 sequence of P1 at nt -18 to -31 (TTGATnnnnCGCAT). FNR is a transcriptional activator for some genes and repressor for others in anaerobic respiration.

(d) There is a putative NarL binding site (TACTGAT) at nt -324 to -330. NarL is the transcriptional activator for several genes in anaerobic respiration when nitrate is present.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(e) There is a putative ArcA binding site (TCTTAATTGC) at nt -83 to -92 between P1 and P2. ArcA is the anaerobic repressor for some genes, it can also be an aerobic activator for other genes under other circumstances.

(f) There are two stem loops. They overlap in a tail-head fashion: CCGGGCTTTGCCCAGG at nt -263 to -278 and CCAGGACCTTTGGGCGTGG at nt -242 to -267. There is significant homology between these loops and a stem loop from $tRNA^{Glu}$ (CCAGGACACCGCCCUU UCACGGCGGUA) and an upstream sequence in GTS (CCAGGATTTGCCGGTTGTCCGG) (Brun *et al.*, 1990). These features may reflect coordinate regulation of protein synthesis and tetrapyrrole synthesis by $tRNA^{Glu}$, GTR and GTS.

(g) There is another inverted repeat: TAACGGCTTTCGGCAAT at nt -60 to -75. This could be a binding site for a regulatory protein.

All these features suggest that control of regulation of *hemA* expression is complex. It may be responsive to conditions of growth such as aerobiosis or anaerobiosis and to other elements of regulation.

Drolet et al (1989) carried out primer extension experiments and identified P1. Verkamp and Chelm (1989) mapped two RNA 5' ends which correspond to P1 and P2 by using S1 nuclease protection assays.

The FNR binding site was localized to a 12 bp region within the *hemA* promoter (-24 to -35) by *in vitro* DNase protection assays under anaerobic conditions. This region only contains a FNR half-site (Melville and Gunsalus, 1996).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Choi et al (1996) investigated *hemA* transcription in both *E. coli* and *Salmonella typhimurium*. During starvation, the expression of *hemA-lacZ* in a *hemA* mutant strain was elevated 2 to 6-fold and the *arcA* gene effect was not clear. They also found that the sequence for high level expression of *hemA* was within 129 bp upstream of the major promoter (P1) transcriptional start site. Mutants defective in P1 had greatly reduced *hemA-lacZ* expression both in the presence and absence of ALA. Primer extension analysis showed that P1 and P2 were not expressed at significantly higher levels in ALA starved cultures. However, they found expression from P32 (their P3) under growth conditions of ALA starvation when one *hemA* probe was used.

McNicholas et al (1996) fused the two putative *hemA* promoters, identified by Verkamp and Chelm (1989) to *lacZ*. They reported that the distal promoter, P2, had "weakly constitutive" promoter activity while the proximal promoter, P1, was the stronger promoter.

Cotter et al (1992) studied the effect of iron limitation on *hemA-lacZ* expression. They concluded that the regulation of *hemA* expression did not depend on iron availability.

4. Energy Generation and Regulation in *E. coli*

When there are several energy sources available, *E. coli* is able to choose the most efficient energy generation system to ensure maximal exploitation. This fine-tuning cellular metabolism is achieved by complex signal transduction systems, such as the global regulatory proteins FNR, Arc, CRP and Nar systems (Iuchi *et al.*, 1991).

Bacterial cells must respond to external signals very appropriately and rapidly in order to survive. In prokaryotes a signal transduction system usually has two components: a sensor

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

protein and a regulatory protein (Stock *et al.*, 1989). The sensor protein which often exists in the cytoplasmic membrane monitors the changes in the enviroment, and usually transfers a phosphoryl group to the regulatory protein which exists in the cytoplasm. Then the activated regulatory protein will have an effect on gene expression. In *E. coli*, 20 different signal transduction systems have been identified (Parkinson and Kofoid, 1992).

Carbon sources provide the energy for bacterial cells. In the process of breaking down the carbon source, energy is stored in the form of ATP produced by electron transport chains with electron acceptors. In aerobic respiration, the electron acceptor is oxygen; while in anaerobic respiration, electron acceptors can be nitrate, nitrite, dimethyl sulfoxide (DMSO), trimethylamine-N-oxide (TMNO) or fumarate. The preferential order for the usage of different electron acceptors is: oxygen>nitrate>nitrite>DMSO>TMNO>fumarate (Gunsalus, 1992) (Figure 6). The most wasteful process of energy generation is fermentation in which there is no exogenous electron acceptor.

Since *E. coli* has the ability to use many different electron acceptors, there must be a complex regulatory system to control this hierarchical expression. The mechanism is still under investigation. The complexity of anaerobic metabolism and its regulatory system are essential for survival in a changing environment (Cole, 1996).

5. Positive and Negative Control of Gene Expression

The regulatory proteins in the signal transduction system are usually DNA binding proteins. The function of a DNA binding protein depends on whether it blocks or stimulates binding of RNA polymerase to a promoter. It can be an activator or a repressor.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(*a*) *Aerobic* (*b*) *Anaerobic*

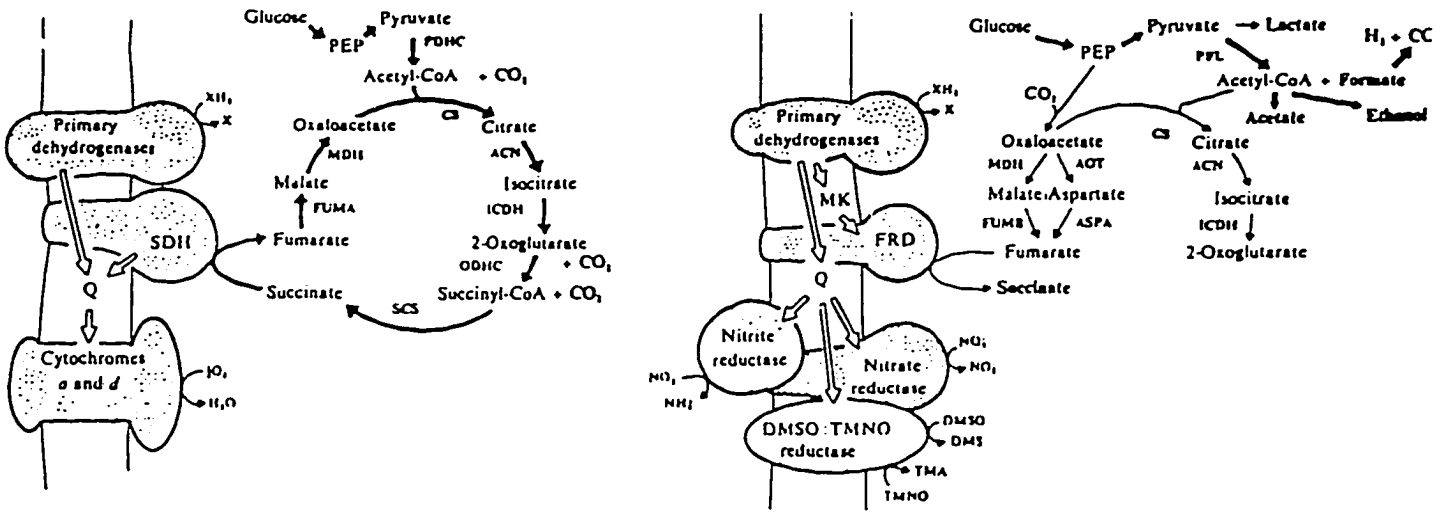

Figure 6. Energy-generating metabolic pathways of *E. coli* during (a) aerobic growth and (b) anaerobic growth in the presence or absence of alternative electron acceptors. Line thickness denotes relative carbon flux (Spiro and Guest, 1991).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A gene under positive control means that its expression needs regulatory protein(s) to bind at a specific site in the upstream region to enhance the initiation of transcription by RNA polymerase. Adhya and Garges summarized several common steps for positive control (Adhya and Gatges, 1990) (Figure 7):

a. A signal, which is normally the intracellular or extracellular concentration of a small molecule, is perceived by a sensor.
b. The signal is transmitted to a regulatory protein (activator protein).
c. The signal transduction changes the conformation of the activator protein, either by noncovalent or covalent modification.
d. The altered activator protein binds to the specific DNA site.
e. The DNA-protein interaction catalyzes the binding or activity of RNA polymerase to facilitate transcription initiation at the promoter.

The only difference between positive control and negative control would be at step e. With the scenario of negative control, the DNA-protein interaction blocks the DNA to which RNA polymerase binds. The repressors can physically prevent RNA polymerase from binding and initiating RNA synthesis at the promoter. In the *E. coli* CytR regulon (cytosine), the CytR repressor binds to operators overlapping the DNA sequence to which RNA polymerase binds; therefore it interferes with the formation of a transcription complex (Mollegaard *et al.*, 1993). Several transcriptional regulators can switch between activator and repressor depending upon both the promoter and the cellular context (Rasmussen *et al.*, 1996). For example, cAMP-CRP can interact with CytR to form a complex which has dual functions. In the proximal position to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Please Note

Page(s) not included with
original material and unavailable
from author or university. Filmed as received.

18

UMI

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The DNA binding site provides specificity for the regulatory protein. A 20 bp specific sequence has about 10^{-12} probability to occur randomly. Almost all known regulatory proteins have consensus sequences to which they bind. One consequence of regulatory protein interacting with DNA may be DNA bending. The function of DNA bending is to bring relevant factors which may be far apart together to form a complex to facilitate transcription initiation (Giladi *et al.*, 1992). The involvement of DNA bending in NarL activation of *narGHJI* operon expression will be discussed below.

6. Oxygen Effect and the Roles of ArcA and FNR in Gene Rgegulation

The ability to sense and adapt to changes in oxygen concentration is critical to the survival of many organisms. In *E. coli*, the Arc system controls the expression of many genes of aerobic metabolism in response to oxygen, while FNR acts as a transcriptional activator for anaerobic metabolism and as a repressor of some genes encoding proteins of aerobic function.

The Arc system is a two-component signal transduction system, although its direct signal has not yet been identified. The corresponding genes were designated *arcA* and *arcB* because they are involved in aerobic respiration control.

ArcB, a histidine protein kinase, is a membrane-bound sensor which monitors oxygen availability (Iuchi and Lin, 1992). In response to oxygen limitation, it is autophosphorylated by ATP or acetylphosphate. A transphosphorylation step converts the cytoplasmic regulatory protein, ArcA, into its active phosphorylated form, ArcA-P (Iuchi, 1993). The kinase activity of ArcB toward ArcA increases progressively during the transition from aerobic to anaerobic growth (Iuchi *et al.*, 1994).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ArcA can receive signals from two different sensor elements, ArcB and CpxA, and behaves in different ways depending on whether it is activated by one sensor or the other (Iuchi *et al.*, 1990; Weber and Silverman, 1988). The DNA binding consensus sequence for ArcA is [A/T]GTTAATTA[A/T] (Table 1) (Lynch and Lin, 1996). There is a putative ArcA binding site, TCTTAATTGC, in the upstream region of *hemA* at -83 bp to -92 bp. ArcA-P is an anaerobic repressor of the majority of target operons encoding enzymes involved in aerobic respiratory pathways, such as *sdhCDAB* (encoding the succinate dehydrogenase complex) and *lctD* (encoding the flavoprotein complex L-lactate dehydrogenase). ArcA-P can also act as an anaerobic transcriptional activator in a few cases for those operons that encode enzymes in microaerobic or fermentative metabolism. In aerobic cells, ArcA-P is also present at a significant level (Iuchi and Lin, 1988; Iuchi and Lin, 1993). Therefore, it may function as an activator aerobically.

The global regulator FNR is encoded by *fnr* (denoting the defects in fumarate and nitrate reduction). When oxygen is limiting, FNR activates synthesis of many enzymes required to generate energy in anaerobic respiration and also represses synthesis of some enzymes involved in aerobic respiration. FNR is a transcriptional activator for *frdABCD* which encodes fumarate reductase (Jones and Gunsalus, 1987), *dmsABC* which encodes dimethylsulfoxide/ trimethylamine N-oxide reductase (Cotter and Gunsalus, 1989), *narGHJI* which encodes nitrate reductase (Stewart, 1982), and *cydAB* which encodes cytochrome d oxidase (Cotter *et al.*, 1990; Cotter and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Operon (site)	DNA sequence (±10 bp)			Match to consensus
Consensus		T T AGTTAATTAA		
sodA	GGTTCATTAT	AGTTAATTAA	ATGATATTGA	10/10
	GGCGGCCGAT	TGTTAATGCC	GCGTAAGCAG	8/10
pfl (A1)	TATAATGCTT	TGTTAGTATC	TCGTCGCCGA	6/10
pfl (A2)	GCAAATACGG	AGTAAATATT	TGATTATCCA	7/10
pfl (A3)	CCGCTAAAAC	AGTTAATTAA	AAGGGAGCAT	10/10
	CGATCGATAT	TGTTACTTTA	TTCGCCTGAT	8/10
pfl (A4)	GATTAGTCTG	AGTTATATTA	CGGGGCGTTT	7/10
cyd (I)	ACCCTGGGTG	AGTTAATTAT	AATATAATTA	10/10
	TATAATTATA	AGTTAACTAA	ATGTTAATAT	9/10
cyd (II)	CTACAAATTA	TGTTAATCAT	GTTTCAATAA	9/10
	ATGTCACTAA	AGTTACCTTA	TTGAAACATG	7/10
cyd (III)	GTTGCAAACA	TGTTAATAAA	AACTCAAATT	8/10
sdh: P	CAGTATAGGC	TGTTCACAAA	ATCATTACAA	7/10
	ACTATATGTA	GGTTAATTGT	AATGATTTTG	8/10
glcA P_1	GTAACAACTT	TGTTGAATGA	TTGTCAAATT	7/10
	TCAACAAAGT	TGTTACAAAC	ATTACCAGGA	6/10
lct P_2	CATAACATTT	AGTTAACCAT	TCATTGTCAT	8/10
	GACAATGAAT	GGTTAACTAA	ATGTTATGTA	8/10
acnA	CAATTTGGGT	TGTTATCAAA	TCGTTACGCG	7/10
aceA	ACTTCCGATG	AGTTAATTGA	TTTCCTGACC	9/10
ald	TGATGATTGA	TGTTAATTAA	CAATGTATTC	10/10
	GTGAATACAT	TGTTAATTAA	CATCAATCAT	10/10
appY	GTTGGTTTTT	TGTTAATTGA	ATGGTTCTAA	9/10
	AGTGAGTTTA	TGTTAATAAA	AAGCATAGTA	9/10
arcA	ACGGTTGAGT	TGTTAAAAAA	TGCTACATAT	8/10
betI	TCATCAGCGG	TGTTTATCTA	TTAAAGCGGT	7/10
betT	GTGTCTTAAT	TGTTACGAAT	TTGATTTTAA	7/10
cyoA-E	GATAATTATT	TGTTAAATAA	TTGTTTTATT	9/10
dad	AAAAGCCGCA	TGTTGAATAA	TATTTTCAAC	8/10
fadB	GTGTATTTTG	TGTTAAAAAT	ATGCAAACAA	8/10
icd	AGCAATTTTT	TGTTAATGAT	TTGTAATTGG	9/10
nuoA-N	ATGCTAATGG	TGTTGATATT	ATGTAAACTA	8/10
mdh	AACATATCTT	AGTTTATCAA	TATAATAAGG	8/10
tsrY	GTTAATAAGG	TGTTAATAAA	ATATAGACTT	9/10

Table 1. The ArcA-P box, or primary sequence DNA binding site consensus for ArcA-P. The 10-bp consensus (5' [A/T]GTTAATTA[A/T] 3') was derived from sequence alignments of regions in template DNAs most strongly protected from DNase I cleavage by Arc-P (or His_6-ArcA-P) binding. For the larger sites of DNase I protection (i.e., all those shown in the top except the *pfl* A1, A2, and A4 sites), only the two best fits to the consensus sequence from within the protected regions are included. In the lower half are shown the best fits to the consensus sequence from the transcriptional regulatory regions of a number of other loci which are thought to be regulated by the Arc system or which are known to be regulated in response to oxygen (Lynch and Lin, 1996)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Gunsalus, 1992), *nirB* which encodes nitrite reductase (Jayaraman *et al.*, 1987), *narK* which encodes NarK, a nitrite export protein (Kolesnikow, 1992), and *fdnGHI* which encodes formate dehydrogenase (Berg and Stewart, 1990). It is also a transcriptional repressor for *cyoABCDE* which encodes cytochrome *o* oxidase in anaerobic respiration (Cotter and Gunsalus, 1992) (Figure 8). The expression of anaerobic genes can only be induced by the activated FNR, suggesting that FNR represents the master-switch which ensures that aerobic respiration is used in preference to anaerobic respiratory metabolism or fermentation (Guest, 1992).

FNR has a helix-turn-helix C-terminal, and it has a N-terminal which includes a four cysteine cluster Cys-X_3-Cys-X_2-Cys-X_5-Cys (Spiro and Guest, 1990). The N-terminal domain of FNR may be involved in the sensory process. This redox/O_2 -sensitive domain may sense the anaerobic state and converts FNR into an active conformation for DNA binding (Unden *et al.*, 1990; Engel *et al.*, 1991).

It is believed that FNR functions as a dimer. However, in most cases FNR has been isolated in the monomeric state (Figure 9). Lazazzerra et al. (1993) purified a FNR mutant protein called FNR-DA154 that was predominantly in a dimeric state when maintained at high protein concentrations (>80 uM). However, when the FNR-DA154 protein was diluted to 8 uM, it dissociated to the monomeric state.

Green and Guest (1993) proposed that the switch of FNR between inactive (aerobic) and active (anaerobic) forms could involve the reduction of a loosely-bound iron cofactor or the reversible binding of Fe^{2+}. Beker et al (1996) suggest that intracellular O_2 is responsible for FNR oxidation. The reaction could be mediated by cellular iron or other redox mediators.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

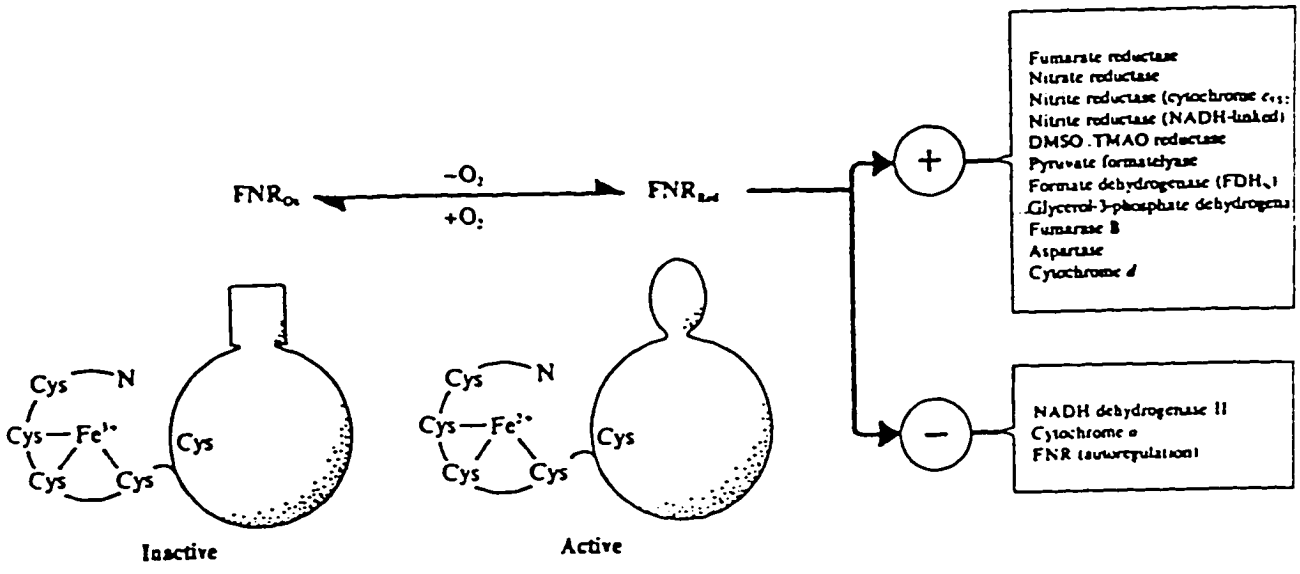

Figure 8. Mechanism for the anaerobic formation of active FNR, and target functions of FNR as an activator (+) or a repressor (-) (Spiro and Guest, 1991).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

FNR contains two [4Fe-4S] clusters per dimer. As cysteines are often the ligands for Fe-S clusters, it seems likely that the four cysteines shown to be essential for FNR function are the ligands for the Fe-S cluster in this protein. The instability of the Fe-S cluster to oxygen regulates the activity of FNR. Since the Fe-S cluster associated with FNR is unstable in the presence of oxygen, it is suggested that the major role of the cluster is in stabilizing an active conformation of FNR under anaerobic conditions (Lazazzera *et al.*, 1993). This conformation could be achieved by a simple reorganization of the N-terminal domain of FNR in the presence of the Fe-S cluster.

Most FNR dependent promoters have an FNR binding site with a consensus sequence: TTGATnnnnATCAA, and centered at position -42 relative to the start sites of transcription (Figure 10). *In vitro* DNase I protection studies were performed to establish the locations of the FNR binding sites at the *narG, narK, dmsA* and *hemA* promoters (Melville and. Gunsalus, 1996). The FNR binding site in *hemA* is the half site, TTGAT at -18 bp to -22 bp.

There are two hypotheses for FNR binding to DNA in response to anaerobiosis (Melville and Gunsalus, 1996). One model proposes that when oxygen is low, the FNR protein binds to the DNA recognition site as a monomer, and upon the subsequent binding of a second FNR monomer at a neighboring DNA half-site, assembles into a stable FNR dimer. In the presence of oxygen, the binding of a FNR monomer at either the first or second half-sites would be substantially reduced due to an unstable conformational state of the protein. The other model suggests that anaerobiosis leads to dimerization of free FNR in solution. The dimeric species then binds to the FNR recognition site at the FNR-dependent promoters.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

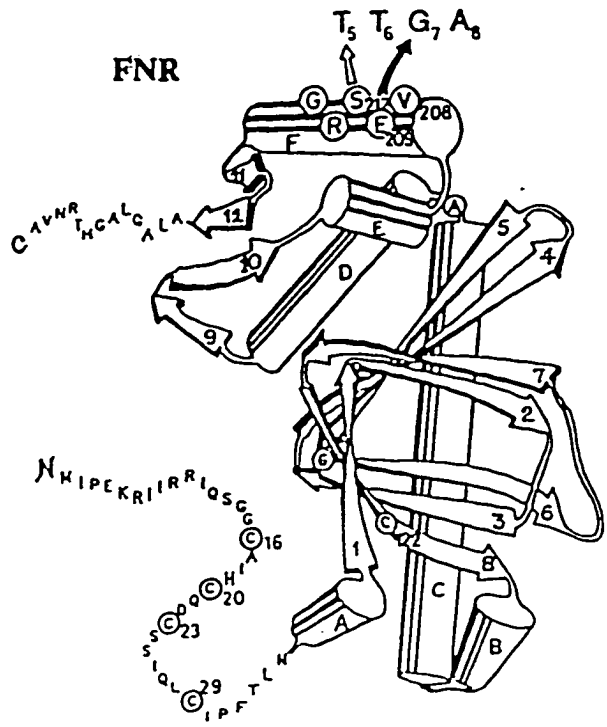

Figure 9. The structure of the FNR monomer (Guest, 1992).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

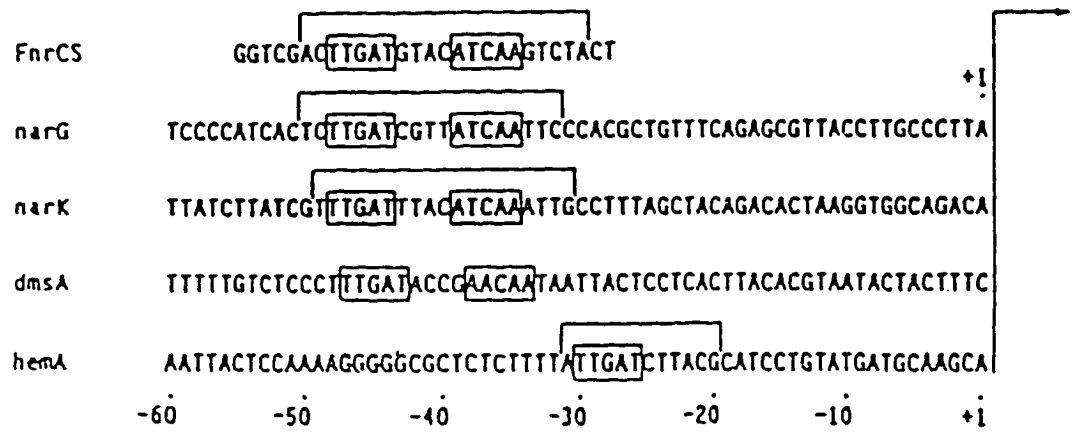

Figure 10. Location of FNR152 recognition sites on the pFNRCS consensus site plasmid at the FNR-activated *narK*, *narG*, and *dmsA* promoters and at the *hemA* promoter that is repressed by FNR. The inverted brackets represent the DNase I-protected regions for the respective promoter elements. Numbering is relative to the start site of transcription (+1). The arrow represents the 5' end of mRNA (Melville and Gunsalus, 1996)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

7. Nitrate Effect and the Role of NarL in Gene Expression

Under anaerobic conditions when nitrate is present, expression of *narGHJI* increases 10-fold, whereas *frdABCD* and *dmsABC* of *E. coli* decrease more than 10-fold (Jones and Gunsalus, 1987; Cotter and Gunsalus, 1989). When fumarate is present, only *frdABCD* expression increases 1.5-fold; other operons exhibit no change in expression. FNR and NarL are involved in the regulation of these genes (Stewart, 1982; Iuchi and Lin, 1987; Kalman and Gunsalus, 1988; Kalman and Gunsalus, 1989). There are two two-component regulatory systems for sensing nitrite and nitrate: NarX:NarL and NarQ:NarP (Kalman and Gunsalus, 1989; Kalman and Gunsalus, 1990; Chiang *et al.*, 1992; Stock *et al.*, 1989). The sensor proteins NarX and NarQ are membrane-bound proteins for detecting nitrate availability; the other members of the system are NarL and NarP which are the response regulators that, once activated, will bind to DNA. NarQ has high amino acid homology to NarX. There is evidence to indicate that like many sensor proteins, NarX and NarQ may be histidine protein kinases (Cavicchioli et al., 1995). Phosphorylated NarL is thought to be the transcriptionally active form (Figure 11). Although NarL and NarP may be mutually redundant as far as gene activation is concerned, only NarL can act as a repressor. NarP can also activate transcription of the *nir* and *fdn* promoters in response to nitrate or nitrite, but it is far less effective than NarL. A minimal consensus for NarL binding is 5'-**TACY*KT-3' (Y=pyrimidine; K=A or C) (Tyson *et al.*, 1993). The position and sequence of NarL binding sites differs from one promoter to another (Figure 12). Some NarL binding sites are inverted repeat sequences but others are direct repeat. There is a putative NarL binding site,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

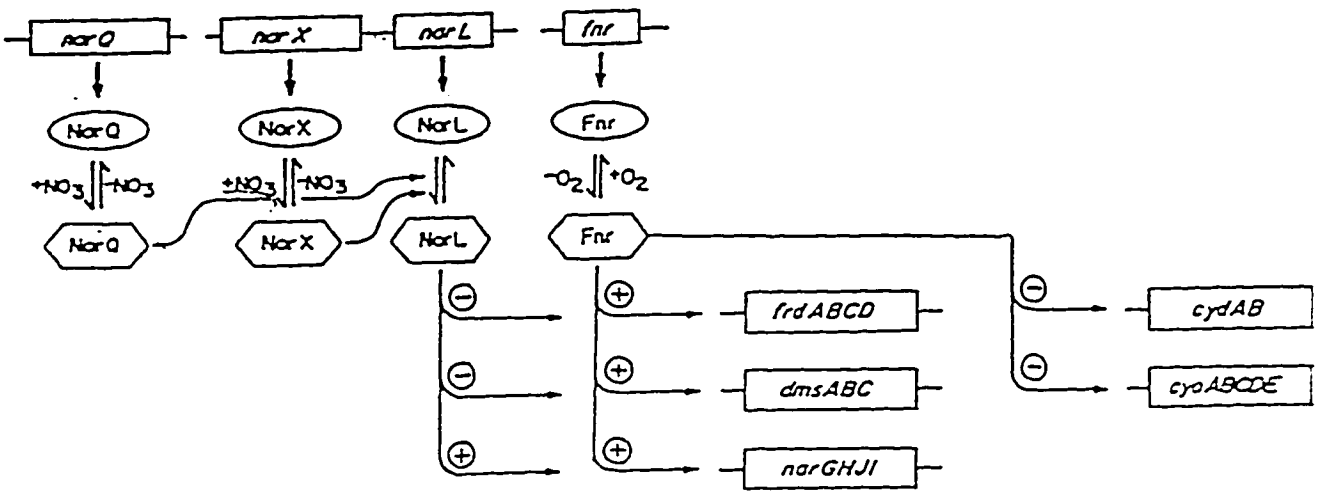

Figure 11. NarL, NarP, NarQ, and NarX dependent control for response to nitrate and molybdate availability. The global regulators: FNR and NarL provide the indicated negative and positive transcriptional control to coordinate enzyme synthesis (Gunsalus, 1992).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TACTGAT, at -324 bp to -330 bp in the 5' upstream region of *hemA*. NarL appears to play different roles in or on the *nirB* and *narGHJI* promoters. At the *narGHJI* promoter, NarL binds about 200 bp upstream from the transcription start site and requires IHF, integration host factor. IHF is essential for promoting DNA bending to enhance transcription (Tyson *et al.*, 1993; Rabin *et al.*, 1992) (Figure 13). These proteins may form a certain complex to facilitate transcription initiation. There is evidence that NarL can directly contact RNA polymerase (Schroder *et al.*, 1993). At the *nirB* promoter, NarL binds adjacent to FNR; FNR-independent expression is minimal. This can be explained if FNR interacts directly with NarL.

8. The Glucose Effect and the Role of CRP in Gene Regulation

E. coli uses different energy-generating modes in response to available carbon sources. When *E. coli* cells are grown with glucose which is the most efficient carbon source, glucose may cause lowering of the cAMP level. Therefore glucose can cause down- or up-regulation of certain genes mediated by cAMP-CRP levels. One role of the cAMP-CRP complex is to ensure that glucose is used in preference to lactose and other carbohydrates. This phenomenon is called catabolite repression. When *E. coli* cells are grown on non-glucose carbon sources, the cAMP concentration inside the cells is very high, cAMP binds to CRP causing a conformational change in CRP, and the cAMP-CRP complex activates genes by binding at specific DNA sequences. Some glucose effects may not be cAMP-CRP mediated. Cra (catabolite repressor/activator) protein which represses sugar catabolic systems and activates sugar anabolic systems in *E. coli* is involved in cyclic AMP-independent catabolite repression (Saier, 1996).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

nirB	1	-79	TATACCCATT	-70
nirB	2	-60	TATACTCCTT	-69
narC	1	-207	AATACTCCTT	-198
narC	2	-197	AATACCCATC	-188
fdnC	1	-114	TCTACCGCTA	-105
fdnC	2	-95	CCTACCTCAA	-104
frdA	1	-9	TGTACCTATA	+1
frdA	2	+11	ACTGCTCCTT	+2
dmsA	1	-39	ATTACTCCTC	-30
dmsA	2	-19	AATACTACTT	-10
dmsA	3	+3	TCTACCTATC	+12
narK	1	-219	TATACCCATA	-210
narK	2	-199	GTTACTCCTT	-190
			--TACY-KT-	

Figure 12. Potential NarL binding sites in different NarL sensitive promoters. 5'-**TACY*KT-3' (Y=pyromidine; K=A or C). The positions are activation of *narGHJI* with respect to the transcription start points (Tyson *et al.*, 1993).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

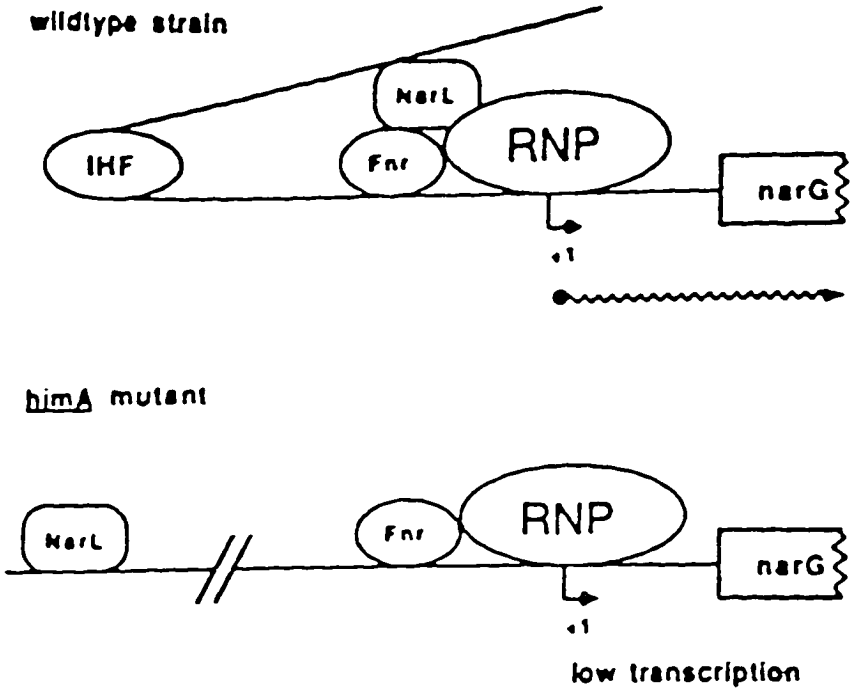

Figure 13. Proposed role of IHF in NarL activation of *narGHJI* operon expression (Shroder *et al.*, 1993)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The cAMP-CRP complex can be a transcriptional activator or a repressor (Rasmussen *et al.*, 1996). All cAMP effects in bacteria appear to be mediated through its binding to CRP except for DnaA (Hughes *et al.*, 1988). cAMP-CRP enhances *lac* operon expression under aerobic conditions. As an activator, cAMP-CRP complex binds to DNA sequences located upstream from what are generally classified as weak promoters and modifies those promoters in a way that enhances promoter recognition by RNA polymerase. The mechanism by which the cAMP-CRP complex activates a given CRP-dependent promoter is only partially understood. Most evidence suggests that CRP contacts RNA polymerase directly. The most likely model for cAMP-CRP activation of transcription is DNA bending (Adhya and Gatges, 1990). The only effect cAMP-CRP is known to have on DNA binding is inducing a bend from linearity (Figure 14).

The sensor for the cAMP level is adenylate cyclase. The nature of the carbon source is sensed by this enzyme probably through the phosphorylation and dephosphorylation of the sugar phosphotransferase system (Roseman and Meadow, 1990).

Inspection of 26 known CRP-binding sites has revealed a palindromic consensus sequence: ---TGTGAnnnnnnTCACA--- (Ebright *et al.*, 1984). Although the consensus sequence is limited to 22 bp, distal DNA sequences also play a role in determining the affinity of CRP for a specific DNA sequence. There is a putative cAMP-CRP binding site with the sequence, TGAGAnnnnnnGCAGT at -144 bp to -159 bp of the upstream region of *hemA*.

It appears that the affinity of the cAMP-CRP complex for different CRP-binding sites is controlled by the DNA sequence. Therefore variations in sequence among CRP-binding sites

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

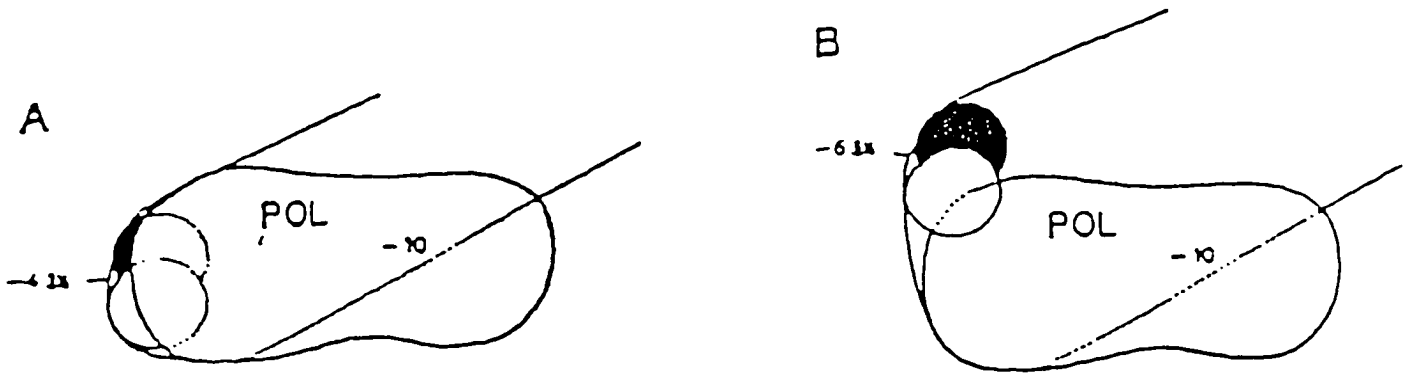

Figure 14. Interaction of CRP and RNA polymerase. A: the CRP site is centered at -41. 5. B: the CRP site is centered at 61. 5 from the transcription start point (Williams *et al.*, 1991).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

provide one method to control CRP activated promoter strength. Constructs which have CRP binding sites at near-integral turns of the helix (i.e., n x 10.5 bp) from the -10 sequence of a promoter are activated by cAMP-CRP, while constructs with binding sites at a half-integral turn of the helix cannot be activated by cAMP-CRP (Gaston *et al.*, 1990). The conclusion drawn from these studies is that CRP can activate transcription when bound to DNA on the same face of the helix within a range of helical turns from the RNA polymerase-binding site. In addition, the data showed that the extent of CRP-dependent promoter activation decreases with increasing CRP-binding-site distance from the -10 sequence of a promoter. Therefore the position relative to the -10 sequence of a promoter also appears to be an important determinant of CRP-dependent promoter strength.

The binding sites of CRP are located at different positions in the 5' upstream regulatory region from the transcription initiation site in different genes. For example, in the *lac* operon, cAMP-CRP binds to a site centered at 61.5 bp upstream from the transcription start point; in the *gal* operon it is centered at 41.5 bp. In several CRP activated promoters, the position of the CRP-binding motif relative to the -10 sequence is very similar (Valentin-Hansen *et al.*, 1991) (Figure 15).

CRP is a 47,238-Da protein. The CRP protein exists as a dimer which is composed of two identical subunits. Each has 209 amino acid residues. For each subunit the N-terminal is responsible for cAMP binding and dimer contacts. The C-terminal contains a helix-turn-helix motif involved in DNA binding (Botsford and Harman, 1992). CRP monomers (Figure 16) contain two potential activating domains which are located on opposite faces. However, when

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CRP binds as a dimer to the inverted repeat consensus sequence, the two potential activating domains are displayed on adjacent faces of different subunits (Williams *et al.*, 1991).

9. The Goal of This Study

In order to systematically study the regulatory elements of the 5' upstream region of *hemA* and their response to different conditions, a series of fragments of this region were cloned in front of a promoterless gene for chloramphenicol acetyltransferase (CAT) in the vector, pKK232-8. The role of the global regulators, ArcA, NarL, FNR, CRP, in *hemA* regulation was investigated by using isogenic strains for the regulator. The effects of oxygen, nitrate, glucose, medium, ALA and heme starvation were studied.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

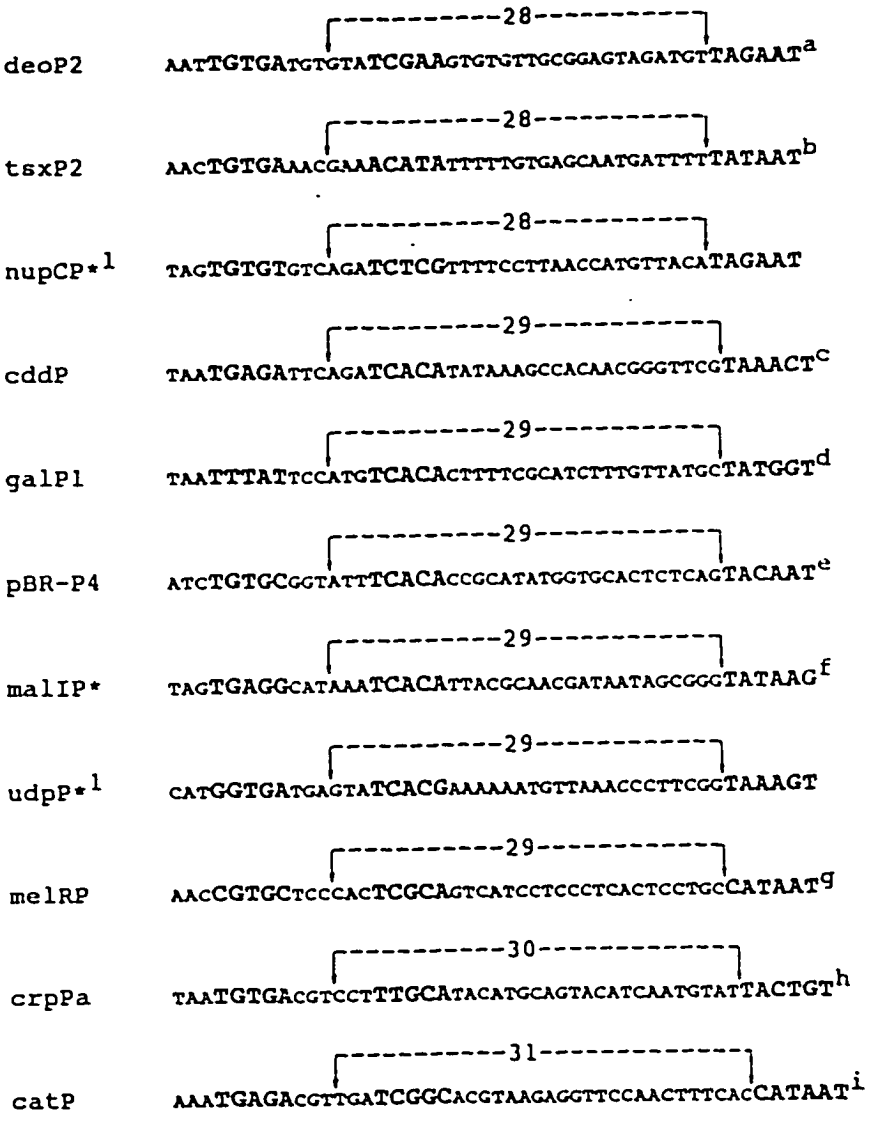

Figure 15. Sequences of CRP binding sites and -10 sequences and the distance between them are shown for 11 promoters that are activated by cAMP-CRP (Valentin-Hansen, 1991).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

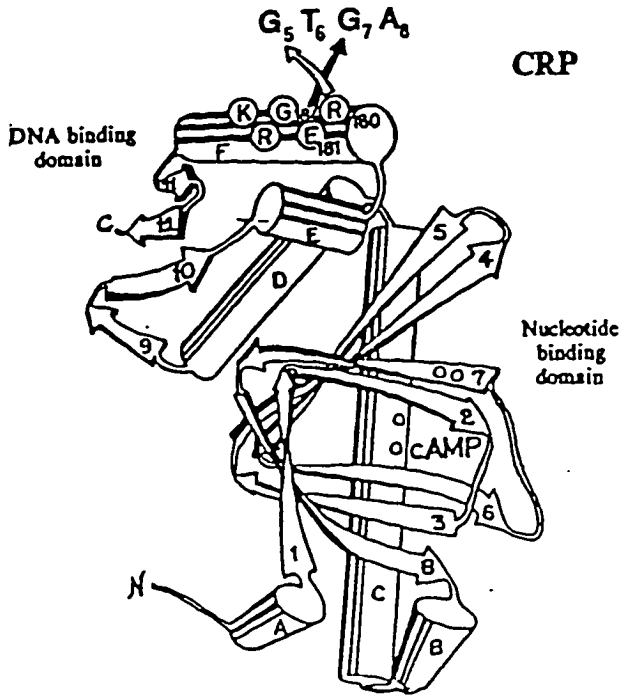

Figure 16. The structure of the CRP monomer (Guest, 1992).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MATERIALS AND METHODS

1. Chemicals and Enzymes

Restriction enzymes were purchased from New England Biolabs (Beverly, MA). Media were prepared from Difco products (Detroit, MI). ALA, hemin, molecular biology grade agarose, ampicillin, kanamycin, Tween 80, and chloramphenicol were purchased from Sigma (St. Louis, MO). Phenol and chloroform were purchased from Fisher Scientific (Pittsburgh, PA). X-gal was purchased from Jersey Labs (Livingston, NJ).

The TA Cloning Kit was purchased from Invitrogen (San Diego, CA). The DIG DNA Labeling and Detection kit, CAT ELISA kit and Quick Spin™ columns were purchased from Boehringer Mannheim (Indianapolis, IN). The GELase™ agarose gel-digesting preparation was purchased from Epicentre Technologies (Madison, WI). The Wizard™ Maxipreps DNA purification system was purchased from Promega (Madison, WI)

The PCR primers used in this study are listed in Table 2. Primers A1, A2 and A5 were purchased from Oligos Etc. (Wilsonville, OR). Primer A4 was purchased from the Hunter College RCMI Sequencing Facility (New York, NY). Primers A6 and A7 were purchased from Keystone Labs (Menlo Park, CA).

2. Strains

Strains used in this study are listed in Table 3.

Plasmids used in this study are listed in Table 4.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2. Primers used in this study

Primers		Sequence and Restriction Enzyme Sites	
A1	5'	TAACGCTGACCTTTATTGTCGACT	3'
		*Sal*I	
A2	5'	AAATAATACCAAGCTTGATAGGGTT	3'
		*Hin*dIII	
A4	5'	AGCCAGCGGTAGCAGGCGGATAAGACGAAA	3'
A5	5'	GCAAGCTTGGCAATTAAGACAACTTGA	3'
		*Hin*dIII	
A6	5'	GCGGATCCCAGAATCTAACGGCTTTCG	3'
		*Bam*HI	
A7	5'	CGGATCCGCGAGTCTGATAGTGATTAAGATTG	3'
		*Bam*HI	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3. Strains used in this study

E. coli Strains	Relevant Genotype and/or Phenotype[a]	Source
HU227	*hemA*, hemin-permeable	Umanoff *et al.* (1988)
JL1268	HU227/ pJL68	Li *et al.* (1989)
RP7947	*pcnB1 zad-981*::mini-*kan*	Liu *et al.*(1989)
RK4353	wild type	Stewart (1982)
WH105	RK4353 but *pcnB1*	This study
VJS1741	RK4353 but *fnr*::Tn*10*	Stewart (1982)
WH106	VJS1741 but *pcnB1*	This study
RK5278	RK4353 but *narL215*::Tn*10*	Stewart (1982)
WH107	RK5278 but *pcnB1*	This study
INVα F'	*endA1 recA1 hsdR17(r-k, m+k) supE44 thi-1 gyrA relA o80 lacZ M15 (lacZYA-argF) deoR*$^+$ F'	Invitrogen

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3. Strains used in this study (continued)

E. coli Strains	Relevant Genotype and/or Phenotype	Source
1100	*relA1 thi-1 bglR11*	CGSC[b]
SP101	1100 but *pcnB1*	This study
PP47(5333)	1100 but *crp-1*	CGSC
SP102	PP47 but *pcnB1*	This study
SP312	Δ *(galK-bioD)76 relA1 spoT1 thi-1 deoC76*	
	Δ *(deoD-trpR)251*	CGSC
SP103	SP312 but *pcnB1*	This study
SP314	SP312 but Δ *(deoD-arcA)253*	CGSC
SP104	SP314 but *pcnB1*	This study

a. Other genetic markers are: HU227: *mel1*, Hfr Cavalli. Strain HU227 was derived from SASX41B, the Sasarman *hemA* mutant. RK4353: *araD139, argF-lac, U169, deoC1, flhD5301, gyrA219, non-9, ptsF25, relA1, rpsL150.*
b. CGSC, Coli Genetic Stock Center, Yale University, New Haven, CT.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 4. Plasmids used in this study

Plasmids	Relevant Genotype and/or Phenotype	Source
pJL68	pTZ19U (*hemA hemM*)	Li *et al.* (1989)
pKK232-8	pBR322 derivative; a promoterless *cat;* translation terminators in all three frames upstream from *cat;* Amp^r	Brosius (1984)
pCR™II	p*lacZα* ColE1 ori; f1 ori	Invitrogen
pWH515	pKK232-8 with 515 bp (-496 to +19) of *hemA*	This study
pWH359	pKK232-8 with 359 bp (-340 to +19) of *hemA*	This study
pWH239	pKK232-8 with 239 bp (-220 to +19) of *hemA*	This study
pWH239*	pKK232-8 with 239 bp (-220 to +19) of *hemA;* 3 base mismatches at -13, -52 and -84 bp	This study
pWH101	pKK232-8 with 101 bp (-82 to +19) of *hemA*	This study
pWH415	pKK232-8 with 415 bp (-496 to -81) of *hemA*	This study

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3. Media and Growth Conditions

LB, YT and M9 minimal media were prepared according to Sambrook et al (1989).

For aerobic growth, liquid cultures were grown in a shaking waterbath at 37°C. Solid medium cultures were grown in the incubator at 37°C.

For anaerobic growth, liquid cultures were grown in 15 ml Klett tubes fitted with a tight rubber cap. A needle was placed through the cap to the middle of the tube and another needle placed through the cap for an outlet. In this manner prepurified nitrogen was bubbled through the sample for 5 min. The cultures were then put into an anaerobic jar with a gas pack CO_2 generator (Becton Dickinson Microbiology Systems). The cultures were grown overnight.

M9 or LB was supplemented with varying combinations of the following: glucose (0. 2%), glycerol (1%), $NaNO_3$ (40mM), fumarate (50mM), and hemin (4ug/ml) as noted. The hemin solution (4 mg/ml) was made by mixing 0.1 g of hemin and 2.5 ml of Tween 80 in 22.5 ml of sterile water with 4 drops of 10 N NaOH. Ampicillin was used in each culture at 50 ug/ml when there was a pKK232-8-derived plasmid.

Cell cultures for ALA or heme starvation experiments were grown in 100 ml LB containing KPO_4 (pH 7.2) and 20 mM sodium pyruvate. After 15 hr at 37°C, the culture was diluted 1:1000 to100 ml LB containing KPO_4 and sodium pyruvate for another 15 hr growth. The cell cultures for the ALA or heme experiments (non-starvation) were grown first under the same conditions as the starvation experiments for 15 hr. Then the cultures were diluted (1:10) into 20 ml LB with ALA at 50 ug/ml or hemin at 4 ug/ml, and grown 3 hr to log phase.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4. Methods

4.1. PCR

PCR was performed in a Perkin Elmer Cetus DNA Thermal Cycler (Model 9810) by following the instructions in the TA Cloning Kit. 25 cycles of amplification were performed using the following profile: 1 min at 94°C, 2 min at 55°C, and 3 min at 72°C. An additional 7 min at 72°C was followed. After the 25 cycles, the tubes were cooled to 4°C.

4.2. TA CLONING

The TA Cloning Kit (Invitrogen) is designed to clone PCR products directly from a PCR reaction without the need for modifying enzymes, purification or restriction digestion. This system takes advantage of the non-template dependent activity of thermostable polymerase used in PCR that adds single deoxyadenosines to the 3'-end of all duplex molecules produced by PCR. Among the characteristics of the TA Cloning vector, pCR™II, are the *lacZ* gene for blue-white color selection, ampicillin and kanamycin resistance genes, and a versatile polylinker. All the procedures in the TA cloning kit for ligation and transformation were followed.

Ten possible transformants (white colonies) were picked. A mini prep DNA isolation (Sambrook *et al.*, 1989) from each of the transformants was used for DNA dot blot detection.

4.3. DNA DOT BLOT

Appropriate probes were prepared by denaturing and labeling PCR products with the DIG DNA Labeling and Detection Kit (Boehringer Mannheim). Unincorporated nucleotides were removed from the labeling reactions by Quick Spin™ columns (Boehringer Mannheim).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2 ul of each plasmid DNA from small scale isolations was spotted on a nylon membrane (Boehringer Mannheim) with appropriate PCR product as the positive control DNA, and pKK232-8 as the negative control DNA. The membrane then was placed on a piece of blotting paper and fixed by UV crosslinking for 3 min in the UV Stratalinker 2400 (Stratagene).

Prehybridization, hybridization, and immunological detection were carried out according to the instructions of the supplier (Boehringer Mannheim).

4.4. SUBCLONING OF THE 5' UPSTREAM REGION OF *hemA* INTO PKK232-8

A strategy was followed to subclone each PCR product into pKK232-8. For example, a 515 bp DNA fragment had been produced by PCR using the primers A1, A2. To facilitate the proper cutting of this fragment by restriction enzymes, this fragment was cloned into the pCR™ II vector. The new plasmid, pCRA, and pKK232-8 were cut by *Sal*I and *Hin*dIII, and ligated to form pWH515. The transformants were selected on 50 ug/ml chloramphenicol LB plates. The cloning strategy is presented in Figure 17.

The plasmid, pJL68, which contains the whole *hemA* gene (including the 5' upstream region) was linearized by *Hin*dIII enzyme digestion at 37°C for 1 hr, to serve as a template for PCR, using primers A1 and A2. The PCR product was 515 bp long and had a *Sal*I site at one end and a *Hin*dIII site at the other end. The PCR product was cloned into the TA cloning vector, pCR™II, to produce pCRA. Colonies which gave positive results on a DNA dot blot with the 515 bp insert as a probe were used for subsequent cloning. Both pKK232-8 and pCRA were cut with *Sal*I and *Hin*dIII, then ligated to form pWH515, thus cloning the insert into pKK232-8 in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

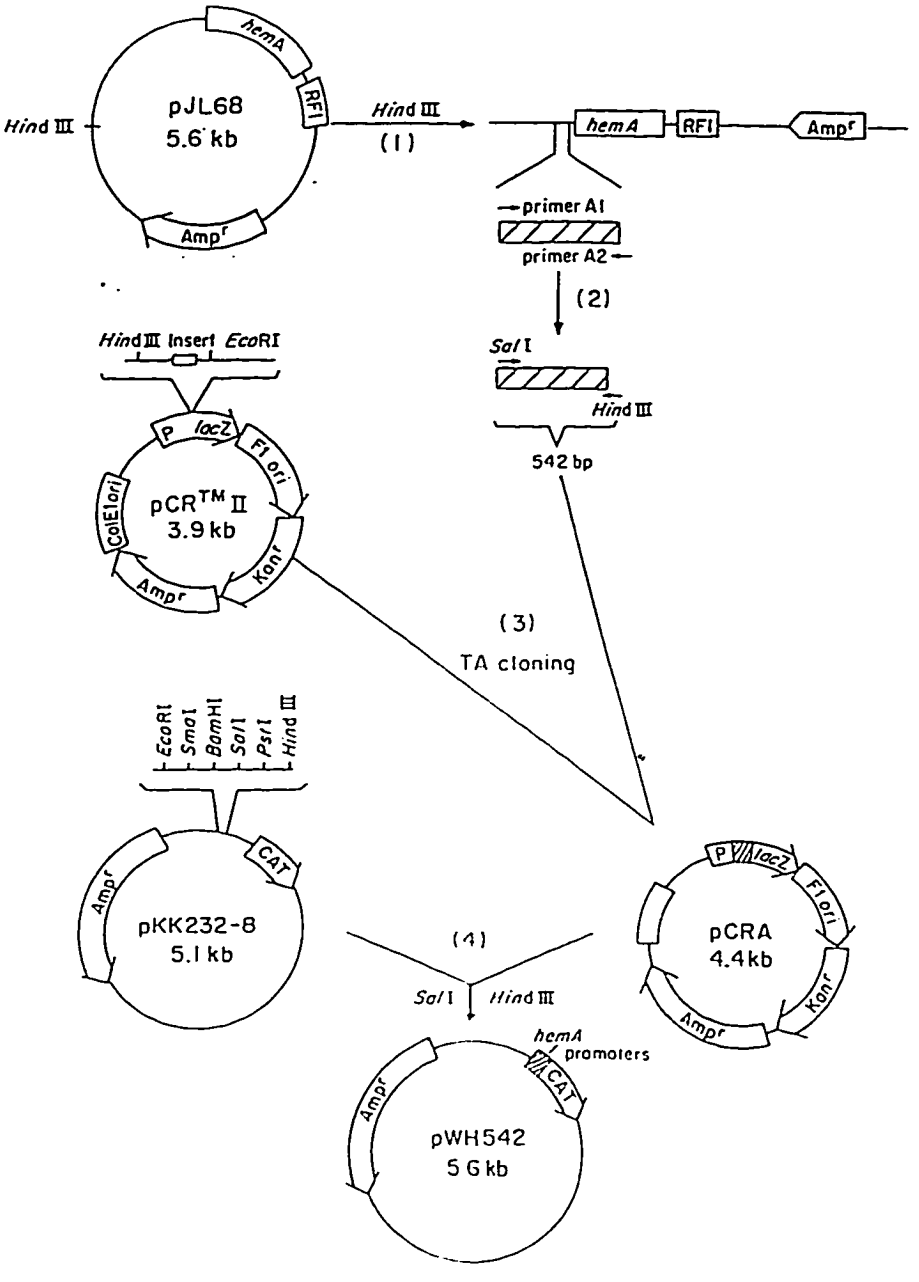

Figure 17. Strategy for subcloning the 5' upstream region of *hemA*.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

front of the reporter gene, *cat*. Transformants selected on LB plates containing 50 ug/ml ampicillin were tested for growth on 10, 25 and 50 ug/ml chloramphenicol. Transformants which could grow on 50 ug/ml of chloramphenicol LB plates were selected for DNA sequencing.

The main differences in cloning among the different fragments are the cloning sites. pWH359 and pWH101 were created by using *Bam*HI and *Hin*dIII with vector and insert. pWH415 was created by using *Sal*I and *Hin*dIII with vector and insert. pWH239 and pWH239* were created by using sticky-blunt end ligation: the inserts had one *Hin*dIII cutting end and one blunt end; the vectors were cut by *Sma*I (*Sma*I created one blunt end) and *Sal*I.

(a) Isolation of DNA from TA cloning plasmids

Plasmid DNA from TA cloning was purified by using Promega Maxipreps DNA Purification System. The DNAs were used for cloning and sequencing.

(b) Restriction Enzyme Digestion

Both vector pKK232-8 and plasmid from TA cloning were cut by appropriate restriction enzymes. A typical restriction digestion protocol is as follows: In an Eppendorf tube the following were mixed. A 20 ul reaction mixture contained 2 ul of 10X buffer (usually supplied with enzyme), 1-4 units of enzyme, about 1 ug of DNA and H_2O to make up a total of 20 ul. The mixture was incubated at 37°C for 2-12 hr for most digestions.

The sizes of restriction fragments and the quality of the digestion were determined by agarose gel electrophoresis. The concentration of agarose (usually 0. 7-1%) used depended on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the size of the fragment to be separated. 10 cm long minigels were run in 0. 5X TBE at 100 V for 2 hr. *Hin*dIII-*Eco*R I digested λ DNA fragments were used as the standards.

The insert was cut out of the gel. Gelase™ agarose gel-digesting preparation (Epicentre Technologies) was used for simple and quantitative recovery of intact DNA from low-melting-point agarose gels.

(c) Ligation

The gelase-purified DNA fragment and pKK232-8 were ligated at a ratio of vector: insert of 1:3. 1 unit T_4 DNA ligase was used for 1 ug of total DNA. The reaction was incubated overnight at 16°C.

(d) Transformation

Competent cells were prepared using the procedure in Advanced Bacterial Genetics (Davis *et al.*, 1982) and was modified by using 2 ul 0. 5M β-mercaptoethanol per tube at the step before putting in DNA samples. The transformants were selected on LB agar plates supplemented with 50 ug/ml ampicillin. Transformants were then streaked onto LB media supplemented with 50 ug/ml ampicillin and 5, 10, 25, or 50 ug/ml chloramphenicol to screen for colonies with resistance to both antibiotics.

(e) DNA sequencing

The insert of pWH515 was sequenced by using the Sequenase Version 2. 0 sequencing kit (U. S. Biochemical Corp). The inserts of pWH359, pWH239, pWH239*, pWH415 and pWH101 were sequenced by Dr. Gary Pestano using an ABI Sequencer (Model 373A) and Prism Kit with

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Taq DNA polymerase and dye terminators. This was carried out in the automated DNA sequencing facility at CCNY.

4.5. CONSTRUCTION OF *pcnB*⁻ STRAINS BY P1 TRANSDUCTION

A mutation at the *E. coli pcnB* locus reduces the copy number of ColE-like plasmids to two to three copies per cell (Liu and Parkinson, 1989). Transduction was used to introduce the *pcnB* mutation to eliminate gene dosage effects.

Strains RK4353, VJS1741, RK5278; SP312, SP314; 1100, and PP47 were grown in YT medium at 37°C overnight. The cells were subcultured (1:50) into fresh YT (5 ml) and were grown to an optical density of 40-50 Klett (about 5 hr). The cultures were centrifuged at 5000 rpm for 10 min at 4°C. 1 ml of 0.01M $CaCl_2$ and 0.01M $MgSO_4$ were added to resuspend the pellet. P1 phage which had been grown on RP7947 was diluted 10 and 100 fold. For each dilution, 100 ul cells and 100 ul P1 were mixed. The tubes were incubated for 20 min at 37°C. 100 ul from each tube was plated on a LB plate with 50 ug/ml kanamycin to select for *pcnB*⁻. The plates were incubated at 30°C for 2 days. These new strains were designated WH105 (RK4353, *fnr*⁺, *narL*⁺), WH106 (VJS1741, *fnr*⁻), WH107 (RK5278, *narL*⁻), SP101 (1100, *crp*⁺), SP102 (PP47, *crp*⁻), SP103 (SP312, *arcA*⁺), SP104 (SP314, *arcA*⁻). Plasmid was transformed into these new strains. The transformants were picked and tested for the respective plasmid by comparing the Maxi-preparations of DNA from the putative *pcnB*⁻ strains with their respective parental strains on agarose gels. The *pcnB*⁻ strains produced less plasmid DNA than their respective parental strains.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Each of the cloned plamids (pWH515, pWH359, pWH239, pWH239*, pWH415, pWH101) was transformed into WH105, WH106, WH107, SP101, SP102, SP103, SP104 respectively.

4.6. CAT ELISA ASSAY

The *E. coli* strains to be tested for CAT expression were grown under different conditions, and then sonicates of the pellets were prepared. The cells were washed with PBS twice, and the pellets were resuspended in 2 ml PBS buffer and sonicated for four 15 sec cycles at 50% output, 80% duty cycle in an Ultrasonic Sonicator #W-37J (Heat Systems). Cell extracts were used to determine protein concentration.

(a) Protein determination

Protein concentration was determined with the Bio-Rad Protein Assay (Bio-Rad, Richmond, CA). BSA was used as protein standard. Five dilutions were used: 0 mg/ml, 0.07 mg/ml, 0.14 mg/ml, 0.28 mg/ml, 0.56 mg/ml. Samples were diluted 3:4, 1:3, 1:2, and 10 µl for each standard and samples was transferred into separate microtiter plate wells. 200 µl of dye solution was added to each well. The sample and reagent were mixed by multi-channel pipet. The microtiter plate was incubated at room temperature for at least 5 min. The absorbance was measured at 595 nm in a Microplate Reader (Bio-Rad Model 450).

(b) CAT ELISA assay

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The instructions for the CAT ELISA Kit (Boehringer Mannheim) were followed. The amounts of protein used in the CAT ELISA assay were 0.1 ug, 0.05 ug, and 0.01 ug per well. Absorbances were read with the Microplate Reader.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

RESULTS

1. Construction of Plasmids with a Reporter Gene Driven By Various Fragments of the Upstream Region of *hemA*

1.1. AMPLIFICATION OF VARIOUS FRAGMENTS OF THE 5' UPSTREAM REGION OF *hemA* BY PCR

In order to study promoter strength and regulation, five fragments containing different regions of the *hemA* 5' upstream region were created by PCR. Figure 18 shows the structures of each PCR product and its primers. A 101 bp fragment contains 82 bp upstream from the initiation of transcription site including the P1 region and the putative FNR binding site (pWH101). A 239 bp fragment contains 220 bp upstream from the initiation of transcription site including P1 and P2 and the putative FNR, Arc and cAMP-CRP binding sites (pWH239). DNA sequencing showed that a second 239 bp fragment created by PCR had one base pair substitution in the -10 sequence of P1 and two other base pair substitutions at -52 bp and -84 bp (Figure 19). A 415 bp fragment contains 415 bp of the upstream region including P2 and P3, the putative NarL, Arc and cAMP-CRP sites and two stem loops (pWH415). A 359 bp fragment, containing 340 bp upstream from the initiation of transcription site including P1 and P2, putative NarL, cAMP-CRP, Arc binding sites and two stem loops (pWH359), was designed to have one base pair substitution in the putative NarL binding site (TA**G**TGATT instead of the original TA**C**TGATT). A 515 bp fragment contains all the features described above. The positions of each primer in the 5' upstream of *hemA* is shown in Figure 20.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

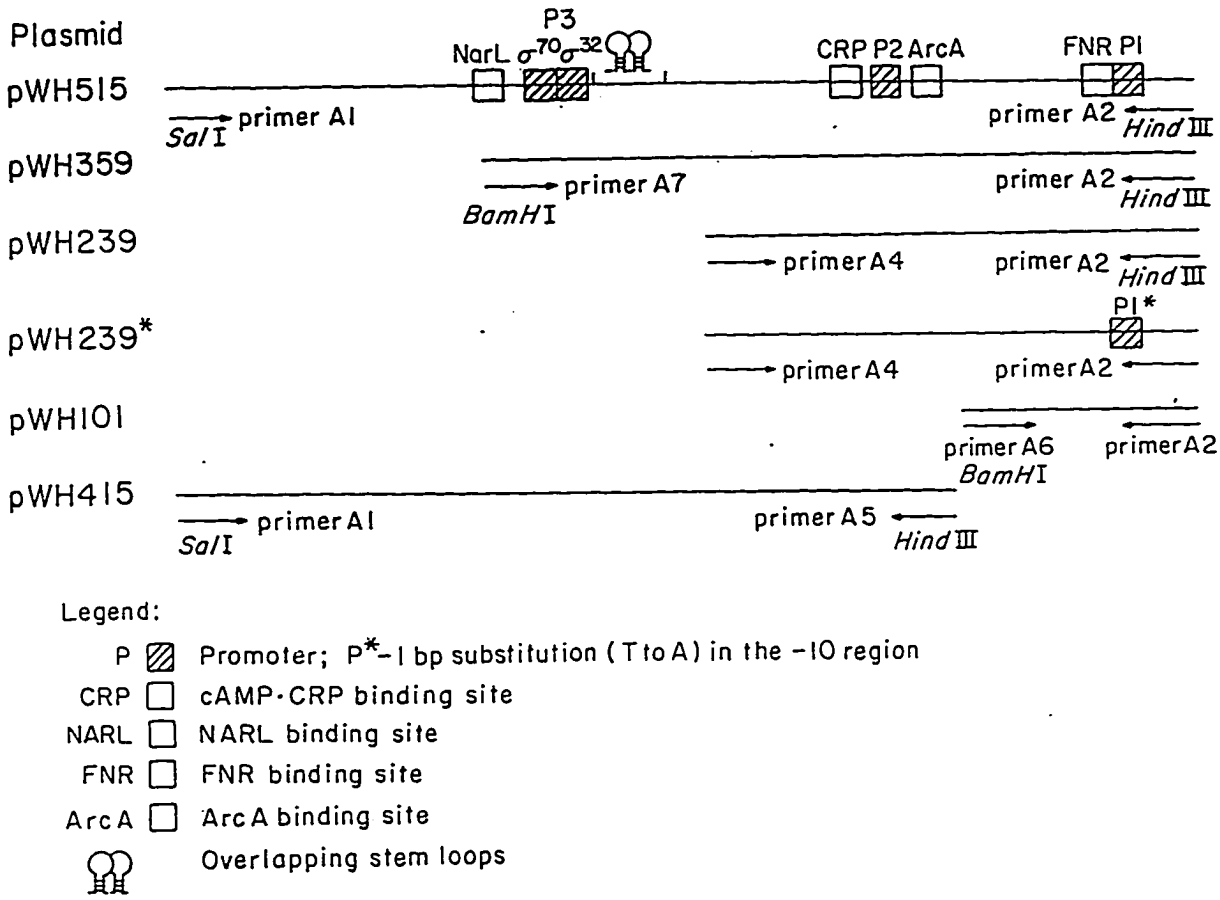

Figure 18. Various fragments of the 5'upstream region of *hemA*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

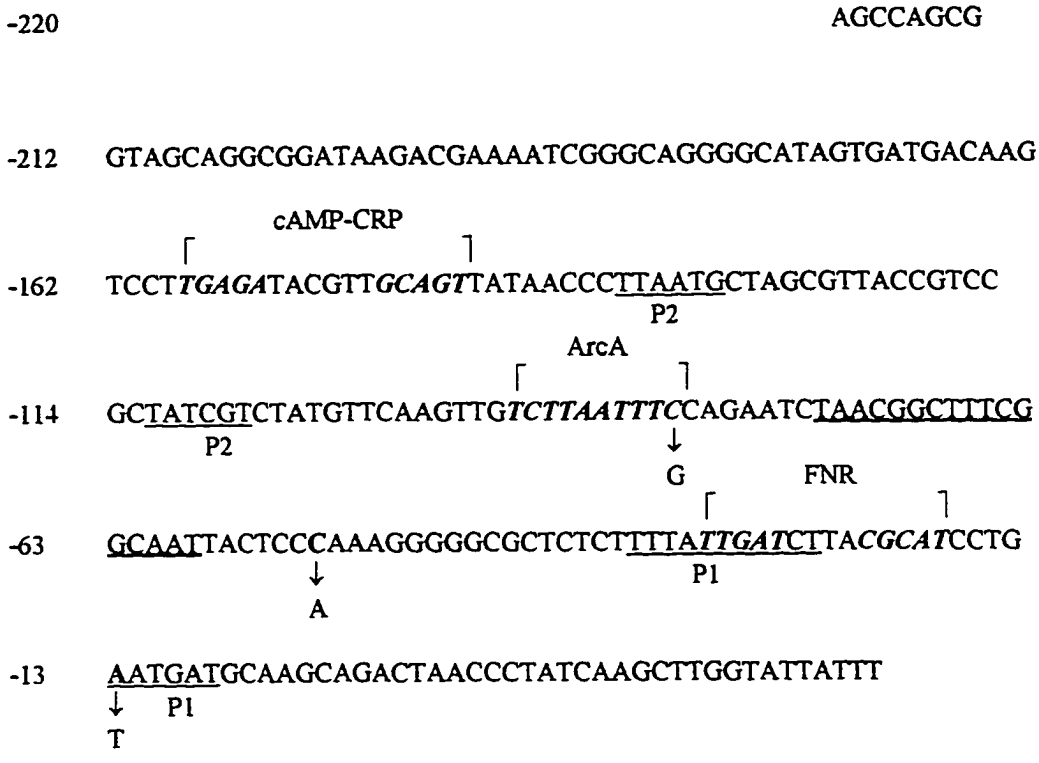

Figure 19. Sequence of the insert in pWH239*. There is one base pair substitution (T to A at P1 (pointed by arrow), and two other one base pair substitutions, at -52 bp and -84 bp.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-513 **TAACGCTGACCTTTATTGTCGACT**AACTGCACGTTACGCGGTTGAGCATT

→ primer A1

-463 CAGCTCCAGTTCCGTGCTGCCCAATGGGTTAGTGAGCAGCAGACGGTAGC

-413 GATCCTGGCCGGTTTGCTGCCAGAAAAAGCGGGCGTACACTTTTTGTTGGT

→ primer A7

-362 CAGAAATATAAGCGAACGCGCC**GCGAGTCTGATACTGATTAAGATTG**CG

-313 CACGTCTTGCTGATGCTGACGCCATTGTGGCGAATCCGGGCTTTTGCCAGG

→ primer A4

-262 ACCTTTGGGCGTGGTAACGGAACAGGCAGTGAGCACAAGAGC**AGCCAGCG**

-212 **GT*AGCAGGCGGATAAGACGAAAA***TCGGGCAGGGGCATAGTGATGACAAG

-162 TCCT***TGAGA***TACGTT***GCAGT***TATAACCCTTAATGCTAGCGTTACCGTCC

→ primer A6

-114 GCTATCGTCTATGTTCAAGTTGTCTTAATTGC**CAGAATCTAACGGCTTTCG**

3' AGTTCAACAGAATTAACGGTTCGAACG 5' ← primerA5

-

-63 GCAATTACTCAAAAGGGGGCGCTCTCTTTTATTGATCTTACGCATCCTG

-13 TATGATGCAAGCAGACTAACCCTATCAACGTTGGTATTATTT

3' TTGGGATAGTTCGAACCATAATAAA 5'← primer A2

Figure 20. The positions of *hemA* primers. The 5' to 3' primers are bold and underlined. When the primer is used for the other direction, it is marked with the 5' and 3' ends.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Since the 5' upstream region of *hemA* that we worked with is short (about 515 bp long), and there are no appropriate restriction enzyme sites for cloning it into the vector, pKK232-8, two primers which have the right restriction enzyme sites were designed. Primers A1 and A2 were used to create the insert in pWH515. The *Hin*dIII site in primer A2 has two mismatches compared to the original sequence; the original sequence in the 5' upstream region of *hemA* is AACGTT. To facilitate cloning, appropriate restriction enzyme sites were added to certain primers such as primers 6 and 7. Primers A2 and A7 were used to produce the 359 bp PCR product. Primers A2 and A4 were used to produce the 239 bp PCR product. Primer A2 and A6 were used to produce the 101 bp PCR product. Primers A1 and A5 were used to produce the 415 bp PCR product.

1.2. SUBCLONING OF VARIOUS FRAGMENTS OF THE 5' UPSTREAM REGION OF *hemA*

Since the two restriction enzyme cutting sites (*Sal*I and *Hin*dIII) are at the very ends of this fragment, direct cloning into the pKK232-8 vector, proved to be difficult. In pKK232-8, the reporter *cat* gene is flanked by efficient transcription terminators which prevent transcription from other pBR322 promoters (pKK232-8 is the derivative of pBR322). In addition the translational stop codons are introduced in all three frames upstream from the initiation codon of the *cat* gene, which prohibits reading through into the *cat* gene. Therefore, *cat* expression is solely driven by the DNA fragment which is in front of *cat*. It provides a powerful tool to study promoters and regulation. The restriction sites, available for cloning, are in the order: *Eco*RI, *Sma*I, *Bam*HI, *Sal*I, *Pst*I and *Hin*d III. Computer analysis for these restriction enzyme sites in the 5' upstream

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

region of *hemA* revealed that there were none in this region except one *Sal*I site. Therefore, PCR was used to create various fragments which had two appropriate restriction enzyme sites at each end, so the fragment could be spliced into the pKK232-8 in the correct orientation. To prevent *cat* expression from the *hemA* ribosome binding site, at the 3' end of this region, the Shine-Dalgarno sequence (ribosome binding site) of *hemA* was cut out after restriction enzyme digestion since it is right after the *Hin*d III site. The PCR product of *hemA* was directly cloned into a TA cloning vector to facilitate proper cutting of the fragment for subsequent subcloning.

The 515 bp PCR fragment inserted into pKK232-8 was sequenced manually and was shown to be correct. Automated sequencing of pWH359 (designed with a 1 bp mismatch at -328 bp), pWH239, and pWH415 showed 100% homology to their inserts sequences. DNA sequencing three times of one strand of pWH239* showed that it has a one base pair mismatch in the -10 sequence of P1 and two other 1 bp substitutions at -52 bp and -84 bp (Figure 19). DNA sequencing of pWH101 from both strands showed 100 % homology to its insert.

2. Regulation of Initiation of Transcription

2.1. PROMOTER STRENGTH OF THE 5' UPSTREAM REGION OF *hemA*

Plasmids: pWH515, pWH359, pWH239, pWH239*, pWH415, pWH101 were used for promoter strength analysis. They were first transformed individually into a *pcnB*⁻ strain, RP7947, so that only two to three copies of the plasmids were present in one cell thus preventing a dosage effect on gene expression. The level of expression of the reporter gene was measured by CAT ELISA assay. The results are shown in Figure 21.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Maximal expression of CAT was observed with the 515 bp fragment (pWH515). Promoter strength dropped with size of the fragment in the order: pWH515, pWH359, pWH239 (Table 5).

The construct which contains P1 and P2 (pWH239) expressed CAT. But a one base pair mutation in the -10 sequence of P1 (pWH239*) reduced the CAT expression to less than 5% of the level of CAT activity with pWH515. Cutting out the first 156 bp (-496 to -340) from the 5' end of the 515 bp fragment had a less dramatic effect on CAT expression but it did lower CAT expression (pWH359). Further cutting of another 120 bp (-340 to -220) lowered CAT expression but also had no dramatic effect on CAT expression (pWH239). But further cutting of another 138 bp (-220 to -82) (pWH101) reduced the CAT expression dramatically to less than 6% of the level with pWH515. pWH415, the structure which covers the first 415 bp from the 5' end of the 515 bp upstream region but is missing P1 showed less than 10% of the level of CAT activity with pWH515. We can conclude that P1 is necessary but not sufficient for the expression of the *hemA* gene. The P1 region is essential since one base pair mutation can reduce CAT expression dramatically. P2 and P3 are weak promoters. The 138 bp upstream of P1 is also important. It appears that the region upstream of P1 is required for high level *hemA* expression. This indicates that there may be some other important regulatory elements involved in these events. RNA polymerase may need looping of DNA and participation of upstream regions in addition to P1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

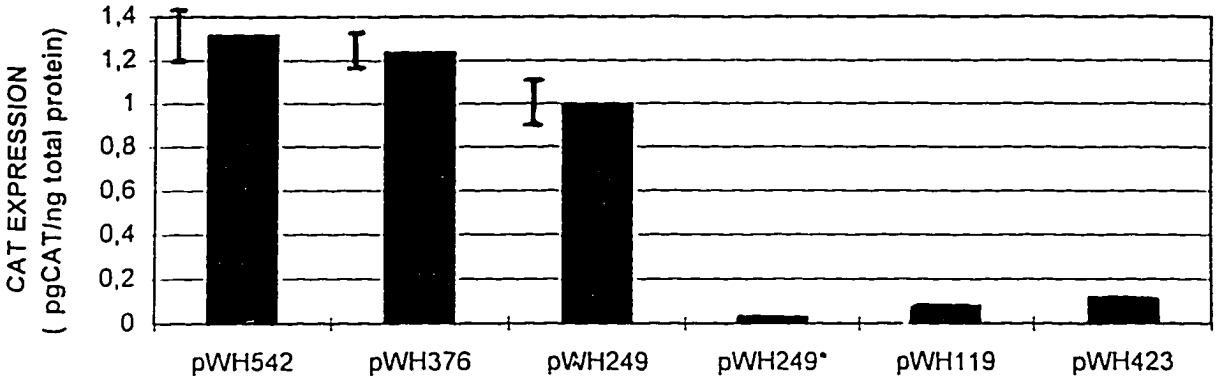

Figure 21. Promoter strength of *hemA*. The plasmids were transformed into the *pcnB*⁻ strain, RP7947, respectively. The transformants were grown overnight in LB medium under aerobic conditions. The bars show CAT expression as determined by CAT ELISA assay. These experiments were repeated four times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 5. Promoter strengths of various fragments of the 5' upstream region of *hemA*

Plasmid	CAT expression (pg CAT/ ng total protein)	Percentage
pWH515	1.32	100%
pWH359	1.24	93.9%
pWH239	1.00	75.8%
pWH239*	0.03	2.3%
pWH101	0.07	5.3%
pWH415	0.13	9.8%

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.2 THE EFFECT OF OXYGEN AND THE ROLES OF ArcA AND FNR

2.2.1. OXYGEN EFFECT

Since the tetrapyrrole biosynthetic pathway of *E. coli* can function under both aerobic and anaerobic conditions, the role of oxygen was tested. Strains with plasmids were grown aerobically and anaerobically.

The results of the effect of oxygen on CAT expression are shown in Figure 22. While the order of promoter strength remained the same, aerobically and anaerobically, the cells grown aerobically showed 1.3-fold more CAT expression.

2 2.2. EFFECT OF ArcA UNDER AEROBIC CONDITIONS

In response to oxygen availability, the transcriptional regulator ArcA controls the expression of many genes of aerobic metabolism. Since oxygen seemed to have an effect on *hemA* expression, the involvement of ArcA in *hemA* expression was tested. The six plasmids: pWH515, pWH359, pWH239, pWH239*, pWH415 and pWH101 were transformed respectively into the isogenic pair of *pcnB*⁻ strains: SP103 (*arcA*⁺) and SP104 (*arcA*⁻). They were grown aerobically and the CAT ELISA assay was performed. The results are shown in Figure 23.

The largest insert produced the greatest expression in both *arcA*⁺ and *arcA*⁻ strains. In addition the experiment showed that the *arcA*⁺ strain had twice as much CAT expression as the *arcA*⁻ strain for pWH515, pWH359 and somewhat less enhancement for pWH 239. With pWH239*, pWH415 and pWH101, very little expression could be detected. ArcA stimulated CAT expression. Thus ArcA appears to be an activator for maximal *hemA* expression.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

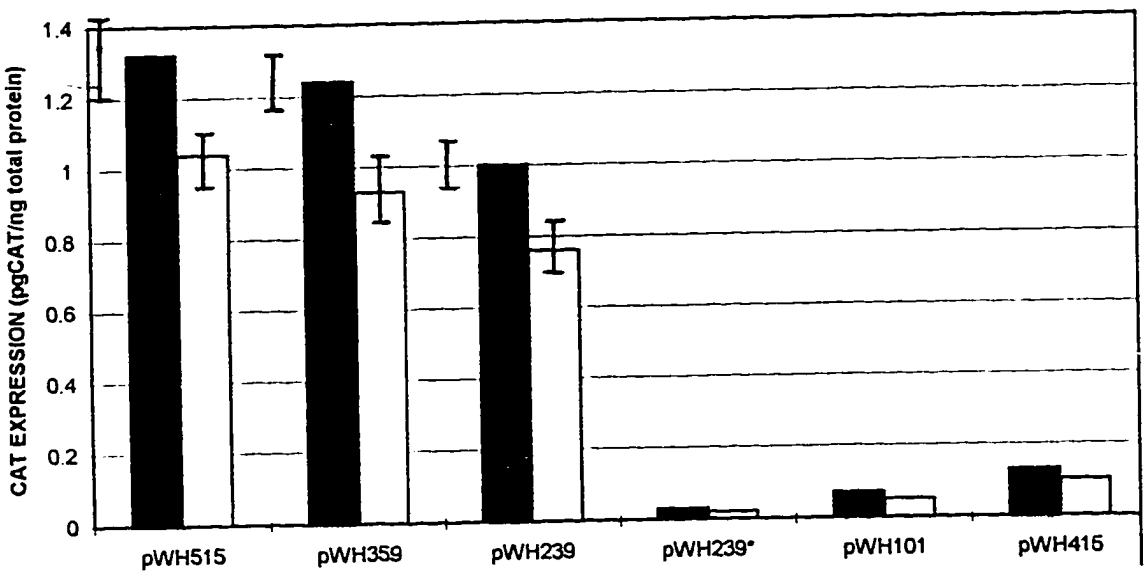

Figure 22. Oxygen effect on promoter strength. The plasmids were transformed into the *pcnB*⁻ strain, RP7947, respectively. The transformants were grown in LB overnight under aerobic and anaerobic conditions. The cells were then harvested and CAT ELISA assay was performed. The black bars show CAT expression for these strains grown with oxygen. The white bars show CAT expression for these strains grown anaerobically. These experiments were repeated three times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

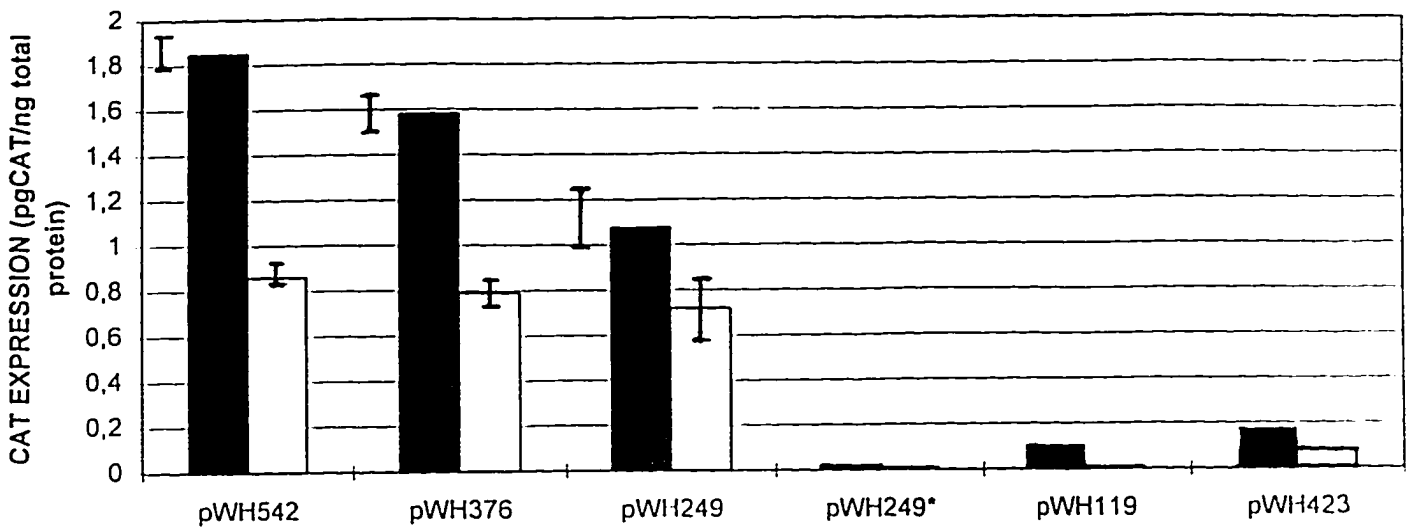

Figure 23. ArcA effect on promoter strength. The plasmids were transformed into SP103 and SP104 strains individually. The transformants were grown in LB medium overnight under aerobic condition. The cells were then harvested and CAT ELISA assay was performed. The black bars show CAT expression for *arcA*$^+$ strains (SP103). The white bars show CAT expression for *arcA*$^-$ strains (SP104). These experiments were repeated three times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.2.3. FNR EFFECT UNDER ANAEROBIC CONDITIONS

In response to oxygen availability, the transcriptional regulator, FNR, controls the expression of many genes required for anaerobic metabolism. The expression of anaerobic genes can only be induced by the activated FNR, suggesting that FNR represents the master-switch which ensures that aerobic respiration is used in preference to anaerobic respiration.

In order to determine if FNR is involved in *hemA* expression, the three plasmids (pWH515, pWH359, pWH239) were transformed into an isogenic pair of *pcnB*⁻ strains: WH105 (*fnr*⁺) and WH106 (*fnr*⁻). They were grown anaerobically and CAT ELISA assay was performed. The results are shown in Figure 24. FNR inhibited CAT expression by 1.8-fold with pWH515, by 1.5-fold with pWH359, and by 1.5-fold with pWH239. Thus, FNR appears to be a repressor of *hemA* expression.

2.3. THE EFFECT OF GLUCOSE AND THE ROLE OF CRP

Growth on glucose is known to alter cyclic AMP levels in many cases (Botsford, 1981). The transcriptional factor CRP can be activated by cAMP binding. But a glucose effect does not always indicate a role for cAMP (Eisenstein, *et al*, 1981).

Work done in our laboratory (Chen, unpublished data) showed that glucose inhibited ALA accumulation in a *hemB* mutant strain. To test if the basis of this glucose effect is on the transcriptional level and if CRP plays any role, the following experiments were performed.

The strains of RP7947 (*pcnB*⁻) which had been transformed with pWH515, pWH359 and pWH239 were grown with glucose (0.2%) and without glucose (1% glycerol) in minimal

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

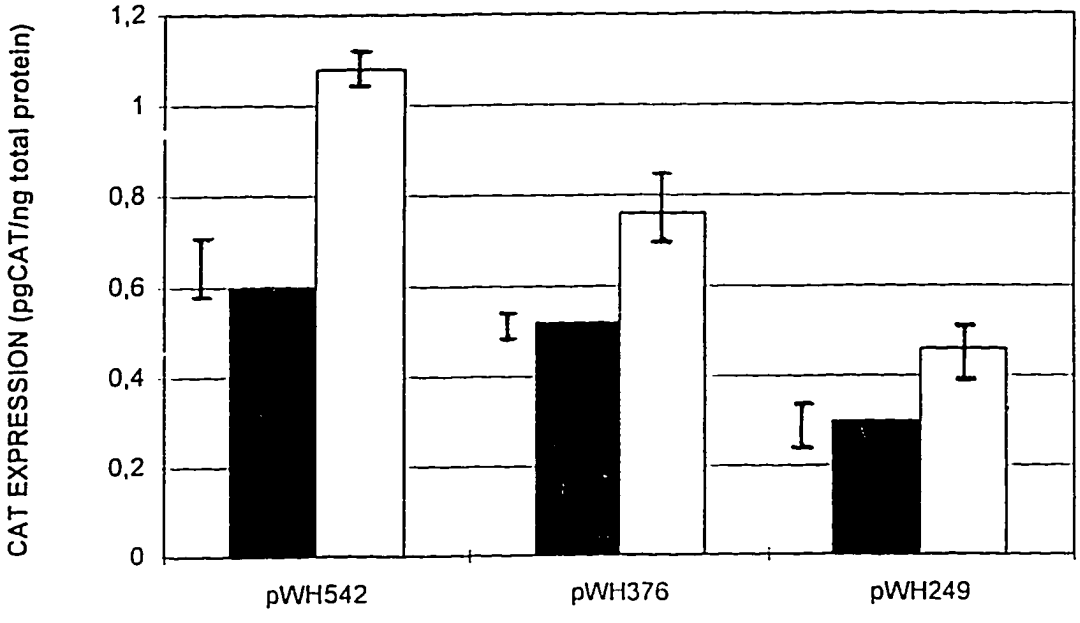

Figure 24. FNR effect on promoter strength. The plasmids were transformed into *fnr*$^{+}$ and *fnr*$^{-}$ strains (WH105 and WH106) respectively. The transformants were grown in LB medium overnight under anaerobic conditions. The cells were then harvested and CAT ELISA assay was performed. The black bars show CAT expression for *fnr*$^{+}$ strains (WH105). The white bars show CAT expression for *fnr*$^{-}$ strains (WH106). These experiments were repeated three times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

medium. Two sets of samples were collected, one at log phase of growth, the other at stationary phase. The results of CAT ELISA assay done on these cultures are shown in Figure 25. As can be seen, in log phase, the differences in CAT expression in cultures grown with and without glucose are not very pronounced (pWH515: 1.1-fold, pWH359: 1.1-fold and pWH239 1.3-fold). However, in the stationary phase, glucose inhibited *hemA*-CAT expression in all cases to about the same extent (pWH515: 2.2-fold, pWH359: 2.2-fold and pWH239: 2.1-fold). This indicates that in stationary phase there is a glucose effect.

In order to find out if the glucose effect is mediated by CRP, the three plasmids were also transformed individually into an isogenic pair of *crp*$^+$ (SP101) and *crp*$^-$ (SP102) strains which are *pcnB*$^-$. The CAT ELISA assay was performed on these cultures grown in minimal medium to stationary phase (Figure 26).

For *crp*$^+$ strains (SP101), glucose inhibited *hemA*-CAT expression in pWH515 by 1.7-fold, in pWH359 by 1.8-fold, and in pWH239 by 1.8-fold. The *crp*$^-$ strains did not show a glucose effect. These results suggest that CRP mediates the glucose effect on *hemA* expression. CRP appears to be an activator for maximal *hemA* expression in the stationary phase.

2. 4. THE EFFECT OF NITRATE AND THE ROLE OF NarL

Since nitrate reductase is a heme protein, and nitrate is an anaerobic electron acceptor, it appeared likely that *hemA* expression might be responsive to growth on nitrate and to the presence of the regulator protein NarL.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

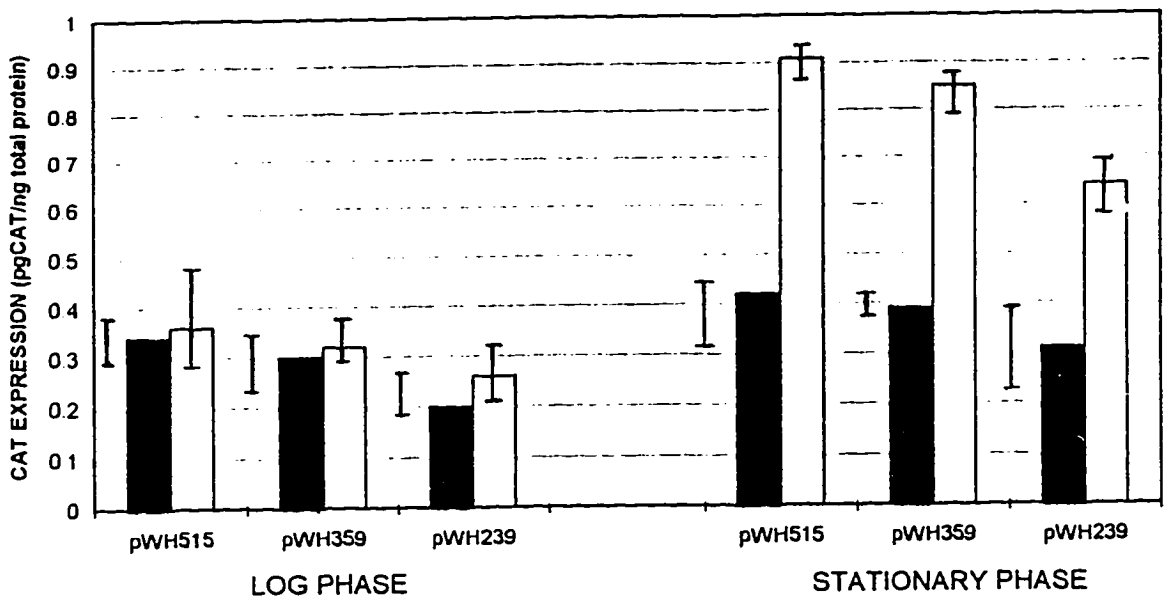

Figure 25. Glucose effect on promoter strength. The plasmids were transformed into the *pcnB*⁻ strain, RP7947, respectively. The black bars show CAT expression for samples grown on glucose. The white bars show CAT expression for samples grown on glycerol. These experiments were repeated three times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

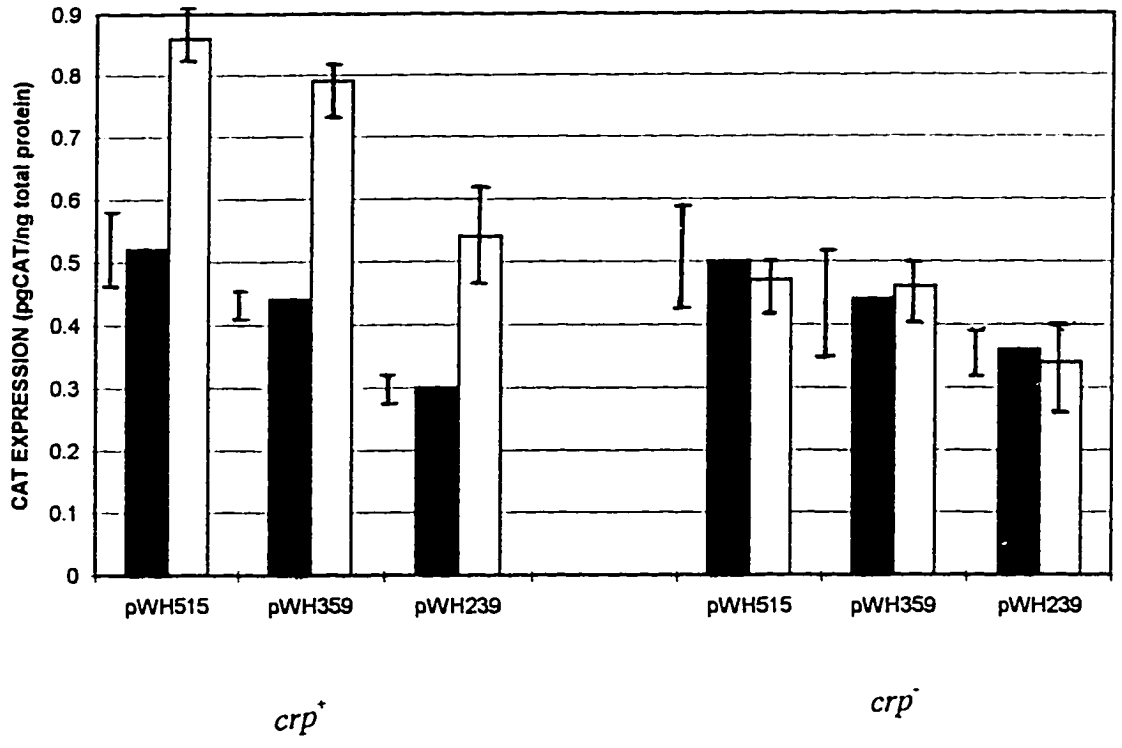

Figure 26. CRP effect on promoter strength. The black bars show CAT expression for samples with glucose. The white bars show CAT expression for samples without glucose. The first six bars show results with *crp*$^{+}$ strains (SP101). The following six bars show results with *crp*$^{-}$ strains (SP102). These experiments were repeated three times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The three plasmids: pWH515, pWH359 and pWH239 were transformed into an isogenic pair of *pcnB*⁻ strains: WH105 (*narL*⁺) and WH107 (*narL*⁻). Cells were grown anaerobically with and without nitrate. The results showed that nitrate induced CAT expression by 1.3-fold in the *narL*⁺ strain, WH105, but in the *narL*⁻ strain, WH107, there was very little difference between cells grown with and without nitrate (Figure 27). Thus it appears that nitrate weakly stimulates *hemA* expression under anaerobic conditions.

There is a putative NarL binding site in the 5' upstream region of *hemA*. A site mutation in the insert of pWH359 has one base substitution (G to C) (at -328 bp). This base is considered a critical base in the putative NarL binding site (TAGTGATT instead of TACTGATT) (Dong *et al.*, 1992). Although the putative NarL binding site is located on the insert of pWH515, is mutated on the insert of pWH359, and is not present on the insert of pWH239, CAT expression from all these plasmids showed the same increase in *narL*⁺ vs. *narL*⁻ host strains. Therefore we cannot determine, at this time, whether NarL is involved in the regulation of the *hemA* gene.

2.4. MEDIUM EFFECT

To test if the medium plays a role in *hemA* regulation, the six plasmids: pWH515, pWH359, pWH239, pWH239*, pWH415, and pWH101 were transformed into the *pcnB*⁻ strain, RP7947, respectively. They were grown in minimal (with glycerol) and rich (LB) medium. The effects on CAT expression are shown in Figure 28.

There was a difference in CAT expression between cells that grew with rich medium and minimal medium-M9 with casamino acids. The cells which were grown on rich medium had

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

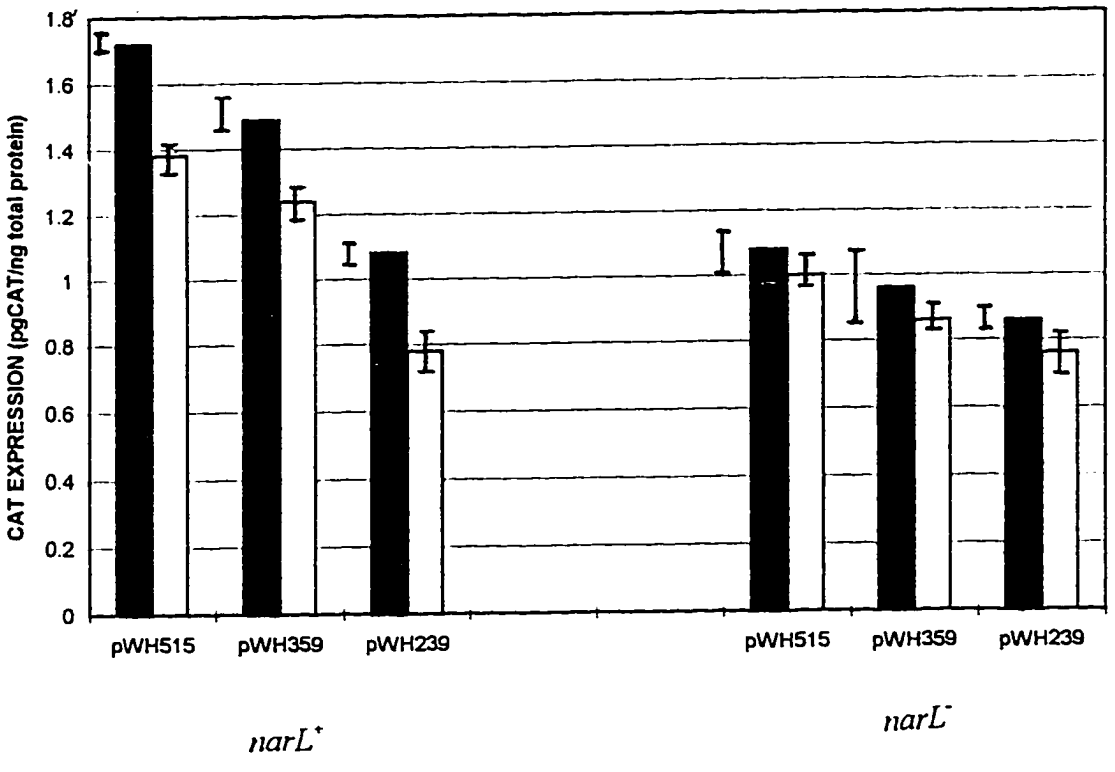

Figure 27. Nitrate and NarL effect on promoter strength. The plasmids were transformed into *narL*$^+$ (WH105) and *narL*$^-$ (WH107) strains respectively. The transformants were grown in LB medium to early log phase, and then 40 mM $NaNO_3$ was added under anaerobic conditions. The cells were then harvested and CAT ELISA assay was performed. The black bars show CAT expression for strains which grew with nitrate. The white bars show CAT expression for strains which grew without nitrate. These experiments were repeated three times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

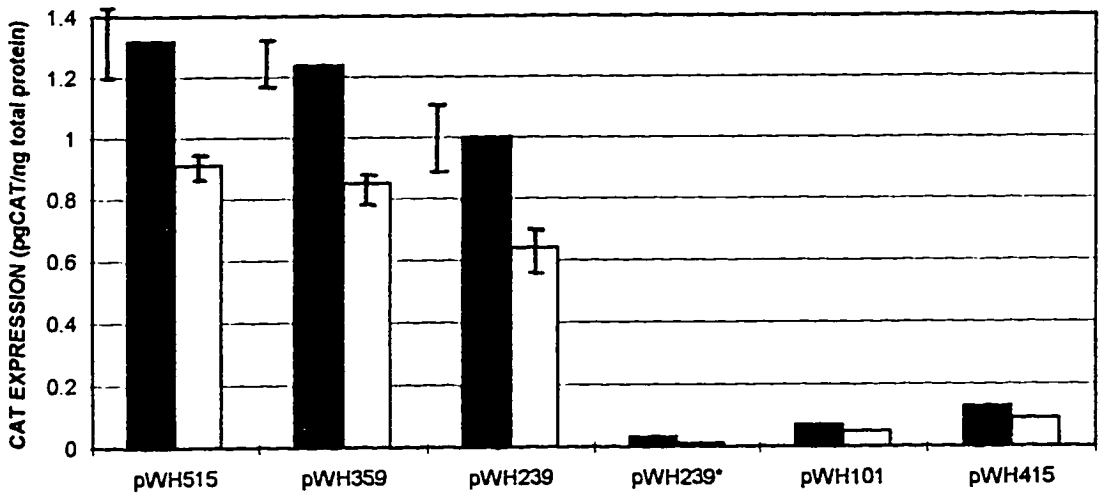

Figure 28. Medium effect on promoter strength. The plasmids were transformed into *pcnB*⁻ strain, RP7947, respectively. The transformants were grown in LB or minimal medium overnight under aerobic conditions. The cells were then harvested and CAT ELISA assay was performed. The black bars show CAT expression for these strains in LB medium. The white bars show CAT expression for these strains in minimal medium. These experiments were repeated three times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

about 1.5-fold more CAT expression than cells which were grown on minimal medium. Again with pWH239*, pWH415 and pWH101, CAT expression was at a low level.

2.5. EFFECT OF END PRODUCTS, ALA AND HEMIN

ALA is the end product of a three step reaction sequence in which the *hemA* gene product, GTR, is involved. Heme is one of the end products of the tetrapyrrole biosynthetic pathway. It is possible that the end products, ALA or heme, would regulate their own synthesis.

To test *hemA* expression in the situations where the cell is starved for ALA or heme, the three plasmids (pWH515, pWH359, pWH239) were transformed into HU227, which is a hemin-permeable *hemA* mutant strain. As seen in Figure 29, when these cells were starved for ALA, the CAT expression increased by 2.6-fold with pWH515. It also increased with pWH359 by 1.9-fold and by 2.1-fold with pWH239. When the cells were starved for heme, the *hemA*-CAT expression increased by 2.5-fold with pWH515, 1.7-fold with pWH359, and 2.5-fold with pWH239. These results suggest that ALA and heme starvation stimulate *hemA* expression and at the same level.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

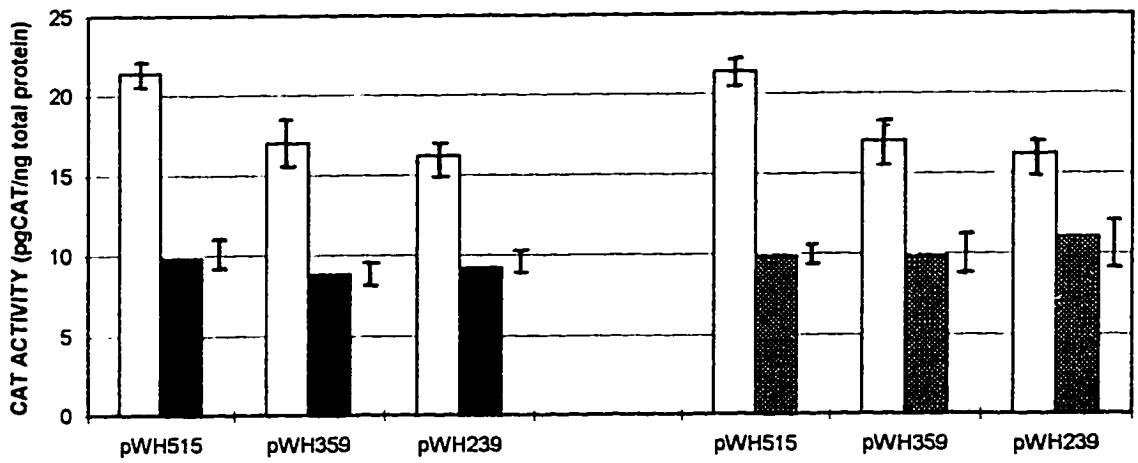

Figure 29. ALA and hemin starvation effect on promoter strength. The plasmids were transformed into HU227, which is a hemin-permeable *hemA* mutant strain, individually. The transformants were grown in LB medium overnight under aerobic conditions. The first set of cells were grown with 20 mM pyruvate and 50 mM KPO_4. The second set of cells were grown with 50 ug/ml ALA. The third set of cells were grown with 4 ug/ml hemin. These cells were then harvested and CAT ELISA assay was performed. The white bars show CAT expression for these strains starved for ALA and hemin. The black bars show CAT expression for those strains grown on ALA and the gray bars show CAT expression for those strains grown on hemin. These experiments were repaeated four times.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

DISCUSSION

Biosynthesis of tetrapyrroles is a regulated process in *E. coli*. Since supplementation with ALA causes accumulation of pathway intermediates, it is logical to suppose that regulation of the pathway occurs at the level of ALA formation. Glutamyl tRNA is also a precursor for protein synthesis, and GSA undergoes conversion to ALA non-enzymatically, so the step which forms GSA is probably the regulated step in ALA synthesis. Because *E. coli* is a facultative bacterium, heme is made for various purposes under many growth conditions, regulation is expected to be complex.

This work focuses on the regulation of *hemA* transcription initiation. By fusing the 5' upstream regulatory region of *hemA* to a promoterless CAT reporter gene, we were able to study *hemA* promoter activity and some of the elements which modulate that activity.

1. Promoter Strength of the 5' Upstream Region of *hemA*

We chose G as +1 (initiation of transcription) for P1 (Figure 5). The P1 and P2 sequences are identical to the ones that Verkamp and Chelm (1989) used. All three promoters are poorly homologous to the consensus sequences at -10 bp (TATAAT) and at -35 bp (TTGACA) upstream from the start site of transcription. P1 has one base mismatch (TATGAT vs. TATAAT) in the -10 sequence. One possible -35 sequence of P1 is better matched to the consensus (TTGATC vs. TTGACA). However, it has a 13 bp spacer which is less than the consensus 17 bp spacer. The other, more distal -35 sequence of P1 has three mismatches to the consensus (TTTATT vs. TTGACA), but it has a 16 bp spacer. P2 has two base pair mismatches in the -10 sequence (TATCGT vs. TATAAT) and it has three base pair mismatches in the -35 sequence

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(TTAATG vs. TTGACA). Thus according to consensus sequence, P1 and P2 seem to be weak promoters, and the affinity between these promoters and RNA polymerase is not expected to be strong.

The third promoter appears to be the most controversial. Sequences have been found that correspond to a σ^{70} and to a σ^{32} RNA polymerase binding site (P3 and P32). We consistently found that fragments containing these sequences are stronger promoters than fragments which lack them (see Table 5). The sequences of the putative σ^{70} and σ^{32} promoters in the 5' upstream region of *hemA* contain a number of mismatches from the actual consensus sequence. The -10 sequence of P32, CGCCAT has one mismatch to the consensus, CCCCAT; and the -35 sequence, ATTGCG has three mismatches to the consensus, CTTGAA. The -10 sequence of P3, GATGCT, has three mismatches to the consensus, TATAAT; and the -35 sequence of P3, TTAAGA, has two mismatches to the consensus, TTGACA. Thus both P3 and P32 appear to be weak promoters.

To study the function of different putative promoters and the role of regulators, different fragments of the 5' upstream region of *hemA* were cloned into a vector in front of a promoterless *cat* gene. pWH515 contains all the putative promoter sites plus putative regulator binding sites for FNR, NarL, CRP and ArcA proteins and two stem loops (Figure 5). pWH359 contains all the features of pWH515, except it is shorter than pWH515 and has one base pair mutation in the putative NarL binding site. pWH239 is shorter than pWH359 and does not contain the putative NarL site and stem loops. pWH239* is pWH239 with 3 base point mutations including one in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-10 sequence of P1 (Figure 19). The two other mutations occur at -52 bp and -84 bp. However, they do not appear to be in any relevant promoter sequences. The -84 bp is part of the putative ArcA binding site, and the mutation T is a mismatch to the consensus A. However, the wild type *hemA* 5' upstream sequence has a non-consensus G in that spot, so that base was not conserved in the original sequence. Therefore we believe that the elimination of promoter activity is due to the mutation in the -10 sequence of P1 alone. pWH101 contains only P1 and a FNR binding site. pWH415 lacks P1 and the putative FNR binding site, but contains all the other features present in pWH515.

The results we obtained show that promoter strength dropped with size in the order of pWH515 > pWH359 > pWH239. Suprisingly, pWH239*, pWH101 and pWH415 had very little CAT expression (Figure 21). This data reveals the complexity of *hemA* gene expression and regulation. Under the conditions of our experiments, the difference in expression between pWH239 and pWH239* suggested that P1 is essential for expression and P2 is a very weak promoter. Since pWH415 did not show a significant level of expression, P3 appears to be a very weak promoter, if it is, in fact, a promoter at all. Although P1 appears to be an essential functional promoter, the results with pWH101 demonstrate that it cannot function by itself. Some elements in the 138 bp sequence upstream are necessary for transcription as shown by results with pWH239. The fact that pWH 515 showed higher level expression than pWH359 suggests that there could be an enhancer in pWH515 which is not present in pWH359.

Verkamp and Chelm carried out S1 mapping and S1 protection experiments for *hemA* (Verkamp and Chelm, 1989). Their results suggested that there were two transcription start

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

points which correspond to P1 and P2 sites. They were not sure, however, whether there were two functional transcriptional start sites or some RNA processing (cleavage) events producing the two different RNA species. Since the probe that they used for mapping was a 447 bp fragment that included 144 bp of the *hemA* coding sequence and 267 bp of upstream sequence from transcription initiation site, there was no information about further upstream promoters. From S1 mapping, they identified two possible initiation of transcription sites, A or G, for P1 (at +1 bp); and one initiation of transcription site for P2 (T at -94 bp).

Drolet et al (1989) carried out primer extension experiments on *hemA*. They found one major band corresponding to P1. Their P1 was in the same region as the promoter identified by Verkamp and Chelm (1989), the only difference was that their -35 sequence was 4 bases shorter (TTTATT vs TTTATTGATC). The initiation of transcription site for this promoter was G.

Choi et al (1996), using primer extension analysis, with a *hemA* mutant, showed that the *hemA* P1and P2 were not significantly induced by ALA starvation, although expression of a *hemA-lacZ* fusion was 2-fold higher in cells that were starved for ALA than in cells grown with ALA. Interestingly, another longer transcript, designed by them as P3 appeared under conditions of ALA starvation. This indicated that there might be another initiation of transcription site upstream from P2. However, a *hemA-lacZ* fusion containing this study's putative P1, P2 and P3/P32 showed the same level of β-galactoside activity as a fusion containing only P1 and the -10 sequence of P2, under conditions of ALA starvation. This result does not correlate with their

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

primer extension results, and they suggested that their P3 transcript identified by primer extension was an artifact.

The results reported here are partially consistent with the results of Choi et al (1996). They constructed two mutants with alterations in the -10 sequence of P1 in *Salmonella typhimurium* and looked at *hemA-lacZ* expression. These mutants showed less than 10% of the *hemA-lac* expression compared to a functional P1. They also found out that all sequences important for expression of *hemA* under the conditions they examined lie within 128 bp upstream from the initiation of transcription site in both *E. coli* and *Salmonella typhimurium*.

We also used promoters containing one base pair mutation in the -10 sequence of P1. Our mutations differed in that the mutation was a T to A change in -13 bp and theirs was a T to A change in -8 bp. Both of these mutations had greatly reduced *hemA* expression. Therefore P1 is the major promoter for *hemA* expression under the conditions that we tested. Also we constructed a smaller fragment containing only 82 bp from the initiation of transcription site. We found that this fragment gave very little *hemA-CAT* expression although it contained P1. The 46 bp (-129 bp to -82 bp) difference between our construct and the construct of Choi et al (1996) seems to be very important for high level expression of *hemA*. Their 128 bp fragment contains P1, and the -10 sequence of P2, but not its -35 sequence, the initiation of transcription site for P1 and P2 and the putative FNR and ArcA binding sites. Our construct (pWH101) contains only P1, the initiation of transcription site for P1 and the putative FNR binding site. Based on our results, we conclude that P1 is necessary but not sufficient for *hemA* expression. The 46 bp upstream of P1 is essential for high level expression. These results imply that some kind of structure has to be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

formed for optimum expression from P1. Since computer analysis shows that P1 is basically a weak promoter, activators may be involved.

McNicholas et al (1997) found that *hemA-lacZ* expression from P2 is 10-15% of the level of expression from their full length fusion containing P1 and P2. They seem to have used the same construct as Choi et al for P1, and observed that P1 yielded 85-90% of expression. Since there is no information about how the constructs were made, and no data for the *hem-lacZ* expression were shown for the promoters, we are not able to compare our results with theirs. The important issue is the construction of the inserts. Both Choi et al and McNicholas et al made their *hemA-lacZ* fusions by using inserts which include part of the *hemA* coding region. It is important to point out that our constructs contained only the upstream region of *hemA*.

2. Possible Regulators of *hemA* Transcription

The efficiency of a promoter sets the basal rate for gene expression. However, regulatory proteins can change the rate of transcription for a particular gene according to the needs of the cell under different conditions. Positive regulators can strengthen the binding of RNA polymerase to achieve maximal expression. Negative regulators can also play a role by blocking or decreasing the affinity of binding of RNA polymerase. The regulators can directly or indirectly (through another regulator) bind to RNA polymerase. They could participate in and facilitate the formation of a three dimensional transcription initiation complex or prevent its formation. The regulatory proteins that we have examined in this study are global regulators; we have not yet found any specific regulator protein for *hemA*.

2.1. Effect of Oxygen and Roles of ArcA and FNR

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

hemA is expressed aerobically and anaerobically. ArcA and FNR are considered to be the master switches between aerobic and anaerobic conditions (Guest, 1992). Our results (Figure 22) show that the level of *hemA* expression is higher aerobically than anaerobically. It is consistent with the fact that ArcA stimulates *hemA* expression under aerobic conditions and FNR represses *hemA* expression under anaerobic conditions. There are reports that *E. coli* cells contain more heme when grown under aerobic conditions than anaerobic conditions (Hino and Ishida, 1973). So far, there is no evidence that oxygen is directly involved in ALA synthesis regulation. Beale (1996) pointed out that the primary effect of oxygen is probably on the induction of various apocytochromes, and heme synthesis probably increases in response to depletion of the pool of free heme in order to supply prosthetic groups to the apocytochromes. It is also possible that other regulators are responsible for the effect of oxygen on *hemA* expression.

As can be seen in Figure 23, the two largest inserts (pWH515 and pWH359) show the strongest stimulation by ArcA. Expression of CAT in pWH239 is also stimulated by ArcA, but to a lesser degree. This is consistent with the observation that the putative ArcA binding site (-92 bp to -83 bp) exists in all three constructs.

Choi et al (1996) did not find a clear effect of ArcA in their ALA starvation experiments of *E. coli*. Expression of *hemA-lacZ* was not increased by overexpression of *arcA* during ALA starvation. However, in the presence of ALA, the *arcA*$^-$ strain had an increase in the expression of *hemA-lacZ* by 1.2-fold over that of *arcA*$^+$ strain. In our experiments, ArcA stimulated CAT expression by 2-fold (Figure 23). The difference between our results and those of Choi et al

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

could be due to the different host strains that we used. They used *hemA* auxotroph strains, and We used *hemA*$^+$ strains grown without ALA supplementation.

The FNR binding site of *hemA* overlaps the -35 sequence of P1. Comparing this binding site (TTGATnnnnCGCAT) to the FNR consensus sequence (TTGATnnnnATCAA) reveals that CGCAT is a poor match to the consensus sequence ATCAA. This may indicate an FNR monomer can bind only to a single half-site. Most FNR binding involves FNR dimers binding to two symmetrical half sites with dyad symmetry. However, Melville and Gunsalus (1996) using *in vitro* DNA footprinting, found that the size of the DNA footprint on the *hemA* 5' upstream region was consistent with an FNR monomer binding to a single half-site.

The results reported here are consistent with these findings. FNR represses *hemA* expression at a level which is not very high, probably because it only binds to the half site. Since the FNR binding site overlaps the -35 sequence of P1, it is logical that it would act as a repressor of *hemA* expression. This repression was found in pWH515, pWH359, and pWH239. Thus, experimental evidence for FNR involvement in *hemA* regulation is drawn from footprinting and its effect on promoter activity.

2.2. Effect of Glucose and the Role of CRP

Previous studies in our laboratory showed that glucose down-regulates ALA formation in a *hemB* mutant (Chen, unpublished data). This study shows that in strains which lack CRP (*crp*$^-$), there is no glucose effect on expression of the reporter gene. This suggests that CRP is involved in *hemA* expression and that the glucose effect is mediated by CRP. However, this effect is observed only in stationary phase cells but not in log phase cells. Stationary phase cells have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

metabolism and features which differ from log phase cells, including steady-state growth, limitation for nutritional sources of either carbon, nitrogen or phosphate, and production of peroxide.

The CRP binding site consists of two symmetrical half sites with dyad symmetry (TGTGAnnnnnnTCACA). The putative CRP binding site of *hemA* (Figure 5) is poorly homologous with the consensus. There is one base pair mismatch in TGAGA and there are three base pair mismatches in GCAGT. This CRP site is probably a single half-site like the FNR binding site in *hemA*. Therefore CRP binding may not be strong in the *hemA* upstream region.

2.3. Effect of Growth Phase

In general, the higher expression of *hemA* was observed in stationary phase cells (data not shown). This may be because *E. coli* cells have more demand for catalases and peroxidases in the stationary phase which correlates with greater *hemA* expression under aerobic growth. Stationary-phase-specific σ factor, σ^S, binds to core RNA polymerase and recognizes a specific promoter sequence (Hengge-Aronis, 1993). The consensus sequence: -10 sequence: CCCCAT and -35 sequence: CTTGAA is the same consensus sequence for the σ^{32} promoter. Inspection of the *hemA* upstream region reveals P32 and a sequence near the P1 region which resemble σ^S promoters. The sequence near P1 has the -10 sequence (+7 bp to +12 bp) with two mismatches to consensus (CGCTAT vs. CCCCAT). The -35 sequence (-17 bp to -12 bp) has two mismatches to consensus (CTTAAT vs. CTTGAA). Whether P32 or the sequence near P1 is a σ^S promoter is unknown. Since P32 was observed by Choi et al (1996) only in ALA starved cells,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and both ALA starved and non-starved cells were stationary phase cells, it is unlikely that P32 is the σ^S promoter. Further investigation has to be done to clarify this point.

Although σ^S and σ^{32} recognize the same consensus sequence, they are encoded by different genes. σ^S is encoded by the *rpoS* and σ^{32} is encoded by *rpoH*. In order to find out if these genes are involved in *hemA* expression, CAT expression could be tested in cells with a mutation in *rpoS* or *rpoH*.

2.4. Effect of Nitrate and Role of NarL

Since nitrate is an anaerobic electron acceptor and nitrate reductase is a heme protein, it is logical to think that nitrate induction might have an effect on heme biosynthesis. A putative NarL binding site in the 5' upstream region of *hemA* has been found (Figure 5).

The two-component signal transduction systems for nitrate and nitrite induction are NarX:NarL and NarQ:NarP. The results in Figure 27 show a small nitrate effect on CAT expression. It seems that nitrate stimulates CAT expression slightly under anaerobic conditions. However a site mutation in the putative NarL binding site had no effect on CAT expression (pWH359). pWH 239 which does not contain the putative NarL binding site had a similar response to nitrate induction as pWH 515. This suggests that the putative NarL binding site that we identified is not the real one. Therefore we cannot determine, at this time, whether NarL is involved in the regulation of the *hemA*. There may be another putative NarL binding site in the 5' upstream region of *hemA*. A deletion of the whole binding site may reveal more information. Since there is no perfect match for the putative NarL binding site in the 5' upstream region of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

hemA, the fact that the nitrate effect is not very pronounced is not surprising. It is possible that NarP instead of NarL is involved.

2.5. Effect of Medium

LB medium is composed of tryptone and yeast extracts, thus containing all the of amino acids, vitamins, nucleotides and other nutrients needed for maximal cell growth. The effect of medium can be complex. Since cells grow faster and better in LB medium than minimal medium, the difference in *hemA* expression could be a reflection of the growth rate. A greater growth rate implies more synthesis. Another possibility is that in rich medium, there is probably more tRNA. *hemA* transcription is enhanced by $tRNA^{Glu}$ (Chen *et al*, 1996).

2.6. Effect of End Products: ALA and Heme Starvation

Overexpression of a rat cytochrome b_5 gene in *E. coli* results in an increase in total cellular heme (Woodard and Dailey, 1995). This suggests that the cellular heme concentration is a factor in regulation of heme synthesis.

HU227 is a heme-permeable *hemA* mutant which requires ALA or heme for growth. Without ALA or heme, HU227 grows fermentatively. This strain was used as a tool to test how *hemA* promoters react to ALA or heme starvation. *hemA*-CAT expression in those cells starved for ALA increased about 2-fold over cells grown with ALA (Figure 29). Similar results were obtained with heme starvation. This suggests that *hemA* is an inducible gene. The mechanism of induction may involve a sensor protein which responds to the free heme pool.

As mentioned above, Choi et al (1996) found that expression of *lacZ* was induced during starvation of a *hemA* auxotroph. The level of induction observed was maximally 6-fold for *E. coli*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and 2 to 4-fold for *S. typhimurium*. The difference between these results and the ones reported here may be due to the fact that they were working with a single copy chromosomal fusion and I was working with a reporter gene on a plasmid with limited copy number. The regulation of a single copy and more than one copy may be different. In addition, since the plasmid that I used has a transcription terminator and stop codons to prevent read through from regions upstream from studied promoter, the reporter gene expression of our plasmid may be lower than that found in the *hemA-lacZ* fusions.

2.7. Concluding Remarks

The *hemA* gene lies in a possible operon with *pfrA*, encoding transcription release factor 1 and *hemK* whose function is unknown. Thus regulation of the initiation of transcription of *hemA* also affects the synthesis of these gene products. Tight regulation of transcription of *hemA* could be deleterious to the cell, destroying the normal process of translation termination. The levels of inhibition or enhancement of CAT expression by ArcA, CRP, NarL and FNR are not dramatic (about 2-fold). These regulators may slightly modulate the ability of the cells to synthesize heme according to the need for heme by the cell in different environments. However, the main control or additional control may be another step.

Major regulation of *hemA* expression might be at steps after initiation of transcription such as at transcription elongation or termination and stability of RNA messages. Although not common in *E. coli*, regulation could be at a translation or post-translation level. The overlapping stem loops that occur in the 5' upstream region of *hemA* and which resemble a stem loop from $tRNA^{Glu}$ and the stem loop in the *gltX* gene, which encodes glutamyl-tRNA synthetase (GTS),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

might be involved in translational regulation of *hemA* (Brun *et al*, 1990). The fact that these stem loops are near the putative ribosome binding site indicates that the secondary structure of *hemA* mRNA could control the availability of the translation initiation site. It is possible that GTR or GTS could bind to these stem loops and change the structure, therefore controlling ribosome binding to *hemA* mRNA and the initiation of translation. On the other hand, it is also possible that GTR could interact with the $tRNA^{Glu}$-like stem loop of the *gltX* mRNA to control this gene expression. Therefore protein synthesis and tetrapyrrole synthesis could be coordinately regulated by $tRNA^{Glu}$ and regulatory elements for ALA synthesis. Recent studies in our laboratory suggest that the regulation of *hemA* by heme occurs at the post-transcriptional step (Knight, unpublished data).

Another reason that we have not observed strong up or down regulation for *hemA* expression could also be that dramatic changes in free heme pool levels would be lethal. The main players in regulation of *hemA* may be specific regulators for *hemA* which would have a dramatic effect on the regulation. This work found some clues about regulation of *hemA* expression. But the whole picture is still a puzzle with a lot of missing pieces.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bibliography

Adhya, S., and S. Gatges. 1990. Positive control. J. Biol. Chem. **265:**10797-10800.

Avissar, Y. J., J. G. Ormerod, and S. I. Beale. 1989. Distribution of δ-aminolevulinic acid pathways among photosynthetic bacteria and related organisms. Arch. Microbiol. **151:**513-519.

Beale, S. I. 1996. Biosynthesis of hemes. *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. 2nd ed. American Society for Microbiology, Washington, D. C. pp. 731-748.

Beale, S. I., and P. Castelfranco. 1974. The biosynthesis of δ-aminolevulinic acid in higher plants. Plant Physiol. **53:**291-296.

Beker, S., G. Holighaus, T. Gabrielczyk, and G. Unden. 1996. O_2 as the regulatory signal for FNR-dependent gene regulation in *Escherichia coli.* J. Bacteriol. **178:**4515-4521.

Berg, B. L., and V. Stewart. 1990. Structural genes for nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. Genetics. **125:**691-702.

Brosius, J. 1984. Plasmid vectors for the selection of promoters. Gene. **27:**151-160.

Brun, Y. V., H. Sanfacon, R. Breton, and J. Lapointe. 1990. Closely-spaced and divergent promoters for an aminoacyl-tRNA synthetase gene and a tRNA operon in *Escherichia coli.* J. Mol. Biol. **214:**845-864.

Bobik, T. A., M. Ailion, and J. R. Roth. 1992. A single regulatory gene integrates control of vitamin B_{12} synthesis and propanediol degradation. J. Bacteriol. **174:**2253-2266.

Bonnefoy, V., and J DeMoss. 1992. Identification of functional cis-acting sequences involved in regulation of *narK* gene expression in *Escherichia coli*. Mol. Microbiol. **6:**3595-3602.

Botsford, J. L. 1981. Cyclic nucleotides in procaryotes. Microbiol. Rev. **45:**620-642.

Botsford, J. L., and J. G. Harman. 1992. Cyclic AMP in prokaryotes. Microbiol. Rev. **56:**100-122.

Cavicchioli, R., I. Schroder, M. Constanti, and R. P. Gunsalus. 1995. The NarX and NarQ sensor-transmitter proteins of *Escherichia coli* each require two conserved histidines for nitrate-dependent signal transduction to NarL. J. Bacteriol. **177:**2416-2424.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chen, M. W., D. Jahn, G. P. O'Neill, and D. Soll. 1990. Purification of the glutamyl-tRNA reductase from *Chlamydomonas reinhardtii* involved in δ-aminolevulinic acid formation during chlorophyll biosynthesis. J. Biol. Chem. **265:**4058-4063.

Chen, W., C. S. Russell, Y. Murooka, and S. D. Cosloy. 1994. 5-Aminolevulinic acid synthesis in *Escherichia coli* requires expression of *hemA*. J. Bacteriol. **176:**2743-2746.

Chen, W., L. Wright, S. Lee, S. D. Cosloy,and C. S. Russell. 1996. Expression of glutamyl-tRNA reductase in *Escherichia coli.* Biochim. Biophys. Acta. **1309:**109-121.

Chiang, R. C., R. Cavicchioli, and R. P. Gunsalus. 1992. Identification and characterization of *narQ*, a second nitrate sensor for nitrate-dependent gene regulation in *Escherichia coli*. Mol. Microbiol. **6:**1913-1923.

Choi, P., L. Wang, D. Archer, and T. Elliot. 1996. Transcription of the glutamyl-tRNA reductase (*hemA*) gene in *Salmonella typhimurium* and *Escherichia coli*: role of the *hemA* P1 promoter and the *arcA* gene product. J. Bacteriol. **178:**638-646.

Cotter, P. A., and R. P. Gunsalus. 1989. Oxygen, nitrate, and molybdenum regulation of *dmsABC* gene expression in *Escherichia coli*. J. Bacteriol. **171:**3817-3823.

Cotter, P. A., V. Chepuri, R. B. Gennis, and R. P. Gunsalus. 1990. Cytochrome o (*cyoABCDE*) and d (*cydAB*) oxidase gene expression in *Escherichia coli* is regulated by oxygen, pH, and the *fnr* gene product. J. Bacteriol. **172:**6333-6338.

Cotter, P. A., S. Darie, and R. P. Gunsalus. 1992. The effect of iron limitation on the expression of the aerobic and anaerobic electron transport pathway genes in *Escherichia coli.* FEMS Microbiol. Lett. **100:**227-232.

Darie, S., and R. P. Gunsalus. 1994. Effect of heme and oxygen availability on *hemA* gene expression in *Escherichia coli*: role of the *fnr*, *arcA*, and *himA* gene products. J. Bacteriol. **176:**5270-5276.

Davis, R. W., D. Botstein and J. R. Roth. 1980. Advanced Bacterial Genetics. Cold Spring Harbor Laboratory, Publisher, Cold Spring Harbor, N.Y. pp 140-141.

Dong, X. R., S. F. Li, and J. A. DeMoss. 1992. Upstream sequence elements required for NarL-mediated activation of transcription from the *narGHJI* promoter of *Escherichia coli.* J. Biol. Chem. **267:**14122-14128.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Drolet, M., L. Peloquin, Y. Echelard, L. Cousineau, and A. Sasarman. Isolation and nucleotide sequence of the *hemA* gene of *Escherichia coli*. Mol. Gen. Genet. **216:**347-352.

Ebright, R. H., P. Cossart, B. Gicquel-Sanzey, and J. Beckwith. 1984. Mutation that alter the DNA sequence specificity of the catabolite gene activator protein of *E. coli*. Nature. **311:**232-235.

Eiglmeier, K., N. Honore, S. Iuchi, E. C. C. Lin, and S. T. Cole. 1989. Molecular genetic analysis of FNR-dependent promoters. Mol. Microbiol. **3:**869-878.

Eisenstein, B. I., and D. C. Dodd. 1982. Pseudocatabolite repression of type 1 firmbriae of *Escherichia coli*. J. Bacteriol. **151:**1560-1567.

Elliott T. 1989. Cloning, genetic characterization, and nucleotide sequence of the *hemA-prfA* operon of *Salmonella typhimurium*. J. Bacteriol. **171:**3948-3960.

Elliott, T. 1992. A method for constructing single copy lac fusions in *Salmonella typhimurium* and its application to the *hemA-prfA* operon. J. Bacteriol. **174:**245-253.

Engel, P., M. Tragester, and G. Unden. 1991. Reversible interconversion of the functional state of the gene regulator FNR from *in vivo* by O_2 and iron availability. Arch. Microbiol. **156:**463-470.

Friedmann, H. C., and R. K. Thauer. 1992. Macrocyclic tetrapyrrole biosynthesis in bacteria. Encyclopedia of Microbiology, vol 3. Academic Press, Inc. pp. 1-18.

Friedmann, H.C., R. K. Thauer, S. P. Gough, and C. G. Kannangara. 1987. Δ-Aminolevulinic acid formation in the *Archaebacterium, Methanobacterium thermoautotrophicum,* requires $tRNA^{Glu}$. Carlsberg Res. Commun. **52:**363-371.

Gaston, K., A. Bell, A. Kolb, H. Buc, and S. Busby. 1990. Stringent spacing requirements for transcription activation by CRP. Cell. **62:**733-743.

Giladi, H., K. Igarashi, A. Ishihama, and A. B. Oppenheim. 1992. Stimulation of the phage λ pL promoter by integration host factor requires the carboxy terminus of the α-subunit of RNA polymerase. J. Mol. Biol. **227:**985-990.

Green, J., and J. R. Guest. 1993. Activation of FNR-dependent transcription by iron: an *in vitro* switch for FNR. FEMS Microbiol. Lett. **113:**219-222.

Grimm, B., A. Bull, and V. Breu. 1991. Structural genes of glutamate 1-semialdehyde aminotransferase for porphyrin synthesis in a *cyanobacterium* and *Escherichia coli*. Mol. Gen. Genet. **225:**1-10.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Guest, J. R.. 1992. Oxygen-regulated gene expression in *Escherichia coli*. J. Gen. Microbiol. **138:**2253-2263.

Gunsalus, R. P. 1992. Control of electron flow in *Escherichia coli*: coordinated transcription of respiratory pathway genes. J. Bacteriol. **174:**7069-7074.

Hengge-Aronis, R. 1993. Survival of hunger and stress: the role of *rpoS* in early stationary-phase gene regulation in *E. coli*. Cell. **72:**165-168.

Hino, S., and A. Ishida. 1973. Effect of oxygen on heme and cytochrome content in some facultative bacteria. Enzyme. **16:**42-49.

Hughes, P., A. Landouls, and M. Kohiyama. 1988. A novel role for cAMP in control of the activity of the *E. coli* chromosome replication initiator protein, DnaA. Cell. **55:**343-350.

Ikemi, M., K. Murakami, M. Hashimoto, and Y. Murooka. 1992. Cloning and characterization of genes involved in the biosynthesis of δ-aminolevulinic acid in *E. coli*. Gene. **121:**127-132.

Ilag, L. L., D. Jahn, G. Eggertsson, and D. Soll. 1991. The *Escherichia coli hemL* gene encodes glutamate semialdehyde amino transferase. J. Bacteriol. **173:**3408-3413.

Iuchi, S., and E. C. C. Lin. 1987. The *narL* gene product activates the nitrate reductase operon and represses the fumarate reductase and trimethylamine *N*-oxide reductase operons in *Escherichia coli*. Proc. Natl. Acad. Sci. USA. **84:**3901-3905.

Iuchi, S., and E. C. C. Lin. 1988. *arcA* (*dye*), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. Proc. Natl. Acad. Sci. USA **85:**1888-1892.

Iuchi, S., Z. Matsuda, T. Fujiwara, and E. C. C. Lin. 1990. The *arcB* gene of *Escherichia coli* encodes a sensor-regulator protein for anaerobic repression of the *arc* modulon. Mol. Microbiol. **4:**715-727.

Iuchi, S., and E. C. C. Lin. 1991. Adaptation of *Escherichia coli* to respiratory conditions: regulation of gene expression. Cell. **66:**5-7.

Iuchi, S., and E. C. C. Lin. 1992. Mutational analysis of signal transduction by ArcB, a membrane sensor protein responsible for anaerobic repression of operons involved in the central aerobic pathway in *Escherichia coli*. J. Bacteriol. **174:**3972-3980.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Iuchi, S. 1993. Phosphorylation/dephosphorylation of the receiver module at the conserved aspartate residue controls transphosphorylation activity of histidine kinase in sensor protein ArcB of *Escherichia coli.* J. Biol. Chem. **268:**23972-23980.

Iuchi, S., and E. C. C. Lin. 1993. Adaptation of *Escherichia coli* to redox environments by gene expression. Mol. Microbiol. **9:**9-15.

Iuchi, S., A. Aristarkhov, J. M. Dong, J. S. Taylor, and E. C. C. Lin. 1994. Effects of nitrate respiration on expression of the Arc-controlled operons encoding succinate dehydrogenase and flavin-linked L-lactate dehydrogenase. J. Bacteriol. **176:**1695-1701.

Jayaraman, P. S., T. Peakman, S. Busby, R. Quincey, and J. Cole. 1987. Location and sequence of the promoter of the gene for the NADH-dependent nitrite reductase of *Escherichia coli* and its regulation by oxygen, the Fnr protein and nitrite. J. Mol. Biol. **196:**781-788.

Jones, H. M., and R. P. Gunsalus. 1987. Regulation of *Escherichia coli* fumarate reductase (*frdABCD*) operon expression by respiratory electron acceptors and the *fnr* gene product. J. Bacteriol. **169:**3340-3349.

Jordan, P. M., B. I. A. Mgbeje, A. F. Alwan, and S. D. Thomas. 1987. Nucleotide sequence of *hemD*, the second gene in the *hem* operon of *Escherichia coli* K12. Nucl. Acids Res. **15:**10583.

Jordan, P. M., B. I. A. Mgbeje, S. D. Thomas, and A. F. Alwan. 1988. Nucleotide sequence for the *hemD* gene of *Escherichia coli* encoding uroporphyrinigen III synthase and initial evidence for a *hem* operon. Biochem. J. **239:**613-616.

Kalman, L. V., and R. P. Gunsalus. 1988. The *frdR* gene of *Escherichia coli* globally regulates several operons involved in anaerobic growth in response to nitrate. J. Bacteriol. **170:**623-629.

Kalman, L. V., and R. P. Gunsalus. 1989. Identification of a second gene involved in global regulation of fumarate reductase and other nitrate-controlled genes for anaerobic respiration in *Escherichia coli.* J. Bacteriol. **171:**3810-3816.

Kalman, L. V., and R. P. Gunsalus. 1990. Nitrate- and molydenum-independent signal transduction mutations in *narX* that alter regulation of anaerobic respiratory genes in *Escherichia coli.* J. Bacteriol. **172:**7049-7056.

Kolesnikow, T., I. Schroder, and R. P. Gunsalus. 1992. Regulation of *narK* gene expression in *Escherichia coli* in response to anaerobiosis, nitrate, iron, and molybdenum. J. Bacteriol. **174:**7104-7111.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Lazazzera, B. A., D. M. Bates, and P. J. Kiley. 1993. DNA binding and dimerization of the Fe-S-containing FNR protein from *Escherichia coli* are regulated by oxygen. Genes Dev. **7:**1993-2005.

Li, J. M., S. D. Cosloy, and C. S. Russell. 1988. 5-Aminolevulinic acid synthesis in *Escherichia coli.* J. Cell Biol. **107:**617a.

(a) Li, J. M., O. Brathwaite, S. D. Cosloy, and C. S. Russell. 1989. 5- Aminolevulinic acid synthesis in *Escherichia coli.* J. Bacteriol. **171:**2547-2552.

(b) Li, J. M., C. S. Russell, and S. D. Cosloy. 1989. Cloning and structure of the *hemA* gene of *Escherichia coli* K-12. Gene **82:**209-217.

Liu, J., and J. S. Parkingson. 1989. Genetics and sequence analysis of the *pcnB* locus, an *Escherichia coli* gene involved in plasmid copy number control. J. Bacteriol. **171:**1254-1261.

Lynch, A. S., and E. C. C. Lin. 1996. Transcriptional control mediated by the ArcA two-component response regulator protein of *Escherichia coli:* characterization of DNA binding at target promoters. J. Bacteriol. **178:**6238-6239.

May, B. K., I. A. Borthwick, G. Srivastava, B. A. Pirola, and W. H. Elliot. 1986. Control of 5-aminolevulinate synthase in animals. Curr. Top. Cell. Regul. **28:**233-261.

McClung. R., J. E. Somerville, M. L. Guerinot, and B. K. Chelm. 1987. Structure of *Bradyrhizobium japonicum* gene *hemA* encoding 5-aminolevulinic acid synthase. Gene. **54:**133-139.

McNicholas, P. M., G. Javor, S. Darie, and R. P. Gunsalus. 1997. Expression of the heme biosynthetic pathway genes *hemCD*, *hemH*, *hemM* and *hemA* of *Escherichia coli*. FEMS Lett. **146:**143-148.

Melville, S., and R. P. Gunsalus. 1996. Isolation of an oxygen-sensitive FNR protein of *Escherichia coli*: interaction at activator and repressor sites of FNR-controlled genes. Proc. Natl. Acad. Sci. USA. **93:**1226-1231.

Miyamoto, K., K Nakahigashi, K. Nishimura, and H. Inokuchi. 1991. Isolation and characterization of visible light-sensitive mutants of *Escherichia coli*. J. Mol. Biol. **219:**393-398.

Mollegaard, N. E., P. B. Rasmussen, P Valentin-Hansen, and P. E. Nielsen. 1993. Characterization of promoter recognition complexes formed by CRP and CytR for repression and by CRP and RNA polymerase for activation of transcription on the *Escherichia coli deo*P2. J. Biol. Chem. **268:**17471-17477.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Murphy, M. J., and L. M. Siegel. 1973. Siroheme and sirohydrochlorin. The basis for a new type of porphyrin-related prosthetic group common to both assimilatory and dissimilatory sulfite reductase. J. Biol. Chem. **248:**6911-6919.

Nakayashiki, T., K. Nishimura, and H. Inokuchi. 1995. Cloning and sequencing of a previously unidentified gene that is involved in the biosynthesis of heme in *Escherichia coli*. Gene. **153:**67-70.

Nishimura, K., T. Nakayashiki, and H. Inokuchi. 1993. Cloning and sequencing of the *hemE* gene encoding uroporphyrinogen III decarboxylase (UPD) from *Escherichia coli* K12. Gene. **133:**109-113.

O'Neill, G. P., M. W. Chen, and D. Soll. 1989. δ-Aminolevulinic acid biosynthesis in *Escherichia coli* and *Bacillus subtilis* involves formation of glutamyl-tRNA. FEMS Microbiol Lett **60:**255-260.

Parkinson, J. S., and E. C. Kofoid. 1992. Communication modules in bacterial signaling proteins. Ann. Rev. Genet. **26:**71-112.

Petricek, M., L. Rutberg, I Schoder, and L. Hederstedt. 1990. Cloning and characterization of the *hemA* region of the *Bacillus subtillis* chromosome. J. Bacteriol. **172:**2250-2258.

Phillipp-Dormston, W. K., and M. Doss. 1975. Over-expression of porphyrins and heme in heterophic bacteria. Z. Naturforsch. **30c:**425-426.

Pontoppidan, B., and C. G. Kannangara. 1994. Purification and partial characterisation of barley glutamyl-tRNA reductase, the enzyme that directs glutamate to chlorophyll biosynthesis. Eur. J. Biochem. **225:**529-537.

Rabin, R. S., L. A. Collins, and V. Stewart. 1992. *In vivo* requirement of integration host factor for *nar* (nitrate reductase) operon expression in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA. **89:**8701-8705.

Rasmussen, P. B., B. Holst, and P. Valentin-Hansen. 1996. Dual-function regulators: the cAMP receptor protein and the CytR regulator can act either to repress or to activate transcription depending on the context. Proc Natl. Acad. Sci. **93:**10151-10155.

Reinbothe, S., and C. Reinbothe. 1996. The regulation of enzymes involved in chlorophyll biosynthesis. Eur. J. Biochem. **237:**323-343.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Rieble, S., and S. I. Beale. 1991. Purification of glutamyl-tRNA reductase from *Synechocystis* sp. PCC 6803. J. Biol. Chem. **266:**9740-9745.

Rivera, M., and F. A. Walker. 1995. Biosynthetic preparation of isotopically labeled heme. Anal. Biochem. **230:**295-302.

Saier, Jr, M. H. 1996. Cyclic AMP-independent catabolite repression in bacteria. FEMS Lett. **138:**97-103.

Sambrook. J., E. F. Frisch, and T. Maniatis. 1989. Molecular Cloning, 2d ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sasarman, A., M. Suredeanu, and T. Horodniceanu. 1968. Locus determining the synthesis of σ-aminolevulinic acid in *Escherichia coli* K12. J. Bacteriol. **96:**1882-1884.

Sasarman, A., A. Nepveu, Y. Echelard, J. Dmetryszn, M. Drolet, and C. Goyer. 1987. Molecular cloning and sequencing of the *hemD* gene of *Escherichia coli* and preliminary data on the *uro* operon. J. Bacteriol. **169:**4257-4262.

Schroder, I., S. Darie, and R. P.Gunsalus. 1993. Activation of the *Escherichia coli* nitrate reductase (*narGHJI*) operon by NarL and Fnr requires integration host factor. J. Biol. Chem. **268:**771-774.

Schroder, I., P. Johansson, and L. Hederstedt. 1994. The *hemX* gene of the *Bacillus subtilis hemAXCDBL* operon encodes a membrane protein, negatively affecting the steady-state cellular concentration of *hemA* (glutamyl-tRNA reductase). Microbiol. **140:**731-740.

Shaw, D. J., D. W. Rice, and J. R. Guest. 1983. Homology between CAP and Fnr, a regulator of anaerobic respiration in *Escherichia coli.* J. Mol. Biol. **166:**241-247.

Spiro, S., and J. R. Guest. 1990. FNR and its role in oxygen-regulated gene expression in *Escherichia coli.* FEMS Microbiol. Rev. **75:**399-428.

Stewart, V. 1982. Requirement of Fnr and NarL functions for nitrate reductase expression in *Escherichia coli* K-12. J. Bacteriol. **151:**1320-1325.

Stock, J. B., A. J. Ninfa, and A. M. Stock. 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. **53:**450-490.

Thomas, S. D., and P. M. Jordan. 1986. Nucleotide sequence of the *hemC* locus encoding porphobilinogen deaminase of *Escherichia coli* K12. Nucleic Acids Res.**14:**6215-6226.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Tragester, M., S. Spiro, A. Duchene, E. Kojro, F. Fahrenholz, J. R. Guest, and G. Unden. 1990. Isolation of intact FNR protein (Mr 30 000) of *Escherichia coli*. Mol. Microbiol. **4(1):**21-27.

Troup, B., M. Jahn, C. Hungerer, and D. Jahn. 1994. Isolation of the *hemF* operon containing the gene for the *Escherichia coli* aerobic coproporphyrinogen III oxidase by *in vivo* complementation of a yeast *HEM13* mutant. J. Bacteriol. **176:**673-680.

Troup, B., C. Hungerer, and D. Jahn. 1995. Cloning and characterization of the *Escherichia coli hemN* gene encoding the oxygen-independent coproporphyrinogen III oxidase. J. Bacteriol. 1995. **177:**3326-3331

Tyson, K. L., A. I. Bell, J. A. Cole, and J. W. Busby. 1993. Definition of nitrite and nitrate response elements at the anaerobically inducible *Escherichia coli nirB* promoter: interactions between FNR and NarL. Mol. Microbiol. **7:**151-157.

Umanoff, H., C. S. Russell, and S. D. Cosloy. 1988. Availability of porphobilinogen control appearance of porphobilinogen deaminase activity in *Escherichia coli* K-12. J. Bacteriol. **170:**4969-4971.

Unden, G., M. Tragester, and A. Duchene. 1990. Effect of positive redox potentials (>+400 mV) on the expession of anaerobic respiratory enzymes in *Escherichia coli*. Mol. Microbiol. **4:**315-319.

Urban-Grimal, D., V. Ribes, and R. Labbe-Bois. 1984. Cloning by genetic complementation and restriction mapping of the yeast HEM1 gene coding for 5-aminolevulinate synthase. Curr. Genet. **8:**327-331.

Valentin-Hansen, P., B. Holst, L. Sogaard-Andersen, J. Matinussen, J. Nesvera, and S. R. Douthwaite. 1991. Design of cAMP-CRP-activated promoters in *Escherichia coli*. Mol. Microbiol. **5(2):**433-437

Verkamp, E., and B. K. Chelm. 1989. Isolation, nucleotide sequence, and preliminary characterization of the *Escherichia coli* K-12 *hemA* gene. J. Bacteriol. **171:**4728-4735.

Verkamp, E. M., D. Jahn, A. Jahn, M. Kumar, and D. Soll. 1992. Glutamyl-tRNA reductase from *Escherichia coli* and *Synechosystis* 6803. J. Biol. Chem. **267:**8275-8280.

Weber, R. F., and P. M. Silver. 1988. The Cpx proteins of *Escherichia coli* K12. Structure of the CpxA polypeptide as an inner membrane component. J. Mol. Biol. **203:**467-478.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Williams, R., A. Bell, G. Sims, and S. Busby. 1991. The role of two surface exposed loops in transcription activation by the *Escherichia coli* CRP and FNR proteins. Nucl. Acids. Res. **19:**6705-6712.

Woodard, S. I., and H. A. Dailey. 1995. Regulation of heme biosynthesis in *Escherichia coli.* Arch. Biochem. Biophys. **316:**110-115.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.