INFORMATION TO USERS

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

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

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

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

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

U·M·I

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

Order Number 9218243

Molecular genetic studies of human delta-aminolevulinic acid dehydratase

Kaya, Angela Haruko, Ph.D.

City University of New York, 1992

Copyright ©1992 by Kaya, Angela Haruko. All rights reserved.

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

MOLECULAR GENETIC STUDIES OF HUMAN DELTA-AMINOLEVULINIC ACID DEHYDRATASE

by

Angela H. Kaya

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

1992

Copyright by

Angela H. Kaya

1992

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

November 11, 1991

11/11/91
Date

Chair of Examining Committee

Executive Officer

Supervisory Committee

The City University of New York

ABSTRACT

MOLECULAR GENETIC STUDIES OF
HUMAN DELTA-AMINOLEVULINIC ACID DEHYDRATASE

by: Angela H. Kaya

Advisor: Professor James G. Wetmur

The ALAD gene has been cloned and sequenced in both directions. The genomic sequence was found to be identical to the previously published cDNA sequence. The gene encodes thirteen exons: eleven coding exons, exons 2-12, and two non-coding exons, 1A and 1B. The zinc binding site was located in exon five. The active lysine was located in exon ten. Nine Alu repetitive elements were found in the gene. Additional types of repetitive elements were not found in the genomic sequence.

The most upstream of the ALAD non-coding exons, exon 1A, is presumed to be governed by a housekeeping promoter. The promoter region of exon 1A contained characteristics of a constitutive promoter. The promoter region of exon 1B contained characteristics of genes which are specifically regulated in erythrocytes.

The $ALAD^2$ allele was cloned and sequenced. The only mutation found was a G to C transversion at nucleotide 177 of the coding sequence. This transversion created an *Msp*I restriction endonuclease site and predicted the substitution of a positively charge lysine by a neutral asparagine in residue 59 of the ALAD 2 subunit. Because of the presence of the restriction endonuclease site, a simple PCR based method was employed for ALAD genotyping. Analysis of a 85 ALAD 1-2 individuals and eight ALAD 2-2 individuals revealed that, in all cases, the $ALAD^2$ allele phenotype corresponded with the

MspI RFLP genotype.

Analysis of a random population of 428 normal Caucasian individuals revealed the frequency of the ALAD^1 and the ALAD^2 allele frequencies were 0.88 and 0.12, respectively. The allele frequencies of a previously identified *Rsa*I RFLP were 0.75 (*Rsa*I$^-$) and 0.25(*Rsa*I$^+$). Individually, the RFLPs were in Hardy-Weinberg equilibrium, however, the two RFLPs taken together were in linkage disequilibrium. The ALAD^2/*Rsa*$^+$ allele was highly under represented.

Bacterial expression of an authentic, soluble, biologically active human ALAD^1 was achieved using a co-cistronic vector derived from pUC9. The ALAD enzymatic activity was over 100 times that found in human erythrocytes. The recombinant protein was identical to human erythrocyte protein by polyacrylamide gel electrophoresis mobility, Western blotting, elution of the homooctamer from a gel filtration column and its sedimentation in a sucrose gradient.

FOREWORD

Portions of this thesis have been presented in the following publications, permission of the publishers has been granted.

Wetmur, J. G., Kaya, A.H., Plewinska, M. and Desnick, R. J. (1991). Molecular Characterization of the Human delta-Aminolevulinate Dehydratase 2 ($ALAD^2$) Allele: Implications for Molecular Screening of Individuals for Genetic Susceptibility to Lead Poisoning. *American Journal of Human Genetics* 49, 757-763.

Astrin, K. H., Kaya, A. H., Wetmur, J. G., and Desnick, R. J. (1991). RsaI polymorphism in the human delta-aminolevulinate dehydratase gene at 9q34. *Nucleic Acids Research* **19**, 4307.

ACKNOWLEDGEMENTS

I wish to express my gratitude and appreciation to the following individuals all without whom this work could not have been completed: Dr. James Wetmur and Dr. Robert Desnick; Dr. Kenneth Astrin, Dr. David Bishop, Mr. Seth Consigli, Mr. John DiMari, Dr. Marie Grace, Mr. Nicholas Hand, Dr. Yiannis Ioannou, Ms. Judy Nussbaum, Dr. Magdalena Plewinska, Dr. Robin Quartin, Mr. Daniel Rosenberg, Dr. Peter Tsai, Mr. Peter Weinstock, and Ms. Daphne Wong-Szeto; and my son, Zachary Alexander Huizar.

DEDICATION

This thesis is dedicated to my parents, Roberta and Albert Kaya.

TABLE OF CONTENTS

APPROVAL PAGE iii

ABSTRACT iv

FOREWORD vi

ACKNOWLEDGEMENTS vii

DEDICATION viii

TABLE OF CONTENTS ix

LIST OF TABLES xii

LIST OF FIGURES xiii

LIST OF ABBREVIATIONS xv

I. Background 1

A. Heme Biosynthesis 1

B. Regulation and tissue specificity of heme biosynthesis 5

C. Molecular analysis of heme biosynthetic enzymes and role in regulation/tissue specificity 7

D. Delta-aminolevulinic acid dehydratase (ALAD) 11

E. Lead toxicity and ALAD 22

II. Objectives 33

III. Materials 35

A. Chemicals and biochemicals 35

B. Vectors 37

C. Cell lines 40

IV. Methods 42

A. Oligonucleotide synthesis 42

B. Cloning 42

C. Sequencing 42

D. RNA isolation from cultured cells 42

E. Primer extension 43

F. Polymerase chain reaction (PCR) 43

G. DNaseI protection 43

H. CAT expression 44

I. Northern analysis 45

J. Preparation of genomic DNA from whole blood 45

K. Starch gel electrophoresis 46

L. Cellulose acetate gel electrophoresis 46

M. Bacterial lysis 47

N. Quantitation of total protein 48

O. ALAD activity assay 49

P. Polyacrylamide electrophoresis 49

Q. Silver stain 49

R. Immunodetection 49

S. Purification of recombinant ALAD 50

T. Sucrose gradient 51

V. Results 52

A. Characterization of ALAD gene 52

B. Isolation and characterization of the ALAD-2 allele 66

C. Amplification method for ALAD genotyping 67

D. Bacterial expression of ALAD-1 and ALAD-2 alleles 71

VI. Discussion 81

A. The ALAD gene 81

B. The ALAD alleles 83

C. Recombinant human ALAD 86

VII. Concluding Remarks 90

A. ALAD genomic sequence 90

B. ALAD polymorphism 90

C. ALAD proteins 91

VIII. Bibliography 92

List of Tables

1. $ALAD^2$ allele and increased blood lead levels 30

2. $ALAD^2$ allele and median blood lead levels 31

3. Chemicals and biochemicals 35

4. Specific activities in purification scheme 79

List of Figures

1. Current concept of heme biosynthesis	2
2. Inborn errors of heme biosynthesis	3
3. Reaction catalyzed by ALAD	12
4. Catalytic site of ALAD	18
5. Zinc binding site of ALAD	19
6. Complete active site of model of ALAD	20
7. Lead exposed children-USA	27
8. Lead workers-FRG	28
9. pUC9 and pUC19 vectors	37
10. pKK223-3 vector	38
11. pSV0 and pSV2 CAT vectors	39
12. Genomic sequence	53
13. Genomic map	61
14. Intron/exon boundaries	62
15. Promoter elements of exon 1A	64
16. Promoter elements of exon 1B	65
17. PCR-based method for ALAD genotyping	68
18. ALAD haplotypes	70
19. pUC9-50 expression vector	72
20. Other pUC9 expression vectors	73
21. Western blot analysis of recombinant ALAD	76

22. Western blot and silver stain of purification scheme 78

List of Abbreviations

ALA	aminolevulinic acid
ALAD	aminolevulinic acid dehydratase
ALAS	aminolevulinic acid synthase
bp	base pair
BSA	bovine serum albumin
CAT	chloramphenicol acetyltransferase
CRIM	cross reacting material
COPRO-OX	coproporphyrinogen oxidase
da	dalton
DEAE	diethylaminoethyl
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid
FEP	free erythrocyte protoporphyrin
HMB	hydroxymethylbilane
HMBS	hydroxymethylbilane synthase
HRP	horse radish peroxidase
IPTG	isopropyl ß-D-thiogalactopyranoside
LB	Luria broth
LINE	long interspersed repetitive element

Lv	levulinate locus
nt	nucleotide
PAGE	polyacrylamide gel electrophoresis
PBG	porphobilinogen
PBGD	porphobilinogen deaminase
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEI	polyethyleneimine
PMSF	phenylmethylsulphonylflouride
RFLP	restriction fragment length polymorphism
SDS	sodium dodecyl sulfate
SSC	sodium citrate
STS	sequence tagged site
TCA	trichloroacetic acid
THE	transposon like element
U	unit
URO Co-S	uroporphyrinogen III co-synthase
UROD	uroporphyrinogen decarboxylase
X-gal	5-bromo-4-chloro-3-indoyl-ß-D-galactoside

I. Background

A. Heme biosynthesis

Heme, ferroprotoporphyrin IX, the prosthetic group of hemoglobin, myoglobin, mitochondrial cytochromes, catalase, peroxidases and tryptophan pyrrolase, is a metal-organic compound, and the end product of the heme-biosynthetic pathway (Figure 1). The porphyrias are disorders which are the result of defects in specific enzymes of the heme-biosynthetic pathway (Figure 2). Porphyrias, both inherited and acquired, are classified as hepatic or erythropoietic depending upon the site of phenotypic expression. In humans, these disorders are characterized by increased accumulation and excretion of chemical intermediates and by various clinical manifestations, in particular, neurologic abnormalities and cutaneous photosensitivity.

Although clinicians and researchers have been intrigued by the porphyrias and their clinical manifestations since the turn of this century, the elucidation of the biosynthesis of heme was not begun until the 1940's and enzymatic deficiencies in the porphyrias were not identified until the mid-1970's. In 1946, Shemin and Rittenberg performed radioactive tracer experiments which demonstrated that glycine and succinic acid acted as precursors of the porphyrins (Shemin and Rittenberg, 1946). Subsequently, they proposed that the product of the condensation of glycine and succinate was the 5-carbon amino ketone δ-aminolevulinic acid (ALA) (Radin, et al, 1950). The monopyrrole porphobilinogen (PBG) is formed from two molecules of ALA. In 1975, it was shown that two enzymes, porphobilinogen deaminase (PBGD) and uroporphyrinogen cosynthase (URO Co-S), were required to catalyze the conversion of four molecules of the PBG into

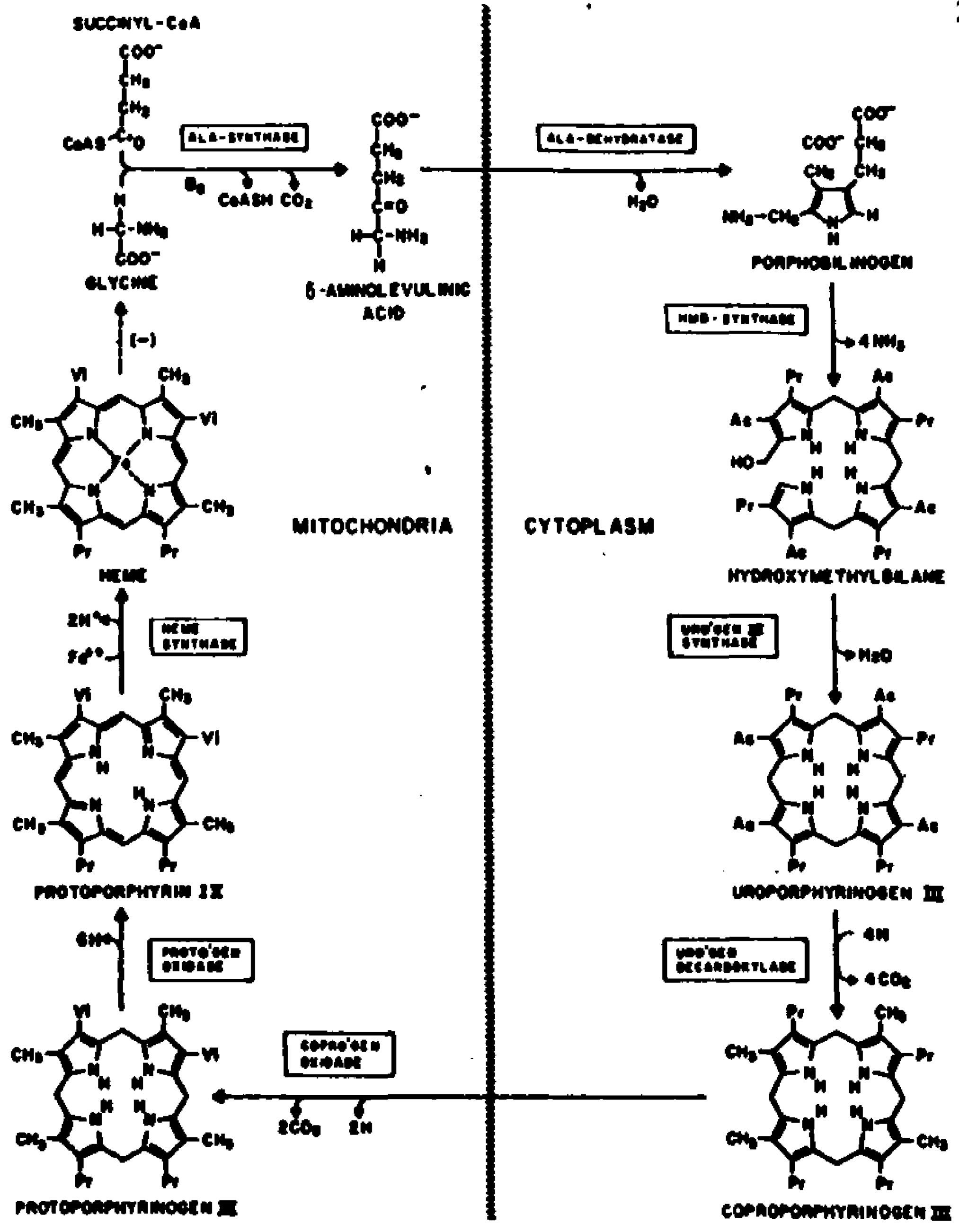


Figure 1. The heme-biosynthetic pathway.

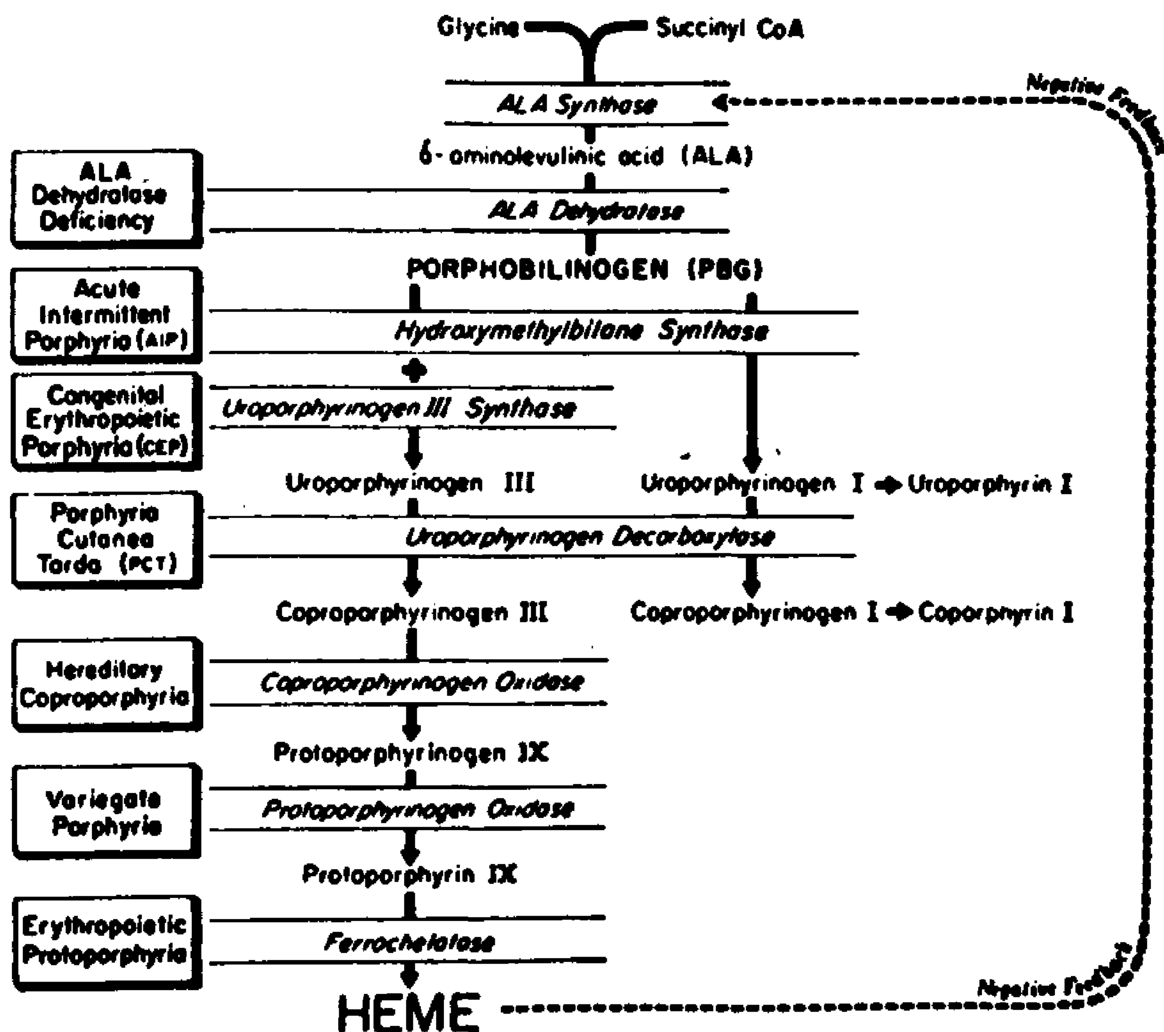


Figure 2. The porphyrias in the heme-biosynthetic pathway.

uroporphyrinogen III, a tetrapyrrole (Battersby and McDonald, 1975). Additional enzymatic steps affect the side chains of the tetrapyrrole. It is now known that eight distinct enzymes are required in order to convert eight moles of glycine and eight moles of succinyl CoA through the porphyrin precursors, the porphyrins (tetrapyrroles) and, finally, into one mole of heme.

As outlined in Figure 1, aminolevulinic acid synthase (ALAS) catalyzes the condensation of glycine and succinyl CoA to form ALA. ALAD then catalyzes the condensation of two molecules of ALA to form a monopyrrole, PBG. Enzyme catalyzed condensation of four molecules of PBG yields a linear tetrapyrrole, hydroxymethylbilane (HMB), which cyclizes spontaneously to yield uroporphyrinogen I. In the presence of both PBGD and URO Co-S, uroporphyrinogen III is formed, an isomer of uroporphyrinogen I. Ring D of the tetrapyrrole is isomerized. Uroporphyrinogen decarboxylase (UROD) catalyzes the sequential removal of the four carboxylic groups of the carboxymethyl side chains of uroporphyrinogen to yield coproporphyrinogen. Coproporphyrinogen oxidase removes the carboxyl group and the two hydrogens from the propionic acid groups of pyrrole rings A and B of coproporphyrinogen to form vinyl groups at these positions. The reaction yields a divinyl compound protoporphyrinogen IX. Protoporphyrinogen oxidase oxidizes protoporphyrinogen IX to protoporphyrin IX. The final step of heme biosynthesis is catalyzed by ferrochelatase which inserts iron into protoporphyrin IX to form ferroprotoporphyrin IX, or heme. Aminolevulinic acid synthase, the first enzyme of the pathway, and coproporphyrinogen oxidase, protoporphyrinogen oxidase and ferrochelatase, the last three enzymes of the pathway,

function in the mitochondria. The four intermediate enzymes are located in the cytosol.

The heme biosynthetic pathway and the associated inherited porphyrias provide a unique opportunity to investigate the genetic organization and regulation of the enzymes which control this essential system. During the past decade, studies of the enzymes have provided insights into their properties and reaction mechanisms, the regulation and control of heme biosynthesis and the environmental influences which cause perturbations in heme biosynthesis and exacerbate the clinical expression of the porphyrias (for reviews see Bottomley and Muller-Eberhard, 1988; Grandchamp and Nordmann, 1988; Kappas et al., 1989).

B. Regulation and tissue specificity of heme biosynthesis

The human heme biosynthetic pathway is inducible and its regulation tissue-specific. All cells require low level expression of the heme biosynthetic genes in order to synthesize cytochromes and hemoproteins. Heme synthesis in hepatic cells is modulated by the end-product repression of ALAS, the first and rate-limiting enzyme in the pathway (Kappas et al., 1989). "Free heme" acts to inhibit the translocation of the cytosolic hepatic ALAS across the mitochondrial membrane and binds a repressor which inhibits the synthesis of the enzyme (Whiting and Granick, 1976; Yamaguchi et al, 1980). Acute porphyric attacks may be induced in individuals expressing up to 50% of normal levels of a heme biosynthetic enzyme (Kappas et al, 1989). The hormonal or environmental inducers significantly increase the concentration of ALA by inducing or stabilizing hepatic ALAS. The other hepatic heme biosynthetic enzymes remain at constitutive levels leading to the accumulation of toxic porphyrin precursors.

Induction of heme and hemoglobin synthesis (Levere et al., 1967) as well as the formation of ALA synthase (Irving et al., 1978) occurs after treatment of cultured chick blastoderm with a number of naturally occurring steroids. Such steroids are also capable of inducing erythroid colonies in human bone marrow cell in culture (Urabe et al., 1979) and of inducing ALA synthase and porphyrins in cultured chick embryo liver *in vitro* (Stephens et al., 1977; Sassa et al., 1979) and in isolated rat liver cell suspension.

The regulation of the erythroid ALAS gene is probably controlled by erythroid specific factors and may be affected by an iron responsive element found in the erythroid cDNA (Cox et al., 1991; Dandekar et al., 1991). Not only ALAS, but also the other enzymes of the heme biosynthetic pathway are increased during erythroid differentiation. Heme pathway enzymes increase in mouse Friend-virus transformed erythroleukemia cells after treatment with DMSO (Granchamp et al., 1985). This result suggests that ALAS may not always be the rate limiting enzyme for heme formation in erythroid cells. In fact, ferrochelatase may have a rate limiting role in heme biosynthesis. Hemoglobin formation does not occur until ferrochelatase activity increases in Friend cells treated with DMSO (Rutherford et al., 1979).

The response of erythroid cells to hemin is opposite to the response of liver and other cell types. Hemin can increase both globin and heme synthesis in undifferentiated Friend cells (Granick and Sassa, 1978). Hemin-mediated induction of ALAS and other enzymes in the pathway, has also been reported in human erythroid cells (Hoffman et al., 1980). The mechanisms responsible for the inducibility and tissue-specific expression of the heme biosynthetic genes are now being elucidated at the molecular level.

C. Molecular analysis of heme biosynthetic enzymes and role in regulation/tissue specificity

All of the enzymes of the heme biosynthetic pathway have been isolated making it possible to obtain information on some of their physicochemical, catalytic and molecular biological properties. cDNAs for six out of the eight enzymes, aminolevulinic acid synthase and dehydratase, porphobilinogen deaminase, uroporphyrinogen III synthase, uroporphyrinogen decarboxylase, coproporphyrinogen oxidase and ferrochelatase have been cloned in mammalian systems. Two enzymes currently have erythroid and non-erythroid isozymes. The erythroid and hepatic isozymes of ALAS (Bishop et al., 1981) arise from two separate genes (Bishop, 1991). Surprisingly, the two isozymes of PBGD arise from a single gene (Grandchamp et al., 1987). Alternative transcription and splicing produces two tissue-specific transcripts. There is also tissue-specific control of expression of other heme biosynthetic genes (Grandchamp and Nordmann, 1988).

ALAS cDNAs have been cloned from liver and erythroid cells in both chickens (Borthwick et al., 1985 and Riddle et al., 1989) and humans (Bishop, 1991). The hepatic sequence has also been determined for the rat (Yamamoto et al., 1988). The predicted human protein contained 586 amino acids. Human, rat and chicken hepatic ALAS shared eighty-nine percent identity on the amino acid level. The erythroid ALAS sequence has also been determined in the mouse (Schoenhaut and Curtis, 1989). Human and mouse erythroid ALAS shared ninety percent identity on the amino acid level. Human hepatic and erythroid ALAS show no homology in the amino-terminal region (196 residues) and only 73% homology in the remainder of the proteins.

The hepatic ALAS genomic sequence has been determined for the chicken (Maguire et al., 1986) The structural gene spans 6.9 kb and contains 10 exons. The promoter region contains a TATA and CAAT box and 4 GC rich regions, SP1 factor binding sites. In humans, the housekeeping gene, designated ALAS1, maps to chromosome band 3p21 (Astrin et al., 1987 and Sutherland et al., 1988). The erythroid specific gene, designated ALAS2, maps to the X chromosome (Bishop et al., 1990; Astrin et al., 1989 and Cox et al., 1990).

The two isozymes of porphobilinogen deaminase (PBGD), the third enzyme in the heme biosynthetic pathway, differ in molecular mass as determined by SDS polyacrylamide gel and immunodetection techniques (Grandchamp et al., 1987). The cDNAs for PBGD have been isolated from both erythroid and non-erythroid sources (Grandchamp et al., 1987; Raich et al., 1986). The non-erythroid cDNA contains an open reading frame with an additional 17 amino acid residues at its amino terminus compared with the erythroid cDNA. The erythroid form of PBGD is restricted to erythropoietic cells in human (Grandchamp, 1987). In mouse erythroleukemia cells, the non-erythroid form predominates in undifferentiated cells, but during maturation of these cells the erythroid specific isoform increases with time and becomes the major form in differentiated cells (Grandchamp and Nordmann, 1988).

The human PBGD gene contains 15 exons over 10 kb of DNA (Chretien et al., 1988). The non-erythroid and erythroid RNAs are produced by alternative splicing of two primary transcripts arising from two promoters. The upstream promoter is the housekeeping promoter and is active in all cells. The erythroid promoter is structurally

homologous to the ß-globin gene promoters. Using a run-on transcription assay in isolated nuclei, Beaumont et al. (1988) found that erythropoietic differentiation brings about a transcriptional activation of the erythroid specific promoter.

In transfection experiments with either erythroid or non-erythroid cells, the PBGD erythroid-specific promoter was correctly transcribed and regulated in erythroid cells and only the non-erythroid, or housekeeping, promoter was active in non-erythroid cells (Raich et al., 1989). DNase I footprinting, gel retardation and methylation interference assays showed the erythroid specific promoter to have three binding sites for an erythroid specific factor already described (NF-E1) and one binding site for second erythroid specific factor (NF-E2) not yet described (Mignotte et al., 1989). NF-E1 binding sites are common to the PBGD promoter, the ß-globin promoter (DeBoer et al., 1988) and the ß-globin 3′ enhancer (Wall et al., 1988).

Uroprophyrinogen III cosynthase, the fourth enzyme in the heme biosynthetic pathway catalyzes the formation of URO III from hydroxymethylbilane, the unstable product of PBGD. A human cDNA clone was isolated from an adult liver expression library (Tsai et al., 1988). An open reading frame of 798 base pairs consisted of 263 amino acids with a predicted molecular weight of 28,607 Da. The predicted amino acid sequence of the liver cDNA is colinear with peptides from purified erythrocyte enzyme representing 30% of the sequence, suggesting the hepatic and erythroid enzymes are the same (Tsai, 1988).

cDNA clones for uroporphyrinogen decarboxylase, the fifth enzyme in the heme biosynthetic pathway, were first obtained in rats (Romeo et al., 1984) and then in humans

(Romeo et al., 1986). The human cDNA contains an open reading frame with a predicted amino acid sequence of 367 residues and a molecular weight of 40,000 Da. The human UROD gene is over 3 kb in length and contains 10 exons (Romana et al., 1987). The promoter region contains a TATA box and an SP1 binding site. Northern analysis (Grandchamp and Nordmann, 1988) and sequencing of cDNAs from erythroid and non-erythroid sources (Romeo, 1986; de Verneuil et al., 1986) show that UROD mRNA is identical in different cell types. However, UROD mRNA levels were substantially greater in erythroid cells than in non-erythroid cells. UROD has two initiation sites (Romana, 1987). A major site was located 15 base pairs upstream from the initial ATG and a minor site is six base pairs from that ATG. Tissue specific enhancement of UROD is presumed by increased transcriptional activity from the single promoter.

Using oligonucleotides made from the amino acid sequence of the previously purified ferrochelatase protein, Brenner and Frasier (1991) cloned the murine ferrochelatase cDNA. The cDNA has an open reading frame of 397 amino acids and shares 47% identity with the predicted amino acid sequence of *Saccharomyces cerevisiae* ferrochelatase (Brenner and Frasier, 1991). Southern blot analysis results are consistent with the presence of a single gene, while Northern blot analysis indicates the presence of two transcripts in all tissues tested. Thus, similar to the UROD case, tissue-specific expression appears to result from different concentrations of both ferrochelatase messages in different tissues.

D. δ-aminolevulinic acid dehydratase (ALAD)

δ-aminolevulinate dehydratase (ALAD, ALA-D, porphobilinogen synthase, PBGS,

EC 4.2.1.24), the second enzyme in the heme biosynthetic pathway, catalyzes the asymmetric condensation of two molecules of 5-aminolevulinate (ALA) to form porphobilinogen (PBG), the monopyrrole precursor of heme, cytochromes and other hemoproteins. The mammalian enzyme has been purified to homogeneity from various species and shown to be a metalloenzyme of about 280 kDa which is composed of eight identical subunits (Wu et al., 1974, Anderson and Desnick, 1979) and eight zinc atoms (Tsukamoto et al., 1979; Bevan et al., 1980; Gibbs and Jordan, 1981; Dent et al., 1990). ALAD enzymes purified from different mammalian tissues are remarkably similar when molecular weight, K_m values and V_{max} values are compared (Shemin, 1972). ALAD from various tissues all cross react with an antibody against the hepatic enzyme (Yamamoto et al., 1986 and Chang et al., 1984). Finally, the amino acid sequences of the amino termini of human erythrocyte and bovine liver ALAD are colinear (reviewed in Wetmur et al., 1986a). Thus, there is no evidence for tissue-specific isozymes in various strains of mice and rats (Chang et al., 1987 and Yamamoto et al., 1987). The structural gene for human ALAD has been assigned to chromosome 9q34 (Potluri et al., 1987) and the full-length cDNA encoding the human enzyme has been cloned and sequenced (Wetmur et al., 1986a and 1986b).

A reaction mechanism has been proposed (Figure 3), based on studies by Nandi and Shemin (1968) using ALAD obtained from the photosynthetic bacterium *Rhodopseudomonas spheroides*, where the keto group of ALA forms a Schiff's base with the ε-amino group of a lysine on the enzyme. Competitive inhibitors of the enzyme also have a keto group γ to the carboxyl carbon. A second ALA molecule condenses with the

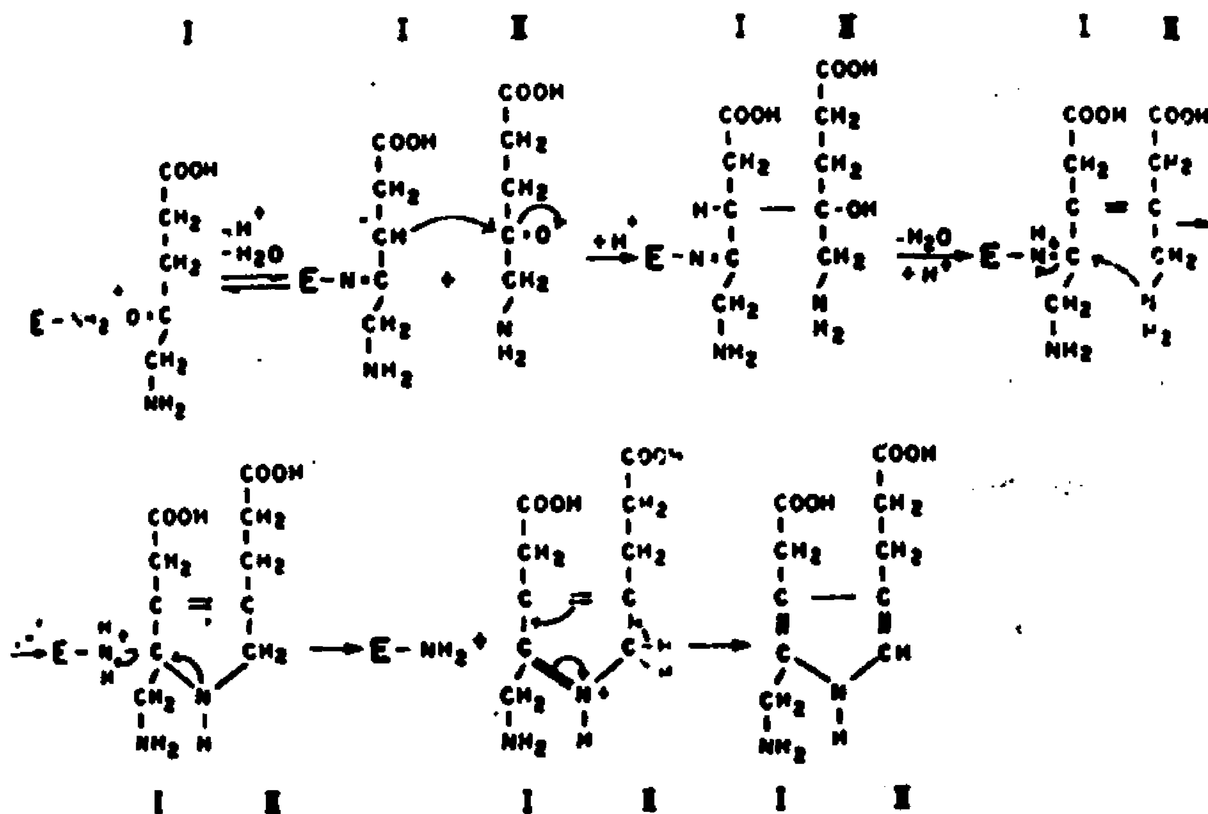


Figure 3. The reaction catalyzed by ALAD.

imine and the resulting aldol loses one molecule of water. In a transamination reaction, the free amino group displaces the ε-amino group of lysine on the enzyme, forming PBG. Jordan and Seehra (1980), through the use of ^{14}C labeling studies in bovine liver, have shown that the ALA which participated in the formation of the Schiff's base ultimately forms the propionic acid side chain of PBG, while the second ALA gives rise to the acetic acid moiety of PBG.

Studies of liver and erythrocyte ALAD have indicated that the enzymatic activity changes during liver development and erythropoietic maturation. ALAD activity in mouse and erythroid cells increases sharply during erythroid differentiation (Sassa, 1976 and 1980; Hoffman et al., 1980), then falls with the maturation of reticulocytes into mature erythrocytes (Sassa and Bernstein, 1977), the activity in erythrocytes being approximately 15 fold less than reticulocytes (Anderson et al., 1977). Similarly, the specific activity of ALAD in fetal rodent liver has been found to be three fold greater than in adult liver; the enzymatic activity was lowest at, or soon after, birth and gradually rose to adult levels during the first month of life. These developmental changes in enzymatic activity parallel those of PBGD and UROD, both of which are known to be regulated at the transcriptional level (Grandchamp et al., 1985). Enzyme protein concentrations are proportional to mRNA levels.

In 1979, a new, autosomal recessive, hepatic porphyria was described by Doss and coworkers in two unrelated German patients. The porphyria is characterized by deficient ALAD activity (3% of normal) which leads to the accumulation of ALA, the porphyrin precursor which presumably is responsible for the clinical manifestations of this

and related porphyrias (Kappas et al., 1989). In addition to the marked plasma and urinary accumulation of ALA, these patients have abnormally elevated levels of erythrocyte protoporphyrin and urinary coproporphyrin III and slightly increased porphyrin intermediates (Doss et al., 1979 and 1980; Thunell et al., 1987, Hassoun et al., 1989).

To date, only four unrelated homozygotes with this disease have been reported. Notably, there is remarkable phenotypic heterogeneity among these patients, including an infant who presented with failure to thrive (Thunell et al., 1987), two adolescent males with neurologic involvement presumably resulting from acute porphyric attacks (Doss et al., 1979 and 1980) and an adult male who developed a subacute polyneuropathy with predominance of motor signs in the upper limbs at age 63 years (Hassoun et al., 1989). The affected homozygotes had ALAD activities ranging from <2 to 12% of normal (Doss et al., 1979, 1980; Thunell et al., 1987, Hassoun et al., 1989) and immunologic studies revealed that from 20 to 33% of cross-reacting immunologic material (CRIM) was present in erythrocytes from the three youngest homozygotes, consistent with the expression of a non-functional enzyme subunit by at least one of the mutant alleles (de Verneuil et al., 1985; Fujita et al., 1987).

The remarkably different ages of onset and clinical manifestations of the four affected homozygotes with ALAD deficient porphyria suggests the occurrence of different mutations responsible for mild to severe alterations of the enzyme protein. The availability of the full-length cDNA encoding human ALAD permits the identification of the molecular nature of these lesions. Two allelic mutations causing the severe

infantile-onset form of ALAD deficient porphyria, the first molecular lesions identified in the ALAD gene have been described. Discrimination of the mutant alleles was facilitated by a common *Rsa*I restriction fragment length polymorphism (RFLP) in the coding region of the gene. The maternal mutation (designated G133R) was a G to A transition of nucleotide 397 which predicted a glycine to arginine substitution at residue 133 at the carboxyl end of the highly conserved zinc-binding site in the enzyme subunit. The paternal mutation was a G to A transition of nucleotide 823 which predicted a valine to methionine substitution of residue 275 (designated V275M). The two missense mutations which occurred at CpG dinucleotides were confirmed in genomic DNA by competitive PCR for the paternal mutation and by the detection of a *Pst*I restriction site created by the maternal mutation. These mutations altered the structure of functional regions of the enzyme subunit.

Studies of the genetic regulation of ALAD in mice indicated that the enzymatic activity was under the control of at least two codominant alleles at a single genetic locus, the Lv or levulinate locus (Doyle and Schimke, 1969 and Coleman, 1966). Synthesis of ALAD was more than three time greater in mice homozygous for the Lv^a allele than those strains homozygous for the Lv^b allele. The heterozygotes (Lv^a/Lv^b) show an intermediate level of enzyme activity (Coleman, 1966). Physicokinetic studies revealed no differences between ALAD from the two alleles. Instead, it was shown that, although the rate of ALAD degradation in the two strains is identical, Lv^a mice have higher rates of ALAD synthesis than for Lv^b mice. Recently, quantitiative Southern hybridization with cloned ALAD probes has revealed gene dose differences among mouse strains, with mouse

strains carrying one (C57BL/6), two (C58) or three (DBA/2) copies of the murine ALAD gene per haploid genome (Bishop et al., 1986). The amount of ALAD is proportional to the amount of mRNA which, in turn, is proportional to the number of ALAD genes.

In the human population, ALAD showed a three to four fold range of enzyme activity, with a greater correlation between enzyme activities among sibs than nonsibs and with very similar enzymatic activities for monozygotic twins (Sassa, 1973). There being no evidence for variable ALAD gene dose in man, the lower enzymatic activities observed in some individuals are most likely due to polymorphic defects in the ALAD gene (Doss et al., 1986).

ALAD is inhibited by lead (see the following section) and other heavy metals. Other inhibitors of ALAD include levulinic acid (Beale, 1970) and succinylacetone (Tschudy et al., 1981), both analogues of ALA, and hemin (Weissberg et al., 1971). As a structural analogue of ALA, succinylacetone is a potent, competitive inhibitor of ALAD from bovine live or human erythrocytes. The K_i of succinylacetone for the bovine enzyme is 10^{-4} of the K_m of the enzyme for ALA (Sassa and Kappas, 1983). As demonstrated by Tschudy et al (1981), succinylacetone acts by binding to the enzyme at the site normally involved in the formation of a Schiff's base with ALA. Interestingly, patients with hereditary tyrosinemia, who excrete large quantities of succinylacetone, exhibit severe secondary ALAD deficiency, and, therefore, excrete large quantities of ALA into their urine (Tschudy et al., 1981). A recent study found that ALAD in the kidney had different pH optimums and kinetic properties than that of the liver enzyme. The authors speculate that the renal ALAD may be a different molecular species (Roth et al, 1990). However,

there is no molecular evidence to support this theory.

Insights into the nature of the active site and zinc-binding domains of the enzyme have been provided by biochemical studies of the purified enzyme (Anderson and Desnick, 1979; Tsukamoto et al., 1979 and 1980; Gibbs et al., 1985; Jaffe and Hanes, 1986) and by analysis of the predicted amino acid sequence from the recently isolated human ALAD full-length cDNA (Wetmur et al., 1986a and 1986b). Sulfhydryl groups are essential for activity and a lysinyl residue is located in the active site of the enzyme (Barnard et al., 1977; Tsukamoto et al., 1980; Jaffe et al., 1984; Jordan and Gibbs, 1985; Gibbs et al., 1985; Gibbs and Jordan, 1987; Jaffe and Markham, 1987; Jaffe et al., 1990). The active lysine, involved in the Schiff's base formation with ALA, has been identified at residue 252 (Gibbs and Jordan, 1986) located between two highly hydrophilic regions (Wetmur et al., 1986b)(Figure 4). The fifteen amino acid sulfhydryl-rich zinc-binding domain (residues 119-133) conforms to the consensus motif (Miller et al., 1985) and is remarkably conserved between species, being identical in man, rat and mouse (Bishop et al., 1986 and 1989; Wetmur et al., 1986b) and having only four amino acid differences with the *Escherichia coli* sequence (Echelard et al., 1988) (Figure 5).

Zinc atoms are essential for ALAD enzymatic activity and stability, and chelation by EDTA rapidly inactivates the enzyme. pH titrations of zinc, cadmium (Schlösser and Beyersmann, 1987; Farant and Wigfield, 1987) and lead (Farant and Wigfield, 1987; Tomokuni and Ichiba, 1988) forms of ALAD demonstrate that zinc atoms are near enough to the active site to affect ionization of essential amino acid residues (Figure 6). ALAD inhibition by lead is reversible *in vitro* by incubation with zinc ions and dithiothreitol

Catalytic site of ALA-D

Human	GGA	GCT	GAC	ATG	CTG	ATG	GTG	AAG	CCG	GGA	ATG	CCC	TAC	CTG	GAC	ATC	GTG	CGG	GAG
	Gly	Ala	Asp	Met	Leu	Met	Val	Lys	Pro	Gly	Met	Pro	Tyr	Leu	Asp	Ile	Val	Arg	Glu
Mouse	GGA	GCT	GAC	ATG	CTC	ATG	GTG	AAG	CCG	GGA	TTG	CCC	TAC	CTG	GAC	ATG	GTG	CGA	GAG
	Gly	Ala	Asp	Met	Leu	Met	Val	Lys	Pro	Gly	Leu	Pro	Tyr	Leu	Asp	Met	Val	Arg	Glu
Rat	GGA	GCT	GAC	ATA	CTC	ATG	GTA	AAG	CCG	GGA	TTA	CCC	TAC	CTG	GAT	ATG	GTG	CAG	GAG
	Gly	Ala	Asp	Ile	Leu	Met	Val	Lys	Pro	Gly	Leu	Pro	Tyr	Leu	Asp	Met	Val	Gln	Glu
E.coli	GGC	GCA	GAC	TGC	CTG	ATG	GTT	AAA	CCT	GCT	GGA	GCG	TAC	CTC	GAC	ATC	GTG	CGT	GAG
	Gly	Ala	Asp	Cys	Leu	Met	Val	Lys	Pro	Ala	Gly	Ala	Tyr	Leu	Asp	Ile	Val	Arg	Glu

Figure 4. Conservation of amino acid sequence of the catalytic site in human, mouse, rat and *E. coli*.

Zinc binding site of ALAD

A	TGT	GAT	GTC	TGC	CTG	TGT	CCC	TAC	ACC	TCC	CAT	GGT	CAC	TGC	GGG
	Cys	Asp	Val	Cys	Leu	Cys	Pro	Tyr	Thr	Ser	His	Gly	His	Cys	Gly
B	TGT	GAT	GTG	TGC	CTG	TGC	CCC	TAC	ACC	TCC	CAT	GGC	CAC	TGC	GGG
	Cys	Asp	Val	Cys	Leu	Cys	Pro	Tyr	Thr	Ser	His	Gly	His	Cys	Gly
C	TGT	GAC	GTC	TGC	TTG	TGC	CCC	TAC	ACC	TCC	CAT	GGC	CAC	TGC	GGC
	Cys	Asp	Val	Cys	Leu	Cys	Pro	Tyr	Thr	Ser	His	Gly	His	Cys	Gly
D	TAC	GAC	ACC	TGC	TTC	TGT	GAA	TAC	ACT	TCT	CAC	GGT	CAC	TGC	GGT
	Ser	Asp	Thr	Cys	Phe	Cys	Glu	Tyr	Thr	Ser	His	Gly	His	Cys	Gly

Figure 5. Conservation of amino acid sequence of the zinc-binding domain in human (A), mouse (B), rat (C) and *E. coli* (D).

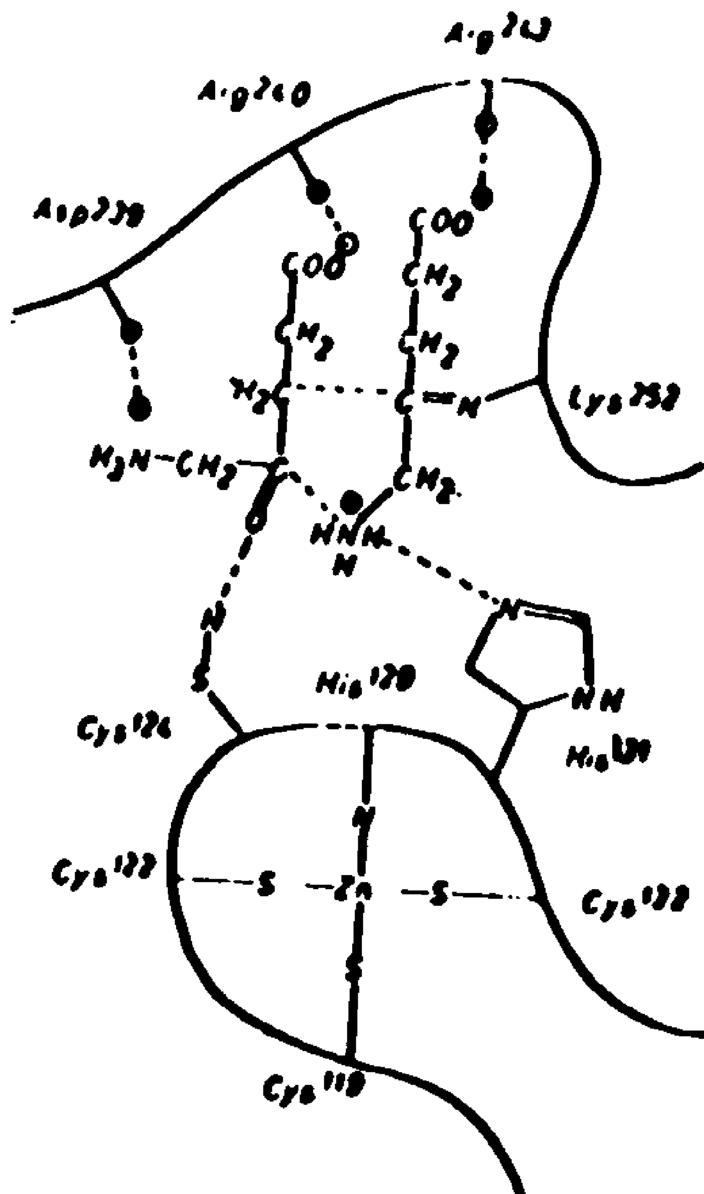


Figure 6. Complete active site model (Schlosser and Beyersmann, 1987).

(DTT) (Finelli et al., 1975; Geisse et al., 1983). The enzymes from bovine and human are maximally active in the presence of zinc and sulfhydryl compounds such as beta-mercaptoethanol, cysteine, reduced glutathione or dithiothreitol (DTT). Removal of zinc by EDTA or other chelators, or oxidation or inhibition of sulfhydryl groups, results in rapid loss of activity. Interestingly, under anaerobic conditions, removal of zinc does not cause loss of activity, implying that zinc may not be required for catalytic activity, but may act to protect essential sulfhydryl groups from oxidation. It has been suggested that four zinc ions are required to prevent oxidation of the sulfhydryl groups in the active site whereas the other four presumably bridge subunits to maintain the stability of the homooctameric structure (Jaffe et al., 1984; Dent et al., 1990).

Molecular studies have revealed interesting characteristics in the cDNA and genomic structure of ALAD. The ALAD cDNA clone was originally detected in an adult liver cDNA library. The library was made by dC tailing of cDNAs with terminal transferase and cloning into a *Pst*I cleaved and dG tailed pKT218 vector. The pKT218 clone was cleaved with *Pst*I and the two insert fragments were subcloned into a pUC9 vector and sequenced (Wetmur et al., 1986b). The ALAD cDNA clones contained an open reading frame of 990 bp which encoded 330 amino acids, a 68bp of upstream untranslated region and an additional 3′untranslated sequence preceding the poly A. The ATG initiation codon was preceded by ACGCC, a functional initiation sequence, while an upstream, in-phase AACTG ATG sequence was entirely non-homologous with the initiation consensus sequence (Kozak, 1984) and, therefore, is presumed to be non-functional. An unusual polyadenylation signal, AGTAAA, was observed which had been

reported only for human c-H-*ras*1 (Reddy, 1983). A cysteine- and histidine-rich domain was observed from amino acid 119 to 132. The distribution of these amino acids conformed with the known consensus sequence for zinc binding (Berg, 1986). A second domain contained several charged amino acids from 239-244 and 262-267 surrounding the active lysine^{252} residue in the catalytic site.

Recent studies in which PCR amplified transcripts from murine liver and DMSO-induced murine erythroleukemia cells revealed two classes of transcripts (Bishop, et al., 1991). The first transcript was found in both cell types and presumed to be the housekeeping transcript. The second transcript was found specifically in erythroid cells. The two transcripts resulted in differential splicing of two primary exons, designated exon 1A and 1B, to exon 2. The genomic sequence of exon 1A, whose 5′ untranslated region is that of the constitutive transcript, contained no TATA box and four GC boxes which may have potential SP1 binding capabilities. The erythroid specific, exon 1B was located 3′ to exon 1A and its genomic region contained a pair of CACCC boxes about a center of symmetry and three NF-E1 or GATA1 consensus binding sites.

E. <u>Lead toxicity and ALAD</u>

It is well known that lead is a potent inhibitor of ALAD activity (for reviews, see Bottomley and Muller-Eberhard, 1988; Kappas et al., 1989). Although lead itself is neurotoxic *in vitro* (Audesirk, 1985), the role of ALA as a neurotoxic agent also has been supported by a variety of studies (for review, see Kappas et al., 1989; Cutler et al., 1985). In fact, the stoichiometric inhibition of human erythrocytic ALAD activity has been used as a sensitive diagnostic indicator of lead exposure. For example, blood lead levels of

15 and 30 μg per deciliter result in 50% and 75% inactivation of ALAD activity, respectively. The lead cations replace the zinc cations in the active site, and this inhibition is reversible *in vitro* by incubation with zinc cations and dithiothreitol (DTT) (Geisse et al., 1983). Inactivation of ALAD activity by lead results in a proportional accumulation of ALA in blood and urine. In fact, several lines of evidence have suggested that ALA is the neuropathologic agent in lead poisoning. First, a variety of *in vitro* studies have demonstrated the neurotoxicity of ALA (Muller and Snyder, 1977; Brennan and Cantrill, 1979 and 1981; Audesirk, 1985; Cutler et al., 1985; Minnema and Michaelson, 1986; Kappas et al., 1989). Second, the clinical manifestations of lead poisoning closely resemble those of the acute neurologic attacks in the hepatic porphyrias, during which the levels of ALA and porphobilinogen are significantly elevated. Third, affected homozygotes with the ALAD deficient porphyria have elevated levels of ALA and chronic neurologic manifestations (Doss et al., 1979). Finally, the role of the ALA accumulation in lead poisoning is supported by the finding that asymptomatic heterozygotes for the ALAD deficient porphyria are prone to acute lead poisoning when exposed to low levels of lead (Doss et al., 1982 and 1984).

Exposure to environmental lead may affect numerous organ systems, including the renal, reticuloendothelial, nervous and reproductive systems. Although the reduced use of leaded gasoline, widespread detection of lead-paint poisoning of children and control of lead exposure in the workplace have combined to substantially reduce the incidence of acute lead poisoning in the United States, exposure to low doses of lead is still quite common. The Second National Health and Nutrition Examination Survey (NHANES-II)

found that 1.5 million preschool children had blood lead levels of 25 µg/dl and above, indicating significant lead absorption.

The best documented long-term consequence of exposure to lead at blood lead levels *below* 40 µg/ml is irreversible neurological damage. In the peripheral nervous system, the outcome measure of ulnar nerve conduction velocity was slowed at blood lead levels of 30-40 µg/dl (Seppäläinen et al., 1983). In the central nervous system, neurobehavioral deficits have been detected in children who experienced blood lead levels at or even below 30 µg/dl (for review, see Lippmann, 1990). Needleman et al. (1979 and 1990) studied neurobehavioral outcomes over 11 years in children divided into groups based on dentine lead levels in deciduous teeth, an integrated measure of lead exposure. The mean blood lead level experienced by the high dentine lead group was 34 µg/dl. Early on, compared to the low lead group, this group showed a 4 point IQ decrement. On follow-up, the high lead group was more likely to exhibit reading disability and to drop out of high school. Neurobehavioral deficits resulting from effects of low level lead exposure, as measured by blood lead as well as dentine lead, have been confirmed in European studies (Winneke et al., 1983; Bergomi et al., 1989; Wineeke et al., 1990).

Prenatal exposure to lead is also significant. Both decreased IQ and shortened attention span have been reported in young children with an average umbilical cord blood lead of only 15 µg/dl compared to another group with 6.5 µg/dl (Bellinger et al., 1987). Lead is mobilized from maternal bone stores during the last trimester of pregnancy and readily crosses the placenta, thereby placing the fetus at risk. McMichael et al., (1988) have measured neurobehavioral outcomes at age 4 in children of lead workers at Port

Pirie, Australia, followed for both prenatal and postnatal exposure to lead, with mean blood lead values of 9.1 μg/dl in mid-pregnancy to 21.2 μg/dl at 2 years. Surprisingly and of great concern was the finding that *no* threshold dose was observed for the effect of lead on these children.

Lead may be teratogenic. Increased placental lead concentrations have been correlated with frequency of stillborns and with mental retardation and congenital anomalies (Wibberley et al., 1977). A multivariant analysis of over 5000 births with 3% congenital malformations demonstrated a significant dose-related correlation between placental lead concentration and congenital anomalies (Needleman et al., 1984). Increased numbers of chromosomal aberrations have been observed in experimental animals exposed to lead and in lead workers (Deknudt et al., 1977).

Human ALAD was shown to be a polymorphic enzyme in 1981 (Battistuzzi et al., 1981). Expression of the two common alleles, designated $ALAD^1$ ($q = 0.9$) and $ALAD^2$ ($q = 0.1$), results in three distinct charge isozyme phenotypes, ALAD 1-1, 1-2, and 2-2. These isozymes, encoded by a single gene localized to chromosome 9q34 (Potluri et al., 1987), are readily separated by starch or cellulose acetate gel electrophoresis (Battistuzzi et al., 1981; Astrin et al., 1987). In a study of 1074 individuals of various ethnic and demographic backgrounds from New York City, ALAD isozyme phenotypes in erythrocytes were determined. The population was 87% Caucasian, 10% American Black, 2% Hispanic and 0.5% Asian. The $ALAD^1$ allele frequency in the total population was 0.89, a frequency similar to that (0.90) reported in a large Italian population (Petrucci et al., 1982; Battistuzzi et al., 1981). It was notable that the $ALAD^2$ allele frequency in the

115 American Black individuals studied was 0.03, whereas those individuals of other ethnic and demographic groups ranged from 0.08 to 0.13. This finding was consistent with the reported absence of the $ALAD^2$ allele in a Black population from Liberia (Benkmann et al., 1983).

The existence of this common polymorphism and the fact that ALAD is markedly inhibited by lead suggested a possible physiologic relationship between the ALAD isozymes and lead poisoning. Although no difference in the activities of the erythrocytic ALAD 1-1, 1-2 or 2-2 isozymes was demonstrated (Battistuzzi et al., 1981; Ostasiewicz and Desnick, unpublished results), individuals with the $ALAD^2$ allele could be more susceptible to lead exposure if the ALAD 2 subunit binds lead more tightly than the ALAD 1 subunit. Thus, $ALAD^2$ heterozygotes and homozygotes might have higher blood and bone lead concentrations as well as higher total body lead stores, making them even more likely to express subclinical and clinical manifestations of chronic low-level (Bellinger et al., 1987; Needleman et al., 1990) or acute lead exposure.

Previous studies of ALAD isozyme phenotypes and blood lead levels only have been reported from two laboratories. The ALAD isozyme types of 1278 blood samples were determined from children who were subjected to low level environmental lead exposure in New York City (Astrin et al., 1987). The blood lead level, measured by the New York City Lead Screening Program, and the ALAD isozyme phenotype were determined in a double-blind fashion (Figure 7). The second population was derived from lead workers in a factory in Germany (Ziemsen et al., 1986) (Figure 8). In a combined analysis of the relationship between ALAD genotypes and blood lead levels in

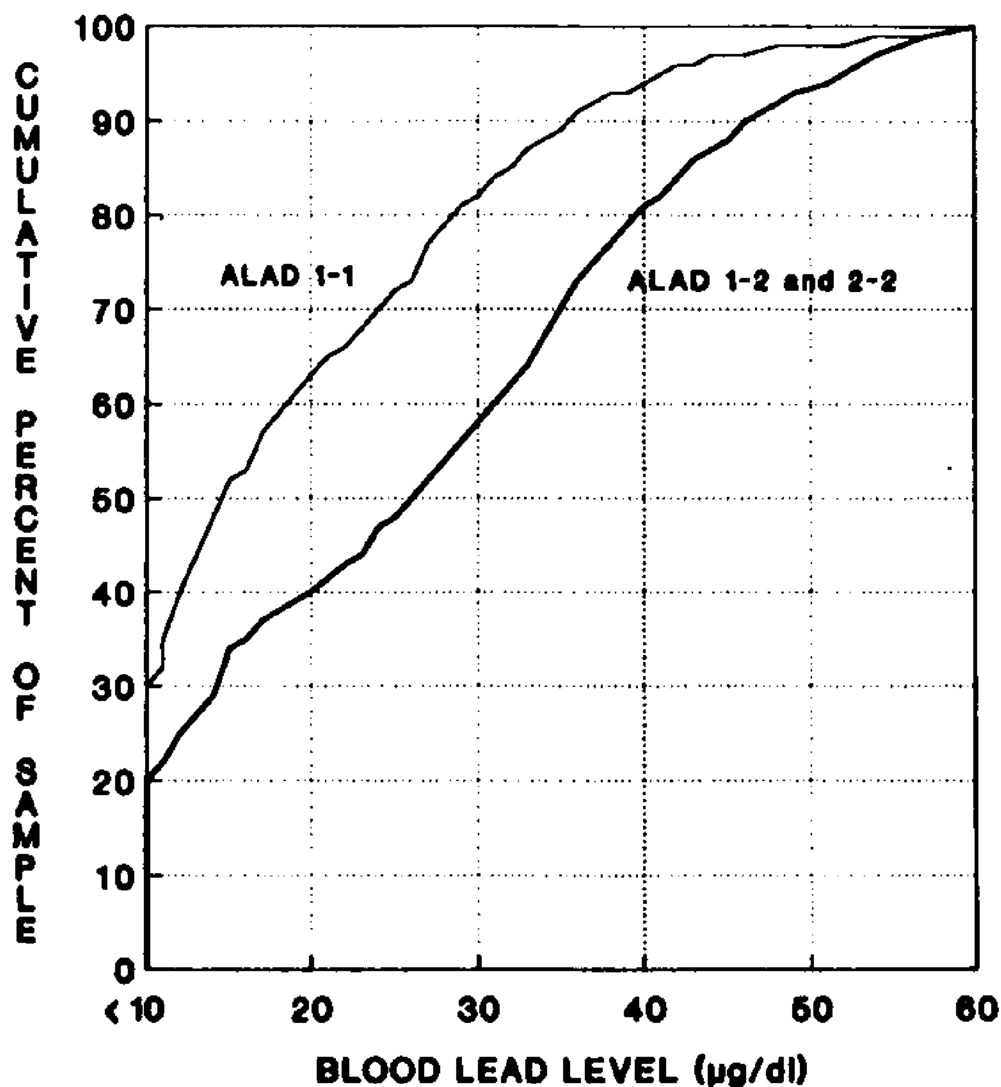


Figure 7. The blood lead level, measured by the New York City Lead Screening Program, and the ALAD isozyme phenotype.

LEAD WORKERS (FRG)

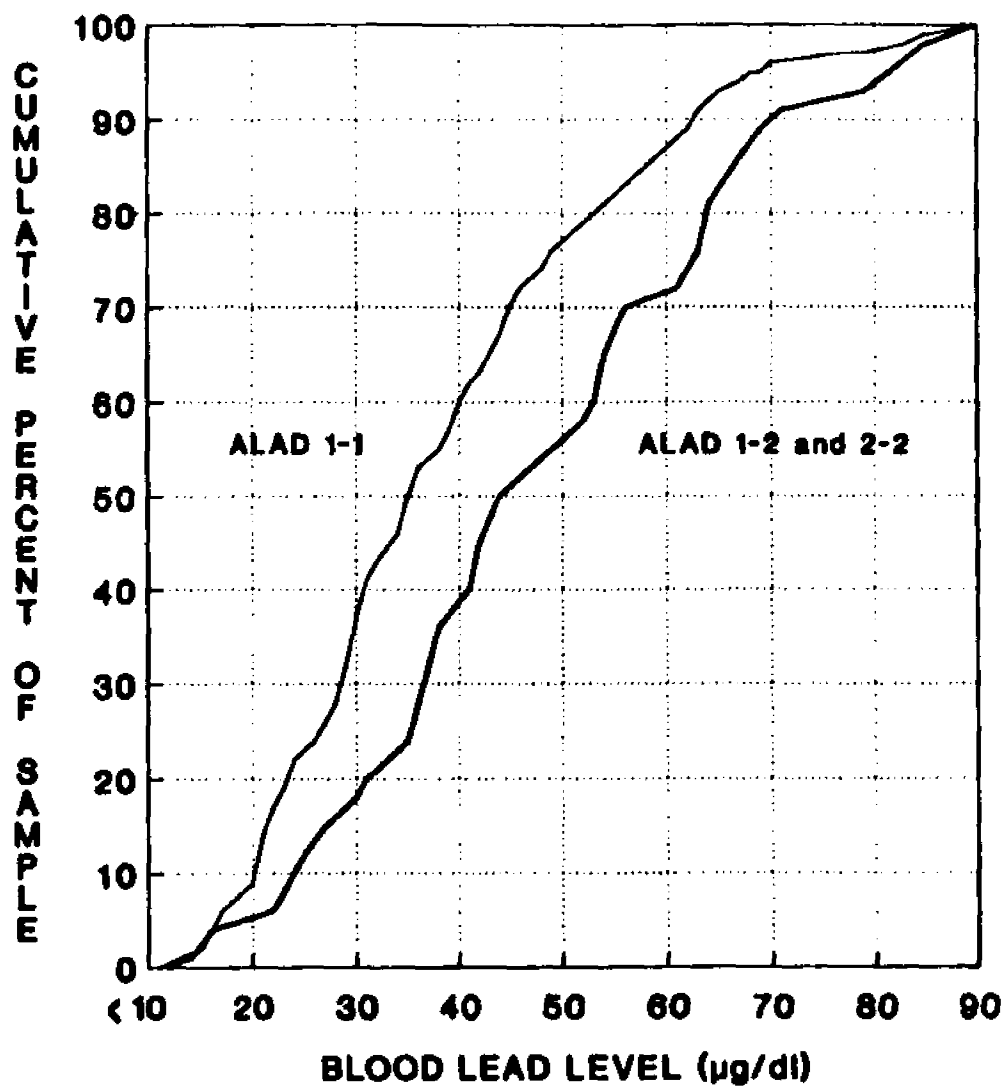


Figure 8. The blood lead level derived from lead workers in a factory in Germany and the ALAD isozyme phenotype.

equivalently exposed individuals (Wetmur et al., in press) one population, designated USA below, consisted of children in New York City whose free erythrocyte protoporphyrin (FEP) levels (Piomelli et al., 1982; Kaul et al., 1983) indicated possible environmental lead exposure. The second population, designated FRG, consisted of German lead factory workers.

The two populations of lead exposed individuals were subdivided into two groups, one homozygous for the $ALAD^1$ allele (phenotype ALAD 1-1) and the other carrying the $ALAD^2$ allele (phenotypes 1-2 and 2-2). The ALAD 1-2 or ALAD 2-2 group was not separated into ALAD 1-2 and ALAD 2-2 subgroups because the small number of ALAD 2-2 individuals. Table 1 shows that for both populations, the groups carrying the $ALAD^2$ allele have increased blood lead levels. Using the t-test, the means of the groups were different at the level of $t = -7.0$, $df = 1276$, $P > 0.0001$ and $t = -2.9$, $df = 200$, $P > 0.004$ in the USA and FRG populations, respectively (SAS Institute, 1985). Table 2 shows that the presence of the $ALAD^2$ allele was associated with an increase in median blood lead level of about 10 µg/dl for individuals in both populations. A two way analysis of variance showed the differences between the two groups to be highly significant ($p < 0.0001$). There was no interaction ($p = .68$), indicating that the difference in the two populations was comparable even though the measurements themselves were different (Wallenstein et al., 1980), consistent with a common mechanism being responsible for the differences between isozyme phenotype groups.

It should be noted that the $ALAD^2$ allele is less frequent in non-Caucasian populations with no $ALAD^2$ alleles found in Liberian Blacks (Benkmann et al., 1983).

TABLE 1

Population	Phenotype	N	Blood Lead (μg/dl) Mean ± SD	Blood Lead (μg/dl) Mean ± SEM
USA	1-1	1136	19.5 ± 11.6	19.5 ± 0.3
	1-2 + 2-2	142	27.1 ± 15.2	27.1 ± 1.3
	Significance: P > 0.0001			
FRG	1-1	160	38.4 ± 16.8	38.4 ± 1.3
	1-2 + 2-2	42	47.0 ± 18.0	47.0 ± 2.8
	Significance: P > 0.004			

TABLE 2

Population	Phenotype	N	Blood Lead (μg/dl) Median	Difference
USA	1-1	1136	15	11
	1-2 + 2-2	142	26	
FRG	1-1	160	35	9
	1-2 + 2-2	42	44	
	No Interaction:		P = 0.68	

In some cases, the ethnic group of the NYC population was known. The USA population had 11% rather than 19% of individuals with the $ALAD^2$ allele because many of the children were Hispanic or American Blacks ($ALAD^2$ allele frequency around 5%). The average blood lead level of American Black children within this population was marginally higher than the blood lead level of the overall population. Thus, exclusion of the American Black children from the analysis would have increased the observed difference in blood lead levels between the two ALAD isozyme phenotype groups, thereby strengthening these conclusions.

In summary, these results indicated that in two very different populations, the presence of the $ALAD^2$ allele, found in approximately 19% of the Caucasian population, led to accumulation of an additional 10 μg/dl blood lead in lead exposed individuals. These findings are consistent with the hypothesis that the $ALAD^2$ gene product binds lead more effectively, and therefore that individuals with the $ALAD^2$ allele may be more susceptible to lead poisoning. This 10 μg/dl represents an additional 40% above the currently permissible blood lead measure of exposure (25 μg/dl).

II. Objectives

δ-aminolevulinic acid dehydratase is the second enzyme in the biosynthetic pathway of heme, a molecule essential for the function of all aerobic cells. Therefore, ALAD is constitutively expressed in all cells, as are all the enzymes of the heme biosynthetic pathway. ALAD is a homooctomer which requires at least four zinc atoms to function and is inhibited by lead and other heavy metals. In differentiating erythroid cells, the enzyme is expressed at high levels, and ALAD in red cells is capable of retaining significant quantities of lead. There are two polymorphic alleles, $ALAD^1$ and $ALAD^2$, which account for three distinct isozyme forms which can be separated by charge. The $ALAD^2$ allele is associated with having a higher susceptibility to lead poisoning. The neurological symptoms of lead poisoning are not unlike those of ALAD deficient porphyria. The overall objective of proposed research is to study the organization and expression of the ALAD gene, the molecular nature of the $ALAD^2$ allele and the properties of recombinant ALAD 1-1 and ALAD 2-2 as they relate to lead poisoning. The specific aims of the research are:

1. Characterization of the entire genomic sequence of ALAD, including intron/exon boundaries, identification of repetitive sites such as Alu sequences, and potential promoter elements. The promoter will be examined by chloramphenicol acetyltransferase (CAT) expression and deletion analysis in non-erythroid and an inducible erythropoietic cell line. Characteristics of ALAD which are derived from these studies will be compared with those of the other enzymes of the heme

biosynthetic pathway.

2. The $ALAD^2$ allele will be characterized on a molecular level. The resultant mutation will be confirmed as the $ALAD^2$ allele in a population study, where ALAD 1-1, 1-2 and 2-2 individuals will first be determined by electrophoretic mobility.

3. The common ALAD 1-1 and the rare ALAD 2-2 polymorphic gene products will be expressed and purified as full length, non-fusion proteins in *E. coli*. A purification scheme will be devised so as to obtain large quantities of both the ALAD 1-1 and ALAD 2-2 isozymes. Purified recombinant enzymes may then be used for quantitative binding and catalytic studies as well as for collaborative studies including x-ray crystallography.

III. Materials

A. Table 3: Chemicals and Biochemicals

Item	Company	Location
Zetabind nylon membrane	AMF Cuno Microfiltration Products Division	Meriden, CT
5-bromo-4-chloro-3-indoyl-ß-D-galactoside	Aldrich	Milwaukee, WI
Centricon ultrafilters	Amicon Corp.	Lexington, MA
A,G,C and T-CPG 0.2 μm columns Reagent kits for DNA synthesizer	Applied Biosystems	Foster City, CA
Whatman 3MM sheets DEAE-Cellulose	Baxter	Edison, NJ
Acrylamide 99% Agarose Biotinylated SDS-PAGE M_r markers Gel filtration standards HRP color development reagent Silver stain kit Urea	Bio-Rad	Melville, NY
cDNA synthesis kit Formamide RNA markers	BRL-Life Technologies, Inc.	Gaithersburg, MD
Anhydrous acetonitrile Bis-tris Cellogel sheets Chromerge N,N-dimethyl formamide Hydrolyzed starch Glacial acetic acid Methanol	Fisher Scientific	Pittsburgh, PA

Kapak, Scotchpak bags Phenol, saturated Isopropyl-B-D-thiogalacto-pyranoside	Jersey Lab and Glove	Livingston, NJ
T4 polynucleotide kinase T4 DNA Ligase Restriction enzymes	New England Biolabs	Beverly, MA
Gene-Amp kit Amplitaq	Perkin Elmer Cetus	Emeryville, CA
Dextran sulphate Poly(dI-dC) Sephacyrl S-300	Pharmacia LKB Biotechnology	Piscataway, NJ
Restriction enzymes RNasin RQ DNase I SP1 human transcription factor Taq polymerase	Promega	Madison, WI
Aminolevulinic acid Ampicillin (anhydrous) Chloramphenicol Dimethylaminobenzaldehyde Dithiothreitol Glass beads, acid washed Glycine Isopropanol Lysozyme Mercuric chloride S1 nuclease	Sigma	St. Louis, MO
PrimeErase Quik columns Strataclean resin	Stratagene	La Jolla, CA
Sequenase DNA polymerase Sequenase sequencing kit	United States Biochemical	Cleveland, OH
Vectastain ABC kit (Rabbit IgG)	Vector Labs	Burlingame, CA

B. Vectors

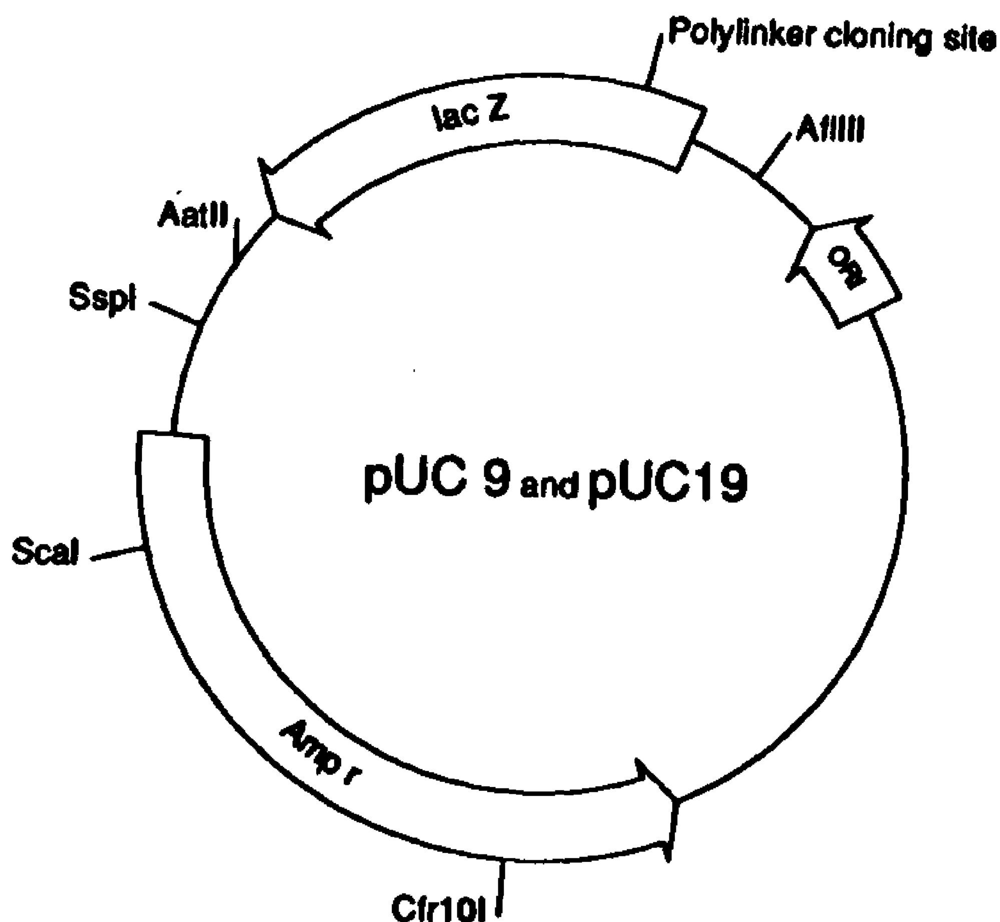


Figure 9. pUC9 and pUC19, *E. coli* plasmids approximately 2.7 kb in length. They are high copy number plasmids and part of a series of related plasmids constructed by Messing et al. (1985). The polylinker cloning sites are located in the *lac*Z gene, clockwise from an *Eco*RI site to a *Hin*dIII site and spanning 36 and 54 base pairs, respectively. pUC9 was used for the original, in-frame, cloning of the ALAD cDNA, for formation of a *lac*Z-ALAD fusion protein. pUC19 was used for most of the cloning of the genomic ALAD fragments and various other cloning projects.

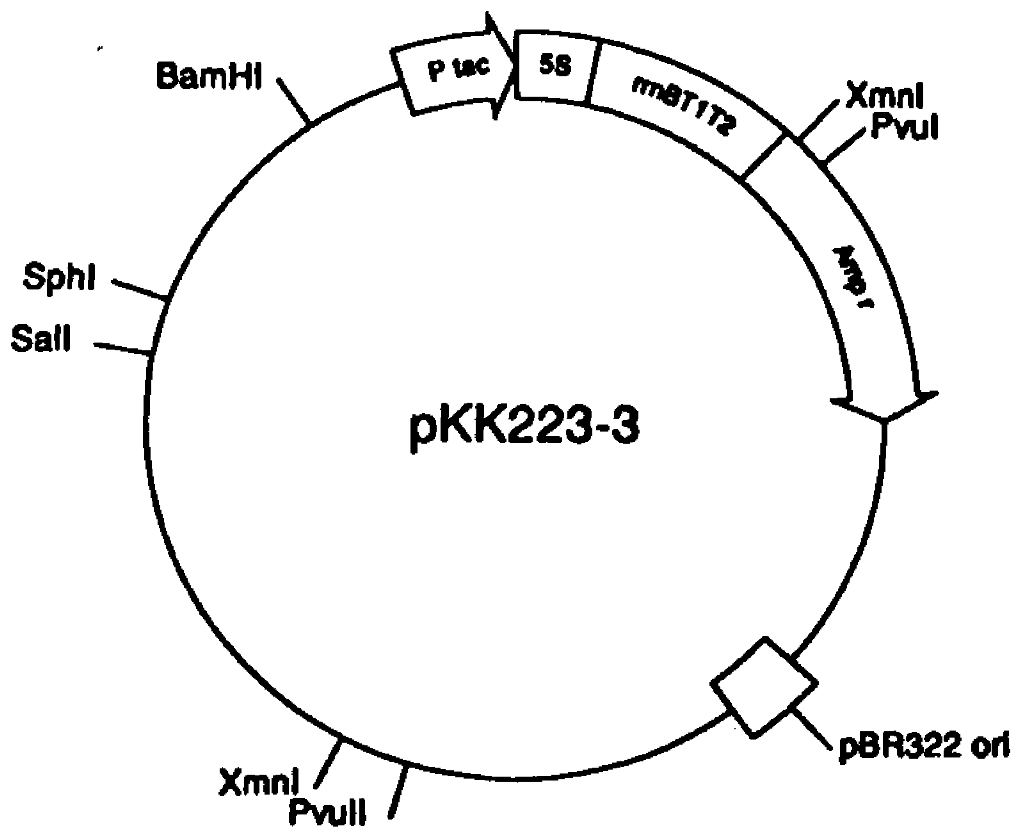


Figure 10. pKK223-3 is a prokaryotic expression vector developed by Brosius and Holy (1984). The vector is approximately 4.6 kb in length. It contains the *tac* promoter and can be induced by the addition of IPTG. Following the *tac* promoter is a multiple cloning site and the strong *rrn*B ribosomal terminator.

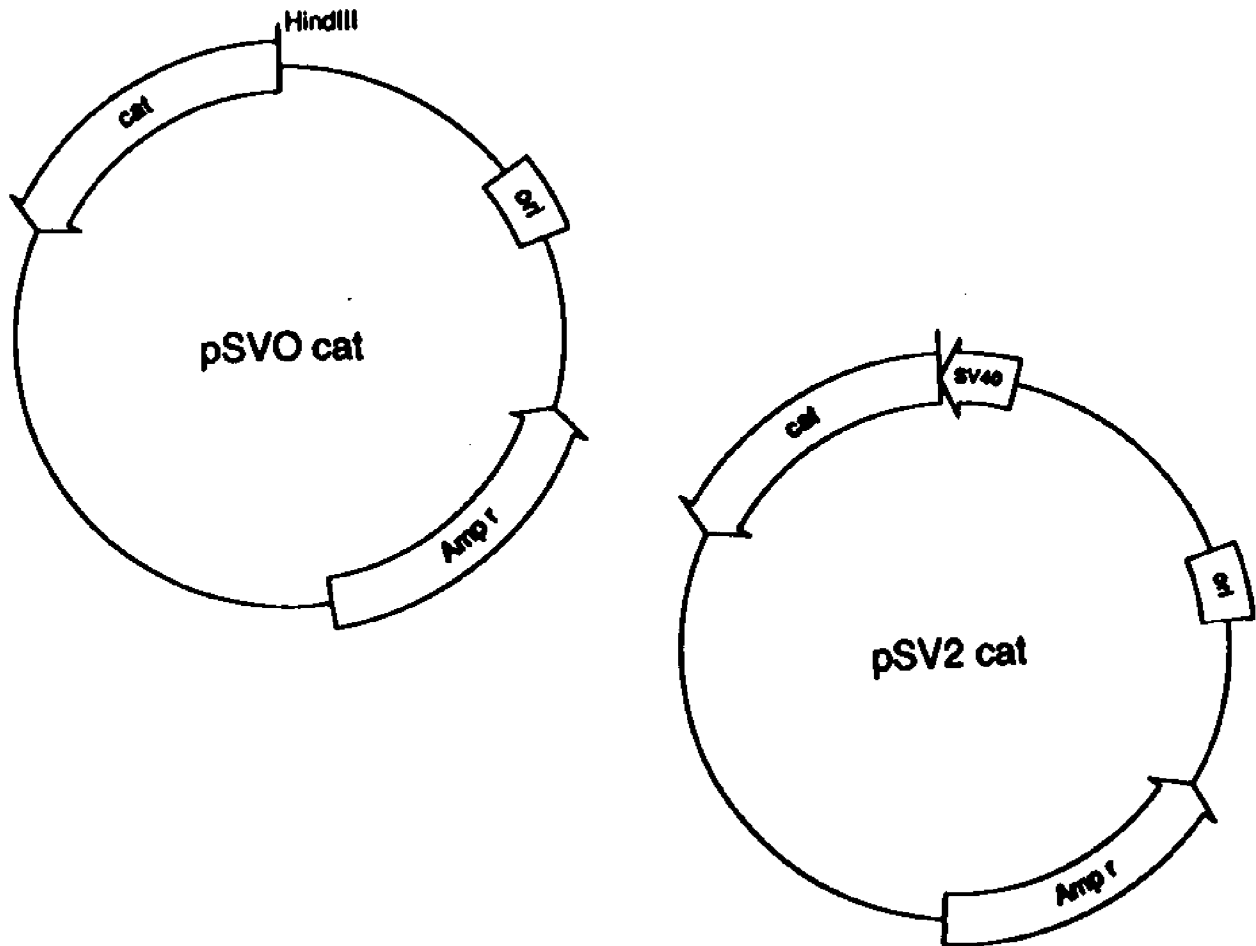


Figure 11. pSV0 and pSV2 cat or CAT vectors were used for the analysis of eukaryotic promoters and enhancers. Both are approximately 5.0 kb in length and are assembled from elements of pBR322, SV40 and the CAT cistron. pSV0 does not contain the SV40 promoter region, but does contain a *Hin*dIII cloning site for insertion of promoter or enhancer regions. pSV2 contains the SV40 early promoter and was used as a control.

C. Cell lines

BL21 - (*hsd*S, *gal* [λcIts857, *ind*1, *S* am7, *nin*5, *lac*UV5-T7 gene*1*]) - a B strain of *E. coli*, was used for expression because of the lack of the *lon* protease gene (Studier and Moffatt, 1986). The BL21 (F^-, *omp*T $r_B^- m_B^-$)strain lacks the *omp*T outer membrane protease which has been shown to degrade proteins during purification (Grodberg and Dunn, 1988). ALAD was found to be more stable in this strain than in the *E. coli* K12 strains described above. This cell line contained the plasmid pLysE - Plasmid which contains the gene for T7 lysozyme, a bifunctional protein used in T7 expression systems to bind T7 RNA polymerase and inhibit transcription (Moffatt and Studier, 1987). The second function of T7 lysozyme is its ability to cut a specific bond in the peptidoglycan layer of the *E. coli* cell wall (Inouye et al., 1973). This second function allowed for a simplified lysis procedure. ALAD expression plasmids were transformed into cells already containing the pLysE plasmid and cells were grown in the presence of two antibiotics, chloramphenicol and ampicillin, except where noted, to maintain both plasmids.

DH5-α - (*sup*E44, Δ-*lac*U169(80*lac*ZΔM15), *hsd*R17, *rec*A1, *end*A1, *gyr*A96, *thi*-1, *rel*A1) - was used primarily for propagation of plasmids. It was used in initial expression studies but abandoned due to lack of productivity of active ALAD.

*hem*B - Cosloy and coworkers (Li et al., 1988) isolated a *hem*B mutant of *E. coli* which

failed to synthesize ALAD and was dependent on exogenously added heme for growth. The pUC9 expression vector was transferred into these mutant cells. The pUC9 (pUC9-50) vector produced the fusion protein this strain, which was then selected for transformants using ampicillin in the absence of heme. All of the transformed bacteria carried the vector and no longer required heme for growth. Thus, the human ALAD fusion protein was biologically active in *E. coli*.

JM103 - (*thi*, *str*A, *sup*E, *end*A, *sbc*B, *hsd*R$^-$, F′ *tra*D36, *pro*AB, *lac*I^q Z ΔM15,Δ(*lac*,*pro*)) - was originally used in the cloning of the ALAD cDNA.

LC137 - (*htp*R^{165}, *lon*^{R9}TS, *lac*, *trp*, *pho*, *rps*L, *sup*C^{TS}, *mal*, *tsx*:Tn*10*) - was used for expression studies only. Plasmid PC-3 was the only expression vector transformed into this strain. The htpR mutant cells are defective in proteolysis (Baker et al., 1984)

All cells were grown in LB media or on LB agar plates according to Sambrook et al (1989). Antibiotics were added to the following concentrations: ampicillin, 50 μg/ml; kanamycin, 25 μg/ml; chloramphenicol, 50 μg/ml and tetracycline, 5 μg/ml. When isopropyl ß-D-thiogalactopyranoside (IPTG) was used for induction of expression, 1 mM final concentration was used. When IPTG and 5-bromo-4-chloro-3-indolyl-ß-D-galactoside (X-gal) were used to distinguish between recombinant plasmid and parental plasmid, concentrations were used in accordance with Sambrook et al (1989).

IV. Methods

A. Oligonucleotide synthesis

Sense and antisense oligonucleotide primers for sequencing, primer extension, or polymerase chain reaction were synthesized on a model 380B DNA synthesizer (Applied Biosystems) by standard phosphoramidite chemistry.

B. Cloning

Growth of bacterial cultures, harvesting and lysis of bacteria, purification of plasmid DNA, ligation reactions, preparation and transformation of competent *E. coli*, screening by hybridization (Southern and Northern analysis), digestion of DNA with restriction enzymes, labeling with polynucleotide kinase and other standard molecular biological techniques were performed according to Sambrook et al. (1989).

C. Sequencing

Double-stranded plasmid template was isolated and sequenced with ALAD specific (Wetmur et al., 1986b) or vector specific oligonucleotides by the dideoxy chain termination method (Sanger, 1981) using the Sequenase DNA Sequencing Kit.

D. RNA isolation from cultured cells

Total RNA was isolated by the method of Chirgwin et al. (1979) with the following modifications from 10^6 cultured lymphoblasts from a normal individual who had the ALAD 2-2 isozyme phenotype. The guanidinium thiocyanate solution contained 25 mM sodium thiocyanate; the Antifoam A and diethylpyrocarbonate in the CsCl solution were reduced to 0.03% and 0.02%, respectively.

E. Primer extension

Oligo-dT primed first-strand cDNA was reverse-transcribed from ~10 μg of total RNA with the BRL cDNA Synthesis Kit (Bethesda Research Laboratories). The entire ALAD[2] cDNA was amplified by the polymerase chain reaction (PCR) using sense (5′-CAGAGCG**AATTC**CCAATGCCCCAGGAGCCC-3′) and antisense (5′-GTTCTA**AAGCTT**GGGCCTGGCACTGTCTCC-3′) oligonucleotide primers synthesized on an Applied Biosystems Model 380B DNA synthesizer. These primers were designed to amplify the region from nucleotide -44 to 994 of the ALAD cDNA which included the entire coding sequence; the primers also contained *Eco*RI or *Hin*dIII sites (boldface) for subcloning into plasmids for sequencing.

F. Polymerase Chain Reaction (PCR)

The PCR reaction was performed in a total volume of 100 μl containing 1 μM of each primer and 10 mM Tris-HCl buffer, pH 8.3, containing 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, and 200 μM of each dNTP (Saiki et al., 1988). The mixture was incubated at 100°C for 5 min to denature the DNA and to inactivate proteases, then 2.5 units of *Taq* polymerase were added. Amplification was performed on a Perkin-Elmer Cetus Thermal Cycler; each of the cycles consisted of denaturation at 94°C for 1 min, primer annealing at 53°C for 1 min, and extension at 72°C for 3 min.

G. DNaseI protection

Probe preparation involved cleaving the plasmid DNA with appropriate restriction endonucleases and gel purification of the DNA band containing the ALAD promoter region. The gel-purified fragment was dephosphorylated with alkaline phosphatase and

labeled at both ends using T4 polynucleotide kinase. A probe labeled at one end was generated by removal of the other end with a specific restriction endonuclease. This terminally-labeled fragment was incubated in the presence of SP1 in a binding buffer consisting of 50 mM Tris-HCl, pH 8.0, 0.1 M KCl, 12.5 mM $MgCl_2$, 1 mM EDTA, 20% glycerol, and 1 mM DTT. Incubation was on ice for 10 minutes. $CaCl_2$ and $MgCl_2$ at room temperature were added to final concentrations of 2.5 mM and 5 mM, respectively. Following incubation at room temperature for 1 minute, 0.6U of RQ1 RNase-free DNase was added. The mixture was mixed thoroughly and incubated at room temperature for an additional 1 minute. The reaction was terminated by the addition of 0.2 M NaCl, 0.03 M EDTA, 1% SDS and 100 μg/ml yeast RNA. The reaction mixture was then phenol extracted (phenol: chloroform:isoamyl alcohol), ethanol precipitated and vacuum dried. The pellet was resuspended in loading dye [0.1 M NaOH:formamide (1:2 v/v), 0.1% xylene cyanol, 0.1% bromophenol blue], heated to 95°C for 2 minutes and loaded onto a 8% polyacrylamide sequencing gel. The gel was run at 50 mAmp until the bromophenol blue was at the bottom of the gel and then dried under a vacuum. The gel was exposed on film for 19-72 hours with an intensifying screen at -70°C.

H. <u>CAT expression</u>

PCR primers were constructed to amplify portions of the ALAD promoter, not including the first splice site. These oligonucleotides contained a *Hin*dIII restriction site to facilitate cloning into the pSV0 vector. Constructs were made by cleaving the PCR products with *Hin*dIII, ligating into the *Hin*dIII digested pSV0 and transforming into DH5-α. Colonies were chosen on the basis of positive hybridization with ALAD specific

probes. Plasmid preparation was performed by standard methods, and isolated DNA was used for the electroporation of COS cells. Standard CAT transient assays were performed by the method of Gorman (1982).

I. Northern analysis

Isolated RNA was run on a formaldehyde (6.7%)/agarose (1.5%) gel in preparation for Northern analysis. The gel was run in a 20 mM sodium phosphate buffer, pH 7.0. The samples were prepared in the same buffer with 3.5 µl formaldehyde and 10 µl of formamide in a total volume of 20 µl. Samples were incubated at 65°C for 15 minutes, then 2 µl of loading buffer (50% glycerol, 1 mM EDTA, 0.25% bromophenol blue, 0.25% xylene cyanol FF) were added to each sample before loading on the gel. RNA was transferred to nitrocellulose filters for standard Northern analysis (Sambrook, 1989). The filter was washed and prehyridized in 50% formamide, 5x SSC, 2X Denhardt's reagent and 0.1% SDS for 4-16 hours. Radiolabeled ALAD cDNA was added directly to the prehybridization solution while incubation continued for another 4-16 hours. Filters were washed at room temperature in 5x SSC, 0.1% SDS for 20 minutes, 2x SSC, 0.1% SDS for 20 minutes, 1x SSC, 0.1% SDS for 20 minutes, and followed by 3 washes of 20 minutes at 68°C in 0.1x SSC and 0.1% SDS. Filters were exposed on film for 24-48 hours with an intensifying screen at -70°C.

J. Preparation of genomic DNA from whole blood

Genomic DNA was isolated from an aliquot of heparinized blood by the rapid method described by Higuchi (1989). For determination of ALAD activity and isozyme phenotype, heparinized whole blood was centrifuged at 500 g for 30 min at 4°C, and the

plasma and buffy coat were removed. The erythrocytes were washed twice with an equal volume of 0.9% sodium chloride, followed by centrifugation at 500 g for 30 min at 4°C. For enzyme assay, the washed, pelleted erythrocytes were lysed by the addition of 1 mM potassium phosphate buffer, pH 6.8, containing 1 mM $MgCl_2$, 1 mM DTT and 0.5% Triton X-100. The mixture was vortexed for 2 min and centrifuged at 15,000 g for 40 min, and the supernatant (lysate) was used to determine ALAD activity (Anderson and Desnick, 1979).

K. Starch gel electrophoresis

For determination of the isozyme phenotype, the washed, pelleted erythrocytes were lysed with an equal volume of 0.25 mM zinc acetate and 20 mM DTT and equalized for lysate protein. The ALAD isozyme phenotypes were determined by starch gel electrophoresis as previously described (Battistuzzi et al., 1981; Petrucci et al., 1982; Astrin et al., 1987). Briefly, the starch gels were 10% with a gel buffer of 0.01 M sodium phosphate, pH 6.8. The bridge buffer was 0.1 M sodium phosphate, pH 6.8. Electrophoresis was carried out at 5 V/cm for 18 hr at 4°C. The gel was stained first with ALA and DTT and then with modified Ehrlich's reagent (Battistuzzi et al., 1981; Astrin et al. 1987).

L. Cellulose acetate gel electrophoresis

Lysates also were analyzed by cellulose acetate gel electrophoresis, which improved the resolution of the isozyme phenotypes. Lysates (0.5 to 4 μl) were applied to the gel and subjected to electrophoresis for 2 hr at 200 V and 4° C with a running buffer of 0.03 M phosphate, pH 7.0. The isozymes were detected by an immunodetection method

described elsewhere.

M. Bacterial lysis

Bacterial cells were inoculated into LB medium and grown with shaking overnight at 37°C for 16 - 20 hours. For maximal expression of ALAD in BL21, 1 mM IPTG was added to the inoculum. LC137 cells were grown at room temperature with shaking in the presence of 1 mM IPTG. All procedures thereafter were performed at 4°C or on ice except where indicated.

Cells were harvested by centrifugation in a Beckman J21C Centrifuge at 6,000 rpm for 30 minutes at 4°C. Pelleted cells were resuspended in lysis buffer which consisted of 0.05 M Bis-Tris, pH 6.8, 0.233 M NaCl, 2 mM EDTA, 1 mM dithiothreitol (DTT), 0.2 mM $ZnCl_2$ and to 0.2 mM phenylmethylsulphonylflouride (PMSF), using 4 ml of lysis buffer per gram of wet cells. Lysozyme was added to 5 mg/ml, and the suspension was incubated on ice for 30 minutes. Sodium deoxycholate (Harris et al., 1986) was then added to a final concentration of 0.5%, and the suspension was incubated at 37°C for 30 minutes with occasional stirring with a glass rod. After the 30 minute incubation, the lysis was complete and the solution was extremely viscous. In order to eliminate the nucleic acids, RQ1 DNaseI was added and incubated at room temperature for 30 minutes. DNaseI was needed in milligram quantities to reduce the concentration of nucleic acids. Therefore a second technique was used. Poly(ethyleneimine) treatment at 20 µl/ml efficiently removed nucleic acids from samples without loss of ALAD enzymatic activity (Zucker et al., 1985). The lysis suspension was centrifuged at 6,000 rpm for 30 minutes at 4°C, and the supernatant was recovered.

Bacteria from a one liter culture were sedimented and washed in 0.07 M Bis-tris, pH 6.8, 0.233 M NaCl, 2 mM EDTA, 0.1 mM $ZnSO_4$ and 0.1 mM DTT. Washed cells were then sedimented and resuspended in the same buffer plus 0.2 mM PMSF and 100 µg/ml lysozyme. The cells were incubated on ice for 30 minutes. Sarkosyl detergent (Frankel et al., 1990; 1991) was added to the lysis buffer to a final concentration of 0.2%, and RQ1 DNaseI was added to reduce the viscosity of the solution. The lysate was centrifuged at 10,000 g for 35 minutes at 4°C. Octyl-glucoside was added to the low speed supernatant to a final concentration of 2% using 25% stock. After incubation on ice for 5 min, $MgCl_2$ and $CaCl_2$ were added to final concentrations of 1.25 and 1.00 mM, respectively. The solution was stirred for 30 minutes at 4°C and then subjected to high speed centrifugation at 40,000 g for 16 hours at 4°C. The resultant supernatant was tested for ALAD enzymatic activity.

For small scale lysis preparations, glass beads were employed to break open the cells. Glass beads ≤ 106 microns in diameter were used as described by Scopes (1987).

Freezing (16 hours at -20°C) and thawing on ice in the presence of 2 mM EDTA, pH 8.0, of cells containing pLysE (BL21) allowed the T7 lysozyme to lyse the cells efficiently. Nucleic acids are also solubilized by this procedure.

N. Quantitation of total protein

In instances where a gross measurement of protein was needed (e.g. aliquots from gel filtration chromatography) absorbance at 280 nm was measured in Bis-Tris buffer at room temperature, except where noted.

The Bradford protein assay (1976) was performed for quantitative protein assays.

Coomassie Blue G was used as a dye, and cuvettes were washed in ethanol between samples. Absorbance was read at 595 nm. Bovine serum albumin was used for standard curves for all assays.

O. ALAD activity assay

Enzymatic activity of ALAD was detected by a modified ALAD assay developed by Sassa (1982). This protocol involved incubation of cell lysate, DTT, and the substrate ALA to final concentrations of 20 mM and 8 mM, respectively, in Bis-tris buffer, pH 6.8 at 37°C for 1 hour. Trichloroacetic acid (TCA) and $HgCl_2$ were then added to precipitate proteins. After 15 seconds at 12,000 rpm, the supernatant was combined with a modified Erhlich's solution which develops a dark pink color in the presence of PBG. Color change was measured with a spectrophotometer at 553 nanometers. Units of ALAD activity are defined to be micromoles of PBG formed per ml per hour.

P. Polyacrylamide gel electrophoresis

Polyacrylamide gel electrophoresis protocols were in accordance with the standard methods of Laemli (1970).

Q. Silver stain

Silver staining of polyacrylamide gels was carried out using the manufacturer's instructions (Biorad), including 3 fixation steps, an oxidation step followed by several washing steps, incubation with the silver reagent, another washing step, 3 developing steps and an incubation in stop solution.

R. Immunodetection

After electrophoresis, the gel was first incubated for 16 hr in 30 mM phosphate

buffered saline (PBS), pH 7.4, with 100 ml of rabbit anti-human ALAD antibody diluted 1:1000 in 2.5% non-fat dried milk in PBS. The polyclonal anti-human ALAD IgG was raised against the human erythrocyte enzyme purified to homogeneity by the method of Anderson and Desnick (1979). The gel was washed with 10 ml of PBS, incubated for 30 minutes with biotinylated goat anti-rabbit IgG, washed for 10 minutes with 100 ml of PBS, and finally incubated for 10 minutes with a preformed complex of avidin and biotinylated horseradish peroxidase provided in the Vectastain ABC kit. Following a 10 minutes wash with 100 ml of PBS, the gel was incubated for 5 minutes in the horseradish peroxidase color reagent and then photographed with Polaroid 084 film.

S. Purification of recombinant ALAD

Solid ammonium sulfate was added to the solution until the ammonium sulfate concentration reached 45% of saturation (Anderson and Desnick, 1979). The mixture was stirred at 4°C for one hour and then centrifuged at 5,000 g for 30 minutes. The precipitate was resuspended in 0.05 M Bis-Tris, pH 6.8, 0.233 M NaCl, 2 mM EDTA, 1 mM dithiothreitol and 0.2 mM ZnCl (lysis buffer).

Sephacryl 300 was used for gel filtration of the bacterial cell lysate. Because the molecular weight cut off of S-300 is 300 kDa, the homooctameric ALAD was expected to elute near the beginning of the column. The column had a total bed volume of 750 ml, a calculated void volume of 200 ml, and was 124 cm in length. The column was washed with several volumes of lysis buffer. Samples were layered carefully over the top of the resin and washed into the column with 2 layers of 20 ml of lysis buffer. Aliquots were collected by a fraction collector beginning immediately after the sample had entered

the column. The flow rate of the column was controlled by a peristaltic pump at 30 cm/hr.

Dry ion exchange cellulose was precycled in 0.5 N HCl and then, 0.5 N NaOH, as described by manufacturer. A column (35 cm) was packed in the standard fashion (Scopes, 1987). Equilibration, sample loading and sample elution were done according to Anderson and Desnick (1979) except for the following modification: Bis-Tris, pH 6.8 replaced potassium phosphate buffer, zinc and DTT were included in all buffers at 0.2 mM and 0.1 mM, respectively, and the 0.12 M NaCl uroporphyrinogen-I-synthase elution step was eliminated.

T. Sucrose gradient

A linear gradient of 5% - 20% sucrose was used. The sample was layered at the top of the gradient and subjected to centrifugation in a SW50.1 rotor at 45,000 rpm for 9 hours in a Beckman L3-50 Preparative Ultracentrifuge. After centrifugation, 3 drop aliquots were collected by making a hole at the bottom of the centrifuge tube. The aliquots were diluted with 400 μl of Bis-Tris buffer and tested for reactivity with the anti-ALAD antibody by Western analysis. Proteins were analyzed by PAGE and silver staining.

V. Results

A. Characterization of ALAD gene

Genomic Charon 30 lambda ALAD clones were isolated by hybridization to the pUC9 ALAD cDNA clones and were subcloned into two separate pUC19 plasmids. These plasmids were characterized by restriction mapping and Southern hybridization, and were completely sequenced in both directions. Analysis of the 5′ putative promoter region revealed a 3′ splice site immediately preceded by an Alu sequence, indicating the presence of another upstream non-coding exon. Further analysis of the original pKT218 cDNA clone revealed a second internal *Pst*I site and a short sequence upstream of this *Pst*I site. This sequence contained a rare *Eag*I site and was presumed to be from the missing first exon. The *Eag*I site was located 40 nt 3′ from the end of the cDNA sequence of pKT218. Thus, the 5′ untranslated region was 131 nt in length. Sequencing of the cloning site of the pKT218 revealed a mutation in the *Pst*I sequence, thus providing an explanation for the *Pst*I subcloning of pALAD 50 which did not contain the first exon with its *Eag*I site.

Screening of the genomic lambda clones revealed the presence of the additional exon with an *Eag*I site approximately 7.5 kb from exon 2. Lambda clones containing exon 1, 5′ upstream sequence of exon 1 and intron 1 were subcloned into pUC19 vectors and sequenced in both directions, thus completing the genomic sequence of ALAD (Figure 12). Sequence analysis of genomic clones of human ALAD established the presence of 13 exons over a span of 13.5 kilobases, including the two 5′ non-coding exons, 1A which is approximately 7.1 kilobases from, and 1B which is approximately 3.2 kilobases from

Figure 12. Genomic sequence of human ALAD. Bold face type indicates ALU sequences or upstream promoter elements.

1	....	....	.AAGCTTTTT	ATATTTCAAC	AGTTTCTCTA
51	CTTATGTCCT	TTTTCGATTA	CAGGGGCCAA	CCCAGAATAC	CACATTGCAT
101	CTAGTGAACT	CACTTTTTAA	GACTGAGATG	AAATTCAAAG	AACATAGAAT
151	TAACCATTTT	TAAATGAGCA	GTTCAGTGGC	ATTTAGAACA	GTCACAATGT
201	TGTGCAACCA	CTACCTCTAT	CTAGTTTCAA	AACATTTTCC	TCAGGCCCTG
251	GCAATCAGCA	GTCTGCTTTC	TTGTTCCTTT	TTTATGCTTT	TCACAGCATT
301	TAGCACAATG	CCATATACCT	ACCTAGAAGA	CACTCAGAAT	AGTTTAGTTT
351	TAATTTTTTA	TTGAAACGAA	ATATACTCAT	AAAAGACACA	AATAAGTGTG
401	GAGCTTAATG	AATTTTCACG	AATTAAACAC	AGCCATGTAA	CTCACACCCA
451	GACCAATAAA	CATTAACTGC	ATCCAAAGAA	ACCTCTTTAG	TCTCCCCCTG
501	CCGCC**TTTTT**	**TTTTTTTTTT**	**TTTTTTTTTT**	**TTTTTTTTCT**	**GGAGACAGGG**
551	**TCTTGCTCTG**	**TCATCCAGGC**	**TGGAGTGGTA**	**CGATCACAGC**	**TTACCGCAGC**
601	**CTTGACCTCC**	**CAGGCTCAAG**	**CAATCCTCCC**	**ATGTCAGCCT**	**CCCAAGTAGC**
651	**TGGGACTACA**	**GGTGTCACCA**	**CCATGCCCAG**	**CTACTTTTTG**	**TAATTTTTGT**
701	**AGAGACTGTT**	**TCGCCATGTT**	**GCCAAGGTTG**	**GTCTTGAACT**	**CCTGGGCTCA**
751	**AGCGATCCAC**	**CAGCCTTGGC**	**CTCCCAAAGT**	**GCTAGGATTA**	**CAGGTGTGAG**
801	**CCACCACACC**	**CAGCC**TAATC	TCCCTTTCAA	TCCTCCCCAT	CTCCCCAGTG
851	GAAATCACTA	TTCTGAGTCC	TAACAACACA	GCTTAATTTT	GCTTGATTGT
901	GTAGTTAATA	TAAATGAAAT	CATATTATAT	ACTCTTTTGT	GTCTCACTTC
951	TTTTACTCAA	CAATATGTTT	AAGATTTATT	CATATTATTG	CATTCACAAT
1001	ACTGAAAGAT	TAAGGTAAAA	TGTTTCAAAA	GGTGGAGGCT	GGGGGGAGAG
1051	TGGGAACAGG	TCTTTCAGTC	AAGAAGAAGA	AAAGACATGG	GGGAAATAAG
1101	CTACACCAAA	CAGCAACATT	TTAAAATTTC	CGTTCTATTT	GGGTTCTTAG
1151	ATGGTTCTAA	TTATTTTTAG	TTACCAATAA	CCTCAGGGAA	GGAATGAGTG
1201	TCAGGAAAGA	AATGTCGGGC	AGGGAATCTA	GGAAAAAACA	AATAATAATT
1251	GCCCCTATGG	AGTGTGGTTA	TTAGCAGGCA	TTGCAGAGAG	AGCCCAGTGT
1301	AACATGGGGG	AAATTATTTA	ATCTAGCTGG	CCAATTGTTT	TCCCAATCCG
1351	TAAAGTGAGA	ATAGTAACAT	CGTTGCTGTC	CAAAATTAGA	CAAATGAATG
1401	TAAAATGGCA	TGGTTACTGT	CCCTGACTCT	TAATCGACAA	AACATTATTT
1451	GCTTTCTATC	CCCTTCCTTG	TTGAATACAG	ATCAGAACTC	CAGTAGCTGG
1501	AAACTAAGCA	GTGTCATATT	TCCCTAGCCT	TACGCCGCTA	GGCCCCTTAA
1551	ATTTAAAATA	TAAAGTATAC	AACAGGTTTG	GGGTACAGAT	TTTGAGAAAT
1601	TTAAAAAAAA	**ATTTTTTTTG**	**AGACAGAGTC**	**TCCCTCTGTC**	**ACCCAGGCTG**
1651	**GAATGCAGTG**	**GCATGATCTC**	**AGCCCATTAC**	**AACCTCCGCC**	**TCCCGGGTTC**
1701	**AAGCGATTCT**	**CCTGCCTCAG**	**CCTCCCGAGT**	**AGCTGGGATT**	**ACAGGAGCCC**
1751	**GCCACCACGC**	**CCGGCTAATT**	**TTTTGTATTT**	**TTAGTAGACG**	**ACAGGGTTTC**
1801	**ACCATGTTGG**	**TCAGGCTGGT**	**CTTGAACTCC**	**TGACCTCAAG**	**AGATCCGCCC**
1851	**GCCTTGGCCT**	**CCCAAAGTGC**	**TGGGATTACA**	**GGCGTGACGA**	**CCGAACCGGG**
1901	**CTGAG**AAATA	AAATCTGAGA	ATAAGTGTAA	TGTAACCCAC	TAGTGTCTCT
1951	ACATTTGAAA	GTCAACTTTT	CAATAACCCC	TGAAGAACCA	CCTGAAAGGC
2001	ATTGCTTCGC	TAGGGCGATG	ATGCCCGATC	GCCATTTTAC	TCATCCAATC
2051	CTACTGGTTT	ACCCAGGCCC	CATAAGCCTC	AAGATCCCTC	TTGCCCTGCA
2101	CATCTCCAAT	CTCCATAAAG	ACCTTTGATC	GGATCTATCA	TTGTACCTAT
2151	CATAGGTCTG	ATGCCCCTAT	CAAGACTTGG	AGTTTTCCTA	AACGCCCATG
2201	TCTTTCACAT	CACATCTCTC	AACTCATAGC	ATTGCCGTAC	CCTGAGAAAT
2251	AAATGAAGCT	GGTACAAATT	TTCTCTGAAA	ACTTGGAAGC	GGCGCTGAAC

```
2301  TCAACAACTT TAGCTTAGTC CGCAAAAATG GACTTAGCTA ATACTGTGTT
2351  AGGCATAATG GTGCTGAACG GCTTTGACAT CCAATTATTC TTTCACCAAG
2401  CCATTAACCA GAGCCAAGTT TGCCCTAGGC TGGGCAAAGG AATGCTGTGA
2451  AGTGAAACCT GGGCTGTAAC CGCCTCCGAG GCCGGCTCAC ACCCAGGGAC
2501  CCTCCCAAGC GGTTCCTGCG CCCCCTCGTG GCCTCCCTCC TGCAGTTCGC
2551  TGCACAAGCG CAGGGCCCGG AGCGGGGACG GGCGGGCCCT CCAGGCCCTC
2601  GACACACCCA ATTGGCCAGG CCTCGTTGCG TGTGGGCGGG GCCGCGGCTG
2651  GAGCGGAGGA GCCCCCCAGA AGACCGCCTT CAGGGGGCGT GGCCTGGACT
2701  GTGGAGGGGG CGGACCGTGG GGAAGCTGCT TCCGGGTGAG CCCCCCCCGC
2751  TCTTACGCGG TCTGT
                        >EXON 1A>
                        GGGAG ACCGGAGCGG GAGACAGCGG TGACAGGAGC
2801  AGCGGCCGGG AGCCCTTAGG GAGGCAG
                                     >INTRON 1A>
                                     GTG AGCGGGGGCT GGATGCGGGG
2851  AGATAGCGGG CCTGGAGGGC GGCTCCAGGG CGAGGCGGGG ACTGTGGCAC
2901  CAAGAGTCCT GCGTCCCCAG ATTGTGCTGT GCGCCTGAAG CCCTGGCGGT
2951  GCAGCCGTGC ACGGAGTCCC GTGGAGCGTT TGTTTGGGCT GGATTATCCC
3001  GCTGCAGCCG AGGTTGGGGG CCCGGGTTCG GGGCTCCGCA CCTCTTCTCT
3051  GGGGTTCACC CTCTCGGATA CCGCCTGGGC AGACTCCATC GATGCAGAAC
3101  CCGAGGTGCG GGTGGTGGTT GGTGGGGGGG TAGGGAGGTG AAAACAATAA
3151  CAGCCGTTTT GTACTAAGCG AAGCGTGCAT TGACCATTCA TTTATTTTAT
3201  TTATTTATTT ATTTTTGAGA CGGAGTTTCG CTCTTGTTGC CCAGGCTGGA
3251  GTGCAATGGC GCGATCTCAG CTCACTGCAA CCTCCGCCTC CCGGGTTCAA
3301  GCGATTCTCC TGCCTCAGCC TGCCAACTAG CTGGGATTAC AGGCGTGTGC
3351  CACCACGCCT GCTAATTTTT GTATTTTTAG TAGAGATGGG GTTTCTCCCT
3401  GTTGGTCAGG CTAGTCTCGA ACTCCCGAGC TCAGGTGATT CGCCCACCTC
3451  GGCCTCCCAA AGTGCAGGGA TTACAGGCGT GAGCCACCAC GCCCGGCCTG
3501  ACCCATTCAT TTCTTACAAC AACCCATTGA GGTCGGTTCT AGTCCCCATT
3551  TACTAAACGA AGAAACTGAG GTCACACAAC TAGAAAGCTA CAGAACTAGG
3601  ATTTGTACCC ACATCTATCA GACTCCAGAG CCCACATACT TAACCATCCT
3651  GCTACCAACA GCAGCCAGGC CCAGAGTCTG GCTCACCCAA AGCTCTCCAC
3701  AAGCCCAACC TGCTGGGAAA CAGATTTGAT ATCTATCCTT AAAGTATGGA
3751  GGTGTCATGG TGGGCACTGA GGACTGGGCT GCTGGAGACC TTCCCCCAGA
3801  TCTGCCGCTG ACCCACTGCG TGATCCTGGG CCTGTGTCTA CCCCTTGAGC
3851  CCAGTTTCAC CCTTGTTTCA GTGATACAAT CAGCAAACTT TAAGAAGACT
3901  CTTAGCCAAG ATATTCTAGG ATTCTGGGAA TGGCAGCTTC TCCTACGTCT
3951  AAATGGCTAA CCCAGAGGCG ATCAGGCTCT TCAAAGCCCC CAGGGCAAGG
4001  CTGGGGTCTC ATTCTTCTCT TTAATCCCTG TGTCTAGCAC TTCAGCAGGC
4051  CCATAGTAGA TATTCAAGAA ATGTGGGTTG GCTGAATAAA CACTTCATCT
4101  CTTTTCAGAA CTGCATGACA GCTGCTCTCC TGGGGCCTTT GTGGGGCTTT
4151  GGATTTCTCA GTTCTGGCCA CCTTCTATCT AGCTGGCTCT CCTGACTGCC
4201  TCCTCCCTGG CCAGTGGCAG AGCCTGCCAG GGCTGGCAGC AGGTCCTAAC
4251  CCTCTGTGCG AAGACCCACA GCCTCTGGAA TCCTCAGCAT AGTACAACTT
4301  GCAGGGGTGG GGGGTGGCAC ATGAAGCTCA TAGCCCCTCC TGGTAGCAGT
4351  GCTCAGGGCC CAACTTCAGT TACTCCCAGT TTGGGGATGT TGCCTTCCAT
4401  GGGAGGAGAT ACTTTCATAT TAAATCTGAA ACACTTCTTA TCCTTCACTG
4451  TCCCCTCTTC CATGAAACCA CCCTTGACTC AGAGACACAG TCCCTCTGAC
4501  AGCACCTGGT CCACCCCTCT ACTCTGGCAG CCCGGTGGTG ATTATTTGT
```

```
4551   TTTTCCTCCA CCACATAACC TGCTGGAAGG CAGGGACTCC ACTTCAGAAA
4601   TCTTGGAAAA CCCTGTGGCC CACAAGGGTT TGAGCTCATA TTCCCATCTG
4651   GTGAGGTGCA GAAGGCACAG AGCTGCGCCA GGACATCTCG TTCCAACCCA
4701   GCCCCACTCC TAATAGTAAT AAAAGTAGCT ACATGTTTAC TGTGCAGCTA
4751   CTTGGTGCCA GACACCTACC TACCTTATCC CTATCAATAA TCTCCACAGC
4801   CTTGCAAGTT AGAGATATCA TGCCTATTTT GTAGATTATG AAATGAACAC
4851   TGAAAGGTCA ATTACTTTGC CCAAAATTAC ACAGCAGAGC ATAGACCAAA
4901   CCTGGGACTG TCTGACCCTG AAACATGGAT CTTTCTGCAC CGCCCTATAC
4951   TCTGCGTCCC AGCAAAGCTA ACCTGTCTGA TGTGGTTGGG CAGAGGTGCT
5001   GGACCTCCTG GCGGCTCGCT CCTGTCTTCA GTGTGGTAAT GTCACAGGCA
5051   TTTAAATTGT GGTTTGCAAG CTCGCTTGAT TTGGGAGAAG TGCCCTCCCA
5101   GTCTGTGAAA TTCTCCCAAA CTCTAAGGGC CTTTTTGCAA AGCAAAACCC
5151   ACCCTGCATT TGGCTCTGCC AATATGTGAT TCAGGGGGTG TTTTCCCACT
5201   CACAGAATGG AGTGGGAAAC TCCACTTATT AAGTACCTCT TACAATGGAA
5251   TGACTCTATG GTGTAGGTAT CAGGGCACTG AAATAGAAGC CAGGAGCTGC
5301   AGGGTCAGTT TATAGCTCTG TATGTCACTC ACTGTGATCT CCAAAAGTCC
5351   CGTTGCCTCT TTATACCCCA TTTTCCACAT TAAAGTGTAA GGGATTAAAC
5401   AGGATTGGCA ATTTTCAAG
                           >CLONE 3.0>
                           C TTTTAATATT TTCAGCACTG ACAGCCGTTT
5451   CTCAAATCAA ATCTGTCTCT GAATCTTAAA ATTCAAAAAC AGATCAAAGT
5501   GGACTGCTGT AGTGAGACTG TTCCCTTTAC TCTAACCCTG CCCCGGTCTC
5551   AGGCTCCTAG GCAGTCCTGC CCCAGGGTTT CTGCTAAGCA GGAGCCAGTT
5601   GTTCCCTCAG CCTCCTAAAA GCGATGGGGA GTTGAGGCCA GTGGAATGCA
5651   GCTTTGGATT CAGAAGCCAG AAAACAACAG AGATTGTTTG TGGGAAGTGA
5701   GTCAGAGAGA TTAGACCTCT CTTCAACATC TGACTAGACT CTAGAATGGC
5751   ATTAAGCAAG TCACTGAACC GCTCAGAGCC TCTGTTTCCT TATCACTAAA
5801   AGTTAGGACA ATTCTTGCAA TGAAATGAGA TACTACTACC AAGTACTTCA
5851   GTGCTGTGCC CAGCACAGAA CGAATACACT CAGTATGTTT TTCTTTTGTT
5901   TCTTCTGTCT CTCCAGCACC CAAAGCAGAA ACAGTATGGG CAGAGAGACG
5951   CCTACTGAAT CAGGTGTTTA TCAGTATTAT GAAATACAGT ATATAAAATT
6001   TAGCATTTGG AGTCCAGAAC AAATTCAAAT TCTGGCTCAG TCACTAAGTA
6051   GCTATGTGGT CTTGGGCAAC TGGCTTACAT TTTGGTAATT CTCTCAATAT
6101   GCAAAAATAT TCTCTCGGCC AGGTGCGGTG GCTCACGCCT GTAATCCCAG
6151   CACTTTGGGA GGCAGAGGCG GGCGGTTCAC TTGAAGTCAG GAGTTTGAGA
6201   CCAGCCTGGC CAACATGGCG AAACCCTGTC TCTACTAAAA ATACAAAAAT
6251   TAGCCAAGCA TGGTCATGCA TGCCTGTAGT CCCATCTATT TGGGAGGCTG
6301   AGGCAGGAGA ATCACTTGAA CCCAGGAGGT GGAGGTTGCA GTGAGCCGAG
6351   ATCGAACCAC TGCACTCCAG CTTGGGTGAC AGAGCCAGAC TCTGTCTCAA
6401   AAAAAAAAAA AAAAAAAAGA AAAGGAAAAA AAACTCCTGG CAGACATGAG
6451   AGCAAAATTG AGTACGAAAT GGGCAGCTGG GACCTCATAT CTACTGAATA
6551   CTGTGTTAGG TGATGGGGCT GGCCCACCCA GGGGTGTGGC AGCAGGGCTG
6601   GCCCCTGTGT GACTTGTGAT AACCCCACCC TACCAAGGAG GAAGACTGGA
6651   TAAAATGGCC TGAGATGGCT GAAGGCAGCT AAAGGGGCCT AAGAGACTAT
6701   GGGGAGAAGG GCAGCCAGTT CCTAGGCAGC CCTGCCCCAC TGTTTCTGAT
6751   AAGCAGGAGG CAGGTGTTCC CTCAGCCTCC TGAAAGCAAT GAGGAGTTGA
6801   GGCCACGGGA ATGCAGCTTT GGATTCCAGA AGCCAAGGAC CTGTTACAAA
6851   GGTGGTTTTC TGGG
                      >POSSIBLE EXON 1B START SITE>
```

AAGCAA ATTAGAGGGA TTAGACCTGG CTTCGACATC
6901 TGACTTGACT CCAACCTACC AAAGGATCTC CCAGTTCACC TGGGAGTTAA
6951 CACACCAGGG TTCTAGTCCT TTTGCTCGCT GTGTGGCTTT GG

>INTRON 1B>

GTAAGTCC
7001 TTGCCCCTCT CTGGACCTTG AGGTCCCATT CATACAAAAA GCAGCCGGAC
7051 TCAGGCCTGC CAGCCCCCAT TTTTGTGGCT CTTCCCCGCC CCTCCCCATC
7101 CTTTGGCTAT GGGATAGTCT CCCAGGCACC TCCAAAAGGA AGTGTCTAAA
7151 ACAGAGCTCT CCACCTCCTC TCCAAGCCCA GCTCCCCACT GCCCCCTTCC
7201 TGCCCTGTCC CTGTCTCCAC TGCCCTTTAG ATCAGGCAGG CAGGCAGGCC
7251 TAGAGCATCC TCCCTGAGAC ACTGCCTGTT GGCTGCCTTA AGACGCCCCT
7301 TCCTTCTCTG CCCGCAGCCT CCCCAGCCTC TGCATTCCTC TGCTTCCTGA
7351 CTGCTCTGCT CAGCCTGCAG GATTGATCTT CCCAAAGCCA GGTGTCAACC
7401 AGGCCACTAA CCCTACCCAC TGCTCAGAAA TCTTCCCACA GCCTCAGCCT
7451 GGCCCCCTCT TATCCATCCC CTCCATACAC ACTTGCCTCA GCTCCTCAAA
7501 ACAGTCCACA GCCATCCTCC TTCCCCACCT TTGCCCCTTT CTTCCCTTCC
7551 TAAGGAAATA TTGGCCCCCT TCTCTCCACC AGGCTGAACC CACTCATCCT
7601 GCAAAACCAA GTTCAACGCC CTGTCATTTT CTTTTTTTTT TTTTTTTTTG
7651 AGATGGAGTC TCACACTGTC GCCCAGGCTA GAGTGCAATG GGTGCACTCT
7701 CGGCTCACTG CAACCTCTGC CTCCTAGGTT CAATCGATTC TTCTGCCTCA
7751 GCCTCCAGAG GGAACTCTGA CTACAGGTGC ACACCACCAC ACCCAGCTAA
7801 TTTGTATTTT ATTTTATTTC TATTTATTTA TTTATTGTGA GATGGAGTTT
7851 CTCTCTTGTT GCCCATGCTG GAGTGCAATG GCACGATCTC GGCTCATTGC
7901 AACCTCCGCC TCCCAAGTTC AAAGGATTCT CCTATCTCAG CTTCTGAAGT
7951 AGCTGGGATT ACAGGCGTGT GCCACCACAC CCGGCTAATT TTTTTTTATT
8001 TTTAGTAGAG ATGGGGTTTC ACCGATGTTG GTCAGGGTTG GTCTCAATCC
8051 TCGACCATAC AAGTGATCCA CCCAACCTCG GCCTCCCAAA GTGCTGGGAT
8101 TACAGGAGTG AGCCACCGTG CCTGGGCTAA TTTTTGTATT TTTAGTAGAG
8151 ACGGAGTTTC ACCACATTGG CCAGCCTGGT CTCAAACTCC AGACCTCAAG
8201 TGATCCACCC GCCTCAGCCT CCCAAAGTGC TGGGGTTACA GGTGTGAGCC
8251 ACTACGCCTG GCGCCTGTCT TCTTCTAATA TGCCTCTTCT GACCTCCTCA
8301 CTCCCATTGG AGCCATGGCT CACTTCTGGT CAGGCTCCAT GCACTCAATC
8351 AGCCCTGCCT GCCAGGACTG TAGTGCCCAC TTGATGAGAA ATCACCTGTG
8401 AAAATGCCAG GCATGGTCTG GCACACGGTA GGCGCTCAAG AAACAAATGT
8451 TAATGCCCCA GCTTCCTTCT GTCTTGTGTC CACTAATTTG GCATTTGGAG
8501 CATATTTTTG TACTTCACGT AATTAAATAT TTACTGTGTG CCAAGGTCAG
8551 TGTGTGACAG AGCACTGTAG GGGAGGCAAA AATGAACCAA CTCTGGACCT
8601 TCCCTGCAGA AGCTCTGTCT ACTGGCATGA GGTAAGACAG GTCCACAGAT
8651 AACTTACATA AGAGAGACTG CAATCTGGAC AAGACAGCCT CAGTGGGAGA
8701 GGGTTGGAGA ACCCTGTAGA TTTGGGTGCG GGGAGGCTGT GGGGAGAGCA
8751 CAGCCTTTCA GGTAAAGGGA CAAGGTTAGG AAAGCAAGAA GGTGTAGCTT
8801 GTGCTCGTGG AGCCCAGGAA ACTCCTTGGG CTAGAAAAGA TTTGGTGGCA
8851 CTCTATACAG TGTATGTGTC TGTCTGTCTA TCTAACATGA ATGGTACTAT
8901 TTTCTCCTAT TGGTGTGCAT TGTACCTAAC AAAAAATCCA TGTCCCAAGT
8951 ACTTTTCTCT TTTCTTTCTT TCTTTTTTTT TTTGGAATGA AGTCTTGCTC
9001 TGTCGCCCAG GCTGGAGTGA AGTGGCGGTA TCTCGGCTTA CTGCAACCTC
9051 CTCCTCCTGG ATTCAAGCAG TTCTCCTGCC TCAGCCTCCC CAGTAGCTGG
9101 GATTACAGGC ACGCACCACC ACGCCCAGCT AATTTTTTTT GTATTTTTAG

9151 **TAGAGGTGGG GTTTAACCGT GTTGGCCAGG TCTGGCTCGA ACTCCTGACC**
9201 **TCAAGTGATC TGCCCATCTC AGCCTCCCAA AGTGCTGAGA TTATAGGTGT**
9251 **GAGCCGCCAT GCCCGGCCAA** CAGTACTTTT CTCTAGAGGT AGAATTGCAG
9301 GGGATTTTTC TCTTTTTGCT TATCTGCACT TTTCTAAAAG TTCTGCCATG
9351 AACACGGGCA TTTTTTGTGG AATAAACAAA GTTTAAATGT TGGCTAGATC
9401 AGTGCTCTTT ATGCATAAAG GAAAGGATTA GTTCTTTTTT TTCTATTTCC
9451 AATTTATTGA TTCTTTGATA AAATGCAATG CAATAAAATT **ATTATTTTTT**
9501 **TTTAGAGACG TCTCCCTCTC TCGCCCAGGC TGGAGTGCAG TGGTGCAATC**
9551 **ATAGCTTACT GCAACCTTGA ACTCCTGGGC TCAAGCGATC CTCTTGCTTT**
9601 **AGCCTCCTGA GTAGCTGGGA CTACAGGTGT GGACCACCAC ACCTGGCTGA**
9651 **TTTTTAAACT TTTTGTAGAT ACAGGGTCTC GCCATGTTGC CCAGGAGTTC**
9701 **GAGACCATCC TGGGAAGCAT GGCAAGACCT CCATCTCTAC AAAAAATTCG**
9751 **AAAATTAGCT GGATGTTGTG GTGCACACCT GCAGTCCCAG CTACTTGGGA**
9801 **GGCTGAGTTG GGAGAAACAG TTGAGCCCGG GAGGTCAAGG CTGCAGTGAG**
9851 **TCGAGATTGC ACCACTGCAC TCCAGCCTGG GCGACAGAGA CCCTGTGTGA**
9901 **AAAAAAAAAA AAGAAGAGAA** TTTTTTTTAA ACAGTCATTG CTTGCTCAGA
9951 TGTTTACTTT AAAAGATAAT AATGAACAAG AAGCAGTCAC ATAAAATACA
10001 AGCCCAAATT TTATATCATT AGATTCTGAT TGTCATGAAA GTTTCTAAAG
10051 ACTTACTTTC ATTTCTCAAC TTACCTTGTT GACCAGCAGG GATTGGTGAA
10101 CCAGGCTGTG AGTAGCATTG GGCTAGAGAG AGGGGAGGCA GGAATCTAGA
10151 AGAGCTGTTT TCCAGATGTG ACCATCTCCT GAGGACAGGG ACCATGTCCT
10201 ATGTGCCACC CATCACCCCC CACAG

>EXON 2>

ACAGA GCCTGCAGCC AATGCCCCAG
10251 GAGCCCTCGG TTCCAACCAA CTGATGCCCC TGTGCCCACT GGCCCACGCC
10301

>BEGIN AMINO ACID CODING SEQUENCE>

ATGCAGCCCC AGTCCGTTCT GCACAGCGGC TACTTCCACC CACTACTTCG
10351 GGCCTGGCAG ACAGCCACCA CCACCCTCAA TGCCTCCAAC CTCATCTACC
10401 CCATCTTTGT CAC

>INTRON 2>

GTGAGTC TCCAAGAATG GCCAGGCCT CTGCTCTGCT
10451 GGTTGGGGTT GGGGTTGGGG AGGGAGTGTT GACTGGAGCG GGCATCAGTA
10501 TGGCTGGGGG TGGCAAAGTG AGCTGTCAGC TTGAAATTCA AGGCACTGGA
10551 AGCAGGCTAC TTGGATTAAG GACAGGAATC TTAGGAACAA AACAAACTTT
10601 GAAAGAACTC ATTCATCCCA TTTGGAAAAT TAGAAGAATA ACCCTTGCCT
10651 GCCATCCTGA GCTCTTGCAG TAAGACAGAA GCTGAGAAGG TGCTCTGTAC
10701 ATTGTAAAGT GCTATGTACC TGTAAGAGAT GGCAGTCATT **GAGGCTGGGC**
10751 **ACGGTGGCTC ACGCCTGTAA TCCCAGCACT TTGGGAGGCT GAGGCAGGCG**
10801 **GATCACGAGG TCAGGAGATC GAGACCATCC TGGCTAATAT GGTGAAACCC**
10851 **TGTCTCTACT AAAAACACAA AGAAATTAGC CAGGCGTGGT GGCGGGTGCC**
10901 **TGTAGTCCCA GCTACTTGGG AGGCTGAGGC AGGAGAATGG CGTGAACCCG**
10951 **GGAGGCGGAG CTTGCAGTGA GCCGAGATTG CACCACTTCA CTCCAGCCTG**
11001 **GGCGACAGAG CCAGACTCCA TCTCAAAAAA AAAAAAAAAA AAAA**GAGATG
11051 GCAATCGTGA TTGTTAATAA TAATGCAGAC ATTTACTGAG TACTTACTAT
11101 CTACCAGGTA CTATGCTAAG CACCTACACA CATTATCTCA TTCAATTCTG
11151 AGAGCATTTG TATGAAGAAG GAGTAGCTAT CCTCTAGAAC ATCAGCTCCA
11201 TGAGGGCAGG GATGTTTGTC TATTTTGTTC ACTGTTGTAT CATCAGGGCC
11251 TAGAACAGTA CTTGGCACAT AATAAGTACT CAATAAATAT TTGTTGAATG

```
11301  AATGAATTAA CCACGCATGA TATAGATGAA GGCCTAAGGC TCAAAGAGAT
11351  GATAGAACTT GGCCACGGTC ACCCAGGCAG TAAGTGGCTG GGATAGAAAG
11401  CAAGGACCTG CCAAATTCAG AGTCCAAGTT CTTAACCACT TAATTCCTTC
11451  CTGTAATTAC CGTTCTTTTA GTACAGTTGC TAGTGTTGTC ACTGTTATTC
11501  TTGTTGTTCC TATTATTATT TCAGGCCCTG GGCTTGGCCA GGCAGGGAAG
11551  CCAGACACTG GATCCCATCC TCCTCCCACC ATCTCCACTT CCATATTTCT
11601  TTCCTGCTTC CCAACCATCC CTCTCAGTCG CCCCCGCACC ACTGGCCCTT
11651  CCCACAGCTA CCAATCCATA TCCCACCCCC GCTCTTGCAG
                                                >EXON 3>
                                                GGATGTTCCT
11701  GATGACATAC AGCCTATCAC CAGCCTCCCA GGAGTGGCCA G
                                                 >INTRON
       3>
                                                 GTAGGAGAC
11751  GTGGAGTTGG GGGGCCAGCG GGTGGTGGAG GGAGAGATTC CACAGGTGGA
11801  AGTGCTGGGA GGCAGAAGCA GACCTAGGAA GTAGAAGATG CGGACAGACA
11851  GACATTAGCT CAGTAGAGGA AAGGGTTTCC CCGGGGCCAG AGCTGTTCCA
11901  CAGTGGAAGG GGCAGCCCCA TAAAGTAAAG AGCTACCCAT CACCCGAGAC
11951  GTCGTGGCAG AGGCTGTTGC AGAAGGGAGC TGAACTGCAG ATGGGAGTTC
12001  AAAAAGAGGG CCTCGAAGGA GCCTTCCACA GCCGAA
                                           >CLONE 3.6>
                                           TTCC GGAGCTCTGC
12051  TACTCAGGGC CTCAGTCTTC CCTCCTATTT AGTGGATGCA TCCCTGCCCC
12101  TTCTGTCCTG GGGGCTTGAG CCCTCCTGGT GCCATATGCA GCTTGGTTTC
12151  TAACAGAGGC ACACAGTGTG GTGGGGTCCG GAGGACCGTT GCCTGGGACC
12201  TGCCTTCCTT CAACCCCTCT ACCCACACCC ACACAG
                                           >EXON 4>
                                           GTAC GGTGTGAAGC
12251  GGCTGGAAGA GATGCTGAGG CCCTTGGTGG AAGAGGGCCT ACGCTGTGTC
12301  TTGATCTTTG GCGTCCCCAG CAGAGTTCCC AAG
                                        >INTRON 4>
                                        GTGAAGA ATCAAAGGAA
12351  GGGCTAAGAA GGGAGGTTGC CTCACGCCCG TAATCCCAGC ACTTTGGGAG
12401  GCCAAAGTGG GTGGATCACT TGAGCCCAGG ATTTTGAGAC CAGCCTGGAC
12451  AACATGGCAA AACCCATCTC TACAAAAAAT ACAAAAGTTA GCTGGGTGTG
12501  GGGGTATGTG CCTGTAGTCC CAGCTACTCG GGAGGTGGAG AGGTGGGAGG
12551  ATTGCTTGAG CCCAGAAAGT CGAGGCTGCA GTGAGCCAAA ATCGCGCCAG
12601  TGCACTCTAG CCTGGGTGAC AGAGCAAGAC CCTGTCTCCA ATACAAACAG
12651  AAAAAGGAAG GGAGGTTGGG CAAAGGTGGA CTGAGGGTCC ACACTGACTG
12701  CACCCTCACT CCCACATTGT GCTGGCCCTG GGGCCACAGG TGAATGGACG
12751  TGGTCTTTGC CCTTAAGTCA GCACCCATGT AGGGTCGGTC CTCTGTGCTT
12801  CCTTATCCAG GGGCTGTGAT GATGAAGGAA GGAGAAGGCC AGGGCTATGC
12851  TCTGTGATGG CTGTCATCCT GCCTTCCAAA GCTACATGTA ATAGACACAC
12901  TGCTTTGTCC CTCCCCTGCC CCTAG
                            >EXON 5>
                            GACGA GCGGGGTTCC GCAGCTGACT
12951  CCGAGGAGTC CCCAGCTATT GAGGCAATCC ATCTGTTGAG GAAGACCTTC
13001  CCCAACCTCC TGGTGGCCTG TGATGTCTGC CTGTGTCCCT ACACCTCCCA
13051  TGGTCACTGC G
```

```
                  >INTRON 5>
                  GTGAGTTCC CTCCCTCCCA CCAGCCCTGC TGCCACCCAC
13101   ACTCCTACTG CCCACTTCTC AACAGGGTGG GGACAGCCAG GGCCCAAGGT
13151   GCTCCCCAAA ACCCAGTCAT CTGTCCTGAA G
                                          >EXON 6>
                                          GGCTCCTGA GTGAAAACGG
13201   AGCATTCCGG GCTGAGGAGA GCCGCCAGCG GCTGGCTGAG GTGGCATTGG
13251   CGTATGCCAA GGCAG
                      >INTRON 6>
                      GTGAG TGAACCACCA GCAGGGATGG GCACCTCTGG
13301   GTCAGGAGGT GGCAGAGTGG CTAGGAGGGC CCCAGAGTTC TGAAGGCCAC
13351   CCTCTGCCCC CCAG
                     >EXON 7>
                     GATGTC AGGTGGTAGC CCCGTCGGAC ATGATGGATG
13401   GACGCGTGGA AGCCATCAAA GAGGCCCTGA TGGCACATGG ACTTGGCAAC
13451   AGG
           >INTRON 7>
           GTAAGGG CAGGGAATGC AGCACAGGGC TGGCAGGAGA TAGTCTGCAC
13501   CAGCCCTGCC CCCGTGTCTG CTAAGAATCA CAGAACTGCC GGGCGTGTTG
13551   GCTCACACCT GTAGTCCCAG CACTTTGGGA GGCTGAGGCA GGTAGATCAC
13601   TTGAGGTCAG GGGTTCAAGA CCAGCCTGGC CAACATGGTG AAACCCCATC
13651   TCTACTAAAA ACACAAAAAT TAGCTGGGCG TGGTGGCAGG CGCCTGCAAT
13701   CCCAGCTACT GGGGAGGCTG AGGCAGGAGA ATCGCTTGAA CCCACGAGGC
13751   AGTGAGCTGA GATCATGCCA CTGCACTTCA GCCTGGATGA CAGAGCTAGA
13801   CTCCATCTCA AAAAAAAAAA GAATCACAGA ACTGAAGACA GTGCTGGATG
13851   AGGCTTTGGG GAACCATTTA AACCTCTGGG CCTCTGCAGG GAAATCAAGC
13901   CCAGCACTCC AACAGGACCA GAACACAGGC AGTCTCCTTC CCAGCCTAGG
13951   TTCTTTCTCT CCCTGCCACA TCACCCTGGG ATACCTGGCA AGGGCCGAAT
14001   AAGCCAAGAC CTCCATTGTC TCCCCATAG
                                        >EXON 8>
                                        G TATCGGTGAT GAGCTACAGT
14051   GCCAAATTTG CTTCCTGTTT CTATGGCCCT TTCCG
                                              >INTRON 8>
                                              GTGAG CAGGGGTGGG
14101   CAGGGGTCTG CTGTGAATCC CTGCCCTTTG GCCCAAAGCT GGAGCCCACC
14151   CTGATGACTC TGCTTTGCAG
                             >EXON 9>
                             GGATGCAGCT AAGTCAAGCC CAGCTTTTGG
14201   GGACCGCCGC TGCTACCAGC TGCCCCCTGG AGCACGAGGC CTGGCTCTCC
14251   GAGCTGTG
                >INTRON 9>
                GT GAGTGACTAG GACTTGAGCC CCACCCTCAG CCCCCTCCTA
14301   GGCACCACCC ACATTATACC CTCATCCCTT AG
                                           >EXON 10>
                                           GACCGGGA TGTACGGGAA
14351   GGAGCTGACA TGCTCATGGT GAAGCCGGGA ATGCCCTACC TGGACATCGT
14401   GCGGGAGGTA AAGGACAAG
                            >INTRON 10>
                            G TGAGCACAGG TACGAGGCAA AGGGGGCTCA
```

```
14451   GGGGGCTGGG ACAGAGTTTT CCACAGACTC TGGAATCTCA GAGTTGGAAG
14501   CAGTTTGCCC TTAAGCATGC ATCCTCTCCT CCCCTTCCCT GCCCAGGAAC
14551   CATCGTGGCC TTCTATGTCG GGGCTTGCAC GAGCCTCAAA CAGCCCTGCT
14601   TTAACAGTTC AAGAGTGGGC CAGGCTGCCA GCCGCAGTAA CCCAGGACAC
14651   GGGGCTCAAG ATGGTCACAG ATTGAGCAGG GGGGAAGGGA CGCTTCCAGA
14701   GCCACATCCA CCCTCCATTT CAGCCTGTCT CCCTGTCTGC TTCCCTGCAG
14751
        >EXON 11>
        CACCCTGACC TCCCTCTCGC CGTGTACCAC GTCTCTGGAG AGTTTGCCAT
14801   GCTGTGGCAT GGAGCCCAGG CCGGGGCATT TGATCTCAAG GCTGCCGTAC
14851   TGGAGGCCAT GACTGCCTTC CGCAGAGCAG
                                         >INTRON 11>
                                         GTAGGCAGGC AAGGGTGGGG
14901   TGTTTTGACC TGCGCCACAG GGACTGATAA GCACTCTGCC TAGATCGGGG
14951   AACGACGTCC TGAGAGCTTG GGATCTTATT CCGGGAATTA CTAGTGATCT
15001   AAACAGACAC ACACTGAGGA AGAGATATGG AACTGCAGCA TAGAACACGG
15051   CCCGGTGAAG CAAGCAGAGC CCTTCATTTT TGGTTGTGAG AACGTGGCAA
15101   GCCACTTCTC TGAACCTCAG TGTCCTCACC CATAACTGGA TAACTGGGGA
15151   TAAGATACCT GGTGCGTGGT TGTCCTGAGG ATTAAATGAA GTAATATCAC
15201   TCCATAAAGG GGACTCATTT TGTTAGAATT GCACACCAGC ATGGGAAGGA
15251   ACTTGCCTCT TATATTTCCT TCACTGTGCA TTTTATTCTT TGGTAAACTG
15301   AGGCCCCAAA AGAGGAAATG ACTTGCCCAA GAAATAGAGT TTCCCAAAGC
15351   TGGGCTCCGT CTCATGTGGT GTGCCCACAG GCTGTGCTTC TTCATGGTAG
15401   CCTTCTTCCC CGCCTGGCCT TCCCATCGCA GAAGGTGTGC TCAGAGCTGA
15451   TCAGCGTCCC CCCAGCAACT TTCTGCATCT CTCCCAACAC AG
                                                     >EXON
        12>
                                                     GTGCTGAC
15501   ATCATCATCA CCTACTACAC ACCGCAGCTG CTGCAGTGGC TGAAGGAGGA
15551   ATGATGGAGA CAGTGCCAGG CCCAAGAACT AGAACTTTAA AACGTTCCCG
15601   GGGCCTCAGA CAAGTGAAAA CCAAAGTAAA TGCTGCTTTT AGAACTGT
                                                            >P
        OLYA IN cDNA>
                                                            GC
15651   CCTCATGCCC TCTTCCTGCT CACATGCTAG CGGGGCCCAG CAGCCCTGGG
15701   TGGTTTTGCC AGCATGCTAA CTCTTGTAAC TCGCAGCTGC ATCCTATGAG
15751   CTCTCCCAAG CTT
```

GENOMIC MAP OF HUMAN ALAD

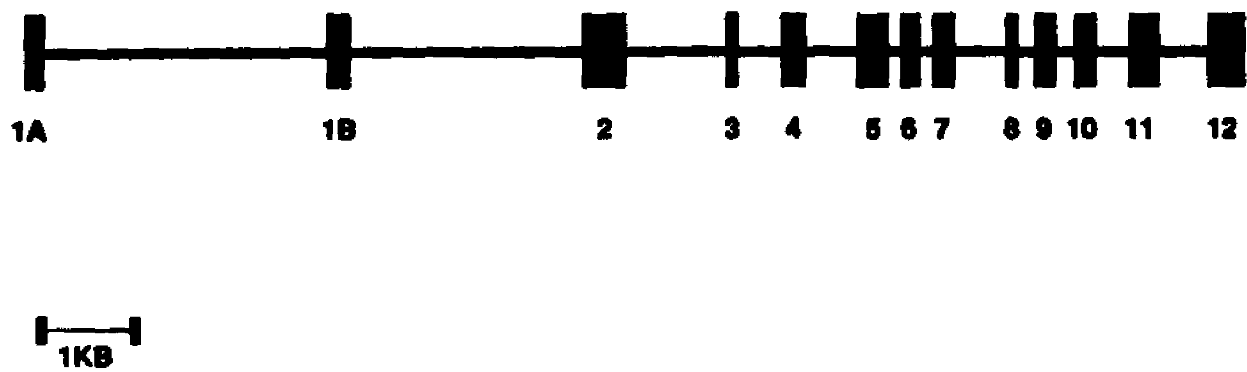


Figure 13. Genomic map of human ALAD.

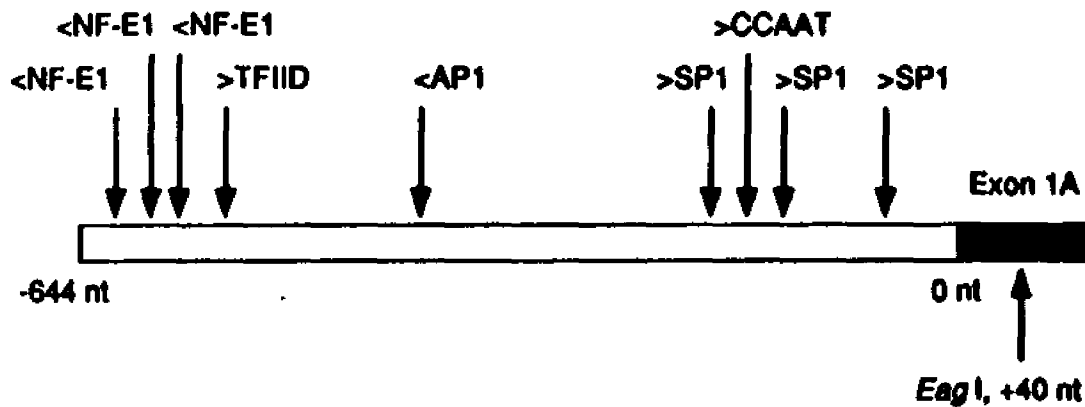
INTRON				5' BOUNDARY		3' BOUNDARY				INTRON SIZE (nt)
1A	GGG	AGG	CAG	gtgagc	///	ggtggttttctggg	AAG	CAA	ATT	3900
1B	TGG	CTT	TGG	gtaagt	///	atcaccccccacag	ACA	GAG	GCC	3200
2	TTT	GTC	AC	gtgagt	///	ccccgctcttgcag	G	GAT	GTT	1286
	phe	val	thr				thr	asp	val	
	36	37	38				38	39	40	
3	GTG	GCC	AG	gtagga	///	ccacacccacacag	G	TAT	GGT	490
	val	ala	arg				arg	tyr	gly	
	53	54	55				55	56	57	
4	GTT	CCC	AAG	gtgaag	///	tcccctgcccctag	GAC	GAG	CGG	642
	val	pro	lys				asp	glu	arg	
	85	86	87				88	89	90	
5	CAC	TGC	G	gtgagt	///	catctgtcctgaag	GG	CTC	CTG	121
	his	cys	gly				gly	leu	leu	
	131	132	133				133	134	135	
6	AAG	GCA	G	gtgagt	///	ctctgtccccccag	GA	TGT	CAG	99
	lys	ala	gly				gly	cys	gln	
	159	160	161				161	162	163	
7	GGC	AAC	AGG	gtaagg	///	ttgtctccccatag	GTA	TCG	GTG	576
	gly	asn	arg				val	ser	val	
	188	189	190				191	192	193	
8	CCT	TTC	CG	gtgagc	///	actctgctttgcag	G	GAT	GCA	85
	pro	phe	arg				arg	asp	ala	
	207	208	209				209	210	211	
9	CGA	GCT	GTG	gtaagt	///	ccctcatcccttag	GAC	CGG	GAT	74
	arg	ala	val				asp	arg	asp	
	236	237	238				239	240	241	
10	AAG	GAC	AAG	gtgagc	///	ctgcttccctgcag	CAC	CCT	GAC	331
	lys	asp	lys				his	pro	asp	
	265	266	267				268	269	270	
11	AGA	GCA	G	gtaggc	///	ctctcccaacacag	GT	GCT	GAC	614
	arg	ala	gly				gly	ala	asp	
	309	310	311				311	312	313	

Figure 14. Exon/intron boundaries of human ALAD.

the second or first coding exon (Figure 13). The exonic sequence in the genomic DNA sequence was identical to the previously published cDNA sequence including the coding region, the upstream ATG and the unusual polyadenylation site. The exon\intron boundaries are shown in Figure 14. The zinc binding site, identified by the presence of a consensus cysteine and histidine rich domain, was located in exon five. The active lysine, known to be involved in ALA binding, as well as surrounding regions which contain several charged amino acids thought to be involved in stabilization of the active site, were located in exon ten.

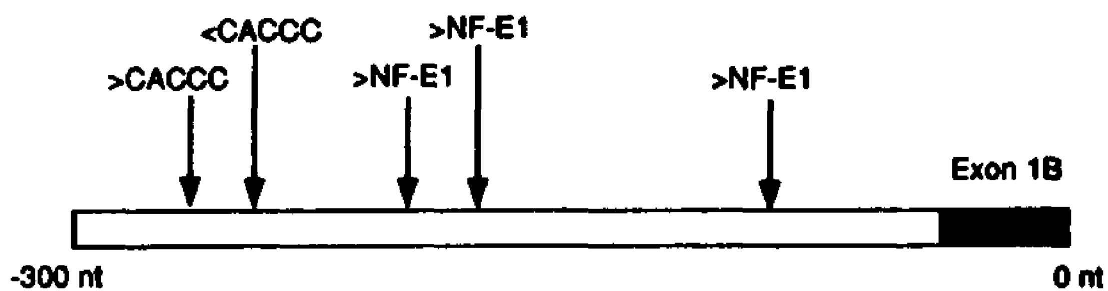
Analysis of the Alu repetitive elements within the ALAD genomic sequence revealed nine Alu repeats scattered throughout the intervening and upstream sequences (Figure 12). The three most 5′ Alu repeats were reversed in orientation, with two located 5′ upstream from the *Eag*I site located in exon 1A and another located in intron 1A. Five Alu repeated elements were located in intron 1B, Alu sequences 5,6 and 9,8 were fused. Alu 5,6,7 and 8 were all reversed in orientation while Alu 9 was not. The three remaining were forward orientation Alu repeats were located in intron 2, intron 4 and intron 7. The Alu sequences are bold faced in the genomic sequence shown in Figure 12. Additional repetitive elements, including LINE (L1) (Long interspersed repetitive elements) and THE (transposon like elements), were not found in the genomic sequence.

The proximal promoter region of exon 1A was extremely GC rich and contained three possible SP1 binding sites and a CCAAT box (Figure 15). However, there was no consensus TATA (TFIID) (Maniatis et al., 1982) box present. A perfect SP1 transcription factor (Kadonaga et al., 1986) binding sequence was 167 base pairs from the *Eag*I site.



Site		Consensus	Direction	Location
NF-E1	TGATAG	(A/T)GATA(A/G)	Reverse	-634
NF-E1	TGATAG	(A/T)GATA(A/G)	Reverse	-624
NF-E1	TGATAG	(A/T)GATA(A/G)	Reverse	-594
TFIID	AATAAAT	TATA(A/T)(A/T)T	Forward	-617
AP1	CGGACTAA	(C/G)TGACT(C/A)A	Reverse	-443
SP1	CGGGCGGGCC	(T/G)GGGCGG(GGC/AAT)	Forward	-184
CCAAT	GCCAAT	GCCAAT	Reverse	-158
SP1	TGGGCGGGGC	(T/G)GGGCGG(GGC/AAT)	Forward	-130
SP1	GGGGGCGGACC	(T/G)GGGCGG(GGC/AAT)	Forward	-67

Figure 15. Location of ALAD promoter elements of exon 1A.



Site		Consensus	Direction	Location
CACCC	CACCC	CACCC	Forward	-281
CACCC	CACCC	CACCC	Reverse	-268
NF-E1	TGATAG	(A/T)GATA(A/G)	Forward	-240
NF-E1	TGATAG	(A/T)GATA(A/G)	Forward	-209
NF-E1	TGATAG	(A/T)GATA(A/G)	Forward	-112

Figure 16. Location of ALAD promoter elements of exon 1B.

The promoter element, CCAAT (Maniatis et al., 1982), which is commonly used to transcribe genes by PolII, was 27 base pairs downstream from this SP1 binding site and upstream of the most distal SP1 site. A cluster of three NF-E1 (Tsai et al., 1989) sites was found in the distal promoter region, approximately 620 bp from the *Eag*I site and 200 bp upstream from a putative TFIID (TATA box) binding site. An AP1 binding site was located between the distal and proximal promoter regions. No acceptable 5′-splice junction sequences were located in the promoter region.

The promoter region of exon 1B contained a pair of CACCC boxes, a feature shared with the globin gene promoters, which are palindromic around an AG nucleotide center (-284bp from the possible start site of exon 1B). Three NF-E1 sites were located at -112, -209 and -240 bp from the possible start site of exon 1B. The NF-E1 sites were all in the forward orientation (Figure 16).

PCR was used to amplify promoter regions of exon 1A. Oligonucleotide primers were designed to have a *Hin*dIII restriction site to facilitate cloning into the pSV0 CAT expression vector. A set of promoter deletions was constructed varying in size from the proximal promoter with the SP1 and CCAAT sites to the complete promoter, including the AP1 site and the distal NF-E1 and TFIID sites. These constructs will be used to transfect mammalian cells for transient CAT assays.

B. <u>Isolation and characterization of the ALAD-2 allele</u>

To determine the coding sequence for the $ALAD^2$ allele, total RNA was isolated from cultured lymphoblasts of an $ALAD^2/ALAD^2$ individual, reverse-transcribed into cDNAs, and then the entire $ALAD^2$ coding region and 44 nucleotides of 5′ untranslated sequence

was PCR-amplified and sequenced. Compared to the sequence of the ALAD^{1} allele (Wetmur et al., 1986b), the ALAD^{2} allele had a single G to C transversion at nucleotide 177 of the coding sequence as the only difference in all three subclones analyzed. This transversion created an *Msp*I restriction endonuclease site and predicted the substitution of a positively-charged lysine by a neutral asparagine in residue 59 (designated K59N) of the ALAD 2 subunit.

C. Amplification method for ALAD genotyping

The *Msp*I restriction site created by the G to C transversion in the ALAD^{2} allele permitted the development of a PCR-based method for ALAD genotyping as illustrated in Figure 17. For amplification of the 916 bp ALAD genomic sequence containing the ALAD^{1}/ALAD^{2} polymorphic site, sense (5′-AGACAGACATTAGCTCAGTA-3′) and antisense (5′-GGCAAAGACCACGTCCATTC-3′) oligodeoxynucleotides were synthesized which corresponded to ALAD intronic sequences. Amplification with an annealing temperature of 55°C and *Msp*I digestion of the 916 bp ALAD genomic fragment from the ALAD^{1} or ALAD^{2} alleles resulted in 582 bp and 511 bp fragments, respectively. The fragments were analyzed by agarose gel electrophoresis followed by staining with ethidium bromide and fluorography (Figure 18, lanes 3-5).

To determine if the nucleotide 177 G to C transversion was present in other individuals who had the ALAD 1-2 or 2-2 isozyme phenotype, genomic DNAs from 85 ALAD 1-2 and eight ALAD 2-2 individuals were amplified and the 916 bp ALAD PCR products were cleaved with *Msp*I. All ALAD 1-2 individuals had both the 582 bp and 511 bp fragments, consistent with heterozygosity for the ALAD^{1} and ALAD^{2} alleles, and

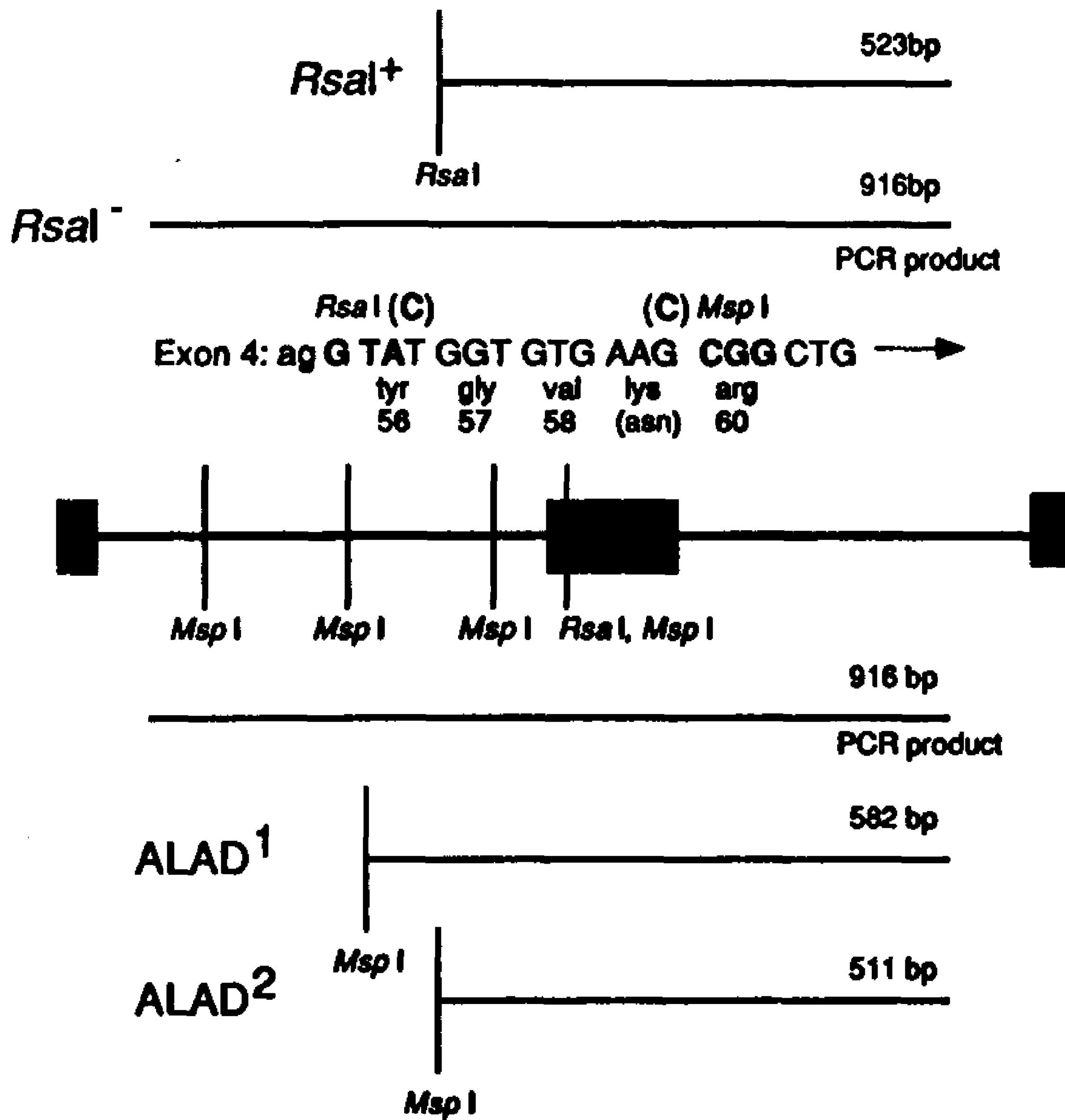


Figure 17. PCR-based method for ALAD genotyping.

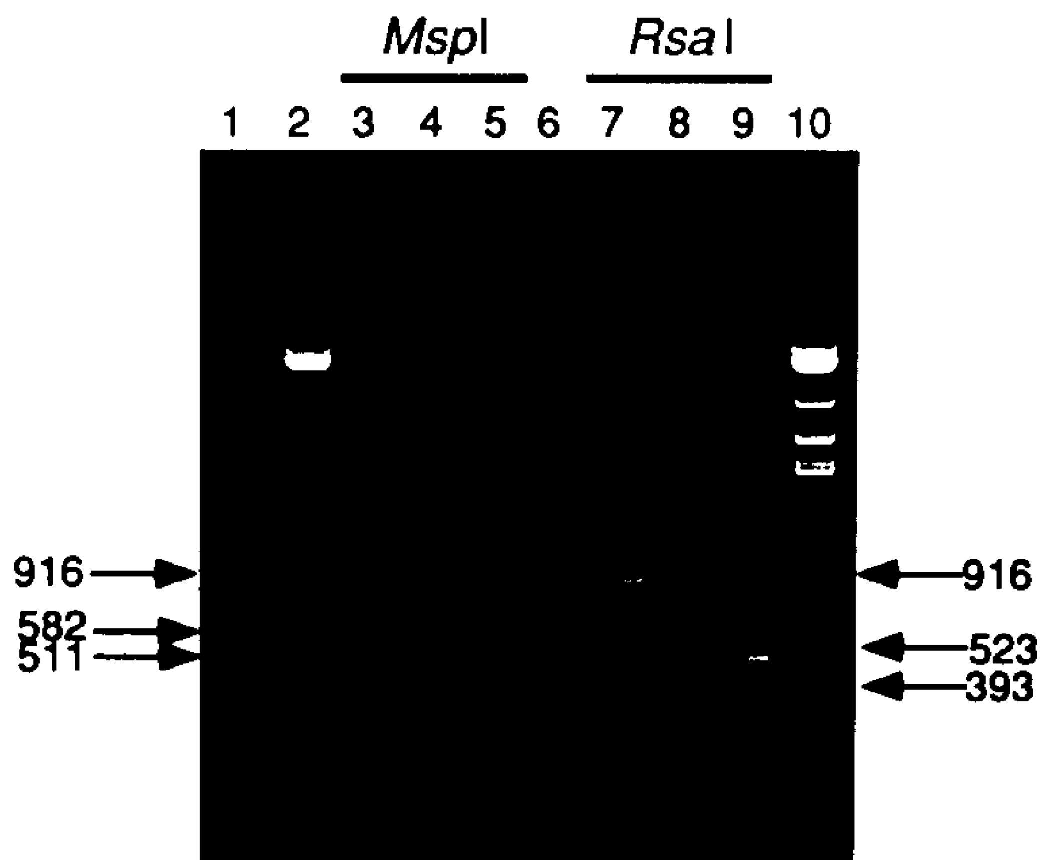


Figure 18. Ethidium bromide stained agarose gel. PCR product from the PCR-based method for ALAD genotyping restricted with *Msp*I and *Rsa*I.

all ALAD 2-2 individuals had only the 511 fragment, consistent with homozygosity for the $ALAD^2$ allele. Of the 87 individuals who had the ALAD 1-1 isozyme phenotype, molecular analyses revealed that all had the $ALAD^1$ alleles.

Since the previously identified *Rsa*I ALAD polymorphism (Astrin et al. 1991) occurred in the coding sequence (nucleotide 168) near the *Msp*I site in the $ALAD^2$ allele, the *Rsa*I$^+$ polymorphism also could be analyzed by digestion of the 916 bp PCR product into 523 and 393 bp fragments (Figure 17 and 18). To control for incomplete *Rsa*I digestion, the Bluescript SK vector, which contains two *Rsa*I cleavage sites, was added to the reaction after amplification of the genomic DNA. Complete digestion of Bluescript SK vector resulted in 1189 and 1770 bp fragments. The polymorphic information content of the *Msp*I/*Rsa*I haplotype was calculated as described by Botstein et al. (1980).

Analysis of a random population of 428 normal Caucasian individuals revealed that the $ALAD^1$ (*Msp*I$^-$) and $ALAD^2$ (*Msp*I$^+$) allele frequencies were 0.88 and 0.12, respectively. The allele frequencies for the *Rsa*I$^-$ and *Rsa*I$^+$ alleles in the same population were 0.75 and 0.25, respectively. Individually, the *Msp*I and *Rsa*I RFLPs were in Hardy-Weinberg equilibrium (X^2 *Msp*I = 3.5, df = 2, $p > 0.10$; X^2 *Rsa*I = 1.88, df = 2, $p > 0.25$). Based on the above frequencies, the expected *Msp*I/*Rsa*I haplotypes would be $ALAD^1$(*Msp*I$^-$)/*Rsa*I$^-$, 0.66; $ALAD^1$/*Rsa*I$^+$, 0.22; $ALAD^2$/*Rsa*I$^-$, 0.09 and $ALAD^2$/*Rsa*I$^+$, 0.03.

However, the two RFLPs were in linkage disequilibrium ($X^2 = 22$, df = 1, $p < 0.001$). Of the 259 $ALAD^1$ alleles studied, 27.4% were *Rsa*I$^+$, whereas only 5% of 101 $ALAD^2$ alleles were *Rsa*I$^+$. The expected number of $ALAD^1$/RsaI$^+$ and $ALAD^2$/*Rsa*I$^+$ were each

25%, thus the $ALAD^2$/*Rsa*I$^+$ allele was highly underrepresented. The PIC for the combined *Msp*I/*Rsa*I haplotype was 0.45, a reasonably informative value (0.5 > PIC > 0.25) (Botstein et al., 1980).

D. Bacterial expression of ALAD-1 and ALAD-2 alleles

Studies aimed at achieving significant expression of the human enzyme in *E. coli* started with the construction of a pUC9 expression vector, named pUC9-50. The pUC9 vector was chosen so that the ALAD sequence would be in frame with the lacZ gene and would produce a fusion protein of ß-galactosidase (9 amino acids) and ALAD. The vector was made by partial *Pst*I and complete *Xma*I digestion of the pKT218 ALAD cDNA plasmid, and cloned into *Pst*I and *Xma*I digested pUC9. The resultant expression vector contained the ALAD cDNA sequence from the *Pst*I site at position -65 to the *Xma*I site at 1040 and the promoter, ribosome binding site and 25 nucleotides of the *lac*Z gene prior to the *Pst*I cloning site. Thus, a fusion protein of ß-galactosidase with ALAD was produced with 31 additional amino acids [(25+68)/3] prior to the N-terminus (Figure 19).

It was necessary to remove the extraneous amino acids from the end terminal of the expressed protein of pUC9-50. The first step was to introduce a mutation (C to G) two base pairs upstream from the initiation codon ATG. This base substitution now allowed the restriction enzyme *Sph*I to recognize and cleave the sequence. This construct was labeled pUC9-50-S (Figure 20).

Double-stranded site-directed mutagenesis was carried out with the use of oligonucleotide primers. The double-stranded DNA was separated into two samples, one of which was linearized by cutting with a single enzyme and the second of which was

LacZ (in pUC-9):

```
ATG ACC ATG ATT ACG CCA AGC TTG GCT (PstI)
```

ALA-D 5'-Untranslated Region:

```
GCA GCC AAT GCC CCA GGA GCC CTC GGT TCC AAC
CAA CTG ATG CCC CTG TGC CCA CTG GCC CAC GCC
```

ALA-D Coding Sequence:

```
  1 ATGCAGCCCC AGTCCGTTCT GCACAGCGGC TACTTCCACC CACTACTTCG GGCCTGGCAG
 61 ACAGCCACCA CCACCCTCAA TGCCTCCAAC CTCATCTACC CCATCTTTGT CACGGATGTT
121 CCTGATGACA TACAGCCTAT CACCAGCCTC CCAGGAGTGG CCAGGTATGG TGTGAAGCGG
181 CTGGAAGAGA TGCTGAGGCC CTTGGTGGAA GAGGGCCTAC GCTGTGTCTT GATCTTTGGC
241 GTCCCCAGCA GAGTTCCCAA GGACGAGCGG GGTTCCGCAG CTGACTCCGA GGAGTCCCCA
301 GCTATTGAGG CAATCCATCT GTTGAGGAAG ACCTTCCCCA ACCTCCTGGT GGCCTGTGAT
361 GTCTGCCTGT GTCCCTACAC CTCCCATGGT CACTGCGGGC TCCTGAGTGA AAACGGAGCA
421 TTCCGGGCTG AGGAGAGCCG CCAGCGGCTG GCTGAGGTGG CATTGGCGTA TGCCAAGGCA
481 GGATGTCAGG TGGTAGCCCC GTCGGACATG ATGGATGGAC GCGTGGAAGC CATCAAAGAG
541 GCCCTGATGG CACATGGACT TGGCAACAGG GTATCGGTGA TGAGCTACAG TGCCAAATTT
601 GCTTCCTGTT TCTATGGCCC TTTCCGGGAT GCAGCTAAGT CAAGCCCAGC TTTTGGGGAC
661 CGCCGCTGCT ACCAGCTGCC CCCTGGAGCA CGAGGCCTGG CTCTCCGAGC TGTGGACCGG
721 GATGTACGGG AAGGAGCTGA CATGCTCATG GTGAAGCCGG GAATGCCCTA CCTGGACATC
781 GTGCGGGAGG TAAAGGACAA GCACCCTGAC CTCCCTCTCG CCGTGTACCA CGTCTCTGGA
841 GAGTTTGCCA TGCTGTGGCA TGGAGCCCAG GCCGGGGCAT TTGATCTCAA GGCTGCCGTA
901 CTGGAGGCCA TGACTGCCTT CCGCAGAGCA GGTGCTGACA TCATCATCAC CTACTACACA
961 CCGCAGCTGC TGCAGTGGCT GAAGGAGGAA TGA
```

ALA-D 3' Untranslated Region:

```
TGGAGACAGT GCCAGGCCCA AGAACTAGAA CTTTCAAAAC GTTCCCGGG (XmaI)
```

Normal ALA-D: 330 Amino Acids.   Expressed ALA-D: 361 Amino Acids.

Figure 19. pUC9-50, an ALD expression vector.

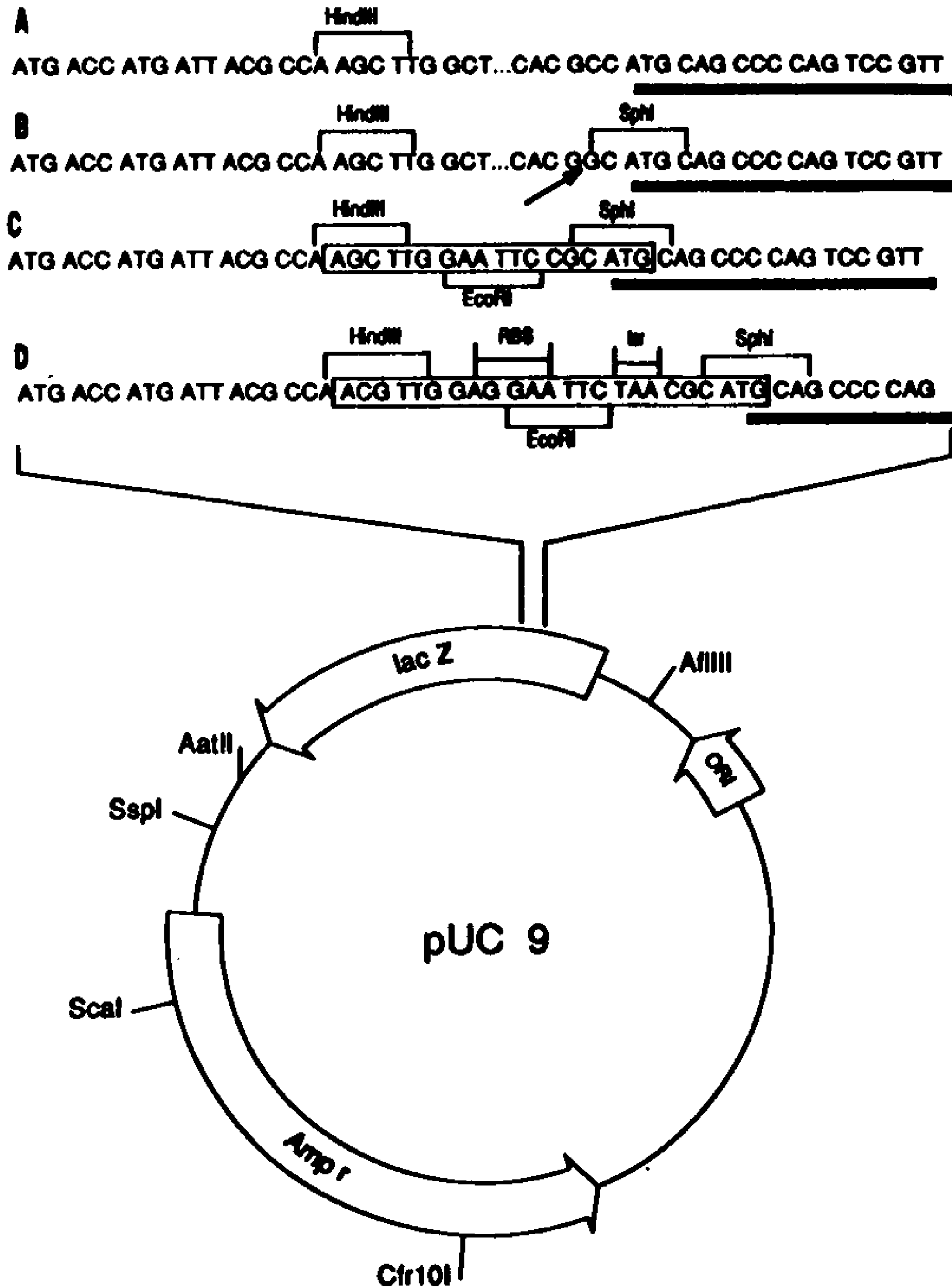


Figure 20. pUC9-50 derived expression vectors.

cleaved with two enzymes surrounding the site at which mutagenesis was to occur. These two aliquots were mixed in the presence of 50% formamide at 100°C. This mixture was allowed to cool to 37°C over 30 minutes. The resultant structures were linear plasmid, plasmid restricted at two sites and "gapped" plasmid which contained a single stranded gap in the area at which site directed mutagenesis was to occur. An oligonucleotide primer was synthesized with a single base substitution at the desired location. This primer was allowed to anneal to the mixture, including the gapped plasmid. The oligonucleotide preferentially bound to the gapped plasmid as it was single-stranded in that area. Space between the oligonucleotide and the gap was filled in with T4 DNA polymerase and the product was transformed into *E. coli*. Only those plasmids that are circular will transform and replicate in *E. coli*. In the case of pUC9-50-S, mutagenesis was indicated by the presence of the unique *Sph*I site created by the site-directed mutagenesis. The sequence was confirmed by dideoxy sequencing.

A fusion vector that did not contain the upstream ATG was created with oligonucleotide linkers which were made specific for a 5′ *Hin*dIII and a 3′ *Sph*I site and which contained an *Eco*RI site. pUC9-50-S was cleaved with *Hin*dIII and *Sph*I and allowed to ligate in the presence of a 6 fold molar excess of linker oligonucleotides (Figure 20). The resultant colonies were screened for the absence of the upstream *Pst*I site and the presence of an *Eco*RI site which had been synthesized into the linker. The resultant construct was called pUC9-50-S-L(Linker). This vector retained the ALAD coding sequence in frame with the ß-galactosidase amino acid sequence. However, the non-ALAD leader sequence was now shortened to 11 amino acids. Thus, these clones

continued to produce a fusion protein with ALAD immunologically cross-reactive with anti-ALAD antibody and possessing ALAD activity.

The polycistronic vector, pUC9-50-PC, was constructed by the addition of oligonucleotide linkers (Figure 20). The sequence of the oligonucleotide linker contained a 5′ overhang of a *Hin*dIII site and the 3′ overhang of a *Sph*I site, a ribosome binding site and an in-frame termination codon for the ß-galactosidase protein. The logic of this construct was that the ribosome would stop production of ß-galactosidase but remain bound to the message by the proximity of the newly introduced ribosome binding site and could initiate synthesis of the ALAD protein. A non-fusion protein product was produced which was also immunologically cross-reactive with anti-ALAD antibody.

Using the *Eco*RI site in pUC-9-50-S-L and the *Xma*I at the 3′ end of the ALAD sequence, the ALAD coding sequence was subcloned into the pKK223-3 expression vector. This vector, pKK-50, placed the initiation codon ATG in proximity to a ribosome binding site present in the pKK223-3 vector.

Expression from all vectors was detected by Western blot analysis (Figure 21). All the vectors except for pKK50 lead to expression of soluble product which cross reacted with the anti-ALAD antibody and was, by PAGE analysis, the appropriate molecular weight. The pKK-50 vector, in fact, did produce the ALAD protein as an aggregate which was insoluble in lysis buffer. The expressed protein had no activity when incubated with substrate. The sarkosyl-octyl glucoside method for bacterial cell lysis might have been used to solubilize aggregates of the recombinant protein expressed in *E. coli*. However, the polycistronic vector expressed large quantities of soluble protein and

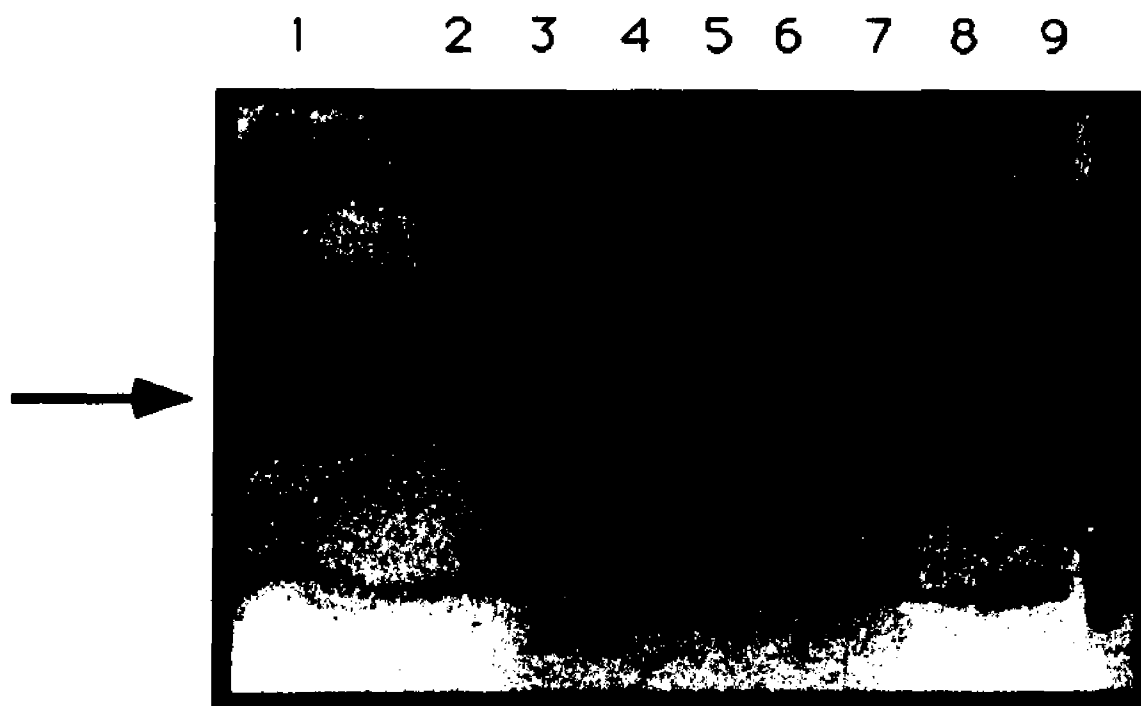


Figure 21. Western blot analysis of recombinant ALAD.

1. Control ALAD
2. pUC9-PC14
3. pUC9-PC12
4. pUC9-PC7
5. pUC9-PC5
6. pUC9-PC3
7. pUC9-50-SL
8. pKK50
9. Biotinylated Standards

in the BL21 strain lysis was simple and did not involve long incubation times.

Immediately following bacterial cell lysis, an ammonium sulfate precipitation was performed (4x purification; Table 4). At this point the lysate, now an ammonium sulfate pellet, was either frozen or resuspended in 10-20 ml of lysis buffer (PMSF was omitted from this buffer) and loaded onto the Sephacryl S-300 column (at least 2.5-fold purification; Figure 22 A, B, C and Table 4). DEAE-cellulose was used to concentrate sample (Figure 22 A, B) but did not purify the enzyme in an appreciable manner. Enzyme always was stored as an ammonium sulfate pellet or in 50% glycerol at -20°C.

The ALAD^2 allele was cloned by PCR with oligonucleotide primers designed to contain an *Sph*I site at the translation start and a *Bam*HI site at the 3′-non-coding region. The template was the original ALAD^2 plasmid which had been produced by reverse transcription of RNA. The ALAD^2 PCR product was cleaved with *Sph*I and *Bam*HI and cloned into pUC19 with a 6 molar excess of polycistronic oligonucleotide adapters. The resultant expression vector contained exactly the same coding sequence as the ALAD^1 polycistronic vector except for the mutation at nucleotide 177 (G to C).

In order to understand the biochemical mechanism by which the ALAD isozyme phenotypes influence lead accumulation in exposed individuals, milligram quantities of the isozymes, ALAD 1-1 and ALAD 2-2, are needed, and the only practical source was recombinant enzyme. To date, authentic, soluble and biologically active human ALAD 1-1 has been produced using a co-cistronic vector in *E. coli* LC137, a *lon*$^-$ and *htp*$^{R-}$ strain. The ALAD enzymatic activity was 0.7 U/mg protein, over 1000 times that found in human erythrocytes (0.45 mU/mg protein). The recombinant protein was identical to

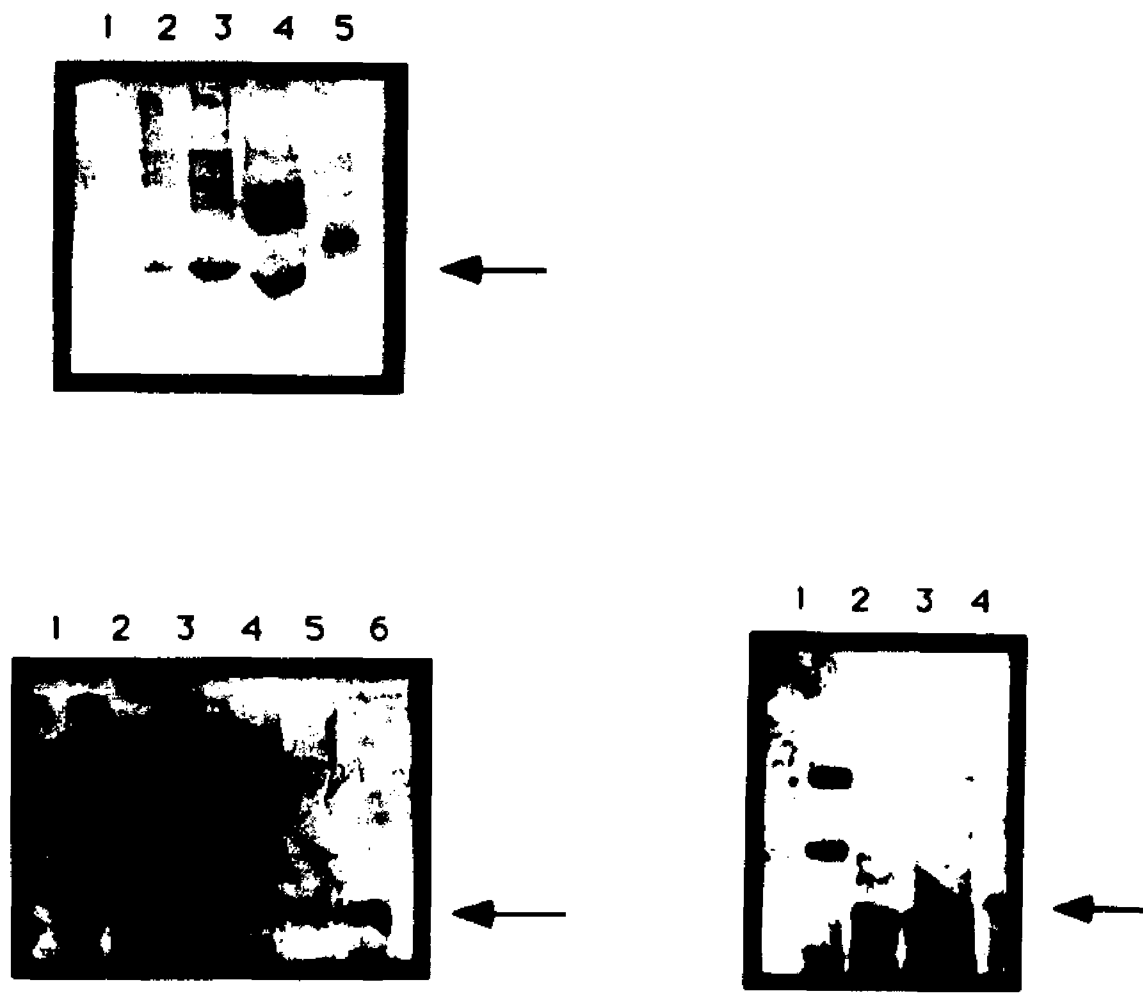


Figure 22. Western blot (A) and silver stain (B,C) analyses of recombinant ALAD in purification scheme

A:
1. PC3 after DEAE
2. PC3 after Sephacryl column
3. PC3 Ammonium Sulfate Precipitate
4. PC3 after lysis
5. Biotinylated Standard

B:
1. Standard
2. PC3 after lysis
3. Ammonium Sulfate Precipitate
4. PC3 after Sephacryl column
5. PC3 after DEAE
6. Marker - Betalactoglobulin

C:
1. Standard
2. PC3 after lysis
3. PC3 Ammonium Sulfate Precipitate
4. PC3 after Sephacryl column

STEPS IN PURIFICATION SCHEME	ALAD CONTENT(units/ml)	PROTEIN CONTENT(mg/ml)	SPECIFIC ACTIVITY(units/mg)
PC3 after lysis	1.916	7.75	0.246
PC3 Ammonium Sulfate Precipitate	8.522	7.75	1.099
PC3 after Sephacryl column	0.709	0.300	2.37

Table 4. Specific activities of recombinant ALAD in purification scheme.

human erythrocyte protein by polyacrylamide gel electrophoresis mobility and Western blotting. The recombinant protein not only eluted at the appropriate fraction from the gel filtration column but also sedimented in a sucrose gradient in a manner expected of a homooctamer, with a subunit size of 35,000 Da.

Thus, an efficient system for synthesis of recombinant human ALAD has been developed. The $ALAD^2$ allele has been cloned into this expression vector and production and purification of large quantities of human ALAD 2-2 has begun. The recombinant human isozymes are to be used to study the differential effects of lead, competing with zinc, on binding to ALAD and on inhibition of ALAD enzymatic activity.

VI. Discussion

A. The ALAD gene

The mechanism and the DNA sequences responsible for the coordinate transcriptional activation of all the heme biosynthesis genes during erythropoiesis are unknown. However, as the genes are isolated and characterized, more information can be elucidated. ALAS is an excellent example of transcriptional activation at a tissue specific level. In this case, there are two genes. The non-erythroid gene is expressed constitutively in all cells and has been designated the nonspecific, or housekeeping, form of ALAS. The non-erythroid promoter contains potential TATA boxes and two potential CAAT boxes, and four GC hexanucleotide boxes are located in the vicinity of the transcription start site (Maguire et al., 1986). The mRNA of the second gene is expressed exclusively in erythroid cells, and since it is located on the X chromosome, it is necessarily under the control of a completely unrelated promoter. However, there are not two genes for every enzyme in the heme biosynthetic pathway.

In the case of PBGD, two distinct mRNAs are formed, one which is only transcribed in erythroid cells. However, this tissue-specific expression is due to alternative splicing of two primary transcripts arising from two promoters (Chretien et al., 1988). One promoter is, therefore, considered a housekeeping promoter and the other is an erythroid-specific promoter. Furthermore, two tissue-specific factors have been found that bind both the PBGD erythroid promoter and the human ß-globin promoter and enhancer (Mignotte et al., 1989). Cis-control regions binding erythroid-specific nuclear factors have been found to be important in globin and non-globin transcriptional activation

(Plumb et al., 1989) The erythroid-specific promoter region contains two Sp1 binding sites and a -230 region consisting of a direct 13 base pair repeat. Neither a classical TATA box nor the CAAT box were found. An ACAACT box was postulated to be the TATA box at -25, and a GGCCAGGCT box was postulated to be the CAAT box at -85 by analogy with the ß-globin CAAT box at -75 with a sequence of GGCCAATCT. The authors proposed that the putative TATA and CAAT boxes, because of their homology to ß-globin promoter elements, were, in fact, promoter elements of PBGD, but may result in lower transcription of PBGD mRNA, due to the lack of total homology. The sequence motif GATAAG is present in the promoter region of mouse α- and ß-globin genes and human PBGD.

UROD is one gene with one promoter which contains a pseudo TATA box and a single Sp1 binding site. Two mRNAs have been found for UROD. Two transcriptional start points were found, a major (90%) and a minor (10%), separated by six base pairs. However, the two initiation starts are used in the same proportion in all tissues tested. The level of UROD mRNA is markedly increased in tissue and cell lines of erythroid origin and is due to tissue specific transcriptional activation of the single UROD gene.

The sequence required for Sp1 binding, (T/G)GGGCGG(GGC/AAT), a factor known to activate transcription in eukaryotic genes, is described by Kadonaga et al.(1986). This region is also sometimes referred to as a GC box and has been found in many housekeeping gene promoters. These Sp1 binding sites or GC rich domains have been found in the two human and one chicken heme biosynthetic genes described so far. In fact, although the UROD gene only has one region, the chicken ALAS has three and

human ALAS gene has four Sp1 elements. The ALAD cluster consists of one binding site with perfect homology over 10 bp, the other three sites being GC rich domains with a high degree of homology to the binding site consensus sequence (Figure 15).

Another interesting characteristic found in many of the promoters described is the presence of clusters of repetitive sequences, unique to each promoter. The chicken ALAS promoter contains a repeat element CCCC(C/T)CATGG, which is reproduced exactly three times between -211 and -180, and a very similar sequence, CCCCTCA, which is repeated twice. In erythroid PBGD, a putative erythroid specific binding site sequence, (C/A)Py(T/A)ATC(T/A)Py, is repeated 3 times at positions -180, -70 and +45. The human ALAD gene also contains its own unique repetitive sequence, CPyGGAGCGGPuPuPu, which is reproduced three times -237, -156 and -33 from the *EagI* site in exon 1A.

The presence of two ALAD transcripts, one erythroid specific and the other constitutive, in the mouse and the presence of two homologous exons 1A and 1B in the human suggest that both human and mouse ALAD gene product is regulated by differential splicing of 5′ untranslated sequence.

As the enzymes of the complete heme biosynthetic pathway are characterized on the molecular level, more information pertaining to regulation will be elucidated. These analyses of the putative promoter sites as well as possible enhancer sites and the binding proteins regulating these genes should provide for an interesting account of coordinated developmental and tissue specific regulation.

B. The ALAD alleles

Reverse-transcription and sequencing of the human ALAD[2] allele from an individual who had the ALAD 2-2 isozyme phenotype revealed a single nucleotide difference with the ALAD[1] sequence, a G to C transversion of coding nucleotide 177, which predicted the substitution of an asparagine for a lysine in residue 59 (K59N) of the enzyme subunit. The replacement of a positively-charged residue by a neutral amino acid would account for the difference in the electrophoretic mobilities of the ALAD 1-1, 1-2 and 2-2 charge isozymes. Since this point mutation created an *MspI* restriction site, a non-radioactive PCR-based method was developed which permitted the precise molecular identification of the ALAD[1] and ALAD[2] alleles using genomic DNA. Since genotyping of individuals whose erythrocyte ALAD isozyme phenotypes were classified as either ALAD 1-2 or 2-2 were unambiguously identified as heterozygous or homozygous for the ALAD[2] allele, the molecular analysis correlated completely with the determination of the isozyme phenotype, indicating that no other mutation studied was responsible for the more electronegative subunit, further demonstrating the reliability of the molecular analysis. Moreover, these findings indicated that the occurrence of another exonic base substitution resulting in the more electronegative migration of the ALAD homooctomer would be rare.

The frequencies of the ALAD[1] and ALAD[2] alleles in a random normal Caucasian population were 0.88 and 0.12, respectively. These results were essentially identical to those previously determined in American and European populations by analysis of the erythrocyte isozyme phenotypes (Battistuzzi et al., 1981; Pettrucci et al., 1982; Benkmann et al., 1983; Astrin et al., 1987).

In addition to man, the ALAD cDNAs for the mouse (Bishop et al., 1989), rat

(Bishop et al., 1986), and *E.coli* (Echelard et al., 1988) enzymes have been isolated and sequenced. The predicted amino acid sequences of the human and rodent enzymes both contain 330 amino acids and are 87% identical, whereas bacterial ALAD has only 37% identity and 53% similarity with the human protein. Interestingly, in the mouse, rat and human sequences, only three predicted amino acids differ from codons 55 to 87, one of which is codon 59, the site of the *Msp*I polymorphism. In fact, the asparagine in position 59 in the human $ALAD^2$ allele also is present in this position in the rat, mouse and bacterial coding sequences. The conservation of this residue across species suggests that $ALAD^2$ was the ancestral human allele. However, the fact that the ALAD 1-2 or 2-2 isozyme phenotype was not detected in an African population (Benkmann et al., 1983) argues against this concept, and implies that $ALAD^2$ allele may have become established in the Caucasian population by selection.

The identification of the *Msp*I site provides a second site which may be used as a sequence tagged site (STS) for the chromosomal region 9q34 as well as a common polymorphism (2pq=0.19) for linkage studies of disease genes mapped to this chromosomal location. The *Msp*I and *Rsa*I polymorphic sites which are located only eight nucleotides apart, are in linkage disequilibrium. However, all four haplotypes were detected, indicating that one of the polymorphisms arose twice or that the fourth haplotype resulted from a rare recombinational or gene conversion event.

The findings that individuals heterozygous or homozygous for the $ALAD^2$ allele have higher blood lead levels when exposed to low or high levels of lead in the environment (Astrin et al., 1987; Ziemsen et al., 1987), suggest that the lead binding and/or stability

of the ALAD 2 subunit is greater than those of the ALAD 1 subunit. Analysis of the predicted amino acid sequence for the ALAD polypeptide reveals the presence of a consensus zinc-binding domain (Wetmur et al., 1986) which may be involved in the active site (Schlösser and Beyersmann, 1987) or in a recently recognized second zinc binding site (Dent et al., 1990). Although residue 59 is not in or adjacent to either of these binding sites, it is possible that the lysine to asparagine substitution in residue 59 could alter the quaternary conformation of the enzyme, thereby increasing the binding affinity for lead. Whatever the physiologic mechanism, the ability to rapidly, non-radioactively and precisely determine the ALAD genotype provides a means to identify individuals who may be at increased risk for lead poisoning due to exposure in the workplace or the environment.

C. Recombinant Human ALAD

Many approaches to the production of recombinant ALAD were pursued. Finally, the polycistronic vector was chosen as the expression vector of choice. Expressed protein from this vector is soluble, unlike the protein product of the pKK-50 vector, and the specific activity was 1000 times the specific activity from unpurified human erythrocytes.

The pKK50 vector expressed aggregates of ALAD. When total cell lysates (before the 6,000 rpm centrifugation step) were analyzed by PAGE and Western transfer, the subunits could be detected with anti-ALAD antibody. These aggregates had activity slightly greater than controls and substantially below the activity of the fusion protein product of the pUC9-50 expression vector. Consequently, the pKK50 vector was not used in further experiments.

A series of pUC9 *lacZ*/ALAD expression vectors were constructed. The pUC9-50-S vector was made by site-directed oligonucleotide mutagenesis which merely introduced an *Sph*I site at the translation initiation site. The vector was made to facilitate subsequent cloning of the ALAD coding sequence into the pKK233-3 vector. The expression vector, pUC9-50-SL was constructed to exclude the 22 amino acids resulting from translation of the ALAD 5′ untranslated region. This vector continued to produce a fusion protein of *lacZ* and ALAD. However, it did not contain the upstream ATG. Expressed protein from this construct continued to react with anti-ALAD antibody and to have enzymatic activity. This observation was consistent with the upstream ATG not being the translation initiation site. The second ATG, with a translation start site which fit the consensus Kozak sequence for translation initiation, coded for a protein which matched the N-terminal amino acid sequence of human erythrocyte and bovine liver ALAD.

The last of the pUC9 expression vectors was the polycistronic vector. This construct contained an adapter sequence which included a termination site for the *lacZ* gene and another ribosome binding site. That expressed product from this vector was no longer a fusion protein of *lacZ* and ALAD was confirmed by the mobility in PAGE and Western blot analysis. The enzymatic activity produced in strains with this vector were as high as found with pUC9-50.

A number of cell lines were examined for maximal expression of ALAD. The original *E. coli* strain used with the pUC9-50 vector was JM103. Although this strain gave high activities, it was not inducible with IPTG. The strain DH5α was used only for plasmid maintenance; activity was low in this strain. Both the LC137 and BL21 strains

gave approximately the same activities with the polycistronic vector.

The BL21 strain contained an additional plasmid, pLysS, which encodes T7 lysozyme which cuts a specific bond in the peptidoglycan layer of the *E. coli* cell wall. Treatments that disrupt the inner membrane, but do not normally cause lysis, induce rapid lysis of these cells. Cells which contain the pLysS plasmid and the pUC9-50-PC vector were easily lysed by freezing and thawing. This step eliminated the incubation steps of the deoxycholate or sarkosyl-octyl glucoside methods of bacterial lysis.

Maximal activity in BL21 cells was obtained with growth in LB medium in the presence of chloramphenicol and ampicillin for 20 hr with shaking at 37°C. Induction with IPTG began at inoculation. Many other variations on these conditions were examined, including induction at various stages of the bacterial growth cycle followed by continued growth for varied time periods. Any of these modifications resulted in lower ALAD activities.

Purification of the enzyme was based on previous work with human erythrocytes (Anderson and Desnick, 1979). The ammonium sulfate precipitation was useful to concentrate cell lysate into a volume suitable for loading onto the gel filtration column, as well as for storage purposes. Cell lysates could be stored as an ammonium sulfate pellet at -20°C without significant loss of ALAD activity.

Aliquots from the gel filtration column were immediately frozen at -20°C in 50% glycerol. One ml aliquots were taken for ALAD activity and Bradford protein quantitation assays and for Western blot and silver staining analysis. Samples containing high ALAD activity within the peak of CRIM positive protein were pooled and stored for

further purification steps. The protein is approximately 50% pure at this stage of purification.

The growth and purification schemes described will also be used to purify the $ALAD^2$ gene product. Purified ALAD 1-1 will be used to study the crystal structure. Both ALAD polymorphic enzymes will be used for biochemical properties, protease inhibitory function, and to investigate the differential effects of lead binding.

VII. Concluding Remarks

A. ALAD genomic sequence

The entire genomic sequence has now been completed in both directions. Repetitive sites, such as Alu sequences and potential promoter elements have been identified. The mechanism and the DNA sequences responsible for the coordinate transcriptional activation of all the heme biosynthesis genes during erythropoiesis are unknown. However, as the genes are isolated and characterized, more information can be elucidated. In an effort to learn more about the developmental regulation of heme, this study concentrated on the characteristics of the ALAD genomic sequence and promoter elements, including cis-control elements which may be important in transcriptional activation.

Additional studies will include proving that these putative promoter elements are important in transcriptional regulation. The promoter will need to be examined by CAT expression and deletion analysis in an inducible erythropoietic cell line. The length and cap site of the message must be determined in both erythroid and non-erythroid sources.

B. ALAD polymorphism

The mutation which causes the charged isozyme polymorphism in ALAD 2-2 individuals has been established and confirmed by a population study. A PCR-based ALAD genotype test has been described which easily distinguishes the presence or absence of the $ALAD^2$ allele from a small blood sample. Understanding the effect of ALAD genotype on both acute and chronic measures of high and low-level lead exposure will provide the basis for identifying genetically susceptible individuals with an increased

health risk if exposed to lead in the workplace or environment.

C. ALAD proteins

Large quantities of recombinant human ALAD 1-1 have been purified from bacterial cells and the cloning and expression of the recombinant human ALAD 2-2 is underway. High concentrations of recombinant ALAD 1-1 and 2-2 will be used to carry out studies of (a) the kinetics and equilibria of lead binding and (b) the kinetics and extent of lead inhibition of ALAD activity.

Competition assays will be carried out by equilibrium dialysis where both ALAD isozymes are simultaneously exposed to radioactive lead (^{210}Pb) or zinc (^{65}Zn) as a function of time, pH, and the concentrations of free zinc, free lead, sulfhydryl reagents, and serum (Jaffe et al., 1991). This study is crucial to proving the hypothesis that the ALAD 2 subunit binds lead tighter than the ALAD 1 subunit, leading to the observed pharmacogenetics of the ALAD polymorphic locus.

The completion of the genomic sequence, identification of putative promoter elements, the RFLP analysis of both the $ALAD^2$ allele (MspI polymorphism) and RsaI polymorphism, the development of PCR-based ALAD genotype test and the production of recombinant enzyme have all contributed to achieving the goals of understanding the regulation of ALAD and the effect of ALAD genotype on both acute and chronic measures of high and low-level lead exposure, genetic toxicology of lead poisoning.

VIII. Bibliography

Anderson, K. E., Sassa, S., Peterson, C. M., and Kappas, A. (1977). Increased erythrocyte uroporphyrinogen-I-synthase, delta-aminolevulinic acid dehydratase and protoporphyrin in hemolytic anemia. *Am. J. Med.* **63**, 359.

Anderson, P. M., and Desnick, R. J. (1979). Purification and properties of δ-aminolevulinate dehydratase from human erythrocytes. *J. Biol. Chem.* **254**, 6924-6930.

Astrin, K. H., Bishop, D. F., Wetmur, J. G., Kaul, B., Davidow, B., and Desnick, R. J. (1987), δ-aminolevulinic acid dehydratase isozymes and lead toxicity. In Mechanisms of Chemically-Induced Porphyrinopathies, Silbergeld, E.K., and Fowler, B.A., Eds. (Vol. 514: *Ann. N. Y. Acad. Sci.*), 23-29.

Astrin, K. H., Kaya, A. H., Wetmur, J. G., and Desnick, R. J. (1991). *Rsa*I polymorphism in the human delta-aminolevulinate dehydratase gene at 9q34. *Nucleic Acids Research* **19**, 4307.

Audesirk, G. (1985). Effects of lead exposure on the physiology of neurons. *Prog. Neurobiol.* **24**, 199-231.

Baker, T. A., Grossman, A. D., and Gross, C. A. (1984). A gene regulating the heat shock response in *Esherichia coli* also affects proteolysis. *Proc. Natl. Acad. Sci. U. S. A.* **21**, 6779-6783.

Barnard, G. F., Itoh, R., Hohberger, L. H., and Shemin, D. (1977). Mechanism of porphobilinogen synthase. *J. Biol. Chem.* **252**, 8965-8974.

Battersby, A. R. and McDonald, E. (1975). In: Porphyrins and Metalloporphyrins, Smith, K.H., ed., Elsevier, Amsterdam, 61.

Battistuzzi, G., Petrucci, R., Silvagni, L., Urbani, F. R., and Caiola, S. (1981). δ-aminolevulinate dehydrase: A new genetic polymorphism in man. *Ann. Hum. Genet.* **45**, 223-229.

Bawden, M. J., Borthwick, I. A., Healy, H. M., Morris, C. P., May, B. K., and Elliot, W. H. (1987). Sequence of the 5-aminolevulinate synthase cDNA. *Nucleic Acids Res.* **15**, 8563.

Beale, S. I. (1970). The biosynthesis of delta-aminolevulinic acid in chlorella. *Plant Physiol.* **45**, 504-506.

Bellinger, D., Leviton, A., Waternaux, C., Needleman, H., and Rabinowitz, M. (1987). Longitudinal analysis of prenatal and postnatal lead exposure in early cognitive development. *N. Engl. J. Med.* **316**, 1037-1043.

Ben-Ezzer, J., Oelsner, H., and Szeinberg, A. (1987). Genetic polymorphism of delta-aminolevulinate dehydrase in several population groups in Israel. *Hum. Hered.* **37**, 229-232.

Benkmann, H.-G., Bogdanski, P., and Goedde, H. W. (1983). Polymorphism of delta-aminolevulinic acid dehydratase in various populations. *Hum. Hered.* **33**, 62-64.

Berg, J. M. (1986). More metal-binding fingers. *Nature* **319**, 264-265.

Bergomi, M., Borella, P., Fantuzzi, G., Vivoli, G., Sturloni, N., Cavazzuti, G., Tampieri, A., and Tartoni, P. L. (1989). Relationship between lead exposure indicators and neuropsychological performance in children. *Dev. Med. Child. Neurol.* **31**, 181-190.

Bevan, D. R., Bodlaender, P., and Shemin, D. (1980). Mechanism of porphobilinogen synthase. Requirement for Zn^{2+} for enzyme activity. *J. Biol. Chem.* **255**, 2030-2035.

Bishop, D. F. (1990). Two different genes encode delta-aminolevulinate synthase in humans: nucleotide sequences of cDNA for the housekeeping and erythroid genes. *Nucleic Acids Res.* **18**, 7187.

Bishop, T. R., Cohen, P. J., Boyer, S. H., Noyes, A. N., and Frelin, L. P. (1986). Isolation of a rat liver δ-aminolevulinate dehydratase (ALAD) cDNA clone: Evidence for unequal ALAD gene dosage among inbred mouse strains. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5568-5572.

Bishop, T. R., Frelin, L. P., and Boyer, S. H. (1986). Nucleotide sequence of rat liver δ-aminolevulinic acid dehydratase cDNA. *Nucleic Acids Res.* **14**, 10115.

Bishop, T. R., Frelin, L. P., and Boyer, S. H. (1990). Isolation and characterization of a rat δ-aminolevulinate dehydratase processed pseudogene. *Genomics* **7**, 629-632.

Bishop, T. R., Hodes, Z. I., Frelin, L. P., and Boyer, S. H. (1989). Cloning and sequence of mouse erythroid δ-aminolevulinate dehydratase cDNA. *Nucleic Acids Res.* **17**, 1775.

Bishop, T. R., Miller, M. W., Frelin, L. P., Dierks, P., and Boyer, S. H. (1991). The gene encoding the second enzyme of the heme biosynthetic pathway, aminolevulinate dehydratase, has two transcriptional promoters. (abstract) International Congress of Human Genetics, Washington D.C.

deBoer, E., Antoniou, M., Mignotte, V., Wall, L., and Grosveld, F. (1988). The human beta-globin promoter; nuclear protein factors and erythroid specific induction of transcription. *EMBO J.* **7**, 4203-4212.

Borthwick, I. A., Srivastava, G., and Day, A. R. (1986). Complete nucleotide sequence of hepatic 5-aminolevulinate synthase. *Eur. J. Biochem.* **150**, 480-484.

Borthwick, I. A., Srivastava, G., Day, A. R., Pirola, B. A., Snoswell, M. A., May, B. K., and Elliot, W. (1985). Complete nucleotide sequence of hepatic 5-aminolevulinate synthase precursor. *Eur. J. Biochem.* **150**, 480-484.

Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction and genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**, 314-331.

Bottomley, S. S., and Muller-Eberhard, U. (1988). Pathophysiology of heme synthesis. *Semin. Hematol.* **25**, 282-302.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* **72**, 248-254.

Brennan, M. J., and Cantrill, R. C. (1979). Delta-aminolaevulinic acid is a potent agonist for GABA autoreceptors. *Nature* **280**, 514-515.

Brennan, M. J., and Cantrill, R. C. (1981). Delta-aminolevulinic acid and amino acid neurotransmitters. *Mol. Cell. Biochem.* **38**, 49-58.

Brenner, D. A., and Frasier, F. (1991). Cloning of murine ferrochelatase. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 849-853.

Brosius, J., and Holy, A. (1984). Regulation of ribosomal RNA promoters with a synthetic lac operator. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 6929-6933.

Chang, A. C. Y., and Cohen, S. W. (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* **134**, 1141.

Chang, C. S., and Sassa, S. (1984). Induction of δ-aminolevulinate dehydratase in mouse Friend virus-transformed erythroleukemia cells during erythroid differentiation. *Blood* **64**, 64-70.

Chang, C. S., Sassa, S., and Doyle, D. (1987). An immunological study of delta-aminolevulinic acid dehydratase: specificity consistent with the phylogeny of species. *Biochim. Biophys. Acta.* **797**, 297.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299.

Chretien, S., Dubart, A., Beaupain, D., Raich, N., Grandchamp, B., Rosa, J., Goossens, M., and Romeo, P. H. (1988). Alternative transcription and splicing of the human porphobilinogen deaminase gene result either in tissue-specific or in housekeeping expression. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6-10.

Coleman, D. L. (1966). Purification and properties of delta-aminolevulinate dehyrdatase from two tissues of two strains of mice. *J. Biol. Chem.* **241**, 5511-5517.

Cornish-Bowden, A. (1974). A simple graphical method for determining the inhibition constants of mixed, uncompetitive and non-competitive inhibitors. *Biochem. J.* **137**, 143-144.

Cox, T. C., Bawden, M. J., Martin, A., and May, B. K. (1991). Human erythroid 5-aminolevulinate synthase: promoter analysis and identification of an iron responsive element in mRNA. *EMBO J.* **10**, 1891-1892.

Cutler, M. G., McLaughlin, M., McNeil, E., and Moore, M. R. (1985). Effects of delta-aminolaevulinic acid on contractile activity in the isolated small intestine of the rabbit. *Neuropharmacol.* **24**, 1005-1009.

Dandekar, T., Stripecke, R., Gray, N. K., Goosen, B., Constable, A., Johansson, H. E., and Hentze, M. W. (1991). Identification of a novel iron-responsive element in murine and human erythroid delta-aminolevulinic acid synthase mRNA. *EMBO J.* **10**, 1903-1909.

Deknudt, G., Manuel, Y., and Gerber, G. B. (1977). Chromosomal aberrations in workers professionally exposed to lead. *J. Toxicol. Environ. Health* **3**, 885-891.

Dent, A. J., Beyersmann, D., Block, C., and Hasnain, S. S. (1990). Two different zinc sites in bovine 5-aminolevulinate dehydratase distinguished by extended X-ray absorption fine structure. *Biochemistry* **29**, 7822-7828.

Dixon, M., Webb, E. C., Thorne, C. J. R., and Tipton, K. F. (1979). Enzymes, 3rd ed. Chapter VII, Academic Press, New York, 332-381.

Doss, M., Becker, U., Sixel, F., Geisse, S., Solcher, H., Schneider, J., Kufner, G., Schlegel, H., and Stoeppler, M. (1982). Persistent protoporphyrinemia in hereditary porphobilinogen synthase (δ-aminolevulinic acid dehydratase) deficiency under low lead exposure. *Klin. Wochenschr.* **60**, 599-606.

Doss, M., Benkmann, H.-G., and Goedde, H. W. (1986). delta-Aminolevulinic acid dehydrase (porphobilinogen synthase) in two families with inherited enzyme deficiency. *Clin. Genet.* **30**, 191-198.

Doss, M., Laubenthal, F., and Stoeppler, M. (1984). Lead poisoning in inherited δ-aminolevulinic acid dehydratase deficiency. *Int. Arch. Occup. Environ. Health* **54**, 55-63.

Doss, M., von Tiepermann, R., and Schneider, J. (1980). Acute hepatic porphyria syndrome with porphobilinogen synthase defect. *Int. J. Biochem.* **12**, 823-826.

Doss, M., von Tiepermann, R., Schneider, J., and Schmid, H. (1979). New type of hepatic porphyria with porphobilinogen synthase defect and intermittent acute clinical manifestation. *Klin. Wochenschr.* **57**, 1123-1127.

Doyle, D., and Schimke, R. T. (1969). The genetic and developmental regulation of hepatic delta-aminolevulinate dehydratase in mice. *J. Biol. Chem.* **224**, 5449-5459.

Echelard, Y., Dymetryszyn, J., Drolet, M., and Sasarman, A. (1988). Nucleotide sequence of the *hem*B gene of *Escherichia coli* K12. *Mol. Gen. Genet.* **214**, 503-508.

Farant, J. P., and Wigfield, D. C. (1987). Interaction of divalent metal ions with normal and lead-inhibited human erythrocytic porphobilinogen synthase in vitro. *Toxicol. Appl. Pharmacol.* **89**, 9-18.

Finelli, V. N., Klauder, D. S., Karaffa, M. A., and Petering, H. G. (1975). Interaction of zinc and lead on δ-aminolevulinate dehydratase. *Biochem. Biophys. Res. Commun.* **65**, 303-311.

Frankel, S., Condeelis, J., and Leinwand, L. (1990). Expression of actin in *Escherichia coli*. Aggregation, solubilization, and function analysis. *J. Biol. Chem.* **265**, 17980-17987.

Frankel, S., Sohn, R., and Leinwand, L. (1991). The use of sarkosyl in generating soluble protein after bacterial expression. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 1192-1196.

Fujita, H., Sassa, S., Lundgren, J., Holmberg, L., Thunell, S., and Kappas, A. (1987). Enzymatic defect in a child with hereditary hepatic porphyria due to homozygous δ-aminolevulinic acid dehydratase deficiency: Immunochemical studies. *Pediatr.* **80**, 880-885.

Fujita, H., Sato, K., and Sano, S. (1982). Increase in the amount of erythrocyte δ-aminolevulinic acid dehydratase in workers with moderate lead exposure. *Int. Arch. Occup. Environ. Health* **50**, 287-297.

Geisse, S., Brüller, H.-J., and Doss, M. (1983). Porphobilinogen synthase (δ-aminolevulinic acid dehydratase) activity in human erythrocytes: Reactivation by zinc and dithiothreitol depending on influence of storage. *Clin. Chim. Acta* **135**, 239-245.

Gellerfors, P. L., Saltzgaber-Muller, J., and Douglas, M. (1986). Selection by genetic complementation and characterization of the gene coding for the yeast porphobilinogen deaminase. *Biochem. J.* **240**, 673-677.

Gibbs, P. N. B., Gore, M. G., and Jordan, P. M. (1985). Investigation of the effect of metal ions on the reactivity of thiol groups in human 5-aminolaevulinate dehydratase. *Biochem. J.* **225**, 573-580.

Gibbs, P. N. B., and Jordan, P. M. (1981). 5-aminolevulinate dehydratase: Zinc-65 binding and exchange with the enzyme from human erythrocytes. *Biochem. Soc. Trans.* **9**, 232-233.

Gibbs, P. N. B., and Jordan, P. M. (1986). Identification of lysine at the active site of human 5-aminolevulinate dehydratase. *Biochem. J.* **236**, 447-451.

Goering, P. L., and Fowler, B. A. (1987). Regulatory roles of high-affinity metal-binding proteins in mediating lead effects on δ-aminolevulinic acid dehydratase. In Mechanisms of Chemically-Induced Porphyrinopathies, Silbergeld, E.K., and Fowler, B.A., Eds. (Vol. 514: *Ann. N. Y. Acad. Sci.*), 235-247.

Gorman, C. M., Moffat, L. F., and Howard, B. H. (1982). Recombinant genomes which express chloramphenicol acetyl-transferase in mammalian cells. *Mol. Cell. Biol.* **2**, 1044-1051.

Grandchamp, B., Beaumont, C., de Verneuil, H., and Nordmann, Y. (1985). Accumulation of porphobilinogen deaminase, uroporphyrinogen decarboxylase, and alpha- and beta-globin mRNAs during differentiation of mouse erythroleukemic cells. Effects of succinylacetone. *J. Biol. Chem.* **260**, 9630-9635.

Grandchamp, B., and Nordmann, Y. (1988). Enzymes of the heme biosynthesis pathway: recent advances in molecular genetics. *Semin. Hematol.* **25**, 303-311.

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

Granick, S., and Sassa, S. (1978). Hemin control of heme biosynthesis in mouse Friend-virus transformed leukemic cells in culture. *J. Biol. Chem.* **253**, 5402-5406.

Grodberg, J., and Dunn, J. J. (1988). ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J. Bacteriol.* **170**, 1245-1253.

Hasnain, S. S., Wardell, E. M., Garner, C. D., Schlösser, M., and Beyersmann, D. (1985). Extended X-ray-absorption-fine-structure investigations of zinc in 5-aminolevulinate dehydratase. *Biochem. J.* **230**, 625-633.

Hassoun, A., Verstraeten, L., Mercelis, R., and Martin, J.-J. (1989). Biochemical diagnosis of an hereditary aminolaevulinate dehydratase deficiency in a 63-year-old man. *J. Clin. Chem. Clin. Biochem.* **27**, 781-786.

Hechemy, K. E., Harris, H. L., Wethers, J. A., Stevens, R. W., Stock, B. R., Reilly, A. A., and Benach, J. F. (1989). Flouroimmunoassay studies with solubilized antigens from Borrelia burgdorfer. *J. Clin. Microbiol.* **27**, 1854-1858.

Higuchi, R. (1989), Simple and rapid preparation of samples for PCR. In PCR technology: Principles and applications for DNA amplification, Erlich, H.A., Ed. Stockton Press, New York, 31-38.

Hoffman, R., Ibrahim, N., Murnane, M. J., Diamond, A., Forget, B. G., and Levere, R. D. (1980). Hemin control of biosynthesis and catabolism in a human leukemia cell line. *Blood* **56**, 567-570.

Inouye, M., and Yee, M. L. (1973). Homogeneity of envelope proteins of *Escherichia coli* separated by gel electrophoresis in sodium dodecyl sulfate. *J. Bacteriol.* **113**, 304-312.

Irving, R. A., Mainwaring, W. I. P., and Spooner, P. M. (1978). The regulation of haemglobin synthesis in cultured chick blastoderms by steroid related to 5beta-androstane. *Biochem. J.* **158**, 8.

Jaffe, E. K., Bagla, S., and Michini, P. A. (1991). Reevaluation of a sensitive indicator of early lead exposure. Measurement of porphobilinogen synthase in blood. *Biol. Trace Element Res.* **28**, 223-231.

Jaffe, E. K., and Hanes, D. (1986). Dissection of the early steps in the porphobilinogen synthase catalyzed reaction: requirements for Schiff's base formation. *J. Biol. Chem.* **261**, 9348-9353.

Jaffe, E. K., and Markham, G. D. (1988). ^{13}C NMR studies of methylene and methine carbons of substrate bound to a 280,000-dalton protein, porphobilinogen synthase. *Biochemistry* **27**, 4475-4481.

Jaffe, E. K., Markham, G. D., and Rajagopalan, J. S. (1990). ^{15}N and ^{13}C NMR studies of ligands bound to the 280,000-dalton protein porphobilinogen synthase elucidate the structures of enzyme-bound product and a Schiff base intermediate. *Biochemistry* **29**, 8345-8350.

Jaffe, E. K., Salowe, S. P., Chen, N. T., and DeHaven, P. A. (1984). Porphobilinogen synthase modification with methylmethanethiosulfonate. *J. Biol. Chem.* **259**, 5032-5036.

Jordan, P. M., and Gibbs, P. N. B. (1985). Mechanism of action of 5-aminolevulinate dehydratase from human erythrocytes. *Biochem. J.* **227**, 1015-1020.

Jordan, P. M., and Seehra, J. S. (1980). ^{13}C NMR as a probe for the study of enzyme-catalysed reactions: mechanism of action of 5-aminolevulinic acid dehydratase. *FEBS-Lett* **114**, 283-286.

Kappas, A., Sassa, S., Galbraith, R. A., Nordmann, Y. (1989).The prophyrias. In: The metabolic basis of inherited disease, 6th ed, McGraw Hill, New York, 1305-1365.

Kadonaga, J. T., and Tijan, R. (1986). Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5889-5893.

Kaul, B., Davidov, B., Eng, Y. M., and Gewirtz, M. H. (1983). Lead, erythrocyte protoporphyrin, and ferritin levels in cord blood. *Arch. Environ. Health.* **38**, 296-300.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**, 8125-8148.

Levere, R. D., and Granick, S. (1967). Control of hemoglobin synthesis in the cultured chick blastoderm. *J. Biol. Chem.* **242**, 1903-1911.

Li, J.-M., Umanoff, H., Proenca, R., Russell, C. S., and Cosloy, S. D. (1988). Cloning of the *Escherichia coli* K-12 *hemB* gene. *J. Bacteriol.* **170**, 1021-1025.

Lippmann, M. (1990). Lead and human health: Background and recent findings. *Environ. Res.* **51**, 1-24.

Maguire, D. J., Day, A. R., Borthwick, I. A., Srivastava, G., Wisley, P. L., May, B. K., and Elliot, W. H. (1986). Nucleotide sequence of the chicken 5-aminolevulinate synthase gene. *Nucleic Acids Res.* **14**, 1379-1391.

Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987). Regulation of inducible and tissue-specific gene expression. *Science* **236**, 1237-1244.

Marmur, J. (1961). A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J. Mol. Biol.* **3**, 208-218.

McMichael, A. J., Baghurst, P. A., Wigg, N. R., Vimpani, G. V., Robertson, E. F., and Roberts, R. J. (1988). Port Pirie cohort study: Environmental exposure to lead and children's abilities at the age of four years. *N. Engl. J. Med.* **319**, 468-475.

Mignotte, V., Wall, L., deBoer, E., Grosveld, F., and Romeo, P. H. (1989). Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. *Nucleic Acids Res.* **17**, 37-54.

Miller, J., MacLachlan, A. D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**, 1609-1614.

Minnema, D. J., and Michaelson, I. A. (1986). Differential effects of inorganic lead and delta-aminolevulinic acid in vitro on synaptosomal gamma-aminobutyric acid release. *Toxicol. Appl. Pharmacol.* **86**, 437-447.

Moffatt, B. A., and Studier, F. W. (1987). T7 lysozyme inhibits transciption by T7 RNA polymerase. *Cell* **49**, 221-227.

Muller, W. E., and Snyder, S. H. (1977). δ-Aminolevulinic acid: Influences on synaptic GABA receptor binding may explain CNS symptoms of porphyria. *Ann. Neurol.* **2**, 340-342.

Murakami, K., and Etlinger, J. D. (1986). Endogenous inhibitor of nonlysosomal high molecular weight protease and calcium-dependent protease. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 7588-7592.

Nandi, D. L., and Shemin, D. (1968). Delta-aminolevulinic acid dehydratase of Rhodopseudomonas spheroides. 3. Mechanism of porphobilinogen synthesis. *J. Biol. Chem.* **243**, 1236-1242.

Needleman, H. L., Gunnoe, C., Leviton, A., Reed, R., Peresie, H., Maher, C., and Barrett, P. (1979). Deficits in psychologic and classroom performance of children with elevated dentine lead levels. *N. Engl. J. Med.* **300**, 689-695.

Needleman, H. L., Rabinowitz, M., and Leviton, A. (1984). The relationship between prenatal exposure to lead and congenital anomalies. *J. Amer. Med. Assoc.* **251**, 2956-2959.

Needleman, H. L., Schell, A., Bellinger, D., Leviton, A., and Allred, E. N. (1990). The long-term effects of exposure to low doses of lead in childhood. *N. Engl. J. Med.* **322**, 83-88.

Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Ramírez-Solis, R., Webster, T. D., Ledbetter, D. H., and Caskey, C. T. (1989). Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 6686-6690.

Petrucci, R., Leonardi, A., and Battistuzzi, G. (1982). The genetic polymorphism of human δ-aminolevulinate dehydratase in Italy. *Hum. Genet.* **60**, 289-290.

Piomelli, S., Seaman, C., Zullow, D., Curran, A., and Davidow, B. (1982). Threshold for lead damage to heme biosynthesis in urban children. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 3335-3339.

Plumb, M., Frampton, J., Wainwright, H., Walker, M., Macloed, K., Goodwin, G., and Harrison, P. (1989). GATAAG; a cis-control region binding an erythroid-specific nuclear factor with a role in globin and non-globin gene expression. *Nucleic Acids Res.* **17**, 73-92.

Potluri, V. R., Astrin, K. H., Wetmur, J. G., Bishop, D. F., and Desnick, R. J. (1987). Human δ-aminolevulinate dehydratase: Chromosomal localization to 9q34 by *in situ* hybridization. *Hum. Genet.* **76**, 236-239.

Radin, N. S., Rittenberg, D., and Shemin, D. (1950). The role of glycine in the biosynthesis of heme. *Journal of Biological Chemistry* **184**, 745.

Raich, N., Mignotte, V., Dubart, A., Beaupain, D., Leboulch, P., Romana, M., Chabret, C., Charnay, P., Papayannopoulou, T., and Goosens, M. (1989). Regulated expression of the overlapping ubiquitous and erythroid transcription units of the human porphobilinogen deaminase (PBG-D) gene introduced into non-erythroid and erythroid cells. *J. Biol. Chem.* **264**, 10186-10192.

Raich, N., Romeo, P. H., Dubart, A., Beaupain, D., Cohen-Solal, M., and Goossens, M. (1986). Molecular cloning and complete primary sequence of human erythrocyte PBG deaminase. *Nucleic Acids Res.* **14**, 5955-5968.

Reddy, E. P. (1983). Nucleotide sequence analysis of the T24 human bladder carcinoma oncogene. *Science* **220**, 1061-1063.

Riddle, R. D., Yamamoto, M., and Engel, J. D. (1989). Expression of delta-aminolevulinate synthase in avian cells: Separate genes encode erythroid-specific and non-specific isozymes. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 792-796.

Romana, M., Dubart, A., Beaupain, D., Chabret, C., Goossens, M., and Romeo, P. H. (1987). Structure of the gene for the human uroporphyrinogen decarboxylase. *Nucleic Acids Res.* **15**, 7343-7356.

Romeo, P.-H., Dubert, A., Grandchamp, B., de Verneuil, H., Rosa, J., Nordmann, Y., and Goossens, M. (1984). Isolation and identification of a cDNA clone coding for rat uroporphyrinogen decarboxylase. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 3346-3350.

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

Roth, K. S., Spencer, P. D., Moses, L. C., and Carter, B. E. (1990). delta-Aminolevulinic acid dehydratase: is there a form unique to renal cortex? *Enzyme* **43**, 17-25.

Rutherford, T., Thompson, G. G., and Moore, M. R. (1979). Heme biosynthesis in Friend erythroleukemia cells: Control by ferrochelatase. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 833-836.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular cloning, a laboratory manual, 2nd ed., Cold Spring Harbor University Press, Cold Spring Harbor.

Sanger, F. (1981). Determination of nucleotide sequences in DNA. *Science* **214**, 1205-1210.

SAS Institute (1985). SAS Users Guide Statistics Version 5. Cary, NC.

Sassa, S. (1976). Sequential induction of heme pathway enzymes during erythroid differentiation of mouse Friend leukemia virus-infected cells. *J. Exp. Med.* **143**, 305-315.

Sassa, S. (1982). Delta-aminolevulinic acid dehydratase assay. *Enzyme* **28**, 133.

Sassa, S., and Bernstein, S. E. (1977). Levels of δ-aminolevulinate dehydratase, uroporphyrinogen I synthase, and protoporphyrin IX in erythrocytes from anemic mutant mice. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 1181-1184.

Sassa, S., Bradlow, H. L., and Kappas, A. (1979). Steroid induction of δ-aminolevulinic acid synthase and porphyrins in liver: Structure-activity studies of the permissive effects of hormones on the induction process. *J. Biol. Chem.* **254**, 10011-10020.

Sassa, S., Granick, S., Bickers, D. R., Levere, R. D., and Kappas, A. (1973). Studies on the inheritance of human erythrocyte delta-aminolevulinate dehydratase and uroporphyrinogen synthase. *Enzyme* **16**, 326-333.

Sassa, S., and Kappas, A. (1983). Hereditary tyrosinemia and the heme biosynthetic pathway. Profound inhibition of delta-aminolevulinic acid dehydratase activity by succinylacetone. *J. Clin. Invest.* **71**, 625-634.

Schlick, E., Mengel, K., and Friedberg, K. D. (1983). The effect of low lead doses in vitro and in vivo on the d-ala-d activity of erythrocytes, bone marrow cells, liver and brain of the mouse. *Arch. Toxicol.* **53**, 193-205.

Schoenhaut, D. S. and Curtis, P. J. (1986). Nucleotide sequence of mouse 5-aminolevulinic acid synthase cDNA and expression of its gene in hepatic and erythroid tissues. *Gene* **48**, 55-63.

Schoenhaut, D.S. and Curtis, P.J. (1989). Structure of a mouse erythroid 5-aminolevulate synthase gene and mapping of erythroid-specific DNAseI hypersensitive sites. *Nucleic Acids Research* **17**, 7013-28.

Scopes, R. K. (1987). Protein purification. Principles and practice, 2nd ed., Springer-Verlag, New York.

Seppäläinen, A. M., Hernberg, S., Vesanto, R., and Kock, B. (1983). Early neurotoxic effects of occupational lead exposure: A prospective study. *Neurotoxicol.* **4**, 181-192.

Shemin, D., and Rittenberg, D. (1946). The biological utilization of glycine for the synthesis of the protoporphyrin of hemoglobin. *Journal of Biological Chemistry* **166**, 621.

Shemin, D. (1972). Delta-aminolevulinic acid dehydratase, Enzymes, 3rd ed. Volume VII. Academic Press, New York.

Somervaille, L. J., Nilsson, U., Chettle, D. R., Tell, I., Scott, M. C., Schütz, A., Mattsson, S., and Skerfving, S. (1989). *In vivo* measurements of bone lead - A comparison of two x-ray fluorescence techniques using three different bone sites. *Phys. Med. Biol.* **34**, 1833-1845.

Srivastava, G., Borthwick, I. A., McGuire, D. J., Elferink, C. J., Bawden, M. J., Mercer, J. F., and May, B. K. (1988). Regulation of 5-aminolevlinate synthase mRNA in different rat tissue. *J. Biol. Chem.* **263**, 5202-5209.

Stephens, J. K., Fischer, P. W., and Marks, G. S. (1977). Porphyrin induction: equivalent effects of 5alphaH and 5betaH steroids in chick embryo liver cells. *Science* **197**, 659-660.

Stubnicer, A. C., Picat, C., and Grandchamp, B. (1988). Rat porphobilinogen cDNA: nucleotide sequence of the erythropoietic form. *Nucleic Acids Res.* **16**, 3102.

Studier, F. W., and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expresssion of cloned genes. *J. Mol. Biol.* **189**, 113-130.

Sutherland, G. R., Baker, E., Callen, D. F., Hyland, V. J., May, B. K., Bawden, M. J., Healy, H. M., and Borthwick, I. A. (1988). 5-Aminolevulinate synthase is at 3p21 and thus not the primary defect in X-linked siderblastic anemia. *Am. J. Hum. Genet.* **43**, 331-335.

Thomas, S. D., and Jordan, P. M. (1986). Nucleotide sequence of the hemC locus encoding porphobilinogen deaminase of *Esherichia coli* K12. *Nucleic Acids Res.* **14**, 6215-6225.

Thunell, S., Holmberg, L., and Lundgren, J. (1987). Aminolaevulinate dehydratase porphyria in infancy. A clinical and biochemical study. *J. Clin. Chem. Clin. Biochem.* **25**, 5-14.

Tomokuni, K., and Ichiba, M. (1988). Comparison of inhibition of erythrocyte pyrimidine 5'-nucleotidase and δ-aminolevulinic acid dehydratase by lead. *Toxicol. Lett.* **40**, 159-163.

Tsai, S. F., Bishop, D. F., and Desnick, R. J. (1988). Human porphobilinogen III synthase: molecular cloning, nucleotide sequence, and expression of a full length cDNA. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 7049-7053.

Tsai, S. F., Martin, D. I. K., Zon, L. I., D'Andrea, A. D., Wong, G. G., Orkin, S. H. (1989). Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* **339**, 446-451.

Tschudy, D. P., Hess, R. A., and Frykholm, B. C. (1981). Inhibition of delta-aminolevulinic acid dehydratase by 4,6-dioxoheptanoic acid. *J. Biol. Chem.* **256**, 9915-9923.

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

Tsukamoto, I., Yoshinaga, T., and Sano, S. (1980). Zinc and cysteine residues in the active site of bovine liver δ-aminolevulinic acid dehydratase. *Int. J. Biochem.* **12**, 751-756.

Urabe, A., Sassa, S., and Kappas, A. (1979). The influence of steroid hormone metabolism on in vitro development of erythroid colonies derived from human bone marrow. *J. Exp. Med.* **149**, 1314-1325.

de Verneuil, H., Doss, M., Brusco, N., Beaumont, C., and Nordmann, Y. (1985). Hereditary hepatic porphyria with δ-aminolevulinate dehydratase deficiency: immunological characterization of the non-catalytic enzyme. *Hum. Genet.* **69**, 174-177.

Vieira, J., and Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259-268.

Wall, L., deBoer, E., and Grosveld, F. (1988). The human beta-globin gene 3'enhancer contains multiple binding sites for an erythroid-specific protein. *Genes. Dev.* **2**, 1089-1100.

Wallenstein, S., Zucker, C. L., and Fleiss, J. L. (1980). Some statistical methods useful in circulation research. *Circulation Res.* **47**, 1-9.

Weissberg, J. B., Lipschutz, F., and Oski, F. A. (1971). Delta-aminolevulinic acid dehydratase activity in circulating blood cells. A sensitive laboratory test for the detection of childhood lead poisoning. *N. Engl. J. Med.* **284**, 565-569.

Wetmur, J. G., Bishop, D. F., Cantelmo, C., and Desnick, R. J. (1986). Human δ-aminolevulinate dehydratase: Nucleotide sequence of a full-length cDNA clone. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 7703-7707.

Wetmur, J. G., Bishop, D. F., Ostasiewicz, L., and Desnick, R. J. (1986). Molecular cloning of a cDNA for human delta-aminolevulinate dehydratase. *Gene* **43**, 123-130.

Wetmur, J. G., Kaya, A.H., Plewinska, M. and Desnick, R. J. (1991). Molecular Characterization of the Human delta-Aminolevulinate Dehydratase 2 ($ALAD^2$) Allele: Implications for Molecular Screening of Individuals for Genetic Susceptibility to Lead Poisoning. *American Journal of Human Genetics* 49, 757-763.

Whiting, M. and Granick, S. (1976). Delta-Aminolevulinic acid Synthase from chick embryo liver mitochondria. I. Purification and some properties. *Journal of Biological Chemistry* **251**, 1347.

Wibberley, D. G., Khera, A. K., Edwards, J. H., and Rushton, D. I. (1977). Lead levels in human placentae from normal and malformed births. *J. Med. Genet.* **14**, 339-345.

Wieder, R., and Wetmur, J. G. (1981). One hundred fold acceleration of DNA renaturation rates in solution. *Biopolymers* **20**, 1537-1547.

Winneke, G., Brockhaus, A., Ewers, U., Kramer, U., and Neuf, M. (1990). Results from the European multicenter study on lead neurotoxicity in children: Implications for risk assessment. *Neurotoxicol. Teratol.* **12**, 553-559.

Winneke, G., Kramer, U., Brockhaus, A., Ewers, U., Kujanek, G., Lechner, H., and Janke, W. (1983). Neuropsychological studies in children with elevated tooth-lead concentrations. II. Extended study. *Int. Arch. Occup. Environ. Health* **51**, 231-252.

Wittmers, L. E., Jr., Wallgren, J., Alich, A., Aufderheide, A. C., and Rapp, G., Jr. (1988). Lead in Bone. IV. Distribution of lead in the human skeleton. *Arch. Environ. Health.* **43**, 381-391.

Wu, W. H., Shemin, D., Richards, K. E., and Williams, R. C. (1974). The quaternary structure of δ-aminolevulinic acid dehydratase from bovine liver. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 1767-1770.

Yamamoto, M., Fugita, H., Watanabe, N., Hayashi, N., and Kikuchi, G. (1986). An immunochemical study of delta-aminolevulinate synthase and delta-aminolevulinate dehydratase in liver and erythroid cells of rat. *Arch. Biochem. Biophys.* **245**, 76-83.

Yamamoto, M., Fujita, H., Watanabe, N., Hayashi, N., and Kikuchi, G. (1987). An immunochemical study of delta-aminolevulinate synthase and delta-aminolevulinate dehydratase in rat liver and erythroid cells. *Arch. Biochem. Biophys.* **245**, 76.

Yamamoto, M., Kure, S., Engel, J. D., and Hiraga, K. (1988). Structure, turnover, and heme-mediated suppression of the level of mRNA encoding rat liver delta-aminolevulinate synthase. *J. Biol. Chem.* **263**, 15973-15979.

Yamamoto, M., Yew, N. S., Federspiel, M., Dodgson, J. B., Hayashi, N., and Engel, J. D. (1985). Isolation of recombinant cDNAs encoding chicken erythroid δ-aminolevulinate synthase. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 3702-3706.

Yamauchi, K., Hayashi, N. and Kikuchi, G. (1980). Translocation of delta-aminolevulinate synthase from the cytosol to the mitochondria and its regulation by hemin in the rat liver. *Journal of Biological Chemistry* **255**, 1746.

Zheng, W., Perry, D. F., Nelson, D. L., and Aposhian, H. V. (1991). Choroid plexus protects cerebrospinal fluid against toxic metals. *FASEB J.* **5**, 2188-2193.

Ziemsen, B., Angerer, J., Lehnert, G., Benkmann, H.-G., and Goedde, H. W. (1986). Polymorphism of delta-aminolevulinic acid dehydratase in lead-exposed workers. *Int. Arch. Occup. Environ. Health* **58**, 245-247.