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.

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.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

STRUCTURAL AND FUNCTIONAL ANALYSIS OF GENES INVOLVED IN THE HEME BIOSYNTHETIC PATHWAY OF *SACCHAROMYCES CEREVISIAE*

by

Celestino Di Flumeri
Department of Microbiology and Immunology
McGill University, Montreal
August, 1997

A Thesis Submitted to the Faculty of Graduate Studies and Research in Partial Fulfillment of the Requirements of the Degree of Doctor of Philosophy

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

0-612-44409-0

Canada

ABSTRACT

The heme biosynthetic pathway in *Saccharomyces cerevisiae* comprises eight enzymatic steps. The *HEM6* gene encodes uroporphyrinogen decarboxylase, a cytoplasmic enzyme catalyzing the fifth step of heme biosynthesis. *HEM6* was cloned by complementation of a heme auxotrophic yeast strain containing a *hem6* mutant allele. An open reading frame of 1086 nucleotides encoding a protein of 362 amino acids was obtained by sequence analysis. Expression of *HEM6* was found to be induced two-fold by lactate, a non-fermentable carbon source, but not regulated by heme. Site-directed mutagenesis of a conserved cysteine residue implicated in catalysis demonstrated that cysteine 52 of the yeast enzyme is not essential for enzymatic activity.

The enzyme catalyzing the sixth step in the heme biosynthetic pathway, coproporphyrinogen oxidase, is encoded by the *HEM13* gene. Transcription of *HEM13* is regulated by the product of the pathway, heme. In the presence of heme, transcription of the *ROX1* gene is induced, and the ROX1 protein represses *HEM13* transcription. ROX1 is a member of the High Mobility Group (HMG) family of DNA-binding proteins. In order to define the mechanism of repression of *HEM13* by ROX1, we synthesized ROX1 protein and derivatives either *in vitro* or in *Escherichia coli* as fusions to the glutathione-S-transferase (GST) protein. All ROX1 derivatives containing intact HMG domains were capable of specific binding to DNA sequences from the *HEM13* promoter region. In contrast, ROX1 proteins which contained deletions within the HMG domain were no longer capable of binding to DNA. In addition, ROX1 was capable of oligomerization, and the amino-terminal 100 amino acids of ROX1 which constitutes the HMG domain was required for oligomerization as well as for DNA-binding.

The *HEM13* promoter contains five possible operator consensus sequences postulated to be binding sites for ROX1. We showed that ROX1 binds to all five sites with differential affinity and that this binding occurs primarily via minor groove contacts. *In vivo* repression was examined by deletion of the five operator sites, either individually or in combination. Three of the operator sites were shown to act in an additive manner in bringing about repression of *HEM13* and a fourth site appeared to be active only when all other operator sites were intact. A random PCR-based selection procedure used to select ROX1 binding sites predicted the optimal ROX1 binding site to consist of the consensus sequence (A/T)TT(T/G)TT.

Finally, a preliminary analysis of the ROX1 protein suggested that the region of ROX1 responsible for repression is located within the carboxy-terminus of the protein. This region of ROX1 may be required for interaction with the general repressor complex SSN6/TUP1. Chromosomal disruption of either of these two genes resulted in derepression of *HEM13* expression in a *ROX1* wild type strain. Therefore, repression of *HEM13* is brought about by multiple regulatory factors interacting at various *cis*-acting elements within the *HEM13* non-coding sequences.

RESUMÉ

La biosynthèse de l'hème chez *Saccharomyces cerevisiae* est composée de huit étapes enzymatiques. Le gène *HEM6* code pour l'uroporphyrinogène décarboxylase, un enzyme cytoplasmique qui catalyse la cinquième étape de la biosythèse de l'hème. Le gène *HEM6* a été cloné par complémentation d'une souche auxotrophe contenant une allèle mutante de *hem6*. Une séquence contenant 1086 nucleotides codant pour une protéine de 362 acides aminés a été obtenue par l'analyse de ce gène. L'expression de *HEM6* a doublé en presence d'acide lactique, un source de carbone non fermentable, mais l'expression de ce gène n'était pas controlée par l'hème. La mutagénèse dirigée d'une cystéine conservée (Cys 52) impliquée dans la catalyse, a demontré que cette cystéine n'est pas essentielle pour l'activité enzymatique chez la levure.

L'enzyme catalysant la sixième étape dans la cascade biosynthétique de l'hème, la coproporphyrinogène oxidase, est codée par le gène *HEM13*. La transcription de ce dernier est controlée par le produit de la cascade, l'hème. En presence d'hème, la transcription du gène *ROX1* est induite et le protèine ROX1 réprime la transcription de *HEM13*. ROX1 est un membre de la famille de protéines High Mobility Group (HMG) qui se lient à l'ADN. Pour definir le mécanisme de répression de *HEM13* par ROX1, nous avons synthétisé la protéine ROX1 *in vitro* ou dans l'*Escherichia coli* joint à la protéine glutathion-S-transférase (GST). Tous les dérivés de ROX1 contenant un domaine HMG intact ont été capables de se lier au promoteur de *HEM13* avec spécificité. Au contraire, les protéines ROX1 contenant des délétions dans le domaine HMG n'étaient plus capables de se lier à l'ADN. En plus, ROX1 a été capable de s'oligomériser, et la région nécessaire à l'oligomérisation, ainsi que celle responsable

pour la liaison à l'ADN, se trouve dans les 100 acides aminés du coté N-terminal, qui constituent le domaine HMG.

Le promoteur du gène *HEM13* contient cinq séquences opérateur conservées que nous pensions être des sites liées par ROX1. Nous avons demontré que ROX1 contacte chacune de ces cinq sites avec une affinité différente et que ces contacts se font principalement dans le sillon mineur de l'ADN. La répression a été étudiée *in vivo* par délétion des cinq sites opérateurs, individuellement ou en combinaisons. Nous avons demontré que trois des cinq sites fonctionnent de manière additive pour réprimer l'expression de *HEM13*; un quatrième site ne fonctionne qu'en présence des trois autres sites. Une méthode utilisant une sélection aléatoire à l'aide de PCR a été utilisée afin de sélectionner des séquences de liaison de ROX1. Les résultats demontrent que la séquence consensus est (A/T)TT(T/G)TT.

Finalement, une analyse préliminaire de ROX1 a indiqué que la région responsable de la répression est située dans la partie carboxy-terminale de ce protéine. Cette région est peut-être nécessaire pour l'intéraction avec le complexe de répression générale SSN6/TUP1. Une insertion dans la copie chromosomique inactivant chaqu'un de ces deux gènes a résulté en une dérepression de l'expression de *HEM13* dans une souche contenant une copie de type sauvage de *ROX1*. Donc, la répression de *HEM13* est causée par de multiples facteurs régulatoires intéragissant avec des éléments situés dans les séquences non-codantes de *HEM13*.

ACKNOWLEDGMENTS

I wish to acknowledge the F.C. Harrison Fellowship Fund, the Alexander McFee Fellowship Fund, the Fonds pour la Formation de Chercheurs et L'Aide à la Recherche (FCAR), and the Natural Sciences and Engineering Research Council (NSERC) of Canada for providing generous financial support.

To Dr. Teresa Keng, I would like to sincerely thank her for the opportunity to work under her supervision. I thank her for her guidance and willingness to help at all times. I acknowledge Dr. Nicholas Acheson for giving me the opportunity to complete my studies under his supervision. I thank him for his insight and valuable advice.

I wish to thank Dr. Dalius Briedis for allowing me to use his lab as my home away from home and Marie-Claude Ouimet for help with translation of the French . I wish to thank many special people for their friendship; Soula Chronopoulos, Francisco Ferreira, Richard McLaughlin, Tina Martino, Marie-Claude Ouimet, Dr. David Scott, Ian Siboo, and Felix Sieder.

To all former and present members of the laboratory; Rami Batal, Dr. John Bertin, François Brabant, Dr. Edward Cherng, Robert Larocque, Yu-Cai Peng, Catherine Richard, Dr. Sophia Ushinsky, and Sandy Weber, thank you for the memories.

I wish to express to Angela De Ciccio a special thank you for a friendship that I cherish dearly. She had the wisdom and innate sense to comprehend what true friendship requires; honesty and trust. I thank her for willingly being the object of our laughter; a characteristic which only enhanced the years we spent together.

I wish to express my sincerest gratitude to Dr. Peter Liston for enriching me with his knowledge and instilling confidence in my scientific ability. I am indebted to him for much of the scientific progress I made during the years we spent together.

I thank Alfredo Staffa for always being genuine and good-natured. His willingness to give freely of himself to forward my cause will not be forgotten. I feel privileged to have spent the last few years together and acquired many fond memories. Thanks for making this experience more tolerable than it otherwise would have been.

I thank my parents, Maddalena and Carmine, for their inspiration and for sacrifices they made along the way. I could not have had two better role models. To my brother Jimmy and sister Emma, thanks for your encouragement and love all the way through to the end. I wish to thank all members of my extended family for their generosity and good wishes.

I thank my wife Adriana for the sacrifices she made in order that I could achieve this goal. Her love, support, and encouragement were of monumental importance during all these years. I thank her for the ability to make me laugh, a quality which was much appreciated during some of the hardships of the last few years. Much of the credit for this accomplishment is rightfully hers as it could not have been done without her.

OBJECTIVES OF STUDY

This research project involved the cloning and characterization of the *HEM6* gene of *Saccharomyces cerevisiae* and an examination of the structure and function of the ROX1 protein as it pertains to its role in regulating *HEM13* expression.

Heme is an essential molecule required by all cell types for diverse functions. In mammalian cells, heme is an integral component of hemoglobin and myoglobin, and thus plays a role in oxygen transport and storage. In addition, heme is also used as a prosthetic group for many enzymes involved in respiration and in removal of toxic oxygen radicals. In *Saccharomyces cerevisiae*, heme is also involved in transcriptional regulation.

We proposed to isolate *HEM6*, the gene encoding the enzyme catalyzing the fifth step of the heme biosynthetic pathway, uroporphyrinogen decarboxylase, by complementation of a yeast mutant defective in the gene encoding this enzyme. The promoter region of the gene was then fused to the β-galactosidase gene of *E.coli* in order to study its regulation by heme and carbon source. In addition, Northern analysis was employed to examine *HEM6* RNA levels in cells grown under various conditions. Various studies utilizing sulphydryl-specific reagents have implicated cysteine residues as being important for the function of uroporphyrinogen decarboxylase. We aimed to examine the role of the conserved cysteine residue by site-directed mutagenesis. Plasmids with a *HEM6* gene containing these mutations were tested for their ability to complement a *hem6* mutant strain 1B. In addition, the transformed strains were

examined for phenotypes associated with *hem6* mutants, such as inability to grow on non-fermentable carbon sources and accumulation of porphyrins.

HEM13 encodes coproporphyrinogen oxidase, a cytoplasmic enzyme catalyzing the sixth step of heme biosynthesis. The expression of *HEM13* is negatively regulated by heme and oxygen at the transcriptional level. This repression requires the product of the *ROX1* gene, a DNA-binding protein containing a High Mobility Group (HMG) motif that is only synthesized in the presence of heme. ROX1 protein is postulated to repress *HEM13* transcription by binding to sequences located upstream of the *HEM13* transcription start site. This repression may require the products of the *TUP1* and *SSN6* genes, which are general repressors of pathway specific genes in yeast.

We aimed to identify the sites required for *HEM13* repression within the promoter region of this gene. These sites were used as probes for gel retardation assays using ROX1 proteins synthesized *in vitro* or in *E. coli*. Deletion derivatives of ROX1 were tested for their ability to bind to *HEM13* specific probes in order to delineate the region of ROX1 responsible for DNA-binding. In addition, the region of ROX1 responsible for oligomerization was delineated by using various GST-ROX1 derivatives and ^{35}S-labelled *in vitro* translated ROX1. The repression domain of ROX1 was delineated by transforming various ROX1 deletion plasmids into a *rox1* mutant strain containing an integrated copy of *HEM13-lacZ*. Levels of *HEM13-lacZ* expression were examined in strains with *ssn6* and *tup1* chromosomal disruptions to assess the role of TUP1 and SSN6 proteins in repression of *HEM13*.

CLAIM OF CONTRIBUTIONS TO KNOWLEDGE

1. *HEM6* was sequenced and shown to encode a protein of 362 amino acids displaying extensive homology to other mammalian uroporphyrinogen decarboxylases.

2. Northern analysis for *HEM6* mRNA and ß-galactosidase assays of strains containing a *HEM6-lacZ* fusion demonstrated a two-fold induction in *HEM6* expression by lactate, and no detectable induction by heme.

3. *HEM6* expression was not regulated by the transcriptional activator complex HAP2/3/4/5, unlike other genes encoding heme biosynthetic enzymes.

4. Site-directed mutagenesis experiments revealed the non-essential nature of the conserved cysteine 52 of the yeast uroporphyrinogen decarboxylase.

5. Bacterially expressed and *in vitro* translated ROX1 proteins were capable of specific DNA-binding to *HEM13* target sequences.

6. Deletion derivatives of ROX1 containing all of the HMG domain were capable of specific DNA-binding to *HEM13* target sequences.

7. Oligomerization of ROX1 protein *in vitro* was demonstrated using affinity-purified ROX1 protein and ROX1 labelled with ^{35}S-methionine.

8. Delineation of the oligomerization domain identified an overlap with the DNA-binding domain within the amino-terminal 100 amino acids of ROX1.

9. Studies employing minor groove specific drugs indicated a preference for binding of ROX1 within the minor groove of the DNA helix.

10. Oligonucleotides containing inosine substitutions employed in gel retardation assays demonstrated a specific preference for binding of ROX1 to the minor groove of the DNA helix with contacts occurring primarily with adenine-thymine base pairs.

11. ROX1 shows differential binding affinity for the five consensus operator sites upstream of the *HEM13* transcription start site.

12. A random PCR-based selection protocol predicted the consensus DNA-binding site for ROX1 to be (A/T)TT(T/G)TT.

13. The TUP1 and SSN6 proteins were shown to be required for ROX1-mediated repression of *HEM13*.

14. The region of ROX1 required for repression was localized to the carboxyl-terminus of the protein.

GUIDELINES REGARDING DOCTORAL THESIS

MANUSCRIPTS AND AUTHORSHIP

Candidates have the option of including, as part of the thesis, the text of one or more papers submitted or to be submitted for publication, or the clearly-duplicated text of one or more published papers. These texts must be bound as an integral part of the thesis.

If this option is chosen, **connecting texts that provide logical bridges between the different papers are mandatory.** *The thesis must be written in such a way that it is more than a mere collection of manuscripts; in other words, results of a series of papers must be integrated.*

The thesis must still conform to all other requirements of the "Guidelines for Thesis Preparation". **The thesis must include:** *a Table of Contents, an abstract in English and French, an introduction which clearly states the rationale and objectives of the study, a review of the literature, a final conclusion and summary, and a thorough bibliography or reference list.*

Additional material must be provided where appropriate (e.g. in appendices) and in sufficient detail to allow a clear and precise judgement to be made of the importance and originality of the research reported in the thesis.

In the case of manuscripts co-authored by the candidate and others, **the candidate is required to make an explicit statement in the thesis as to who contributed to such work and to what extent.** *Supervisors must attest to the accuracy of such statements at the doctoral oral defense. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to make perfectly clear the responsibilities of all the authors of the co-authored papers.*

* * *

In accordance with the above guidelines, I hearby state that:

Chapter 3 of this thesis is the text of a published manuscript entitled "Molecular Analysis of *HEM6* (*HEM12*) in *Saccharomyces cerevisiae*, the Gene for Uroporphyrinogen Decarboxylase", by Celestino Di Flumeri, Robert Larocque, and Teresa Keng. This manuscript was published in *Yeast* **9:** 613-623 (1993). Robert Larocque was responsible for construction and testing for complementation of the *hem6* mutant strain by subclones of *HEM6*. In addition, Robert Larocque was involved with initial phases of the *HEM6* sequencing effort. The majority of the *HEM6* sequencing and all other experiments were designed in consultation with Dr. Keng and performed by myself. The manuscript was written by myself, and edited by Dr. Keng.

Chapter 4 of this thesis is the text of a published manuscript entitled "A Conserved Cysteine Residue in Yeast Uroporphyrinogen Decarboxylase is not Essential for Enzymatic Activity", by Celestino Di Flumeri, Nicholas H. Acheson, and Teresa Keng. This manuscript was published in *Canadian Journal of Microbiology* **43**: 792-795.

All experiments were designed by myself in consultation with Dr. Keng, and performed by myself. The manuscript was written by myself, and edited by Drs. Acheson and Keng.

Chapter 5 of this thesis is the text of a published manuscript entitled "The HMG Domain of the ROX1 Protein Mediates Repression of *HEM13* Through Overlapping DNA-Binding and Oligomerization Functions", by Celestino Di Flumeri, Peter Liston, Nicholas H. Acheson, and Teresa Keng. This manuscript was published in *Nucleic Acids Research* **24:** 808-815 (1996). Peter Liston was responsible for construction of plasmid pGEX-ROX' 176-368 and for valuable suggestions. All other experiments were conducted by myself in consultation with Drs. Acheson and Keng. The manuscript was written by myself, and edited by Drs. Acheson and Keng.

Chapter 6 of the thesis represents a manuscript which is in preparation for submission entitled "ROX1 Represses Transcription of *HEM13* by Binding to Multiple Upstream Operator Sites with Differential Affinity". Alfredo Staffa was reponsible for design of the random selection procedure. Sophia Ushinsky was responsible for plasmid constructions required for *HEM13* deletion analysis as well as valuable advice. Teresa Keng was responsible for most of the plasmid constructions and all experiments concerning figure 20 of the manuscript. All other experiments were designed and performed by myself in consultation with Drs. Keng and Acheson. The manuscript was written by myself and edited by Drs. Acheson and Keng.

Chapter 7 contains data which are unpublished. All experiments were designed and performed by myself in consultation with Drs. Keng and Acheson.

TABLE OF CONTENTS

Page

Abstract ..... **i**
Resumé ..... **iii**
Acknowledgments ..... **v**
Objectives of study ..... **vii**
Claim of contribution to knowledge ..... **ix**
Guidelines regarding doctoral thesis ..... **xi**
Table of contents ..... **xiv**
List of figures ..... **xix**
List of tables ..... **xxi**

CHAPTER 1. INTRODUCTION **1**
I. Heme biosynthetic pathway **1**
1. δ-Aminolevulinate synthase 3
2. δ-Aminolevulinate dehydratase 7
3. Porphobilinogen (PBG) deaminase 8
4. *Uroporphyrinogen III synthase* 9
5. *Uroporphyrinogen decarboxylase* 9
6. *Coproporphyrinogen oxidase* 11
7. *Protoporphyrinogen oxidase* 14
8. Ferrochelatase 16

II. Regulatory functions of heme **17**

III. Heme-dependent transcription factors **22**
1. HAP1 (CYP1) protein 22
A) Structure of HAP1 23
B) HAP1 DNA-binding characteristics 26
2. HAP2/3/4/5 complex 29
A) Functions of the different subunits 30

i) HAP2 30
ii) HAP3 30
iii) HAP4 31
iv) HAP5 31
B) Regulation of HAP2, HAP3, HAP4, and HAP5 31
C) Genes under HAP2/3/4/5 control 32

IV. Heme-repressed genes **33**

V. Heme-regulated repression factors **35**
1. ROX1 35
A) Cloning of ROX1 35
B) Regulation of ROX1 expression 36
C) ROX1 binding site 40
D) Structure of ROX1 protein 45
i) DNA binding domain 45
ii) Glutamine-rich domain 49
iii) Repression domain 49

VI. Repression by ROX1 requires the general repressors SSN6 and TUP1 ..... **50**
1) SSN6 (CYC8) 50
2) TUP1 51
3) TUP1 and SSN6 function as a complex 52
4) Role of SSN6/TUP1 in repression by ROX1 54

VII. Regulation of *HEM13* expression. **55**
1) Involvement of heme in HEM13 expression. 56
2) Function of ROX1 in repression of HEM13. 57
3) Function of HAP1 in HEM13 expression. 58
4) No role of HAP2/3/4/5 complex in HEM13 expression. 59
5) HAP1 regulates ROX1 expression. 59
6) Function of TUP1 and SSN6 in repression of HEM13. 60
7) HEM13 cis-acting sequences 61

A) Activation sequences 63
B) Negative sites 65
8) Trans-acting factors binding to the HEM13 UAS 66
9 Trans-acting factors involved in HEM13 repression via hypoxic operator sites 67

VIII. Focus of the present study **68**

CHAPTER 2. MATERIALS AND METHODS **69**
I. Cell cultures **69**
1. Strains **69**
2. Media and growth conditions **70**
3. Yeast and bacterial transformation **71**
4. Characterization of heme auxotrophs **72**

II. Nucleic acids **73**
1. Extraction and enzymatic manipulations of DNA **73**
2. Plasmid constructs **73**
A. lacZ fusion constructs 73
B. HEM6 site-directed mutant constructs 74
C. In vitro transcription constructs 75
D. Disruption constructs 76
E. GST-ROX1 fusion constructs 76
F. Operator site deletion constructs 77
G. In vivo ROX1 deletion constructs 79
3. DNA sequencing and analysis **80**
4. Northern blot analysis **81**
5. *In vitro* transcription and translation **81**

III. Protein analysis **82**
1. Expression and purification of GST fusion proteins **82**
2. Assay for protein-protein interaction **83**

IV. Electrophoretic mobility shift assays **84**

1. Gel electrophoresis conditions and preparation of probes 84
2. Delineation of DNA-binding domain of ROX1 84
3. Operator site binding affinity 85
4. Minor groove binding drugs 86
5. Oligonucleotide substitution 86

V. Random PCR-based selection of optimal ROX1 binding sites 87

VI. Enzymatic assays 89
1. β-galactosidase assays 89

CHAPTER 3. Molecular analysis of *HEM6* (*HEM12*) in *Saccharomyces cerevisiae*, the gene for uroporphyrinogen decarboxylase 90
Abstract 91
Introduction 92
Results 95
Isolation of HEM6 95
Nucleotide sequence of HEM6 97
Expression of HEM6 100
Discussion 106

CHAPTER 4. A conserved cysteine residue in yeast uroporphyrinogen decarboxylase is not essential for enzymatic activity 111
Abstract 112
Introduction 113
Results 115
Discussion 118

CHAPTER 5. The HMG domain of the ROX1 protein mediates repression of *HEM13* through overlapping DNA-binding and oligomerization functions 120
Abstract 121
Introduction 122
Results 125

Synthesis of ROX1 protein in vitro 125
Specific DNA binding by ROX1 protein synthesized in vitro 125
Delineation of the DNA-binding domain of ROX1 127
Expression of GST-ROX1 fusion proteins in E. coli 130
DNA binding by GST-ROX1 fusion proteins 132
ROX1 protein-protein interactions 133
Discussion **139**

CHAPTER 6. The *Saccharomyces cerevisiae* ROX1 protein represses transcription of *HEM13* by binding to multiple upstream sites with differential affinity **145**

Abstract **146**
Introduction **147**
Results **151**
Binding of GST-ROX1 1-101 protein to HEM13 sites 151
Functional analysis of hypoxic operator sites 154
ROX1 binds to HEM13 site 3 primarily through minor groove contacts 157
PCR-mediated selection of optimal ROX1 binding sites 161
Discussion **164**

CHAPTER 7. Repression of *HEM13* is mediated by the carboxy-terminal region of ROX1 and requires the SSN6/TUP1 complex **172**

Abstract **173**
Introduction **174**
Results **177**
Discussion **182**

CHAPTER 8. SUMMARY AND CONCLUSIONS **185**

LITERATURE CITED **201**

LIST OF FIGURES

		Page
Figure 1	Functions of heme in *Saccharomyces cerevisiae*	2
Figure 2	Heme biosynthetic pathway in *Saccharomyces cerevisiae*	4
Figure 3	Structure of the HAP1 protein	24
Figure 4	Structure of the ROX1 protein	46
Figure 5	Regulatory sequences of *HEM13*	62
Figure 6	Localization of *HEM6* within the plasmid pTK600	96
Figure 7	Sequencing strategy of the *HEM6* coding region and the flanking regions	98
Figure 8	Nucleotide sequence of *HEM6* and the translated amino acid sequence	99
Figure 9	Alignment of derived amino acid sequences for uroporphyrinogen decarboxylase from different sources	101
Figure 10	Northern blot analysis of *HEM6* RNA	103
Figure 11	*In vitro* translation of ROX1 protein	126
Figure 12	Binding of *in vitro* translated ROX1 to a DNA fragment containing a specific ROX1 binding site	128
Figure 13	Delineation of the DNA binding domain of ROX1	129
Figure 14	Expression and purification of GST-ROX1 fusion proteins	131
Figure 15	DNA binding by the ROX1 portion of the GST-ROX1 fusion protein	134
Figure 16	Determination of ROX1-ROX1 protein interactions	135

Figure 17 Delineation of the DNA binding domain and oligomerization domain of ROX1 137

Figure 18 Nucleotide sequence of the *HEM13* promoter 152

Figure 19 Affinity of ROX1 for different repeats upstream of *HEM13* transcription start site 153

Figure 20 Effect of deletion of hypoxic operator sites on *HEM13-lacZ* expression 156

Figure 21 Effect of minor groove binding reagents on binding of ROX1 1-101 to DNA 159

Figure 22 Effect of substitution of A-T base pairs in the ROX1 core consensus with I-C or G-C base pairs 160

Figure 23 Alignment of selected ROX1 binding sites 163

Figure 24 Model for heme-dependent repression of *HEM13* expression in the presence of oxygen 200

LIST OF TABLES **Page**

Table 1 Genes regulated by heme and oxygen 19

Table 2 Hypoxic operator sequences 43

Table 3 Regulation of *HEM6-lacZ* expression.....104

Table 4 Effects of site-directed mutagenesis of conserved cysteine residues.....116

Table 5 Effects of *SSN6* and *TUP1* mutations on *HEM13-lacZ* expression.....178

Table 6 Effects of carboxy-terminal deletions of ROX1 on *HEM13-lacZ* expression.....180

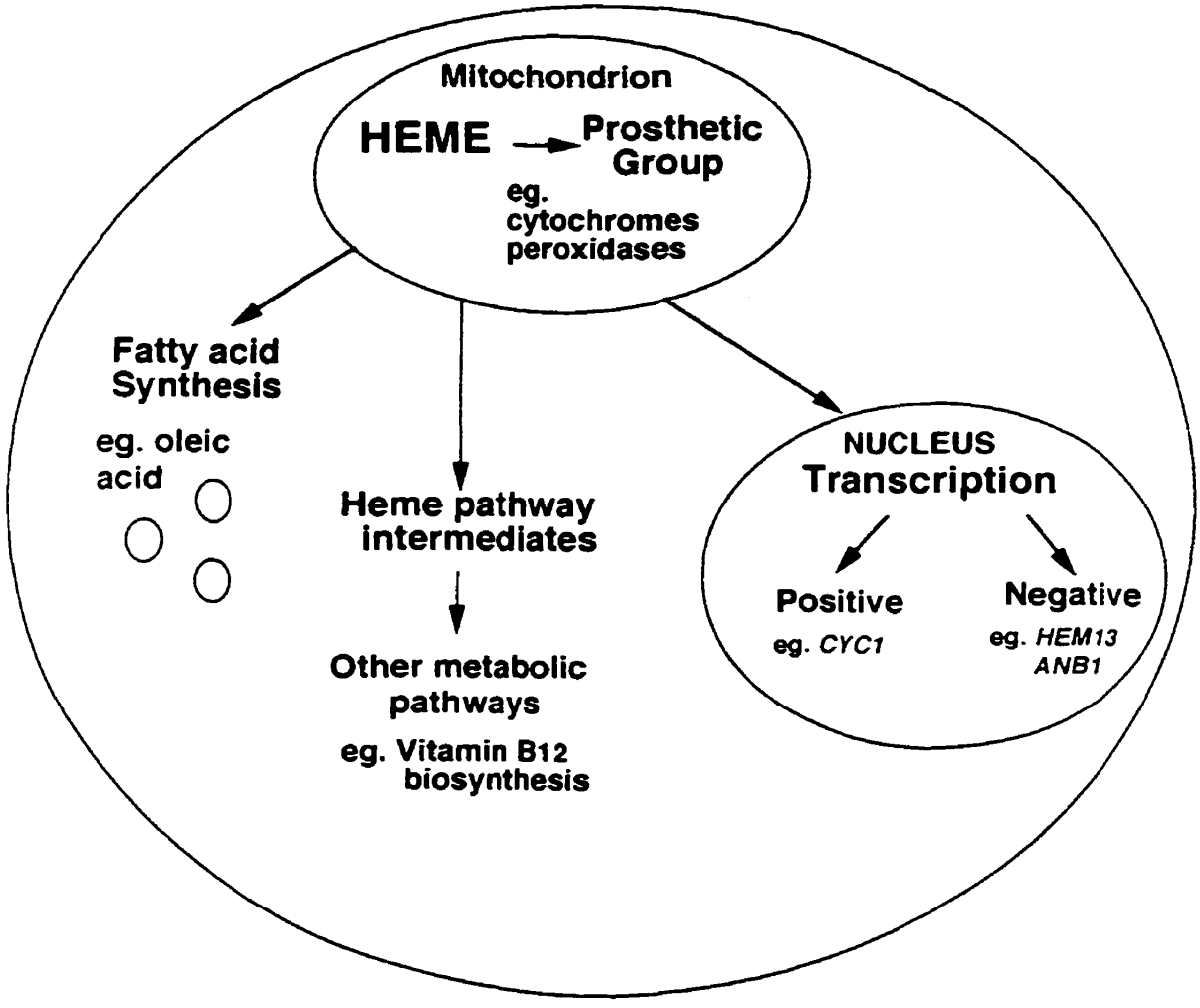
CHAPTER 1. INTRODUCTION

I. Heme biosynthetic pathway

Heme is an essential molecule which serves a multitude of cellular functions in many organisms ranging from bacteria to mammalian cells. It is a constituent of many hemoproteins found within different cellular compartments including mitochondria and the cytosol. These hemoproteins include hemoglobin and myoglobin which are required for oxygen transport and storage, respiratory cytochromes located within the mitochondria which make up the electron transport chain, catalases and peroxidases found within the peroxisomes which are involved in the detoxification of metabolic byproducts, and tryptophan pyrrolase which breaks down tryptophan. In addition, the heme biosynthetic pathway provides intermediates for the synthesis of other physiologically important molecules such as siroheme, which is required for the synthesis of methionine. In *Saccharomyces cerevisiae*, heme can also localize to the nucleus where it plays an important regulatory role in gene expression (Figure 1). Synthesis of heme is the result of a complex series of enzymatic steps occurring in both the cytosol and the mitochondrion. The mechanism by which heme and its precursors move across the various intracellular compartments is presently not defined, although recent identification of a candidate mitochondrial membrane protein may provide insight into this phenomenon (Taketani *et al.*, 1994).

In most procaryotes and eucaryotes, the heme biosynthetic pathway consists of eight well-conserved enzymatic steps that convert glycine and succinyl CoA to heme. The first and last three enzymatic steps in the pathway are usually catalyzed by enzymes

Figure 1. Functions of heme in *Saccharomyces cerevisiae*. Heme has a multitude of functions in the cell. It is synthesized in the mitochondrion where it acts as a prosthetic group for cytochromes and peroxidases. In the cytosol, heme is also involved in fatty acid and vitamin B_{12} biosynthesis. Although a heme carrier protein remains to be identified, heme can traverse the nuclear membrane and act as a transcriptional regulatory molecule.


Mitochondrion
HEME
Prosthetic Group
eg. cytochromes peroxidases
Fatty acid Synthesis
eg. oleic acid
Heme pathway intermediates
Other metabolic pathways
eg. Vitamin B12 biosynthesis
NUCLEUS
Transcription
Positive
eg. CYC1
Negative
eg. HEM13
ANB1

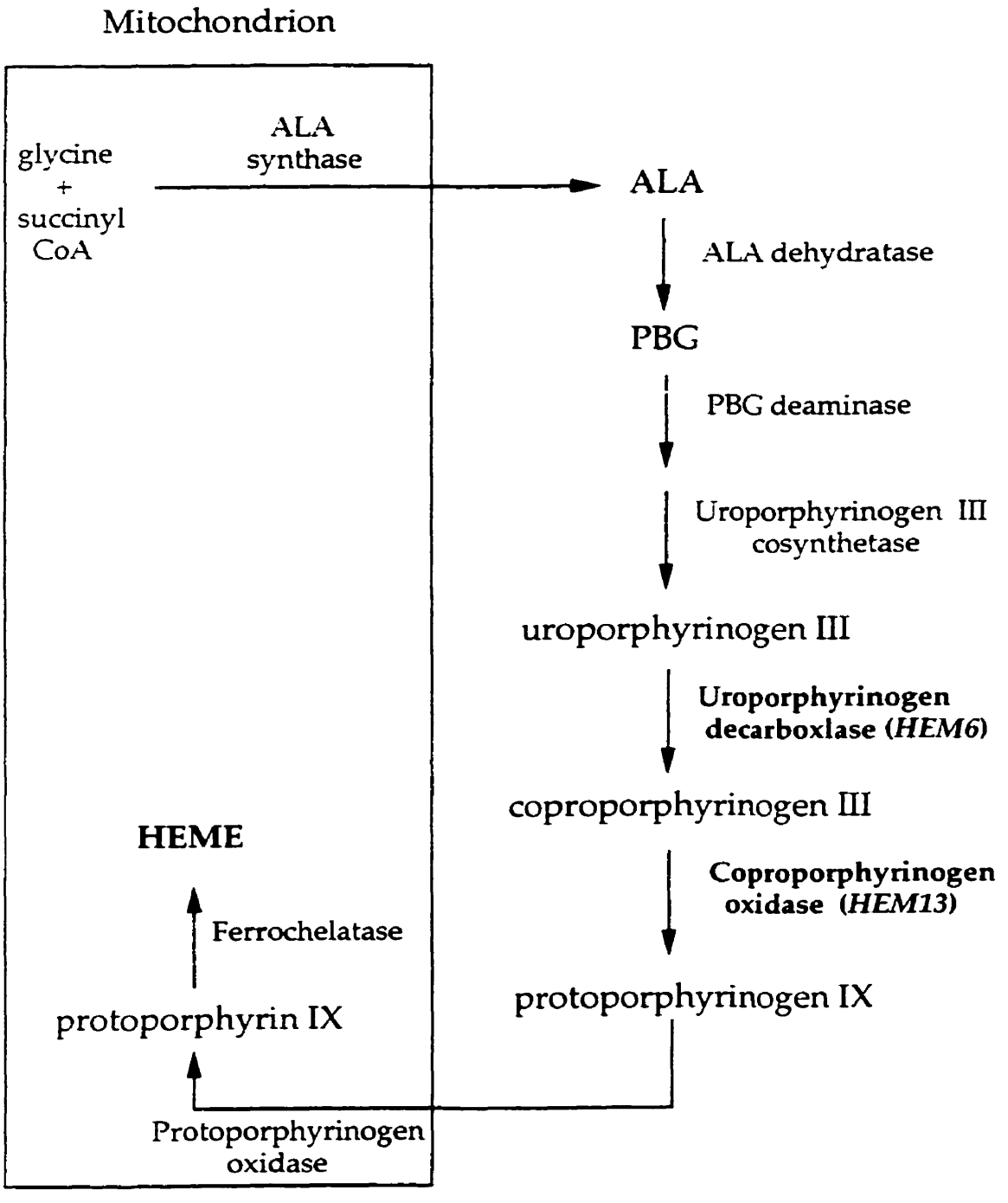
located within the mitochondrion, while the remaining four enzymes are located within the cytoplasmic compartment in higher eucaryotic organisms. In the yeast *Saccharomyces cerevisiae*, only the first and last two enzymes of the pathway are compartmentalized within the mitochondrion (Figure 2). The enzyme coproporphyrinogen oxidase is cytoplasmic in yeast (Camadro *et al.*, 1986), whereas in higher eucaryotes this enzyme is located on the outer surface of the inner mitochondrial membrane (Taketani *et al.*, 1994). Regardless of their ultimate intracellular location, enzymes studied to date are encoded by nuclear genes and synthesized in the cytosol. Subsequent to their synthesis, three of the yeast enzymes and four of the enzymes in higher eucaryotes are transported to the mitochondrion. An alternate pathway occurs in plants, algae, and some bacteria, such as *Escherichia coli*, all of which are devoid of the first enzyme of the pathway, δ-aminolevulinate synthase. In these organisms, δ-aminolevulinate is synthesized via the C-5 pathway from glutamate (Jahn *et al.*, 1992).

Isolation of yeast mutants defective in all of the eight steps in the heme biosynthetic pathway (Gollub *et al.*, 1977; Grimal and Labbe-Bois, 1980; Urban-Grimal and Labbe-Bois, 1981; Kurlandzka and Rytka, 1985; Keng *et al.*, 1992; Amillet and Labbe-Bois, 1995) has enabled the cloning of genes encoding these enzymes as well as a detailed analysis of the various mechanisms of regulation of these genes.

1) δ-Aminolevulinate (ALA) synthase

The first step in the biosynthesis of heme in *Saccharomyces cerevisiae* is catalyzed by δ-aminolevulinate synthase and involves the condensation of glycine and

Figure 2. Heme biosynthetic pathway in *Saccharomyces cerevisiae*. The biosynthesis of heme in yeast comprises eight enzymatic steps. The first and last two enzymatic steps occur in the mitochondrion and the remaining five enzymatic reactions are catalyzed by cytoplasmic enzymes. The enzymes highlighted in bold, uroporphyrinogen decarboxylase and coproporphyrinogen oxidase, are encoded by the *HEM6* and *HEM13* genes respectively, and comprise the main focus of this study.


Mitochondrion
glycine
+
succinyl
CoA
ALA
synthase
ALA
ALA dehydratase
PBG
PBG deaminase
Uroporphyrinogen III
cosynthetase
uroporphyrinogen III
Uroporphyrinogen
decarboxlase (HEM6)
coproporphyrinogen III
Coproporphyrinogen
oxidase (HEM13)
protoporphyrinogen IX
HEME
Ferrochelatase
protoporphyrin IX
Protoporphyrinogen
oxidase

succinyl CoA to form δ-ALA. The *HEM1* gene (Arrese *et al.*, 1983; Bard and Ingolia, 1984), which encodes this enzyme, has been cloned. The enzyme is synthesized as a precursor protein with a presequence of approximately 3 kiloDaltons (kDa). Within the presequence, the amino-terminal nine amino acids are sufficient to direct the protein to within the mitochondrial matrix (Keng *et al.*, 1986; Urban-Grimal *et al.*, 1986). The mature protein has a molecular weight of 53 kDa and is functional as a homodimer. Unlike the δ-ALA synthases from higher eucaryotes, yeast δ-ALA synthase is not feedback inhibited by heme, the end product of the pathway. In heme-deficient mutants, normal levels of δ-ALA synthase activity were observed, indicating that there is no correlation between enzyme activity and heme levels (Urban-Grimal and Labbe-Bois, 1981). Overexpression of *HEM1* in yeast resulted in an increase in δ-ALA synthase activity and increased intracellular levels of δ-ALA, yet the levels of heme remained constant, suggesting that formation of δ-ALA by δ-ALA synthase is not the rate-limiting step in heme biosynthesis (Arrese *et al.*, 1983).

Expression of the gene was examined using a *HEM1-lacZ* fusion and was shown to be essentially constitutive as a result of an intricate interplay between two different regulatory elements controlling its transcription (Keng and Guarente, 1987). *HEM1* expression is induced at most 2-fold in a non-fermentable carbon source such as lactate, and unaffected by growth in heme. However, transcription of *HEM1* is activated by the HAP2/3/4/5 complex which binds to the sequence TCATTGGT in the *HEM1* promoter. Activation by the HAP2/3/4/5 complex is normally increased in a non-fermentable carbon source, such as lactate, and is heme-dependent. However, in *HEM1* this activation

is counteracted by the presence of a negative regulatory element which represses transcription of *HEM1* during growth in lactate and other elements that make expression heme-independent.

Mammalian cells contain two genes coding for δ-ALA synthase, ALAS1 and ALAS2, which are found on different chromosomes, and are differentially regulated. ALAS1 encodes a housekeeping isozyme which is ubiquitously expressed and found primarily in the liver. The liver δ-ALA synthase provides heme for various cytochrome P_{450} proteins (May *et al.*, 1995). The level of this enzyme is feedback inhibited by the end product of the pathway, heme. This repression by heme occurs via multiple mechanisms: heme reduces the half-life of ALAS1 mRNA by reducing its stability. Heme also has been shown to inhibit import of the enzyme into the mitochondrion. This feedback inhibition of ALAS1 expression is consistent with the low requirement for heme by the liver. ALAS2 is expressed in erythrocytes and required for the production of hemoglobin. Regulation of ALAS2 occurs on several levels, with an integral mechanism involving translational regulation of ALAS2 mRNA. This is achieved through binding of an Iron Responsive Element (IRE) binding protein to the IRE element found within the 5' untranslated region of the ALAS2 mRNA (Ferreira and Gong, 1995). Binding of the IRE binding protein to the IRE prevents translation of ALAS2 mRNA when iron pools are low. Translation of ALAS2 mRNA takes place if the level of free intracellular iron is sufficient to modify the IRE binding protein. This mechanism ensures that intermediates of the pathway, including protoporphyrin, are not made in excess when the iron pools are low. Unlike ALAS1, ALAS2 does not seem to be feedback inhibited by heme.

2) δ-ALA dehydratase

The second step in the heme biosynthetic pathway in all organisms is catalyzed by a cytoplasmic enzyme, δ-ALA dehydratase, also known as porphobilinogen synthase. This enzyme catalyzes the condensation of two molecules of δ-ALA to form the pyrrole porphobilinogen with the subsequent loss of two molecules of water. In mammalian cells both the erythroid-specific and housekeeping isozyme are encoded by a single structural gene. This is accomplished by the use of different promoters and alternative splicing. The human enzyme has a predicted molecular mass of approximately 36 kDa and its predicted amino acid sequence shows a high degree of conservation with δ-ALA dehydratases from various species. There is a strict requirement for zinc as a cofactor for the enzyme (Tsukamoto *et al.*, 1979).

In yeast, the enzyme has been postulated to be the rate-limiting step of the heme biosynthetic pathway. The *HEM2* gene coding for δ-ALA dehydratase has been cloned and its open reading frame predicts a protein of 342 amino acids with a predicted molecular weight of 37, 837 Da. The amino acid sequence of the yeast enzyme shows 52% homology with that of the human enzyme, with highly conserved regions occurring in the region of the protein postulated to bind zinc (Myers *et al.*, 1987). The regulation of the *HEM2* gene has been examined and there seems to be no regulation by either carbon source or heme levels (Pinkham and Keng, 1994). Overexpression of the cloned *HEM2* gene will enable purification of the enzyme which should lead to a better understanding of the enzyme's catalytic properties.

3) Porphobilinogen (PBG) deaminase

This enzyme catalyzes the third step in the heme biosynthetic pathway, namely the condensation of four molecules of PBG to form preuroporphyrinogen, a linear hydroxymethylbilane, which then spontaneously cyclizes to form the symmetric cyclic tetrapyrrole uroporphyrinogen I. In mammalian cells, two porphobilinogen deaminase isozymes are synthesized from distinct mRNAs (Grandchamp *et al.*, 1987); one is expressed exclusively in erythroid cells, and the housekeeping isozyme is expressed in all other cells. The two porphobilinogen deaminase isozymes differ in their amino-terminal sequences as a result of differential splicing (Straka *et al.*, 1990). Expression of the two isozymes is differentially regulated from different promoters. The erythroid promoter bears homology to the ALAS promoter and contains sites for binding by the GATA-1 and erythroid-2 nuclear factor (NFE-2) (Ponka and Schulman, 1993).

In yeast, *HEM3*, the gene coding for PBG deaminase, has been cloned (Keng *et al.*, 1992) and characterized. *HEM3* encodes a protein with an open reading frame of 327 amino acids and a predicted molecular weight of 36,650 Da. Regulation of the gene has been examined, and it was found that its transcription is similar to that of *HEM1* in that it is unaffected by either carbon source or heme levels but is under control of the products of the *HAP2/3/4/5* genes. Expression of *HEM3-lacZ* was found to be reduced significantly in a *hap2* mutant strain. Inspection of the promoter region of *HEM3* revealed the presence of the sequence TTATTGGT, homologous to the known binding site for the HAP2/3/4/5 complex.

4) Uroporphyrinogen III synthase

Uroporphyrinogen III synthase, a cytoplasmic enzyme, catalyzes the fourth step of the pathway that culminates in the formation of heme, the conversion of uroporphyrinogen I to uroporphyrinogen III. Human cDNA clones encoding this enzyme have been isolated (Tsai *et al.*, 1988), and the DNA sequence predicts a protein of 263 amino acids with a molecular weight of approximately 28 kDa.

The reaction catalyzed by uroporphyrinogen III synthase is not well understood, but in *Saccharomyces cerevisiae,* it is believed that the enzyme does not catalyze the rate-determining step of the pathway. The gene for this enzyme has recently been cloned (Amillet and Labbe-Bois, 1995), and should provide the tools for purification of the enzyme and subsequent understanding of the catalytic properties of this enzyme. The predicted amino acid sequence of the yeast enzyme encoded by *HEM4* shows a low level of identity with the amino acid sequence of uroporphyrinogen III synthases from other sources (18-23%).

5) Uroporphyrinogen decarboxylase

Uroporphyrinogen decarboxylase is a cytosolic enzyme in both procaryotic and eucaryotic organisms. It catalyzes the fifth step of the heme biosynthetic pathway. This reaction involves the sequential decarboxylation of the four acetate side chains of uroporphyrinogen to yield coproporphyrinogen. The gene for this enzyme has been cloned from a variety of sources as diverse as bacteria and humans (Romeo *et al.*, 1986; Romana *et al.*, 1987; Sasarman *et al.*, 1987; Marczynski and Shapiro, 1992; Blattner *et al.*, 1993). The cDNA for human uroporphyrinogen decarboxylase shows a striking

similarity to the rat cDNA. The cDNAs from these two species show 85 and 90% identity at the DNA and amino acid levels, respectively. Only one species of mRNA for the human enzyme was found by Northern analysis of RNA derived from various tissues, suggesting that the same enzyme exists in all tissues. However, expression is substantially increased in erythroid tissues, although control elements in erythropoetic cells have not been identified (Romana *et al.*, 1987).

The gene for the yeast uroporphyrinogen decarboxylase has been cloned and encodes a protein of 362 amino acids with a predicted molecular weight of 41 kDa (Garey *et al.*, 1992). The yeast enzyme also shows a high degree of conservation with both rat and human decarboxylases. The amino acid identities between the yeast and human enzymes is 51%, whereas the identity of the yeast enzyme to the rat enzyme is 50%. The yeast enzyme has been purified to homogeneity and was shown experimentally to have a molecular weight of 38 kDa (Felix and Brouillet, 1990). No cofactor was required for catalytic activity of the enzyme, although, as has been shown for uroporphyrinogen decarboxylases from other sources, activity was sensitive to sulphydryl-specific reagents. These experiments predicted the involvement of one or more cysteines as integral components of the enzyme's active site. This assumption has been tested directly for the human enzyme (Wyckoff *et al.*, 1996) by creating site-directed mutants of various cysteines within the open reading frame of the cDNA coding for the enzyme. The mutant enzymes were expressed and purified and were found to have no decreases in enzyme activity. These experiments demonstrate quite conclusively that *in vitro*, cysteine residues are not by themselves essential for the reaction catalyzed by uroporphyrinogen decarboxylase.

6) Coproporphyrinogen oxidase

Coproporphyrinogen oxidase catalyzes the conversion of coproporphyrinogen III to protoporphyrinogen IX, the sixth step in the biosynthesis of heme. In eucaryotes examined to date, the reaction catalyzed by this enzyme must proceed in the presence of oxygen. However, both *E. coli* and *Salmonella typhimurium* have two genes encoding coproporphyrinogen oxidase: *hemF* which encodes an enzyme which functions under aerobic conditions, and *hemN* which encodes a coproporphyrinogen oxidase which functions under anaerobic conditions (Xu *et al.*, 1992; Troup *et al.*, 1995). The enzyme is inserted in the mitochondrial membrane in humans due to the presence of a leader peptide. Human cDNA clones for coproporphyrinogen oxidase have been isolated (Martasek *et al.*, 1994; Taketani *et al.*, 1994). Unlike the genes for δ-ALA dehydratase and porphobilinogen deaminase, which have different promoter elements and encode both erythroid and non-erythroid isozymes, the gene for coproporphyrinogen oxidase has a single coding region with multiple transcription initiation sites (Delfau-Larue *et al.*, 1994). This would suggest that the same enzyme is expressed in both erythroid cells as well as in other tissues, but that the regulation in the different cell types is diverse. The promoter of the gene for this enzyme contains elements usually found in promoters of genes which are expressed in erythroid cells, namely GATA-1 binding boxes, Sp1 sites, and CACCC boxes (Grandchamp *et al.*, 1995). Mouse coproporphyrinogen oxidase is transcriptionally induced during red blood cell differentiation in culture (Kohno *et al.*, 1993). This indicates that synthesis of this enzyme may be enhanced when there is increased demand for heme, such as during erythroid differentiation, when heme is

required for the synthesis of hemoglobin. Interestingly, this phenomenon is not restricted to mammalian cells. In photosynthetic organisms and plants, coproporphyrinogen oxidase activity is increased when heme levels are limiting (Madsen *et al.*, 1993; Hill and Merchant, 1995). Thus, the activity of coproporphyrinogen oxidase is upregulated under conditions in which heme is limiting, presumably in order for the enzyme to efficiently synthesize whatever level of heme that is required.

In yeast, coproporphyrinogen oxidase is encoded by the *HEM13* gene, and catalyzes the oxidation of coproporphyrinogen III to protoporphyrinogen IX. However, the yeast enzyme lacks a leader peptide and is therefore cytoplasmic in nature. The cytoplasmic nature of the yeast enzyme poses a problem regarding supply of the product of the enzyme reaction, protoporphyrinogen IX, to the next enzyme in the pathway which is inserted in the inner mitochondrial membrane. Therefore, protoporphyrinogen IX must cross the outer mitochondrial membrane to come into contact with protoporphyrinogen IX oxidase in the inner mitochondrial membrane. The enzyme may interact with the mitochondrial surface, presumably at contact sites between the two membranes. The activation of coproporphyrinogen oxidase by phospholipids and detergents makes this hypothesis plausible. The enzyme has been purified to homogeneity and functions as a homodimer with subunit molecular weight of 35 kDa (Camadro *et al.*, 1986).

The first study describing coproporphyrinogen oxidase regulation indicated that enzyme activity was induced 10-fold under conditions in which the cells were grown anaerobically (Miyake and Sugimura, 1968). The authors concluded that heme represses the synthesis of the enzyme. In subsequent studies, the activity of the yeast enzyme has been shown to be increased in heme biosynthetic mutants (Urban-Grimal and Labbe-

Bois, 1981; Rytka *et al.*, 1984) regardless of the precise block in the pathway. The increase in enzyme activity could be partially reversed by addition of heme to the growth media (Zagorec and Labbe-Bois, 1986). These data suggest that synthesis of coproporphyrinogen oxidase is negatively regulated by heme as had been postulated by Miyake and Sugimura. It was further demonstrated that the activity of the enzyme was increased 40-fold in wild-type cellular extracts prepared from anaerobic cultures as compared to levels from extracts prepared aerobically (Zagorec and Labbe-Bois, 1986). This observation indicated that both heme and oxygen could negatively regulate the activity of the enzyme, however the effects of both were not additive as no additional increase in enzyme activity was seen in a heme-deficient mutant strain that was grown anaerobically. Thus, all the data indicate that both heme and oxygen negatively regulate coproporphyrinogen oxidase activity, and oxygen may exert its effects on enzyme activity via heme. The mechanism of this regulation by heme was examined by looking at steady state levels of coproporphyrinogen oxidase protein with an antiserum raised against the yeast enzyme. In addition, *HEM13* mRNA levels were examined by Northern blots. Both mRNA and protein levels decreased in a wild type strain grown aerobically, or a heme-deficient strain supplemented with heme, indicating that heme and oxygen exert their effects on the synthesis of coproporphyrinogen oxidase at the transcriptional level (Zagorec *et al.*, 1988).

The gene encoding yeast coproporphyrinogen oxidase, *HEM13*, has been cloned and encodes a protein of 328 amino acids with a predicted molecular weight of 37, 673 Da (Zagorec *et al.*, 1988). Its molecular mass and amino-terminal sequence were in agreement with data obtained from the purified enzyme. The sequence of the 5' non-

coding region was also obtained and early deletion studies predicted that sequences within the first 500 nucleotides upstream of the transcription start site were required for aerobic heme-sufficient repression of *HEM13*. A more detailed analysis of the mechanism of repression of *HEM13* expression, including *cis*-acting elements and *trans*-acting factors will be discussed later.

7) Protoporphyrinogen oxidase

The seventh and penultimate step of the heme biosynthetic pathway is catalyzed by protoporphyrinogen oxidase, an inner mitochondrial membrane enzyme. It oxidizes protoporphyrinogen IX to yield protoporphyrin IX in an oxygen-dependent manner. However, unlike coproporphyrinogen oxidase, protoporphyrinogen oxidase is not induced under oxygen limitation, perhaps as a consequence of its maximal velocity being much lower than that of coproporphyrinogen oxidase. The cDNA encoding this enzyme in humans has recently been cloned and shows 86% homology to the cDNA for mouse protoporphyrinogen oxidase (Nishimura *et al.*, 1995). Little information is available concerning the regulation of the human gene, although in murine erythroleukemia cells, it was shown that expression of the last three enzymes in the pathway (including protoporphyrinogen oxidase) is induced during differentiation of these cells (Taketani *et al.*, 1995).

The yeast enzyme has a molecular weight of 55 kDa (Camadro *et al.*, 1994). However, *in vitro* translation of total RNA and subsequent immunoprecipitation of newly synthesized protoporphyrinogen oxidase revealed a protein with a molecular weight of 58 kDa. This is consistent with the presence of a leader peptide for import of this enzyme to

within the mitochondrion. A *hem14* mutant has been isolated which lacks cytochromes and accumulates protoporphyrin (Urban-Grimal and Labbe-Bois, 1981; Camadro *et al.*, 1982). Unexpectedly, the mutant makes normal levels of immunodetectable protein and mRNA as determined by *in vitro* translation/immunoprecipitation. Upon further analysis, it was demonstrated that this mutant was also impaired in some facet of iron metabolism, as it had lower levels of free iron in the mitochondrion compared to wild type cells. The structural gene for protoporphyrinogen oxidase has recently been cloned by complementation of the *hem14* mutant allele (Camadro and Labbe, 1996). *HEM14* encodes a protein of 539 amino acids with a predicted molecular weight of 59,665 Da. The protein contains an ADP-βαβ-binding fold similar to those found in other flavoproteins such as fumarate reductase, lipoamide dehydrogenase, and succinate dehydrogenase. Studies utilizing protoporphyrinogen oxidase overexpressed in yeast demonstrated that the amino-terminal mitochondrial targeting sequence is not cleaved during import into the mitochondrion. In addition, the enzyme has an absorption spectrum which is typical of flavoproteins. The *hem14* mutant allele was cloned and sequenced and found to contain two mutations. One resulted in a leucine to proline change at amino acid 422 and the other a lysine to glutamic acid change, at amino acid 424. The mutations are located in a domain similar to that in monoamine oxidase known to be involved in the covalent binding of FAD to the enzyme. The availability of the overexpressed protein will allow for a more detailed examination of the structural properties of the enzyme and delineation of its active site.

8) Ferrochelatase

The final step leading to the biosynthesis of heme is catalyzed by ferrochelatase, a mitochondrial enzyme which attaches ferrous iron to protoporphyrin IX resulting in formation of heme. The same enzyme also utilizes zinc as a substrate for the formation of Zn-protoporphyrin. In humans, there are two alternate mRNA species which are produced by utilization of different poly A signals. Thus, the erythroid and non-erythroid forms of ferrochelatase differ only at their 3' ends. Several potential Sp1 binding sites and binding sites for NF-E2 and GATA-1 have been identified in the promoter for the gene encoding the human ferrochelatase (May *et al.*, 1995).

In yeast, ferrochelatase activity was shown to be increased 2- to 4-fold in cells grown in a non-fermentable carbon source as opposed to glucose grown cells or anaerobically grown cells (Labbe-Bois and Labbe, 1990). This suggested that ferrochelatase was subject to catabolite repression. These changes in ferrochelatase activity were attributed partly to increases in enzyme synthesis and enzyme activity, and also to changes in the lipid environment of the enzyme within the membrane (Labbe-Bois, 1990). The yeast enzyme has been purified and has a molecular mass of 40 kDa, although a precursor form of the enzyme of 44 kDa could also be detected by immunoprecipitation after *in vitro* translation (Camadro and Labbe, 1988). This is consistent with cleavage and removal of a leader peptide during import of the protein into the mitochondrion. The enzyme requires fatty acids to be active *in vitro*, as addition of fatty acids to ferrochelatase isolated from repressed cells resulted in induction of ferrochelatase activity. Ferrochelatase does not appear to be the rate-limiting step in heme biosynthesis since overexpression in yeast did not lead to any changes in the levels

of heme or its intermediates (Labbe-Bois, 1990). The gene for yeast ferrochelatase, *HEM15*, has been cloned and encodes a protein with a predicted molecular weight of 44,545 Da. Transcription of *HEM15* mRNA was unaffected by changes in levels of heme (Pinkham and Keng, 1994). In addition, there is no regulation by the products of the *HAP2/3/4/5* genes, although there is increased transcription of *HEM15* in the presence of ethanol which correlates with the increase in ferrochelatase activity seen under the identical conditions. Finally, the transcription of *HEM15* mRNA was found to be induced when the cellular levels of iron are low. This scenario may necessitate increased ferrochelatase levels in order that whatever iron is available is utilized optimally for synthesis of heme.

II. Regulatory Functions of Heme

In *Saccharomyces cerevisiae*, heme plays an essential role in regulation of expression of genes, especially of those encoding hemoproteins. Regulation can occur at the transcriptional or translational level. In *Saccharomyces cerevisiae*, the most well characterized mechanism by which heme regulates gene expression is on the transcriptional level. Upon synthesis in the mitochondrion, heme is transported to the nucleus where it can act as a positive transcriptional effector molecule.

Two of the enzymes in the heme biosynthetic pathway, coproporphyrinogen oxidase and protoporphyrinogen oxidase, have a strict requirement for oxygen. Therefore, heme can be synthesized only when oxygen is available and can serve as an indicator of oxygen tension. Many genes that are positively or negatively regulated by oxygen are regulated by heme in a similar fashion. In such cases, heme mediates the

effects of oxygen on gene expression. Heme-mediated gene expression can be studied using *hem1* mutant strains. Such strains lack δ-aminolevulinate synthase, which catalyzes the first step of the heme biosynthetic pathway. These strains cannot synthesize heme unless supplemented with δ-ALA. In this situation, cells take up δ-ALA and convert it to heme. Alternatively, *hem1* mutants can be supplemented with Tween 80 (oleic acid), ergosterol, and methionine which represent products the synthesis of which require heme or heme biosynthetic intermediates. The supplementation of Tween 80, ergosterol, and methionine enables *hem1* mutant strains to synthesize integral cellular components, while remaining heme-deficient. The use of *hem1* mutant strains makes possible the examination of regulation of many genes by heme. The activation of transcription by heme occurs via the action of positive regulators, while repression of oxygen- and heme-repressed genes occurs through negative regulators (Table 1).

A well characterized example of transcriptional activation by heme involves the *CYC1* gene. *CYC1* encodes iso-1-cytochrome *c*, a nuclear encoded protein component of the electron transport chain. Transcriptional regulation of *CYC1* was first examined by constructing a fusion of the promoter region of *CYC1* and the ATG start codon to the *E. coli lacZ* gene (Guarente and Mason, 1983). Expression of β-galactosidase from this fusion is induced 200-fold when the cells are shifted from glucose to lactate, a non-fermentable carbon source. Both basal level expression in glucose media and induced transcription in lactate media are heme-dependent (Guarente and Mason, 1983; Guarente *et al.*, 1984). Deletion analysis of sequences between -312 and -178 elicited a dramatic decrease in β-galactosidase expression in the presence of heme.

Table 1. Genes regulated by heme and oxygen

Gene	Enzyme/protein	Regulatory factor	Reference
Induced by heme			
CYC1	Iso-1-cytochrome-*c*	HAP1, HAP2/3/4/5	1-5
CYT1	Cytochrome c_1	HAP1	6
CYB2	Cytochrome b_2	HAP1, HAP2/3/4/5	7
COX4	Cytochrome oxidase, subunit IV	HAP2/3/4/5	8
COX5a	Cytochrome oxidase, subunit Va	HAP2/3/4/5	9
COX6	Cytochrome oxidase, subunit VI	HAP2/3/4/5	10, 11
COR1	QH_2-cytochrome c oxidoreductase, subunit I	?	12
COR2	QH_2-cytochrome c oxidoreductase, subunit II	HAP1, HAP2/3/4/5	13
CTA1	Catalase A	?	14
CTT1	Catalase T	HAP1	14, 15
SOD2	Manganous superoxide dismutase	HAP1, HAP2/3/4/5	16, 17
HMG1	3-Hydroxy-3-methylglutaryl CoA^c reductase, isozyme	HAP1	18
TIF51A	eIF-5A, isolog	?	16
ROX1	Heme-induced transcription repressor	HAP1	19, 20
Repressed by heme			
COX5b	Cytochrome oxidase, subunit V_b, isolog	ROX1	21
HMG2	3-hydroxy-3-methylglutaryl CoA reductase, isozyme	?	18
ERG11	Cytochrome P-450 lanosterol 14α-demethylase	HAP1, ROX1	22, 23
HEM13	Coproporphyringen oxidase	HAP1, ROX1	20, 23, 24
ANB1	eIF-5A, isolog	HAP1, ROX1	25, 26

Table 1 continued

References

1. Guarente and Mason, 1983.
2. Guarente *et al.*, 1984.
3. Pfeifer *et al.*, 1987.
4. Olesen *et al.*, 1987.
5. Forsburg and Guarente, 1989a.
6. Schneider and Guarente, 1991.
7. Lodi and Guiard, 1991.
8. Forsburg and Guarente, 1989b.
9. Trueblood *et al.*, 1988.
10. Trawick *et al.*, 1989.
11. Trawick *et al.*, 1992.
12. Meyers *et al.*, 1987.
13. Dorsman and Grivell, 1990.
14. Hortner *et al.*, 1982.
15. Winkler *et al.*, 1988.
16. Lowry and Zitomer, 1984.
17. Pinkham *et al.*, 1997.
18. Thorsness *et al.*, 1989.
19. Lowry and Zitomer, 1988.
20. Keng, 1992.
21. Hodge *et al.*, 1989.
22. Turi and Loper, 1992.
23. Verdiere *et al.*, 1991.
24. Zagorec *et al.*, 1986.
25. Mehta and Smith, 1989.
26. Lowry *et al.*, 1990.

Therefore, the upstream activation site (UAS) which is required for *CYC1* induction by heme is located in the 300 bp upstream of the transcription start site. The UAS is divided into UAS1, which mediates induction by heme in glucose media, and UAS2, which mediates induction in response to non-fermentable carbon sources. This distinction in UASs was corroborated with the isolation of *trans*-acting factors HAP1 and HAP2/3/4/5 which bind to UAS1 and UAS2 respectively.

Transcriptional activation by heme has been observed for a number of other respiratory genes encoding hemoproteins besides *CYC1*. These include *CYT1* (encoding cytochrome c_1), *CYB2* (encoding cytochrome b_2), and *CTT1* (encoding catalase T) (Winkler *et al.*, 1988; Lodi and Guiard, 1991; Schneider and Guarente, 1991) (Table 1). These genes are regulated in a manner similar to *CYC1*, such that transcription is induced by heme and non-fermentable carbon sources. HAP1 protein has been demonstrated to bind to the regulatory regions of these genes. The binding sites for HAP1 in *CTT1*, *CYT1*, and *CYB2* are similar to each other and to that of *CYC1* (Winkler *et al.*, 1988; Lodi and Guiard, 1991; Schneider and Guarente, 1991). Transcription of *CYT1* and *CYB2* is induced in the presence of a non-fermentable carbon source. These genes are also under the control of the HAP2/3/4/5 complex, in manner similar to that observed with *CYC1*.

Besides transcriptional control, heme is also involved in regulation of translation, protein stability and protein localization. The best studied example of translational control by heme in yeast involves regulation of translation of the *CTT1* mRNA, encoding Catalase T. In an *in vitro* translation system devoid of heme, translation of Catalase T mRNA occurred with a reduced efficiency. Upon addition of heme to the translation extract, an increased amount of Catalase T protein was detected (Hamilton *et al.*, 1982).

This translational regulation by heme in yeast parallels that seen in higher eucaryotic cells. Regulation of protein synthesis in reticulocytes is affected by heme levels. Heme deficiency leads to the activation of the hemin regulated inhibitor (HRI), a protein kinase which phosphorylates eIF2α at serine 51 and subsequently inhibits translation initiation (Pathak *et al.*, 1988).

A role for heme in controlling protein stability was demonstrated by analysis of strains with mutations in the *cyc3* locus. Mutants in *cyc3* expressed wildtype levels of *CYC1* mRNA, encoding iso-1-cytochrome *c*. However, no iso-1-cytochrome *c* could be detected in these mutants suggesting an instability of the protein in the *cyc3* background (Matner and Sherman, 1982). The *CYC3* gene was subsequently shown to encode the enzyme cytochrome *c* heme lyase which catalyzes the attachment of heme to apo-iso-1-cytochrome *c* as it enters the mitochondrion. Formation of the holoenzyme renders the protein stable (Dumont *et al.*, 1987). Cytochrome *c* heme lyase is also required for transport and proper localization of cytochrome *c* to the mitochondrial intermembrane space (Nicholson *et al.*, 1988).

III. Heme-dependent transcription factors

1) HAP1 (CYP1) protein

The heme-dependent transcription factors include the HAP1 activator and HAP2/3/4/5 activation complex. These factors were uncovered in the studies on the regulation of *CYC1*, encoding iso-1-cytochrome *c*. *CYC1* expression is heme-dependent and is induced in the presence of non-fermentable carbon sources. HAP1 is also required

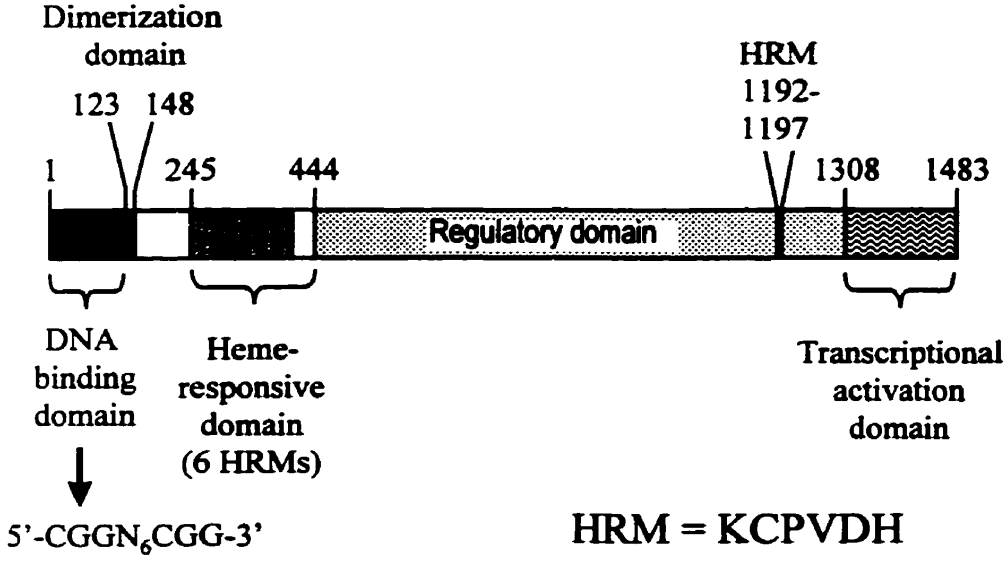
for the expression of *CYC7*, encoding iso-2-cytochrome *c* as well as a number of other genes (see Table 1). The *CYP1* gene was initially identified in a selection for mutations which caused increased expression of *CYC7* (Clavilier *et al.*, 1969). Subsequently, *HAP1* was identified in a screen for mutants with decreased expression of *CYC1* (Guarente *et al.*, 1984). The two genes were found to be allelic and the designation *HAP1* became the accepted designation.

A) Structure of HAP1

The *HAP1* gene has been sequenced and encodes a protein of 1483 amino acids with a predicted molecular weight of 164 kDa (Creusot *et al.*, 1988; Verdiere *et al.*, 1988; Pfeifer *et al.*, 1989). The protein is thought to bind to DNA and activate transcription in the presence of heme. HAP1 protein contains several functional domains which have been localized to different regions of the protein (Figure 3). The first 148 amino acids of HAP1 represents the DNA-binding domain of HAP1. Within this region (residues 64-84), a strong homology to the cysteine-rich zinc finger domains of other yeast transcription activators such as GAL4 protein exists. The importance of the zinc finger domain for DNA-binding by HAP1 protein was demonstrated by mutagenesis of two cysteine residues within the proposed zinc finger domain. When Cys-64 or Cys-81, postulated to be residues required for zinc binding, were mutagenized, DNA-binding activity of HAP1 protein was abolished (Pfeifer *et al.*, 1989).

Adjacent to the zinc finger domain is a dimerization domain found within amino acid residues 123-148. This domain contains a coiled-coiled structure similar to that found within the GAL4 dimerization domains. This region of HAP1 was found to be

Figure 3. Structure of the HAP1 protein. *HAP1* encodes a 1483 amino acid protein with many functional domains. Within the first 148 amino acids of the protein are a zinc finger DNA-binding domain as well as a dimerization domain. The region from amino acid 245-444 contains six copies of the HRM sequence responsible for heme binding. The regulatory region of HAP1 is located within amino acids 447-1308 and contains another copy of the HRM motif (amino acid 1192-1197). The transcriptional activation domain of HAP1 is located within amino acids 1308-1483.


Dimerization domain
123
148
1
245
444
HRM 1192-1197
1308
1483
Regulatory domain
DNA binding domain
Heme-responsive domain (6 HRMs)
Transcriptional activation domain
5'-CGGN6CGG-3'
HRM = KCPVDH

required for DNA-binding and transcriptional activation, as these functions of HAP1 are carried out by protein dimers (Zhang *et al.*, 1993). Overlapping this region (residues 127-189) is a domain of HAP1 which contains a stretch of 12 glutamine residues which are postulated to function as a hinge region separating two functional domains, in this case the dimerization and heme regulatory domains.

The region between amino acids 245 and 444 is characterized by the presence of six repeats of the heme regulatory motif (HRM) containing the amino acids Lys/Arg-Cys-Pro-Val/Ile-Asp-His. This motif is also found within the mammalian δ-aminolevulinate synthase (ALAS) and HRI kinase, which inactivates translational initiation factor eIF2α (Wek, 1994). The HRM domain was shown to be able to associate with heme and the association of the HRM with heme results in changes in both physical and chemical properties of HAP1 (Zhang and Guarente, 1995). The cysteine residue in HRM is absolutely essential for heme binding and the adjacent proline increases the affinity of HRM for heme. These changes in HAP1 conformation upon heme binding may permit binding of HAP1 to DNA. Deletion of the HRM results in a HAP1 molecule which is able to constitutively activate transcription independently of heme. HAP1 derivatives deleted for amino acids 245-444 result in constitutive, heme-independent activation of *CYC1 in vivo*. In contrast, a HAP1 derivative containing the amino-terminal 445 amino acid residues was able to bind to DNA, and this DNA-binding ability was heme-dependent.

Within amino acids 447-1308 of HAP1 protein exists a region that confers regulation. Within this region of the protein there is another copy of the HRM repeat

between amino acid residues 1192-1197. This copy of the HRM is thought to modulate the activation of HAP1 in the presence of heme in a positive manner. When fused to a heterologous activation domain, this HRM was capable of conferring heme dependence to this alternate activation domain (Zhang and Guarente, 1995).

The transcriptional activation domain of the protein is located at the carboxy-terminal end of HAP1 (1308-1483) (Pfeifer *et al.*, 1989). This region of the protein is highly acidic, a feature conserved in transcriptional activation domains from yeast to mammals (Guarente and Birmingham-McDonogh, 1992). Deletion of this region of the protein results in loss of transcriptional activation by HAP1 *in vivo*, although the protein is fully capable of binding to DNA *in vitro*.

B) HAP1 DNA-binding characteristics

Most of the data regarding the mechanism by which HAP1 recognizes its target sequence involve analysis of the regulatory regions of two genes, *CYC1* and *CYC7*, which encode iso-1-cytochrome *c* and iso-2-cytochrome *c* respectively. Approximately 95% of cytochrome *c* within the cell is encoded by *CYC1*. Iso-2-cytochrome *c* represents only 5% of the cytochrome *c* found within the cell. The difference in the level of expression of these two genes can partly be accounted for by the ability of HAP1 protein to recognize both *CYC1* and *CYC7* promoters with similar affinities, yet activate transcription from the promoters differentially. The HAP1 binding sequences in the respective promoters display obvious differences. HAP1 binds to *CYC1* through two sites known as the UAS1A and UAS1B, which were identified by point mutations to be important for UAS1 activity *in vivo* (Lalonde *et al.*, 1986). The binding sites in UAS1A

and UAS1B constitute imperfect direct repeats and DNA binding studies suggest that binding of HAP1 to UAS1A is weaker than that observed at UAS1B (Kim *et al.*, 1990). Cooperativity of binding of HAP1 proteins may be required for formation of the complex at UAS1A, since no HAP1-UAS1A complexes can be detected in the absence of complex formation at UAS1B. In addition, the binding of HAP1 to UAS1 is stimulated by the addition of heme (Pfeifer *et al.*, 1987).

Binding of HAP1 to *CYC7* is also stimulated by addition of heme to extracts. The *CYC7* binding site is composed of 9 bp direct repeats which bear no similarity to the *CYC1* UAS. Despite the dissimilarity, *CYC1* and *CYC7* sites have similar relative affinities for HAP1 and the two binding sites compete equally well for HAP1 protein. Closer inspection of the different binding sites reveals the presence of a GC base pair in *CYC7* at the same position where *CYC1* contains a CG base pair. If the GC base pair was substituted with a CG base pair in the *CYC7* site, expression from *CYC7* driven by HAP1 was increased to levels seen for *CYC1* (Cerdan and Zitomer, 1988).

The ability of HAP1 to distinguish between the two binding sites was better understood with the isolation of the *hap1-18* mutant, which contains a Ser-63 to Arg-63 substitution in the DNA-binding domain. Ser-63 is located immediately amino-terminal to the zinc finger domain of HAP1. In strains containing this mutant allele, transcription from *CYC7* is greatly enhanced, whereas that from *CYC1* is significantly reduced (Verdiere *et al.*, 1986). The increased expression of *CYC7* in *hap1-18* is not as a result of increased affinity of HAP1-18 for the *CYC7* site. When extracts prepared from the *hap1-18* allele were used in gel retardation assays, binding by the HAP1-18 protein to *CYC1* UAS1 was significantly reduced, whereas binding to *CYC7* UAS was unaffected (Pfeifer

et al., 1987). Since transcription from *CYC7* was affected in the *hap1-18* allele, the implication is that an amino acid in the DNA-binding domain, Ser-63, affects the function of the activation domain.

The fact that expression from *CYC1* and *CYC7* is different despite the fact that HAP1 protein has similar affinities for the sites, suggests that HAP1 has differential properties when bound to the respective sites. This observation was further tested by constructing derivatives of HAP1 with deletions of different regions. These derivatives were tested for their ability to affect transcription from both *CYC1* and *CYC7* UAS sites. Interestingly, the expression from *CYC1* increased as the size of the HAP1 derivatives decreased (Kim *et al.*, 1990). In contrast, expression from *CYC7* decreased as the size of the HAP1 derivatives decreased. One of these derivatives of HAP1 consisted of the amino-terminal DNA-binding domain fused directly to the carboxy-terminal activation domain. This derivative was deleted for all of the HRM between 245-444, and all of the regulatory domain between amino acids 447-1308. It bound to both *CYC1* and *CYC7* UASs with equal affinity, but *in vivo,* activated transcription from *CYC1* and not from *CYC7*. If the activation domain of HAP1 was replaced with that of GAL4, the HAP1-GAL4 chimeric fusion was able to activate expression at both *CYC1* and *CYC7* UASs. This experiment suggests that the ability of the HAP1 deletion derivative to distinguish between *CYC1* and *CYC7* promoters is defined by the activation domain of HAP1. The conformation of the DNA-binding domain of HAP1 differs when the protein is bound to *CYC1* and *CYC7*, and this difference in conformation affects the ability of HAP1 to activate transcription. This hypothesis is further supported by isolation of positive control mutants (Turcotte and Guarente, 1992). These mutants mapped to amino acid

residues flanking the zinc finger of HAP1, yet they selectively affect transcription from *CYC7* and not from *CYC1*. The ability of HAP1 to bind to different sequences may reflect a requirement for different levels of HAP1-mediated transcription of these different genes. As such, a certain degree of flexibility must exist within HAP1 which allows it to recognize dissimilar sequences with similar affinities, and to selectively activate transcription to different degrees from these promoters. Although the mechanisms mediating its function is not fully understood, HAP1 plays a central role in coordinating the expression of respiratory genes with heme availabilty.

2) HAP2/3/4/5 complex

When yeast cells are grown in media containing glucose, they preferentially ferment this sugar to derive energy. When shifted to a non-fermentable carbon source, such as lactate, the cells adapt to aerobic growth by induction of genes encoding enzymes required for respiration. The HAP2/3/4/5 complex is responsible for the regulation of these respiratory genes. This complex was initially identified by isolation of a *hap2* mutant, which is unable to allow derepression of *CYC1* expression on a non-fermentable carbon source (Guarente *et al.*, 1984). The *cis*-acting element responsible for this derepression of *CYC1* expression was localized to UAS2, to a region with the sequence 5'-TGATTGGT-3' (Oleson *et al.*, 1987). Subsequently, the *hap3*, *hap4*, and *hap5* mutations were isolated, and strains with these mutations were found to have phenotypes similar to the *hap2* mutant (Hahn *et al.*, 1988; Forsburg and Guarente, 1989a; McNabb *et al.*, 1995).

A) Functions of the different subunits

All four genes encoding the subunits in the complex have been cloned and carefully studied.

i) HAP2

HAP2 encodes a 265 amino acid DNA-binding protein with a basic carboxy-terminus and a run of glutamine residues between amino acids 120-133 (Pinkham *et al.*, 1987). Amino acids 154-218 comprise a 65 amino acid core which is essential for both DNA-binding and assembly of the complex. Within this region, there is a 44 amino acid domain required for association with other subunits and a 21 amino acid domain required for DNA-binding (Olesen and Guarente, 1990).

ii) HAP3

HAP3 encodes a 144 amino acid protein, which is required for DNA-binding activity of the complex. HAP2 binding to UAS2 of *CYC1* cannot occur in the absence of HAP3, even when HAP2 is overexpressed. HAP2 and HAP3 association can occur in the absence of DNA (Hahn and Guarente, 1988). The HAP3 DNA-binding domain consists of a seven amino acid component which together with HAP2 is sufficient for recognition of the 5'-TGATTGGT-3' sequence in UAS2 of *CYC1* (Xing *et al.*, 1993).

iii) HAP4

HAP4, the third component of the complex is a 554 amino acid protein which interacts with HAP2 and HAP3 via its amino-terminus (Forsburg and Guarente, 1989a). The carboxy-terminal 120 amino acids of HAP4 contain an acidic domain reminiscent of a transcriptional activation domain. Deletion of this region of HAP4 inactivates the transcriptional activation ability of the HAP2/3/4/5 complex. This activation function can be restored by fusion of the amino-terminus of HAP4 to the GAL4 activation domain. A *hap4* deletion mutation can be suppressed by fusing the activation domain of GAL4 to HAP2, demonstrating that HAP4 is not required for DNA-binding activity of HAP2 and HAP3, but provides the transcription activation domain of the HAP2/3/4/5 complex (Olesen and Guarente, 1990).

iv) HAP5

HAP5 encodes a 216 amino acid protein which is required for the DNA-binding activity of HAP2 and HAP3. HAP2 and HAP3 expressed and purified in *E. coli* are unable to bind to DNA *in vitro* unless bacterially-expressed HAP5 is also added to the binding mixture (McNabb *et al.*, 1995). Thus, the HAP2/3/4/5 complex contains a transcriptional activation domain located within the HAP4 protein, and a heterotrimeric DNA-binding activity in HAP2, HAP3, and HAP 5.

B) Regulation of HAP2, HAP3, HAP4, and HAP5

The regulation of *HAP2*, *HAP3*, and *HAP4* expression has been examined; *HAP2*, and *HAP4* expression is induced about five-fold in the presence of lactate (Pinkham and

Guarente, 1985; Forsburg and Guarente, 1989a). However, transcription of *HAP3* is unaffected by carbon source (Hahn *et al.*, 1988).

C) Genes under HAP2/3/4/5 control

The inability of strains containing mutations in *hap2*, *hap3*, and *hap4* to respire is due to the fact that many genes regulated by the HAP2/3/4/5 complex encode cytochromes and subunits of the oxidative phosphorylation enzyme complexes which are required for respiration. In addition to *CYC1*, the genes encoding cytochromes regulated by HAP2/3/4/5 include *COX4*, *COX5a*, *COX6*, *CYT1*, and *COR2* (Trueblood *et al.*,1988; Forsburg and Guarente, 1989b; Trawick *et al.*, 1989; Dorsman and Grivell, 1990, Schneider and Guarente, 1991; Trawick *et al.*,1992) (Table 1). These genes represent a class of HAP2/3/4/5-regulated genes which are heme-regulated and induced by growth in non-fermentable carbon sources. Analysis of the UASs of several of these HAP2/3/4/5 regulated genes revealed a consensus motif which conforms to the sequence 5'-TNATTGGT-3'. One of the best studied of these genes, *COX6*, encodes cytochrome oxidase subunit VI. The promoter of *COX6* contains a single copy of the HAP2/3/4/5 recognition sequence. Although expression of *COX6* is reduced 10-fold in a *hap2* or *hap3* mutant strain, gel retardation assays failed to detect HAP2/3/4/5 binding to the *COX6* UAS (Trawick *et al.*, 1992). The second class of genes regulated by the HAP2/3/4/5 complex are induced by growth in non-fermentable carbon sources but are not heme-regulated. This second class of genes encode enzymes of the tricarboxylic acid (TCA) cycle, and include *CIT1* encoding citrate synthase, *ACO1* encoding aconitase, and *LPD1*

encoding dehydrolipoyl dehydrogenase. Transcript accumulation and β-galactosidase expression of fusion constructs of these genes are reduced 5- to 10-fold in *hap2* or *hap3* mutants (Bowman *et al.*, 1992; Rosenkrantz *et al.*, 1994). The last class of HAP2/3/4/5-regulated genes are those encoding heme biosynthetic pathway enzymes. These genes are neither heme-regulated nor lactate-inducible. Examples of this class of gene include *HEM1* encoding δ-aminolevulinate synthase, and *HEM3* encoding porphobilinogen deaminase (Keng and Guarente, 1987; Keng *et al.*, 1992).

IV. Heme-repressed genes

Oxygen serves as an electron acceptor for a variety of essential metabolic reactions in the cell. These include respiration, and fatty acid and heme biosynthesis. Although cells can derive energy solely through fermentation and bypass the oxygen requirement in respiration, the need for oxygen in sterol and fatty acid biosynthesis cannot be bypassed. Therefore, it makes sense for cells grown under hypoxia to maximize the use of limiting oxygen under these conditions. Maximizing the use of limiting molecular oxygen under hypoxic conditions is achieved by increasing the levels of enzymes which use oxygen as a substrate. Thus, as the concentration of oxygen falls, the reaction rate of these enzymes increases. Since oxygen is required as a substrate for two of the enzymes in heme biosynthesis, heme can be synthesized only when oxygen is available. Thus, heme is a good indicator of oxygen tension , and most genes which are repressed by oxygen are similarly repressed by heme. One of the prototypical heme-repressed genes is the *ANB1* gene which encodes an isoform of the translation initiation

factor eIF-5a (Mehta and Smith, 1989; Lowry *et al.*, 1990). Other genes include *COX5b* encoding subunit Vb of cytochrome *c* oxidase (Hodge *et al.*, 1989), *HMG2* encoding 3-hydroxy-3 methylglutaryl CoA reductase (Thorsness *et al.*, 1989), and the *OLE1* gene encoding Δ9 fatty acid desaturase (Stukey *et al.*, 1990) (Table 1). Interestingly, many of the genes which are repressed by heme have homologues whose expression is activated by heme. For example, the *COX5b* homologue *COX5a*, requires heme for expression (Trueblood *et al.*, 1988).

Much of the mechanism of repression of hypoxic genes by heme was deduced from studies on the *ANB1* gene. *ANB1* expression is repressed by a protein encoded by the *ROX1* gene (Lowry and Zitomer, 1988). In the presence of oxygen, heme is made and *ROX1* transcription is induced (Lowry and Lieber, 1986). The ROX1 repressor binds to an operator site in *ANB1* leading to repression of *ANB1* transcription (Balasubramanian *et al.*, 1993). Under hypoxic or anaerobic conditions, heme levels are reduced. Consequently, transcription of *ROX1* is not induced. The absence of the repressor results in derepression of *ANB1* expression. Regulation of *ANB1* expression by ROX1 as presented here represents a very simplistic picture of an otherwise complex pathway involving several other proteins. In the following sections, we will examine ROX1 structure and function, as well as the other regulatory factors involved in repression of hypoxic genes, in a more detailed fashion. The role of ROX1 as it pertains to regulation of *HEM13* expression will also be examined.

V. Heme-regulated repression factors

1) ROX1

The *Saccharomyces cerevisiae* ROX1 protein belongs to the High Mobility Group (HMG) family of DNA-binding proteins and represses hypoxic genes such as *ANB1* and *COX5b* in the presence of heme (Lowry and Zitomer, 1988; Hodge *et al.*, 1989). The mechanism of repression by ROX1 has recently become an area of intense interest. ROX1 repression of hypoxic genes involves binding of ROX1 protein to hypoxic operator sites located within the regulatory sequences of these genes (Balasubramanian *et al.*, 1993; Di Flumeri *et al.*, 1996). In addition, recent data clearly shows that ROX1 repression of hypoxic genes requires other proteins which do not themselves contact DNA. This general repressor complex consists of the TUP1 and SSN6 proteins which are recruited to the target genes by binding of ROX1 protein to the regulatory regions of these genes (Deckert *et al.*, 1995a). The mechanism by which ROX1, in conjunction with TUP1 and SSN6, represses transcription is the focus of ongoing research. There is evidence both for mechanisms involving formation of a repressive chromatin structure and for interference with the basal transcription machinery (Cooper *et al.*, 1994; Herschbach *et al.*, 1994).

A) Cloning of ROX1

Mutations in *ROX1* (regulation by oxygen) were initially identified in a selection for mutants which would allow for constitutive expression of a heme-repressed gene, *ANB1*, (Lowry and Zitomer, 1984). The mutation *rox-1-b3* was recessive and affected expression of heme-repressed genes such as *ANB1* such that they were fully induced

under conditions where heme was present. Subsequently, a number of other heme-repressed genes have been shown to require ROX1 for repression under heme-sufficient conditions. These include *HEM13* and *COX5b* which respectively encode coproporphyrinogen oxidase and subunit Vb of cytochrome oxidase (Hodge *et al.*, 1989; Keng, 1992). The mechanism of regulation of *HEM13* by ROX1 will be reviewed in subsequent sections and is the subject of research included in chapters 5 and 6 of this thesis. The *rox1-b3* mutant strain was transformed with a yeast DNA library and the *ROX1* gene was cloned by complementation of the *rox1-b3* phenotype. This enabled the construction of a null mutant strain in which the entire *ROX1* gene was deleted. The phenotype of the *rox1* deletion strain was identical to that of the *rox1-b3* mutant (Lowry and Zitomer, 1988).

B) Regulation of ROX1 expression

Under conditions when oxygen is available, hypoxic genes in yeast are repressed through the action of the ROX1 repressor protein. The observed negative control of hypoxic genes by oxygen is mediated through the action of heme which is only synthesized when oxygen is available. This was demonstrated conclusively for the *ANB1* gene in experiments which examined mRNA levels under different growth conditions. High *ANB1* mRNA levels were observed under anaerobic conditions in the absence of heme. However, no *ANB1* transcript could be detected in anaerobically grown cells when heme was added (Lowry and Lieber, 1986). When cells containing the *rox1-b3* allele were grown anaerobically, heme no longer repressed expression of *ANB1*. Heme was shown to be required for the synthesis of ROX1. Under aerobic conditions, the *ROX1*

transcript was present but was completely eliminated after 90 min. of anaerobic growth. Heme was the signal necessary for aerobic induction, as the loss of *ROX1* expression in anaerobic cells could be reversed by the addition of heme (Lowry and Zitomer, 1988). Thus, expression of *ROX1* is transcriptionally activated by heme. Subsequent experiments pointed to the involvement of the HAP1 protein in ROX1-mediated repression of hypoxic gene expression. In *hap1* mutant strains, *ROX1* mRNA levels are greatly reduced (Keng, 1992). Therefore, HAP1 induces the transcription of *ROX1* in the presence of heme, and the requirement for heme for repression can be bypassed altogether if *ROX1* expression is rendered heme-independent. ROX1 repressor synthesized when transcription of the *ROX1* coding region was driven by the promoter of the yeast galactose inducible gene *GAL1* was still capable of repressing hypoxic gene expression in the absence of heme (Keng, 1992; Zitomer *et al.*, 1997a). Therefore, heme is only required for synthesis of ROX1 protein, and not for its function.

Regulation of cellular ROX1 levels must be finely tuned in order to avoid overexpression of *ROX1* which could lead to inability to turn on required hypoxic genes as cells are shifted from aerobic to hypoxic conditions. Morever, hypoxic genes such as *OLE1* encoding Δ9 fatty acid desaturase and *ERG11* encoding lanosterol-14α-demethylase encode functions that are also required aerobically. A low level of expression of these genes must be ensured in aerobic conditions.

One mechanism for achieving this fine balance between repression and activation of hypoxic genes such that hypoxic genes can be turned on quickly as the oxygen tension decreases is via the short half-life of the ROX1 protein. Hypoxia is a transient state.

Therefore, the cellular response to lack of oxygen must be rapid. Under conditions when oxygen becomes limiting, *ANB1* expression is fully induced in less than a generation time (Lowry and Zitomer, 1988). This would suggest that the rapid accumulation of ANB1 protein in the cell when oxygen is limiting is due to rapid degradation of ROX1. To examine this possibility, ROX1 was tagged with a *c-myc* epitope and overexpressed in a yeast strain grown aerobically to induce *ROX1* transcription. Subsequent transfer to hypoxic conditions resulted in rapid turnover of ROX1 protein with a dramatic decrease after 30 minutes and complete loss of detection of ROX1 with a *c-myc* monoclonal antibody after 60 minutes (Zitomer *et al.*, 1997a). Thus, the quick adaptation of the cell to hypoxic conditions, that is, induction of the hypoxic genes, is achieved via the short half-life of the ROX1 protein.

A further layer of regulation of *ROX1* expression was revealed when a ROX1 binding site was located upstream of the *ROX1* gene suggesting that ROX1 is involved in its own regulation. Gel mobility shift assays confirmed that ROX1 does bind its own regulatory region (Deckert *et al.*, 1995b). Expression of a *ROX1-lacZ* fusion in a *rox1* deletion strain resulted in an increase in levels of β-galactosidase activity under aerobic conditions of almost 5-fold when compared to expression in a strain with a wild type copy of *ROX1*. In addition, when a *ROX1-lacZ* fusion devoid of the ROX1 binding sites was transformed into a *ROX1* wild type strain, levels of β-galactosidase activity were greatly increased, demonstrating that ROX1 does bind to its regulatory region and repress its own synthesis.

The regulation of *ROX1* expression is made even more complex when the TUP1 and SSN6 proteins are factored into the picture. Previously it was demonstrated that deletion of the *TUP1* gene resulted in increased expression of the *ROX1* gene (Zhang *et al.*, 1991). More recently, the repression of hypoxic genes by ROX1 has been shown to require the protein products of the TUP1 and SSN6 genes (Balasubramanian *et al.*, 1993; Amillet *et al.*, 1995). When expression of *ANB1-lacZ* or *HEM13-lacZ* was measured in a *tup1* or *ssn6* mutant strain which carried a wild type copy of *ROX1*, derepression of both reporter gene fusions occurred under repressing conditions. The fact that ROX1 auto represses its own expression combined with the fact that TUP1 is required for ROX1-mediated repression of hypoxic genes suggests a role for TUP1 in auto repression of *ROX1*. This was examined by measuring *ROX1-lacZ* expression in a Δ*tup1* mutant strain as well as in a Δ*tup1* Δ*rox1* strain. Under aerobic conditions the Δ*tup1* strain accumulated two-fold more β-galactosidase than the wild type strain (Deckert *et al.*, 1995b). The increase was similar to that seen in the Δ*tup1* Δ*rox1* strain, indicating that TUP1 and ROX1 act through the same pathway for ROX1 self repression. Since TUP1 and SSN6 function as a complex, the effect of deletion of *SSN6* on *ROX1* repression was also examined. β-galactosidase expression in the Δ*ssn6* strain was equivalent to that seen in the Δ*rox1* strain and about 3-fold higher than that in the Δ*tup1* strain. Therefore, both the products of the *TUP1* and *SSN6* genes are required for ROX1 protein to self repress its expression under aerobic conditions. In fact, the importance of ROX1 self repression was seen when *ROX1* expression was driven by the galactose-inducible *GAL1* promoter. In the presence of galactose, transformants containing this construct produced extremely

high levels of *ROX1* mRNA and protein compared to wild type cells. These transformants were unable to grow in liquid galactose media or plates containing galactose (Deckert *et al.*, 1995b). Thus, although *ROX1* expression is required for represson of hypoxic genes under aerobic conditions, the levels must be finely tuned. This fine tuning prevents complete repression of certain hypoxic genes such as *HEM13*, *ERG11*, and *OLE1*, whose products are required at low levels during aerobic growth for heme and membrane biosynthesis.

C) ROX1 binding site

ROX1 repression of hypoxic genes was first examined with respect to the *ANB1* gene. In an attempt to locate *cis*-acting regulatory DNA sequences within the *ANB1* regulatory region, plasmids with an *ANB1-lacZ* fusion with various deletions in the upstream noncoding region were individually introduced into a *ROX1* wild type strain and assayed for β-galactosidase expression under aerobic conditions. Two negative regulatory or operator sites, A and B, were shown to be necessary for repression by ROX1 (Lowry *et al.*, 1990). Each operator is composed of two copies of a 12 base pair operator sequence containing the consensus 5'-YYYATTGTTCTC-3'. Deletion of operator A resulted in a 50-fold increase in *ANB1* expression under repressing conditions, whereas deletion of operator B resulted in a 5-fold increase in *ANB1* expression. Operators A and B are located upstream of both the TATA box and transcription initiation sites of *ANB1* and downstream of the poly dA:dT upstream activation sequences of *ANB1*. In *Saccharomyces cerevisiae*, these poly dA:dT sequences are

thought to mediate constitutive expression of the gene upstream of which they are located (Struhl, 1985). Thus, *ANB1* is regulated both constitutively by positive elements and negatively by oxygen and heme via the operator elements. The essential nature of the operator consensus sequence in repression was demonstrated by insertion of an oligonucleotide of the sequence 5'-CCCATTGTTCTC-3' into a plasmid containing an *ANB1-lacZ* fusion devoid of both operator regions A and B (Lowry *et al.*, 1990). Repression of the *ANB1-lacZ* fusion was partially restored when this plasmid was transformed into a *ROX1* wild type strain. In addition, when two copies of the oligonucleotide were placed upstream of *ANB1-lacZ*, the observed repression was four-fold higher than repression in a construct containing only one copy of the oligonucleotide. The same oligonucleotides did not confer repression when placed upstream of a *GAL1-lacZ* fusion construct. However, when larger inserts of 70 bp comprising operator A or 82 bp comprising operator B were placed upstream of the *GAL1-lacZ* fusion, 80-fold and 13-fold repression of *GAL1-lacZ* expression were observed, respectively. These observations suggest that repression by the consensus operator sequence 5'-CCCATTGTTCTC-3' requires flanking sequences. The importance of the operator consensus was independently demonstrated by the creation of multiple point mutations within the consensus operator sequence of operator A (Mehta and Smith, 1989). A fusion of a 300 bp region of *ANB1* regulatory sequences containing operator A was fused upstream of *CYC1-lacZ* deleted for its own regulatory sequences. Expression from this fusion was found to be repressed under aerobic conditions. Mutations within the *ANB1* sequences which caused elevated expression of the fusion under aerobic conditions were isolated and analyzed. A good number of the mutants were located

within the operator consensus sequence. In particular, mutation of the highly conserved T in the eighth position of the operator consensus sequence resulted in highly elevated expression of the reporter construct under repressing conditions.

Another heme-repressed gene, *COX5b*, also contains one copy of the consensus operator sequence within a 44 bp region that functions to repress expression of this gene under aerobic conditions (Hodge *et al.*, 1990). Deletion of this region resulted in high level expression of *COX5b* under aerobic conditions. This region was also able to repress transcription from a heterologous gene. In a construct where this region was positioned downstream from the *CYC1* upstream activation sequences (UAS) or the *LEU2* UAS, transcription driven by these UASs was decreased 5- to 7-fold. Another copy of the consensus operator sequence is also found within the upstream non-coding sequences of *COX5b*. However, this copy is located downstream of the TATA box and its ability to repress transcription has not been investigated.

A comparison of the regulatory sequences from different heme-repressed, ROX1-regulated genes has led to the designation of the hypoxic operator sequence as having the consensus 5'-YYYATTGTTCTC-3' where Y represents a pyrimidine (Table 2) (Zitomer *et al.*, 1997b). These operator elements are found in single copy in genes such as *ERG11* which encodes lanosterol 14α-demethylase (Turi and Loper, 1992), *OLE1* which encodes Δ9 fatty acid desaturase (Stukey *et al.*, 1990), and *CPR1* which encodes NADPH-cytochrome P450 reductase (Turi and Loper, 1992). In other cases, two or more copies of this operator consensus sequence has been found. For instance, in the 200 bp upstream of the *AAC3* coding region which encodes the mitochondrial ADP/ATP

Table 2. Hypoxic operator sequences

Gene	Enzyme/protein		Operator[a]		Reference
HEM13	Coproporphyrinogen III oxidase	-476 -238 -185	TCAATTGTTTAG TGCTTTGTTCAA CCCATTGTTCTC	-465 -249 -174	1
ERG11	Cytochrome P450, lanosterol 14α-demethylase	-358	CCTATTGTGCAT	-347	2
CPR1	NADPH-cytochrome P450 reductase	-95	TCATTTGTTCCT	-84	2
HMG2	3-Hydroxy-3-methylglutaryl CoA reductase	-282	CGCATTGTTTTG	-271	Yeast Genome Database
SUT1	Sterol uptake	-243 -342	GTTTTTGTTCCT AGCTTTGTTCTT	-232 -331	3
OLE1	Δ9 fatty acid desaturase	-272	CCTATTGTTACG	-261	4
COX5b	Subunit Vb of cytochrome c oxidase	-228	TGTATTGTTCGA	-217	5
CYC7	Iso-2-cytochrome c	-333	CCTATTGTATTA	-322	6
AAC3	ADP/ATP translocator	-197 -145	TTCATTGTTTGG TCCATTGTTCTT	-186 -134	7
ANB1	eIF5A	-316 -285 -218 -197	TCCATTGTTCGT CCTATTGTTCTC TCCATTGTTCTC CTCATTGTTGCT	-305 -274 -207 -186	8,9,10
ROX1	Heme-induced transcription repressor	-397 -364	CCTATTGTTGCT CGTATTGTCTTG	-386 -353	11
	Consensus sequence		YYYATTGTTCTC[b]		

[a]Sequences are numbered with the first base in the coding strand immediately 5' to the ATG initiation codon as -1.

[b]Y represents pyrimidine.

Table 2 continued

References

1. Amillet *et al.*, 1996.
2. Turi and Loper, 1992.
3. Bourot and Karst, 1995.
4. Stukey *et al.*, 1990.
5. Hodge *et al.*, 1990.
6. Wright and Zitomer, 1984.
7. Sabova *et al.*, 1993.
8. Balasubramanian *et al.*, 1993.
9. Zitomer *et al.*, 1997ai
10. Lowry *et al.*, 1990.
11. Deckert *et al.*, 1995b.

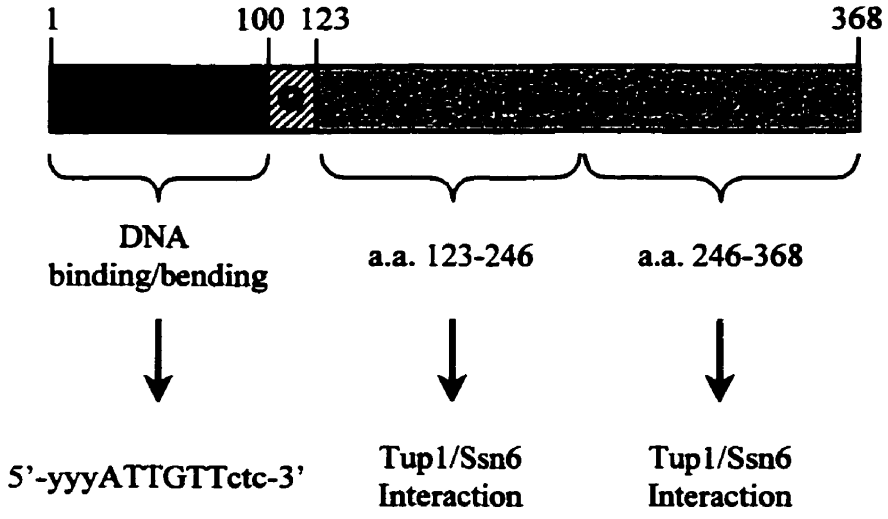
translocator, two copies of the consensus operator sequence can be found. When they are deleted, increased aerobic expression was observed (Sabova *et al.*, 1993). The *SUT1* gene which is involved in sterol uptake also contains two copies of the hypoxic consensus operator sequence (Bourot and Karst, 1995). Some heme-repressed genes contain multiple consensus operator sequences which contribute to repression, under aerobic conditions, to different degrees. As already mentioned, the *ANB1* gene contains four copies of this hypoxic consensus operator sequence within the first 316 bp upstream of the translation initiation codon. Operators A and B are composed of two copies each of the consensus operator sequence and contribute differentially to repression. Most of the repression of *ANB1* is conferred by operator A which when deleted results in a 10-fold higher derepression of *ANB1* expression than deletion of operator B. The *HEM13* gene which encodes coproporphyrinogen oxidase is the only other hypoxic gene which contains multiple copies of this operator consensus sequence. The analysis and the functionality of these elements will be discussed under a subsequent heading and is also the subject of some of the research data presented in chapters 5 and 6.

D) Structure of ROX1 protein

i) DNA binding domain (HMG domain)

Sequence analysis of *ROX1* revealed a coding region of 368 amino acids. The amino-terminal one third of the protein contains an HMG domain found in the HMG class of non-histone chromatin proteins (Balasubramanian *et al.*, 1993) (Figure 4). ROX1 was expressed in *E. coli* as a fusion to the maltose binding protein (MBP) and purified using affinity chromatography (Balasubramanian *et al.*, 1993). Purified ROX1 protein

Figure 4. Structure of the ROX1 protein. Amino acids 1-100 of ROX1 contain a high mobility group (HMG) domain required for both DNA-binding activity and the ability of ROX1 to bend the DNA helix by 90°. The binding sequence depicted is a consensus sequence derived by alignment of the regulatory regions of hypoxic genes regulated by ROX1. Y in the consensus sequence represents a pyrimidine. Q refers to the glutamine-rich domain located between amino acids 100-123. The repression domain is located within the carboxy-terminus of ROX1 between amino acids 123-368 and can be sub-divided into two-halves both of which confer repression. The repression domain has been implicated in protein-protein interactions with the TUP1/SSN6 complex.


1
100
123
368
DNA
binding/bending
a.a. 123-246
a.a. 246-368
5'-yyyATTGTTctc-3'
Tup1/Ssn6
Interaction
Tup1/Ssn6
Interaction

was used in gel mobility shift assays to demonstrate that ROX1 is a member of the sequence specific HMG proteins as it bound to oligonucleotides containing the core sequence 5'-ATTGTT-3' with sequence specificity. These data also conclusively proved that the consensus operator elements 5'-YYYATTGTTCTC-3' found in the *ANB1* promoter was a ROX1-responsive element. The ability of ROX1 to bind to DNA was dictated by the amino-terminal third of the protein, as any deletion derivative containing the amino- terminal 100 amino acids was capable of specific DNA binding (Deckert *et al.*, 1995a). Therefore, the HMG domain within the first 100 amino acids of the protein allows ROX1 to bind to DNA specifically.

Computer modelling studies predict that the HMG domain of ROX1 shares a structure similar to the HMG domain structure of SRY, a transcription factor involved in male determination. The tertiary structure of the protein-DNA complex for the HMG domain of SRY has been deduced by nuclear magnetic resonance (Werner *et al.*, 1995). The HMG domain of SRY is shaped like the letter C and isoleucine 68 sticks out into the C and intercalates between the two TA base pairs at position 2 and 3 of the consensus SRY binding site 5'-(A/T)TTGTT-3'. This causes distortion of the DNA helix which aids in bending of the helix. ROX1 contains an isoleucine at amino acid 18 which is in a homologous position. In addition, the ROX1 binding site also contains the core 5'-ATTGTT-3' found in the SRY binding site. Mutational analysis of the HMG domain of *ROX1* reinforced the prediction that the structure of the ROX1 HMG domain is similar to that of other HMG domain proteins. Missense mutations were isolated which mapped to the HMG domain of ROX1 and resulted in dramatic decreases in DNA-binding ability (Deckert *et al.*, 1995a). Most of the mutations were in residues that are proposed to

contact the DNA helix, or in amino acids which are presumed to be essential for the structural integrity of the HMG domain. Similar mutations were identified in SRY which affected residues required for contacting the DNA or for maintaining HMG structural conformation. As with other HMG proteins, ROX1 is capable of inducing a bend within the DNA helix upon binding to the consensus operator sequence (Deckert *et al.*, 1995a). A gel-retardation assay whereby the consensus operator is located at various positions along a labeled DNA fragment allows for calculation of the bending angle a protein induces upon DNA to which it is bound. In this type of assay, known as a circular permutation assay, it was shown that ROX1 induces a bending angle of 90° upon the helix. The ability of HMG proteins to bend DNA has been postulated to occur as a consequence of their being architectural proteins, proteins that induce conformational changes in DNA that are essential for their function. The SRY HMG domain also induces a 90° bend in the DNA helix to which it is bound (Ferrari *et al.*, 1992). A mutation in SRY which caused a 20° decrease in bending angle but did not affect the affinity of the protein for DNA was isolated. This mutation led to a severe sex-reversal phenotype, and strongly argues in favor of an essential role of DNA bending in protein function (Pontiggia *et al.*, 1994). This hypothesis seems tenable for ROX1, as repression by this protein is thought to be a complex process which involves multiple protein factors whose interaction to form a higher nucleoprotein structure would be enhanced by distortion of the helix brought about by bending. Interestingly, when the ROX1 carboxy-terminal domain, required for repression, was fused to the GAL4 DNA-binding domain, repression of an *ANB1-lacZ* fusion was only half that of the intact ROX1 (Deckert *et al.*,

1995a). DNA bending may be required for optimal repression, but that remains to be determined.

ii) Glutamine-rich domain

ROX1 contains a glutamine-rich domain between amino acids 102 and 123, in which 16 out of 22 residues in this region are glutamines. The function of this domain of ROX1 is not understood although deletion of this region does not affect the ability of ROX1 to repress *ANB1* expression (Deckert *et al.*, 1995a). Glutamine rich domains have been proposed to be involved in protein-protein interactions in eucaryotic transcriptional activator proteins (Courey *et al.*, 1989). This domain may be involved in protein-protein interaction with an as yet unidentified factor involved in ROX1 mediated repression of hypoxic genes.

iii) Repression domain

The ROX1 protein contains a third domain separable from both the DNA binding domain and the glutamine rich domain. This domain of the ROX1 protein is located within the carboxy-terminal two-thirds of the protein and is referred to as the repression domain. The repression domain contains redundant information. Deletion of amino acids 100-245 resulted in a 2.2-fold reduction in repression of an *ANB1-lacZ* fusion, whereas a deletion of amino acids 247-368 resulted in a derepression of 2.8-fold of the same reporter construct (Deckert *et al.*, 1995a). Thus, deletion of either half of the repression domain does not dramatically affect the ability of the the remaining protein to repress transcription of *ANB1-lacZ.* On the other hand, deletion of both halves of the

repression domain leads to a 17-fold derepression of *ANB1-lacZ* expression. This effect is equal to the derepression seen in a *rox1* deletion strain in which the entire *ROX1* coding region has been removed. To determine if the ROX1 repression domain could repress transcription from a heterologous gene, a fusion of either redundant half or the entire repression domain to the DNA-binding domain of the GAL4 transcriptional activator was constructed. Fusion proteins of the complete repression domain, or either half of the domain to GAL4 could repress transcription of an *ANB1-lacZ* fusion that contains GAL4 binding sites (Deckert *et al.*, 1995a).

Interestingly, despite their functional redundancy, there is no obvious repeated or homologous sequence within the two halves of the repression domain. As repression by ROX1 necessitates the function of the TUP1/SSN6 complex, the repression domain may represent the region required for interaction with this general repression complex.

VI. Repression by ROX1 requires the general repressors SSN6 and TUP1

1) SSN6 (CYC8)

CYC8 was first identified in a selection protocol for mutations which allowed elevated expression of *CYC7,* encoding iso-2-cytochrome *c* (Rothstein and Sherman, 1980). Mutations in *SSN6* were later isolated by their ability to allow constitutive expression of glucose-repressible-genes (Carlson *et al.*, 1984). *SSN6* plays a negative role in *SUC2* (encoding invertase) expression as *ssn6* mutants constitutively express *SUC2* (Trumbly, 1986). Subsequently, it was demonstrated that these 2 genes were allelic. Mutants in *SSN6* (*CYC8*) display pleiotropic phenotypes such as extreme

clumpiness, α-specific mating defects, and an inability to maintain minichromosome plasmids stably. The gene encoding the SSN6 protein has been cloned (Schultz and Carlson, 1987) and its predicted amino acid sequence contains 10 copies of a repeated motif called the tetratricopeptide repeat (TPR) at the amino-terminus This motif was shown to be essential for SSN6 function (Schultz *et al.*, 1990), and has been proposed to be required for protein-protein interactions.

2) TUP1

Mutations in *TUP1* were identified in a variety of screens for different phenotypes. Mutants were first isolated as yeast strains which had acquired the ability to take up dTMP from the media, hence the designation *tup1* (Wickner, 1974). Subsequently, a number of other alleles which were shown to be allelic to *tup1* were isolated. These alleles conferred various phenotypes on the respective strains such as flocculance (Fujita *et al.*, 1990), mating type defects in *MATα* cells (MacKay, 1983), non-sporulation of homozygous diploids (Rothstein and Sherman, 1980), and inability to maintain minichromosome plasmids stably (Thrash-Bingham and Fangman, 1989). In addition, *tup1* mutants have been isolated which are defective in repression of *ANB1* under aerobic conditions (Zhang *et al.*, 1991).

The *TUP1* gene encodes a protein of 713 amino acids which shows sequence similarity to the family of proteins known as the transducins, that constitute β-subunits of G-proteins, although TUP1 itself is not believed to be part of a heteromeric G-protein complex (Williams and Trumbly, 1990; Zhang *et al.*, 1991). The β-transducins are

characterized by the presence of a 40 amino acid motif (WD-40) which contains highly conserved tryptophan-aspartate residues, or simply WD. Proteins with WD repeats are involved in diverse biological functions including signal transduction, gene expression, and cell cycle progression. The function of the WD repeats is currently being examined, although they have been postulated to be important for protein-protein interactions. TUP1 protein contains six or seven WD repeats; deletion of a single repeat results in the same phenotype as complete deletion of the *TUP1* gene with respect to derepression of the *SUC2* gene encoding invertase (Williams and Trumbly, 1990).

3) TUP1 and SSN6 function as a complex

The observation that *tup1* and *ssn6* mutants possess similar pleiotropic phenotypes suggested that the two genes function at similar points in regulation. The phenotypes of many of the *tup1* and *ssn6* mutant alleles suggest that the regulatory pathways controlled by their gene products involve repression mechanisms. Thus, TUP1 and SSN6 may be components of the general repression machinery which contribute to repression mediated by the pathway-specific repressor. Examples of these families of specific repressors include MIG1 protein for glucose-repressible genes, ROX1 protein which represses hypoxic genes, and MATα2 protein which represses **a**-specific genes.

Evidence linking these two proteins to a common regulatory pathway was obtained by co-immunoprecipitation studies which demonstrated that SSN6 and TUP1 form a high-molecular weight complex (Williams *et al.*, 1991). Recently, it was demonstrated that the SSN6-TUP1 repressor complex is composed of four subunits of

TUP1 and one subunit of SSN6 (Varanasi *et al.*, 1996). Fusion of both SSN6 and TUP1 to a LexA DNA-binding domain demonstrated that both proteins could repress transcription of reporter genes with upstream LexA binding sites (Keleher *et al.*, 1992). Repression by SSN6 was dependent on the presence of TUP1, whereas LexA-Tup1 repression was independent of SSN6 (Tzamarias and Struhl, 1994). This led to the proposal that repression is mediated directly by TUP1 and that SSN6 may play a role in recruitment of the SSN6/TUP1 complex to the promoter by interaction with the specific repressors. Functional dissection of the TUP1 protein demonstrated that the amino-terminal 200 amino acids (derivative N200) were partly sufficient for repression of both *SUC2* and *ANB1* genes by the pathway specific repressors MIG1 and ROX1 respectively (Tzamarias and Struhl, 1994) Deletion of all six or seven WD repeats at the extreme carboxy-terminus of TUP1 had a minimal effect on glucose repression and hypoxic gene expression (Tzamarias and Struhl, 1994). However, repression conferred by this TUP1 allele was not complete and suggests that although not essential, the WD motifs seem to affect TUP1 function and are likely to be important. In addition to the repression domain, the amino-terminal 200 residues of TUP1 also contained the domain required for interaction with SSN6 between amino acids 1-72. Therefore, both the repression and SSN6-interaction domains of TUP1 can be found in the first 200 amino acids of the protein. The TUP1/SSN6 complex has no DNA-binding ability and it is proposed that the complex is recruited to the promoter targeted for repression by interaction of the SSN6 protein with the pathway-specific repressor (Tzamarias and Struhl, 1995). Different combinations of the ten TPR repeats of SSN6 were required for interaction with different pathway specific repressor proteins. For instance, TPR motifs 8-10 were

required for interaction with the MIG1 repressor whereas TPR motifs 4-7 were required for interaction with ROX1 protein. In addition, it was demonstrated that TPR motifs 1-3 are necessary and sufficient for the interaction of SSN6 with TUP1 (Tzamarias and Struhl, 1995). However, recruitment of the TUP1/SSN6 complex to the **a**-specific gene operator was found to be through interaction of the MATα2 repressor with the WD repeats of the TUP1 protein (Komachi *et al.*, 1994). Taken together, these results indicate that the SSN6-TUP1 repressor complex can be recruited to pathway specific promoters via different mechanisms. The variety of different protein-protein interactions possible may explain the lack of any consensus motif in the specific repressors which could interact with SSN6 and TUP1. The mechanism of repression by SSN6/TUP1 is thought to be mediated by TUP1 through formation of a repressive chromatin structure, or alternatively through the ability of TUP1 to interfere with some component of the basal transcription machinery (Cooper *et al.*, 1994; Herschbach *et al.*, 1994).

4) Role of SSN6/TUP1 in repression by ROX1

Transcription activation in eucaryotes often involves the binding of activator proteins to DNA sequences which often lie hundreds or thousands of base pairs from the TATA box, where RNA polymerase II binds. In addition, some genes require multiple activators which can act individually or in concert with each other to activate transcription. Such complex promoters would argue against simple repression mechanisms whereby a repressor precludes binding of an activator by sharing an overlapping binding site.

The structure of both ROX1 protein and the promoters of hypoxic genes would suggest that repression of these genes by ROX1 is mechanistically complex. ROX1 has separate DNA-binding and repression domains, and binding to DNA alone is not sufficient to repress transcription of hypoxic genes. In addition, the SSN6 and TUP1 proteins are required for ability of ROX1 to repress transcription of its target genes. Strains containing either *tup1* or *ssn6* deletions showed a 13-fold increase in *ANB1-lacZ* expression under aerobic repressing conditions (Zhang *et al.*, 1991; Balasubramanian *et al.*, 1993). When cells were transformed with a plasmid lacking both operators A and B, no additional derepression was observed in the *ssn6* mutant strain, indicating that SSN6 acts at these sites. This is consistent with the hypothesis that SSN6-TUP1 is a general repressor complex which acts in conjugation with specific DNA-binding proteins, in this case, ROX1.

VII. Regulation of *HEM13* expression

HEM13 encodes the sixth enzyme in the heme biosynthetic pathway, coproporphyrinogen oxidase, which catalyzes the oxygen-dependent conversion of coproporphyrinogen III to protoporphyrinogen IX. Early studies measuring enzymatic activity demonstrated that coproporphyrinogen oxidase activity is increased under anaerobic conditions (Miyake and Sugimura, 1968). Later, it was shown that regulation by oxygen is exerted at the transcriptional level via intracellular heme levels (Zagorec *et al.*, 1988).

The mechanism behind oxygen/heme control of *HEM13* expression has since come under investigation, and a clearer picture is beginning to emerge although many

questions remain unanswered. Here we review what is known about regulation of *HEM13* expression, with particular emphasis on the *cis*- and *trans*-acting regulatory elements that are involved in regulation.

1) *Involvement of heme in HEM13 expression*

To assess the effect of heme on *HEM13* expression, a heme biosythetic mutant defective in the first step of the pathway was used. *hem1* mutants fail to synthesize δ-ALA and one can control the intracellular heme levels by addition of high or low concentrations of δ-ALA to the media. Mutants grown in high concentrations of δ-ALA synthesize heme, whereas growth in low δ-ALA concentrations renders cells relatively heme-deficient. Alternatively, *hem1* mutants can be grown in the complete absence of δ-ALA supplementation when Tween 80, ergosterol, and methionine are added to the media. RNA was isolated from *hem1* mutant strains which were grown under heme-deficient or heme-sufficient conditions, and analyzed by Northern blot using a probe for *HEM13*. The results indicated that growth in the presence of δ-ALA resulted in low *HEM13* mRNA levels, whereas growth in the absence of δ-ALA resulted in induction of *HEM13* expression (Keng, 1992). These results were confirmed when assays for β-galactosidase expression from a *HEM13-lacZ* fusion were performed. In this case, expression was derepressed 20-fold in the absence of heme.

2) Function of ROX1 in repression of HEM13

Examples of repression of hypoxic genes are well documented, and in most cases implicate the product of the *ROX1* gene (Hodge *et al.*, 1989, Lowry and Lieber, 1986). Strains carrying a *Δrox1* allele show induction of expression of hypoxic genes such as *ANB1* and *COX5b* under repressing aerobic conditions. Therefore, repression of *HEM13* expression by oxygen would likely involve the ROX1 protein. To test this hypothesis, a *Δrox1* strain was transformed with a *HEM13-lacZ* fusion and assayed for β-galactosidase expression. There was a dramatic increase in *HEM13* expression in a *Δrox1* strain. In addition, Northern analysis demonstrated a marked increase in *HEM13* mRNA levels in the *Δrox1* strain irrespective of growth of the strain in the presence or absence of heme (Keng, 1992). These experiments confirm the requirement of ROX1 for *HEM13* repression under aerobic or heme-sufficient conditions. It remained to be determined what the precise role of heme in repression of *HEM13* was. Is heme required for the function of ROX1, or alternatively is heme solely required to turn on *ROX1* expression? To examine these possibilities, the coding region of *ROX1* was put under the control of the *GAL10* promoter. In this scenario, transcription of *ROX1* is induced by galactose and independent of heme. When *HEM13* expression was examined in a strain containing the *pGAL10-ROX1* fusion in the presence of galactose, *HEM13* was constitutively repressed. This repression occurred both under repressing conditions (+heme) or inducing conditions (-heme). Therefore, the only role of heme in repression of *HEM13* is in the induction of *ROX1* expression. This same conclusion was made independently using a

GAL1-ROX1 fusion in which *ROX1* was placed under the control of the *GAL1* promoter (Zitomer *et al.*, 1997a).

3) Function of HAP1 in HEM13 expression

HAP1 encodes a 1483 amino acid transcriptional activator required for the expression of many respiratory enzymes whose genes are induced in the presence of oxygen or heme. These genes include *CYC1*, encoding iso-1-cytochrome *c*, *CYT1*, encoding cytochrome c_1, and *CTT1*, encoding catalase T. It has been demonstrated that the DNA-binding activity of HAP1 is dependent on heme. Examination of the *HEM13* upstream non-coding region revealed a region that bears homology to a HAP1 binding site between nucleotides -366 to -351. Therefore, it appeared possible that HAP1 may be involved in regulation of *HEM13* expression. This hypothesis was examined by testing the effect of a *hap1* mutation on *HEM13* expression. *HEM13* mRNA levels were increased in a *hap1* mutant strain in the presence of heme (Keng, 1992). Interestingly, in a *hap1* mutant in the absence of heme, a condition under which *HEM13* is usually fully derepressed, *HEM13* mRNA levels were increased but not nearly as much as in a *HAP1* strain in the absence of heme. These results were also observed in β-galactosidase assays. Expression of a *HEM13-lacZ* fusion was derepressed three-fold in a *hap1* mutant grown in the presence of heme. However, expression from *HEM13-lacZ* in a *hap1* mutant grown in the absence of heme was three-fold lower than in the isogenic *HAP1* strain. Therefore, HAP1 is required not only for repression of *HEM13* expression in the

presence of heme, but is also required for activation of *HEM13* expression in the absence of heme.

4) No role of HAP2/3/4/5 complex in HEM13 expression

The HAP2/3/4/5 complex is involved in the expression of a number of heme- and oxygen-regulated genes. Its role in *HEM13* expression was assessed by examining *HEM13* mRNA levels in a *hap2* mutant strain. A *hem1 hap2* mutant strain grown in the absence of heme had the same level of derepression of *HEM13* expression as the isogenic *hem1* mutant strain in the absence of heme (Keng, 1992). Therefore, *HEM13* regulation does not involve the HAP2/3/4/5 complex.

5) HAP1 regulates ROX1 expression

Northern analysis and β-galactosidase assays have demonstrated that both HAP1 and ROX1 are involved in repression of *HEM13* expression in the presence of heme. In addition, the transcription of *ROX1* and the DNA-binding ability of HAP1 both require heme. Therefore, one could envisage a cascade that involves heme, HAP1, and ROX1 all exerting their effects in a coordinate fashion. Perhaps, in the presence of heme, the DNA-binding activity of HAP1 is stimulated leading to *ROX1* induction. Subsequent to its synthesis, ROX1 could act to repress *HEM13* expression. To assess whether HAP1 is required for *ROX1* induction, *ROX1* mRNA levels were examined in a *hap1* mutant. In the *hap1* mutants tested, *ROX1* mRNA levels were greatly reduced in the presence of heme in comparison to the levels seen under the same growth conditions in a *HAP1* strain

(Keng, 1992). Thus, HAP1 and ROX1 function to repress *HEM13* expression within the same cascade, and induction of *ROX1* in the presence of heme requires HAP1. However, this heme- and HAP1-dependent expression of *ROX1* can be bypassed by expression of *ROX1* as a fusion to the *GAL* promoter. In this case, *ROX1* expression is galactose dependent, and *ROX1* transcripts can be observed in a *hap1* mutant in both the presence and absence of heme.

Further evidence suggesting a role for HAP1 in transcriptional activation of ROX1 comes from the isolation of *hap1-43*, a mutant allele of *HAP1* in which the Gly-235 of HAP1 is changed to Asp-235 (Ushinsky and Keng, 1994). This mutation occurs in a region of HAP1 bordering the HRM. The mutant HAP1 protein no longer requires heme to bind and activate expression of HAP1-activated genes such as *CYC1*. Expression of these genes becomes constitutive. In strains with the *hap1-43* allele, *ROX1* is also constitutively induced, resulting in constitutive repression of *HEM13* expression.

6) Function of TUP1 and SSN6 in repression of HEM13

The *TUP1* and *SSN6* genes encode proteins which are involved in repression of diverse genes in *Saccharomyces cerevisiae* (Schultz and Carlson, 1987; Williams and Trumbly, 1990). These proteins have been shown to form a high-molecular weight complex which interact with specific repressors that have DNA-binding activity (Varanasi *et al.*, 1996). The binding of the pathway specific repressors to their cognate DNA recruits the TUP1/SSN6 complex, which in turn represses expression from the promoter by setting up a repressive chromatin structure or by directly interfering with the basal transcription machinery. A few of the specific repressors known to require the

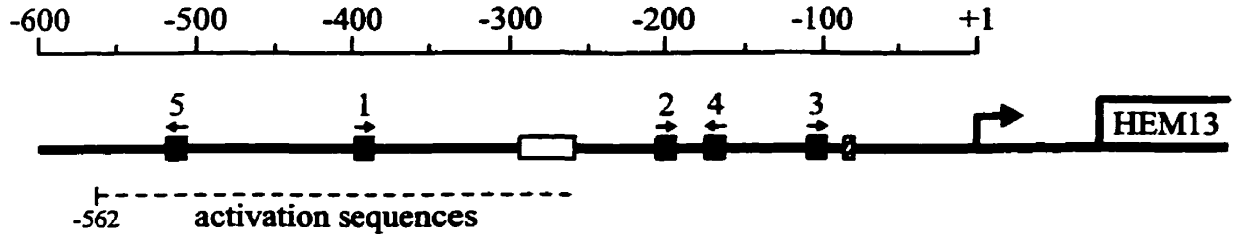
TUP1/SSN6 complex include MIG1 (Treitel and Carlson, 1995), a zinc-finger protein involved in catabolite repression, and α2 (Keleher *et al.*, 1992), a homeodomain protein which regulates cell-type-specific genes. Recently, a *TUP1* homologue in *Candida albicans* has been shown to be involved in repression of genes responsible for initiating filamentous growth (Braun and Johnson, 1997).

The TUP1/SSN6 complex has also been demonstrated to be involved in repression of *ANB1*, a hypoxic gene repressed in the presence of heme through the action of the ROX1 protein. In order to examine if SSN6/TUP1 was involved in ROX1-dependent repression of *HEM13* expression, a set of isogenic strains carrying deleted alleles of either *SSN6* and *TUP1* were constructed. Expression of *HEM13* from these strains was determined by measuring coproporphyrinogen oxidase activity in aerobically grown cells. In *ssn6* and *tup1* deleted strains, a dramatic derepression of *HEM13* expression occurred; coproporphyrinogen oxidase activity was increased 10-fold in the *tup1* mutant strain and 35-fold in the *ssn6* mutant strain (Amillet *et al.*, 1995).

7) HEM13 cis-acting sequences

Deletion analyses have revealed that the elements involved in regulation of *HEM13* expression in the absence and presence of heme are numerous, and the probability that multiple *trans*-acting factors participate in regulation increases with new experimental evidence. Recent evidence suggests that the regulatory regions encompassing the *HEM13* promoter include upstream activation sites (UAS), induction sequences, and repression sites (Figure 5). All three types of regulatory elements will be

Figure 5. Regulatory sequences of *HEM13*. The activation sequences required for full level expression of *HEM13* in the absence of heme span a large region beginning at position -562 upstream of the transcription initiation site (denoted by arrow). The UIS1 sequences, located between nucleotides -293 to -259, are required for induction of *HEM13* expression in the absence of heme. Five hypoxic operator sequences (black boxes), sharing homology to the consensus hypoxic operator sequence derived from the alignment of the regulatory regions of hypoxic genes, are depicted. The TATA box is denoted by a hatched box.



TATA box (-86 to -80)

ROX1 repression sites
- 1 -399 to -388
- 2 -205 to -194
- 3 -110 to -99
- 4 -174 to -163
- 5 -518 to -507

Upstream induction sequences (UIS) -293 to -259

discussed in the following section, with particular emphasis on the effects of deletion of such sites and possible *trans*-acting factors interacting with these sites.

A) Activation sequences

Previous data suggest that the sequences involved in derepression of *HEM13* expression are located upstream of nucleotide -334 relative to the transcription start site (Zagorec *et al.*, 1988). Deletion of sequences upstream of -334 led to an 8-11-fold reduction in coproporphyrinogen oxidase activity under inducing conditions.

To more precisely define the activation sequences, *HEM13-lacZ* fusion plasmids with 5' or internal deletions were constructed and the expression of β-galactosidase from these plasmids was assessed. A plasmid containing 1.45 kilobase (kb) of upstream non-coding and 34 codons of coding sequence of *HEM13* fused to the *lacZ* gene was used as the wild type construct. When transformed into a *hem1* mutant strain, expression from this *HEM13-lacZ* fusion construct is derepressed 15-fold in the absence of heme. Results of 5' deletions indicated that a deletion upstream of position -562 did not alter the regulation of expression. Such a construct still exhibited a 17-fold derepression of *HEM13-lacZ* expression in the absence of heme. Further incremental deletions to position -292 resulted in gradual loss of derepression of the *HEM13-lacZ* fusion. The fusion containing 292 bp of non-coding sequence upstream of the transcription initiation site was derepressed 24-fold less than the wild type *HEM13-lacZ* fusion in the absence of heme. A further deletion of 34 nucleotides to -258 resulted in complete loss of β-galactosidase activity. Therefore, sequences required for activation of *HEM13* expression

are located within a 300 bp region located between nucleotides -562 and -259 upstream of the transcription start site (Keng *et al.*, unpublished observations).

To better define the sites involved in *HEM13* activation, internal deletions within the *HEM13* noncoding region were constructed in *HEM13-lacZ* fusion plasmids and expression of β-galactosidase activity from the different constructs was determined under both repressing and inducing conditions. Expression of the wild-type construct was induced 12-fold in the absence of heme (Keng *et al.*, unpublished observations). The HAP1 protein has been demonstrated to be required for both activation of *HEM13* expression in the absence of heme, and its repression in the presence of heme (Keng, 1992). However, deletion of sequences between -379 and -331 which contain homology to the HAP1 binding site, did not have any effect on *HEM13* expression. However, internal deletion of sequences from -326 to -259 resulted in low level constitutive repression of β-galactosidase activity. Further analysis revealed that the region between -293 and -259 is required for induction of *HEM13* expression in the absence of heme. However, this fragment does not function as an activation sequence when placed upstream of a heterologous gene, and is, hence, designated an upstream induction site (UIS1). UIS1 by itself cannot activate expression, but is essential for derepression of *HEM13* in the absence of heme.

An independent study determined that the sequences required for activation of *HEM13* expression were located between nucleotides -593 and -417, and between nucleotides -292 and -102 (Amillet *et al.*, 1995). This location is consistent with the sequences found to be required for *HEM13* expression by Keng *et al.* Deletion of

sequences between nucleotides –593 and -417 resulted in reduced expression under inducing conditions, whereas deletion of sequences between -292 to -102 impaired full expression under inducing conditions without affecting aerobic basal expression (Amillet *et al.*, 1995).

B) Negative sites

Repression of *HEM13* transcription by oxygen and heme is mediated by the product of *ROX1* (Keng, 1992). ROX1 has been shown to bind to hypoxic operator sequences containing the consensus 5'-CCCATTGTTCTC-3' in other ROX1-repressed genes (Table 2). Examination of the *HEM13* promoter revealed the presence of three such sites located at -110 to -99 (R1), -174 to -163 (R2), and -399 to -388 (R3) upstream of the transcription initiation site. The functionality of these sites in repression of *HEM13* was examined in mutagenesis experiments (Amillet *et al.*, 1996). Each site was altered individually or in combination by site-directed mutagenesis and the effect of the mutations on *HEM13-lacZ* expression was determined. Multiple base pair substitutions within the core ROX1 binding sequence 5'-ATTGTT-3' were introduced into each site. Plasmids carrying these mutations upstream of a *HEM13-lacZ* fusion were introduced into a wild-type strain and Δ*rox1* strain and β-galactosidase expression from the plasmids was determined. Mutating site R2 with a core sequence of 5'-TTTGTT-3' had no effect on expression. However, expression was increased 1.3-1.5-fold and 3.6-3.8-fold when R1 and R3 were mutated, respectively (Amillet *et al.*, 1996). Unlike R2, both R1 and R3 contain sequences that are identical to the core consensus sequence 5'-ATTGTT-3'

derived from alignment of operator sequences from other ROX1-regulated genes. The effect of mutating both R1 and R3 in the same construct was additive and resulted in a 6-7-fold derepression of *HEM13-lacZ* expression. Elimination of both of these sites resulted in the same derepression seen when expression of *HEM13-lacZ* was tested in a strain where the trans-acting factor (ROX1) is genetically eliminated. Thus, sites R1 and R3 are negative genetic elements which confer heme-mediated repression to *HEM13*.

8) Trans-acting factors binding to the HEM13 UAS

Deletion studies demonstrated that *HEM13* sequences required for basal level expression and full induction in the absence of heme are located between nucleotides -592 and -259 upstream of the transcription initiation site. Electrophoretic mobility shift assays employing a 190 bp probe encompassing part of the sequences between -592 and -259 and protein extracts that were prepared from wild type cells grown under aerobic, anaerobic, or heme-deficient conditions, yielded a single complex in all cases. The same complex, complex A, was also seen when extracts were prepared from Δ*hap1*, Δ*hap1*Δ*rox1*, and Δ*rox1* mutant cells suggesting that neither HAP1 nor ROX1 plays a role in the formation of this complex (Amillet *et al.*, 1996). Binding assays with overlapping DNA fragments suggested that the protein binding site was located around position -259 upstream of the transcription start site. The protected region from DNase I footprinting was consistent with formation of complex A within this region. The identity and function of complex A awaits further characterization.

9) Trans-acting factors involved in HEM13 repression via hypoxic operator sites

To date, a number of hypoxic genes which are negatively regulated by oxygen and heme have been reported. These genes all contain one or more copies of the consensus hypoxic operator sequence 5'-CCCATTGTTCTC-3'. These sites are known to be involved in repression of *ANB1* and *HEM13* expression by oxygen and heme (Lowry *et al.*, 1990; Amillet *et al.*, 1996). The hypoxic operator sites found in the gene coding for the mitochondrial ADP/ATP translocator, *AAC3*, have also been demonstrated to be required for repression of this gene (Sabova *et al.*, 1993). Mutation analysis of the operator sites of other hypoxic genes will conclusively demonstrate the importance of these elements in repression.

Evidence implicating ROX1 as the *trans*-acting factor that interacts with these hypoxic operator sites has been obtained for the *ANB1* gene (Balasubramanian *et al.*, 1993). ROX1 binds specifically to an oligonucleotide containing the consensus sequence 5'-CCCATTGTTCTC-3'. Three sites with the consensus operator sequence were found in *HEM13* by alignment with the regulatory regions of other heme-repressed genes. Mutagenesis of two of these sites result in increased *HEM13* expression (Amillet *et al.*, 1996). Therefore, these sites are ideal candidates for binding by ROX1 protein. The binding of ROX1 to the hypoxic sites within the *HEM13* promoter is the subject of data presented in chapters 5 and 6 of this thesis.

VIII. Focus of the present study

The focus of this work describes the cloning and examination of the regulation of expression of *HEM6*, the gene coding for uroporphyrinogen decarboxylase. The mechanisms concerning the mode of expression of this gene are unclear, and data obtained in this study may provide some insight as to how this gene is regulated. Also, we directly address the importance of cysteine residues within the catalytic site of uroporphyrinogen decarboxylase by analysis of mutant alleles of *HEM6* in which these cysteine residues are altered.

HEM13 expression is repressed in the presence of heme by the product of the *ROX1* gene. We wish to examine the mechanism by which ROX1 regulates expression of *HEM13*. Other hypoxic genes repressed by heme and ROX1, contain one or more copies of a hypoxic consensus operator sequence. We propose to locate the functional hypoxic consensus operator sequences located within the regulatory sequences of *HEM13*. ROX1 binding to these operator sequences will be directly tested. In addition, the affinity of ROX1 for different operator consensus sequences will be examined and correlated with the effect of deletion of the operator sites on *HEM13* expression. We also propose to delineate the domains of ROX1 required for DNA-binding and repression. Finally, the role of accessory factors such as SSN6 and TUP1 in ROX1-mediated repression of *HEM13* will be examined.

CHAPTER 2. MATERIALS AND METHODS

I. Cell cultures

1. Strains

Saccharomyces cerevisiae strains BWG1-7a (*MAT***a** *leu2-3 leu2-112 his4-519 ade1-100 ura3-52)*, BWG9a-1 (*MATα ade6 his4-519 ura3-52)* and DBY746 (*MATα his3Δ1 leu2-3 leu2-112 ura3-52 trp1-289*) were used as the wild type *HEM6* strains. BWG1-7a was the parent of the *hem6* mutant strains 1B and 17B while BWG9a-1 was the parent strain of the *hem6* mutant strain 13. BWG1-7a and the isogenic *hem1* disrupted strain 1-7ahem1::LEU2 (Prezant *et al.*, 1987) were used for studies on the regulation of *HEM6* expression. LGW1 is isogenic to BWG1-7a but contains the *hap2-1* mutation (Guarente *et al.*, 1984). Strains DBY746hem6::URA3 and DBY746Δhem1 were derived from strain DBY746. DBY746hem6::URA3 contains a *hem6::URA3* disruption integrated at the *HEM6* locus and was derived by a one step gene disruption method. DBY746 was transformed with a linear 3.1 kb *Sph*I-*Ssp*I fragment of DNA from pAMH6::URA3 which contains the *HEM6* coding region interrupted by the *URA3* gene. Strain CDF22 was derived from strain TKY22 (*MAT***a** *leu2-3 leu2-112 ura3-52 ade1 trp1::hisG Δhem1)* (Keng, 1992) and contains an integrated *HEM13-lacZ* fusion. Plasmid YIpHEM13-lacZ containing a *HEM13-lacZ* fusion and *TRP1* marker was linearized with *Xba*I and introduced into TKY22 by integrative transformation to yield strain CDF22. Strains CDF23, CDF24, and CDF25 were derived from strain CDF22 (*MAT***a** *leu2-3 leu2-112 ura3-52 ade1 Δhem1 HEM13-lacZ).* CDF23 contains a

rox1::LEU2 disruption integrated at the *ROX1* locus and was derived by a one step gene disruption method by transformation of CDF22 with a linear 5.0 kb *Bam*HI-*Sph*I fragment of DNA from pAM rox1::LEU2 in which 1.6 kb of *ROX1* coding sequence was replaced with the *LEU2* gene. CDF24 contains a *ssn6::LEU2* disruption at the *SSN6* locus. The yeast strain CDF24 was constructed by one-step gene replacement of the chromosomal copy of *SSN6* of strain CDF22 with a 3.3 kb *Hin*dIII-*Pst*I fragment from pCK36 in which the *SSN6* gene is disrupted with the entire *LEU2* coding region. CDF25 contains a *tup1::LEU2* disruption at the *TUP1* locus and was constructed by replacement of the chromosomal copy of *TUP1* of strain CDF22 with a 6.4 kb *Pvu*II fragment from plasmid pDSB in which the *TUP1* gene is disrupted with the entire *LEU2* coding region. Correct integration of the disrupted alleles was verified by Southern blot analysis of chromosomal DNA.

Escherichia coli strains MC1061 and DH5α were used for plasmid propagation. *Escherichia coli* strain SG935, containing a mutation in the Lon protease, was used for expression of GST-ROX1 fusion proteins.

2. Media and growth conditions

Saccharomyces cerevisiae cells were grown at 30°C in Yeast Peptone Dextrose (YPD) Medium or Yeast Minimal Medium (YMM) as described (Sherman *et al.*, 1986). YPD media consisted of 1% yeast extract, 2% Bacto-peptone, whereas YMM media consisted of 0.67% yeast nitrogen base without amino acids. All supplements and amino acids were added to a final concentration of 40 μg/ml. Glucose or lactate was added as a

carbon source to a final concentration of 2% when required. Media were solidified by addition of 2% agar.

The *hem1* deletion strains are defective in the enzyme δ-aminolevulinate (δ-ALA) synthase and cannot synthesize δ-ALA. Therefore, *hem1* deletion strains were grown in media supplemented with 50 μg/ml of δ-ALA to mimic heme-sufficient conditions, and 0.5% (vol/vol) Tween 80, 20 μg/ml ergosterol, and 40 μg/ml methionine (TEM) to mimic heme-deficient conditions. Tween 80-ergosterol was made up as a solution of 0.2% (wt/vol) ergosterol in Tween 80-ethanol (1:1). The *hem1* deletion strain 1-7ahem1::LEU2 and isogenic strains cannot grow in TEM, and were grown in 0.5 μg/ml δ-ALA to mimic heme-deficient conditions. When required, heme deficient strains 1B, 17B, and 13 were grown in minimal or complete media supplemented with heme, which was made up in 0.1 M NaOH and added to a final concentration of 50 μg/ml.

E. coli strains were propagated at 37°C in LB medium consisting of 1% Bacto-tryptone, 0.5% yeast extract, 1% NaCl. Ampicillin, when required, was added at a concentration of 100 μg/ml. For growth on solid media, Bacto-agar was added at 3%.

3. Yeast and bacterial transformation

Yeast cells were transformed using the spheroplast method (Hinnen *et al.*, 1978) or rendered competent with lithium acetate (Ito *et al.*, 1983). Bacterial transformations were performed as described (Hanahan, 1983).

4. Characterization of heme auxotrophs

Three mutants, 1B, 17B and 13, were isolated and characterized in this study as follows: strains BWG1-7a and BWG9a-1 were mutagenized with ethyl methanesulfonate as described (Sherman *et al.*, 1986). Colonies that were heme auxotrophs of opposite mating types were first crossed and divided into different complementation groups. The precise block in the pathway of each complementation group was determined. The mutants were divided into "early" and "late" mutants. Early mutants were blocked in steps needed for synthesis of uroporphyrinogen III. Because this molecule is needed for the synthesis of siroheme, the prosthetic group for sulfite reductase, mutants that fail to make this compound are auxotrophic for methionine (Mattoon *et al.*,1979). All mutants blocked after uroporphyrinogen III synthesis are methionine prototrophs. The mutants that were methionine prototrophs were deduced to contain blocks in the heme biosynthetic pathway after uroporphyrinogen III synthesis (Mattoon *et al.*, 1979). These mutants were used to extract porphyrin intermediates which accumulated as a result of the mutation in one of the enzymatic steps. These intermediates were then extracted and separated and identified by thin layer chromatography on Silica Gel G plates (Gollub *et al.*, 1977). During this process, uroporphyrinogen III, coproporphyrinogen III and protoporphyrinogen IX are respectively converted to uroporphyrin III, coproporphyrin III and protoporphyrin IX. Mutants 1B, 17B and 13 were found to belong to the same complementation group and were also found to accumulate uroporphyrin III but not coproporphyrin III or protoporphyrin IX (data not shown). These mutants were thus deduced to be blocked in the uroporphyrinogen decarboxylase step in the heme biosynthetic pathway.

II. Nucleic acids

1. Extraction and enzymatic manipulations of DNA

Yeast genomic DNA was isolated as described (Sherman *et al.*, 1986). Small scale preparations of plasmid DNA for analytical purposes were prepared from 5 ml overnight cultures of *E. coli* by the alkaline lysis procedure. Similarly, large scale stocks of plasmid DNA were prepared from 500 ml overnight cultures of *E. coli* by centrifugation through cesium chloride gradients (Sambrook *et al.*, 1989).

Restriction endonucleases and DNA modifying enzymes were used as recommended by the suppliers (New England Biolabs, Pharmacia Canada, Boehringer-Mannheim Canada and Bethesda Research Laboratories) or according to standard procedures (Sambrook *et al.*, 1989). Gel electrophoresis, recovery of DNA fragments from agarose gels and ligations were also done following standard protocols (Sambrook *et al.*, 1989). Oligonucleotide primers were synthesized by either the Regional DNA Synthesis Laboratory, University of Calgary, Canada, the Sheldon Biotechnology Centre, McGill University, Canada, or Bethesda Research Laboratories.

2. Plasmid constructs

A. lacZ fusion constructs

Plasmid YEpH6-lacZ was constructed by inserting a 1.06 kb *Sph*I - *Bsa*BI fragment from YEpH6 into the vector YEp357 (Myers *et al.*, 1986) which had been digested with *Sph*I and *Sma*I. The resultant plasmid, YEpH6-lacZ, contains 878 bp of upstream non-coding sequences and 183 bp of the *HEM6* coding region, corresponding to the first 61 amino acids, fused in frame to the *lacZ* gene.

B. HEM6 site-directed mutant constructs

Three independent mutations were constructed using the Muta-Gene Kit (Bio-Rad Laboratories): 1) $H6^{S26}$, which has converted cysteine 26 of *HEM6* to serine, 2) $H6^{S52}$, which has converted cysteine 52 of *HEM6* to serine, and 3) $H6^{A52}$, which has converted cysteine 52 to alanine. A 1370 base pair *Bam*HI-*Sph*I fragment from pAMH6-2.3 (Di Flumeri *et al.*, 1993) was inserted into the similarly digested phagemid vector pTZ19U. The resulting pTZH6 plasmid was used as a template for mutagenesis.

Mutagenic oligonucleotides were 5'-ATTATCCAGGATGGCGGTCTC-3' for $H6^{S26}$, 5'-GCATCCCTGGAAGTTTGAAAG-3' for $H6^{S52}$, and 5'-GCATCCCTGGCAGTTTGAAAG-3' for $H6^{A52}$ (altered codon underlined). Phagemids carrying potential mutations were screened by DNA sequence determination. Plasmids $pAMH6^{S26}$ and $pAMH6^{S52}$ were constructed by isolating a 377 base pair *Nco*I-*Bsp*DI fragment containing the desired mutation from pTZH6 and replacing the corresponding wild type fragment in pAMH6-2.3. Plasmid $pAMH6^{A52}$ was constructed by replacing an 1100 base pair *Eco*RV-*Sph*I fragment in pAMH6-2.3 with the same fragment isolated from mutagenized pTZH6. The resultant plasmids $pAMH6^{S26}$, $pAMH6^{S52}$, and $pAMH6^{A52}$ were digested with *Sph*I and *Kpn*I to release 1890 base pair fragments which were used to replace the corresponding wild type fragment in YEpH6 (Di Flumeri *et al.*, 1993). These manipulations yielded the plasmids $YEpH6^{S26}$, $YEpH6^{S52}$, and $YEpH6^{A52}$. Constructions were verified by DNA sequencing. To generate single-copy versions of these plasmids, YEpH6, $YEpH6^{S26}$ and $YEpH6^{S52}$

were digested with *Hind*III and *Sal*I. The 2.7 kb fragments were ligated to vector YCp50 cut at the unique *Hind*III and *Sal*I sites to yield YCpH6, YCpH6^{S26} and YCpH6^{S52}. YCpH6^{A52} was constructed by cloning a 2.3 kb *Eco*RI-*Hind*III fragment of pAMH6^{A52} into unique *Eco*RI and *Hind*III sites of YCp50.

C. In vitro transcription constructs

pAMH6-2.3 contains the entire *HEM6* coding region on a 2.3 kb *Sph*I-*Bgl*II fragment cloned into the unique *Sph*I and *Bam*HI sites of pAM18. Plasmid pSPROX1, used for *in vitro* transcription of *ROX1*, was generated by polymerase chain reaction using the following primers: 5'-GGAATTCATGAATCCTAAATCCTC-3' (primer 1), which introduces an *Eco*RI site (underlined) immediately upstream of the *ROX1* initiation codon, and 5'-GGTAGTCCACTTAAAGATCTGG-3' (primer 2), which corresponds to sequences within the *ROX1* open reading frame (ORF) and contains a *Bgl*II site (underlined). The amplified product was digested with *Eco*RI and *Bgl*II and ligated into the similarly digested plasmid pmini-ROX1. The pmini-ROX1 plasmid was constructed by inserting a 2.0 kB *Xba*I-*Hin*dIII fragment with the entire *ROX1* coding region into the transcription vector pAM18. Inserting the amplified product into the *Eco*RI and *Bgl*II sites of pmini-ROX1 places the *ROX1* AUG codon in close proximity to the SP6 promoter. The ligated product, pSPROX1, contains an SP6 promoter directly upstream of the *Eco*RI site which marks the beginning of the *ROX1* coding region.

D. Disruption constructs

Plsmid pAMH6::URA3 was constructed by ligation of a 1.2 kb Klenow-treated *Hin*dIII fragment from Yep24 containing the entire *URA3* coding region into vector pAMH6 linearized with *Eco*RV. This resulted in insertion of the coding region of *URA3* within the *HEM6* coding region. Plasmid pAM rox1::LEU2 was constructed by ligation of a 2.0 kb *Hpa*I-*Sal*I blunt-ended fragment from plasmid pAA101 containing the entire *LEU2* coding region into vector pAMROX1 *Bam*HI-*Sph*I digested with *Bst*EII-*Eco*RV and treated with Klenow fragment. This replaced 1.6 kb of *ROX1* coding sequences with a 2.0 kb fragment containing the entire *LEU2* marker gene. Plasmids pDSB and pCK36 were used for *SSN6* (*CYC8*) and *TUP1* disruptions, respectively, and were kind gifts from R. Trumbly (Medical College of Ohio). Both pDSB and pCK36 contain the entire *LEU2* coding region inserted within the *SSN6* and *TUP1* coding regions, respectively.

E. GST-ROX1 fusion constructs

Plasmid pGEX-ROX1 Full Length (FL) was constructed by cloning an *Eco*RI-*Hin*dIII fragment containing the entire *ROX1* coding region from the plasmid pSPROX1 into the *Eco*RI and *Aat*II sites of a pGEX-2T plasmid (Pharmacia) modified to contain five extra nucleotides before the *Eco*RI restriction site. Plasmid pGEX-ROX1 176-368 was constructed by digesting pGEX-ROX1 FL with *Bam*HI and *Bgl*II and religating the vector, thereby generating an in-frame fusion of glutathione-S-transferase (GST) to the carboxy-terminal half of the *ROX1* ORF. Plasmid pGEX-ROX1 1-176 was constructed by digesting pGEX-ROX1 FL at the unique *Bgl*II site and blunt ending with the Klenow fragment of DNA polymerase. The blunt-ended DNA was then circularized by ligation.

This plasmid produces a GST fusion protein containing the first 176 amino acids of ROX1 followed by 9 amino acids coded by an alternative reading frame before a stop codon is reached. Plasmid pGEX-ROX1 1-101 was generated by polymerase chain reaction using primer 1 and 5'-GCGGATCCTCACTCGATTTCCTTCAA-3' (primer 3), which introduces an in-frame stop codon just after codon 101 in ROX1. The amplified product was cloned into the PCRII vector using the TA Cloning Kit (In Vitrogen). An *Eco*RI-*Bam*HI fragment from this plasmid was subsequently ligated into the *Eco*RI-*Bgl*II digested vector pGEX-ROX1 FL, generating an in-frame fusion of GST to the amino-terminal 101 amino acids of ROX1. Plasmid pGEX-ROX1 1-50 was generated as follows: plasmid pGEX-ROX1 FL was digested with *Bam*HI and *Bgl*II. The *Bam*HI-*Bgl*II fragment was purified and digested with *Rsa*I, which cleaves the DNA 150 bp downstream from the *ROX1* start codon, generating a blunt-ended fragment. The purified *Bam*HI-*Rsa*I fragment was then ligated into the plasmid pGEX2T cleaved with *Bam*HI and *Sma*I. This construction resulted in an in-frame fusion of GST to the amino-terminal 50 amino acids of ROX1.

F. Operator site deletion constructs

Deletions of each of the five possible negative regulatory sites were constructed with the Mutagene Kit (Bio-Rad), using oligonucleotides which contained sequences flanking either side of each of the hypoxic sites. Plasmid YCp13Z was used as a parental plasmid to construct deletions of the five consensus operator sites. This plasmid contains 1450 base pairs (bp) of upstream non-coding sequence and 34 codons of *HEM13* fused to the *lacZ* gene of *E. coli* on a centromeric plasmid (Richard, 1990). An

800 bp *Nru*I-*Bam*HI fragment containing the five putative hypoxic operator sites from YCp13Z was isolated and ligated into the *Hinc*II-*Bam*HI sites of the phagemid vector pTZ18U. Uracil-containing single-stranded DNA from the resulting pTZHEM13 plasmid was used as a template for mutagenesis. Site 1 was deleted using oligonucleotide ROX1Δ1, complementary to the 15 bp upstream and 12 bp downstream of site 1. Site 2 was deleted using the oligonucleotide ROX1Δ2, complementary to the 12 nucleotides upstream and 12 nucleotides downstream of the site. Site 3 was deleted with oligonucleotide ROX1Δ3, complementary to the 12 bp upstream and 12 bp downstream of the site. Site 4 was deleted using the oligonucleotide ROX1Δ4, complementary to the 11 bp upstream and 13 bp downstream of the site. Site 5 was deleted with oligonucleotide ROX1Δ5, complementary to the 12 bp upstream and 12 bp downstream of the site. DNA sequencing of single-stranded DNA derived from the respective mutagenized plasmids was carried out to ensure proper deletion of the hypoxic operator sites.

The mutated sequences were introduced into *HEM13-lacZ* fusion plasmids by cutting double-stranded pTZHEM13 containing the respective mutations with *Sph*I, treating the DNA with Klenow fragment and then digesting with *Bam*HI. The 800 bp fragment for each individual mutation was isolated and ligated to the 10 kb fragment isolated from plasmid pBL101 that had been digested with *Sma*I and *Bam*HI to remove all *CYC1* sequences. The resulting plasmids, YCp13ZΔ1, YCp13ZΔ2, YCp13ZΔ3, YCp13ZΔ4 and YCp13ZΔ5, each contained a precise deletion of the 12 base pairs corresponding to hypoxic operator sites 1 through 5 respectively.

YCp13ZΔ2Δ3, YCp13ZΔ3Δ4, and YCp13ZΔ1Δ5 were generated respectively by mutagenizing pTZHEM13 deleted for site 2, 3, or 1 with oligonucleotides ROX1Δ3, ROX1Δ4, or ROX1Δ5. YCp13ZΔ3Δ4Δ5 was made by mutagenizing pTZHEM13 deleted for sites 3 and 4 with oligonucleotide ROX1Δ5. In each case, the mutagenized fragments were sequenced to confirm that they contained the appropriate deletions and were individually introduced into the pBL101 backbone to generate a *HEM13-lacZ* fusion as previously described. To generate plasmid YCp13ZΔ1Δ2, YCp13ZΔ1 was digested with *Sac*I and the 8.5 kb fragment was ligated to the 2.3 kb *Sac*I fragment isolated from YCp13ZΔ2. YCp13ZΔ1Δ3 and YCp13ZΔ1Δ4 were constructed in the same manner, except that the 2.3 kb *Sac*I fragments isolated from YCp13ZΔ3 or YCp13ZΔ4 were used. Similarly, YCp13ZΔ3Δ5 was constructed using the *Sac*I backbone from YCp13ZΔ5 and the *Sac*I fragment from YCp13ZΔ3. YCp13ZΔ1Δ2Δ3 and YCp13ZΔ1Δ3Δ4, each containing deletions of 3 hypoxic operator sites, were constructed with the 8.5 kb *Sac*I backbone from YCp13ZΔ1 and the 2.3 kb *Sac*I fragment from either YCp13ZΔ2Δ3 or YCp13ZΔ3Δ4. YCp13ZΔ1Δ3Δ5 was constructed using the *Sac*I backbone from YCp13ZΔ1Δ5 and the *Sac*I fragment from YCp13ZΔ3.

G. In vivo ROX1 deletion constructs

Plasmid pRSROX1 was constructed by digestion of plasmid pAMROX1 *Bam*HI-*Sph*I with *Hin*dIII which released a 2.8 kb fragment of DNA containing the entire *ROX1* open reading frame. This 2.8 kb fragment was subsequently cloned into the *Hin*d III site of the yeast expression plasmid pRS316 (Sikorski and Hieter, 1989) to yield plasmid

pRSROX1. This plasmid contains an ARS-CEN sequence and *URA3* selectable marker. Plasmid pRSROX1 1-176 was constructed by digesting pAMROX1 *Bam*HI-*Sph*I with *Bgl*II, treated with Klenow enzyme to generate blunt ends, and digested with *Hin*dIII to release a 1.8 kb fragment. This 1.8 kb fragment was ligated to vector pRS316 which was digested with *Bgl*II, treated with Klenow enzyme to generate blunt ends, and digested with *Hin*dIII. This resulted in a plasmid containing the coding region for the first 176 amino acids of the ROX1 protein. Plasmid pRSROX1 1-286 was generated by ligating a 2.1 kb fragment obtained by digestion of pAMROX1 *Bam*HI-*Sph*I with *Ssp*I and *Hin*dIII into the similarly digested vector pRS316. This plasmid contains sequences coding for the first 286 amino acids of the ROX1 protein.

3. DNA sequencing and analysis

DNA was sequenced with the dideoxy-chain termination method (Sanger *et al.*, 1977) using the enzymes Sequenase (United States Biochemical Corp.) or T7 DNA polymerase (Pharmacia Canada) for chain elongation and termination. The M13 phages mp18 and mp19 were used to generate single-stranded templates for sequencing reactions. Subclones were prepared by direct cloning of specific restriction fragments into M13 RF vectors. Alternatively, sequencing of double-stranded DNA templates was carried out according to the manufacturers' specifications. Reaction products were labeled using [α-^{35}S]dATP and were run on 8% acrylamide-8 M urea gels. The DNA sequences were analyzed using the DNA Strider program (Marck, 1988), or the PC/Gene program (Intelligenetics, Mountain View, CA).

4. Northern blot analysis

Total RNA was isolated from exponentially grown yeast cells using glass bead lysis (Guarente and Mason, 1983). The RNA samples (20 μg) were separated on 1% agarose gels containing N-morpholinopropanesulfonic acid-formaldehyde buffer and transferred to Nytran membranes (Schleicher and Schuell) (Sambrook *et al.*, 1989). The hybridization and washing conditions were as described previously (Shackelford and Varmus, 1987). The *HEM6* probe was transcribed from plasmid pAMH6-2.3 which had been linearized with *Hind*III, using SP6 RNA polymerase and [α-^{32}P]-CTP. After probing for *HEM6* message, the filters were stripped and re-probed for *ACT1* mRNA. Actin RNA was detected with a RNA probe synthesized with SP6 RNA polymerase from the plasmid pSPactin which contains a 3.3 kb *Eco*RI-*Bam*HI *ACT1* fragment in pSP65. The levels of actin mRNA level served as internal controls for the amount of RNA loaded in each lane. The levels of *HEM6* and *ACT1* mRNA were quantified by densitometric scanning with a Bio-Rad Model 620 video densitometer using the I-D Analyst II data analysis software.

5. *In vitro* transcription and translation

The pSPROX1 plasmid, in which the *ROX1* coding region is placed downstream of the SP6 promoter, was used to generate transcripts which were subsequently translated *in vitro*. The ROX1 FL template was generated by digesting pSPROX1 with *Hin*dIII, while templates for the derivatives ROX1 1-286, ROX1 1-175, ROX1 1-100, and ROX1 1-58 were generated by digesting pSPROX1 with *Ssp*I, *Bgl*II, *Taq*I, and *Hpa*II,

respectively. The DNA templates were then extracted with phenol, and precipitated with ethanol. RNA was synthesized using SP6 RNA polymerase and 5 μg of each linear template in a 50 μl volume according to the manufacturer's instructions (Promega). Reactions were incubated 60 min. at 40°C and terminated by the addition of RQ1 DNase, followed by incubation at 37°C for 15 min. The reactions were extracted with phenol and the RNA precipitated with ethanol and resuspended in 10 μl of distilled water that had been treated with diethylpyrocarbonate. Aliquots of each RNA (2 μl) were then added to wheat germ extracts (Promega) together with 37.5 μCi of ^{35}S-methionine (Amersham) for *in vitro* translation. 5 μl aliquots of each sample were then analyzed on denaturing 15% SDS-polyacrylamide gels. Gels were fixed in isopropanol and acetic acid and treated with Amplify fluorographic reagent (Amersham) for 30 min. before they were dried and exposed to X-ray film.

III. Protein analysis

1. Expression and purification of GST fusion proteins

Expression of GST fusion proteins was performed essentially as described (Smith and Johnson, 1988), with some modifications. Saturated overnight cultures of *E. coli* strain SG935 containing GST-ROX1 fusion expression plasmids were diluted 10-fold in LB ampicillin medium and the cultures were grown for 90 min. at 32°C. Expression of the fusion proteins was induced with 0.15 mM isopropyl-thio-galactoside and the cultures were grown an additional 3 hrs. at 32°C. Cells were harvested by centrifugation at 5,000 rpm for 5 min. and lysed by pulse sonication of 30 seconds twice. Triton-X-100 was

added to a final concentration of 1% (v/v) and the suspension was centrifuged for 5 min. at 10,000 rpm. A 50% slurry of glutathione Sepharose beads (Pharmacia) was added to the supernatant and protein binding was allowed to proceed for 15 min. at 4°C. Following binding, the beads were washed three times with phosphate buffered saline (PBS) and the suspension was centrifuged at 500 rpm for 1 min. to collect the beads. GST fusion proteins were left bound to the glutathione Sepharose beads, or alternatively were eluted from the beads with a solution of 50 mM Tris, pH 8.0 and 20 mM glutathione. The GST portion of the fusion protein bound to glutathione Sepharose beads was removed by cleavage with thrombin. Proteins were then analyzed on a 10% or 15% SDS-polyacrylamide gel.

2. Assay for protein-protein interaction

GST or GST-ROX1 derivatives immobilized on glutathione Sepharose beads were incubated for 2 hrs. at 4°C on a rotating platform with *in vitro* translated ^{35}S-labeled ROX1 protein derivatives in binding buffer (120 mM NaCl, 50 mM Tris, pH 8.0, and 0.5% Nonidet P-40) in a total volume of 200 µl. The beads were then washed four times with 1 ml of binding buffer before bound proteins were eluted by boiling in 30 µl of SDS-polyacrylamide gel electrophoresis sample loading buffer (Sambrook *et al.*, 1989). The eluted proteins were analyzed on 10% or 15 % SDS-polyacrylamide gels.

IV. Electrophoretic mobility shift assays

1. Gel electrophoresis conditions and preparation of probes

DNA binding assays were performed in a total volume of 15 μl containing DNA binding buffer (4 mM Tris, pH 8.0, 4 mM $MgCl_2$, 100 mM KCl, 12% glycerol), 100 ng of poly dI:dC (unless otherwise stated), and 0.5-1.5 ng of ^{32}P-labeled oligonucleotides containing the ROX1 binding site. Complementary oligonucleotides were annealed and labeled by filling in recessed 3' ends with [α-^{32}P]dATP and the Klenow fragment of DNA polymerase. Binding reactions were incubated at room temperature for 30 min. before the samples were analyzed on pre-run 6% polyacrylamide gels in TBE buffer (90 mM Tris, 90 mM H_3BO_3, 2.5 mM EDTA). Gels were run at 4°C at 20 mAmps, dried, and exposed to X-ray film or phosphimager screen. Phosphoimager (Molecular Dynamics) analysis was carried out to quantitate binding levels.

2. Delineation of DNA-binding domain of ROX1

For delineation of the DNA-binding domain of ROX1, 1 μl aliquots of ROX1 derivatives that had been synthesized *in vitro*, 50-100 ng of GST-ROX1 fusion protein derivatives, or ROX1 derivatives released from fusion proteins by thrombin cleavage were employed.

Two double-stranded oligonucleotides with different sequences were used in these studies. The first, RS33, consisted of complementary oligonucleotides

5'-AATTCGCTTTGCCCATTGTTCTCGTTTCGAAAG-3' and

5'-AATTCTTTCGAAACGAGAACAATGGGCAAAGCG-3', while the other, RS32, consisted of complementary oligonucleotides

5'-GCTTGCTTTGCCCATTGTTCTCGTTTCGAAAG-3' and

5'-CTTTCGAAACGAGAACAATGGGCAAAGC-3'. In addition, RNS31, composed of complementary oligonucleotides 5'-CCGGCCGCGGTCCGACGCGTGCGCGCGACGT-3' and 5'-CGCGCGCACGCGTCGGACCGCGGCCGGAGCT-3', was used as a non-specific competitor. Both RS33 and RS32 contain ROX1 binding sequences (underlined) from -111 to -100 upstream of the transcriptional start site of *HEM13* .

3. Operator site binding affinity

For operator affinity binding experiments, 70 000 cpm of the following double-stranded ^{32}P-labeled oligonucleotides were used: operator consensus site 1was generated by annealing complementary oligonucleotides

5'- AATTCTTAATTTCAATTGTTTAGAAAGTGCCTG-3' and

5'-AATTCAGGCACTTTCTAAACAATTGAAATTAAG-3'; operator consensus site 2 was generated by annealing complementary oligonucleotides

5'-AATTCTCGCCTTTTCTGGTTCTCCCACCAATAG-3' and

5'-AATTCTATTGGTGGGAGAACCAGAAAAGGCGAG-3'; operator consensus site 3 was generated as described above for RS33; operator consensus site 4 was generated by annealing complementary oligonucleotides

5'- AATTCGTCTTATGCTTTGTTCAAGCTGGAGCGG-3' and

5'-AATTCCGCTCCAGCTTGAACAAAGCATAAGACG-3'; operator consensus site 5 was generated by annealing complementary oligonucleotides

5'-AATTCATTAGGCGTATTGTTTTTCCCAACAGTG-3' and

5'-AATTCACTGTTGGGAAAAACAATACGCCTAATG-3'.

For operator site binding affinity experiments 50 ng of poly (dI.dC) was employed as a non-specific competitor. GST and GST-ROX1 1-101 proteins were used at a concentration of 50-100 ng.

4. Minor groove binding drugs

For minor groove binding drug experiments, 70 000 cpm of double-stranded ^{32}P-labeled RS33 DNA consisting of complementary oligonucleotides

5'-AATTCGCTTTGCCCATTGTTCTCGTTTCGAAAG-3' and

5'-AATTCTTTCGAAACGAGAACAATGGGCAAAGCG-3' was used (consensus operator site underlined). Stock solutions of actinomycin D (5 mM), chromomycin A_3 (1 mM), and distamycin A (10 mM) (Sigma) were made in 10 mM Tris, pH 7.6. Stock solutions were stored at 4^{0}C and were diluted immediately before use. GST or GST-ROX1 1-101 were used at a concentration of 50-100 ng.

5. Oligonucleotide substitution

For oligonucleotide substitution experiments, 70 000 cpm of the following double-stranded ^{32}P-labeled oligonucleotides were used: RS33 as described above; IC-1, which consists of complementary oligonucleotides

5'-AATTCGCTTTGCCCITTGTTCTCGTTTCGAAAG-3' and

5'-AATTCTTTCGAAACGAGAACAACGGGCAAAGCG-3'; GC-1, which consists of complementary oligonucleotides

5'-AATTCGCTTTGCCCGTTGTTCTCGTTTCGAAAG-3' and

5'-AATTCTTTCGAAACGAGAACAACGGGCAAAGCG-3'; IC-23, which consists of complementary oligonucleotides

5'-AATTCGCTTTGCCCACCGTTCTCGTTTCGAAAG-3' and

5'-AATTCTTTCGAAACGAGAACIITGGGCAAAGCG-3'; GC-23, which consists of complementary oligonucleotides

5'-AATTCGCTTTGCCCACCGTTCTCGTTTCGAAAG-3' and

5'-AATTCTTTCGAAACGAGAACGGTGGGCAAAGCG-3'; IC-56, which consists of complementary oligonucleotides

5'-AATTCGCTTTGCCCATTGCCCTCGTTTCGAAAG-3' and

5'-AATTCTTTCGAAACGAGIICAATGGGCAAAGCG-3'; and GC-56, which consists of complementary oligonucleotides

5'-AATTCGCTTTGCCCATTGCCCTCGTTTCGAAAG-3' and

5'-AATTCTTTCGAAACGAGGGCAATGGGCAAAGCG-3'.

For oligonucleotide substitution experiments, 50-100 ng of either GST or GST ROX1 1-101 were used.

V. Random PCR-based selection of optimal ROX1 binding sites

Binding site selections were performed using a 60 base pair oligonucleotide which contained random sequence at 12 contiguous positions and was flanked on the 5' side by sequences derived from the T3 promoter and a *Sal*I site, and on the 3' side by sequences derived from the T7 promoter and a *Bam*H1 site. The oligonucleotide sequence was

5'-CTCGGAATTAACCCTCACGTCGACN$_{12}$GGATCCGTCGTATTACAATTCACT-3'

(*Sal*I and *Bam*HI sites underlined). Ten pmol of this oligonucleotide was rendered double stranded by PCR amplification using T3 and T7 oligonucleotide primers whose sequence was 5'-CTCGGAATTAACCCTCAC-3' and 5'-AGTGAATTGTAATACGAC-3', respectively. PCR cycles were as follows: 1 min at 94°C, 30 sec at 55°C, and 1 min 30 sec at 72°C for 3 cycles. 0.5 ng of the double-stranded oligonucleotide was then incubated with approximately 1 μg of purified GST or GST-ROX1 1-101 protein bound to glutathione Sepharose beads (Pharmacia) at 4°C for 20 min in ROX1 binding buffer (4 mM Tris, pH 8.0, 4 mM $MgCl_2$, 100 mM KCl, 12% glycerol) containing 10 ng/μl of poly (dI.dC) in a total volume of 100 μl. The beads were spun down and washed three times with 500 μl of lysis buffer (0.5% NP-40, 120 mM NaCl, 50 mM Tris, pH 8). DNA that remained bound to beads containing GST or GST-ROX1 1-101 was then used for PCR amplification using the same conditions as above except that amplification proceeded for 20 cycles. The binding, washing, and amplification procedure was repeated three more times for a total of four selections. After the fourth selection, the PCR products were digested with *Sal*I and *Bam*HI and cloned into the same sites in the plasmid Bluescript SK+ (Stratagene). Double-stranded sequencing of the inserts was performed as per the manufacturer's instructions (Pharmacia) using the T3 primer.

VI. Enzymatic assays

1. β-galactosidase assays

Assays were carried out on cultures that were grown to an optical density (OD_{600}) of approximately 1. The cells were permeabilized with sodium dodecyl sulfate and chloroform as described (Guarente, 1983). β-galactosidase activity is reported in Miller Units. All assays were performed in duplicate on two independent transformants of each plasmid, and the results are averages of at least four determinations. The results for each of the four determinations for each plasmid were within 20% of each other.

CHAPTER 3

MOLECULAR ANALYSIS OF *HEM6* (*HEM12*) IN *SACCHAROMYCES CEREVISIAE*, THE GENE FOR UROPORPHYRINOGEN DECARBOXYLASE

Celestino Di Flumeri, Robert Larocque, and Teresa Keng

Department of Microbiology and Immunology, McGill University,
3775 University Street, Montreal, Quebec, Canada, H3A 2B4

This chapter is derived from a manuscript published
in *Yeast*, **9**, 613-623 (1993).

ABSTRACT

HEM6 (*HEM12*) in *Saccharomyces cerevisiae* encodes uroporphyrinogen decarboxylase, the fifth enzyme in the heme biosynthetic pathway. The *HEM6* (*HEM12*) gene was cloned by complementation of heme auxotrophy of a *hem6* mutant. Sequence analysis revealed an open reading frame of 1086 nucleotides. The predicted amino acid sequence of *HEM6* (*HEM12*) shows extensive homology to those reported for uroporphyrinogen decarboxylase from mammalian sources. Expression of *HEM6* (*HEM12*) was investigated and was found to increase two-fold in a non-fermentable carbon source. However, *HEM6* (*HEM12*) transcription was unaffected by heme or by intermediates in the heme biosynthetic pathway. In addition, *HEM6* (*HEM12*) expression is not regulated by the transcriptional activator complex HAP2-3-4, as has been shown for some genes encoding heme biosynthetic enzymes.

INTRODUCTION

In the yeast *Saccharomyces cerevisiae*, heme biosynthesis involves eight enzymatic steps which convert glycine and succinylCoA to heme (reviewed in Labbe-Bois and Labbe, 1990). Uroporphyrinogen decarboxylase (EC 4.1.1.37), also known as porphyrinogen carboxy-lyase, is a cytosolic enzyme which catalyzes the fifth step of the heme biosynthetic pathway. This enzyme sequentially removes the four carboxyl groups from the acetate side chains of uroporphyrinogen to yield coproporphyrinogen. The enzyme has been partially purified from bacterial (Hoare and Heath, 1959), flagellate (Juknat *et al.*, 1989), yeast (Felix and Brouillet, 1990), avian (Tomio *et al.*, 1970), and mammalian (Romeo and Levin, 1971; De Verneuil *et al.*, 1983; Elder *et al.*, 1983; Straka and Kushner, 1983; Mukerji and Pimstone, 1987) sources. The purified enzymes, with the exception of that from chick, are all monomers with relative molecular masses of 40 - 50 kDa. The cDNAs for the rat and human genes have been cloned and sequenced (Romeo *et al.*, 1984; Romeo *et al.*, 1986; Romana *et al.*, 1987). The deduced amino acid sequences of these two mammalian enzymes are strikingly conserved, showing 90% identity.

Several yeast mutants which are defective in heme biosynthesis have been isolated (Gollub *et al.*, 1977; Urban-Grimal and Labbe-Bois, 1981; Rytka *et al.*, 1984; Kurlandzka *et al.*, 1988). In particular, several mutants which are defective in uroporphyrinogen decarboxylase have been characterized. In an initial screen for mutants which accumulated high concentrations of porphyrins, *pop3* mutants were isolated by Pretlow and Sherman (1967). Subsequently, a *hem6* mutant was isolated by Gollub *et al.* (1977), as a heme auxotroph which accumulated uroporphyrin III. The *hem6* mutant was

deduced to be defective in uroporphyrinogen decarboxylase. Using this protocol designed by Gollub *et al.*, we have isolated more *hem6* mutants (Keng *et al.*, 1992) The *hem6-1* mutant was subsequently shown to be allelic to *pop3* (Arrese *et al.*, 1982). Mutants which are not heme auxotrophs but are defective in uroporphyrinogen decarboxylase activity have been isolated in other independent studies (Urban-Grimal and Labbe-Bois, 1981; Rytka *et al.*, 1984; Kurlandzka and Rytka, 1985). Urban-Grimal and Labbe-Bois (1981) isolated a *hem12-1* mutant strain that was respiration-deficient, lacked cytochromes and catalase activity and that accumulated high levels of porphyrins. Subsequent studies identified other *hem12* mutants that were UV-fluorescent in ethanol-containing medium due to accumulated porphyrins. All *hem12* mutants were shown to have either decreased or undetectable levels of uroporphyrinogen decarboxylase activity (Rytka *et al.*, 1984; Kurlandzka *et al.*, 1988). *HEM12* has been mapped to chromosome IV and is located between *TRP1* and *HEM13* (Kurlandzka *et al.*, 1988). Although it is likely that *HEM6* and *HEM12* both encode uroporphyrinogen decarboxylase and are allelic, this has not been demonstrated.

In yeast, heme plays an important role in the regulation of gene expression. Heme is required for the expression of aerobically induced genes such as *CYC1*, encoding iso-1-cytochrome *c* (Hortner *et al.*, 1982; Guarente and Mason, 1983), and for the repression of hypoxic genes such as *ANB1* which encodes an isoform of eIF-5A (Lowry and Zitomer, 1984). The intracellular levels of heme are believed to increase in response to the presence of oxygen or a non-fermentable carbon source (Mattoon *et al.*, 1979). Such changes in heme levels can mediate the effects of oxygen and carbon source on gene expression. Thus, it becomes important to study how heme levels and, in particular, how

expression of genes encoding heme biosynthetic enzymes are regulated.

We have begun an analysis of the genes encoding various enzymes of the heme biosynthetic pathway in yeast. We are particularly interested in the regulation of expression of these genes. Expression of *HEM1*, encoding δ-aminolevulinate (δ-ALA) synthase, the first enzyme in the pathway and *HEM3*, encoding porphobilinogen deaminase, was found to be unregulated by heme levels or by carbon source (Keng and Guarente, 1987; Keng *et al.*, 1992). However, expression of both these genes was found to be dependent upon the HAP2-3-4 transcriptional activator complex. Expression of *HEM13*, encoding coproporphyrinogen oxidase, was found to be repressed by heme (Zagorec *et al.*, 1988) and its regulation was found to involve both the *HAP1* and *ROX1* gene products (Verdiere *et al.*, 1991; Keng, 1992).

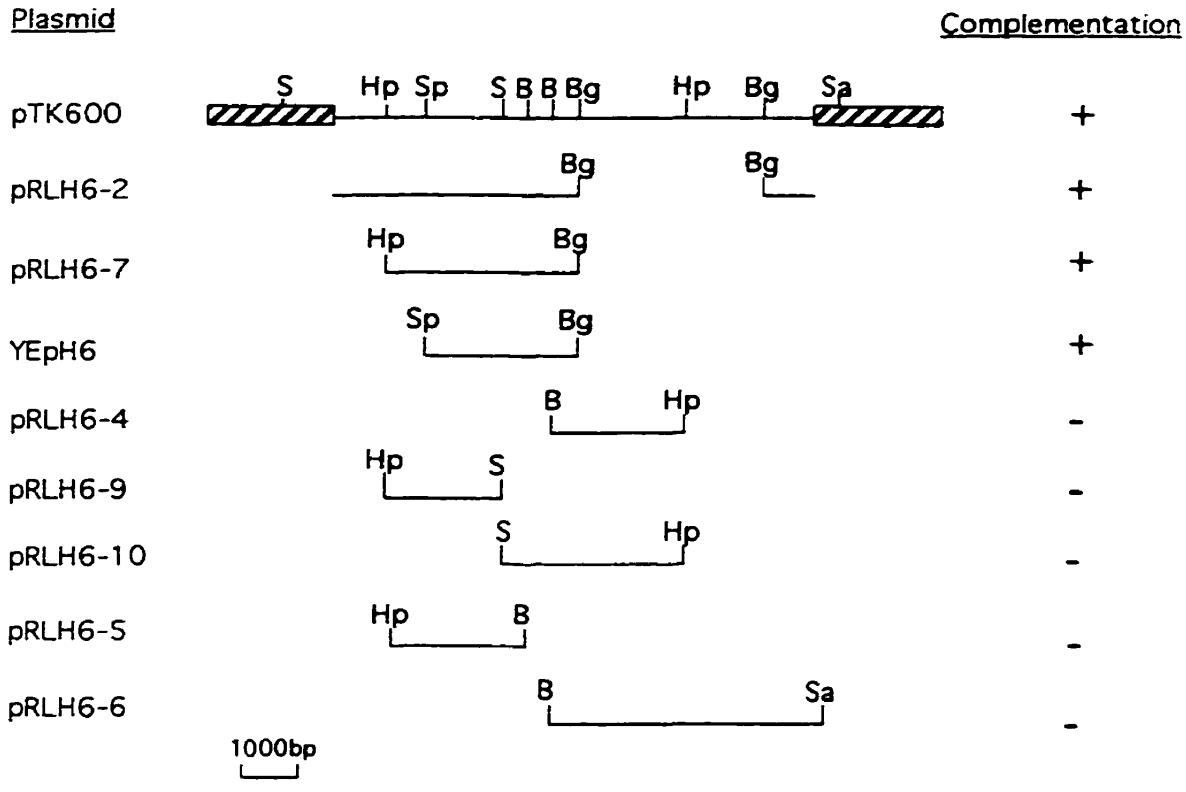
We report here the sequence analysis and the regulation of expression of the yeast *HEM6* gene encoding uroporphyrinogen decarboxylase. Our analysis indicates that transcription of *HEM6* is not affected by heme but is induced approximately two-fold by a non-fermentable carbon source. Expression is not affected by the *hap2* mutation.

RESULTS

Isolation of *HEM6*

The *HEM6* gene was isolated by complementation of heme auxotrophy of the *hem6* mutant strain 13. This mutant strain was transformed with 1 μg each of three plasmid pools representing yeast genomic DNA cloned into the yeast shuttle vector YEp24 (Carlson and Botstein, 1982). Approximately 10^3 -10^4 transformants were obtained per μg of DNA and approximately 0.1 - 1% of the transformants had become heme prototrophs. Loss of plasmid, as indicated by loss of uracil prototrophy, was found to be always associated with loss of heme prototrophy. The same plasmid, pTK600, was isolated from a number of transformants of strain 13 that exhibited heme prototrophy. pTK600 was found to contain an insert of 7.3 kb and was also found to be able to complement the heme auxotrophy in two other *hem6* mutant strains, 1B and 17B. A restriction map of the pTK600 plasmid is shown in figure 6. In order to localize the *HEM6* coding region within the 7.3 kb insert, pTK600 was used to provide restriction fragments for the construction of various subclones (Figure 6). In addition, a deletion derivative (pRLH6-2) of pTK600 was constructed in which a 2.8 kb *Bgl*II fragment within the insert was removed. Both the subclones and deletion derivatives were separately transformed into the *hem6* mutant strains 1B and 17B and tested for their abilities to complement the heme auxotrophy of these strains. Two plasmids were found that could complement the heme auxotrophies of both *hem6* mutants, pRLH6-7 which contains a 3.0 kb *Hpa*I-*Bgl*II fragment and YEpH6 which contains a 2.3 kb *Sph*I-*Bgl*II fragment. The deletion plasmid and the other subclones were unable to confer heme

Figure 6. Localization of *HEM6* within the plasmid pTK600. A partial restriction map of the 7.3 kb yeast genomic DNA in plasmid pTK600 is shown. All subclones were constructed in vector YEp24 and transformed into *hem6* mutant strains 1B and 17B. One deletion plasmid, pRLH6-2, was also constructed. All plasmids were tested for their ability to complement the heme auxotrophy of 1B and 17B (+ indicates complementation; - indicates lack of complementation). B, *Bam*HI; Bg, *Bgl*II; Hp, *Hpa*I; S, *Stu*I; Sa, *Sal*I; Sp. *Sph*I.


Plasmid
Complementation
pTK600
S
Hp
Sp
S
B
B
Bg
Hp
Bg
Sa
+
pRLH6-2
Bg
Bg
+
pRLH6-7
Hp
Bg
+
YEpH6
Sp
Bg
+
pRLH6-4
B
Hp
-
pRLH6-9
Hp
S
-
pRLH6-10
S
Hp
-
pRLH6-5
Hp
B
-
pRLH6-6
B
Sa
-
1000bp

prototrophy to the 1B and 17B strains. Based on this analysis, the *HEM6* gene was localized to a 2.3 kb region between the *Sph*I and *Bgl*II sites. In addition, a strain containing the deleted/ disrupted allele of the cloned gene failed to complement the *hem6* mutant strains 1B and 17B, providing further evidence that the cloned gene is *HEM6* (data not shown).

Nucleotide Sequence of *HEM6*

The nucleotide sequence of a 2.1 kb region between the *Sph*I and *Bgl*II sites was determined on both strands as described in Materials and Methods. Figure 7 illustrates the direction and extent of sequencing on the two DNA strands. Within the 2.1 kb region, a single open reading frame of 1086 bp, beginning at position +1 with an ATG codon and ending at position 1087 with an stop codon, was found (Figure 8). The open reading frame could potentially encode a protein of 362 amino acids with a predicted molecular weight of 41,322 Daltons, a value that is consistent with what had been observed for purified yeast uroporphyrinogen decarboxylase (Felix and Brouillet, 1990). The codon adaptation index of 0.16 for the *HEM6* coding region suggests that the enzyme is expressed at low levels (Sharp and Li, 1987).

Analysis of the 5'-flanking region revealed the presence of a single potential TATA-like sequence, TATATA, at position -268 upstream of the initiation codon. We also found the sequence CAATTGGT between positions -566 and -559. This sequence differs by only one base pair from the consensus HAP2-3-4 binding sequence, TNATTGGT (N can be any nucleotide) (Forsburg and Guarente, 1988). The region downstream of the *HEM6* stop codon contains the tripartite termination sequence

Figure 7. Sequencing strategy of the *HEM6* coding region and the flanking regions. The hatched box represents the open reading frame associated with *HEM6* , beginning with an ATG codon and terminating with a TAA codon. The arrows denote the direction and extent of sequencing. Approximately 90% of the total sequence and 95% of the sequence corresponding to the open reading frame was determined on both strands.

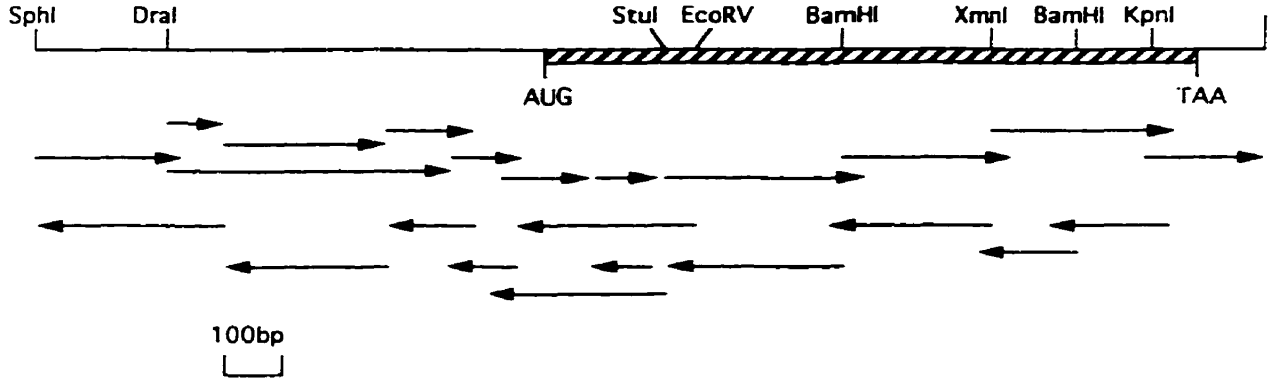

SphI
DraI
StuI
EcoRV
BamHI
XmnI
BamHI
KpnI
AUG
TAA
100bp

Figure 8. Nucleotide sequence of *HEM6* and the translated amino acid sequence. The DNA sequence of the *HEM6* coding region is represented by upper case letters, and begins with an ATG codon at position +1. The star denotes the stop codon at position +1087. The deduced amino acid sequence of *HEM6* is shown below the DNA sequence. The presumptive TATA box at position -268 and the tripartite termination sequence beginning at position +1122 are underlined.

```
gcgatatgtgtcggccggaaaaaaaaagcacgcatccggcaagcacactagtaagtctaa -781
tgcaggatgaagaaacgaagaaggtgccatgagaaggaaaataaaactcaatgggacgga -721
acggcagcatttttaacaaattattttggattcatccgtaacaatataacttggtacgg -661
gaacggcatgttttaaatgcacagcggtgccttgtttcgccgtggattgttaaaggctct -601
ttttcggacgccacatattcctatgctgcaccagcaattggtgtcttcggatcatactac -541
ttcgcactagtttatattcaaggcttttaatggagcttcgctcttaaaagaaaaaacttc -481
tttacttccatgcgcgcgtttttggcgtacgtagcactgcgaaagagtgggtgatatata -421
gagaacatgttgtagtatacgtacgcgggcaattgtggtgacagtatctctttctgttag -361
caagtaatatatccttggcctcaacctgtcaatagtgttagtttatccgccatcttggcg -301
ctatcatacacgatctttttcaatgtttctcgtatataacaggacagggttcggggtttttt -241
ggcggcaggagaggaaagcaaaaaaaaaatgaggcaagagttttaattttgccagcgaaat -181
gttctgccattttcatctttaccgctcgaattagtgatgttccagacgaaggtaacatag -121
agtgatcgataggttcttcaattcatcaagcattatctcagaattgtgtttggtaagctc -61
tcttcgagaaaaggaagaacaaaaaaaaaaaaaaataaagcaaagaacaggttataacgct -1

ATGGGTAACTTTCCAGCTCCAAAAAACGATTTGATATTGAGAGCCGCAAAGGGTGAAAAA +60
 M  G  N  F  P  A  P  K  N  D  L  I  L  R  A  A  K  G  E  K
GTCGAGAGACCGCCATGCTGGATAATGCGCCAGGCAGGTCGTTACCTGCCGGAATATCAC +120
 V  E  R  P  P  C  W  I  M  R  Q  A  G  R  Y  L  P  E  Y  H
GAGGTGAAAAACAATCGTGATTTCTTTCAAACTTGCAGGGATGCGGAAATTGCTTCTGAG +180
 E  V  K  N  N  R  D  F  F  Q  T  C  R  D  A  E  I  A  S  E
ATTACTATCCAGCCGGTAAGACGCTATAGAGGCCTCATTGATGCTGCTATTATTTTTAGT +240
 I  T  I  Q  P  V  R  R  Y  R  G  L  I  D  A  A  I  I  F  S
GATATCTTAGTTATTCCGCAAGCCATGGGTATGAGGGTCGAGATGCTCGAAGGTAAAGGT +300
 D  I  L  V  I  P  Q  A  M  G  M  R  V  E  M  L  E  G  K  G
CCACATTTCCCAGAACCTTTAAGAAATCCGGAAGACCTCCAAACGGTATTAGACTACAAG +360
 P  H  F  P  E  P  L  R  N  P  E  D  L  Q  T  V  L  D  Y  K
GTTGATGTTTTGAAAGAGTTAGATTGGGCTTTCAAGGCAATCACCATGACAAGGATCAAG +420
 V  D  V  L  K  E  L  D  W  A  F  K  A  I  T  M  T  R  I  K
TTGGATGGTGAGGTTCCCTTATTTGGCTTTTGCGGGGGACCTTGGACTCTAATGGTTTAT +480
 L  D  G  E  V  P  L  F  G  F  C  G  G  P  W  T  L  M  V  Y
ATGACGGAAGGCGGTGGATCCCGTCTTTTCAGATTTGCCAAACAATGGATTAACATGTAT +540
 M  T  E  G  G  G  S  R  L  F  R  F  A  K  Q  W  I  N  M  Y
CCAGAGCTTTCTCACAAATTATTACAAAAAATCACTGATGTGGCCGTGGAGTTTCTGAGT +600
 P  E  L  S  H  K  L  L  Q  K  I  T  D  V  A  V  E  F  L  S
CAGCAAGTCGTGGCGGGTGCTCAAATACTACAAGTTTTTGAAAGTTGGGGTGGTGAGCTT +660
 Q  Q  V  V  A  G  A  Q  I  L  Q  V  F  E  S  W  G  G  E  L
TCGTCTGTAGATTTTGATGAGTTTTCCCTACCATATTTAAGACAAATTGCCGAAAGAGTG +720
 S  S  V  D  F  D  E  F  S  L  P  Y  L  R  Q  I  A  E  R  V
CCTAAAAGATTGCAAGAATTAGGTATCATGGAACAGATTCCTATGATCGTTTTTGCGAAA +780
 P  K  R  L  Q  E  L  G  I  M  E  Q  I  P  M  I  V  F  A  K
GGGTCGTGGTATGCTTTGGACAAGCTATGCTGTTCAGGATTTGACGTTGTTTCGTTGGAC +840
 G  S  W  Y  A  L  D  K  L  C  C  S  G  F  D  V  V  S  L  D
TGGTCCTGGGACCCAAGAGAAGCGGTAAAAATAAACAAGAACCGTGTCACCTTGCAGGGC +900
 W  S  W  D  P  R  E  A  V  K  I  N  K  N  R  V  T  L  Q  G
AACCTGGATCCTGGCGTCATGTATGGTTCTAAAGAGGTAATAACAAAGAAAGTTAAACAG +960
 N  L  D  P  G  V  M  Y  G  S  K  E  V  I  T  K  K  V  K  Q
ATGATTGAGGCTTTTGGAGGTGGGAAGTCCCGCTACATTGTTAATTTCGGTCACGGTACC +1020
 M  I  E  A  F  G  G  G  K  S  R  Y  I  V  N  F  G  H  G  T
CACCCTTTCATGGATCCAGACGTCATCAAGTTTTTCTTGGAGGAGTGCCACAGAATTGGT +1080
 H  P  F  M  D  P  D  V  I  K  F  F  L  E  E  C  H  R  I  G
TCGAAGTAAtgtggaaatattgattcttatagtgttataagtgataatatacgtttatgt +1140
 S  K  *
aaataatatctttttcacaacgttcaaaagtggccttttttctttttcgatattctgta +1200
tatttattacgatttttttttttttcagatatttatttttt                    +1260
```

TGA...TATGT...A/T rich...TTT between nucleotides +1122 and +1153, similar to what had been found in other yeast genes (Zaret and Sherman, 1982).

The deduced amino acid sequence of the yeast *HEM6* gene was compared to those derived for the rat and human uroporphyrinogen decarboxylase (Romeo *et al.*, 1986; Romana *et al.*, 1987). Alignment of the three protein sequences revealed an overall identity of 46.4% amongst the three species distributed throughout the lengths of the proteins (Figure 9). There is an additional 30.6% similarity represented by conservative amino acid changes. The degree of identity between the yeast enzyme and the rat enzyme is 50.3%, whereas that for the yeast and human uroporphyrinogen decarboxylases is 50.8%. This high degree of homology between the three proteins indicates that *HEM6* is most likely the structural gene for the enzyme uroporphyrinogen decarboxylase.

Expression of *HEM6*

The cloned *HEM6* gene enabled us to study the regulation of *HEM6* expression on the transcriptional level. The cloned gene was used to generate labeled probes for Northern blot analysis. Total RNA was isolated from the wild-type strain BWG1-7a or an isogenic strain 1-7ahem1::LEU2 which contains a disruption in the *hem1* gene. To investigate the effects of carbon source on *HEM6* expression, BWG1-7a was grown in the presence of glucose or the non-fermentable carbon source, lactate. To examine the effects of heme or intermediates of the heme biosynthetic pathway on transcription of *HEM6*, 1-7ahem1::LEU2 was grown in the presence of high (50 μg/ml) or low levels (0.5 μg/ml) of δ-ALA. *HEM1* encodes δ-ALA synthase, the first enzyme of the heme

Figure 9. Alignment of derived amino acid sequences for uroporphyrinogen decarboxylase from different sources. The amino acid sequence deduced from the *HEM6* DNA sequence was aligned for maximal similarity to the predicted amino acid sequences of rat (Romana *et al.*, 1987) and human (Romeo *et al.*, 1986) uroporphyrinogen decarboxylases. Residues which are identical in all three sequences are indicated with a star. Residues which are conserved amongst the three sequences are marked with a ^.

```
YEAST   MG-------NFPAPKNDLILRAAKGEKVERPPCWIMRQAG 33
RAT     ---NGLGLQNFPELKNDTFLRAAWGEETDYTPVWCMRQAG 37
HUMAN   MEANGLGPQGFPELKNDTFLRAAWGEETDYTPVWCMRQAG 40
                ^***^ *** ^***** **^^^ ^* * *****

YEAST   RYLPEYHEVKNNRDFFQTCRDAEIASEITIQPVRRYRGLI 73
RAT     RYLPEFRETRAAQDFFSTCRSPEACCELTLEPVRRFP--L 75
HUMAN   RYLPEFRETRAAQDFFSTCRSPEACCELTLQPLRRFP--L 78
        ******^^^^^^^^^^ ******  ^^^^^^^^^^^  ^

YEAST   DAAIIFSDILVIPQAMGMRVEMLEGKGPHFPEPLRNPEDL 113
RAT     DAAIIFSDILVVPQALAMEVTMVPGKGPSFPEPLREERDL 115
HUMAN   DAAIIFSDILVVPQALGMEVTMVPSKGPSFPEPLREEQDL 118
        **********^*****^^* *^*^  *** *******^  **

YEAST   QTVLDYKVDVLKELDWAFKAITMTRIKLDGEVPLFGFCGG 153
RAT     ERLRD-PAAVASELGYVFQAITLTRQQLAGRVPLIGFAGA 154
HUMAN   EALRD-PEVVASELGYVFQAITLTRQRLAGRVPLIGFAGA 157
        ^ ^ *    * ^**^^^^^*****^** ^*^* ******* *^

YEAST   PWTLMVYMTEGGGSRLFRFAKQWINMYPELSHKLLQKITD 193
RAT     PWTLMTYMVEGGSFKTMAQAKRWLYQKPVASHKLLGILTH 194
HUMAN   PWTLMTYMVEGGGSSTMAQAKRWLYQRPQASHQLLRILTD 197
        ******^******^ ^ ^  ****^   *  **^**  ^*^

YEAST   VAVEFLSQQVVAGAQILQVFESWGGELSSVDFDEFSLPYL 233
RAT     ALVPYLIGQVAAGAQALQLFESHAGHLGSELFSKFALPYI 234
HUMAN   ALVPYLVGQVVAGAQALQLFESHAGHLGPQLFNKFALPYI 237
        ^ * ^*  **^***** **^**** ^*^*^^  *^^*^***^

YEAST   RQIAERVPKRLQELGIMEQIPMIVFAKGSWYALDKLCCSG 273
RAT     RDVAKRVKAGLQKAG-LTRMPMIIFAKDGHFALEELAQAG 273
HUMAN   RDVAKQVKARLREAG-LAPVPMIIFAKDGHFALEELAQAG 276
        *^^*^^*   *^^ * ^^^^^^^^^^^^ ^*****  ^*

YEAST   FDVVSLDWSWDPREAVKINKNRVTLQGNLDPGVMYGSKEV 313
RAT     YEVVGLDWTVAPKKAREPVGKTVTLQGELDPCALYASEEE 313
HUMAN   YEVVGLDWTVAPKKARECVGKTVTLQGNLDPCALYASEEE 316
        ^^*******^ ^*^^* ^   ^ *******^*** ^^*^^*

YEAST   ITKKVKQMIEAFGGGKSRYIVNFGHGTHPFMDPDVIKFFL 353
RAT     IGRLVQQMLNDFGPQ--RYIANLGHGLYPDMDPEHVGAFL 351
HUMAN   IGQLVKQMLDDFGPH--RYIANLGHGLYPDMDPEHVGAFV 354
        *^^ *^*******    *****^**** ^* ***^ ^  *^

YEAST   EECHRIGS----K 362
RAT     DAVHKHSRLLRQN 364
HUMAN   DAVHKHSRLLRQN 367
        ^^ *^ ^^    ^
```

biosynthetic pathway (Gollub *et al.*, 1977; Mattoon *et al.*, 1979). Mutants defective in this enzyme are able to grow only if the media is supplemented with δ-ALA. Thus, by controlling the amount of δ-ALA available to *hem1* mutants, one can control the levels of heme and heme biosynthetic intermediates present in the cell. After probing with the *HEM6* probe, each membrane was stripped and re-probed for the *ACT1* (actin) message. *HEM6* RNA values were normalized to the amount of *ACT1* message present in each lane (Figure 10).

The *HEM6* probe identified a message of approximately 1.4 kb that co-migrates with the *ACT1* mRNA. The results of the hybridization experiments indicate that the level of *HEM6* mRNA is elevated 1.5 to 2-fold in the presence of the non-fermentable carbon source, lactate (Figure 10, top panel). Examination of *HEM6* mRNA levels in cells grown in 50 μg/ml or 0.5 μg/ml δ-ALA revealed that the transcription of *HEM6* is not altered by the levels of heme or heme biosynthetic intermediates present in the cell (Figure 10, top panel). The densitometric tracing confirmed that equal amounts of *HEM6* RNA were present in cells grown either in low or high heme conditions (Figure 10, bottom panel).

HEM6 expression was also examined by measuring β-galactosidase activity expressed from a *HEM6-lacZ* fusion. The fusion was introduced into the same wild type strain BWG1-7a, as well as the *hem1* mutant strain 1-7ahem1::LEU2. The β-galactosidase assays of the transformed strains supported the results obtained from Northern analysis. In the *hem1* mutant strain, altering the level of δ-ALA did not alter the level of *HEM6-lacZ* expression (Table 3). On the other hand, the level of β-galactosidase

Figure 10. Northern blot analysis of *HEM6* RNA. Total RNA was isolated from the wild type strain BWG1-7a grown in glucose (lane 1) or lactate (lane 2) and the *hem1* mutant strain 1-7ahem1::LEU2 grown in 50 μg/ml δ-ALA (lane 3) or 0.5 μg/ml δ-ALA (lane 4). The top panel represents a Northern blot with 20 μg of total RNA in each lane. The samples were run on a 1% agarose gel in MOPS-formaldehyde buffer, transferred to Nytran membrane and probed with *HEM6* probe as described in Materials and Methods. The filters were then stripped and reprobed with *ACT1* probe. The bottom panel depicts the quantification of *HEM6* RNA levels in each lane by scanning densitometry and normalization to the amounts of *ACT1* RNA present.

Glucose
Lactate
50μg/ml δAla
0.5μg/ml δAla
HEM6
ACTIN

Relative Intensities
0
2
4
6
8
10
12
Glucose
Lactate
50μg/ml δAla
0.5μg/ml δAla

Table 3. Regulation of *HEM6-lacZ* expression

Fusion*	BWG1-7a (*HAP2HEM1*)		1-7ahem1::LEU2 (*HAP2hem1*)		LGW1 (*hap2HEM1*)
	Glucose	Lactate	50 μg/ml δ-ALA	0.5 μg/ml δ-ALA	Glucose
HEM6-lacZ	10.5 ± 0.4	21.2 ± 2.9	12.5 ± 0.7	11.9 ± 0.7	8.0 ± 0.6
CYC1-lacZ	182 ± 5.0	786 ± 23	402 ± 33	15 ± 4.4	116 ± 3.0

*Plasmids used were YEpH6-lacZ (*HEM6-lacZ*) and pLGΔ312 (*CYC1-lacZ*).

activity in the wild type strain was reproducibly induced 2-fold in cells grown in the presence of the non-fermentable carbon source, lactate, when compared to cells grown in glucose. This finding is consistent with the 1.5 to 2-fold increase in *HEM6* mRNA levels that we observed in cells grown in lactate.

When the sequences upstream of the presumptive *HEM6* TATA sequence were examined, the sequence CAATTGGT was found between positions -566 and -559. This sequence is homologous to the binding site for the HAP2-3-4 heteromeric transcription complex (Pinkham and Guarente, 1985; Hahn *et al.*, 1988; Forsburg and Guarente, 1989; Olesen and Guarente, 1990). This binding sequence TNATTGGT is found upstream of the coding regions of several genes, including *HEM1* and *HEM3* (Keng and Guarente, 1987; Keng *et al.*, 1992). In addition, normal expression of *HEM1* and *HEM3* was found to be dependent on the HAP2-3-4 complex. Therefore, β-galactosidase levels expressed from the *HEM6-lacZ* fusion were also examined in a *hap2* mutant strain, LGW1. *HAP2* encodes one subunit of the HAP2-3-4 heteromeric transcription complex. The level of expression of *HEM6-lacZ* in strain LGW1 was found to be 8 units, and approximates the 10 units of activity found in a wild type strain grown under the same conditions. These results indicate that the HAP2-3-4 trimeric transcription complex is not involved in *HEM6* expression.

DISCUSSION

In this paper we report the isolation and sequence analysis of *HEM6*, the structural gene for the enzyme uroporphyrinogen decarboxylase, from *Saccharomyces cerevisiae.* We have demonstrated that the cloned gene could complement the heme auxotrophy of a *hem6* mutant strain. Moreover, a yeast strain which contained a disrupted copy of the cloned gene in its genome behaved as a heme auxotroph which failed to complement a *hem6* mutant. Sequence analysis of the cloned gene revealed an open reading frame of 362 amino acids which has a predicted molecular weight of 41 kDa, a size which is consistent with the observed molecular weight of 38 kDa for the purified enzyme (Felix and Brouillet, 1990). Most importantly, the predicted amino acid sequence derived from the cloned gene showed a significant degree of homology with the deduced amino acid sequence for the rat (Romana *et al.*, 1987) and human (Romeo *et al.*, 1986) uroporphyrinogen decarboxylase; overall a 46.4% amino acid identity with an additional 30.6% similarity were observed. Our sequence is identical to that recently reported by Garey *et al.* (1992). This provides the first direct evidence that *HEM6* and *HEM12* are allelic. Studies on the expression of *HEM6* using Northern analysis and β-galactosidase assays demonstrated that expression of the *HEM6* gene was not regulated by changes in intracellular heme levels and that transcription of *HEM6* was induced at most two-fold in the presence of the non-fermentable carbon source, lactate. This two-fold increase in *HEM6* expression is consistent with the two-fold increase of uroporphyrinogen decarboxylase activity observed in the presence of a non-fermentable carbon source (Kurlandzka *et al.*, 1988). The relatively constitutive expression of *HEM6* resembles

those of *HEM1* (Keng and Guarente, 1987) and *HEM3* (Keng *et al.*, 1992), which encode two early enzymes in the heme biosynthetic pathway. Expression of both these genes is also not significantly affected by heme levels and carbon source. However, both *HEM1* and *HEM3* gene expression is under control of the HAP2-3-4 trimeric regulatory complex. Although analysis of the upstream non-coding sequences of *HEM6* revealed the presence of the sequence CAATTGGT between positions -566 and -559, which differs only by one base from the consensus HAP2-3-4 binding sequence TNATTGGT, *HEM6* expression was found to be not affected in a *hap2* mutant strain. While this does not rule out the possibility that the HAP2-3-4 complex may play a minor role in *HEM6 (HEM12)* expression in glucose, we believe that this possibility is unlikely. Alignment of the consensus sequences from multiple HAP-2-3-4-regulated genes reveals that the T in the first position of the consensus sequence TNATTGGT is very highly conserved and is found in all HAP-2-3-4 regulatory sequences identified to date (Keng and Guarente, 1987; Forsburg and Guarente, 1988; Repetto and Tzagoloff, 1989; Repetto and Tzagoloff, 1990; Bowman *et al.*, 1992; Trawick *et al.*, 1992). Moreover, site-directed mutagenesis of the consensus sequence in *LPD1* from TCATTGGC to GGATCCGC resulted in a large reduction of expression of *LPD1* under derepressing conditions (Bowman *et al.*, 1992). Given the high degree of conservation and the importance of a T at the first position of the consensus sequence and the absence of this residue in the sequence in *HEM6* (*HEM12*), and that the *hap2* mutation does not significantly affect expression of *HEM6* (*HEM12*) under all conditions that were tested, we conclude that *HEM6* (*HEM12*) expression is not under control of the HAP2-3-4 complex.

Studies with purified uroporphyrinogen decarboxylase from many sources have shown that sulfhydryl specific reagents such as N-ethylmaleimide, *p*-chloromercuribenzoate, and 5' 5'-dithiobis(2-nitrobenzoate) inhibit enzyme activity (Tomio *et al.*, 1970; De Verneuil *et al.*, 1983; Elder *et al.*, 1983; Kawanishi *et al.*, 1983; Straka and Kushner, 1983; Koopmann *et al.*, 1987). This inhibition can be reversed in many cases by incubation with dithiothreitol and can be prevented by pre-incubation with the substrate of the enzyme. Uroporphyrinogen decarboxylase from yeast has also been shown to be inhibited by sulfhydryl specific reagents, and as has been found for enzymes from other sources, the inactivation can be reversed by the addition of dithiothreitol (Felix and Brouillet, 1990). Taken together, these results suggest that one or more cysteines may comprise the active site of the enzyme. Due to the high degree of homology amongst the primary structures of the yeast and mammalian enzymes, it is reasonable to postulate that the mechanism of catalysis of the three homologous enzymes would be identical. Both the human and yeast enzymes contain 6 cysteine residues, while the rat enzyme contains 4 cysteine residues. Only one cysteine residue, that at position 52 of the yeast enzyme, was found to be conserved in all three species. However, a partial amino acid sequence of the putative *hemE* gene, encoding uroporphyrinogen decarboxylase, of the cyanobacterium *Synechococcus* sp. PCC7942 (Kiel *et al.*, 1990), revealed that the cysteine conserved in the eucaryotes was not found in the bacterial sequence. Alignment of the deduced partial amino acid sequences of the bacterial and eucaryotic enzymes indicate that there are no conserved cysteines amongst the amino termini of the homologous enzymes. However, there are two tandem cysteine codons found at positions 270-271 of yeast uroporphyrinogen decarboxylase and tandem cysteine

codons are also found at positions 65-66 of the human enzyme and at an equivalent position in the rat enzyme. Although the complete sequence of the bacterial *hemE* gene is unavailable at the moment, these tandem cysteines may be involved in the active site of the enzyme. Site-directed mutagenesis experiments should determine the importance of these cysteine residues at the active site of the enzyme. The enzyme has also been shown to be inhibited by diethylpyrocarbonate (Battle *et al.*, 1986) and ferric iron (Felix and Brouillet, 1990), suggesting the involvement of histidine residues in changing the conformation of the active site.

The isolation of *HEM6* (*HEM12*) and the availability of null mutants of *hem6* (*hem12*) will allow for a more detailed and precise analysis of the mechanism of action of uroporphyrinogen decarboxylase. The delineation of particular amino acid residues required for catalytic activity is now possible through the construction of more mutations of specific residues within the open reading frame.

Connecting text

The first manuscript of this thesis described the cloning of the *HEM6* gene on a 7.3 kb DNA fragment by complementation of the heme auxotrophy of a *hem6* mutant. *HEM6* was further localized to a 2.3 kb region on this fragment. The DNA sequence of this region was determined and was found to contain an open reading frame of 1086 nucleotides encoding a protein of 362 amino acids with a predicted molecular weight of 41 kDa. We examined the regulation of this gene by heme and carbon source, utilizing a *HEM6-lacZ* fusion as well as Northern blot analysis on *HEM6* mRNA levels. There was no regulation by heme and a 2-3-fold induction of *HEM6* expression in the presence of lactate, a non-fermentable carbon source.

Studies utilizing sulphydryl-specific reagents predicted the involvement of one or more cysteine residues as an integral component of the active site of the enzyme encoded by *HEM6*, uroporphyrinogen decarboxylase (Felix and Brouillet, 1990). Alignment of the amino acid sequence of the yeast enzyme and those of rat and human enzymes predicted that cysteine 52 of the yeast enzyme represents the only conserved cysteine in all three species. To determine whether this cysteine residue plays a critical role in uroporphyrinogen decarboxylase activity, we individually mutated the conserved cysteine 52 and the non-conserved cysteine 26 to serine or alanine. Plasmids containing these mutations were tested for their ability to complement the heme auxotrophy of *hem6* mutant strain 1B. In addition, strain 1B containing these mutant plasmids was tested for the ability to grow on a non-fermentable carbon source such as lactate and for accumulation of porphyrins. The results of these analyses is the subject of the manuscript contained within chapter 4.

CHAPTER 4

A CONSERVED CYSTEINE RESIDUE IN YEAST UROPORPHYRINOGEN DECARBOXYLASE IS NOT ESSENTIAL FOR ENZYMATIC ACTIVITY

Celestino Di Flumeri, Nicholas H. Acheson, and Teresa Keng

Department of Microbiology and Immunology, McGill University,
3775 University Street, Montreal, Quebec, Canada, H3A 2B4

This chapter is derived from a manuscript published in
the *Canadian Journal of Microbiology*, **43**, 792-795 (1997).

ABSTRACT

Uroporphyrinogen decarboxylase catalyzes the fifth step of heme biosynthesis in *Saccharomyces cerevisiae*. Studies utilizing sulfhydryl-specific reagents suggest that the enzyme requires a cysteine residue within the catalytic site. This hypothesis was tested directly by site-directed mutagenesis of highly conserved cysteine 52 to serine or alanine. Plasmids containing these mutations were able to complement a *hem6* mutant strain. In addition, properties associated with decreased uroporphyrinogen decarboxylase activity were not detected in the mutant strain transformed with these mutant plasmids. These results suggest that the conserved cysteine 52 by itself is not essential for enzymatic activity.

INTRODUCTION

Heme biosynthesis in *Saccharomyces cerevisiae* requires eight unique enzymatic steps (Tait, 1978; Labbe-Bois and Labbe, 1990), the fifth of which occurs in the cytoplasm, and involves the sequential decarboxylation of the four acetate side chains of uroporphyrinogen to yield coproporphyrinogen. This reaction is catalyzed by uroporphyrinogen decarboxylase (EC4.1.1.37), also known as porphyrinogen carboxy-lyase. The enzyme has been partially purified from bacteria (Hoare and Heath, 1959), and more recently, it has been purified to homogeneity from mammals (De Verneuil *et al.*, 1983; Elder *et al.*, 1983; Mukerji and Pimstone, 1987) and yeast (Felix and Brouillet, 1990). Characterization of the purified enzyme from various sources suggests that the enzyme is monomeric in nature, ranging in size from 40-50 kDa. Although a single protein is believed to catalyze the four sequential decarboxylation steps, uncertainty remains as to whether single or multiple active sites exist within the same protein.

Studies with purified uroporphyrinogen decarboxylase from many sources have demonstrated that sulfhydryl-specific reagents such as N-ethylmaleimide, *p*-chloromercuribenzoate, and 5', 5'-dithiobis (2-nitrobenzoate) inhibit enzyme activity (Tomio *et al.*, 1970; De Verneuil *et al.*, 1983; Elder *et al.*, 1983; Kawanishi *et al.*, 1983; Straka and Kushner, 1983; Koopmann and Battle, 1987; Felix and Brouillet, 1990). Uroporphyrinogen decarboxylase from yeast is also inhibited by sulfhydryl-specific reagents (Felix and Brouillet, 1990), suggesting that one or more cysteines may be a component of the active site of the enzyme.

The importance of a conserved cysteine residue within the active site of the enzyme was directly examined by making substitutions of this residue within the open

reading frame of the gene coding for the yeast enzyme, *HEM6*. Mutagenesis of yeast cysteine 52, the only conserved cysteine found among the cloned genes for the bacterial (Marczynski and Shapiro, 1992; Blattner *et al.*, 1993; Nishimura *et al.*, 1993), yeast (Garey *et al.*, 1992; Di Flumeri *et al.*, 1993), and mammalian enzymes (Romeo *et al.*, 1984; Romeo *et al.*, 1986; Romana *et al.*, 1987), indicated that this residue does not play an essential role in enzymatic activity.

RESULTS

Examination of the amino acid sequence of *HEM6* revealed the presence of six cysteine residues. Only one cysteine residue, that at position 52 of the yeast enzyme, was found to be conserved among bacterial, yeast, and mammalian species. This cysteine was mutagenized to structurally similar amino acids, serine and alanine, by site-directed mutagenesis. We also mutagenized a non-conserved cysteine at position 26 to a serine residue. The mutant alleles were cloned into low-copy (YCp50) and high-copy (YEp24) vectors. The ability of the mutant alleles to complement the heme auxotrophy of the *hem6* mutant strain 1B was examined. We also tested the ability of the transformed strains to grow on a non-fermentable carbon source and for accumulation of porphyrin intermediates.

The mutant plasmids as well as the homologous plasmids containing the wild type allele of *HEM6* were introduced into the *hem6* mutant strain 1B. For the wild type plasmids YEpH6 and YCpH6, colonies which were converted to uracil prototrophs were all found to be able to grow in the absence of heme, indicating that the wild type *HEM6* allele present on the plasmids was able to complement the *hem6* mutation in the heme auxotrophic strain (Table 4). Each mutant plasmid was able to complement the *hem6* mutation in strain 1B whether the mutation was on a low or high copy number plasmid (Table 4). These observations indicate that neither cysteine 26 nor the conserved cysteine 52 was essential for uroporphyrinogen decarboxylase activity. While it remains possible that mutation of cysteine 52 may have resulted in partial inactivation of the enzyme which could not be detected in our assays, we could not detect any defect in growth on lactate, a non-fermentable carbon source, or accumulation of porphyrin intermediates in

Table 4. Effects of site-directed mutagenesis of conserved cystine residues*

Plasmid	Growth in absence of heme:# + Glucose	+ Lactate	Porphyrin Accumulation†
YCp50	–	–	yes
YCpH6	+	+	no
YCpH6^{S26}	+	+	no
YCpH6^{S52}	+	+	no
YCpH6^{A52}	+	+	no
Yep24	–	–	yes
YEpH6	+	+	no
YEpH6^{S26}	+	+	no
YEpH6^{S52}	+	+	no
YEpH6^{A52}	+	+	no

*Plasmids were transformed independently into *hem6* mutant strain 1B and transformants were plated onto medium containing glucose and heme. Three colonies of each mutant were streaked either onto medium devoid of heme but containing glucose, or onto medium devoid of both heme and glucose but containing lactate.

#+ = growth; – = no growth.

†Porphyrin accumulation was determined by exposing three colonies of strain 1B independently transformed with mutant plasmids to ultraviolet light and examining them visually for fluorescence.

the transformed cells (Table 4). Analysis of other *hem6* mutant alleles has revealed that decreases in uroporphyrinogen decarboxylase activity are associated with decreased respiratory competence and accumulation of porphyrin intermediates (Urban-Grimal and Labbe-Bois, 1981; Kurlandzka *et al.*, 1988). The *hem6* mutant strain transformed with parental vectors YCp50 or YEp24 remained heme auxotrophic, suggesting that the heme prototrophy observed with the strain transformed with the mutant plasmids was not due to reversion of the *hem6* allele in 1B. In addition, when the *hem6* mutant strain was cured of the mutant plasmids, it no longer exhibited characteristics of the wild type *HEM6* strain (data not shown). This further proved that reversion of the *hem6* strain to heme prototrophy had not occurred.

If the active form of yeast uroporphyrinogen decarboxylase were multimeric, it is possible that two subunits mutated at different sites could form a functional multimeric enzyme with at least one intact active site. However, complementation by subunit mixing is unlikely to account for our results, since Felix and Brouillet (1990) showed that the yeast uroporphyrinogen decarboxylase is a monomer.

DISCUSSION

Our observations suggest that the conserved cysteine is dispensable for the reaction catalyzed by yeast uroporphyrinogen decarboxylase. Enzyme assays using purified human uroporphyrinogen decarboxylase which contained mutations in each of the six cysteine residues in that enzyme demonstrated that no cysteine residue by itself is critical for enzyme function (Wyckoff *et al.*, 1996). These *in vitro* results are in agreement with our *in vivo* complementation data with the yeast enzyme. However, it has been proposed that the enzyme has two or more active sites, with each catalyzing sequential decarboxylation reactions (De Verneuil *et al.*, 1980). If this were true, the mutation of a cysteine residue at one of the active sites may be insufficient to inactivate the enzyme. Addressing this possibility would necessitate the construction of site-directed mutants of multiple cysteines within the yeast enzyme. Two tandem cysteine residues found at positions 270-271 of the yeast enzyme are also found at positions 65-66 of the human enzyme and at an equivalent position in the rat and mouse (Wu *et al.*, 1996).enzymes. These cysteines could be a target for mutagenesis experiments trying to address this question.

Connecting text

The work presented in the previous chapter dealt with the role of cysteine residues within the catalytic site of uroporphyrinogen decarboxylase. Plasmids containing mutations of non-conserved cysteine 26 or the conserved cysteine 52 within the open reading frame of *HEM6* were able to complement a *hem6* mutant strain indicating that cysteine 52 is not required for enzymatic activity. In addition, *hem6* mutant 1B transformed with these plasmids showed no respiratory defect or porphyrin accumulation.

We proceeded to investigate the regulation of the gene for another heme biosynthetic enzyme, *HEM13*. This gene codes for the enzyme coproporphyrinogen oxidase. Transcription of *HEM13* is repressed by oxygen and heme. In the presence of heme, transcription of *ROX1* is induced and this protein presumably represses *HEM13* transcription. We therefore investigated the mechanism of repression of *HEM13* by ROX1 protein. We utilized ROX1 synthesized by *in vitro* translation and ROX1 produced in *E. coli* as a fusion to the glutathione-S-transferase (GST) protein to look at specific DNA-binding to an oligonucleotide from a region of the *HEM13* promoter previously shown to be important for *HEM13* repression. We also made use of both ROX1 preparations to test for the ability of ROX1 to oligomerize. The results concerning these experiments comprise chapter 5 of this thesis.

CHAPTER 5

THE HMG DOMAIN OF THE ROX1 PROTEIN MEDIATES REPRESSION OF *HEM13* THROUGH OVERLAPPING DNA BINDING AND OLIGOMERIZATION FUNCTIONS

Celestino Di Flumeri, Peter Liston, Nicholas H. Acheson, and Teresa Keng

Department of Microbiology and Immunology, McGill University,
3775 University Street, Montreal, Quebec, Canada, H3A 2B4

This chapter is derived from a manuscript published in
Nucleic Acids Research, **24**, 808-815 (1996).

ABSTRACT

The *ROX1* gene of *Saccharomyces cerevisiae* encodes a protein required for the repression of genes expressed under anaerobic conditions. ROX1 belongs to a family of DNA binding proteins which contain the high mobility group motif (HMG domain). To ascertain whether the HMG domain of ROX1 is required for specific DNA binding, we synthesized a series of ROX1 protein derivatives, either *in vitro* or in *E. coli* as fusions to the glutathione-S-transferase (GST) protein, and tested them for their ability to bind to DNA. Both ROX1 proteins that were synthesized *in vitro* and GST-ROX1 fusion proteins containing the intact HMG domain were able to bind to specific target DNA sequences. In contrast, ROX1 proteins which contained deletions within the HMG domain were no longer capable of binding to DNA. The oligomerization of ROX1 *in vitro* was demonstrated using affinity purified GST-ROX1 protein and ROX1 labeled with ^{35}S-methionine. Using various ROX1 protein derivatives, we were able to demonstrate that the domain required for ROX1-ROX1 interaction resides within the amino-terminal 100 amino acids which constitute the HMG domain. Therefore, the HMG domain is required for both DNA binding activity and oligomerization of ROX1.

INTRODUCTION

In the yeast *Saccharomyces cerevisiae*, a family of genes has been identified whose expression is coordinately regulated by oxygen. Transcription of genes in this family is repressed in the presence of oxygen and heme (Lowry and Lieber, 1986; Zagorec *et al.*, 1988; Hodge *et al.*, 1989). These genes are dispersed throughout the genome and include *HEM13* encoding coproporphyrinogen oxidase, an enzyme which catalyzes the sixth step in the heme biosynthetic pathway. Other members of this family include *ANB1* and *COX5b*, respectively encoding an isoform of translational factor eIF-5A and an isoform of cytochrome oxidase subunit V.

Repression of these genes in the presence of oxygen and heme is mediated by the product of the *ROX1* gene. Mutations in *ROX1* were first identified based on their ability to allow constitutive expression of the anaerobic gene *ANB1*, and *ROX1* was thus shown to encode a repressor (Lowry and Zitomer, 1984). ROX1 is also required for the repression of *COX5b* and *HEM13* by oxygen and heme (Hodge *et al.*, 1989; Keng, 1992). Transcription of *ROX1* is induced in the presence of oxygen or heme (Lowry and Lieber, 1986). Thus, levels of the ROX1 repressor are increased under aerobic conditions and serve to repress expression of the oxygen- and heme-repressed genes. Analysis of the upstream regulatory sequences of *ANB1* revealed the target for repression mediated by ROX1 protein to be two operator sites each containing the consensus sequence YYYATTGTTCTC; deletion of these operator sites resulted in high level expression of *ANB1* in the presence of oxygen (Lowry *et al.*, 1990). Inspection of the upstream non-coding regions of other ROX1-regulated genes revealed the presence of two copies of the operator consensus in *COX5b* (Hodge *et al.*, 1990) and five copies of this sequence in

HEM13 (Keng *et al.*, unpublished observations). Deletion of these sites in *COX5b* and in *HEM13* also resulted in constitutive expression of these genes (Hodge *et al.*, 1990; Keng *et al.*, unpublished observations).

DNA sequence analysis of *ROX1* revealed that it encodes a protein of 368 amino acids, and that the amino-terminal portion of ROX1 protein shows homology to the high mobility group (HMG) class of proteins (Balasubramanian *et al.*, 1993). The HMG1 and HMG2 proteins were first described as acid soluble, non-histone components of mammalian chromatin (Goodwin *et al.*, 1973).

Although these proteins are quite abundant, their function is not yet fully understood. These proteins contain two repeats, A and B, which share significant homology with each other, and an acidic domain which can interact with histone H1. The function of repeats A and B was revealed when Jantzen *et al.* (1990) identified a DNA binding motif in hUBF, a transcription factor for human RNA polymerase I, by sequence alignment with HMG1. On the basis of its homology with the A and B repeats of HMG1, this novel DNA binding motif was named the HMG domain. This domain is composed of an 85 amino acid stretch containing many basic amino acid residues. Since the initial discovery of this novel motif, a number of HMG domain proteins have been identified (Grosschedl *et al.*, 1994). These include the *Schizosaccharomyces pombe* regulatory proteins Ste11, Mat-a1, and Mat-Mc (Kelly *et al.*, 1988; Staben and Yanofsky, 1990; Sugimoto *et al.*, 1991), the mammalian sex determining factor SRY (Gubbay *et al.*, 1990; Sinclair *et al.*, 1990), the lymphoid specific transcription factors LEF-1 (TCF-1α) and TCF-1 (Oosterwegel *et al.*, 1991; Travis *et al.*, 1991; van de Wetering *et al.*, 1991;

Waterman *et al.*, 1991), the mitochondrial transcription factor mtTF1 (Parisi and Clayton, 1991), the yeast ARS binding protein ABF-2 (Diffley and Stillman, 1992), as well as the ROX1 protein (Balasubramanian *et al.*, 1993). Comparison of the sequences of the various members of this family of proteins has revealed the existence of subfamilies of HMG proteins (Laudet *et al.*, 1993). One subfamily is made up of the HMG1 and HMG2 proteins, the ARS binding protein ABF-2, UBF, and the mtTF-1 protein. These proteins are present in all cell types, contain multiple HMG domains, and recognize DNA with no sequence specificity. The second subfamily includes proteins that show tissue-specific expression, interact specifically with restricted DNA sequences, and contain only a single HMG domain. Members of this subfamily include the lymphoid factors LEF-1 and TCF-1, the sex-determining protein SRY, and the fungal transcription factors Ste11, Mat-Mc and Mat-a1.

In order to delineate the functional domains of the ROX1 protein, we expressed the proposed HMG domain of ROX1 both *in vitro* in wheat germ extracts, and *in vivo* as a fusion to glutathione-S-transferase (GST) in *E. coli.* We asked if the HMG domain portion of ROX1 was capable of binding specifically to an oligonucleotide from the *HEM13* promoter which we have shown to be necessary for repression of *HEM13* . Our analysis indicates that ROX1 protein derivatives containing the entire HMG domain, be they *in vitro* translated, GST-ROX1 fusion proteins, or ROX1 derivatives separated from the GST moiety by thrombin cleavage, were able to bind to DNA with sequence specificity. We also demonstrate that the HMG domain of ROX1 is required for the ability of ROX1 to form oligomers *in vitro*.

RESULTS

Synthesis of ROX1 protein *in vitro*

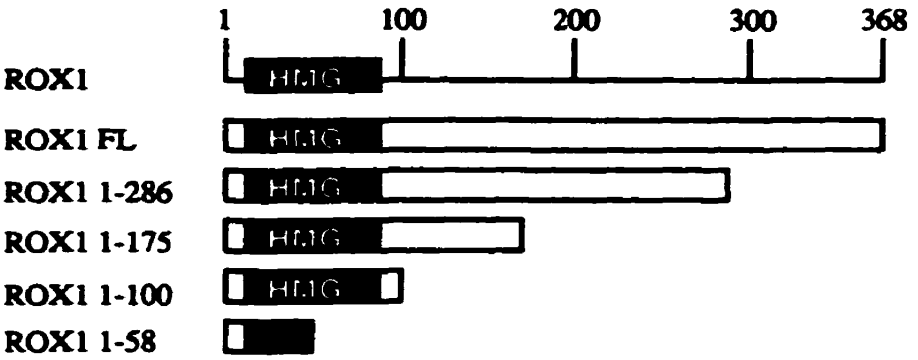
In order to define the region of the ROX1 protein necessary for DNA binding, we generated derivatives of ROX1 with carboxy-terminal deletions using *in vitro* transcription/translation systems. The *ROX1* gene was inserted into a plasmid downstream of an SP6 promoter, and deletions were generated by digesting the DNA with different restriction endonucleases (Figure 11A). The linearized DNA templates were subsequently transcribed with SP6 RNA polymerase, and the transcripts were translated in a wheat germ extract. Aliquots of the ^{35}S-methionine labeled proteins were analyzed by SDS-polyacrylamide gel electrophoresis to verify that synthesis had occurred and that the proteins were produced in similar amounts (Figure 11B). From the autoradiogram, each labeled protein was found to be of a size consistent with its predicted molecular weight. Proteins of lower molecular weight that were also observed in some cases were probably due to premature termination of translation. No labeled protein was observed when no exogenous RNA was added to wheat germ extract (Figure 11B).

Specific DNA binding by ROX1 protein synthesized *in vitro*

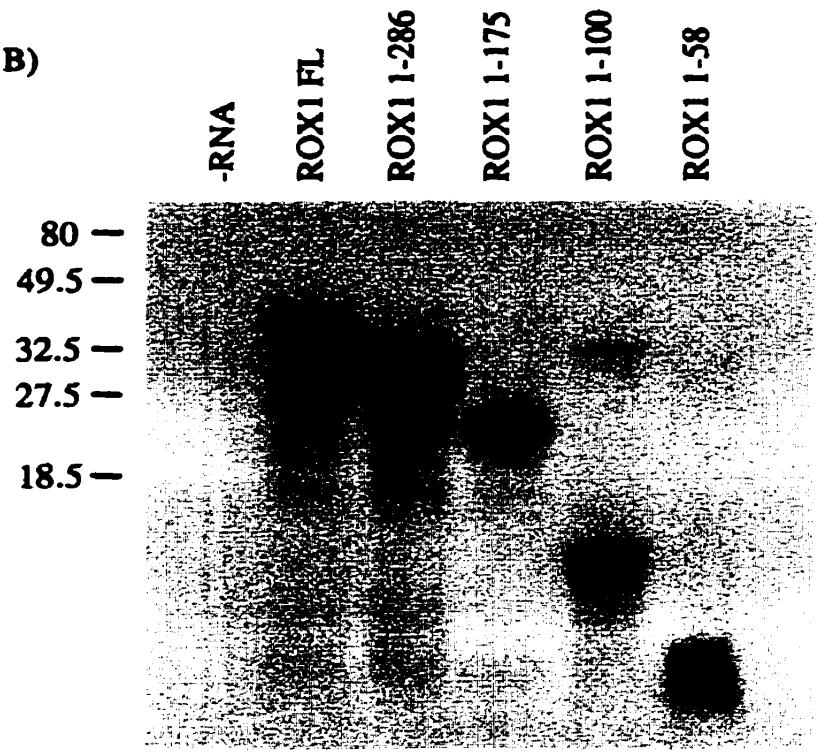
The full length ROX1 protein that was synthesized *in vitro* was tested for its ability to interact with DNA in a sequence-specific manner by an electrophoretic mobility shift assay. A labeled 32 base pair (bp) double stranded DNA, RS32, containing sequences from the *HEM13* promoter defined to be involved in repression, was used as a probe.

Figure 11. *In vitro* translation of ROX1 protein. (A) Schematic diagram summarizing the structure of full length ROX1 protein, as well as various deletion derivatives. The 368 amino acid open reading frame is indicated by a numbered line, with the HMG domain homology indicated by a box. The various deletion mutants are indicated below the numbered line, and are named according to the amino acids found in the ROX1 protein. (B) SDS-polyacrylamide gel analysis of *in vitro* translated ROX1. Templates encoding full length *ROX1* and the various deletion derivatives of *ROX1* were transcribed with SP6 RNA polymerase. The resulting transcripts were translated *in vitro* in wheat germ extracts in the presence of ^{35}S-methionine. Aliquots of the translation products were run on a 15% denaturing gel which was enhanced by fluorography, dried and exposed to X-ray film. The sizes of the molecular mass markers are indicated in kDa.

A)


1
100
200
300
368
ROX1
HMG
ROX1 FL
HMG
ROX1 1-286
HMG
ROX1 1-175
HMG
ROX1 1-100
HMG
ROX1 1-58

B)


-RNA
ROX1 FL
ROX1 1-286
ROX1 1-175
ROX1 1-100
ROX1 1-58
80
49.5
32.5
27.5
18.5

The left panel in figure 12 depicts an autoradiogram of binding reactions performed in the presence of increasing amounts of unlabeled competitor DNA with the same sequence as the labeled probe. In the absence of specific competitor, protein-DNA complexes were detected (Figure 12, lane 2). The complexes were barely detectable when a ten-fold excess of the unlabeled specific DNA was added to the reaction (Figure 12, lane 3), and disappeared when a 50- or 100-fold excess of the unlabeled DNA was included in the binding mixture (Figure 12, lanes 4 and 5). In contrast, no detectable decrease in complex formation was observed when an excess of an unrelated 31 bp fragment, RNS31, was added to the reactions as a non-specific competitor (Figure 12, lanes 8-10). No complex formation was detected when wheat germ extract incubated in the absence of *ROX1* RNA was added to labeled RS32 DNA (Figure 12, lane 1). These experiments demonstrate that ROX1 protein synthesized *in vitro* is able to form a protein-DNA complex with a specific target DNA.

Delineation of the DNA-binding domain of ROX1

To delineate precisely the region of ROX1 required for sequence specific DNA binding activity, we made use of the various ROX1 polypeptides generated from *in vitro* translation of truncated templates shown in figure 11. These polypeptides were tested for their ability to bind to the labeled RS32 DNA probe. ROX1 polypeptides which contain 175, 286, or 368 amino acids gave rise to multiple complexes with identical mobilities on these gels (Figure 13, lanes 2-4). The ROX1 1-100 protein, which contains the entire putative HMG domain, was capable of strong binding to the RS32 DNA. However, in this case, the protein-DNA complexes migrated as a single band (Figure 13, lane 5).

Figure 12. Binding of *in vitro* translated ROX1 to a DNA fragment containing a specific ROX1 binding site. Aliquots of the translation product of the full length *ROX1* transcript were incubated with a labeled DNA fragment containing a specific ROX1 binding site (RS32). Incubations were carried out in the presence of the indicated molar excesses of the identical unlabeled DNA as competitor (specific competitor) or an unlabeled DNA (RNS31) without a specific ROX1 binding site (non-specific competitor). DNA binding and electrophoresis were performed as outlined in Materials and Methods. The -RNA lanes contain labeled DNA fragment (RS32) with a specific ROX1 binding site incubated with wheat germ extract without added RNA.

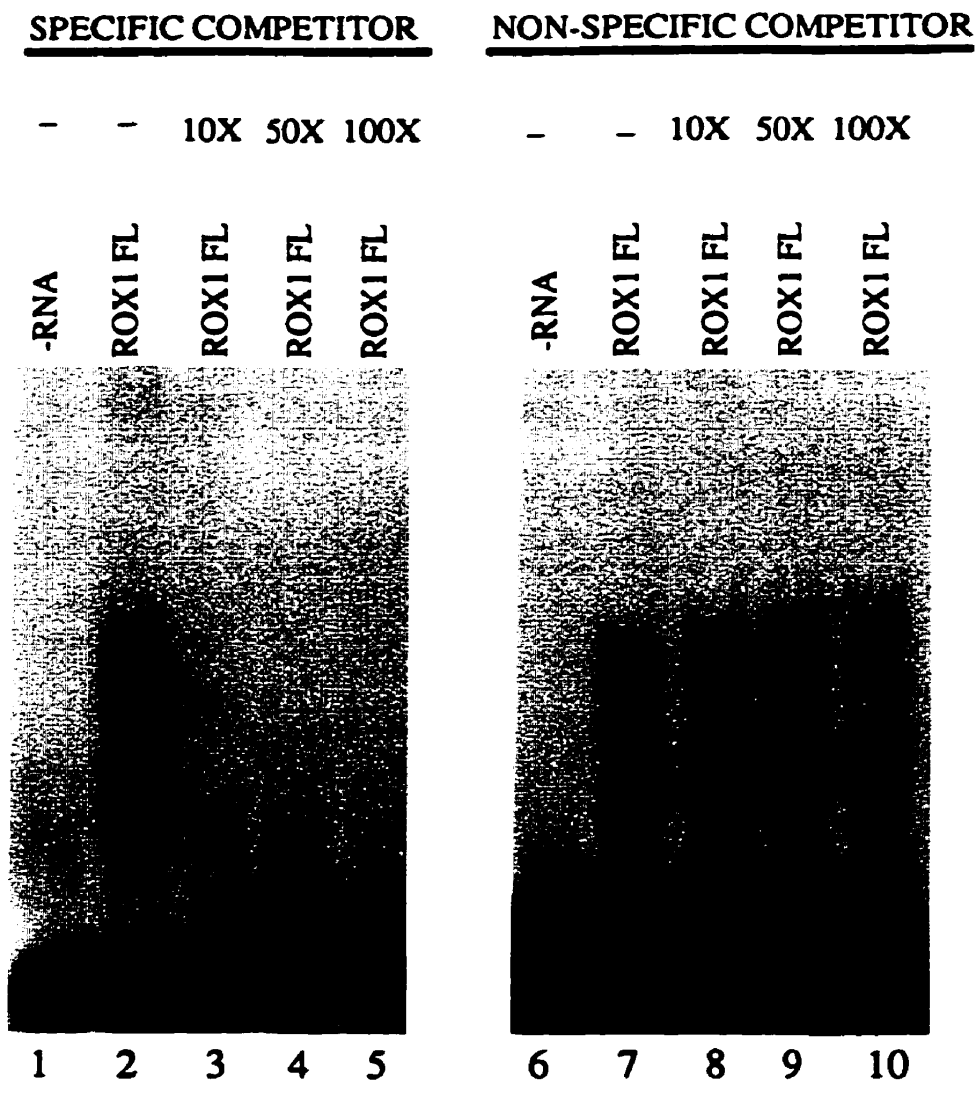
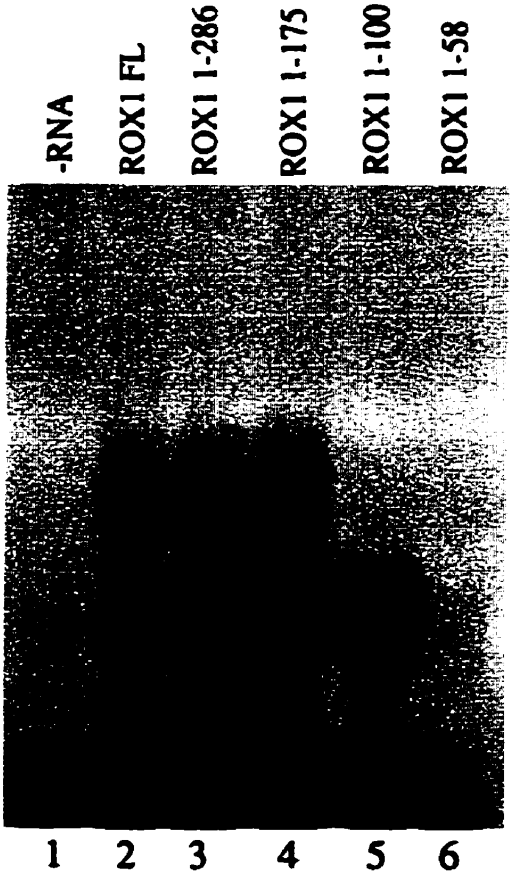

SPECIFIC COMPETITOR
NON-SPECIFIC COMPETITOR
- - 10X 50X 100X
- - 10X 50X 100X
-RNA
ROX1 FL
ROX1 FL
ROX1 FL
ROX1 FL
-RNA
ROX1 FL
ROX1 FL
ROX1 FL
ROX1 FL
1 2 3 4 5
6 7 8 9 10

Figure 13. Delineation of the DNA binding domain of ROX1. 1 µl aliquots of the respective ROX1 derivatives synthesized in *in vitro* transcription/translation systems were incubated with a labeled RS32 probe containing the specific ROX1 binding site. The reaction mixtures were analyzed by electrophoretic mobility shift assays as described in Materials and Methods. -RNA indicates a reaction in which the labeled RS32 DNA was incubated with wheat germ extract into which no RNA had been added.


-RNA
ROX1 FL
ROX1 1-286
ROX1 1-175
ROX1 1-100
ROX1 1-58
1
2
3
4
5
6

ROX1 1-58, which lacks the carboxy-terminal 35 amino acids in the proposed HMG domain, was not able to bind to DNA (Figure 13, lane 6). Taken together, these results suggest that the region of ROX1 required for DNA binding resides within the 100 amino-terminal residues, and that the proposed HMG domain within this region of ROX1 is required for DNA binding by the protein.

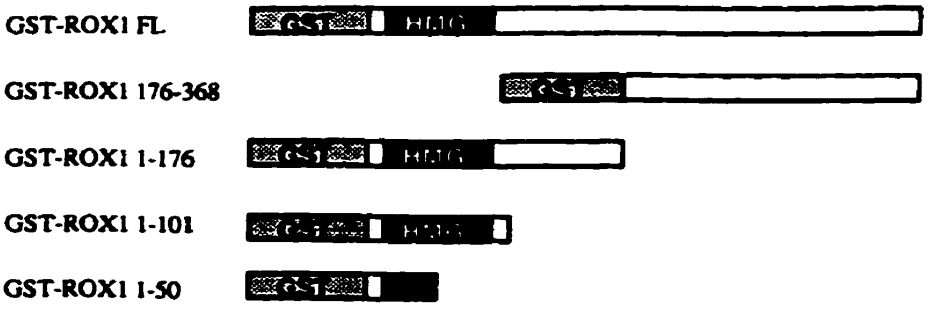
Expression of GST-ROX1 fusion proteins in *E. coli*

In order to obtain a more homogeneous source of ROX1 protein, we expressed ROX1 as a fusion to the glutathione-S-transferase (GST) protein in *E. coli*. The entire *ROX1* open reading frame was cloned in frame into the pGEX-2T plasmid to generate the plasmid GST-ROX1 FL. This plasmid was subsequently used to construct fusion proteins between GST and different regions of the ROX1 protein (Figure 14A). Plasmids containing the different fusions were introduced into *E. coli* strain SG935, which harbors a mutation in the Lon protease. Fusion proteins were purified using a single step affinity purification with glutathione Sepharose beads.

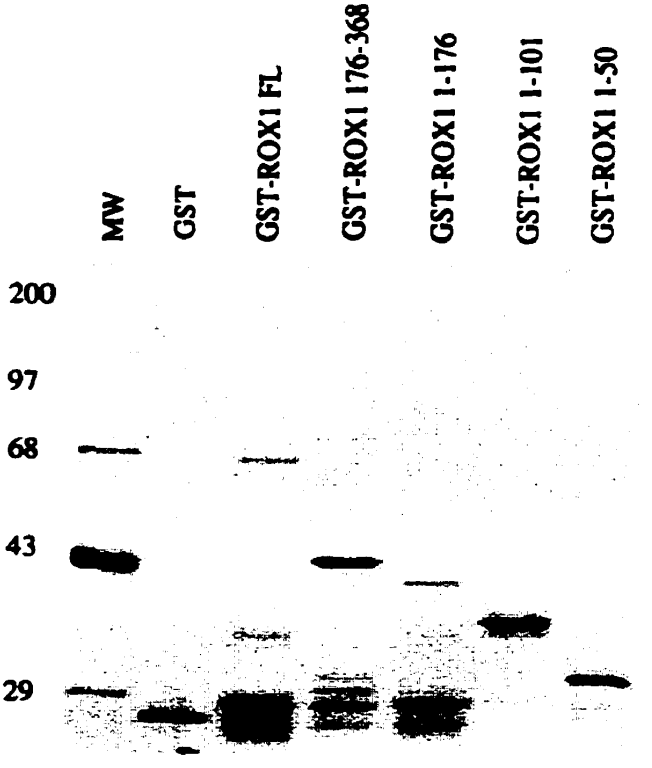
Purified proteins were separated by SDS-polyacrylamide gel electrophoresis to verify the expression and size of each protein. Figure 14B shows expression of the 26 kDa GST protein alone as well as that of various GST-ROX1 fusion proteins. The full length ROX1 protein fused to GST migrated with an apparent molecular weight of 68 kDa, a size consistent with the predicted molecular weight of 66 kDa. The GST-ROX1 amino-terminal fusion proteins, GST-ROX1 1-176, GST-ROX1 1-101, and GST-ROX1 1-50, also gave the expected molecular sizes (Figure 14B). The carboxy-terminal fusion protein GST-ROX1 176-368 migrated with an apparent molecular weight of 43 kDa,

Figure 14. Expression and purification of GST-ROX1 fusion proteins. (A) Various portions of the *ROX1* open reading frame were fused in frame to glutathione-S-transferase. The nomenclature for each construct refers to the amino acids of the ROX1 protein contained in the fusion protein. The shaded portion represents the GST portion of the fusion protein, while the remaining region refers to the amino acids of ROX1. The HMG domain is shown as a black box. (B) Proteins were expressed in *E. coli* and purified using glutathione Sepharose beads. Aliquots of the partially purified proteins were separated by SDS-polyacrylamide gel electrophoresis and stained with coomassie brilliant blue. Molecular mass markers are indicated in kDa.

A)


GST-ROX1 FL
GST-ROX1 176-368
GST-ROX1 1-176
GST-ROX1 1-101
GST-ROX1 1-50

B)


MW
GST
GST-ROX1 FL
GST-ROX1 176-368
GST-ROX1 1-176
GST-ROX1 1-101
GST-ROX1 1-50
200
97
68
43
29

closely approximating the size predicted by its open reading frame. An advantage of this system is that these proteins are made at higher levels than the equivalent proteins synthesized by *in vitro* translation. This was verified by coomassie blue staining of proteins synthesized by both methods.

To demonstrate that the proteins bound to glutathione Sepharose were in fact fusion proteins between GST and ROX1, and that these proteins were devoid of *E. coli* derived proteins, we performed immunoblot analysis on the expressed proteins using a rabbit polyclonal antibody that was raised against the GST-ROX1 1-50 protein (data not shown).

DNA binding by GST-ROX1 fusion proteins

Gel mobility shift assays performed with the various fusion proteins demonstrated that GST-ROX1 fusion proteins containing a minimum of 101 amino-terminal amino acids were capable of binding to the labeled RS33 DNA fragment containing the same ROX1 binding site from the *HEM13* promoter as used with *in vitro* synthesized ROX1 proteins (data not shown). In addition, the inability of GST-ROX1 176-368 fusion protein to bind to the RS33 DNA fragment ruled out the possibility that a second independent DNA binding domain existed in ROX1 at the carboxy terminus.

To demonstrate that the ROX1 portion of the GST-ROX1 fusion protein is responsible for the observed sequence-specific DNA binding activity and that the GST moiety of the fusion protein is not altering the ability of ROX1 to bind to DNA, we performed gel mobility shift assays with ROX1 proteins that were separated from the GST portion (Figure 15). Thrombin cleaves the fusion protein at the junction between

the GST and ROX1 portions. Cleavage was verified by the detection by immunoblot analysis of smaller proteins corresponding to the ROX1 portion and the GST moiety (data not shown). The multiple protein-DNA complexes seen with the ROX1 FL and ROX1 1-176 proteins resembled the pattern of complexes obtained with equivalent ROX1 proteins that were synthesized *in vitro* and used in gel mobility shift assays (compare Figure 15, lanes 2 and 3 with Figure 13, lanes 2 and 4). In addition to the fast-migrating complex formed with the equivalent ROX1 protein synthesized by *in vitro* translation, ROX1 1-101 generated by cleavage of GST-ROX1 1-101 formed a second, slower migrating protein-DNA complex (compare Figure 13, lane 5 with Figure 15, lane 4).

ROX1 protein-protein interactions

To determine if ROX1 is capable of oligomerization, we employed ^{35}S-labeled ROX1 proteins synthesized by *in vitro* transcription/translation, as well as GST-ROX1 fusion proteins, in a GST "pull-down" assay. ^{35}S-labeled ROX1 1-175 protein was incubated with different GST-ROX1 fusion proteins bound to glutathione Sepharose beads. After incubation, the beads containing the GST-ROX1 fusion proteins were isolated by centrifugation. Any ^{35}S-labeled ROX1 1-175 protein interacting with the GST-ROX1 fusion proteins will bind to the beads and can be detected by SDS-polyacrylamide gel electrophoresis. To eliminate the possibility that ^{35}S-labeled ROX1 1-175 protein could be interacting with the GST-ROX1 fusion proteins via the GST moiety, we incubated labeled ROX1 protein with GST alone. No ROX1 1-175 protein was bound to GST alone (Figure 16A, lane 1).

Figure 15. DNA binding by the ROX1 portion of the GST-ROX1 fusion protein. GST-ROX1 FL, GST-ROX1 1-176, and GST-ROX1 1-101 were treated with thrombin which released the ROX1 portion. The released ROX1 proteins were tested for their ability to bind to the labeled RS33 DNA containing the ROX1 binding site.

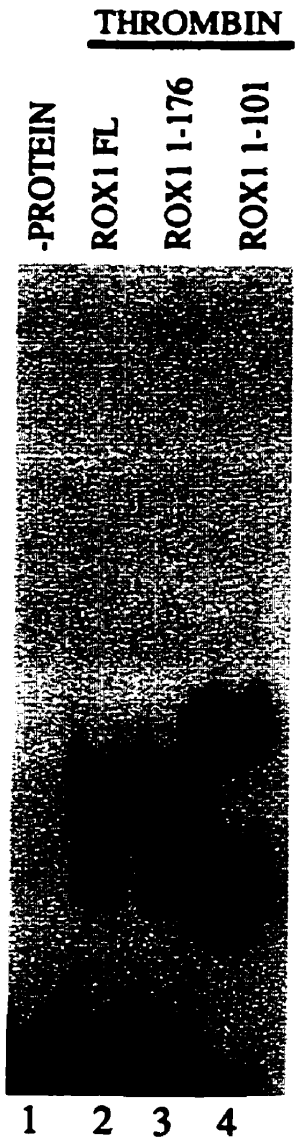
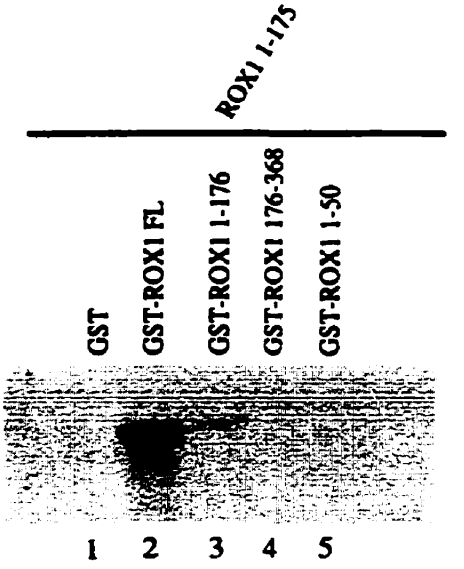
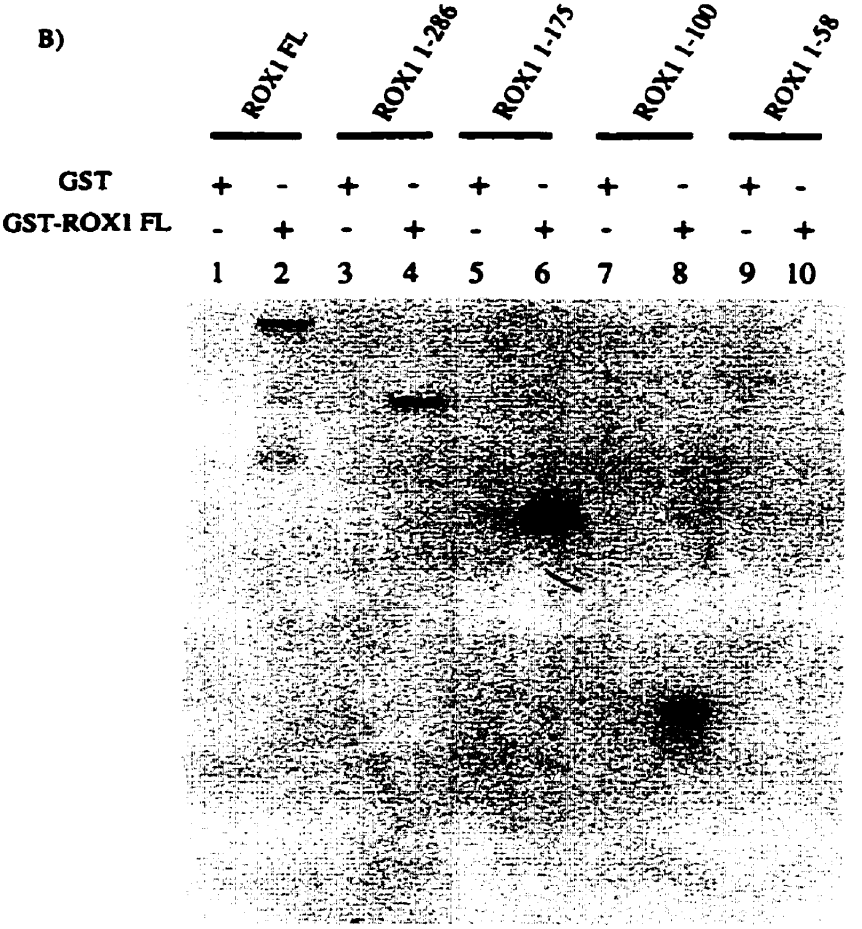

THROMBIN
-PROTEIN
ROX1 FL
ROX1 1-176
ROX1 1-101
1
2
3
4

Figure 16. Determination of ROX1-ROX1 protein interactions. (A) GST protein, as well as the various GST-ROX1 fusion proteins immobilized on glutathione Sepharose beads, were tested for their abilities to interact with ^{35}S-labeled ROX1 1-175 protein as described in Materials and Methods. Labeled protein retained by interaction with GST-ROX1 proteins on the beads was visualized by fluorography after separation on an SDS-polyacrylamide gel. (B) GST-ROX1 FL protein, or GST protein alone, was immobilized on glutathione Sepharose beads and tested for its ability to interact with ^{35}S-labeled ROX1 deletion derivatives. The - symbol indicates the absence of GST protein or GST-ROX1 FL protein in the reactions and the + symbol indicates the presence of the protein in the reactions. Odd numbered lanes contain GST protein while even numbered lanes contain GST-ROX1 FL protein. Labeled proteins bound to the immobilized protein on beads were detected by fluorography.

A)


ROX1 1-175
GST
GST-ROX1 FL
GST-ROX1 1-176
GST-ROX1 176-368
GST-ROX1 1-50
1 2 3 4 5

B)

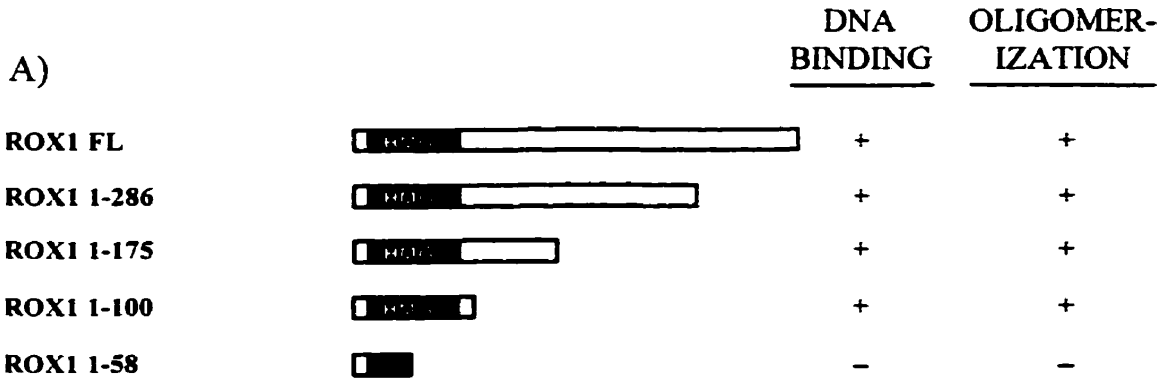

ROX1 FL
ROX1 1-286
ROX1 1-175
ROX1 1-100
ROX1 1-58
GST + - + - + - + - + -
GST-ROX1 FL - + - + - + - + - +
1 2 3 4 5 6 7 8 9 10

When GST-ROX1 FL or GST ROX1 1-176 was used in binding reactions with labeled ROX1 1-175 protein, specific interactions between the labeled protein and GST-ROX1 fusion protein could be detected (Figure 16A, lanes 2 and 3). GST-ROX1 1-50, which contains only the amino-terminal 50 residues of ROX1 fused to GST, did not interact with ^{35}S-labeled ROX1 1-175 protein (Figure 16A, lane 5); neither did GST-ROX1 176-368 protein (Figure 16A, lane 4).

In order to ensure that these interactions were specific, experiments testing binding between GST or GST-ROX1 FL protein and a variety of ^{35}S-labeled ROX1 proteins were performed. GST alone bound to Sepharose beads was unable to bind specifically to any of the *in vitro* translated derivatives of ROX1 (Figure 16B, odd numbered lanes). This indicates that interactions of the GST-ROX1 fusion proteins with labeled ROX1 proteins is due to the ROX1 moiety specifically interacting with ^{35}S-labeled ROX1. When GST-ROX1 FL protein was incubated with ^{35}S-labeled ROX1 FL protein, a specific interaction could be detected (Figure 16B, lane 2). In order to more precisely map the region of ROX1 responsible for this interaction, GST-ROX1 FL was tested for its ability to interact with various truncated ^{35}S-labeled ROX1 proteins. Labeled ROX1 proteins containing at least 100 amino-terminal amino acids were found to specifically interact with the full length GST-ROX1 fusion protein (Figure 16B, lanes 4, 6, 8). However, ^{35}S-labeled ROX1 1-50 protein was unable to interact with the GST-ROX1 FL protein (Figure 16B, lane 10). Taken together, these results indicate that ROX1 can form oligomers and that the region of ROX1 protein required for this function is the HMG domain, which is also required for DNA binding.

Figure 17. Delineation of the DNA binding domain and oligomerization domain of ROX1. (A) The results of the DNA binding assays and oligomerization assays for *in vitro* synthesized ROX1 and its derivatives are summarized. (B) A summary of the results of the DNA binding assays and oligomerization assays performed with GST-ROX1 fusion proteins. + indicates ability to bind DNA or ability to interact with ROX1. – indicates the absence of DNA binding activity and the absence of oligomerization as detected by our assays. N.D. indicates not determined.

A)


DNA BINDING
OLIGOMER-IZATION
ROX1 FL
+
+
ROX1 1-286
+
+
ROX1 1-175
+
+
ROX1 1-100
+
+
ROX1 1-58
–
–

B)

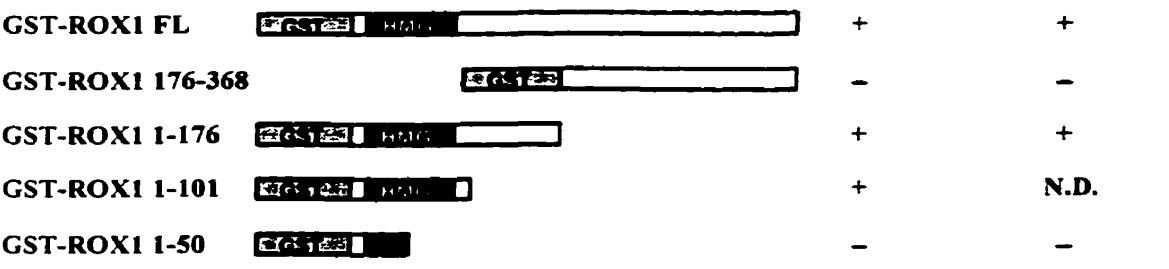

GST-ROX1 FL
+
+
GST-ROX1 176-368
–
–
GST-ROX1 1-176
+
+
GST-ROX1 1-101
+
N.D.
GST-ROX1 1-50
–
–

Figure 17 summarizes the DNA binding and oligomerization properties of the various ROX1 protein derivatives synthesized either *in vitro* or as GST fusions in *E. coli.* In figure 17A, we see that all *in vitro* synthesized ROX1 proteins which contain the entire HMG domain are capable of both DNA binding and oligomerization. The smallest such protein is ROX1 1-100, which contains a deletion of the carboxy-terminal 268 amino acids of ROX1 and yet still retains both the ability to bind DNA and to oligomerize. As summarized in figure 17B, identical results were obtained with GST-ROX1 fusion proteins. In this instance, the GST-ROX1 176-368 protein, lacking the HMG domain, is unable to oligomerize or to bind DNA. While all GST-ROX1 fusion proteins containing the intact HMG domain are capable of DNA binding, a deletion within the HMG domain, either with *in vitro* synthesized protein or with the ROX1 fusion protein, abolishes both functions of ROX1.

DISCUSSION

In this report, we show that the full length ROX1 repressor is capable of sequence specific binding to an operator region in the regulatory region of *HEM13*. This DNA binding activity of ROX1 is mediated by the HMG domain at its amino terminus. We also demonstrate that the HMG domain is required for the oligomerization of ROX1 *in vitro*.

ROX1 is a sequence specific DNA-binding protein. Full length ROX1 synthesized *in vitro* in wheat germ extracts was able to specifically bind to a fragment of DNA containing sequences required for repression of *HEM13*. Neither a 100-fold excess of a non-homologous DNA fragment nor a vast excess of poly dI:dC could compete for binding to ROX1. DNA binding assays utilizing truncated versions of ROX1 synthesized *in vitro* indicated that the amino-terminal 100 amino acids of the ROX1 protein are essential for DNA binding. This region includes the HMG domain which lies between residues 9 and 93 (10). *In vitro* translated ROX1 1-58, which is missing 35 amino acids of the HMG domain, and GST-ROX1 1-50, which is missing 43 amino acids of the HMG domain, are both unable to bind to DNA fragments containing the operator site.

Interestingly, all ROX1 derivatives containing a minimum of 175 amino-terminal amino acids gave rise to multiple protein-DNA complexes which migrated with the same mobilities as complexes formed with full length ROX1 protein. The formation of multiple complexes with these derivatives could be due to the ability of ROX1 to form oligomers *in vitro* and the multiple protein-DNA complexes detected may represent complexes of oligomers of ROX1 with DNA. Alternatively, this phenomenon could be due to the presence of protease-sensitive sites within the ROX1 protein. The different

ROX1 protein derivatives would be degraded to form distinct species which could then form multiple specific protein-DNA complexes. However, we believe the latter possibility to be improbable, given the fact that the multiple protein-DNA complexes appear whether we utilize *in vitro* synthesized ROX1 or *E. coli* derived ROX1 proteins.

The oligomerization of ROX1 was investigated using both GST-ROX1 fusion proteins and *in vitro* labeled ROX1. Our analysis indicates that ROX1 protein is capable of oligomerization and that the HMG domain at the amino terminus of ROX1 is required for formation of oligomers. The ROX1 protein contains a stretch of 22 amino acids from residues 102 to 123 which contains 16 glutamine residues. Such stretches of glutamine residues are a feature common to eukaryotic transcriptional activator proteins and are thought to be required for protein-protein interactions (Courey *et al.*, 1989). In the protein interaction assay, the ROX1 1-100 *in vitro* translated protein was capable of interacting with GST-ROX1 FL protein, indicating that the polyglutamine tract in ROX1 is not required for ROX1 oligomerization.

It is interesting to note that while the *in vitro* translated ROX1 1-100 protein gave rise to a single protein-DNA complex in the DNA binding assay, the *E.coli* produced, thrombin cleaved ROX1 1-101 protein gave rise to a protein-DNA complex of lower mobility in addition to one of the same mobility as that of the *in vitro* synthesized protein. The formation of the additional complex may be explained by the different concentrations of the two proteins used in the DNA binding assays. The concentration of *E. coli* produced ROX1 1-101 in the DNA binding assays was higher than that of *in vitro* translated ROX1 1-100 and the presence of a higher concentration of protein in the binding assay would favor the formation of oligomeric species. When the DNA binding

assay was carried out with lower concentrations of ROX1-101 protein, only the complex with a higher mobility was detected (data not shown).

Although the precise mechanism of repression by ROX1 protein remains unclear, it must involve the TUP1 and SSN6 (CYC8) proteins. In strains deleted for *TUP1* or *SSN6,* expression of ROX1-regulated genes is observed under repressing conditions (Balasubramanian *et al.*, 1993; Deckert *et al.*, 1995b). In particular, we have constructed *ssn6::LEU2* and *tup1::LEU2* disrupted strains each containing an integrated copy of a *HEM13-lacZ* fusion at the *TRP1* locus. When expression of this fusion was tested under repressing conditions, derepressed levels of *HEM13-lacZ* activity were detected (data not shown). This suggests that the TUP1/SSN6 complex is an integral component of the machinery required for repression of *HEM13*. The TUP1 and SSN6 proteins are also required for activity of a number of other DNA binding repressor proteins that function to regulate a wide range of activities including cell type and catabolite repression (Nehlin and Ronne, 1990; Mukai *et al.*, 1991; Keleher *et al.*, 1992). While they do not themselves contact DNA, TUP1 and SSN6 are believed to form a complex that is recruited by specific DNA binding repressors to the promoter (Schultz *et al.*, 1990; Williams and Trumbly, 1990; Williams *et al.*, 1991; Keleher *et al.*, 1992; Tzamarias and Struhl, 1994; Treitel and Carlson, 1995; Tzamarias and Struhl, 1995). Affinity chromatography experiments have detected only a tenuous SSN6-ROX1 interaction (Tzamarias and Struhl, 1995). We have also failed to detect any interaction of SSN6 with ROX1 (data not shown). This may be due to the complexity of the interaction. For example, interaction of ROX1 with SSN6 may only occur in the presence of TUP1. Alternatively, ROX1 binding to its cognate DNA may be required for interaction with

SSN6. Interaction of SSN6 with ROX1 may occur through what we believe is the repression domain of ROX1 which is located at the carboxy terminus. A strain containing a *rox1::LEU2* disruption, as well as an integrated copy of a *HEM13-lacZ* fusion, was transformed with plasmids containing different deletion derivatives of *ROX1*. Ability of the these derivatives to restore repression to the strain was examined. A *ROX1* construct containing a deletion of sequences coding for the 82 carboxy-terminal amino acids, which is fully capable of *in vitro* DNA binding, was unable to repress expression of *HEM13-lacZ in vivo* (data not shown). Thus, we believe that ROX1 can function to recruit the TUP1/SSN6 complex to repress transcription through a yet unidentified domain of ROX1 at the carboxy terminus. In addition, the ability of ROX1 to repress transcription of genes must be intimately associated with its ability to bind DNA. Deletion of binding sites for ROX1 upstream of the *HEM13* transcription start site results in a large increase in expression of *HEM13* under repressing conditions (Keng *et al.*, unpublished observations).

In this study, we have demonstrated that ROX1 is a specific DNA binding protein and that the HMG domain at the amino terminus of ROX1 functions in DNA binding. In addition, we postulate that repression by ROX1 is mediated by an uncharacterized motif found in the carboxy-terminal portion of ROX1.

Similar results have been obtained in a report published while our paper was under revision (Deckert *et al.*, 1995a). These authors show that the first 100 amino acids of ROX1 are sufficient for DNA binding. In addition, they also show that the carboxy-terminus of ROX1 is required for repression. They also demonstrate that mutants in the

HMG domain of ROX1, which bind with reduced affinity, show a reduced ability to repress target genes.

Connecting text

The work presented in chapter 5 demonstrated that ROX1 is involved in repression of *HEM13* in the presence of heme by specifically binding to an upstream operator sequence found in the *HEM13* promoter. In addition, we showed that the region of ROX1 required for DNA-binding resides within the 100 amino-terminal amino acids which constitute the HMG domain of ROX1. We examined the ability of ROX1 to oligomerize, and found that this domain also maps to the 100 amino-terminal amino acids. Therefore, the domains required for ROX1 DNA-binding and oligomerization overlap within the amino-terminus of ROX1.

Chapter 6 deals with a further examination of ROX1 function, in particular, the differential binding of ROX1 to the five potential hypoxic operator sites in *HEM13*. We examined the contribution of each of the five sites to repression of *HEM13* by deleting them individually or in combination, and looking at the effects of the deletions on expression from a *HEM13-lacZ* fusion. The ability of ROX1 to bind to the minor groove of the DNA helix was assessed by employing minor groove binding drugs and oligonucleotides in which A-T base pairs were substituted with I-C base pairs. Finally, we utilized a random PCR-based selection protocol to predict optimal ROX1-binding sites from a pool of random sequences. The results of these experiments comprise the contents of chapter 6.

CHAPTER 6

THE *SACCHAROMYCES CEREVISIAE* ROX1 PROTEIN REPRESSES TRANSCRIPTION OF *HEM13* BY BINDING TO MULTIPLE UPSTREAM SITES WITH DIFFERENTIAL AFFINITY

Celestino Di Flumeri, Alfredo Staffa, Sophia Ushinsky, Nicholas H. Acheson, and Teresa Keng

Department of Microbiology and Immunology, McGill University,
3775 University Street, Montreal, Quebec, Canada, H3A 2B4

This chapter is derived from a manuscript
in preparation.

ABSTRACT

The *HEM13* gene encodes a cytoplasmic oxygen-requiring enzyme, coproporphyrinogen oxidase, which catalyzes the sixth step of the heme biosynthetic pathway in *Saccharomyces cerevisiae*. Transcription of *HEM13* is repressed in the presence of oxygen and heme, through the action of the ROX1 repressor protein. The *ROX1* gene encodes a 368 amino acid protein belonging to the high mobility group (HMG) family of DNA binding proteins. This report describes an analysis of the *HEM13* promoter region, with particular emphasis on locating sites required for *HEM13* repression. In addition, we examined the binding characteristics of ROX1 to these hypoxic operator sites. Five such sites were detected in a region upstream of the transcription start site of *HEM13*, encompassing 419 base pairs of DNA. Mobility gel shift assays demonstrated that the affinity of ROX1 for each of these sites varied substantially. Binding of ROX1 to the site with the strongest affinity occurred primarily through the minor groove of the DNA helix. Deletion analysis demonstrated that four of these sites functioned to repress expression of *HEM13 in vivo*. A PCR-based procedure was employed to select ROX1 binding sites from a pool of random DNA sequences. The selected binding sites contained the core consensus (A/T)TT(T/G)TT, which is identical to the core consensus derived from alignment of regulatory regions of genes repressed by ROX1.

INTRODUCTION

The transcription of a number of genes in the yeast *Saccharomyces cerevisiae* is regulated by complex pathways able to discern oxygen tension in the cell. Genes transcribed preferentially in the presence of oxygen encode mainly cytochromes and respiratory enzymes. These include *CYC1*, which encodes iso-1-cytochrome *c* (Hortner *et al.*, 1982; Guarente and Mason, 1983); *CTT1*, which encodes catalase T (Hortner *et al.*, 1982), and *COX5a,* which encodes the subunit Va isozyme of cytochrome oxidase (Trueblood *et al.*, 1988).

Another set of genes, preferentially transcribed in the absence of oxygen, are referred to as hypoxic genes. Examples of these genes include *ANB1*, *COX5b*, and *HMG2* (Lowry and Lieber, 1986; Trueblood *et al.*, 1988; Hodge *et al.*, 1989; Thorsness *et al.*, 1989). The *ANB1* gene encodes an isoform of the protein encoded by the *TIF51A* gene, translational initiation factor eIF-5A. The *COX5b* and *HMG2* genes encode an isozyme of the subunit V of cytochrome oxidase, and the sterol biosynthetic enzyme 3-hydroxy-3-methylglutaryl coenzyme A reductase, respectively.

Hypoxic genes are regulated by a pathway in which the cellular machinery senses the levels of oxygen and shuts down transcription of these genes when oxygen levels are ample. Heme can mediate the effects of oxygen on gene expression as heme biosynthesis is strictly dependent on the presence of oxygen (Labbe-Bois and Labbe, 1990). Therefore, the effects of oxygen on expression of the hypoxic genes can in many cases be regulated via changes in intracellular heme levels (Zitomer and Lowry, 1992; Pinkham and Keng, 1994). The repression of hypoxic genes is mediated through the action of the

ROX1 repressor. Expression of *ROX1* is induced in the presence of heme (Lowry and Zitomer, 1988) by the HAP1 activator protein (Keng, 1992).

ROX1 is a member of the family of high mobility group (HMG) DNA-binding proteins (Grosschedl *et al.*, 1994). The gene was initially identified in a screen for factors able to render expression of *ANB1* constitutively high under repressing conditions (Lowry and Zitomer, 1984). It was subsequently cloned and sequenced, and was found to encode a 368-amino acid protein with many interesting features (Balasubramanian *et al*, 1993). The 100 amino-terminal residues of ROX1 constitute the highly basic HMG domain, required for specific binding of ROX1 to its target DNA (Deckert *et al.*, 1995a; Di Flumeri *et al.*, 1996). The HMG domain of the protein is also responsible for the ability of ROX1 to produce a significant bend in the DNA helix upon binding to its cognate DNA (Deckert *et al.*, 1995a), and to oligomerize *in vitro* in the absence of DNA (Di Flumeri *et al.*, 1996). Recent experiments suggest that an as yet unidentified domain exists in the carboxy-terminal half of ROX1 which when deleted leads to loss of repression of hypoxic genes *in vivo* irrespective of the ability of the truncated protein to bind to DNA (Deckert *et al.*, 1995a; Di Flumeri *et al.*, 1996). Although the hypoxic genes are activated by a variety of different transcriptional activators, all are repressed through the action of the ROX1 protein. As a result, a common feature of the regulatory regions of all hypoxic genes is the presence of one or more copies of a hypoxic operator site first identified in *ANB1* (Lowry *et al.*, 1990). These sites are also found in the promoter region of the *COX5b* gene (Hodge *et al.*, 1990) and represent binding sites for ROX1 protein. The sites contain the core consensus sequence YYYATTGTTCTC, and

their deletion or mutation resulted in high level expression of *ANB1* and *COX5b* under repressing conditions (Mehta and Smith, 1989; Hodge *et al.*, 1990; Lowry *et al.*, 1990).

Repression of transcription of hypoxic genes in yeast requires the presence of at least two other proteins, SSN6 (CYC8) and TUP1, which function in a complex (Williams *et al.*, 1991; Tzamarias and Struhl, 1994). Repression of a-specific mating type genes, glucose-repressed genes, and hypoxic genes all require the TUP1/SSN6 complex in conjunction with the pathway-specific repressors α2-Mcm 1, Mig 1, and ROX1 (Trumbly, 1986; Schultz and Carlson, 1987; Trumbly, 1988; Williams and Trumbly, 1990; Keleher *et al.*, 1992; Balasubramanian *et al.*, 1993; Amillet *et al.*, 1995; Di Flumeri *et al.*, 1996). The TUP1 and SSN6 proteins have no known DNA-binding ability and are thought to be anchored to specific genes targeted for repression by the pathway-specific repressors which bind to DNA (Nehlin *et al.*, 1991; Keleher *et al.*, 1992; Treitel and Carlson, 1995). The TUP1/SSN6 complex has been proposed to repress transcription either by interference with basal transcription or by production of a repressive chromatin structure (Cooper *et al.*, 1994; Herschbach *et al.*, 1994).

Another protein whose gene is transcriptionally repressed when oxygen or heme is present is coproporphyrinogen oxidase. It is encoded by *HEM13* and catalyzes the sixth step of the heme biosynthetic pathway in yeast. Transcription of this gene is stimulated 40-50 fold in oxygen- or heme-depleted cells (Zagorec and Labbe-Bois, 1986; Zagorec *et al.*, 1988). Like *ANB1* and *COX5b*, *HEM13* repression in the presence of heme is mediated by the ROX1 protein. Expression of *HEM13* is derepressed in a *rox1* mutant strain in the presence of heme (Keng, 1992). ROX1 was shown to be capable of

binding in a sequence-specific manner to a site upstream of *HEM13* homologous to the operator sites found upstream of *ANB1* and *COX5b* (Di Flumeri *et al.*, 1996). Three such operator sites were detected within the *HEM13* promoter (Amillet *et al.*, 1995). Two of these sites were found to account partly for repression of *HEM13* by the ROX1 protein; the third had no effect on *HEM13* expression when mutagenized (Amillet *et al.*, 1996).

To better understand the mechanism by which heme and ROX1 regulate *HEM13* expression, we carried out an extensive deletion analysis to identify operator sites within the *HEM13* regulatory region. We identified a total of five possible hypoxic operator sites in *HEM13*. Three of these are the same as those reported by Amillet *et al.* (1995). Each of these five sites was examined for its function *in vivo* either individually or in combination with other sites. We also tested if the ability of each site to repress expression *in vivo* corresponded to its affinity for ROX1 protein *in vitro*. Oligonucleotide substitution experiments were conducted to determine if ROX1, like other HMG proteins, is capable of binding to the minor groove of the DNA helix. Finally, DNA sequences that bind to ROX1 were selected from a random pool to determine the optimal ROX1 binding site.

RESULTS

We analyzed the upstream non-coding sequences of *HEM13* and identified five sites with homology to those required for repression of *ANB1* (Figure 18). Three of these sites, sites 1, 3, and 4, were also identified by Amillet *et al.* (1995). We proceeded to analyze the function of these five sites and their ability to bind to ROX1 protein.

Binding of GST-ROX1 1-101 protein to *HEM13* sites

We previously demonstrated that ROX1, as well as a GST fusion protein containing the 101 amino-terminal amino acids of ROX1 (ROX1 1-101), can bind specifically to *HEM13* site 3 (Di Flumeri *et al.*, 1996). To determine if the other four repeats were also functional ROX1 binding sites we used double-stranded 33-bp ^{32}P-labeled oligonucleotides corresponding to each of these sites in mobility gel shift analysis. Each of the five oligonucleotides was incubated with 50-100 ng of GST or GST-ROX1 1-101 protein that had been purified by affinity to glutathione-Sepharose (Di Flumeri *et al.*, 1996). In figure 19A, the sequences of the five sites are shown with the core binding sequence shown in bold. The underlined nucleotides depict deviations from the core consensus sequence. The left panel in figure 19B depicts an autoradiogram of binding reactions performed in the presence of GST protein. No specific protein-DNA complexes were noted in lanes 1 through 5, as expected. In contrast, specific protein-DNA complexes were noted when each of the oligonucleotides was incubated with GST-ROX1 1-101 protein (Figure 19B, lanes 6-10).

Site 3 had the highest affinity for binding to ROX1; phosphorimager analysis of this gel showed that the ratio of bound to unbound DNA was 23. Sites 1 and 5 had

Figure 18. Nucleotide sequence of the *HEM13* promoter (Zagorec *et al.*, 1988). The *Nru*I site indicated above the *HEM13* sequence was used in construction of the deletion plasmids. Putative ROX1 binding regulatory sites identified by homology to the consensus YYYATTGTTCTC, are boxed and referred to as sites 1 through 5. The TATA box (underlined) and transcription start site (+1) are indicated. The beginning of the coding region of the gene is shown in bold. Numbers on the right indicate the distance in nucleotides upstream of the transcription start site.

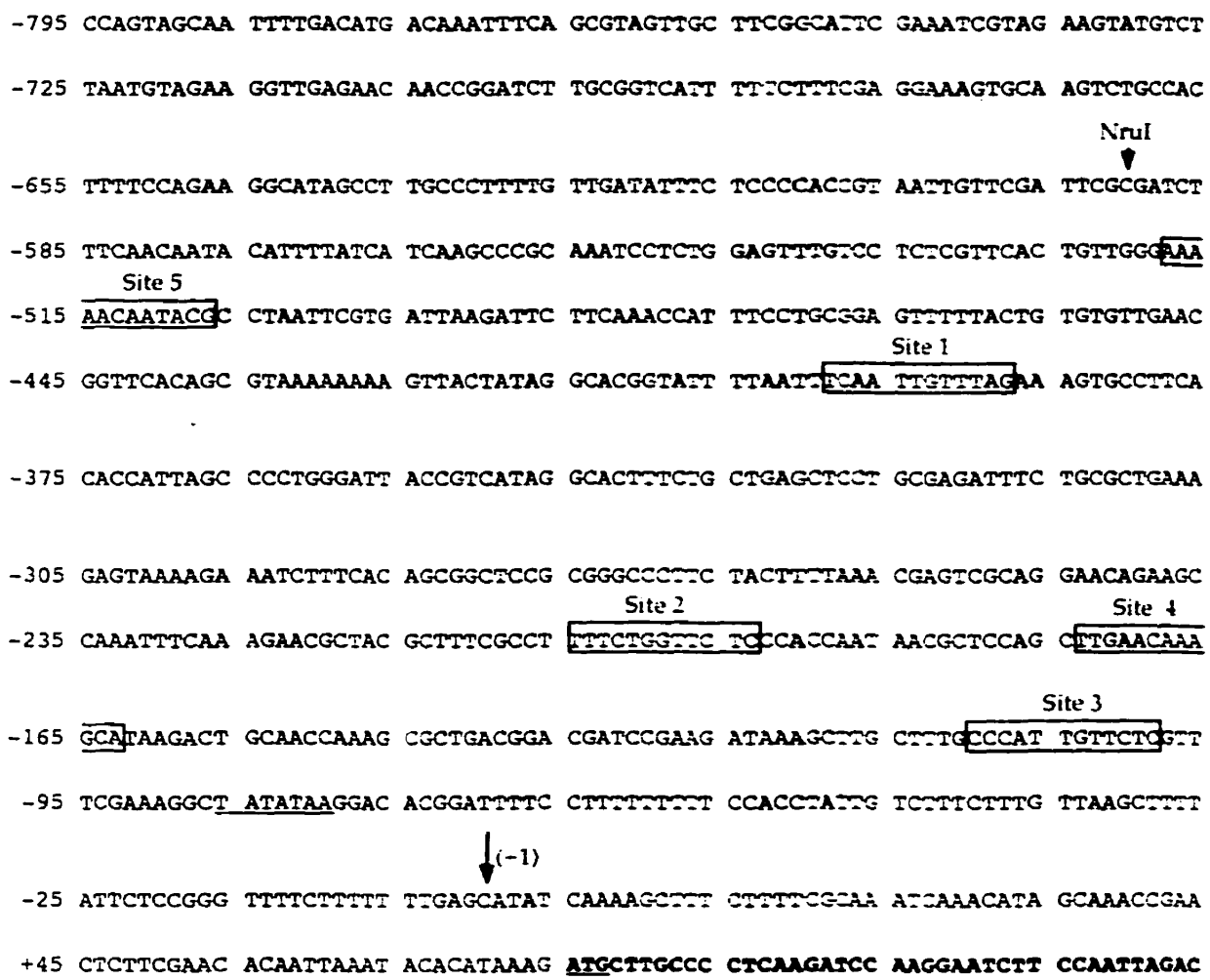
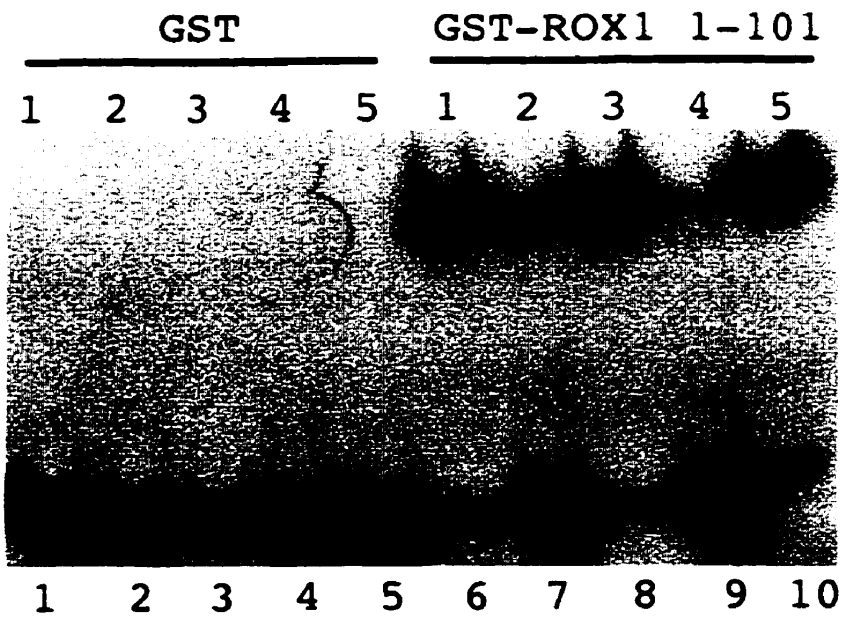

-795 CCAGTAGCAA TTTTGACATG ACAAATTTCA GCGTAGTTGC TTCGGCATTC GAAATCGTAG AAGTATGTCT
-725 TAATGTAGAA GGTTGAGAAC AACCGGATCT TGCGGTCATT TTTCTTTCGA GGAAAGTGCA AGTCTGCCAC
NruI
-655 TTTTCCAGAA GGCATAGCCT TGCCCTTTTG TTGATATTTC TCCCCACCGT AATTGTTCGA TTCGCGATCT
-585 TTCAACAATA CATTTTATCA TCAAGCCCGC AAATCCTCTG GAGTTTGTCC TCTCGTTCAC TGTTGGGAAA
Site 5
-515 AACAATACGC CTAATTCGTG ATTAAGATTC TTCAAACCAT TTCCTGCGGA GTTTTTACTG TGTGTTGAAC
Site 1
-445 GGTTCACAGC GTAAAAAAAA GTTACTATAG GCACGGTATT TTAATCTCAA TTGTTTAGAA AGTGCCTTCA
-375 CACCATTAGC CCCTGGGATT ACCGTCATAG GCACTTTCTG CTGAGCTCCT GCGAGATTTC TGCGCTGAAA
-305 GAGTAAAAGA AATCTTTCAC AGCGGCTCCG CGGGCCCTTC TACTTTTAAA CGAGTCGCAG GAACAGAAGC
Site 2
Site 4
-235 CAAATTTCAA AGAACGCTAC GCTTTCGCCT TTTCTGGTTC TCCCACCAAT AACGCTCCAG CTTGAACAAA
Site 3
-165 GCATAAGACT GCAACCAAAG CGCTGACGGA CGATCCGAAG ATAAAGCTTG CTTTGCCCAT TGTTCTCGTT
-95 TCGAAAGGCT ATATAAGGAC ACGGATTTTC CTTTTTTTTT CCACCTATTG TCTTTCTTTG TTAAGCTTTT
(-1)
-25 ATTCTCCGGG TTTTCTTTTT TTGAGCATAT CAAAAGCTTT CTTTTCGCAA ATCAAACATA GCAAACCGAA
+45 CTCTTCGAAC ACAATTAAAT ACACATAAAG ATGCTTGCCC CTCAAGATCC AAGGAATCTT CCAATTAGAC

Figure 19. Affinity of ROX1 1-101 for different sites upstream of *HEM13* transcription start site. A) The sequence of the five operator sites and their positions upstream of *HEM13* is shown together with the consensus sequence derived from comparison of regulatory regions of ROX1-regulated genes. The sequences in bold represent the ROX1 core binding sequence. Underlined nucleotides represent changes from the consensus core binding sequence. Sites 4 and 5 are shown in inverse orientation in order to emphasize their homology to the consensus sequence. B) Binding affinity was tested using 50-100 ng of either GST or GST-ROX1 1-101 protein. End-labeled double-stranded oligonucleotides representing each of the sites upstream of *HEM13* were incubated with the respective proteins in the presence of 50 ng of poly dI:dC. Complexes were resolved on 6% native polyacrylamide gels at 4°C; gels were dried and exposed to X-ray film.

A)

HEM13 Site	1	-399 to -388	TCA**ATTGTT**TAG
HEM13 Site	2	-205 to -194	TTT**CTGGTT**CTC
HEM13 Site	3	-110 to -99	CCC**ATTGTT**CTC
HEM13 Site	4	-163 to -174	TGC**TTTGTT**CAA
HEM13 Site	5	-507 to -518	CGT**ATTGTT**TTT
Consensus hypoxic sequence			YYY**ATTGTT**CTC

B)



somewhat lower affinities, with ratios of bound to unbound DNA of approximately 10. Sites 2 and 4 were both bound weakly by ROX1, with bound:unbound ratios of less than 1. The strong affinity of binding for site 3 was not surprising as its sequence perfectly matches the consensus sequence YYYATTGTTCTC. Sites 1 and 5 both contain a core ATTGTT sequence identical to that found in the consensus operator but deviate from the consensus in the flanking nucleotides; this may explain why they are bound with reduced affinity by ROX1. *HEM13* sites 2 and 4 showed weak binding probably because their sequences deviate from the core consensus sequence at positions 1 and 3 and at position 1, respectively. These positions are highly conserved in the operators of genes repressed by ROX1 such as *COX5b* and *ANB1*. Therefore, all five ROX1 operator consensus sites identified in the *HEM13* regulatory sequences are capable to different degrees of being bound by ROX1 protein. Much of the specificity of the ROX1-DNA interaction is conferred by the core sequence ATTGTT.

Functional analysis of hypoxic operator sites

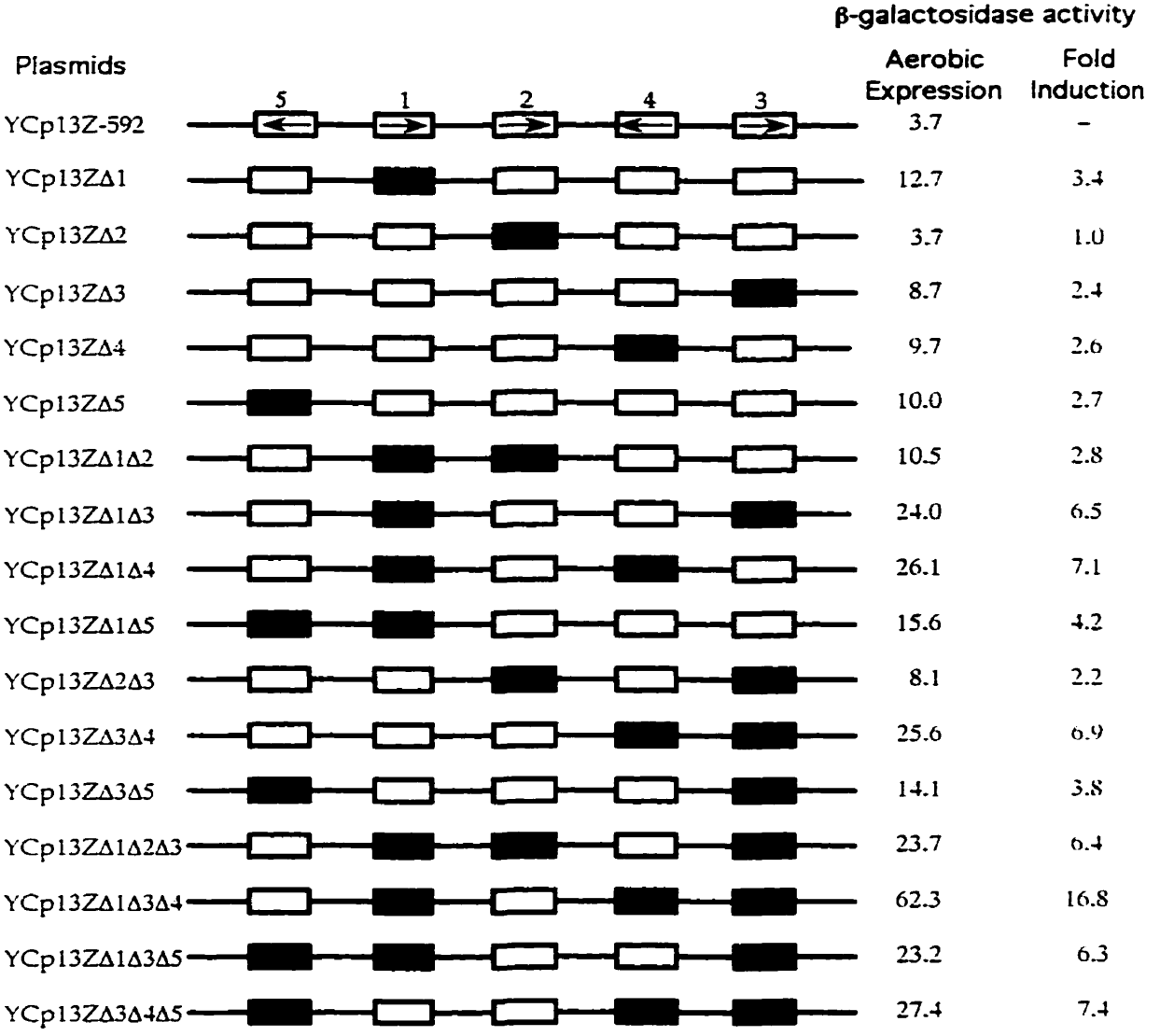
We were interested in determining if the five ROX1 operator sites in *HEM13* represent functional operator sites *in vivo*. Each of these sites was deleted individually and in combination using oligonucleotides which flank each site on both sides. This strategy resulted in precise deletions of the complete 12 bp site; the alternative strategy of site-directed mutagenesis of particular nucleotides within each site could allow residual binding and repressive ability and therefore might not give a clear indication of the importance of each site. The various deletions were constructed as derivatives of plasmid YCp13Z-591. Our analysis has indicated that there are no *HEM13* regulatory sequences

upstream of position -591 (Keng *et al.*, unpublished observations).

Derivatives of YCp13Z-591 carrying the various single or multiple deletions of each site were transformed into the DBY746Δhem1 strain, and the transformants assayed for β-galactosidase activity under repressing conditions (presence of δ-ALA). Expression of *HEM13-lacZ* from YCp13Z-591 was found to be 3.7 units under repressing conditions (Figure 20). Deletion of sites 1, 3, 4, or 5 individually resulted in an average increase of *HEM13* expression under repressing conditions (+δ-ALA) of two- to three-fold (Figure 20). Deletion of site 2 had no effect on *HEM13* expression, indicating that this region plays no role in repression of *HEM13* expression despite its ability to bind to DNA *in vitro*.

To determine if the sites could function independently or in a cooperative manner, we deleted two or three sites in various combinations and measured the expression of *lacZ* driven by these modified promoters (Figure 20). Deletion of site 2 in combination with either site 1 or site 3 (Δ1Δ2 and Δ2Δ3) had no greater effect than deletion of sites 1 or 3 alone, further indicating the lack of repressive effect of site 2 on *HEM13* expression. Deletion of sites 1 and 3, 1 and 4, or 3 and 4 together (Δ1Δ3, Δ1Δ4, Δ3Δ4) led to a 6.5- to 7-fold higher level of expression than that directed by the wild type promoter. This is equivalent to the product of the effects of individual deletions of sites 1, 3, or 4 (2.4- to 3.4-fold), and indicates that these three sites act independently to direct repression of *HEM13*. This conclusion is strengthened by results from triple deletion of sites 1, 3, and

Figure 20. Effect of deletion of hypoxic operator sites on *HEM13-lacZ* expression. Proposed hypoxic operator sites were deleted individually or in combination as outlined in Materials and Methods. The effects of these deletions on expression were monitored by assaying expression from a fusion of *HEM13* to the β-galactosidase gene. Black boxes denote the hypoxic operator site deleted in the respective construct. Arrows indicate the orientation of each site. β-galactosidase activity is shown as Miller units and represents the average of assays done in duplicate on at least two independent transformants of each plasmid. The results for each of the four β-galactosidase determinations for each individual plasmid were within 20% of each other.


β-galactosidase activity
Plasmids
Aerobic Expression
Fold Induction
5
1
2
4
3
YCp13Z-592
3.7
–
YCp13ZΔ1
12.7
3.4
YCp13ZΔ2
3.7
1.0
YCp13ZΔ3
8.7
2.4
YCp13ZΔ4
9.7
2.6
YCp13ZΔ5
10.0
2.7
YCp13ZΔ1Δ2
10.5
2.8
YCp13ZΔ1Δ3
24.0
6.5
YCp13ZΔ1Δ4
26.1
7.1
YCp13ZΔ1Δ5
15.6
4.2
YCp13ZΔ2Δ3
8.1
2.2
YCp13ZΔ3Δ4
25.6
6.9
YCp13ZΔ3Δ5
14.1
3.8
YCp13ZΔ1Δ2Δ3
23.7
6.4
YCp13ZΔ1Δ3Δ4
62.3
16.8
YCp13ZΔ1Δ3Δ5
23.2
6.3
YCp13ZΔ3Δ4Δ5
27.4
7.4

4 (Δ1Δ3Δ4), which showed a further 2.5-fold increase in expression compared with the above pairwise deletions of these sites.

In contrast, deletion of site 5 in conjunction with sites 1 or 3 (Δ1Δ5 and Δ3Δ5) led to only marginally increased expression (3.8-fold to 4.2-fold) compared with single deletions of sites 1, 3, or 5 (2.4- to 3.4-fold). Furthermore, deletion of site 5 in conjunction with double deletions of sites 1 and 3 or 3 and 4 (Δ1Δ3Δ5 and Δ3Δ4Δ5) did not increase expression above the levels directed by the double deletion mutants 1 and 3 or 3 and 4. These results indicate that site 5 can repress *HEM13* gene expression only if sites 1, 3, and 4 are intact (see Discussion).

In summary, sites 1, 3, 4, and 5 all contribute approximately equally to repression of the expression of *HEM13*. Site 2 had no measurable effect on repression. Therefore, even though all five sites can interact with ROX1 protein *in vitro*, not are all functional *in vivo*. Moreover, the relative affinities of ROX1 for the operator consensus sites *in vitro* (Figure 19) are not by themselves indicative of the relative contribution of each site to repression of *HEM13* expression *in vivo* (Figure 20).

ROX1 binds to *HEM13* site 3 primarily through minor groove contacts

A number of the proteins of the HMG family have been shown to contact the DNA helix via minor groove interactions (Starr and Hawley, 1991; Giese *et al.*, 1992; Dooijes *et al.*, 1993). To test whether ROX1 makes contacts with the DNA helix through minor groove contacts we employed drugs known to bind to DNA within the minor groove. The drugs employed were actinomycin D, which binds through contacts with G-

C base pairs; chromomycin A_3, which binds through contacts with A-T and G-C base pairs; and distamycin A, which binds through contacts with A-T base pairs.

ROX1 1-101 and a DNA probe containing site 3 were incubated in the presence of increasing concentrations of each drug and complexes were resolved on acrylamide gels. Actinomycin D inhibited ROX1 protein-DNA complex formation at concentrations between 10 and 100 μM (Figure 21, lanes 5 and 6). In contrast, both chromomycin A_3 and distamycin A inhibited binding at concentrations between 1 and 10 μM (Figure 21, lanes 10 and 11, and 16 and 17, respectively). Therefore, ROX1, like other members of the HMG family of DNA binding proteins, makes contacts with its cognate binding site primarily through the minor groove.

To directly assess the binding of ROX1 in the minor groove of the helix, we used the method developed by Starr and Hawley (1991). Synthetic double-stranded oligonucleotides were generated in which the various A-T and T-A base pairs within the core ATTGTT ROX1 binding sequence of site 3 were selectively replaced with I-C and C-I base pairs respectively (Figure 22A). These substitutions alter the major groove of the DNA without affecting the minor groove (42). The same A-T and T-A base pairs within the ROX1 binding site were also selectively replaced with G-C and C-G base pairs. In oligonucleotides with these substitutions, both the minor and major grooves of the helix are altered. Both the I-C and G-C substituted oligonucleotides, as well as the wild type oligonucleotides, were examined for binding by GST-ROX1 1-101 fusion protein. GST-ROX1 1-101 bound to the I-C substituted oligonucleotides nearly as well as to the wild type oligonucleotide (Figure 22B, compare lanes 3, 4, 6, and 8). In

Figure 21. Effect of minor groove binding reagents on binding of ROX1 1-101 to DNA. ROX1 binding to site 3 was examined in the presence of increasing concentrations of minor groove binding drugs actinomycin D, distamycin A, or chromomycin A_3. A ^{32}P-labeled double-stranded oligonucleotide including site 3 was incubated with 50-100 ng of either GST or GST-ROX1 1-101 protein in the presence of 50 ng of poly (dI:dC) and the indicated amount of drug. Complexes were resolved on 6% native polyacrylamide gels at 4°C, which were dried and exposed to X-ray film. Lane 1 depicts DNA probe incubated in the absence of both drug and ROX1 protein. Lanes 2, 8, and 14 represent binding of ROX1 to site 3 DNA probe in the absence of drug. These lanes were used as controls to show maximal binding of ROX1 to site 3. As controls to ensure that no distortion of the DNA was occuring due to the presence of drug, DNA probe was incubated with 100 μM of each drug in the absence of ROX1 (Lanes 7, 13, and 19).

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Actinomycin D | Distamycin A | Chromomycin A_3

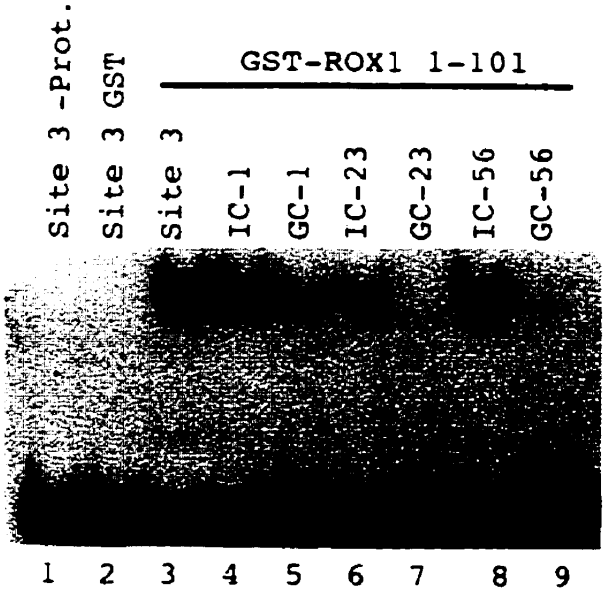
Conc. (μM) 0.1 1.0 10 100 100 | 0.1 1.0 10 100 100 | 0.1 1.0 10 100 100

Figure 22. Effect of substitution of A-T base pairs in the ROX1 core consensus with I-C or G-C base pairs. A) Sequence of the double stranded oligonucleotide probes comprising nucleotides from - 110 to - 99 upstream of the transcriptional start site of *HEM13* (site 3). The ROX1 core binding sequence is shown in bold, and nucleotide substitutions are underlined. B) DNA binding to substituted oligonucleotides was carried out by incubating 50-100 ng of GST or GST-ROX1 1-101 with ^{32}P-labeled oligonucleotides in binding buffer, in the presence of 50 ng of poly (dI:dC). Reactions were incubated at room temperature for 30 min and complexes were resolved on 6% polyacrylamide gels at 20 mAmps at 4°C. Gels were subsequently dried and exposed to X-ray film.

A)

Site 3	5'CCC**ATTGTT**CTC 3' 3'GGG**TAACAA**GAG 5'
IC-1	5'CCC**I**TTGTTCTC 3' 3'GGG**C**AACAAGAG 5'
GC-1	5'CCC**G**TTGTTCTC 3' 3'GGG**C**AACAAGAG 5'
IC-23	5'CCC**ACC**GTTCTC 3' 3'GGGT**II**CAAGAG 5'
GC-23	5'CCC**ACC**GTTCTC 3' 3'GGGT**GG**CAAGAG 5'
IC-56	5'CCC**ATTGCC**CTC 3' 3'GGGTAAC**II**GAG 5'
GC-56	5'CCC**ATTGCC**CTC 3' 3'GGGTAAC**GG**GAG 5'

B)



contrast, substitution of A-T and T-A base pairs with G-C and C-G base pairs greatly reduced or totally abolished ROX1-DNA complex formation (Figure 22B, compare lanes 3, 5, 7, and 9).

The results of the mobility shift experiments in the presence of various DNA binding drugs, together with the experiments with base-substituted binding sites, suggest that ROX1 contacts its DNA target primarily through the minor groove. In particular, the relative affinity of ROX1 for the base-substituted binding sites suggest that ROX1 contacts at least three of the five A-T and T-A base pairs situated within the core consensus binding site ATTGTT.

PCR-mediated selection of optimal ROX1 binding sites

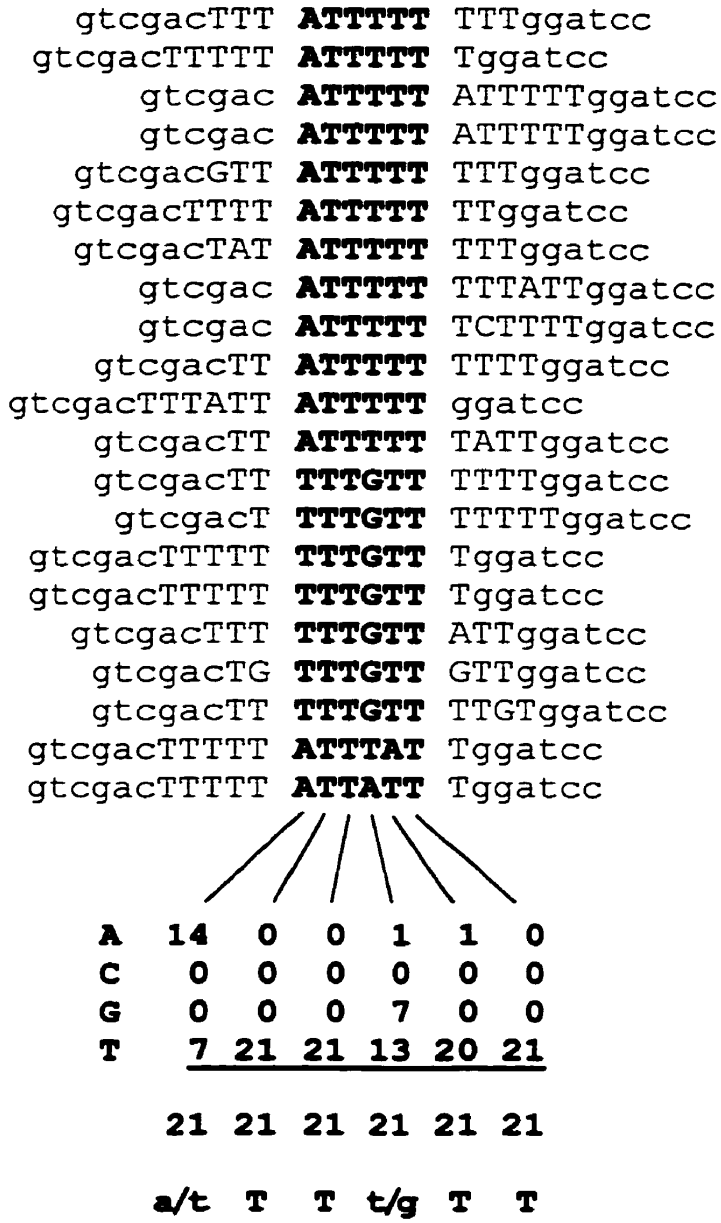
A consensus binding site has been derived for all the hypoxic genes repressed by ROX1. When the ROX1 binding sites within the promoter regions of *ANB1*, *COX5b*, and *HEM13* were aligned, the consensus sequence YYYATTGTTCTC was derived (Y indicates a pyrimidine) (Zitomer *et al.*, 1997). In order to determine if this is the optimal binding site for ROX1, we made use of a 60 bp oligonucleotide which contained random nucleotides at 12 contiguous positions. This oligonucleotide mixture, containing a possible 4^{12} binding sites, was rendered double-stranded and subsequently incubated with purified GST-ROX1 1-101 bound to glutathione Sepharose beads. Oligonucleotides bound to GST-ROX1 1-101 were amplified by PCR, then used in further rounds of binding and amplification to enrich for sequences which are specifically bound to ROX1. This procedure was repeated for a total of four selections, at which time the PCR products were digested with *Sal*I and *Bam*HI and cloned into the similarly digested Bluescript SK+

vector. Binding was specific to ROX1, as there was no amplification of DNA after any of the four cycles when the random oligonucleotide was incubated with GST protein (data not shown). In addition, neither the random unselected oligonucleotide, nor any of the selected sequences from each of the four rounds that specifically bound to GST-ROX1 1-101, bound to GST in a gel shift assay. In contrast, DNA selected after each round of enrichment was bound by GST-ROX1 1-101 in a gel mobility shift assay, whereas the random unselected oligonucleotide showed no binding to GST-ROX1 1-101 (data not shown).

The sequences of 21 individual clones and the derived consensus sequence are shown in figure 23. All 21 sequences were very A/T rich, as expected for DNA bound by an HMG protein (Grosschedl *et al.*, 1994). T residues were found predominantly in one strand; surprisingly, the T-rich strand was the same DNA strand (downstream of the *SalI* site, reading in the 5' to 3' direction) in all 21 clones, as shown in figure 23. This implies that the sequences flanking the selected 12-nucleotide region influenced the binding of ROX1 protein.

The sequences were aligned to best fit the previously-derived consensus sequence ATTGTT. The most common best fit over 6 nucleotides was ATTTTT, which was present a total of 14 times, in 12 out of the 21 clones. The next most common best fit was TTTGTT, present in 7 of the clones. This sequence is present in the hypoxic operator site 4 of the *HEM13* regulatory region. The remaining 2 clones contained an ATTTAT and an ATTATT sequence. None of the clones contained an ATTGTT, even though that was one of the consensus sequences derived from this alignment (Figure 23, bottom).

Figure 23. Alignment of selected ROX1 binding sites. Upper case letters represent the random 12 bp region of the starting 60 bp oligonucleotide. The bold uppercase letters represent the alignment around the core consensus region, whereas the plain uppercase letters represent the flanking regions of the consensus sequence. Lower case letters denote the non-random portion of the oligonucleotide including the *Bam*HI and *Sal*I restriction endonuclease sites. The frequency of occurrence of each nucleotide at the particular position in the 12 bp consensus is shown underneath the aligned sequence. The overall consensus sequence (A/T)TT(T/G)TT is derived from the alignment of the sequences.


gtcgacTTT ATTTTT TTTggatcc
gtcgacTTTTT ATTTTT Tggatcc
gtcgac ATTTTT ATTTTTggatcc
gtcgac ATTTTT ATTTTTggatcc
gtcgacGTT ATTTTT TTTggatcc
gtcgacTTTT ATTTTT TTggatcc
gtcgacTAT ATTTTT TTTggatcc
gtcgac ATTTTT TTTATTggatcc
gtcgac ATTTTT TCTTTTggatcc
gtcgacTT ATTTTT TTTTggatcc
gtcgacTTTATT ATTTTT ggatcc
gtcgacTT ATTTTT TATTggatcc
gtcgacTT TTTGTT TTTTggatcc
gtcgacT TTTGTT TTTTTggatcc
gtcgacTTTTT TTTGTT Tggatcc
gtcgacTTTTT TTTGTT Tggatcc
gtcgacTTT TTTGTT ATTggatcc
gtcgacTG TTTGTT GTTggatcc
gtcgacTT TTTGTT TTGTggatcc
gtcgacTTTTT ATTTAT Tggatcc
gtcgacTTTTT ATTATT Tggatcc
A 14 0 0 1 1 0
C 0 0 0 0 0 0
G 0 0 0 7 0 0
T 7 21 21 13 20 21
21 21 21 21 21 21
a/t T T t/g T T

DISCUSSION

In this report we describe the identification of five consensus hypoxic operator sites within the upstream regulatory region of *HEM13*. We demonstrate that ROX1 directly contacts each of these sites with different affinity and that binding to site 3 occurs through minor groove contacts with A-T base pairs within the core binding site. The consequences of deletion of ROX1 binding sites on expression of *HEM13 in vivo* suggest that repression of *HEM13* requires four of these sites, three of them (1, 3, and 4) acting independently of each other, and one (site 5) active only in concert with the other three. In addition, we show that A-T rich sequences are selected by binding to the ROX1 protein from a random pool of sequences.

Previous studies on the promoter regions necessary for repression of hypoxic genes suggested a requirement for one or more copies of a consensus sequence YYYATTGTTCTC. Experiments on *ANB1* demonstrated that deletion of sequences between -313 and -186 upstream of the transcription start site rendered expression of this gene constitutively high under repressive conditions (Lowry *et al.*, 1990). A plasmid containing the consensus hypoxic operator sequence CCCATTGTTCTC upstream of *ANB1* caused repression of this gene, indicating that this 12 bp sequence is an integral component of the operator. Other experiments also demonstrated that a sequence in the *COX5b* promoter containing the consensus operator site could cause repression of a heterologous gene (Hodge *et al.*, 1990). Subsequently, it was shown that a *trans*-acting factor, ROX1, could bind specifically to an oligonucleotide containing the consensus sequence (Balasubramanian *et al.*, 1993) and an oligonucleotide containing a natural site (site 3) found within the *HEM13* promoter (Di Flumeri *et al.*, 1996). Studies examining

the binding to all individual natural sites within either the *ANB1*, *COX5b*, or *HEM13* promoters were not previously performed.

In this study, we examined binding of ROX1 to all five possible operator sites identified within the promoter region of *HEM13*. In addition, we examined the contribution of each individual site to repression. Binding to site 3, the site that most closely resembled the consensus sequence, was strongest. Two other sites, 1 and 5, contain the core consensus ATTGTT sequence but vary from the consensus sequence in the flanking regions. These two sites were bound by ROX1 with a decreased affinity compared to site 3. Site 4, which has the core sequence TTTGTT, had a much reduced affinity for ROX1. Surprisingly, site 2, which has the core sequence CTGGTT, was bound by ROX1 protein with low affinity, in spite of two substitutions in the highly conserved core motif ATTGTT.

We determined if any of these identified consensus sites mediates the effects of oxygen and heme on expression of *HEM13 in vivo*. Each consensus site was deleted alone, as well as in various combinations with one another, and the effects of these deletions on *HEM13* expression were determined. Sites 1, 3, 4, and 5 were functional and each site made a similar contribution to repression of *HEM13* expression *in vivo*; deletion of any one of these sites resulted in a 2.5- to 3.5-fold increase in aerobic expression of a *HEM13-lacZ* fusion (Figure 20). The contribution of each individual site towards repression cannot be correlated with the affinity of ROX1 to that site. Deletion of site 3, which showed the highest affinity for ROX1, did not result in a larger derepression of *HEM13* expression than deletion of site 4, which had a reduced affinity for ROX1.

These observations are at odds with those of Amillet *et al.* (1996), who demonstrated that mutagenesis of site 1 resulted in only a 1.3- to 1.5-fold increase in expression of *HEM13* while mutagenesis of site 3 resulted in a 3.6- to 3.8-fold increase. More importantly, those authors reported that site 4 did not function in repression of *HEM13* expression, and they did not test the effect of site 5. In our experiments, site 4 was found to be fully functional; a deletion of this site resulted in a 2.7-fold increase in expression over that of the wild type promoter. Moreover, a combined deletion of site 4 with site 1 or site 3 (Figure 20, plasmids YCp13ZΔ1Δ4, YCp13ZΔ3Δ4) resulted in 24 to 26 units of expression, which represents a further increase in *HEM13* expression of approximately 2.5-fold over that observed with deletion of site 1 or site 3 alone. Furthermore, a deletion of site 4 in conjunction with a double deletion of sites 1 and 3 (Figure 20, YCp13ZΔ1Δ3Δ4) resulted in 62.3 units of expression, an overall increase of 16.8-fold over that observed with the wild-type promoter. The differences between the observations reported here and those made by Amillet *et al.* (1996) may be due to the way the different sites were analyzed. We made precise deletions of the consensus sites while Amillet *et al.* (1996) mutagenized the sites by multiple base substitutions.

Site 5, with the core sequence ATTGTT, showed a relatively high affinity for the ROX1 protein and its deletion resulted in a 2.7-fold increase in *HEM13* expression. Interestingly, deletion of site 5 in conjunction with sites 1 and 3 or sites 3 and 4 did not have any effect on expression above that observed upon deletion of sites 1 and 3 or sites 3 and 4 by themselves. Therefore, it would appear that when *HEM13* is fully repressed (when sites 1, 3, and 4 are present and bound by ROX1), site 5 contributes to repression.

However, when sites 1 and 3, 3 and 4, or 1, 3, and 4 are deleted and *HEM13* is largely derepressed, site 5 makes no contribution to repression. These results suggest that occupancy of sites 1, 3, and 4 by ROX1 may bring the ROX1-bound site 5 located some 500 bp upstream of the transcription start site, in closer proximity to the basal transcription machinery and enable site 5 to contribute to repression of *HEM13* expression.

Site 2, which had a low affinity for ROX1, but a higher affinity than site 4, did not appear to mediate the effects of oxygen and heme on expression of *HEM13 in vivo*. Deletion of site 2 by itself had no effect on expression of *HEM13* and deletion of site 2 in conjunction with other functional sites (Figure 20, YCp13ZΔ2, YCp13ZΔ1Δ2, YCp13ZΔ1Δ2Δ3) did not result in additional effects on *HEM13* expression.

Experiments utilizing minor groove binding drugs suggest that ROX1, like many of the HMG proteins, binds DNA through contacts with the minor groove. The substitution of specific A-T base pairs within the core ATTGTT binding site with I-C base pairs provided a more direct method for assessing minor groove contacts and indicated which A-T pairs were involved. ROX1 was still capable of binding to oligonucleotides when adenine residues on either position 1, positions 2 and 3, or positions 5 and 6 were substituted by inosine residues, but not when they were substituted with guanine residues. However, binding was not as strong as binding to the unsubstituted oligonucleotide, suggesting that determinants in the major groove may also make contributions to the binding affinity of ROX1.

In vitro selection of ROX1 binding sites showed that A-T rich sequences with a

consensus (A/T)TT(T/G)TT were the favoured binding sites. The sequences ATTTTT and TTTGTT appeared with highest frequency. The sequence TTTGTT does appear in the *HEM13* promoter (site 4), but is bound with reduced affinity by ROX1 compared to the sequence ATTGTT. Interestingly, the sequence ATTGTT was not selected, although this is the consensus sequence derived from comparison of the regulatory regions of ROX1-responsive genes. A similar observation was made with regards to Sox-5, another member of the HMG protein family. In a gel mobility shift assay, Sox-5 bound with equal affinity to either ATTGTT or TTTGTT, whereas in a selection protocol similar to the one reported here, the selected sequence ATTGTT was found in 17 of 21 clones, while no clones were recovered that had the sequence TTTGTT. This difference was attributed to differences in specificity which were significant in the selection procedure but not in a gel shift assay (Denny *et al.*, 1992). This difference in specificity between the selection procedure and the gel shift assay may also have resulted in our not selecting binding sites with the core sequence of ATTGTT. In addition, if binding of ROX1 to the core region were affected by the flanking regions, then the use of an oligonucleotide with a random region of only 12 bp in the selection protocol may have been insufficient. This could also explain the observation that all clones which were selected had the T-rich sequence on the same DNA strand, strengthening the hypothesis that the sequences flanking the core ROX1 binding site are important determinants for binding of ROX1 to its cognate site. Although the sequence ATTTTT does not appear within any of the sites identified in the *HEM13* promoter, a site-directed mutant within an *ANB1* site that resulted in a substitution from ATTGTT to ATTTTT retained partial function (Mehta and Smith, 1989), suggesting it was still capable of being bound by ROX1.

Analysis of the deletion data in conjunction with the DNA binding data suggests no strict correlation between affinity of ROX1 for different hypoxic operator sites and functionality of these sites *in vivo*. Instead, the data would suggest a more complex scenario taking into account the ability of ROX1 to bend DNA (Deckert *et al.*, 1995a) and oligomerize *in vitro* (Di Flumeri *et al.*, 1996). DNA bending by ROX1 could facilitate interaction of ROX1 molecules bound at different repeats with the formation of a complex nucleoprotein structure. This complex could include SSN6 and TUP1, which presumably could interact with the carboxy terminus of ROX1 (Deckert *et al.*, 1995a). This nucleoprotein complex can mediate repression of *HEM13* presumably through interaction of TUP1/SSN6 with the basal transcription machinery.

Connecting text

Chapter 6 of this thesis dealt with a further analysis of the function of the ROX1 protein with particular emphasis on locating the hypoxic operator sites involved in repression of *HEM13* expression. Five potential sites were found and were deleted individually and in combination with each other. The effects of the deletions on repression of *HEM13* expression were tested by examining the level of β-galactosidase expression from a *HEM13-lacZ* fusion. Three of these sites were found to act in an additive manner in bringing about repression of *HEM13*, and a fourth was found to be active only when the other sites were intact. We also tested the affinity of ROX1 for the different sites and found that affinity does not necessarily correlate with the *in vivo* ability of the site to repress *HEM13* expression. Utilizing minor groove binding drugs and I-C substituted oligonucleotides, the binding of ROX1 to site 3 was found to occur via the minor groove of the DNA helix. Finally, a random PCR-based selection procedure predicted the optimal ROX1-binding site to contain the consensus 5'-(A/T)TT(T/G)TT-3'.

Chapter 7 of this thesis focuses on the repression domain of ROX1. To delineate the repression domain of ROX1, we constructed carboxy-terminal deletion derivatives of ROX1 and determined the effect of these derivatives on expression of β-galactosidase from a *HEM13-lacZ* fusion in a *rox1::LEU2* disrupted strain. In addition, we assessed the role of the SSN6 and TUP1 proteins in ROX1-mediated repression of *HEM13* expression. β-galactosidase expression from a *HEM13-lacZ* fusion was measured in a

strain containing chromosomal disruptions of either *SSN6* or *TUP1* and a wild type copy of *ROX1*. The results of these experiments are the focus of chapter 7.

CHAPTER 7

Repression of *HEM13* is mediated by the carboxy-terminal region of ROX1 and requires the SSN6/TUP1 complex

Celestino Di Flumeri, Nicholas H. Acheson, and Teresa Keng

Department of Microbiology and Immunology, McGill University,
3775 University Street, Montreal, Quebec, Canada, H3A 2B4

This chapter is derived from a manuscript submitted to *Nucleic Acids Research.*

ABSTRACT

The *ROX1* gene encodes a protein belonging to the High Mobility Group (HMG) family of DNA-binding proteins. ROX1 functions as a repressor of hypoxic gene expression in the presence of oxygen and heme. *HEM13*, encoding coproporphyrinogen oxidase, the sixth step of the heme biosynthetic pathway, is a hypoxic gene. ROX1 mediates repression of *HEM13* expression by binding to operator sites upstream of the *HEM13* coding region. DNA binding by ROX1 is mediated through the HMG domain, an 80 amino acid region rich in basic residues, which is localized to the extreme amino-terminal end of ROX1.

In an effort to delineate the region of ROX1 required for repression, we made use of a yeast strain carrying an integrated *HEM13-lacZ* fusion as well as a disrupted allele of *ROX1*. A plasmid containing a wild type copy of *ROX1* was fully capable of restoring repression of *HEM13-lacZ* expression in the presence of heme. However, yeast strains transformed with plasmids missing either 92 or 192 amino acids from the carboxy-terminus of ROX1 were only able to partially repress expression from the *HEM13-lacZ* fusion. This result suggests that the carboxy-terminus of ROX1 is involved in repression of *HEM13*.

The TUP1 and SSN6 proteins form a heteromeric complex which is recruited to promoters targeted for repression by pathway specific DNA-binding proteins. We examined the requirement for the TUP1/SSN6 complex with respect to *HEM13* repression. Isogenic strains carrying disrupted alleles of either *TUP1* or *SSN6* genes showed marked derepression of *HEM13-lacZ* expression suggesting a role for the TUP1/SSN6 complex in *HEM13* expression.

INTRODUCTION

Transcriptional repression in eucaryotes is an integral part of the genetic circuitry required for cells to regulate expression of genes. Cells have evolved a variety of mechanisms to regulate transcription of genes in a negative fashion. Repressors can function by interfering at various points of the transcriptional process. For example, a repressor binding to an operator site may prevent binding of an activator to an overlapping binding site. A more common mechanism by which repressors function involves a direct interaction with components of the basal transcription machinery.

In *Saccharomyces cerevisiae*, SSN6 (CYC8) and TUP1 are involved in repression of genes which are regulated by oxygen, cell type, and glucose. Mutants of *ssn6* were first isolated as yeast strains with elevated expression of iso-2-cytochrome *c* encoded by the *CYC7* gene (Rothstein and Sherman, 1980). Later, other *ssn6* mutant yeast strains deficient in repression of glucose repressible genes were isolated (Carlson *et al.*, 1984). The isolation of *tup1* mutants occured in a screen for mutants which could take up dTMP from the growth media, hence the designation *tup1* (Wickner, 1974). The *ssn6* and *tup1* mutants both have similar pleiotropic phenotypes that include flocculance, an inabilty to maintain minichromosome plasmids stably, defects in sporulation of homozygous diploids, and mating type defects in MATα cells. These diverse phenotypes are a result of the inability of *ssn6* and *tup1* mutants to repress different classes of genes.

Both SSN6 and TUP1 are large proteins, having open reading frames of 966 and 713 amino acids, respectively (Schultz and Carlson, 1987; Trumbly, 1988; Williams and Trumbly, 1990). The SSN6 protein contains ten copies of a tetratricopeptide repeat

(TPR) in the amino terminus of the protein which is the functional domain of the protein (Schultz and Carlson, 1990). The TUP1 protein contains six or seven copies of a sequence similar to that found in the β-transducin class of proteins. This sequence is characterized by high conservation of tryptophan and aspartate residues within forty amino acid stretches and is referred to as the WD repeat. The TPR and WD repeats are thought to mediate protein-protein interactions. The TUP1 and SSN6 proteins were first shown to associate in a high-molecular weight complex by coimmunoprecipitation (Williams *et al.*, 1991), and recently it has been demonstrated that this complex involves a ratio of four molecules of TUP1 for every molecule of SSN6 (Varanasi *et al.*, 1996).

The SSN6-TUP1 repressor complex appears to interact with a diverse group of pathway-specific repressor proteins: 1) ROX1, a repressor of hypoxic genes (Zitomer *et al.*, 1997a); 2) MIG1, a zinc-finger protein responsible for repression of glucose-repressible genes (Treitel and Carlson, 1995); 3) MATα2 repressor, a homeobox protein required for repression of **a**-specific genes (Keleher *et al.*, 1992); and 4) Rgt1 repressor which represses genes encoding glucose transporters (Özcan and Johnson, 1995). The SSN6 and TUP1 proteins have no ability to bind to DNA. The pathway-specific repressors contact the regulatory sequences of the genes targeted for repression, and subsequently recruit the SSN6/TUP1 complex to the promoter (Tzamarias and Struhl, 1995). TUP1 fused to the DNA-binding domain of LexA is capable of repressing transcription of a reporter gene with a LexA binding site independently of SSN6, whereas SSN6 requires TUP1 for repression (Keleher *et al.*, 1992; Tzamarias and Struhl, 1994). This suggests that repression is mediated directly by TUP1. The SSN6 protein is thought

to be involved in interactions with the specific repressor proteins, thus anchoring the TUP1-SSN6 complex to the promoter targeted for repression (Tzamarias and Struhl, 1995). Once anchored to the promoter, the TUP1/SSN6 complex is proposed to repress transcription of the target gene by nucleosome positioning and/or interference with the basic transcription apparatus (Cooper *et al.*, 1994; Herschbach *et al.*, 1994).

Repression of transcription of *HEM13* in the presence of heme requires the product of the *ROX1* gene (Keng, 1992). Under conditions of heme availability, ROX1 binds to hypoxic operator sites within the *HEM13* promoter (Di Flumeri *et al.*, 1996). The putative involvement of the SSN6/TUP1 complex in ROX1 repression of *HEM13* expression was examined by assaying expression of a *HEM13-lacZ* fusion in *tup1* and *ssn6* mutant strains. In addition, we delineated the region of ROX1 required for repression by looking at the effect of carboxy-terminal deletion derivatives of ROX1 on repression of a *HEM13-lacZ* fusion.

RESULTS

The TUP1 and SSN6 proteins have been demonstrated to be required for repression of genes mediated by pathway specific repressors such as MIG1 (Treitel and Carlson, 1995), ROX1 (Balasubramanian *et al.*, 1993), and α2 (Keleher *et al.*, 1992). Deletion of *SSN6 or TUP1* resulted in derepression of *ANB1*, a hypoxic gene under ROX1 control, under aerobic conditions (Zhang *et al.*, 1991; Balasubramanian *et al.*, 1993). *HEM13* encodes an enzyme required for heme biosynthesis; its expression is also repressed under aerobic conditions by the ROX1 protein (Keng, 1992).

We were interested in determining if ROX1-mediated repression of *HEM13* required the products of the *SSN6* and *TUP1* genes. Thus, we constructed yeast strains which were deleted for the chromosomal copy of *SSN6* or *TUP1* and assayed expression of a *HEM13-lacZ* fusion which had been integrated into the genome. Assay for expression from the wild type strain CDF22 in the presence of heme resulted in 4 units of β-galactosidase activity (Table 5). This is consistent with synthesis of the ROX1 repressor in the presence of heme and its repression of expression from *HEM13-lacZ.* Expression of *HEM13-lacZ* from an otherwise isogenic strain containing a disruption of *ROX1* (a *rox1::LEU2* allele) was increased 18-fold to 72 units. This finding was also in agreement with previous reports on *HEM13* repression by ROX1 (Keng, 1992). In other isogenic strains containing either disrupted *SSN6* or *TUP1* genes (*ssn6::LEU2* or *tup1::LEU2* alleles), a 7- to 8-fold derepression of *HEM13-lacZ* expression was observed under repressing conditions. This suggests that both TUP1 and SSN6 proteins are

Table 5. Effects of *SSN6* and *TUP1* mutations on *HEM13-lacZ* expression

Strain	Genotype	β-galactosidase activity	
		+heme (50 μg/ml δ-ALA)	Fold induction
CDF22	*Δhem1 HEM13-lacZ*	4	–
CDF23	*Δhem1 HEM13-lacZ rox1::LEU2*	72	18
CDF24	*Δhem1 HEM13-lacZ ssn6::LEU2*	28	7
CDF25	*Δhem1 HEM13-lacZ tup1::LEU2*	30	8

required for repression of *HEM13* by ROX1. This observation is consistent with the effect of deletion of *SSN6* and *TUP1* on expression of other hypoxic genes which are also under ROX1 control. In the case of *ANB1*, deletion of either gene resulted in a 13-fold increase in expression of an *ANB1-lacZ* fusion (Zhang *et al.*, 1991; Balasubramanian *et al.*, 1993). Recently, Amillet *et al.* (1995) also examined the effects of *TUP1* and *SSN6* deletions on *HEM13* expression. In aerobically grown *tup1* and *ssn6* mutant strains, coproporphyrinogen oxidase activity was found to be increased 10-fold and 35-fold, respectively (Amillet *et al.*, 1995).

To delineate the region of ROX1 required for repression, we individually introduced plasmids containing carboxy-terminal deletions of ROX1 into strain CDF23, which contains a *rox1::LEU2* mutation. In strain CDF22, which has a wild type copy of *ROX1*, the integrated *HEM13-lacZ* fusion expressed 4 units of β-galactosidase activity in the presence of heme. In strain CDF23, 78 units of β-galactosidase activity were expressed, representing a 20-fold increase. Introduction of the wild type copy of *ROX1* on the single copy plasmid pRSROX1 into strain CDF23 restored repression of the *HEM13-lacZ* fusion (Table 6). Plasmid pRSROX1 1-176, which makes a ROX1 protein lacking 192 amino acids at the carboxy-terminus, allowed expression of 46 units of β-galactosidase activity. Therefore, loss of these 192 amino acids resulted in an 11-fold increase in *HEM13-lacZ* expression. Plasmid pRSROX 1-276, which makes a ROX1 protein lacking 92 amino acids at the carboxy-terminus, allowed expression of an equivalent level of *HEM13-lacZ* (43 units). Interestingly, neither construct allowed

Table 6. Effects of carboxy-terminal deletions of ROX1 on *HEM13-lacZ* expression

Strain	Genotype	Plasmid	β-galactosidase activity	
			+heme (50 μg/ml δ-ALA)	Fold induction
CDF22	*Δhem1 HEM13-lacZ*	pRS316	4	–
CDF23	*Δhem1 HEM13-lacZ rox1::LEU2*	pRS316	78	20
CDF23	*Δhem1 HEM13-lacZ rox1::LEU2*	pRSROX1	4	–
CDF23	*Δhem1 HEM13-lacZ rox1::LEU2*	pRSROX1 1-276	43	11
CDF23	*Δhem1 HEM13-lacZ rox1::LEU2*	pRSROX1 1-176	46	11

complete derepression of *HEM13-lacZ* expression. These results suggest that the carboxy-terminal 92 amino acids of ROX1 contains repressing activity. However, since a ROX1 derivative with the amino-terminal 276 residues still retains some ability to repress *HEM13* expression, the entire repression domain cannot reside solely within the carboxy-terminal 92 amino acids.

DISCUSSION

In this study we demonstrated an involvement of the TUP1 and SSN6 proteins in the repression of *HEM13* expression by ROX1. In addition, we showed that the carboxy-terminal region of ROX1 is involved in mediating repression of *HEM13*.

The DNA-binding domain of ROX1 resides within the amino-terminal 100 amino acids of the protein, and contains homology to the high mobility group family of DNA-binding proteins. The protein has been demonstrated to bind to the promoters of both the *ANB1* and *HEM13* genes (Balasubramanian *et al.*, 1993; Di Flumeri *et al.*, 1996). The repression of *ANB1* by ROX1 requires the products of the *SSN6* and *TUP1* genes (Balasubramanian *et al.*, 1993). Therefore, we examined the effect of *TUP1* and *SSN6* deletions on *HEM13* expression. Deletion of either gene resulted in derepressed levels of *HEM13* expression. The increase in expression is similar to what had been observed for *ANB1* expression in *tup1* or *ssn6* mutant strains (Zhang *et al.*, 1991; Balasubramanian *et al.*, 1993). Amillet *et al.* (1995) found that deletion of *TUP1* resulted in a 10-fold induction of coproporphyrinogen oxidase activity, a result which is in agreement with our analysis. However, in their assay, deletion of *SSN6* led to a 35-fold induction of coproporphyrinogen oxidase activity. Considering that ROX1 and the TUP1/SSN6 complex are postulated to function in the same pathway, this result is unexpected. It has been previously shown that *SSN6* and *TUP1* are also required for aerobic auto-repression of *ROX1* (Deckert *et al.*, 1995b). An *SSN6* deletion resulted in *ROX1-lacZ* expression that was 3-fold higher than that observed with a *TUP1* deletion. (Deckert et al., 1995b). This result was attributed to an indirect effect of the *tup1* and *ssn6* mutations. Strains with mutant alleles of these genes are extremely flocculant. Consequently, differences in

ROX1 expression between the *tup1* and *ssn6* mutants could be attributed to subtle differences in flocculance, and thereby aeration. This may also account for the differences in coproporphyrinogen oxidase activity between the *tup1* and *ssn6* mutants observed by Amillet *et al.* (1995).

Deletion analysis revealed that part of ROX1 required for repression of *HEM13* expression is located within the carboxy-terminal 92 amino acids, between residues 277-368. Similar studies looking at repression of *ANB1* demonstrated that a ROX1 deletion construct missing the carboxy-terminal 192 amino acids resulted in a 16-fold derepression of *ANB1-lacZ* expression. Our analysis is in agreement with this data. The expression of ROX1 1-176 in our experiments resulted in an 11-fold increase in *HEM13-lacZ* expression. This observation leads us to conclude that the region between amino acids 177 and 368 contain part of the sequences required for *HEM13* repression. However, the ROX1 derivative with a deletion of the 92 amino acids (pRSROX1 1-276) allowed the same level of *HEM13* expression as ROX1 1-176, suggesting that the region required for repression can be further delineated to the carboxy-terminal 92 amino acids. Interestingly, in the analysis of the regions of ROX1 required for repression of *ANB1*, the repression domain was found to be redundant and can be divided into two regions, both of which contribute to repression. These two regions were found to lie between amino acids 123-246 and 246-368 (Deckert *et al.*, 1995a) of ROX1. Deletion of either one of these regions by itself (Δ100-245 or Δ247-368) had minimal effect on *ANB1* expression, and resulted in only a 2-3-fold increase in expression. Thus, both halves of the repression domain on their own retain strong repressive ability with respect to *ANB1* expression. In

the case of *HEM13*, most of the repression appears to be carried out by the extreme carboxy-terminal 92 amino acids, between amino acids 277 and 368.

If repression of *HEM13* were mediated through the carboxy-terminal 92 amino acids of ROX1, it should be possible to confer this repressive ability to a heterologous DNA-binding domain. Fusion of either amino acids 124-368 or 246-368 to the DNA-binding domain of GAL4 resulted in a hybrid protein which could repress expression of an *ANB1-lacZ* fusion containing two copies of the GAL4 binding site (Deckert *et al.*, 1995a). The repression by GAL4-ROX1 was also shown to require the SSN6 protein, since expression of *ANB1-lacZ* was derepressed in the Δ*ssn6* strain. Thus, it is tempting to speculate that ROX1 binds to the hypoxic operator sites upstream of *HEM13*, and through its carboxy-terminal region recruits the TUP1/SSN6 complex to the promoter. It is possible that there is not an interaction between ROX1 and TUP1/SSN6, but that rather a novel protein interfaces between them. This will require additional studies with highly purified proteins.

CHAPTER 8. SUMMARY AND CONCLUSIONS

Heme biosynthesis in *Saccharomyces cerevisiae* requires the coordinate synthesis of porphyrin intermediates by eight enzymatic activities localized to either the cytoplasmic or mitochondrial compartments. The pathway in *Saccharomyces cerevisiae* is identical to that found in other eucaryotic systems with one exception; coproporphyrinogen oxidase is cytoplasmic in yeast but mitochondrial in other eucaryotic systems (Camadro *et al.*, 1986). Yeast mutants defective in all eight enzymatic steps have been isolated, allowing cloning of the genes which encode the enzymes and subsequent analysis of their gene expression. We were interested in the genes encoding enzymes involved in heme biosynthesis and in examining how synthesis and activity of these enzymes are regulated.

A number of heme auxotrophs were isolated which belong to different complementation groups and encode different enzymes in the pathway (Keng *et al.*, 1992). One of the genes isolated by complementation of these mutant strains, *HEM3*, encodes porphobilinogen deaminase, a cytoplasmic enzyme catalyzing the third step in the pathway leading to heme biosynthesis. Transcription of this gene was found to be unaffected by heme or carbon source. *HEM3* expression requires the HAP/2/3/4/5 complex, and a binding site for this complex, 5'-TTATTGGT-3', is located within the *HEM3* regulatory sequences (Keng *et al.*, 1992).

In this thesis, we report the cloning of the *HEM6* gene via complementation of strain 13, a heme auxotroph which was isolated by ethyl methanesulfonate mutagenesis of strain BWG9a-1. This mutant strain was found to accumulate uroporphyrin III, derived

from an intermediate that is the substrate for uroporphyrinogen decarboxylase, and thus deduced to contain a mutation in *HEM6*. A plasmid containing a 7.3 kb fragment of yeast genomic DNA (pTK600) was able to complement the heme auxotrophy in strain 13. A restriction fragment of 2.3 kb subcloned from plasmid pTK600 was also capable of complementing the heme auxotrophy of this mutant strain. We deduced that the coding region of the *HEM6* gene resided on this 2.3 kb fragment and proceeded to determine the DNA sequence of this region. A single open reading frame of 1086 nucleotides coding for a protein of 362 amino acids was obtained (Di Flumeri *et al.*, 1993). This corresponds to a protein with a predicted molecular weight of 41 kDa, a value consistent with the molecular weight observed for purified yeast uroporphyrinogen decarboxylase (Felix and Brouillet, 1990). The deduced amino acid sequence of this protein was aligned with those of rat and human uroporphyrinogen decarboxylases and was found to be highly homologous to those sequences.

Expression of *HEM6* was examined both by looking at steady-state RNA levels, and by meausuring β-galactosidase expression from a *HEM6-lacZ* fusion. Cells were grown in the presence or absence of heme, as well as in different carbon sources. As with expression of *HEM3*, there was no regulation of *HEM6* expression by heme levels. However, when cells were grown in lactate, a 2-fold increase in mRNA levels was noted above those seen in cells grown in glucose. This observation is in agreement with previous studies that measured uroporphyrinogen decarboxylase activity in non-fermentable carbon sources (Kurlandzka *et al.*, 1988). Expression from the *HEM6-lacZ* fusion was also induced 2-fold in lactate grown cells. Thus, expression of *HEM6* seems to be relatively constitutive in the same manner as that of *HEM1* (Keng and Guarente,

1987) and *HEM3* (Keng *et al.*, 1992). However, *HEM6* differs from both *HEM1* and *HEM3* in that both *HEM1* and *HEM3* require the HAP2/3/4/5 complex for expression while *HEM6* does not. β-galactosidase activity was reduced 5 to 6-fold when either *HEM1-lacZ* or *HEM3-lacZ* constructs were introduced into a *hap2* mutant strain (Keng and Guarente, 1987; Keng *et al.*, 1992). Although the upstream regulatory sequences of *HEM6* contain a sequence between positions -566 to -559 5'-(CAATTGGT)-3' which bears homology to the HAP2/3/4/5 binding site, expression of a *HEM6-lacZ* fusion was not affected in a *hap2* mutant strain. Thus, we conclude that *HEM6* expression does not require the HAP2/3/4/5 complex.

Uroporphyrinogen decarboxylases from many sources have been purified and this has enabled the study of the kinetics and co-factor requirements for this enzyme. Many experiments with sulphydryl-specific reagents have suggested that cysteine residues are important for the activity of this enzyme. In these studies, purified uroporphyrinogen decarboxylases from different sources were incubated with N-ethylmaleimide or *p*-chloromercuribenzoate and assayed for enzymatic activity. In all cases, the enzyme activity was inhibited by these compounds, and thus it was concluded that one or more cysteines was important for the catalytic activity of the enzyme (Tomio *et al.*, 1970; De Verneuil *et al.*, 1983; Elder *et al.*, 1983; Kawanishi *et al.*, 1983; Straka and Kushner, 1983; Koopmann and Battle, 1987). Uroporphyrinogen decarboxylase from yeast has also been shown to be inhibited by sulphydryl-specific reagents (Felix and Brouillet, 1990). These reagents presumably modify cysteine residues at the active site and prevent the thiol groups from catalyzing the conversion of uroporphyrinogen to coproporphyrinogen. However, these reagents may also indirectly cause enzyme

inhibition by binding to non-essential cysteine residues and changing the conformation of the active site. These studies did not address these possibilities. We examined the importance of cysteine residues in a direct fashion by site-directed mutagenesis of the most conserved cysteine amongst uroporphyrinogen decarboxylases from different species. Cysteine 52 of the yeast enzyme is conserved in the *E. coli*, rat, and human uroporphyrinogen decarboxylases. This cysteine was mutagenized to structurally similar amino acids, serine or alanine. In addition, cysteine 26, a non-conserved cysteine residue was changed to a serine residue. The effects of these mutations on uroporphyrinogen decarboxylase activity were tested by the ability of plasmids containing the mutations to complement a *hem6* mutant yeast strain 1B. Both low- and high-copy vectors with the different mutations complemented the heme auxotrophy of strain 1B (Di Flumeri *et al.*, 1997). These observations suggested that the conserved cysteine 52 was not essential for uroporphyrinogen decarboxylase activity. It remains possible that our assay was not sensitive enough to detect a partial inactivation of the enzyme which could still result in complementation of the *hem6* mutant strain. We tested the transformants for phenotypes associated with decreased uroporphyrinogen decarboxylase activity and could not detect any defects in growth on a non-fermentable carbon source, or accumulation of porphyrin intermediates.

These data suggest that the conserved cysteine 52 is dispensable for the activity of yeast uropophyrinogen decarboxylase. The effect of mutation of cysteine residues in the human uropophyrinogen decarboxylase has also been studied (Wyckoff *et al.*, 1996). The human enzyme contains six cysteine residues, none of which were involved in disulphide bonds. All six residues were candidates for mediating substrate binding or participating

in the active site. Each of the six cysteines was mutagenized to serine. The mutagenized proteins were expressed in *E. coli* and purified. All six mutant proteins were assayed for the ability to decarboxylate the substrate uroporphyrinogen to coproporphyrinogen. The results demonstrated that all mutant proteins, including the variant mutated in the highly conserved cysteine, retained 40% of wild type activity (Wyckoff *et al.*, 1996). Therefore, it was concluded that no single cysteine residue is essential for activity of the human enzyme. Assuming that both yeast and human enzymes share the same mechanism for catalysis, our *in vivo* studies of the yeast enzyme support these findings.

If the active form of yeast uroporphyrinogen decarboxylase were multimeric, it is possible that two subunits mutated at different sites could form a functional multimeric enzyme with at least one intact active site. Complementation by subunit mixing is unlikely to account for our results, since Felix and Brouillet (1990) showed that the yeast uroporphyrinogen decarboxylase is a monomer.

It has been proposed that the enzyme has two or more active sites, with each catalyzing sequential decarboxylation reactions (De Verneuil *et al.*, 1980). If this were true, the mutation of one cysteine residue at one of the active sites may be insufficient to inactivate the enzyme. Addressing this possibility would necessitate mutating multiple cysteines within the yeast enzyme. Both our analysis and that of Wyckoff *et al.*(1996) examined only the effect of mutating single cysteine residues.

The complexity of the reaction catalyzed by uroporphyrinogen decarboxylase has made it difficult to define structure-function relationships for the enzyme. In addition, the number of active sites and the mechanism of decarboxylation by the enzyme has not been determined. Structure studies of the enzyme by X-ray crystallography could possibly

reveal how this enzyme recognizes its substrate and catalyzes its conversion to coproporphyrinogen.

HEM13 encodes the sixth enzyme in the heme biosynthetic pathway in yeast, coproporphyrinogen oxidase, which catalyzes the oxygen-dependent conversion of coproporphyrinogen to protoporphyrinogen. *HEM13* is a member of a group of genes that are negatively regulated by oxygen and heme, and that include *ANB1*, *COX5b*, and *HMG2* (Lowry and Lieber, 1986; Hodge *et al.*, 1989; Thorsness *et al.*, 1989). The transcription of this gene has been shown to be repressed by heme through the action of the product of the *ROX1* gene (Keng, 1992). In the presence of heme, HAP1 activates transcription of *ROX1*, and the ROX1 protein represses *HEM13* transcription. In the absence of heme, *ROX1* is not transcribed, resulting in the induction of *HEM13* expression.

In an effort to understand the mechanism by which ROX1 negatively regulates *HEM13* expression, we undertook an analysis of both the properties of the ROX1 protein, as well as the regulatory region of *HEM13*. ROX1 protein was synthesized either by *in vitro* transcription and translation in wheat germ extracts or in *E. coli* as a fusion to the GST protein. These alternate methods of ROX1 synthesis allowed us to delineate the domains of ROX1 required for binding, repression, and oligomerization. Analysis of the *HEM13* promoter was carried out to determine the regulatory sequences required for repression by heme. These regulatory sequences were examined for their *in vivo* ability to repress *HEM13* transcription and their *in vitro* ability to bind ROX1. Finally, a preliminary analysis to determine the role of SSN6/TUP1 in ROX1-mediated repression of *HEM13* was carried out.

All hypoxic genes studied to date contain one or more copies of the consensus operator sequence 5'-YYYATTGTTCTC-3'. However, only the sequence found in the *ANB1* gene has been demonstrated to be a ROX1 binding site (Balasubramanian *et al.*, 1993). We were interested in determining the possible interaction of ROX1 with sequences homologous to the consensus operator sequence found within the *HEM13* regulatory region. ROX1 synthesized in wheat germ extracts was incubated with a 32 base pair fragment (RS32) containing a sequence from the *HEM13* promoter which showed homology to the consensus operator sequence derived from alignment of hypoxic genes. Full length *in vitro* translated ROX1 bound this fragment specifically as neither a 100-fold excess of a non-homologous DNA fragment nor an excess of poly(dI:dC) could compete for binding to ROX1. Thus, repression of *HEM13* by ROX1 is achieved through direct contact of ROX1 with sequences found upstream of the *HEM13* transcription start site. To delineate the region of ROX1 required for binding to RS32, we constructed a series of ROX1 deletion derivatives. Truncated derivatives of ROX1 synthesized *in vitro* in wheat germ extracts were tested for their ability to bind to RS32. The 100 amino-terminal amino acids of ROX1 were found to be essential and sufficient for DNA-binding. This region contains the HMG motif of ROX1, between amino acids 9-93. Full length ROX1, ROX1 1-286, and ROX1 1-175 were all capable of binding to DNA and gave rise to multiple protein-DNA complexes of identical mobilities. ROX1 1-100 protein was also capable of protein-DNA complex formation with RS32, which migrated as a single band. ROX1 1-58, which was missing 35 amino acids of the HMG domain was no longer capable of binding to DNA. In summary, these experiments suggested that

the region of ROX1 responsible for DNA-binding was located within the amino-terminal 100 amino acids of ROX1.

All ROX1 derivatives containing a minimum of 175 amino-terminal amino acids gave rise to multiple protein-DNA complexes which migrated with the same mobilities as the complexes formed with full-length ROX1 protein. This was somewhat surprising as one would have expected the mobilities of the complexes formed with the smaller derivatives, such as ROX1 1-175, to be faster than that of full length ROX1. It may be that ROX1 is very sensitive to proteases and the full length protein was easily degraded to a size similar to the smaller derivatives, thus giving rise to protein complexes that migrate with similar mobilities as those formed by the smaller proteins. Another possibility is that ROX1 is able to form oligomers *in vitro*. The multiple protein-DNA complexes detected could represent complexes containing different oligomerization states of ROX1. To assess these possibilities, we proceeded to obtain a more homogeneous source of ROX1 protein. ROX1 derivatives were expressed in *E.coli* as fusions to the GST protein. The GST moiety of the fusion proteins was removed by thrombin cleavage and the resultant ROX1 proteins were tested for their ability to bind to RS33, a DNA fragment containing the same hypoxic operator site from the *HEM13* promoter as fragment RS32. The multiple protein-DNA complexes observed with full length ROX1 and ROX1 1-176 resembled those formed with the equivalent *in vitro* translated proteins. The ROX1 1-101 protein formed a second slower migrating complex in addition to the faster migrating one seen with *in vitro* translated ROX1 1-100 protein. This difference in the number of complexes formed was probably due to the different concentrations of the *in vitro* translated and *E. coli* produced ROX1 proteins used in the binding assays. The

concentration of *E. coli* ROX1 1-101 protein used in the binding assays was higher than that of ROX1 1-100 produced in wheat germ extract. The higher concentration of the *E. coli* synthesized protein would favor the formation of oligomeric species seen in those binding assays. The inability of ROX1 176-368 to bind to DNA ruled out the possibility that a second independent DNA-binding domain existed in the carboxy-terminus of ROX1.

To determine if ROX1 could oligomerize, and to map a possible oligomerization domain within ROX1, we made use of both the *in vitro* translated ROX1 and GST-ROX1 fusion proteins. Our analysis indicates that ROX1 protein is capable of oligomerization and that the HMG domain at the amino terminus of ROX1 is required for formation of oligomers. The ROX1 protein contains a stretch of 16 glutamine residues adjacent to the HMG domain from amino acids 102 to 123. Such stretches of glutamine residues are a feature common to eucaryotic transcriptional activator proteins and are thought to be required for protein-protein interactions (Courey *et al.*, 1989). In the protein interaction assay, the ROX1 1-100 *in vitro* translated protein was capable of interacting with GST-ROX1 FL protein, indicating that the polyglutamine tract in ROX1 is not required for ROX1 oligomerization. These experiments do not rule out a possible interaction of this polyglutamine tract with other as yet unidentified proteins containing the same motif. The interaction assay was carried out in the absence of DNA, and would predict that the oligomerization of ROX1 does not require the protein to be bound to DNA. The *in vivo* ability of ROX1 to oligomerize may be tested using the yeast two-hybrid assay for protein-protein interaction. The physiological significance of such an interaction remains to be determined.

In conclusion, we have demonstrated in chapter 5 of the thesis that ROX1 represses transcription of *HEM13* by binding to at least one of the hypoxic operator sites upstream of the transcription start site of this gene. In addition, the ROX1 domains required for DNA-binding and oligomerization were mapped and reside within the amino-terminal 100 amino acids of the protein.

We wished to carry out a more extensive analysis of the sites required for *HEM13* repression within the regulatory sequences of the gene. We identified five possible operator sites with homology to the 5'-YYYATTGTTCTC-3' consensus operator sequence. We examined both binding of ROX1 to all five possible operator sites, and the contribution of each site to repression *in vivo*. ROX1 was capable of binding to all five sites with different affinities. Binding to site 3, the site that most closely resembled the consensus sequence, was strongest. Two other sites, 1 and 5, contain the core consensus 5'-ATTGTT-3' sequence but vary from the consensus sequence in the flanking regions. These two sites were bound by ROX1 with a decreased affinity compared to site 3. Site 4, which has the core sequence 5'-TTTGTT-3', had a much reduced affinity for ROX1. Surprisingly, site 2, which has the core sequence 5'-CTGGTT-3', was bound by ROX1 protein with low affinity, in spite of two substitutions in the highly conserved core motif 5'-ATTGTT-3'.

We determined whether any of these identified consensus sites mediates the effects of oxygen and heme on expression of *HEM13 in vivo*. Each consensus site was deleted alone, as well as in various combinations with one another, and the effects of these deletions on *HEM13* expression were determined. Sites 1, 3, 4, and 5 were functional and each site made a similar contribution to repression of *HEM13* expression

of any one of these sites resulted in a 2.5- to 3.5-fold increase in aerobic expression of a *HEM13-lacZ* fusion. The contribution of each individual site towards repression cannot be correlated with the affinity of ROX1 to that site. Deletion of site 3, which showed the highest affinity for ROX1, did not result in a larger derepression of *HEM13* expression than deletion of site 4, which had a reduced affinity for ROX1.

Combined deletions of the operator sites indicated that sites 1, 3, and 4 act in an additive manner to repress *HEM13* expression. A combined deletion of site 4 with site 1 or site 3 resulted in 24 to 26 units of expression, which represents a further increase in *HEM13* expression of approximately 2.5-fold over that observed with deletion of site 1 or site 3 alone. Furthermore, a deletion of site 4 in conjunction with a double deletion of sites 1 and 3 resulted in 62.3 units of expression, an overall increase of 16.8-fold over that observed with the wild-type promoter and a 2.5-fold increase over that observed with a construct with two of these sites deleted. These data suggest that the three sites function in an additive manner in bringing about repression of *HEM13*.

Site 5, with the core sequence 5'-ATTGTT-3', showed a relatively high affinity for the ROX1 protein and its deletion resulted in a 2.7-fold increase in *HEM13* expression. Interestingly, deletion of site 5 in conjunction with sites 1 and 3 or sites 3 and 4 did not have any effect on expression above that observed upon deletion of sites 1 and 3 or sites 3 and 4 by themselves. Therefore, it would appear that when *HEM13* is fully repressed (when sites 1, 3, and 4 are present and bound by ROX1), site 5 contributes to repression. However, when sites 1 and 3, 3 and 4, or 1, 3, and 4 are deleted and *HEM13* is largely derepressed, site 5 makes no contribution to repression. These results suggest that occupancy of sites 1, 3, and 4 by ROX1 may bring the ROX1-bound site 5

located some 500 bp upstream of the transcription start site, in closer proximity to the basal transcription machinery and enable site 5 to contribute to repression of *HEM13* expression. This suggestion is supported by the observation that ROX1 is able to bend the DNA helix by 90° (Deckert *et al.*, 1995a). It may be that when sites 1, 3, and 4 are occupied, ROX1 bound at these sites can bend the helix and bring site 5 in the correct juxtaposition required for repression.

Site 2, which had a low affinity for ROX1, but a higher affinity than site 4, did not appear to mediate the effects of oxygen and heme on expression of *HEM13 in vivo*. Deletion of site 2 by itself had no effect on expression of *HEM13* and deletion of site 2 in conjunction with other functional sites did not result in additional effects on *HEM13* expression. Thus, ROX1 protein binding *in vitro* does not translate into a function in repression of *HEM13* expression *in vivo*.

Experiments utilizing minor groove binding drugs suggest that ROX1, like many of the HMG proteins, binds DNA through contacts with the minor groove. The substitution of specific A-T base pairs within the core 5'-ATTGTT-3' binding site with I-C base pairs provided a more direct method for assessing minor groove contacts and indicated the specific A-T pairs involved. ROX1 was still capable of binding to oligonucleotides when adenine residues on either position 1, positions 2 and 3, or positions 5 and 6 were substituted by inosine residues, but not when they were substituted with guanine residues. However, binding was not as strong as binding to the unsubstituted oligonucleotide, suggesting that determinants in the major groove may also make contributions to the binding affinity of ROX1.

In vitro slection of ROX1 binding sites showed that A-T rich sequences with a consensus 5'-(A/T)TT(T/G)TT-3' were the favoured binding sites. The sequences 5'-ATTTTT-3' and TTTGTT appeared with highest frequency. The sequence 5'-TTTGTT-3' does appear in the *HEM13* promoter (site 4), but is bound with reduced affinity by ROX1 compared to the sequence 5'-ATTGTT-3'. Interestingly, the sequence 5'-ATTGTT-3' was not selected, although this is the consensus sequence derived from comparison of the regulatory regions of ROX1-responsive genes.

The function of SSN6 and TUP1 in repression of *HEM13* expression was assessed by examining the effects of deletion of either of these genes on *HEM13-lacZ* expression. Both *ssn6* and *tup1* deleted strains showed a marked increase in *HEM13-lacZ* expression under aerobic conditions (7-8-fold). Thus, both factors are required for repression of *HEM13* and presumably act through the same pathway as ROX1, although we have not examined *HEM13* expression in Δ*tup1* Δ*ssn6*, Δ*tup1* Δ*rox1*, or Δ*ssn6* Δ*rox1* strains.

The domain of ROX1 required for *HEM13* repression was delineated by determining the level of β-galactosidase expression from a *HEM13-lacZ* fusion in a Δ*rox1* strain transformed with ROX1 deletion derivatives. Deletion of 92 amino acids from the carboxy-terminus of ROX1 resulted in an 11-fold derepression of *HEM13-lacZ* expression. We conclude that repression by ROX1 requires these 92 amino acids. It is tempting to speculate that the carboxy-terminal 92 amino acids of ROX1 are required to recruit the SSN6-TUP1 complex to the *HEM13* promoter. Recruitment of the TUP1/SSN6 complex by pathway specific repressors to promoters targeted for repression has been demonstrated for MATα2 protein. The WD repeats of TUP1 were shown to be

capable of interacting with MATα2 protein and causing repression of a-specific genes (Komachi *et al.*, 1994). Other pathway specific repressor proteins recruit the SSN6/TUP1 complex presumably through interactions with the TPR motifs of SSN6. Repression of *ANB1* expression by ROX1 required TPR motifs 4-7 of SSN6 (Tzamarias and Struhl, 1995). However, an attempt to define a physical interaction of SSN6 and ROX1 was inconclusive. Recently, indirect evidence for ROX1-SSN6 interaction has been obtained. Addition of SSN6 to ROX1 protein-DNA complexes stabilized the complexes (Zitomer *et al.*, 1997a). The implication is that ROX1 activity decays and the decay is prevented by SSN6 protein.

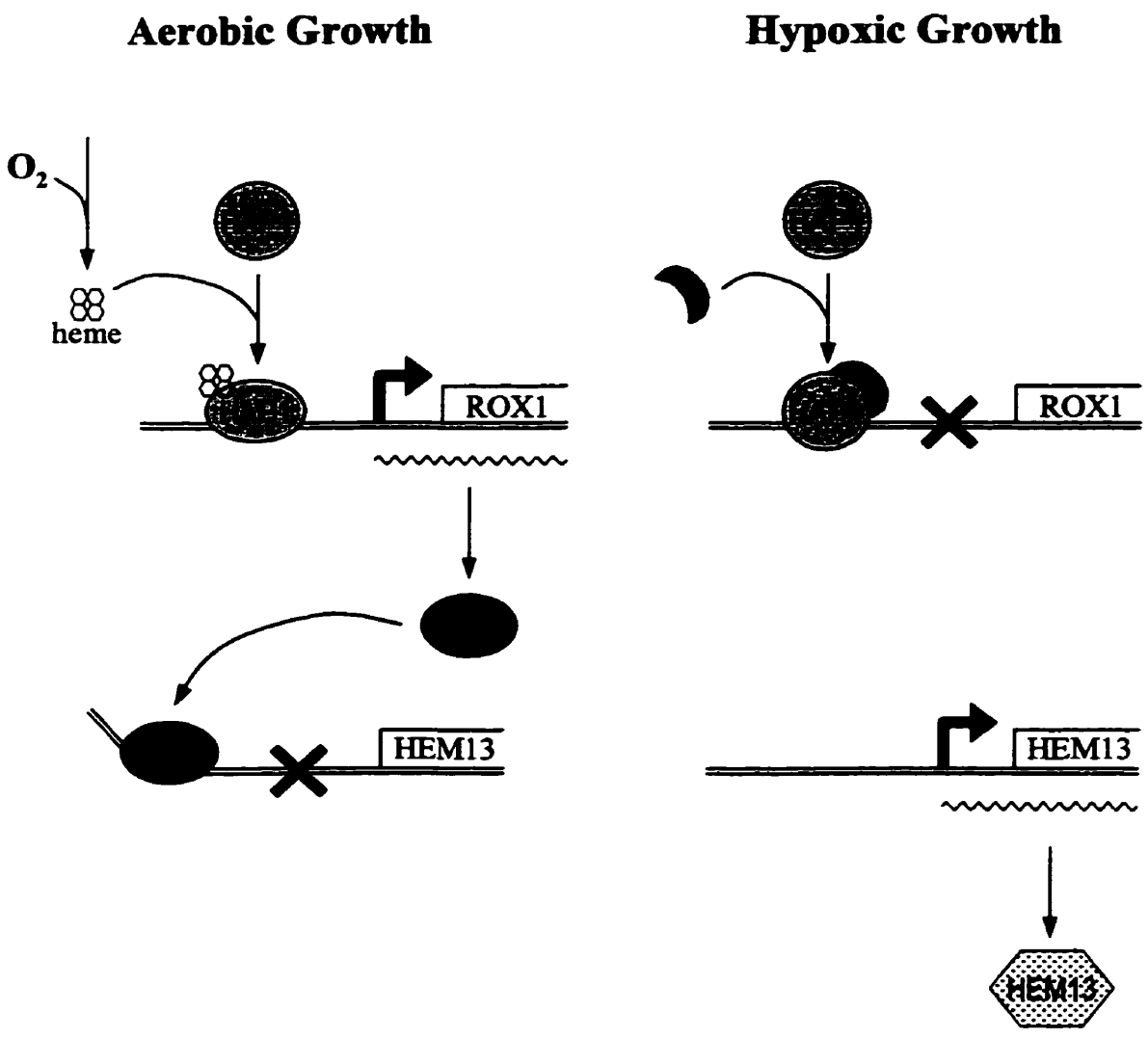
To summarize what is known about *HEM13* repression by heme and ROX1, we examine data that has been obtained in this study as well as that of others. In the presence of oxygen, heme is synthesized and binds to the transcriptional activator HAP1. When heme is bound, HAP1 activates expression of *ROX1*. The ROX1 repressor then binds to its hypoxic operator sites upstream of the *HEM13* promoter through its amino-terminal HMG domain (Deckert *et al.*, 1995a; Di Flumeri *et al.*, 1996). Five hypoxic operator sites can be observed; all of them are bound by ROX1 but only four of them are functional *in vivo* (Di Flumeri *et al.*, Manuscript in preparation). Binding of ROX1 to the DNA occurs primarily through the minor groove of the DNA helix, a characteristic which may distort the helix and facilitate the 90° bend imposed by ROX1 (Deckert *et al.*, 1995a; Di Flumeri *et al.*, 1996). Bending by ROX1 bound to sites 1, 3, and 4 may put site 5 in a context which enables it to contribute to repression. Sites 1, 3, and 4 act additively to repress *HEM13* transcription.

ROX1 protein requires TUP1 and SSN6 for repression (Zhang *et al.*, 1991; Balasubramanian *et al.*, 1993). It is postulated that the bound ROX1 molecules recruit the SSN6-TUP1 complex through interaction of the carboxy-terminus of ROX1 with TPR motifs 4-7 of SSN6 (Tzamarias and Struhl, 1995). Once recruited to the promoter, the TUP1 protein induces the formation of a repressive chromatin structure or interferes with the basal transcription apparatus, leading to the shut down in transcription from the *HEM13* promoter. Possible targets for TUP1 include the SRB10 and SRB11 proteins, which form a cyclin-kinase complex that is part of the yeast RNA polymerase II holoenzyme (Wahi and Johnson, 1995). To ensure that levels of ROX1 in the cell do not become exceedingly high, ROX1 also simultaneously binds to its own promoter, and through the mechanism just described for *HEM13* represses its own synthesis (Deckert *et al.*, 1995b).

In the absence of oxygen, heme levels decrease and HAP1 is associated with other factors which change its specificity (Fytlovich *et al.*, 1993), resulting in the shut down in transcription of *ROX1*. Consequently, the levels of ROX1 in the cell are reduced, and the hypoxic genes are induced. The repression and induction schemes for *HEM13* expression are simplified in figure 24.

The study of the mechanism of repression of hypoxic gene expression has revealed a complexity far beyond what was expected. Many protein factors are involved, some of which have been defined, others are still elusive. The data reported in this thesis and by others, has led to a better understanding of how genes are turned off in response to certain stimuli.

Figure 24. Model for heme-dependent repression of *HEM13* expression in the presence of oxygen. Cells grown in the presence of oxygen synthesize heme, which modulates the activity of the HAP1 protein. HAP1 activates transcription of *ROX1*, resulting in the synthesis of ROX1, which subsequently binds to the hypoxic operator sites within the *HEM13* regulatory region and represses its transcription. Alternatively, hypoxia results in diminished levels of heme which causes HAP1 to bind an unknown cellular factor. This heteromeric complex represses *ROX1* transcription, leading to *HEM13* expression and synthesis of coproporphyrinogen oxidase.


Aerobic Growth
Hypoxic Growth
O_2
heme
ROX1
ROX1
HEM13
HEM13
HEM13

LITERATURE CITED

Amillet, J-M., and R. Labbe-Bois. 1995. Isolation of the gene *HEM4* encoding uroporphyrinogen III synthase in *Saccharomyces cerevisiae*. *Yeast* **11:** 419-424.

Amillet, J-M., N. Buisson, and R. Labbe-Bois. 1995. Positive and Negative elements involved in the differential regulation by heme and oxygen of the *HEM13* gene (coproporphyrinogen oxidase) in *Saccharomyces cerevisiae*. *Curr. Genet.* **28:** 503-511.

Amillet, J-M., N. Buisson, and R. Labbe-Bois. 1996. Characterization of an upstream activation sequence and two Rox1p-responsive sites controlling the induction of the yeast *HEM13* gene by oxygen and heme deficiency. *J. Biol. Chem.* **271:** 24425-24432.

Arrese, M., E. Carvajal, S. Robison, A. Sambunaris, A. Panek, and J. Mattoon. 1983. Cloning of the δ-aminolevulinic acid synthase structural gene in yeast. *Curr. Genet.* **7:** 175-183.

Arrese, M.R., D. Vojensky, and J.R. Mattoon. 1982. Characterization of a cytochrome oxidase-deficient yeast mutant which accumulates porphyrins. *Biochem. Biophys. Res. Comm.* **107:** 848-855.

Balasubramanian, B., C.V. Lowry, and R.S. Zitomer. 1993. The ROX1 repressor of the *Saccharomyces cerevisiae* hypoxic genes is a specific DNA-binding protein with a high-mobility-group motif. *Mol. Cell. Biol.* **13:** 6071-6078.

Bard, M., and T.D. Ingolia. 1984. Plamid-mediated complementation of a δ-aminolevulinic acid-requiring *Saccharomyces cerevisiae*. *Gene* **28:** 195-199.

Battle, A. M. del C., R.E. de Salamanca, S. Chinaro, S.G. Afonso, and A.M. Stella. 1986. Photodynamic inactivation of red cell uroporphyrinogen decarboxylase by porphyrins. *Int. J. Biochem.* **18:** 143-147.

Blattner, F.R., V. Burland, G. III. Plunkett, H.J. Sofia, and D.L. Daniels. 1993. Analysis of the *Escherichia coli genome*. IV. DNA sequence of the region from 89.2 to 92.8 minutes. *Nucleic Acids Res.* **21:** 5408-5417.

Bourot, S. and F. Karst. 1995. Isolation and characterization of the *Saccharomyces cerevisiae* SUT1 gene involved in sterol uptake. *Gene* **165:** 97-102.

Bowman, S.B., A. Zaman, L.P. Collinson, A.J.P. Brown, and I.W. Dawes. 1992. Positive regulation of the *LPD1* gene of *Saccharomyces cerevisiae* by the HAP2/HAP3/HAP4 activation system. *Mol. Gen. Genet.* **231:** 296-303.

Braun, B.R. and A.D. Johnson. 1997. Control of filament formation in *Candida albicans* by the transcriptional repressor TUP1. *Science* **277:** 105-109.

Camadro, J.M. and P. Labbe. 1988. Purification and properties of ferrochelatase from the yeast *Saccharomyces cerevisiae.* Evidence for a precursor form of the protein. *J. Biol. Chem.* **263:** 11675-11682.

Camadro, J.M., and P. Labbe. 1996. Cloning and characterization of the yeast *HEM14* gene coding for protoporphyrinogen oxidase, the molecular target of diphenyl ether-type herbicides. *J. Biol. Chem.* **271:** 9120-9128.

Camadro, J.M., D. Urban-Grimal, and P. Labbe. 1982. A new assay for protoporphyrinogen oxidase - evidence for a total deficiency in the activity in a heme-less mutant of *Saccharomyces cerevisiae. Biochem. Biophys. Res. Comm.* **106:** 724-730.

Camadro, J.M., D. Urban-Grimal, and P. Labbe. 1986. Purification and properties of coproporphyrinogen oxidase from the yeast *Saccharomyces cerevisiae. Eur. J. Biochem.* **156:** 579-587.

Camadro, J.M., F. Thome, N. Brouillet, and P. Labbe. 1994. Purification and properties of protoporphyrinogen oxidase from the yeast *Saccharomyces cerevisiae.* Mitochondrial location and evidence for a precursor form of the protein. *J. Biol. Chem.* **269:** 32085-32091.

Carlson, M., and D. Botstein. 1982. Two differentially regulated mRNAs with different 5' ends encode secreted and intracellular forms of yeast invertase. *Cell* **28:** 145-154.

Carlson, M., B.C. Osmond, L. Neigeborn, and D. Botstein. 1984. A suppressor of *snf1* mutations causes constitutive high-level invertase synthesis in yeast. *Genetics* **107:** 19-32.

Cerdan, M.E., and R.S. Zitomer. 1988. Oxygen-dependent upstream activation sites of *Saccharomyces cerevisiae* cytochrome *c* genes are related forms of the same

sequence. *Mol. Cell. Biol.* **8:** 2275-2279.

Clavilier, L., G. Pere-Aubert, and P.P. Slonimski. 1969. Mise en evidence de plusieurs loci independants impliques dans la synthese de l'iso-2-cytochrome *c* chez la levure. *Mol. Gen. Genet.* **104:** 195-218.

Cooper, J.P., S.Y. Roth, and R.T. Simpson. 1994. The global transcriptional regulators, SSN6 and TUP1, play distinct roles in the establishment of a repressive chromatin structure. *Genes Dev.* **8:** 1400-1410.

Courey, A.J., D.A. Holtzman, S.P. Jackson, and R. Tjian. 1989. Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. *Cell* **59:** 827-836.

Creusot, F., J. Verdiere, M. Gaisne, and P.P. Slonimski. 1988. CYP1 (HAP1) Regulator of oxygen-dependent gene expression in yeast. I. overall organization of the protein sequence displays several novel structural domains. *J. Mol. Biol.* **204:** 263-276.

De Verneuil, H., B. Grandchamp, and Y. Nordmann. 1980. Some kinetic properties of human red cell uroporphyrinogen decarboxylase. *Biochim. Biophys. Acta* **611:** 174-186.

De Verneuil, H., S. Sassa, and A. Kappas. 1983. Purification and properties of uroporphyrinogen decarboxylase from human erythrocytes. A single enzyme catalyzing the four sequential decarboxylations of uroporphyrinogens I and III. *J. Biol. Chem.* **258:** 2454-2460.

Deckert, J., A.M. Rodriguez Torres, J.T. Simon, and R.S. Zitomer. 1995a. Mutational Analysis of ROX1, a DNA-Bending repressor of hypoxic genes in *Saccharomyces cerevisiae. Mol. Cell. Biol.* **15:** 6109-6117.

Deckert, J., R. Perini, B. Balasubramanian, and R.S. Zitomer. 1995b. Multiple elements and auto-repression regulate ROX1, a repressor of hypoxic genes in *Saccharomyces cerevisiae. Genetics* **139:** 1149-1158.

Delfau-Larue, M.H., P. Martasek, and B. Grandchamp. 1994. Coproporphyrinogen oxidase: gene organization and description leading to exon 6 skipping. *Human. Mol. Genet.* **3:** 1325-1330.

Denny, P., S. Swift, F. Connor, and A. Ashworth. 1992. An *SRY*-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein. *EMBO J.* **11:** 3705-3712.

Di Flumeri , C., P. Liston, N.H. Acheson, and T. Keng. 1996. The HMG domain of the ROX1 protein mediates repression of *HEM13* through overlapping DNA binding and oligomerization functions. *Nucleic Acids Res.* **24:** 808-815.

Di Flumeri, C., A. Staffa, S. Ushinsky, N.H. Acheson, and T. Keng. In preparation. The *Saccharomyces cerevisiae* ROX1 protein represses transcription of *HEM13* by binding to multiple upstream sites with differential affinity.

Di Flumeri, C., N.H. Acheson, and T. Keng. 1997. A conserved cysteine residue in yeast uroporphyrinogen decarboxylase is not essential for enzymatic activity. *Can. J. Microbiol.* **43:** 792-795.

Di Flumeri, C., R. Larocque, and T. Keng. 1993. Molecular analysis of *HEM6* (*HEM12*) in *Saccharomyces cerevisiae*, the gene for uroporphyrinogen decarboxylase. *Yeast* **9:** 613-623.

Diffley, J.F., and B. Stillman. 1992. DNA binding properties of an HMG1-related protein from yeast mitochondria. *J. Biol. Chem.* **267:** 3368-3374.

Dooijes, D., M. van de Wetering, L. Knippels, and H. Clevers. 1993. The *Schizosaccharomyces pombe* mating-type gene *mat*-Mc encodes a sequence-specific DNA-binding high mobility group box protein. *J. Biol. Chem.* **268:** 24813-24817.

Dorsman, J.C., and L.A. Grivell. 1990. Expression of the gene encoding subunit II of yeast QH2: cytochrome *c* oxidoreductase is regulated by multiple factors. *Curr. Genet.* **17:** 459-464.

Dumont, M.E., J.F. Ernst, D.M. Hampsey, and F. Sherman. 1987. Identification and sequence of the gene encoding cytochrome *c* heme lyase in the yeast *Saccharomyces cerevisiae*. *EMBO J.* **6:** 235-241.

Elder, G.H., J.A. Tovey, and D.M. Sheppard. 1983. Immunochemical evidence for a single protein with decarboxylase activity in human erythrocytes and liver. *Biochem. J.* **215:** 45-55.

Felix, F., and N. Brouillet. 1990. Purification and properties of uroporphyrinogen decarboxylase from *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **188:** 393-403.

Ferrari, S., V.R. Harley, A. Pontiggia, P.N. Goodfellow, R. Lovell-Badge, and M.E. Bianchi. 1992. SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J.* **11:** 4497-4506.

Ferreira, G.C., and J. Gong. 1995. 5-Aminolevulinate synthase and the first step of heme biosynthesis. *J. Bioenerg. Biomemb.* **2:** 151-159.

Forsburg, S.L., and L. Guarente. 1988. Mutational analysis of upstream activation sequence 2 of the *CYC1* gene of *Saccharomyces cerevisiae*: a *HAP2-HAP3* responsive site. *Mol. Cell. Biol.* **8:** 647-654.

Forsburg, S.L., and L. Guarente. 1989a. Identification and characterization of HAP4: a third component of the CCAAT-bound HAP2/HAP3 heteromer. *Genes Dev.* **3:** 1166-1178.

Forsburg, S.L., and L. Guarente. 1989b. Communication between the mitochondria and the nucleus in regulation of cytochrome genes in the yeast *Saccharomyces cerevisiae*. *Annu. Rev. Cell Biol.* **5:** 153-180.

Fujita, A., S. Matsumoto, S. Kuhara, Y. Misumi, and H. Kobayashi. 1990. Cloning of the yeast *SFL2* gene: its disruption results in pleiotropic phenotypes characteristic for *tup1* mutants. *Gene* **89:** 93-99.

Fytlovich, S., M. Gervais, C. Agrimonti, and B. Guiard. 1993. Evidence for an interaction between the CYP1 (HAP1) activator and a cellular factor during heme-dependent transcriptional regulation in the yeast *Saccharomyces cerevisiae*. *EMBO J.* **12:** 1209-1218.

Garey, J. R., R. Labbe-Bois, A. Chelstowska, J. Rytka, L. Harrison, J. Kushner, and P. Labbe. 1992. Uroporphyrinogen decarboxylase in *Saccharomyces cerevisiae-HEM12* gene sequence and evidence for two conserved glycines essential for enzymatic activity. *Eur. J. Biochem.* **205:** 1011-1016.

Giese, K., J. Cox, and R. Grosschedl. 1992. The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69:** 185-195.

Gollub, E. G., K.P. Liu, J. Dayan, M. Adlersberg, and D.B. Sprinson. 1977. Yeast mutants deficient in heme biosynthesis and a heme mutant additionally blocked in cyclization of 2, 3-oxidosqualene. *J. Biol. Chem.* **252:** 2846-2854.

Goodwin, G.H., C. Sanders, and E.W. Johns. 1973. A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. *Eur. J. Biochem.* **38:** 14-19.

Grandchamp, B., H. De Verneuil, C. Beaumont, S. Chretien, O. Walter, and Y. Nordmann. 1987. Tissue-specific expression of porphobilinogen deaminase. Two isoenzymes from a single gene. *Eur. J. Biochem.* **162:** 105-110.

Grandchamp, B., J. Lamoril, and H. Puy. 1995. Molecular abnormalities of coproporphyrinogen oxidase in patients with hereditary coproporphyria. *J. Bioenerg. Biomemb.* **27:** 215-219.

Grimal, D., and R Labbe-Bois. 1980. An enrichment method for heme-less mutants of *Saccharomyces cerevisiae* based on photodynamic properties of Zn-protoporphyrin. *Mol. Gen. Genet.* **178:** 713-716.

Grosschedl, R., K. Giese, and J. Pagel. 1994. HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet.* **10:** 94-100.

Guarente, L. 1983. Yeast promoters and *lacZ* fusions designed to study expression of cloned genes in yeast. *Methods Enzymol.* **101:** 181-191.

Guarente, L. and O. Birmingham-McDonogh. 1992. Conservation and evolution of transcriptional mechanisms in eukaryotes. *Trends Genet.* **8:** 27-32.

Guarente, L., and T. Mason. 1983. Heme regulates transcription of the *CYC1* gene of *Saccharomyces cerevisiae* via an upstream activation site. *Cell* **32:** 1279-1286.

Guarente, L., B. Lalonde, P. Gifford, and E. Alani. 1984. Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *Saccharomyces cerevisiae*. *Cell* **36:** 503-511.

Gubbay, J., J. Collignon, P. Koopman, B. Capel, A. Economou, A. Munsterberg, N. Vivian, P. Goodfellow, and R. Lovell-Badge. 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically-expressed genes. *Nature* **346:** 245-250.

Hahn, S., and L. Guarente. 1988. Yeast HAP2 and HAP3: transcriptional activators in a heteromeric complex. *Science* **240:** 317-321.

Hahn, S., J. Pinkham, R. Wei, R. Miller, and L. Guarente. 1988. The *HAP3* regulatory locus of *Saccharomyces cerevisiae* encodes divergent overlapping transcripts. *Mol. Cell. Biol.* **8:** 655-663.

Hamilton, B., R. Hofbauer, and H. Ruis. 1982. Translational control of catalase synthesis by hemin in the yeast *Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA* **79:** 7609-7613.

Hanahan, D. 1983. Studies on the transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166:** 557-580.

Herschbach, B.M., M.B. Arnaud, and A.D. Johnson. 1994. Transcriptional repression directed by the yeast α2 protein *in vitro*. *Nature* **370:** 309-311.

Hill, K.L., and S. Merchant. 1995. Coordinate expression of coproporphyrinogen oxidase and cytochrome *c*6 in the green alga *Chlamydomonas reinhardtii* in response to changes in copper availability. *EMBO J.* **14:** 857-865.

Hinnen, A., J.B. Hicks, and G.R. Fink. 1978. Transformation of yeast. *Proc. Natl. Acad. Sci. USA* **75:** 1929-1933.

Hoare, D.S., and H. Heath. 1959. The biosynthesis of porphyrins from porphobilinogen by *Rhodopseudomonas spheroides 2.* The partial purification and some properties of porphobilinogen deaminase and uroporphyrinogen decarboxylase. *Biochem. J.* **73:** 679-690.

Hodge, M.R., G. Kim, K. Singh, and M.G. Cumsky. 1989. Inverse regulation of the yeast *COX5* genes by oxygen and heme. *Mol. Cell. Biol.* **9:** 1958-1964.

Hodge, M.R., K. Singh, and M.G. Cumsky. 1990. Upstream activation and repression elements control transcription of the yeast *COX5b* gene. *Mol. Cell. Biol.* **10:** 5510-5520.

Hortner, H., G. Ammerer, E. Hartter, B. Hamilton, J. Rytka, T. Bilinski, and H. Ruis. 1982. Regulation of synthesis of catalases and iso-1-cytochrome *c* in *Saccharomyces cerevisiae* by glucose, oxygen, and heme. *Eur. J. Biochem.* **128:** 179-184.

Ito, H., Y. Jukuda, K. Murata, and A. Kimura. 1983. Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153:** 163-168.

Jahn, D., E. Verkamp, and D. Soll. 1992. Glutamyl-transfer RNA: a precursor of heme and chlorophyll biosynthesis. *Trends in Biochem. Sci.* **17:** 216-218.

Jantzen, H.M., A. Admon, S.P. Bell, and R. Tjian. 1990. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature* **344:** 830-836.

Juknat, A.A., A. Seubert, S. Seubert, and H. Ippen. 1989. Studies on uroporphyrinogen decarboxylase of etiolated *Euglena gracilis Z. Eur. J. Biochem.* **179:** 423-428.

Kawanishi, S., Y. Seki, and S. Sano. 1983. Uroporphyrinogen decarboxylase. Purifcation, properties, and inhibition by polychlorinated biphenyl isomers. *J. Biol. Chem.* **258:** 4285-4292.

Keleher, C.A., M.J. Redd, J. Schultz, M. Carlson, and A.D. Johnson. 1992. Ssn6-Tup1 is a general repressor of transcription in yeast. *Cell* **68:** 709-719.

Kelly, M., J. Burke, M. Smith, A. Klar, and D. Beach. 1988. Four mating-type genes control sexual differentiation in the fission yeast. *EMBO J.* **7:** 1537-1547.

Keng, T. 1992. HAP1 and ROX1 form a regulatory pathway in the repression of *HEM13* transcription *in Saccharomyces cerevisiae. Mol. Cell. Biol.* **12:** 2616-2623.

Keng, T., and L. Guarente. 1987. Multiple regulatory systems result in constitutive expression of the yeast *HEM1* gene. *Proc. Natl. Acad. Sci. USA* **84:** 9113-9117.

Keng, T., C. Richard, and R. Larocque. 1992. Structure and regulation of yeast *HEM3*, the gene for porphobilinogen deaminase. *Molec. Gen. Genet.* **234:** 233-243.

Keng, T., C. Richard, and S. Ushinsky. Unpublished observations.

Keng, T., E. Alani, and L. Guarente. 1986. The nine amino-terminal residues of δ-aminolevulinic acid synthase direct δ-galactosidase into the mitochondrial matrix. *Mol. Cell. Biol.* **6:** 355-364.

Kiel, J.A.K.W., J.M. Boels, G. Beldman, and G. Venema. 1990. Nucleotide sequence of the *Synechococcus* sp. PCC7942 branching enzyme gene (*glgB*): expression in *Bacillus subtilis*. *Gene* **89:** 77-84.

Kim, K.S., K. Pfeifer, L. Powell, and L. Guarente. 1990. Internal deletions in the yeast transcriptional activator HAP1 have opposite effects at two sequence elements. *Proc. Natl. Acad. Sci. USA* **87:** 4524-4528.

Kohno, H., T. Furukawa, T. Yoshinaga, R. Tokunaga, and S. Taketani. 1993. Coproporphyrinogen oxidase: purification, molecular cloning and induction of mRNA during erythroid differention. *J. Biol. Chem.* **268:** 21359-21363.

Komachi, K., M.J. Redd, and A.D. Johnson. 1994. The WD repeats of Tup1 interact with the homeo domain protein Alpha2. *Genes Dev.* **8:** 2857-2867.

Koopmann, G. E., and A. M. Del C. Battle. 1987. Biosynthesis of porphyrins in *Rhodopseudomonas palustris-VI* The effect of metals, thiols, and other reagents on the activity of uroporphyrinogen decarboxylase. *Int. J. Biochem.* **19:** 373-377.

Kurlandzka, A., and J Rytka. 1985. Mutants of *Saccharomyces cerevisiae* partially defective in the last three steps of the haem biosynthetic pathway: isolation and genetical characterization. *J. Gen. Microbiol.* **131:** 2909-2918.

Kurlandzka, A., T. Zoladek, J. Rytka, R. Labbe-Bois, and D. Urban-Grimal. 1988. The effects *in vivo* of mutationally modified uroporphyrinogen decarboxylase in different *hem12* mutants of baker's yeast (*Saccharomyces cerevisiae*). *Biochem. J.* **253:** 109-116.

Labbe-Bois, R. 1990. The ferrochelatase from *Saccharomyces cerevisiae*. Sequence, disruption, and expression of its structural gene *HEM15*. *J. Biol. Chem.* **265:** 7278-7283.

Labbe-Bois, R., and P. Labbe. 1990. Tetrapyrrole and heme biosynthesis in the yeast *Saccharomyces cerevisiae*. In: *Biosynthesis of Heme and Chlorophylls*. H.A. Dailey, ed., McGraw-Hill Publishing, New York, pp. 235-285.

Lalonde, B., B. Arcangioli, and L. Guarente. 1986. A single *Saccharomyces cerevisiae* upstream activation site (UAS1) has two distinct regions essential for its activity. *Mol. Cell. Biol.* **6:** 4690-4696.

Laudet, V., D. Stehelin, and H. Clevers. 1993. Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Res.* **21:** 2493-2501.

Lodi, T., and B. Guiard. 1991. Complex transcriptional regulation of the *Saccharomyces cerevisiae CYB2* gene encoding cytochrome *b2*: CYP1 (HAP1) Activator binds to the *CYB2* upstream activation site UAS1-B2. *Mol. Cell. Biol.* **11:** 3762-3772.

Lowry, C.V., and R. Lieber. 1986. Negative regulation of the *Saccharomyces cerevisiae ANB1* gene by heme, as mediated by the *ROX1* gene product. *Mol. Cell. Biol.* **6:** 4145-4148.

Lowry, C.V., and R.S. Zitomer. 1984. Oxygen regulation of anaerobic and aerobic genes mediated by a common factor in yeast. *Proc. Natl. Acad. Sci. USA* **81:** 6129-6133.

Lowry, C.V., and R.S. Zitomer. 1988. *ROX1* encodes a heme-induced repression factor regulating *ANB1* and *CYC7* of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **8:** 4651-4658.

Lowry, C.V., M.E. Cerdan, and R.S. Zitomer. 1990. A hypoxic consensus operator and a constitutive activation region regulate the *ANB1* gene of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **10:** 5921-5926.

MacKay, V.L. 1983. Cloning of yeast *STE* genes in 2μm vectors. *Methods Enzymol.* **101:** 325-343.

Madsen, O., L. Sandal, N.N. Sandal, and K.A. Marcker. 1993. A soybean coproporphyrinogen oxidase gene is highly expressed in root nodules. *Plant Mol. Biol.* **24:** 35-43.

Marck, C. 1988. 'DNA Strider'; a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Res.* **16:** 1829-1836.

Marczynski, G.T., and L. Shapiro. 1992. Cell-cycle control of a cloned chromosomal origin of replication from *Caulobacter crescentus*. *J. Mol. Biol.* **226:** 959-977.

Martasek, P., J.M. Camadro, M.H. Delfau-Larue, J.B. Dumas, J.J. Montagne, H. De Verneuil, P. Labbe, and B. Grandchamp. 1994. Molecular cloning,

sequencing, and functional expression of a cDNA encoding human coproporphyrinogen oxidase. *Proc. Natl. Acad. Sci. USA* **91:** 3024-3028.

Matner, R.R. and F. Sherman. 1982. Differential accumulation of two apo-iso-cytochromes *c* in processing mutants of yeast. *J. Biol. Chem.* **257:** 9811-9821.

Mattoon, J.R., W.E. Lancashire, H.K. Sanders, E. Carvajal, D.R. Malamud, R.C. Braz, and A.D. Panek. 1979. Oxygen and catabolite regulation of hemoprotein biosynthesis in yeast. In: *Biochemical and Clinical Aspects of Oxygen.* W.J. Caughey, ed., Academic Press, NY, pp. 421-435.

May, B.K, S.C. Dogra, T.J. Sadlon, C.R. Bhasker, T.C. Cox, and S.S. Bottomley. 1995. Molecular regulation of heme biosynthesis in higher vertebrates. *Prog. Nuc. Acid Res. & Mol. Biol.* **51:** 1-51.

McNabb, D.S., Y. Xing, and L. Guarente. 1995. Cloning of yeast *HAP5*: a novel subunit of a heterotrimeric complex required for CCAAT binding. *Genes. Dev.* **9:** 47-58.

Mehta, K.D., and M. Smith. 1989. Identification of an upstream repressor site controlling the expression of an anaerobic gene (*ANB1*) in *Saccharomyces cerevisiae. J. Biol. Chem.* **264:** 8670-8675.

Miyake, S. and T. Sugimura. 1968. Coproporphyrinogenase in a respiration-deficient mutant of yeast lacking all cytochromes and accumulating coproporphyrin. *J. Bacteriol. Chem.* **96:** 1997-2003.

Mukai, Y., S. Harashima, and Y. Oshima. 1991. AAR1/TUP1 protein, with a structure similar to that of the beta subunit of G proteins is required for a1-alpha 2 and alpha 2 repression in cell type control of *Saccharomyces cerevisiae. Mol. Cell. Biol.* **11:** 3773-3779.

Mukerji, S.K., and N.R. Pimstone. 1987. Evidence for two uroporphyrinogen decarboxylase isoenzymes in human erythrocytes. *Biochem. Biophys. Res. Commun.* **146:** 1196-1203.

Myers, A.M., A. Tzagoloff, D.M. Kinney, and C.J. Lusty. 1986. Yeast shuttle and integrative vectors with multiple cloning sites suitable for construction of *lacZ* fusions. *Gene* **45:** 299-310.

Myers, A.M., M.D. Crivellone, T.J. Koerner, and A. Tzagoloff. 1987. Characterization of the yeast *HEM2* gene and transcriptional regulation of *COX5* and *COR1* by heme. *J. Biol. Chem.* **202:** 16822-16829.

Nehlin, J.O., and H. Ronne. 1990. Yeast MIG1 repressor is related to the mammalian early growth response and Wilms' tumour finger proteins. *EMBO J.* **9:** 2891-2898.

Nehlin, J.O., M. Carlberg, and H. Ronne. 1991. Control of yeast *GAL* genes by Mig1 repressor: a transcriptional cascade in the glucose response. *EMBO J.* **10:** 3373-3377.

Nicholson, D.W., C. Hergersberg, and W. Neupert. 1988. Role of cytochrome *c* lyase in the import of cytochrome *c* into mitochondria. *J. Biol. Chem.* **263:** 19034-19042.

Nishimura, K., S. Taketani, and H. Inokuchi. 1995. Cloning of a human cDNA for protoporphyrinogen oxidase by complementation *in vivo* of a *hemG* mutant of *Escherichia coli*. *J. Biol. Chem.* **270:** 8076-8080.

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

Olesen, J. T., and L. Guarente. 1990. The *HAP2* subunit of yeast CCAAT transcriptional activator contains adjacent domains for subunit association and DNA recognition: model for the HAP2/3/4 complex. *Genes Dev.* **4:** 1714-1729.

Olesen, J., S. Hahn, and L. Guarente. 1987. Yeast HAP2 and HAP3 activators both bind to the *CYC1* upstream activation site, UAS2, in an interdependent manner. *Cell* **51:** 953-961.

Oosterwegel, M., M. van de Wetering, D. Dooijes, L. Klomp, A. Winoto, K. Georgopoulos, F. Meijlink, and H. Clevers. 1991. Cloning of murine TCF-1, a T cell-specific transcription factor interacting with functional motifs in the CD3-epsilon and T cell receptor alpha enhancers. *J. Exp. Med.* **173:** 1133-1142.

Özcan, S., and M. Johnston. 1995. Three different regulatory mechanisms enable yeast hexose transporter (HXT) genes to be induced by different levels of glucose. *Mol. Cell. Biol.* **15:** 1564-1572.

Parisi, M.A., and D.A. Clayton. 1991. Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. *Science* **252:** 965-969.

Pathak, V.K., D. Schindler, and J.W.B. Hershey. 1988. Generation of a mutant form of protein synthesis initiation factor eIF-2 lacking the site of phosphorylation by eIF-2 kinases. *Mol. Cell. Biol.* **8:** 993-995.

Pfeifer, K., K-S. Kim, S. Kogan, and L. Guarente. 1989. Functional dissection and sequence of yeast HAP1 activator. *Cell* **56:** 291-301.

Pfeifer, K., T. Prezant, and L. Guarente. 1987. Yeast HAP1 activator binds to two upstream activtion sites of different sequence. *Cell* **49:** 19-27.

Pinkham, J. L., and L. Guarente. 1985. Cloning and molecular analysis of the *HAP2* locus: a global regulator of respiratory genes in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **5:** 3410-3416.

Pinkham, J.L., and T. Keng. 1994. Heme-mediated gene regulation in *Saccharomyces cerevisiae*. In: *Metal Ions and Fungi*. Winkleman, G. and D.R. Winge, ed., Marcel Dekker Inc., New York, pp. 455-501.

Pinkham, J.L., J.T. Olesen, and L. Guarente. 1987. Sequence and nuclear localization of the *Saccharomyces cerevisiae* HAP2 protein, a transcriptional activator. *Mol. Cell. Biol.* **7:** 578-585.

Pinkham, J.L., Z. Wang, and J. Alsina. 1997. Heme regulates *SOD2* transcription by activation and repression in *Saccharomyces cerevisiae*. *Curr. Genet.* **31:** 281-291.

Ponka, P., and H.M. Schulman. 1993. Regulation of heme biosynthesis: distinct regulatory features in erythroid cells. *Stem Cells* **11:** 24-35.

Pontiggia, A., R. Rimini, V.R. Harley, P.N. Goodfellow, R. Lovell-Badge, and E. Bianchi. 1994. Sex-reversing mutations affect the architecture of the SRY-DNA complexes. *EMBO J.* **13:** 6115-6124.

Pretlow, T.P., and F. Sherman. 1967. Porphyrins and zinc porphyrins in normal and mutant strains of yeast. *Biochim. Biophys. Acta* **148:** 629-644.

Prezant, T., K. Pfeifer, and L. Guarente. 1987. Organization of the regulatory region of the yeast *CYC7* gene: multiple factors are involved in regulation. *Mol. Cell. Biol.* **7:** 3252-3259.

Repetto, B., and A. Tzagoloff. 1989. Structure and regulation of *KGD1*, the structural gene for yeast α-ketoglutarate dehydrogenase. *Mol. Cell. Biol.* **9:** 2695-2705.

Repetto, B., and A. Tzagoloff. 1990. Structure and regulation of *KGD2*, the structural gene for yeast dihydrolipoyl transsuccinylase. *Mol. Cell. Biol.* **10:** 4221-4232.

Richard, C. 1990. Deletion analysis of the sequences that regulate *HEM13* expression. M.Sc. thesis. McGill University, Montreal, Quebec, Canada.

Romana, M., P. LeBoulch, and P.H. Romeo. 1987. Rat uroporphyrinogen decarboxylase cDNA: nucleotide sequence and comparison to human uroporphyrinogen decarboxylase. *Nucleic Acids Res.* **17:** 7211.

Romeo, G., and E.Y. Levin. 1971. Uroporphyrinogen decarboxylase from mouse spleen. *Biochim. Biophys. Acta* **230:** 330-341.

Romeo, P.H., A. Dubart, B. Grandchamp, H. De Verneuil, J. Rosa, Y. Nordmann, and M. Goossens. 1984. Isolation and identification of a cDNA clone coding for rat uroporphyrinogen decarboxylase. *Proc. Natl. Acad. Sci. USA* **81:** 3346-3350.

Romeo, P.H., N. Raich, A. Dubart, D. Beaupain, M. Pryor, J. Kushner, M. Cohen-Solal, and M. Goossens. 1986. Molecular cloning and nucleotide sequence of a complete human uroporphyrinogen decarboxylase. *J. Biol. Chem.* **261:** 9825-9831.

Rosenkrantz, M., C.S. Kell, E.A. Rennell, and L.J. Devenish. 1994. The HAP2, 3, 4 transcriptional activator is required for derepression of the yeast citrate synthase gene, *CIT1*. *Mol. Microbiol.* **13:** 119-131.

Rothstein, R.J. and F. Sherman. 1980. Genes affecting the expression of cytochrome c in yeast: genetic mapping and genetic interactions. *Genetics* **94:** 871-889.

Rytka, J., T. Bilinski, and R. Labbe-Bois. 1984. Modified uroporphyrinogen decarboxylase activity in a yeast mutant which mimics porphyria cutanea tarda. *Biochem. J.* **218:** 405-413.

Sabova, L., I. Zeman, F. Supek, and J. Kolarov. 1993. Transcriptional control of

AAC3 gene encoding mitochondrial ADP/ATP translocator in *Saccharomyces cerevisiae* by oxygen, heme and ROX1 factor. *Eur. J. Biochem.* **213:** 547-553.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: a Laboratory Manual,* second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

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

Sasarman, A., A. Nepveu, Y. Echelard, J. Dymetryszyn, M. Drolet, and C. Goyer. 1987. Molecular cloning and sequencing of the *hemD* gene of *Escherichia coli* K-12 and preliminary data on the Uro operon. *J. Bacteriol.* **169:** 4257-4262.

Schneider, J.C., and L. Guarente. 1991. Regulation of the yeast *CYT1* gene encoding cytochrome *c*1 by HAP1 and HAP2/3/4. *Mol. Cell. Biol.* **11:** 4934-4942.

Schultz, J. and M. Carlson. 1987. Molecular analysis of *SSN6*, a gene functionally related to the SNF1 protein kinase of *Saccharomyces cerevisiae. Mol. Cell. Biol.* **7:** 3637-3645.

Schultz, J., and M. Carlson. 1987. Molecular analysis of *SSN6*, a gene functionally related to the SNF1 protein kinase of *Saccharomyces cerevisiae. Mol. Cell. Biol.* **7:** 3637-3645.

Schultz, J., L. Marshall-Carlson, and M. Carlson. 1990. The N-terminal TPR region is the functional domain of SSN6, a nuclear phosphoprotein of *Saccharomyces cerevisiae. Mol. Cell. Biol.* **10:** 4744-4756.

Shackelford, G.M., and H.E. Varmus. 1987. Expression of the proto-oncogene *int-1* is restricted to postmeiotic male germ cells and the neural tube of mid-gestational embryos. *Cell* **50:** 89-95.

Sharp, P.M., and W.-H. Li. 1987. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15:** 1281-1295.

Sherman, F., G.R. Fink, and J.B. Hicks. 1986. *Methods in yeast genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sinclair, A.H., P. Berta, M.S. Palmer, J.R. Hawkins, B.L. Griffiths, M.J. Smith, J.W. Foster, A.-M. Frischauf, R. Lovell-Badge, and P.N. Goodfellow. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346:** 240-244.

Smith, D.B., and K.S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67:** 31-40.

Staben, C., and C. Yanofsky. 1990. *Neurospora crassa* a mating-type region. *Proc. Natl. Acad. Sci. USA* **87:** 4917-4921.

Starr, D.B., and D.K. Hawley. 1991. TFIID binds in the minor groove of the TATA box. *Cell* **67:** 1231-1240.

Straka, J.G., and J.P. Kushner. 1983. Purifcation and characterization of bovine hepatic uroporphyrinogen decarboxylase. *Biochemistry* **22:** 4664-4672.

Straka, J.G., J.M. Rank, and J.R. Bloomer. 1990. Porphyria and porphyrin metabolism. *Ann. Rev. Med.* **41:** 457-469.

Struhl, K. 1985. Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl. Acad. Sci. USA* **82:** 8419-8423.

Stukey, J.E., V.M McDonough, and C.E. Martin. 1990. The *OLE1* gene of *Saccharomyces cerevisiae* encodes the Δ9 fatty acid desaturase and can be functionally replaced by the rat stearoyl-CoA desaturase gene. *J. Biol. Chem.* **265:** 20144-20149.

Sugimoto, A., Y. Iino, T. Maeda, Y. Watanabe, and M. Yamamoto. 1991. *Schizosaccharomyces pombe ste11*$^{+}$ encodes a transcription factor with an HMG motif that is a critical regulator of sexual development. *Genes and Dev.* **5:** 1990-1999.

Tait, G.H. 1978. In: *Heme and Hemoproteins.* F. De Metteis, and W.N. Aldridge, eds. Springer-Verlag, New York. pp. 1-48.

Taketani, S., H. Kohno, T. Furukawa, T. Yoshinaga, and R. Rokunaga. 1994. Molecular cloning, sequencing and expression of cDNA encoding human coproporphyrinogen oxidase. *Biochim. Biophys. Acta* **1183:** 547-549.

Taketani, S., T. Yoshinaga, T. Furukawa, H. Kohno, R. Tokunga, K. Nishimura, and H. Inokuchi. 1995. Induction of terminal enzymes for heme biosynthesis during differentiation of mouse erythroleukemia cells. *Eur. J. Biochem.* **230:** 760-765.

Thorsness, M.W., W. Schafer, L. D'Ari, and J. Rine. 1989. Positive and negative transcriptional control by heme of genes encoding 3-hydroxy. C-3-methylglutaryl coenzyme A reductase in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **9:** 5702-5712.

Thrash-Bingham, C. and W.L. Fangman. 1989. A yeast mutation that stabilizes a plasmid bearing a mutated *ARS1* element. *Mol. Cell. Biol.* **9:** 809-816.

Tomio, J.M., R.C. Garcia, L.C. San Martin de Viale, and M. Grinstein. 1970. Porphyrin biosynthesis VII. Porphyrinogen carboxy-lyase from avian erythrocytes. Purification and properties. *Biochim. Biophys. Acta* **198:** 353-363.

Travis, A., A. Amsterdam, C. Belanger, and R. Grosschedl. 1991. LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor α enhancer function. *Genes Dev.* **5:** 880-894.

Trawick, J.D., N. Kraut, F.R. Simon, and R.O. Poyton. 1992. Regulation of yeast *COX6* by the general transcription factor ABF1 and separate HAP2- and heme-responsive elements. *Mol. Cell. Biol.* **12:** 2302-2314.

Trawick, J.D., R.M. Wright, and R.O. Poyton. 1989. Transcription of yeast *COX6*, the gene for cytochrome *c* oxidase subunit VI, is dependent on heme and on the *HAP2* gene. *J. Biol. Chem.* **264:** 7005-7008.

Treitel, M.A., and M. Carlson. 1995. Repression by SSN6-TUP1 is directed by MIG1, a repressor/activator protein. *Proc. Natl. Acad. Sci. USA* **92:** 3132-3136.

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.* **177:** 3326-3331.

Trueblood, C.E., R.M. Wright, and R.O Poyton. 1988. Differential regulation of the two genes encoding *Saccharomyces cerevisiae* cytochrome *c* oxidase subunit V by heme and the *HAP2* and *REO1* genes. *Mol. Cell. Biol.* **8:** 4537-4540.

Trumbly, R.J. 1986. Isolation of *Saccharomyces cerevisiae* mutants constitutive for invertase synthesis. *J. Bacteriol.* **166:** 1123-1127.

Trumbly, R.J. 1988. Cloning and characterization of the *CYC8* gene mediating glucose repression in yeast. *Gene* **73:** 97-111.

Tsai, S.F., D.F. Bishop, and R.J. Desnick. 1988. Human uroporphyrinogen synthase III synthase: Molecular cloning, nucleotide sequence, and expression of a full-length cDNA. *Proc. Natl. Acad. Sci. USA* **85:** 7045-7053.

Tsukamoto, I., T. Yoshinaga, and S. Sano. 1979. The role of zinc with special reference to the essential thiol groups in delta-aminolevulinic acid dehydratase of bovine liver. *Biochim. Biophys. Acta* **570:** 167-178.

Turcotte, B., and L. Guarente. 1992. HAP1 positive control mutants specific for one of two binding sites. *Genes & Dev.* **6:** 2001-2009.

Turi, T.G., and J.C. Loper. 1992. Multiple regulatory elements control expression of the gene encoding the *Saccharomyces cerevisiae* cytochrome P_{450}, lanosterol 14 α-demethylase *(ERG11). J. Biol. Chem.* **267:** 2046-2056.

Tzamarias, D., and K. Struhl. 1994. Functional dissection of the yeast CYC8-TUP1 transcriptional corepressor complex. *Nature* **369:** 758-761.

Tzamarias, D., and K. Struhl. 1995. Distinct TPR motifs of Cyc8 are involved in recruiting the Cyc8-Tup1 corepressor complex to differentially regulated promoters. *Genes Dev.* **9:** 821-831.

Urban-Grimal, D., and R. Labbe-Bois. 1981. Genetic and biochemical characterization of mutants of *Saccharomyces cerevisiae* blocked in six different steps of heme biosynthesis. *Molec. Gen. Genet.* **183:** 85-92.

Urban-Grimal, D., C. Volland, T. Garnier, P. Dehoux, and R. Labbe-Bois. 1986. The nucleotide sequence of the *HEM1* gene and evidence for a precursor form of the mitochondrial 5-aminolevulinate synthase in *Saccharomyces cerevisiae. Eur. J. Biochem.* **156:** 511-519.

Ushinsky, S.C., and T. Keng. 1994. A novel allele of *HAP1* causes uninducible expression of *HEM13* in *Saccharomyces cerevisiae. Genetics* **136:** 819-831.

van de Wetering, M., M. Oosterwegel, D. Dooijes, and H. Clevers. 1991. Identification of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *EMBO J.* **10:** 123-132.

Varanasi, U.S., M. Klis, P.B. Mikesell, and R.J. Trumbly. 1996. The Cyc8 (Ssn6)-Tup1 corepressor complex is composed of one Cyc8 and four Tup1 subunits. *Mol. Cell. Biol.* **16:** 6707-6714.

Verdiere, J., F. Creusot, and L. Guarente. 1986. The overproducing *CYP1* mutation and the underproducing *hap1* mutations are alleles of the same gene which regulates in *trans* the expression of the structural genes encoding iso-cytochromes c. *Curr. Genet.* **10:** 339-342.

Verdiere, J., M. Gaisne, and R. Labbe-Bois. 1991. *CYP1 (HAP1)* is a determinant effector of alternative expression of heme-dependent transcription in yeast. *Mol. Gen. Genet.* **228:** 300-306.

Verdiere, J., M. Gaisne, B. Guiard, N. Defranoux, and P.P. Slonimski. 1988. CYP1 (HAP1) regulator of oxygen-dependent gene expression in yeast. 2. Missense mutation suggests altenative Zn fingers as discriminating agents of gene control. *J. Mol. Biol.* **204:** 277-282.

Wahi, M. and A.D. Johnson. 1995. Identification of genes required for alpha-2 repression in *Saccharomyces cerevisiae. Genetics* **140:** 79-90.

Waterman, M.L., W.H. Fischer, K.A. Jones. 1991. A thymus-specific member of the HMG protein family regulates the human T cell receptor Cα enhancer. *Genes Dev.* **5:** 656-669.

Wek, R.C. 1994. eIF-2 kinases: regulators of general and gene-specific translation initiation. *Trends Biochem. Sci.* **19:** 491-496.

Werner, M.H., J.R. Huth, A.M. Gronenborn, and G.M.Clore. 1995. Molecular basis of human 46X, Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell* **81:** 705-713.

Wickner, R.B. 1974. Mutants of *Saccharomyces cerevisiae* that incorporate deoxythymidine-5'-monophosphate into deoxyribonucleic acid in vivo. *J. Bacteriol.* **117:** 252-260.

Williams, F.E., and R.J. Trumbly. 1990. Characterization of *TUP1*, a mediator of glucose repression in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **10:** 6500-6511.

Williams, F.E., U. Varanasi, and R.J. Trumbly. 1991. The CYC8 and TUP1 proteins involved in glucose repression in *Saccharomyces cerevisiae* are associated in a protein complex. *Mol. Cell. Biol.* **11:** 3307-3316.

Winkler, H., G. Adam, E. Mattes, M. Schanz, A. Hartig, and H. Ruis. 1988. Co-ordinate control of synthesis of mitochondrial and non-mitochondrial hemoproteins: a binding site for the HAP1 (CYP1) protein in the UAS region of the yeast catalase T gene *(CTT1). EMBO J.* **7:** 1799-1804.

Wright, C.F. and R.S. Zitomer. 1984. A positive regulatory site and a negative regulatory site control the expression of the *Saccharomyces cerevisiae* CYC7 gene. *Mol. Cell. Biol.* **4:** 2023-2030.

Wu, C., W. Xu, C.A. Kozak, and R.J. Desnick. 1996. Mouse uroporphyrinogen decarboxylase: cDNA cloning, expression, and mapping. *Mammalian Genome* **7:** 349-352.

Wyckoff, E.A., J.D. Phillips, A.M. Sowa, M.R. Franklin, and J.P. Kushner. 1996. Mutational analysis of human uroporphyrinogen decarboxylase. *Biochim. Biophys. Acta* **1298:** 294-304.

Xing, Y., J.D. Fikes, and L. Guarente. 1993. Mutations in yeast HAP2/HAP3 define a hybrid CCAAT box binding domain. *EMBO J.* **12:** 4647-4655.

Xu, K., J. Delling, and T. Elliott. 1992. The genes required for heme synthesis in *Salmonella typhimurium* include those encoding alternative functions for aerobic and anaerobic coproporphyrinogen oxidase. *J. Bacteriol.* **174:** 3953-3963.

Zagorec, M., and R. Labbe-Bois. 1986. Negative control of yeast coproporphyrinogen oxidase synthesis by heme and oxygen. *J. Biol. Chem.* **261:** 2506-2509.

Zagorec, M., J.M. Buhler, I. Treich, T. Keng, L. Guarente, and R. Labbe-Bois. 1988. Isolation, sequence, and regulation by oxygen of the yeast *HEM13* gene coding for coproporphyrinogen oxidase. *J. Biol. Chem.* **263:** 9718-9724.

Zaret, K.S., and F. Sherman. 1982. DNA sequence required for efficient transcription termination in yeast. *Cell* **28:** 563-573.

Zhang, L., and L. Guarente. 1995. Heme binds to a short sequence that serves a regulatory function in diverse proteins. *EMBO J.* **14:** 313-320.

Zhang, L., O. Bermingham-McDonough, B. Turcotte, and L. Guarente. 1993. Antibody-promoted dimerization bypasses the regulation of DNA binding by the heme domain of the yeast transcriptional activator HAP1. *Proc. Natl. Acad. Sci. USA* **90:** 2851-2855.

Zhang, M., L.S. Rosenblum-Vos, C.V. Lowry, K.A. Boakye, and R.S. Zitomer. 1991. A yeast protein with homology to the β-subunit of G proteins is involved in control of heme-regulated and catabolite-repressed genes. *Gene* **97:** 153-161.

Zitomer, R.S., and C.V. Lowry. 1992. Regulation of gene expression by oxygen in *Saccharomyces cerevisiae*. *Micrbiol. Rev.* **56:** 1-11.

Zitomer, R.S., M.P. Limbach, A.M. Rodriguez-Torres, B. Balasubramanian, J. Deckert, and P.M. Snow. 1997a. Approaches to the study of ROX1 repression of the hypoxic genes in yeast *Saccharomyces cerevisiae*. *Methods: A Companion to Methods in Enzymology* **11:** 279-288.

Zitomer, R.S., P. Carrico, and J. Deckert. 1997b. Regulation of hypoxic gene expression in yeast. *Kidney Int.* **51:** 507-513.