EXPRESSION AND CHARACTERIZATION OF TWO RECOMBINANT MAMMALIAN METALLOPROTEINS: BOVINE MICROSOMAL CYTOCHROME b_5 AND HUMAN SERUM TRANSFERRIN (N LOBE)

By

Walter David Funk

B.Sc. (Hon) 1984
University of British Columbia

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

Ph.D.

in

THE FACULTY OF MEDICINE

DEPARTMENT OF BIOCHEMISTRY

We accept this thesis as conforming to the required standard

The University of British Columbia

November, 1990

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Biochemistry

The University of British Columbia
2075 Westbrook Place
Vancouver, Canada
V6T 1W5

August 1, 1990

Abstract

Two separate systems were developed for the expression of recombinant metalloproteins. A synthetic gene encoding the lipase-solubilized form of bovine liver microsomal cytochrome b_5 was designed and assembled for expression in *E. coli*. Analysis of the initial recombinant cytochrome revealed differences in several physical characteristics of the molecule compared to the authentic bovine liver species, including a reduction potential that was lower by 17 mV. Further studies showed the primary sequence of the initial recombinant differed from the authentic protein in the amidation status of three residues which, when corrected yielded a recombinant protein identical in behaviour to the authentic protein.

The participation of Ser64 in the stabilization of the oxidized form of cytochrome b_5 was investigated using site-directed mutagenesis to alter this residue to Ala, which was predicted to ablate a hydrogen bond formed between the protein and heme-propionate 7. Spectroelectrochemical analysis of this variant showed that the reduction potential had been shifted downwards by 7 mV, in contrast to predictions from a structural model describing the red/ox behaviour of cytochrome b_5 (Argos and Mathews,1975). The role of heme carboxylates in determining the reduction potential was confirmed for both the wild-type and Ala64 variants by heme replacement studies using the esterified derivative of protoporphyrin IX, suggesting that the presence of free carboxylates contributes to the stabilization of the oxidized species. In addition, constructions for the expression of the trypsin-solubilized form of bovine liver microsomal cytochrome b_5 and the erythrocytic form of human cytochrome b_5 are described.

A tissue culture cell system was developed for the expression of the N-terminal half molecule of human serum transferrin. The recombinant molecule (hTF/2N) was secreted at high levels from selected eukaryotic cells, and displayed high identity with

the proteolytically-derived molecule from authentic human serum transferrin as judged by sequence analysis, electrophoretic mobility and iron binding capacity. A construction for the expression of the C-terminal half molecule was assembled but failed to express recombinant protein when introduced into tissue culture cells.

The production of these two heterologous expression systems allows for high-level recovery of recombinant protein and provides a convenient approach to structure-function studies employing site-directed mutagenesis techniques.

TABLE OF CONTENTS

TABLES

FIGURES

Abbreviations Used

A	Absorbance
Å	Ångstrom
ALA	Aminolevulenic acid
ATP	Adenosine triphosphate
BHK	Baby Hamster Kidney
b.p.	before present
bp	base pair
C.D.	Circular dichroism
cDNA	complementary deoxyribonucleic acid
CoA	Coenzyme A
Da	Dalton
DHFR	Dihydrofolate reductase
DME	Protoporphyrin IX dimethylester
DMEM	Dulbecco's Modified Essential Medium
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylenediamine tetraacetic acid
EPR	Electron paramagnetic resonance
Erb_5	Erythrocytic form of cytochrome b_5
ExoIII	Exonuclease III
FPLC	Fast Protein Liquid Chromatography
g	Universal gravitational constant
hGH	Human growth hormone
HPLC	High Pressure Liquid Chromatography
IPTG	isopropylthiogalactoside
IRE	Iron responsive element

K_a	Association constant
kbp	kilobase pair
kDa	kiloDalton
Klenow	large fragment of *E. coli* DNA Polymerase I
LF	Lactoferrin
Lpb_5	Lipase-solubilized form of hepatic microsomal cytochrome b_5
mRNA	messenger ribonucleic acid
mV	millivolt
melanoTF	melanotransferrin (p97)
MT	metallothionein
MTX	methotrexate
NadChol	sodium deoxycholate
NaDodSO4	sodium dodecylsulphate
NADH	nicotinamide adenine dinucleotide (reduced)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced)
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
O.D.	Optical density
OMb_5	Mitochondrial Outer Membrane Cytochrome b_5
OTF	Ovotransferrin
OTTLE	Optically Transparent Thin Layer Electrode
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
RNA	Ribonucleic acid
rRNA	ribosomal Ribonucleic acid

SCE	Standard Calomel Electrode
SHE	Standard Hydrogen Electrode
SSC	Standard Saline Citrate
SV40	Simian Virus 40
TF	Transferrin
TFA	Trifluoroacetic acid
TF/2C	Carboxy-terminal half molecule of transferrin
TF/2N	Amino-terminal half molecule of transferrin
Tp	Trypsin
TPCK	L-1-tosylamide-2-phenyl-ethylchloromethyl ketone
Tris	Tris(hydroxymethyl)aminomethane

ACKNOWLEDGEMENTS

I'd like to thank all the little people.....I mean the really large and important people with whom I have had the pleasure of working with over the past five years. This project involved literally dozens of people and I deeply appreciate the attention and patience awarded to me by co-workers. In particular, huge thanks to Bob Woodworth and Anne Mason for their achievements on the transferrin project and Marcia Mauk and Paul Barker for their guidance through the realms of protein chemistry. The spectacular graphics in this dissertation were produced by the lightning keystrokes of the Brayer lab members; thanks very much to Terrance Lo, Michael Murphy and Albert Berghius for great screenplay.

As a longstanding member of the MacGillivray lab, I've seen many come and go (and some come back again). Very special thanks to Marlys Koschinsky for sharing the excitement and comradery of science and to Bruce Ritchie and Dave Banfield for much needed laughs. I'd like to wish each and every lab member (past, present and future) a cheery sink full.

Having two research supervisors has been the best of two worlds. Thank you Grant for sharing with me the beauty of precision. The lesson may not yet be completed, but will never be forgotten. Ross; you have given me opportunities when I really needed them and have shown me not only great science, but even greater friendship. Thanks forever.

Pamela; your patience, encouragement and understanding mean everything. Thank you for letting me dream.

I. INTRODUCTION

The past decade has seen the development of powerful methods for the expression of recombinant proteins in heterologous systems. Typically, these systems rely on the introduction and overexpression of cloned DNA molecules in either bacterial or tissue culture hosts and often allow the recovery of large amounts of protein that may normally occur with extremely low natural abundance. Combining these techniques with the ability to manipulate DNA molecules precisely through the process of *in vitro* mutagenesis (Zoller and Smith,1983) has revolutionized the manner in which the structural and functional properties of proteins are studied. Non-genetic approaches have been limited not only by the natural abundance and availability of a given protein but also by the accuracy and specificity of chemical modification techniques that have been commonly used to probe protein structures. Genetic manipulations in concert with the capacity to synthesize extended-length oligonucleotides chemically now allows for the assembly of virtually any imaginable coding sequence, such that the term "synthetic biology" has now been coined to describe these new approaches.

This thesis reports the development and initial characterization of systems for the expression of recombinant forms of two mammalian iron-containing proteins. Cytochrome b_5 is a small (16 kDa) heme protein found in the microsomal fraction and on the inner mitochondrial membrane of liver, where it acts as an essential participant in a wide variety of electron transfer reactions. In both hepatic forms, this cytochrome exists as a soluble heme-binding core attached by a short linker sequence to a membrane-anchoring domain. Serum transferrin (78 kDa) is the primary carrier of iron in plasma and as such is central to maintaining metal homeostasis. This protein is

composed of two highly similar lobes, each containing a single iron-binding center. Our initial attempts have been directed toward expression of these two proteins in forms that are amenable to purification and characterization. To this end, a bacterial expression system for the production of the soluble domain of bovine cytochrome b_5 and a tissue culture expression system for production of the N-terminal lobe of human serum transferrin have been developed.

A review of previous literature concerning both of these proteins is presented below to illustrate our interest in the investigation of both proteins through molecular genetic techniques.

A. Cytochrome b_5

Cytochromes b_5 are acidic cytochromes that possess a non-covalently-bound protoheme IX prosthetic group with histidyl residues providing the proximal and distal ligands to the heme iron. Cytochromes b_5 are found in most mammalian tissues in the microsomal and mitochondrial fractions. The hepatic form was discovered first and has been studied most thoroughly. This protein plays a central role in reduction/oxidation (red/ox) processes at the endoplasmic reticulum (Oshino,1982) and can be described in terms of three structural domains. The water soluble heme-binding domain consists of approximately 88 amino acid residues and contains an abundance of acidic residues that are concentrated near the edge of the heme-binding cleft. A short linker peptide of approximately 10 residues joins the soluble domain to a membrane-anchoring C-terminal domain of approximately 40 predominantly hydrophobic residues.

The existence of other forms of cytochrome b_5 in mammalian cells has been conclusively demonstrated only in the past decade. Erythrocytic cytochrome b_5 was described as a soluble form of the hepatic molecule which was originally thought to be derived by proteolytic cleavage of the linker peptide during maturation of the red blood cell (Passon *et al.*,1972). Recent analysis of the amino acid (Kimura,1984) and mRNA sequence (Steggles *et al.*,1989) of this form shows it to differ by one amino acid residue from the soluble domain of its hepatic counterpart leading to speculation of the existence of at least separate exons, if not separate genes, for the two forms. Nisimoto (1977) and Ito (1980) purified a third member of the cytochrome b_5 family from the outer membrane of mitochondria, OMb_5. Amino acid sequence analysis of OMb_5 (Lederer *et al.*,1983) has shown it to share approximately 58% sequence identity with microsomal

cytochrome b_5. No reports concerning the molecular genetics of this protein have appeared.

In the following review, the term "cytochrome b_5", or its abbreviation "b_5", will apply to the microsomal cytochrome unless otherwise indicated.

1. Historical

The initial observation of non-hemoglobin pigmented substances in tissues was made by MacMunn in the late 19th century. With little support from the scientific community of the time, it remained for Keilin in 1925 to "re-discover" these coloured proteins, which he subsequently termed cytochromes, signifying coloured pigments (Keilin,1925). Keilin described four primary absorption bands for this pigment, the A band (615nm-593nm), B band (567.5nm-561nm), C band (554.4nm-546 nm) and D band (532nm-511nm) and went on to demonstrate that the cytochrome absorption pattern was due to the presence of three distinct components, which he termed a',b' and c'. Considerable variability was observed in the B band of cytochromes from different organisms and tissues, leading to the identification and definition of cytochrome *b* (Keilin,1925), cytochrome b_1 (Keilin,1925), cytochrome b_2 (Bach *et al.*,1946), cytochrome b_3 (Martin and Morton, 1955), and cytochrome b_4 (Egami *et al.*,1953). Each member of the *b*-type family was characterized by activity towards red-ox agents and spectroscopic behavior, with each member containing a protoheme or related non-formyl heme non-covalently bound by the protein (Wainio,1982).

The first description of what has since been commonly termed cytochrome b_5 was made by Keilin and Hartree (1940) who described a *b*-type cytochrome in pig kidney, which they called b_1. Later, Yoshikawa (1951) reported a similar protein in

liver, calling it cytochrome *b'*, while Strittmatter and Ball (1954) referred to a similar enzyme from rat liver microsomes as cytochrome *m*. Cytochrome b_5 was initially used as a descriptor for a protein from *Ceropia* silkworm larvae (Pappenheimer and Williams,1954), though it now appears likely that the silkworm cytochrome was actually cytochrome c_1 of the respiratory chain (Wainio,1982). Chance and Williams (1954) recognized the spectroscopic similarity between cytochrome *m* and the silkworm protein and renamed the microsomal protein cytochrome b_5.

Microsomal b_5 was subsequently shown to be present in the liver of guinea pig, rabbit and calf and could be released from microsomal preparations with a pancreatic lipase preparation (Strittmatter and Velick,1956). The same study showed b_5 to be distinct from mitochondrial b_1 and identified the prosthetic group subsequently as iron protoporphyrin IX. Detergent solubilization of microsomal membranes allowed identification of the hydrophobic C-terminal membrane-binding domain (Ito and Sato,1968).

Passon and co-workers (1972) identified a soluble form of cytochrome b_5 in mammalian erythrocytes. The purified cytochrome was shown to participate in the methemoglobin reductase system and was proposed to arise from hydrolytic cleavage of a membrane-bound precursor during the maturation of the erythrocyte. However, recent evidence from nucleotide sequence analysis of erythrocytic b_5 mRNA has shown conclusively that this protein is distinct from the hepatic species and is translated from a different mRNA (see IIB4, Nucleotide Sequence Analysis).

The existence of a mitochondrial species of b_5 was first proposed by Raw and Mahler (1959), who identified an NADH-cytochrome *c* reductase activity in the outer membrane. However several years passed before this outer membrane protein, OMb_5,

was purified and shown to be unequivocally distinct from the microsomal species (Nisimoto *et al.*,1977; Ito,1980).

2. **Biosynthesis and Physiological Role**

Descriptions of *b*-type cytochrome in the microsomal fraction of hepatocytes were reported as early as 1954 (Chance and Williams). Cytochrome b_5 has been detected immunologically on both the rough and smooth endoplasmic reticulum (Remacle *et al.*,1976), and a NADH-reduceable cytochrome was also reported in nuclear and Golgi membrane fractions (Ichikawa and Yamano,1970; Fleischer *et al.*,1971). These findings suggested that cytochrome b_5 may be distributed on most cellular membranes, with the notable exception of the plasma membrane (Fleischer *et al.*,1971). With the exception of heart and skeletal muscle, almost all tissues yield detectable amounts of cytochrome b_5 in their microsomal fractions, though relative levels of the protein vary from 0.03 nmol/mg protein in the rat testis to 1.0 nmol/mg protein in the liver (see Oshino,1982). Mitochondrial b_5 has been localized to the outer face of the outer membrane (Fukushima and Sato,1973) while erythrocytic b_5 is strictly a cytosolic protein (Passon *et al.*,1972).

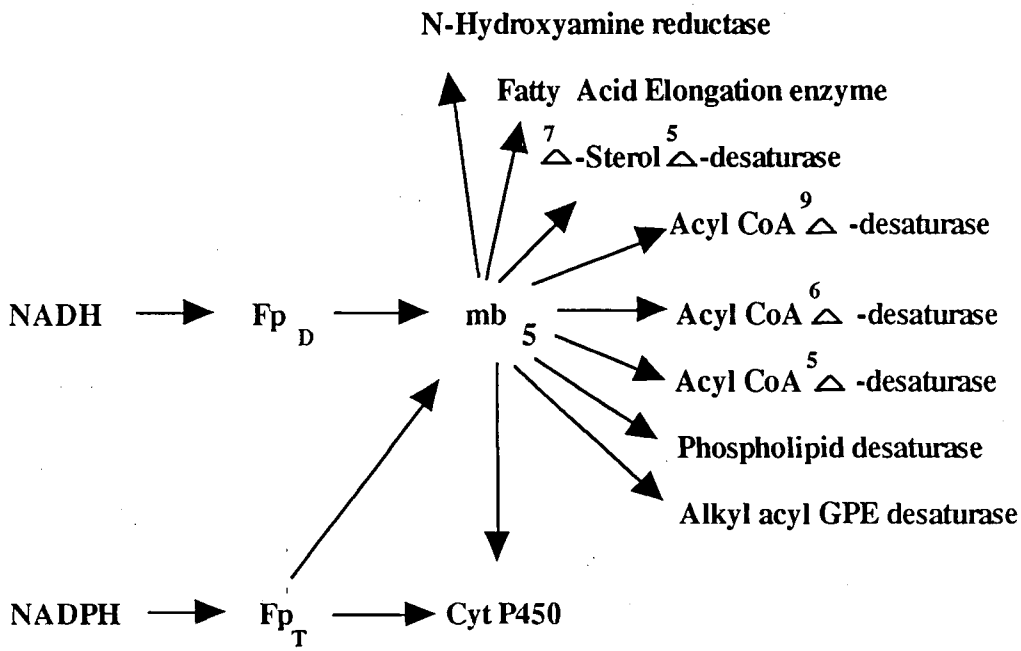
Cytochrome b_5 is synthesized on free ribosomes (Rachubinski *et al.*,1980; Okada *et al.*,1982) after which the protein relocates to membrane surfaces where binding occurs. The ability of cytochrome b_5 to insert spontaneously into many artificial and natural membrane systems (Rogers and Strittmatter,1975) suggests that the observed concentration of b_5 on microsomal membranes *in vivo* may be dependent on transport or exchange processes as not yet described. Little is known concerning the mechanism of heme insertion, though the ability of apo-b_5 to assimilate heme spontaneously suggests that a specific heme "insertase" is not required. Apo-b_5 is easily hydrolyzed by

proteases, and the protein undergoes an increase in structural order upon binding heme (Huntley and Strittmatter, 1972). The half-life of hepatic b_5 *in vivo* has been estimated to be approximately 100 hr. (Omura *et al.*,1967) while the heme moiety has a half life of just 55 hr. (Druyan *et al.*,1968), indicating that the heme group may undergo exchange reactions with other proteins.

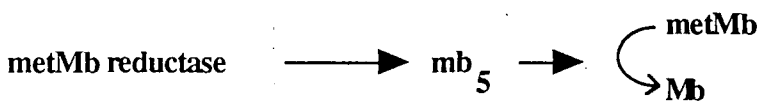
Hepatic cytochrome b_5 plays a central role in the electron transfer systems of the microsomal membrane. In general, cytochrome b_5 mediates one electron transfer reactions between two flavoproteins, NADH-cytochrome b_5 reductase and NADPH-cytochrome *c* reductase that participate in a variety of metabolic reactions as shown in Figure 1. Both reductase molecules share a similar organization with b_5 in that they consist of a soluble catalytic domain which can be released from membranes by proteolysis (Spatz and Strittmatter,1973).

Both the hydrolytic and native forms of cytochrome b_5 reductase are capable of reducing the hydrolytic and native forms of b_5, and reactions between the native forms of the two proteins are accelerated under conditions which favour protein-protein interaction (Rogers and Strittmatter, 1973). The catalytic cycle of this enzyme involves the transfer of two electrons from the reductase to two molecules of b_5 in sequential reaction steps (see Mathews and Czerwinski, 1976 for review). Reduction of b_5 by NADPH-cytochrome b_5 reductase requires the reductase to be in its native membrane-bound state (Bilimoria and Kamin,1973). However, the cytochrome P450 family appears to be the principal electron acceptor of cytochrome *c* reductase (Jansson and Schenkman,1977), thus the physiological significance of cytochrome *c* reductase in b_5-dependent reactions remains unclear.

A)


N-Hydroxyamine reductase
Fatty Acid Elongation enzyme
7 △-Sterol 5 △-desaturase
Acyl CoA 9 △ -desaturase
NADH
Fp D
mb 5
Acyl CoA 6 △ -desaturase
Acyl CoA 5 △ -desaturase
Phospholipid desaturase
Alkyl acyl GPE desaturase
NADPH
Fp T
Cyt P450

B)


metMb reductase
mb 5
metMb
Mb

C)

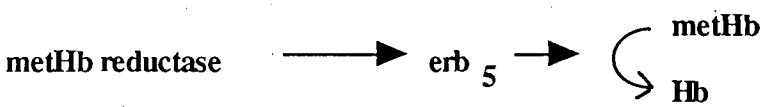

metHb reductase
erb 5
metHb
Hb

Figure 1: Metabolic Pathways That Involve Cytochrome b_5 as a Mediator of Electron Transfer (modified from Oshino, 1982).

Panel A: Liver microsomal cytochrome b_5 (mb_5) mediates single electron transfers between membrane-bound NADH-cytochrome b_5 reductase (Fp_D) and several fatty acid desaturase systems, fatty acid elongation enzyme N-hydroxyamine reductase and the Cytochrome P450 family of enzymes (see Oshino,1982). *In vitro*, cytochrome b_5 may also shuttle electrons from NADPH-cytochrome *c* reductase (Fp_T) to the same cytochrome P450 enzymes though the physiological significance of this observation is weak due to the high relative activity of Fp_D and the abundance of NADH in intact liver.

Panel B: An NADH-dependent reductase activity has been identified in the muscle tissues of several organisms (see Livingston *et al.*,1985 for example). This reductase activity is required to maintain myoglobin (Mb) in a reduced (ferrous) state in order to bind oxygen reversibly, and *in vitro*, this reaction is increased 100-fold by the presence of cytochrome b_5.

Panel C: Methemoglobin (MetHb) reductase is the term commonly used to describe the soluble form of NADH-cytochrome b_5 reductase found in the erythrocyte cytoplasm. This reductase system maintains hemoglobin in the ferrous state to ensure cooperative oxygen binding and utilizes the soluble erythrocytic form of cytochrome b_5 (erb_5).

The principal cytochrome b_5 oxidase activities can be classified either as desaturases or as members of the cytochrome P450 family. Acyl CoA desaturase activities catalyze the Δ9-, Δ6- and Δ5- desaturation of long-chain fatty acids at the expense of molecular oxygen and NAD(P)H or ascorbate (Stoffel,1961; Marsh and James,1962). The involvement of cytochrome b_5 in these reactions has been inferred because b_5 is the only microsomal protein that can be reduced by all three reductants (Holloway and Katz,1972) and, desaturation is specifically inhibited by the presence of anti-b_5 antibodies (Jansson and Schenkman,1977). Other desaturase processes that probably require the participation of cytochrome b_5 include alkylacylglycerophosphoryl-ethanolamine desaturase (Paltauf and Holasek,1973), phospholipid desaturase (Pugh and Kutes,1977) and Δ7-sterol Δ-5 desaturase (Reddy *et al.*,1977).

The cytochrome P450 family catalyzes the hydroxylation of a wide spectrum of drugs and xenobiotics and as such plays a central role in detoxification. Cytochrome b_5 is required for the maximal activity of many P450 reactions, including the oxidation of chlorobenzene (Lu *et al.*,1974) and 3,4-benzpyrene (West et al.,1974) and the demethylation of benzphetamine (Imai and Sato,1977) and *p*-nitroanisole (Sugiyama *et al.*,1979). In addition, hepatic cytochrome b_5 has been implied as an intermediary in many other microsomal reactions that use NAD(P)H as a source of reducing equivalents. These processes include fatty acid elongation (Keyes *et al.*,1979) and N-hydroxyamine reduction (Kadlubar and Ziegler,1974).

The soluble erythrocytic form of cytochrome b_5 functions as a shuttle in the methemoglobin reductase system (Passon *et al.*,1972), transferring electrons from methemoglobin reductase to methemoglobin (see Figure 1). Hemoglobin must be maintained in the reduced state to bind oxygen and thus erythrocytic b_5 is essential for normal respiration. Both the erythrocytic b_5 reductase and erythrocytic b_5 are soluble molecules, and initial studies suggested these molecules are produced from the hydrolysis of membrane-bound precursor forms during the maturation of erythrocytes (Slaughter and Hultquist,1979). Analysis of cDNA sequences for the erythrocytic forms of cytochrome b_5 show that this molecule is translated from a distinct erythrocytic mRNA, with a predicted sequence slightly different from that of hepatic b_5 (A. Steggles, personal communication). The existence of separate genes for the hepatic and erythrocytic forms of b_5 may be pertinent to the condition of methemoglobinemia (Hegesh *et al.*,1986), in which low levels of erythrocytic b_5 activity accompany high levels of methemoglobin. These patients may maintain normal hepatic b_5 expression and function, while a genetic lesion specifically affects the expression of erythrocytic b_5.

The existence of a cytochrome b_5-like molecule in the outer membrane of the mitochondrion was first described by Raw and Mahler (1959) as a component of a rotenone-insensitive cytochrome *c* reductase system. The term outer membrane, or OMb_5 was used to describe this hemoprotein (Fukushima and Sato,1972) and sequence analysis confirmed that OMb_5 is closely related to the hepatic and erythrocytic b_5 species (Lederer *et al.*,1983). The rat microsomal and mitochondrial proteins exhibit a 58% sequence identity. The basis for the localization of OMb_5 in the mitochondrial membrane remains enigmatic because small differences in the sequences of membrane-binding domains are not known to provide membrane binding specificity. The presence of membrane-specific protein-protein interactions may be a factor in compartmentalizing membrane proteins (Okada *et al*,1982).

3. Structural Considerations

i) Isolation of Soluble and Membrane-Binding Forms

The release of a soluble heme-binding domain of microsomal cytochrome b_5 can be achieved with a variety of proteolytic activities. Pioneering work by Strittmatter and Velick (1956) used a pancreatic lipase preparation for this purpose. The active proteolytic agent in this preparation may be cathepsin-type proteases released during the isolation of microsomes. Kajihara and Hagihara (1968) used trypsin and subtilisin in their preparations. These proteins produced b_5 derivatives with different N- and C-termini as a result of differences in protease substrate specificity (Figure 2). By comparison, erythrocytic forms of cytochrome b_5 maintain their N-terminal residues and end at position 97.

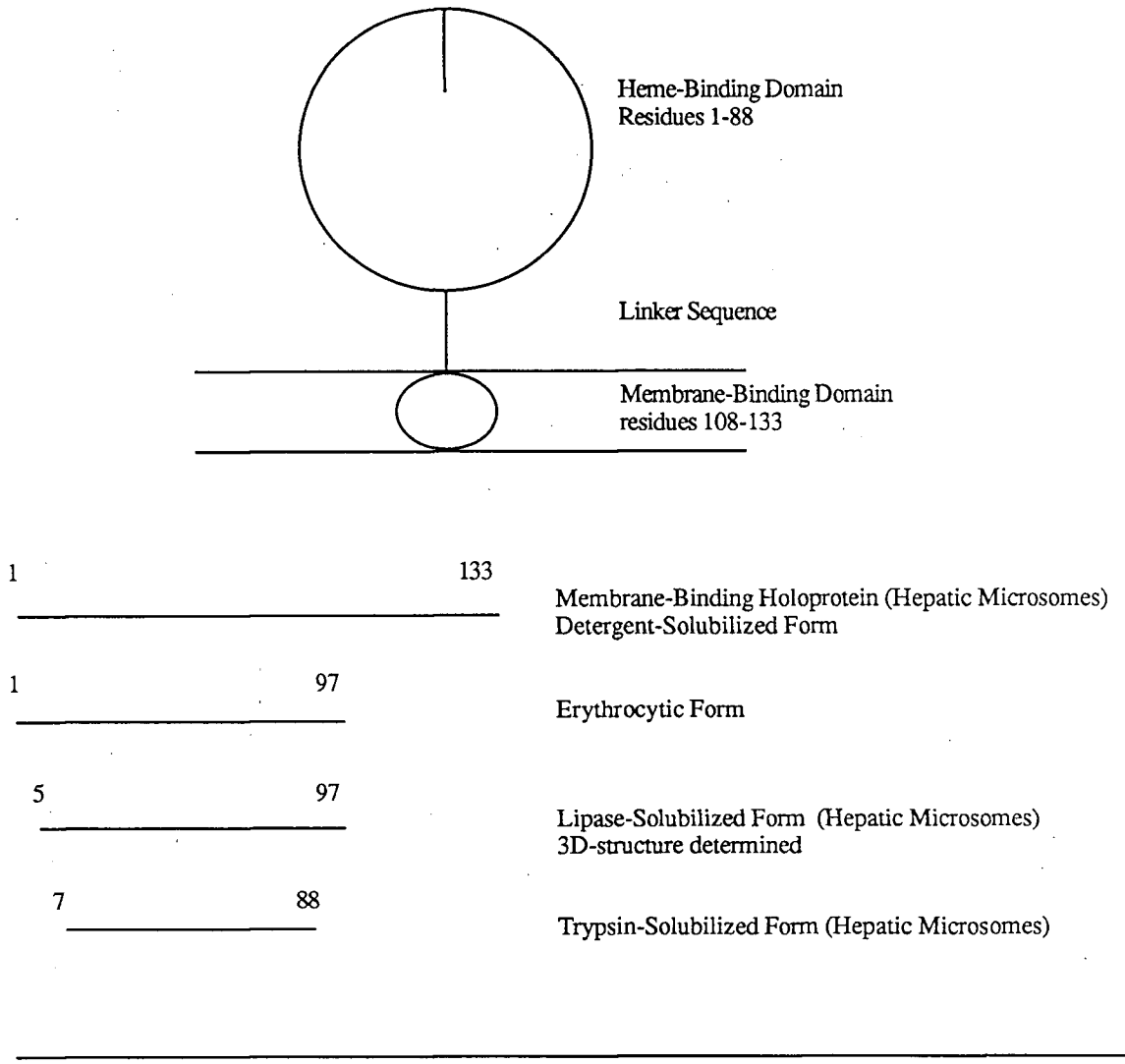


Figure 2: Isolated Forms of Bovine Cytochrome b_5 (modified from Abe et al.,1985)

The hepatic form of cytochrome b_5 is a microsomal protein consisting of a cytoplasmic heme-binding domain, joined to a membrane anchoring domain by a short linker sequence. The erythrocytic form maintains an intact heme-binding domain along with a portion of the linker peptide. Isolation of the soluble domain by using pancreatic lipase preparations results in the removal of four residues from the N-terminus while trypsin solubilization hydrolyzes the heme-binding domain at both the N- and C-termini.

Detergent solubilization of microsomal preparations with Triton X-100 releases the intact cytochrome (Spatz and Strittmatter,1971). The functional properties of the catalytic domain are not affected by the presence of the non-polar domain, though aggregation of the intact protein into octamers occurs in the absence of detergent (Oshino,1978; Waino,1978).

The insertion of b_5 into lipid bilayers has been shown to occur as a "loose" complex in which the protein can readily exchange between lipid vesicles. Membrane bound b_5 is susceptible to C-terminal cleavage by carboxypeptidase Y (Enoch and Strittmatter, 1979; Leto *et al.*,1980; Takagaki *et al.*,1983). Reconstitution into liposomes in the presence of detergent produces a complex in which the C-terminus is protected from carboxypeptidase Y activity, and transfer between vesicles is not observed. This observation suggests formation of a "tight" complex in which the C-terminus is now found in the lumen of the vesicle under these conditions. Studies on the topology of b_5 in physiological membranes indicates that the C-terminal hexapeptide can be released from microsomal preparations by the action of trypsin (Ozols,1989), in agreement with membrane-binding models in which the hydrophobic domain loops into the bilayer and re-emerges on the same cytosolic face. Recent studies in which the hydrophobic domain sequence of cytochrome b_5 was fused to the C-terminus of β-galactosidase and subsequently expressed in *E. coli* have demonstrated that this domain directs attached proteins to membrane surfaces (George *et al.*,1989), and therefore, probably maintains a conformation which is independent of the soluble domain of cytochrome b_5.

```
                    1                   10                  20                  30
                    *                   *                   *                   *
Human(l)[a]         A E Q S D E A V K Y Y T L E E I Q K H N H S K S T W L I L H
Human(e)[b]         A E Q S D K A V K Y Y T L E E I Q K H N H S K S T W L I L H
Bovine(l)[c]        A E E S S K A V K Y Y T L E E I Q K H N N S K S T W L I L H
Bovine(e)[d]        A E E S S K A V K Y Y T L E E I Q K H N N S K S T W L I L H
Rat(l)[e]           A E Q S D K E V K Y Y T L E E I E K H K D S D S T W V I L H
Chicken(l)[f]   G S S E A G G E A W R G R Y Y R L E E V Q K H N N D Q S T W I I V H

                              40                  50                  60
                              *     #             *                   *
Human(l)            H K V Y D L T K F L E E H P G G E E V L R E Q A G G D A T E N F E D
Human(e)            H K V Y D L T K F L E E H P G G E E V L R E Q A G G D A T E N F E D
Bovine(l)           Y K V Y D L T K F L E E H P G G E E V L R E Q A G G D A T E N F E D
Bovine(e)           Y K V Y D L T K F L E E H P G G E E V L R E Q A G G D A T E N F E D
Rat(l)              H K V Y D L T K F L E E H P G G E E V L R E Q A G G D A T E N F E D
Chicken(l)          H R I Y D I T K F L D E H P G G E E V L R E Q A G G D A T E N F E D

                         70                  80                  90
                    #    *                   *                   *
Human(l)            V G H S T D A R E M S K T F I I G E L H P D D R P K L N K P P E T L
Human(e)            V G H S T D A R E M S K T F I I G E L H P D D T P K L N K P P E P
Bovine(l)           V G H S T D A R E L S K T F I I G E L H P D D R S K I T K P S E S I
Bovine(e)           V G H S T D A R E L S K T F I I G E L H P D D R S K I T K P S E S
Rat(l)              V G H S T E A R E L S K T Y I I G E L H P D D R S K I S K P S E T L
Chicken(l)          V G H S T D A R A L S E R F I I G E L H P D D R P K L Q K P A E T L

                    100                 110                 120                 130
                     *                   *                   *                   *
Human(l)            I T T I D S S S S W W T N W V I P A I S A V A V A L M Y R L Y M A E D
Bovine(l)           I T T I D S N P S W W T N W L I P A I S A L F V A L I Y H L Y T S E N
Rat(l)              I T T V E S N S S W W T N W V I P A I S A L V V A L M Y R L Y M A E X
Chicken(l)          I T T V Q S N S S S W S N W V I P A I A A I I V A L M Y R S Y M S E
```

[a]predicted from cDNA sequence; Yoo and Steggles,1988.
[b]protein sequence; Abe et al.,1985.
[c]predicted from cDNA sequence; Cristiano and Steggles,1989.
[d]protein sequence; Douglas and Hultquist,1978.
[e]protein sequence; Ozols and Heinemann,1982.
[f]predicted from cDNA sequence; Zhang and Somerville,1988.
#axial histidyl residues

Figure 3: Primary Sequence Comparison of Selected Hepatic and Erythrocytic Cytochrome b_5 Species.

Published primary sequences were compared and residues at variance with the majority of mammalian residues are underlined.

---

ii) Amino Acid Sequence Analysis

The amino acid sequences of both the hepatic and erythrocytic forms of cytochrome b_5 have been determined for several species, using both chemical and cDNA sequencing techniques. The results of sequence analysis of several species presented in Figure 3 clearly show the extraordinary sequence conservation seen for this protein. Sequence identity between avian and mammalian species is approximately 75%, indicating that the protein must be under considerable evolutionary constraint. The divergence of these two lineages has been estimated to have occurred 200 million years ago (Doolittle and Feng,1987). A total of 54 positions are invariant in all known microsomal species. Significant differences are seen primarily at the extreme N- and C-termini and in the hinge region connecting the hydrophilic and hydrophobic domains of the protein. Sequence analysis has established that the N-termini of most species are acetylated. Only the porcine and chicken have free amino termini (Ozols,1989). As the first 15 residues are identical for porcine, horse and rat cytochromes, the lack of acetylation of the porcine molecule is surprising. The estimated molecular weight of intact bovine liver microsomal b_5 determined from its amino acid sequence is 15,233 daltons (including heme), while the lipase and trypsin solubilized forms of the same species have molecular weights of 11,250 (residues 5-97) and 10,082 daltons (residues 7-88) respectively.

The initial report of the sequence of the erythrocytic and hepatic forms of bovine b_5 concluded the sequence to be identical over residues 1-97 (Slaughter *et al.*,1982). The authors suggested that the erythrocytic form is generated from a full-length precursor that is identical to the hepatic form. Lysosomal proteases released during the maturation of red blood cells were postulated as being responsible for this conversion. This model has since been modified by the sequence determination of the C-termini of erythrocytic b_5 from human, pig (Kimura *et al.*,1984) and rabbit (Schafer and Hultquist,1983). Residue 97 was identified as Pro and Ser for the human and porcine erythrocytic forms respectively, while the corresponding liver forms of both species had Thr at this position. This led to speculation that the two forms were encoded by distinct transcripts, either as the result of alternate splicing of exons from a single gene, or from the presence of two separate genes. Recent evidence provided by Southern analysis of chicken genomic DNA favours the latter of these two models (see IA4, Nucleotide Sequence Analysis).

The complete amino acid sequence of the heme-binding domain of rat liver outer mitochondrial membrane cytochrome b_5 has also been determined (Lederer *et al.*,1983). This sequence differs from that of the microsomal protein at 38 of 91 residues. When the non-conservative changes are mapped to the crystal structure of the calf microsomal b_5 (Mathews *et al.*,1972), most replacements are found at the surface of the molecule.

iii) Structural Analysis

The crystal structure of the lipase-solubilized bovine liver microsomal b_5 was first solved by Mathews et al. (1971a) to a resolution of 2.8Å and later refined to 2.0Å by the same workers (1971b) (see Fig 4). The protein has an approximately cylindrical

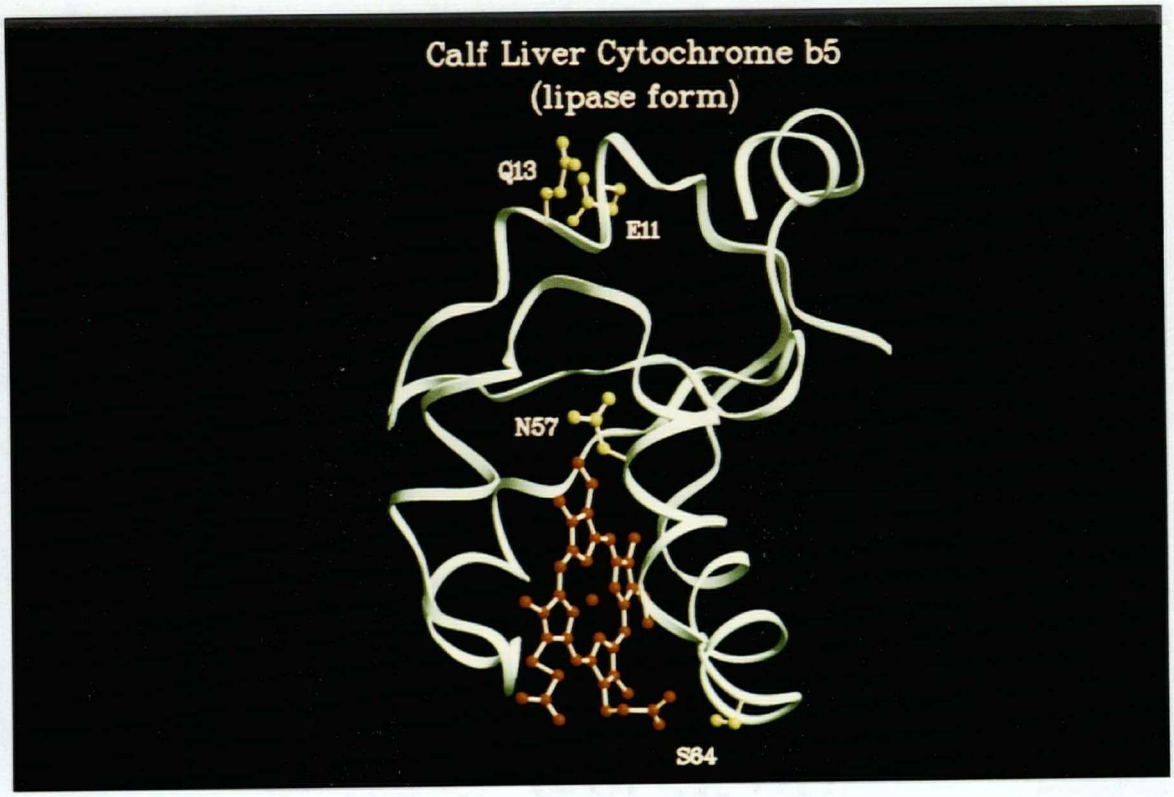


Figure 4: Ribbon-representation of the Crystal Structure of Recombinant Lipase-solubilized Cytochrome b_5.

The crystal structure of a triple mutant form of recombinant bovine cytochrome b_5 (Q11,E13,D57) is represented in ribbon fashion (Funk *et al.*,1990b). The three mutant side chains and Ser64 are shown as ball and stick projections. The soluble domain consists of two smaller subdomains; the heme-binding domain (residues 21-78) is formed by four adjacent α helices and a β-pleated sheet, while a smaller hydrophobic domain (upper in this representation) is comprised of residues 1-20 and 79-88.

---

shape, with a 35Å length and 30Å diameter. Two helical subdomains that are separated by a β-pleated sheet can be identified. The larger subdomain (residues 21-78) holds the heme moiety which sits in a distinctive crevice formed by four helices that run in anti-parallel directions. The floor of the binding pocket is formed by the β-pleated sheet

which separates the two domains. The heme group is bound non-covalently to the protein, with the heme iron coordinated to histidine residues 39 and 63. Several non-covalent interactions with 11 hydrophobic amino acid residues stabilize the binding orientation of the heme group. Compared to other heme proteins such as hemoglobin, the structural rigidity of the tightly packed heme crevice accounts for the low reactivity of b_5 towards binding of oxygen and carbon monoxide (Mathews and Czerwinski,1976).

The heme-binding domain is highly acidic with acidic groups outnumbering basic groups by 20 to 9 (Mathews,1985). The negatively charged residues are concentrated at the heme edge, where the estimated net charge is -10 (including the propionate groups). Complex formation between cytochrome b_5 and several reductases and oxidases is strongly dependent upon electrostatic interactions between the negatively charged face of b_5 and corresponding positively charged surfaces on the other component. Based on the manipulation of high-resolution crystal structures, models of the complexes of b_5 with cytochrome *c* (Salemme,1976), methemoglobin (Poulos and Mauk,1983), myoglobin (Livingston *et al.*,1985), and cytochrome $P450_{cam}$ (Stayton *et al.*,1989) have been proposed which feature charge complementarity at the protein interface between acidic residues of b_5 and lysyl residues of the binding partner. In agreement with such modelling, formation of the cytochrome *c*:cytochrome b_5 complex is dependent on pH and ionic strength (Mauk *et al.*,1982). Reduction of the negative charge on the binding face of b_5 by protein engineering (Rodgers *et al.*,1988) or neutralization of the lysyl residues of cytochrome *c* by chemical modification (Stoneheurner *et al.*,1979) greatly reduces the stability of the complex.

Although hemoglobin and cytochrome b_5 share little sequence identity, the overall structural organization of their heme binding regions are quite similar (Rossman

and Argos,1975). Both molecules bind heme in a non-covalent manner, with the propionate groups extending from the opening of the heme crevice. In particular, the four α-helices that frame the heme moiety of b_5 can be accurately superimposed onto similar helices in both the A and B chains of hemoglobin.

Other proteins that share structural homology with the heme-binding domain of b_5 include yeast flavocytochrome b_2 (Jacq and Lederer,1974) and mammalian sulfite oxidase (Cohen and Fridovich,1971). Upon proteolytic digestion, both proteins release a "core" fragment of approximately 11 kDa which binds heme non-covalently and shares similar spectroscopic properties with b_5. Guiard and Lederer (1979) have compared the amino acid sequences of the heme-binding cores of the three proteins and have defined a common "b_5 fold" domain which has been subsequently identified in a fourth protein, nitrate reductase from *Neurospora crassa* (Le and Lederer,1983). Each fold contains two domains that are separated by a β-pleated sheet. A total of 13 invariant residues have been identified, including two axial histidines to which the heme is coordinated. A plausible scheme for the evolution of these proteins has been proposed in which coding sequence for a primordial heme binding fragment has been recruited for fusion with other enzymatic coding sequences.

The heme group is almost entirely buried within cytochrome b_5 and only the propionate groups are significantly exposed to solvent (Stellwagen,1978). One propionate extends fully into the solvent while the other is only partially exposed, folding back into the protein to form hydrogen bonds with the hydroxyl group of Ser 64 and a main-chain amide (see Figure 5). Two conformations are available for the heme moiety, related by a rotation of 180° about the α-γ meso axis (Figure 6). Analysis of the crystal structure of oxidized b_5 indicated that the two heme orientations occur with a ratio of 4:1

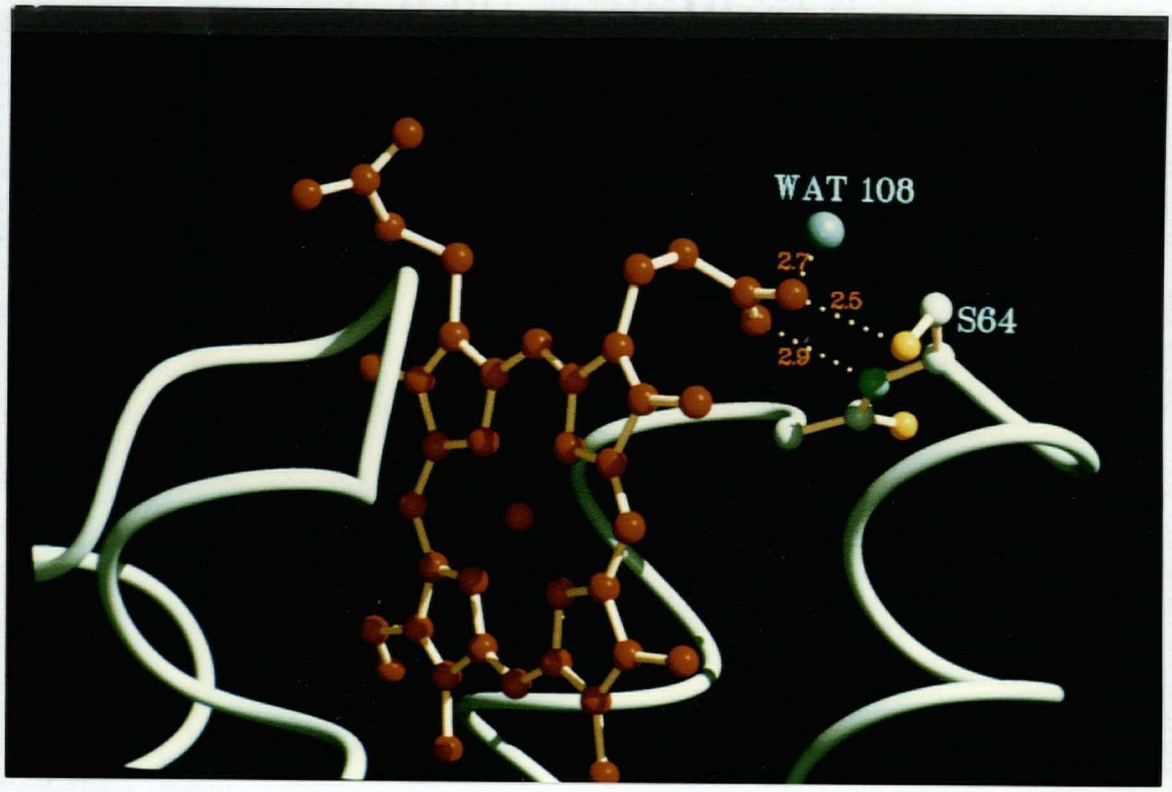


Figure 5: Ser64 Hydrogen-bonding Network

The hydrogen-bonding pattern at residue 64 of the recombinant cytochrome b_5 is virtually identical to that reported by Mathews *et al.* (1971b) for the authentic bovine liver molecule. Heme propionate 7 forms two hydrogen bonds with Ser64; one carboxylate oxygen bonds with the side chain hydroxyl while the other bonds to the main chain amide of Ser64. The bond lengths (Å) of these interactions, as well as a third hydrogen bond to a water molecule in the crystal structure are depicted by dotted lines.

(Mathews,1980). Proton NMR studies have also found a significant bias of conformation for protoporphyrin-containing b_5 which is absent in b_5 reconstituted with deuterohemin or isopemptohemin. These results suggest that the interactions between

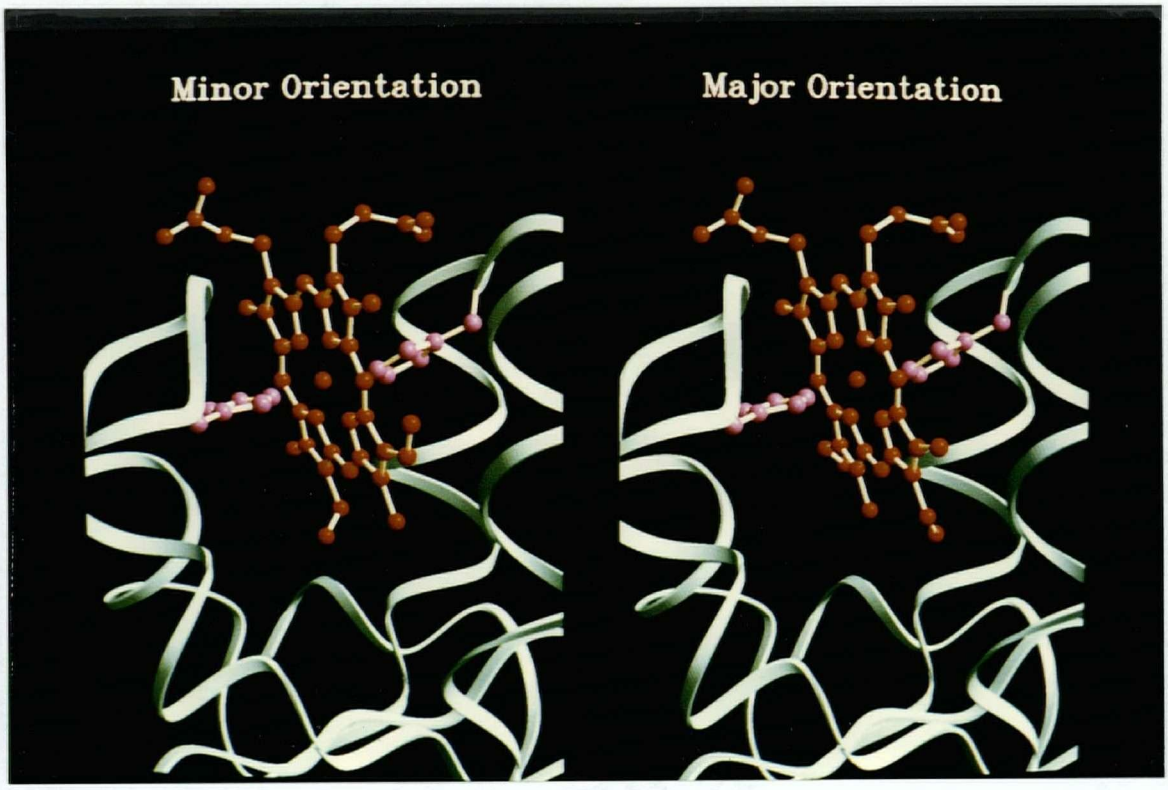


Figure 6: Heme Orientational Disorder in Cytochrome b_5

Two orientations of the heme prosthetic group are available in *b*-type cytochromes owing to the absence of covalent bonds between the heme and protein. In the recombinant cytochrome b_5 molecule, rotation about the α-γ heme axis results in a dominant orientation estimated at 60% *versus* 40% for the alternate orientation (Funk *et al.*,1990b).

---

the 4-vinyl group and the apo-protein may be responsible for the observed effect (LaMar *et al.*,1981). Studies of freshly reconstituted cytochrome b_5 by Walker *et al.* (1988) have shown that the two conformers differ in reduction potential by 27 mV (+0.8 *versus* -26.2 mV *vs.* SHE) with the interconversion occurring with a half life of 12 hr.

The same heme reorientation takes almost 100 times longer in the NADH-reduced form of the protein (LaMar *et al.*,1978), thus the physiological relevance of this finding may be minimal.

The lower third of the heme-binding fragment constitutes a separate hydrophobic core domain that has been highly conserved in all b_5 species. Although no clear functional role has been described for this domain, evidence from NMR studies of apocytochrome b_5 suggests that a stable native core centered about Trp22 exists in the absence of heme (Moore and Lecomte,1990). The hinge region and membrane anchoring domain have not been defined structurally, though Chou-Fasman analysis of the C-terminal primary sequence predicts the linker (91-97) to be α-helical (Fleming *et al.*,1978), while hydropathy plotting and proteolytic protection studies strongly suggest that residues 103-127 insert into the membrane, with residues 127-133 re-emerging into the cytosol (Ozols,1989)

X-ray diffraction analysis of the dithionite-reduced protein (ferrocytochrome b_5) showed virtually no difference in the positions of main chain and side chain atoms (Argos and Mathews,1975). The only significant change in electron density occurs near the carboxylate of the buried propionate. This density was interpreted as indicating the binding of a cation upon reduction. The authors provided a model for this process in which the formal charge of +1 for the oxidized heme is neutralized by Coulombic interaction with the buried propionate, whereas upon reduction the heme loses charge and the heme propionate is free to attract a cation from solution (see Figure 7). In concordance with this model, the reduction potential of b_5 is sensitive to ionic strength, pH and temperature (Reid *et al.*,1982), and elimination of the heme carboxylate charge

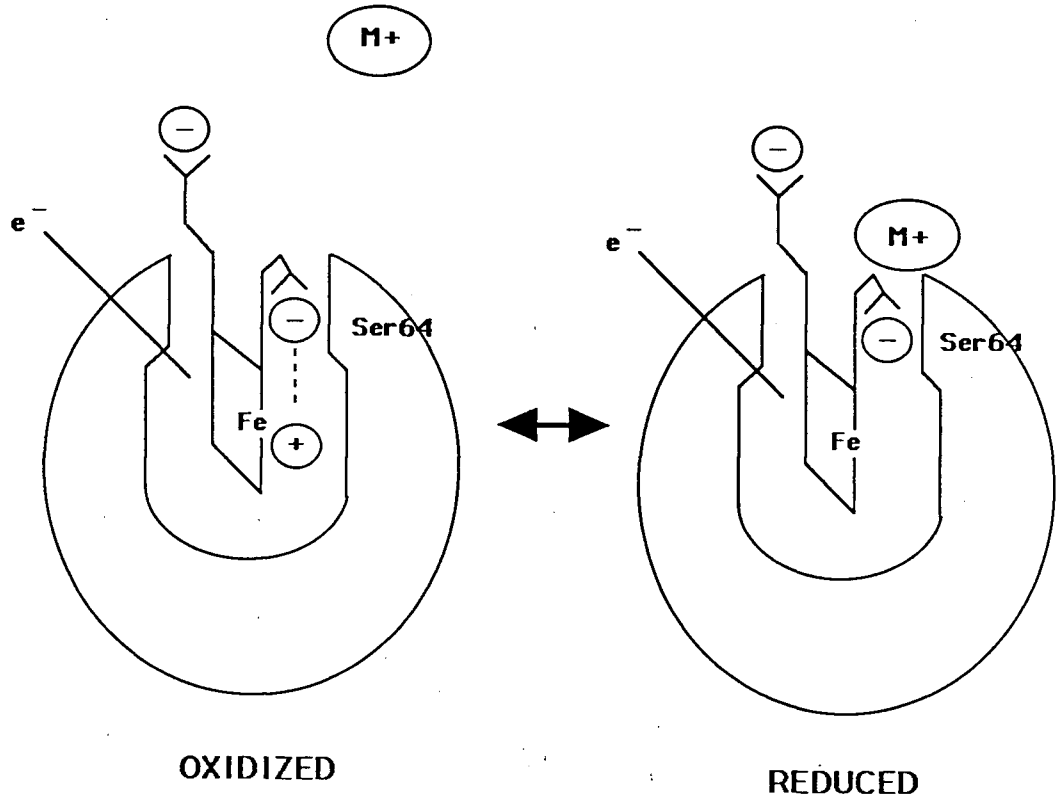


Figure 7: A Model for the Reduction/Oxidation Equilibrium of Cytochrome b_5 (from Argos and Mathews, 1975).

The reduction of cytochrome b_5 by dithionite does not result in an appreciable rearrangement of the protein structure. This model suggests that the formal positive charge of the ferric atom in the oxidized species is stabilized by the carboxylate charge of heme propionate 7, which in turn is favoured by hydrogen bonds to Ser64 (see Figure 5). Upon reduction, the carboxylate charge is released from the iron center and is free to bind a cation from solution. Thus disruption of a hydrogen bond from Ser64 would be predicted to destabilized the carboxylate charge and the oxidized protein, shifting the equilibrium towards the reduced cytochrome and raising the reduction potential.

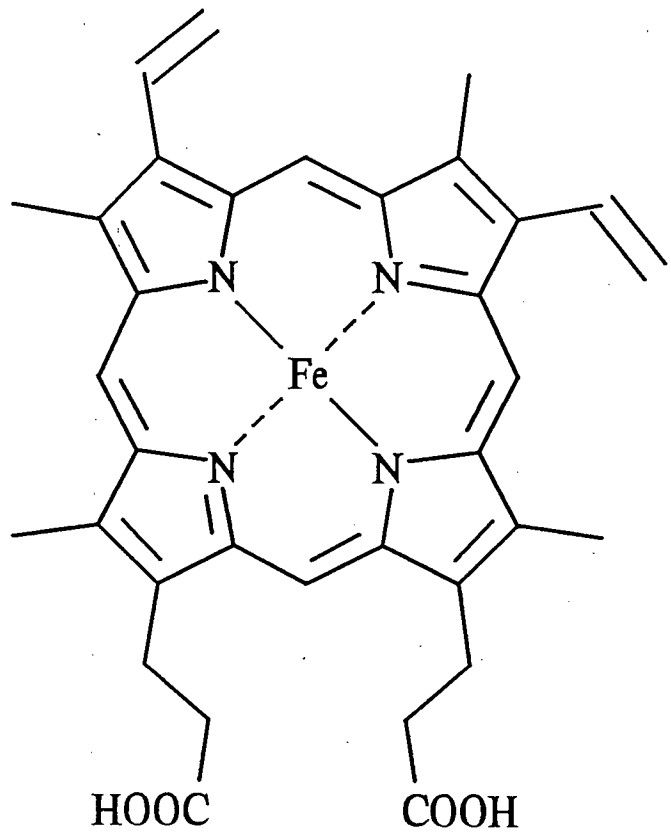


Figure 8: Structure of Ferriprotoporphyrin IX

Heme: R=H
Dimethylester (DME) heme: R=CH_3

by substitution of dimethylesterprotoporphyrin IX for protoheme IX in b_5 increases the reduction potential by 60 mV (Reid *et al.*,1984). The physical properties of soluble cytochrome b_5 (low molecular weight, high solubility and stability) make it ideal for NMR studies (Keller and Wuthrich,1980; Reid *et al.*,1987; Burch *et al.*,1990). The complete assignment of proton resonances of trypsin-solubilized hepatic cytochrome b_5 from pig and calf liver has now been reported (Veitch *et al.*,1988; Guiles *et al.*,1990), and the resulting connectivity assignments are in excellent agreement with the crystallographic model reported by Mathews (1972). The anticipated availability of a solution structure of cytochrome b_5 together with assignments for cytochrome c (Wand *et al.*,1989) should prove a powerful tool for investigating protein-protein interactions and charge transfer.

4. Nucleotide Sequence Analysis

The elucidation of the nucleotide sequences of cDNAs encoding hepatic cytochrome b_5 species has recently been achieved. Initial reports of the sequence for rabbit (Daruish *et al.*,1988) and chicken (Zhang and Somerville,1988) cDNAs demonstrated that the high degree of identity between species was, not surprisingly, conserved at the nucleotide level. The rabbit and chicken cDNAs exhibit a 76% nucleotide identity. The chicken cDNA sequence showed a suprising difference in that 5 additional amino acids were predicted at the N-terminus, while the C-terminus was one residue shorter than all mammalian b_5 species. Thus, the chicken protein is estimated to consist of 138 residues, compared to 134 residues for all mammalian species. However the N-terminal sequence identified from chemical analysis of the protein (Ozols,1989) starts at the third residue predicted from the cDNA sequence. Evidently, a

proteolytic event removes the first two residues to yield a mature protein of 136 amino acids.

During the course of studies reported in this dissertation, Steggles and co-workers used the rabbit cDNA as a cross-hybridization probe to isolate the corresponding hepatic cDNAs from bovine (Cristiano and Steggles,1989) and human sources (Yoo and Steggles,1988). These studies have correctly identified the N-terminal sequence of bovine b_5, clarified the amidation status of several residues in this protein, provided the sequence of the hydrophobic region of the human protein and corrected the assignment of several residues.

Southern analysis using the full-length chicken cDNA as a hybridization probe indicates the presence of multiple bands, while a probe spanning only the hydrophobic domain hybridizes to unique bands (Zhang and Somerville,1988). The authors suggest these results can be explained by the presence of two cytochrome b_5 genes, only one of which encodes a hydrophobic domain. Alternatively, the hybridization pattern could result from a single gene in which exons encoding the polar and non-polar domains are separated by a 5.3 kbp intron. The characterization of several partially processed mRNAs for human b_5 (Yoo and Steggles,1989) shows that a splice junction occurs between residues 86 and 87. This observation is consistent with occurrence of a gene in which functional domains are commonly separated by introns (Gilbert,1978; Rogers, 1985). Processed pseudogenes for human b_5 have also been identified (Yoo and Steggles,1989) and will no doubt complicate the search for active members of the b_5 gene family. The presence of a distinct transcript encoding the erythrocytic form of cytochrome b_5 has now been confirmed by sequence analysis of an erythrocytic cDNA (A. Steggles, personal communication).

B. The Transferrins

The transferrins constitute a family of related iron binding proteins which share a common overall molecular structure, consisting of a two-lobed, single-chain molecule in which each half of the protein maintains a similar metal-binding site. The three principal members of the transferrin family of proteins include serum transferrin, which functions as the primary carrier of iron in the circulation (Shade and Caroline,1946), lactoferrin, a major protein constituent of mammalian milk, tears and other secretions and ovotransferrin, one of the most abundant proteins in avian egg whites. The latter two species presumably function as anti-microbial agents by limiting the supply of free iron (Alderton et al.,1946).

The two iron binding sites display high similarity in both structure and function. When compared at the level of amino acid and nucleotide sequence, the two lobes show high identity with each other, suggesting that an ancestral gene encoding a protein with a single metal binding site underwent a tandem duplication event to give rise to the present two lobed molecule (MacGillivray *et al*,1983). Two-sited transferrin-related proteins have been found in all chordate species examined to date. The organization and chromosomal localization of the genes encoding transferrin family members strongly suggests that an ancestral gene encoding a two site transferrin molecule has undergone successive gene amplification events followed by divergence of both sequence and function to generate the present suite of transferrin-related genes (Bowman *et al.*,1988).

This chapter describes in greater detail the structural and functional properties of the transferrins with particular emphasis on the molecular details made available during the past 10 years from studies on transferrin-encoding gene structures and the solution of the crystal structures for both the human lactoferrin and rabbit serum transferrin molecules. Throughout the remainder of this dissertation, the following abbreviations will be used:

TF: serum transferrin

LF: lactoferrin

OTF: ovotransferrin

melanoTF: melanotransferrin, also p97 and melanoma-related transferrin

TF/2N TF/2C: N and C lobe half molecules of transferrin

Small case letters prefixing any of the above abbreviations will designate the species of origin (i.e. hTF: human serum transferrin).

1. Historical

The presence of non-heme iron in plasma was described as early as 1925 (Fontes and Thivolle, 1925). The globulin fraction of plasma was later shown to be the principal carrier of iron (Starkenstein and Harvalik,1933), though the direct demonstration of a specific iron-binding protein was accomplished only after the development of a precise method for fractionating plasma constituents (Cohn *et al.*,1946). Independent studies by Holmberg and Laurell (1945) and Schade and Caroline (1946) were the first to demonstrate directly the capacity of a specific β-globulin protein to bind ferric iron, which led to the application of some early synonyms for transferrin; siderophilin (Schade *et al.*, 1949) and β1-metal-binding globulin (Surgenor *et al.*,1949). Holmberg and Laurell (1947) distinguished an iron-specific binding protein separate from the copper-binding component of plasma and were the first to suggest the name transferrin as such, a term which has superseded the earlier descriptors.

Transferrins have been found in all chordate species examined to date. Typically, chordates maintain a similar suite of plasma proteins and even primitive members of these phyla, such as the nurse shark and hagfish (MacGillivray,1977, Aisen *et al.*, 1972) , display a two-sited transferrin protein of molecular weight *c.a.* 80 kDa. An examination of the prochordate sea squirt, *Pyura haustor*, by Yang *et al.* (1985)

demonstrated that TF-related sequences are found in the genome of this organism. Specific hybridization signals were identified on Southern blots of genomic DNA from the sea squirt, using either the N- or C-terminal coding regions of the hTF cDNA as a probe. These results may correlate with the presence of a "half-sized" TF-like iron binding protein (40 kDa) described in the related sea squirt *Pyura stolonifera* (Martin, *et al.*,1984) and may represent the single-lobed iron binding protein postulated as the predecessor of modern day transferrins. Since the divergence of human and prochordate lineages is estimated at 400-500 million years b.p. (Doolittle,1984), transferrin-like sequences appear to have been remarkably well conserved throughout evolution.

An iron-binding component of avian egg whites which inhibited microbial growth was first described by Schade and Caroline (1946), and later identified as conalbumin (Alderton *et al.*,1946). Williams (1962) later recognized the similarities between conalbumin and chicken serum TF as the result of comparative studies employing gel electrophoresis, amino acid and carbohydrate composition analysis and immunoelectrophoresis. The term ovotransferrin (OTF) was suggested to reflect these similarities. More recently, the complete identity of the primary sequences of OTF and chicken serum TF has been confirmed by studies showing only a single TF gene within the genome (Lee *et al.*,1978), the only recognized difference between the two species being the carbohydrate composition (Williams, 1962).

The presence of an iron-binding factor in the milk of certain mammalian species was confirmed by the isolation of a salmon-pink coloured protein from bovine milk (Polis and Shmukler, 1953) and from human milk (Johansson, 1958). This protein was shown to be immunologically distinct from serum TF (Johansson, 1960; Montreuil *et.al.*, 1960) though it displayed a similar absorption spectrum and mode of iron binding. Schultze and Heremans (1960) encouraged the use of the term lactoferrin to describe the protein, reflecting the lack of antigenic similarities with serum TF, and thus replaced the terms lactotransferrin, lactosiderophilin and the "red protein". However the recent

solution of the crystal structure for human LF (Anderson *et al.*,1987) and rabbit serum TF (Bailey *et al.*,1988) has shown that the two molecules share a high degree of structural similarity, especially at the iron binding sites while analysis of the amino acid sequence of human lactoferrin (Metz-Boutigue *et al.*,1984) and the partial nucleotide sequence of the corresponding cDNA (Rado *et al.*, 1987) shows a high degree of sequence conservation between these two proteins.

In addition to lactating mammary gland, many other tissues express high levels of LF. Immunological detection has shown LF to be present in tears, bronchial secretions, seminal fluid and cervical mucus (Masson and Heremans,1966) where it likely serves as a bacteriostatic agent (Arnold et al.,1977). Lactoferrin has also been detected in the neutrophil lineage of human and guinea pig leukocytes (Masson et al.,1969) where it may serve a similar role in the acidic environment of these cells.

A recent addition to the transferrin family of proteins is p97, initially identified as a cell surface glycoprotein which is over-expressed in human melanomas (Woodbury *et al.*, 1980; Brown *et al.*, 1980). p97 displays a protease-sensitive domain arrangement similar to other transferrins and also binds iron (Brown *et al.*, 1982) but is not secreted; rather p97 exists as an integral membrane protein (Brown *et al.*, 1981). The predicted amino acid sequence, deduced from the complete cDNA sequence (Rose *et al.*,1986) shows a typical signal sequence followed by two internally homologous domains, typical of other transferrins. A stretch of hydrophobic residues at the carboxy-terminus of the protein likely comprises a membrane-anchoring domain. A cellular function for p97 has not yet been identified as it is difficult to reconcile an iron transport role for an integral membrane protein, however the limited tissue distribution of the protein (melanomas, nevi and fetal intestine) suggests p97 may serve as a useful marker for these tissues.

2. Physiological Role and Expression

Iron is of central importance to the chemistry of life. The ubiquitous requirement for iron is seen in reactions which include respiratory electron flow, the fixation of nitrogen and oxygen and ligand binding by proteins such as hemoglobin and myoglobin. The reversible reduction of ferric iron is often utilized by these systems to effect energy conservation and transfer, though under physiological conditions, Fe^{III} is the stable state. As noted by Aisen and Liebman (1973), the low solubility product of $Fe(OH)_3$ (10^{-36} M) dictates that free iron in solution cannot exceed 10^{-17} M. As such, organisms have adapted several siderophilic molecules to perform the function of maintaining a high concentration of iron in soluble form; in chordates transferrins are the principal iron-binding proteins. In addition, the chelation of iron prevents iron-intoxication and minimizes urinary loss of iron (Putnam,1975).

The primary function of serum transferrin is to transport ferric iron from acquisition and storage organs to tissues where it is required. In human serum, transferrin binds almost all available iron, though this amount (3-4 mg) represents less than 0.1% of the body iron stores and leaves serum TF only thirty percent saturated (Gitlow and Beyers,1952). The concentration of TF in serum is traditionally represented as a total iron-binding capacity (TIBC) and is calculated as the number of micrograms of iron bound by 100 mL of serum (Laurell,1960).

The most studied model of iron transport remains the transferrin-reticulocyte system. Hemoglobin iron turnover accounts for almost 30 mg per day and reticulocytes express an estimated 300,000 TF receptors per cell in order to accommodate demand (van Bockxmeer and Morgan,1979). The TF receptor molecule is well characterized, identified first by classical membrane biochemistry (see Newman *et al.*,1982) and later sequenced by characterization of genomic and cDNA clones (Kuhn *et al.*,1984; McClelland *et al.*,1984). The receptor consists of a homodimer of two disulfide-linked monomers of molecular weight 95 kDa each. Diferric TF binds with an apparent

affinity of 10^9M^{-1}, with apoTF showing a 30-fold weaker affinity (Young and Aisen,1981). The complex is clustered into clathrin-coated pits (Watts,1985), likely as a result of interactions between the cytoplasmic tail of the receptor with coat proteins (Iacopetta *et al.*,1988). Invagination of these pits forms vesicles which then fuse with endosomes where conditions of low pH (5.5) cause the release of iron from TF (see Kuhn *et al.*,1990). Transferrin is a true carrier molecule in that TF:TF receptor complexes are recycled to the cell surface (Harding *et al.*,1983; Hopkins and Trowbridge,1983).

Studies on several cultured cell types indicates that the limiting factor in cellular iron uptake is the number of TF receptor molecules on the membrane surface (Ciehanover *et al.*,1983; Iacopetta and Morgan,1983; Hebbert and Morgan,1985). Regulation is achieved by modulating the receptor mRNA activity through so-called iron responsive elements (IRE) in the 3' untranslated region of the mRNA (Casey *et al.*,1988). IREs are palindromic sequences which bind a regulatory factor(s), leading to a stabilization of the mRNA molecule (Mullner *et al.*,1989; Rouault *et al.*,1988). This binding activity is iron-sensitive and thus links the translation of TF receptor mRNA to the level of available iron (Kuhn *et al.*,1990).

In contrast, TF expression is only slightly sensitive to iron levels, showing a 2-fold increase in mRNA as a result of increased gene transcription (McKnight *et al.*,1980a;1980b). However, OTF gene regulation by steroid hormones is well documented in the chick oviduct, where estrogen has been shown to effect an 8-fold rise in TF mRNA levels by increasing the rate of TF gene transcription in the primary oviduct, though the same treatment increases TF synthesis in the liver by only 2-fold, suggesting a tissue-specific component must also be involved (Oka and Schimke,1969; McKnight and Palmiter,1979). In the Sertoli cells of rat testis, follicle stimulating hormone, insulin, retinol and testosterone are all capable of increasing TF mRNA levels. Homology searches of the upstream region of the TF gene have identified several

potential hormone response elements (Lucero *et al.*,1986; Adrian *et al.*,1986) though their status as regulators of gene activity is still unclear.

The liver is the primary site of serum protein synthesis, including transferrin (Morgan,1983). The half life of the protein is approximately 8 days, suggesting a daily turnover of one gram (Jarnum and Lassen,1961). Recent studies have used cell culture model systems in attempting to identify factors responsible for the liver-specificity of TF synthesis. Promoter analysis of the TF gene has shown the region from -650 to the transcriptional start site as essential for hepatocyte specific activity and several protein-binding domains close to the start site have been demonstrated by footprint analysis (Brunel *et al.*,1988).

In addition, TF is highly expressed in the Sertoli cells of the testis, the fetal yolk sac and the choroid plexus, all examples in which TF is synthesized in tissues adjacent to physiologic enclosures (i.e. the testes ducts, placenta and cerebrospinal system, respectively) in order to maintain homeostasis within the compartment (see Funk and MacGillivray,1990; Bowman *et al.*,1988). TF expression has also been demonstrated in T-lymphocytes, submaxillary gland, ovary, spleen, bone marrow, lymph nodes and thymus (Lum *et al.*,1986; Thorbecke *et al.*,1973; Phillips and Thorbecke,1966; Meek and Adamson,1985)

Transferrin has also been demonstrated to act as a potent antimicrobial agent (Schade and Caroline,1946) though LF is a much more likely candidate for this role, owing to its wide distribution in body fluids, tighter iron binding and greater acid stability (Mason *et al.*,1965,1966). In concordance with this role, lactoferrin has been identified in many biological fluids including tears, bronchial secretions, seminal fluids and cervical mucus, and thus represents a major secreted protein of endothelial cells (Masson and Heremans,1966; Arnold *et al.*,1977).

3. Fundamental Properties

i) Physiochemical and Conformational Characteristics

The relatively high concentration of transferrin in serum and the central role played by the protein in iron transport led to intensive studies on the isolation and characterization of serum TF soon after the initial description by Holmberg and Laurell (1945). Serum TF belongs to the β-globulin fraction, having a typical serum concentration of 250 mg/100mL. Some of the earliest purification methods involved crystallization of TF directly from ethanol fractions of serum (Surgeoner *et al.*,1949; Koechlin, 1952). Later, methods were developed for large-scale purification, involving the use of ammonium sulfate, ethanol or 2-ethoxy-6,9-diaminoacridine lactate (Rivanol) precipitations, ion exchange and gel exclusion chromatography or combinations of these techniques (see for example Inman *et al.*,1961; Schultze and Schwick, 1957).

Serum TF was readily identified by electrophoretic methods, such as the discontinuous buffer starch gel system described by Poulik (1957) or in combination with radioautography using ^{59}Fe (Giblett *et al.*, 1959). However, estimates of the molecular weight of serum transferrin differed widely during this period. Sedimentation diffusion and sedimentation equilibrium methods were often employed, resulting in initial estimates between 68-90 kDa for human serum TF (Schultze and Schwick, 1957; Charlwood, 1963). Other methods, such as gel filtration, intrinsic viscosity, iron-binding capacity and $NaDodSO_4$ gel electrophoresis arrived at values within this range (see Putnam, 1975), though it was not until more accurate sedimentation equilibrium measurements became available that a value of approximately 77 kDa was generally accepted (Mann *et al.*, 1970; Aisen *et al.*, 1970). Other accepted physical properties of serum TF include: a partial specific volume of 0.725 (Oncley *et al.*, 1947), a diffusion coefficient of 5.4 Fick units (Bezkorovainy and Grohlich, 1967), and a frictional ratio (f/fo) of 1.37 (Oncley *et al.*,1947). The latter determination suggests an asymmetric shape for the molecule, in agreement with later studies which determined the protein

shape as ellipsoidal, with a major to minor axis ratio of 2:1 (Rosseneu-Moutreff *et al.*, 1971).

Evidence from many studies show a significant conformational change in transferrin molecules upon binding of iron. Ferric complexes of TF and OTF are more resistant to denaturation and have reduced susceptibility to proteases (Azari and Feeney 1958; 1961; Yeh *et al.*,1979) and hydrogen-tritium exchange (Ulmer,1969). The major to minor axis ratio increases from 2.0:1 to 2.5:1 (Rosseneu-Montreff *et al.*,1971) which confirms recent crystallographic evidence showing an open configuration for the metal free N lobe of apoLF (Baker *et.al.*,1989). As well, saturation with iron lowers the isoelectric point of TF from 5.80 to 5.45 (Keller and Pennell,1959).

ii) Carbohydrate Composition

The carbohydrate content of human serum TF was initially reported by Heimburger *et al.* (1964) and later studies involving neuraminidase treatment demonstrated the presence of terminal sialic acid residues (Blumberg and Warren,1961; Poulik, 1961). The complete sequence of the oligosaccharide was later reported by Jamieson *et al.* (1971) though later investigations involving chemical and enzymatic studies (Spik *et al.*,1975) and nuclear resonance spectroscopy (Dorland *et al.*,1977) and mass spectroscopy (Karlsson *et al.*,1978) have identified a different biantennary structure (Figure 9). The number of heteropolysaccharide chains varies between transferrins of different species (Graham and Williams, 1975). Human serum TF contains two identical chains, linked to asparagine residues 415 and 608 in the C lobe of the protein (MacGillivray *et al.*,1977).

As yet no functional role has been ascribed to the carbohydrate moieties of TF. Enzymic removal of terminal sialic residues with neuraminidase does not result in gross changes in conformation (Bezkorovainy,1966), nor does it interfere with iron binding

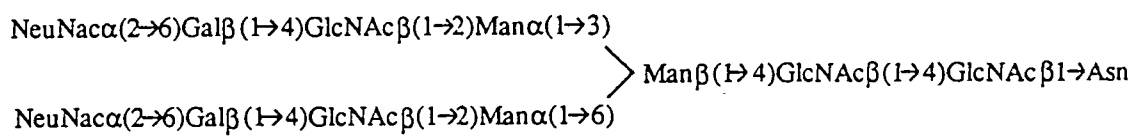


Figure 9: Structure of the Glycan Chains of Human and Rabbit Serum Transferrin.

As determined by Leger *et al.* (1978), the predominant carbohydrate structure of serum transferrin is doubly branched (biantennary), with each branch terminating with a sialic acid residue. The two carbohydrate chains of human serum TF are both attached to the C lobe via asparagyl linkages to residues 415 and 608.

(Morgan *et al.*,1967). Transferrin differs from most other plasma glycoproteins in that neuraminidase treatment does not enhance clearance by the liver (Ashwell and Morell,1974). It has been suggested that oligosaccharide may function to enhance the secretion of proteins to which they are attached, however treatment of rat hepatocytes with tunicamycin causes only a slight decrease in TF secretion even though only 25% of TF is glycosylated under these conditions (Struck *et al.*,1978). Attempts to demonstrate a role for the carbohydrate moiety in the binding of TF to its receptor have not been successful.

iii) Amino Acid Sequence Analysis

The amino acid composition of human serum TF has been the subject of many investigations (Parker and Bearn,1962; Roop and Putnam,1967; Palmour and Sutton,1971; Sutton and Brew,1974). These studies indicated the protein consisted of approximately 630 amino acid residues, with a calculated molecular weight of 76 kDa. Other significant findings from composition analysis studies include a complete absence of free sulfhydryls and a substantially greater proportion of basic amino acids in LF as

compared to serum TF, a fact reflected in the higher pI of LF (pI=5.3 *versus* 5.8 for human serum TF) (Putnam, 1975).

Amino-terminal sequence analysis from several species provided clear evidence of the high degree of sequence conservation in the transferrin family (see Graham and Williams,1975; Jolles *et al.*,1976). Partial sequences for tryptic and cyanogen bromide fragments from human TF, OTF and LF were complicated by the presence of multiple half-cystine residues (see Putnam, 1975), however MacGillivray *et.al.* (1977) were able to show conclusively the presence of a high degree of internal sequence identity between the two lobes of human serum TF. The same group later reported the complete primary sequence of the protein which indicated a remarkable 40% identity between residues of the N and C lobes (MacGillivray *et al.*,1983), providing indisputable evidence for at least one gene duplication event in the evolutionary past of TF. The complete primary sequence of human LF (Metz-Boutigue *et al.*,1984) similarly showed a high degree of conservation (37% between the two lobes) and an even higher degree of identity between the complete human serum TF and LF sequences (59%) (Figure 10).

iv) Metal Binding Characteristics

The binding of two atoms of ferric iron by each TF molecule involves the concomitant binding of two carbonate anions (Schade *et al.*,1949) resulting in a salmon-pink complex characterized by an absorption maximum at 465 nm (Schade and Caroline,1944,1946; Cohn,1947). As the precise nature of the iron binding capacity of melanotransferrin has yet to be resolved, the following discussion will be restricted to serum TF, OTF and LF.

The available crystal structures for human LF (Anderson *et al.*,1987;1989) and rabbit serum TF (Bailey *et al.*,1988) have clearly identified the protein ligands to the iron (consisting of two tyrosyl, one histidyl and one aspartyl residue) and have confirmed a bidentate carbonate anion as the bound anion (see IB6. Iron Binding Site and

	hSTF	rSTF	hLF	cOTF
hTF	41	--	--	--
rTF*	81	N.D.	--	--
hLF*	59	57	37	--
cOTF	51	50	49	33

* 3D-crystallographic structures for human lactoferrin (Anderson et al.,1987) and rabbit serum transferrin (Bailey et al.,1988) have been reported.

Figure 10: Amino Acid Sequence Identity Between Transferrin Family Members.

Primary sequence data for the mature forms of human serum transferrin (hTF) (Yang et al.,1984), human lactoferrin (hLF) (Metz-Boutigue et al.,1984), chicken ovotransferrin (cOTF) (Jeltsch and Chambon,1981) and the first 397 amino acids of mature rabbit serum transferrin (rTF) (D.K. Banfield, unpublished results) were compared using the method of Meyers and Miller (1988) and are reported as percentage amino acid identities. Comparisons of the N and C lobes of individual species appear on the diagonal.

Structural Considerations). Previous physiochemical studies however have implicated additional residues as being involved in the iron binding process, several of which are described below.

The iron:TF complex is sensitive to pH, being stable in the range from 7.5 to 10, but unstable below 6.5 (Ehrenberg and Laurell,1955). The binding affinity is extremely tight; under physiological conditions the K_a of the N lobe binding site for ferric iron estimated at $10^{22}M^{-1}$ with the C lobe site having approximately five-fold greater affinity (Aisen and Leibman,1978). Indeed this binding is so strong that iron bound at either site is effectively locked, such that kinetic, rather than thermodynamic considerations govern the relative occupancy of the two sites (Aisen and Brown,1975a). The binding of iron to lactoferrin has been estimated to be 260-fold stronger than serum TF, and

displays considerable stability at acidic pH (2-4.5) (Aisen and Leibman,1972), being fully dissociated only at pH 2 (Johansson,1960).

Transferrin also binds a variety of transition and lanthanide series metals, including Cu^{+2}, Zn^{+2} (Cohn,1947), Cr^{+3}, Mn^{+2}, Co^{+2}, Ni^{+2}, Cd^{+2}, (Tan and Woodworth,1969; Aisen *et al.*, 1969), Ga^{+3} (Woodworth *et al.*,1970), Tb^{+3}, Eu^{+3}, Er^{+3}, Ho^{+3}, Nd^{+3}, Pr^{+3} (Luk, 1971). The relative affinity of several of these metals was determined by Tan and Woodworth (1969) as:

$$Fe^{+3}\ Cr^{+3} >\ Cu^{+2}\ Mn^{+2} > Co^{+2} > Cd^{+2}\ Zn^{+2}\ Ni^{+2}$$

Complexes with chromium, manganese and cobalt show binding of a molar equivalent of carbonate. Though the ferric complex is likely the only one of physiological significance, reports of TF in serum binding chromium (Hopkins and Schwartz,1964), zinc (Harris,1983) and manganese (Aisen *et al.*,1976) have been reported. Ferrous iron is bound very weakly to TF, if at all (Gaber and Aisen,1970).

Prior to the resolution of detailed crystal structures, much emphasis in TF research was placed on the identification of the binding site ligands. These studies used many different techniques, including NMR spectroscopy, resonance-enhanced Raman spectroscopy, proton titration, EPR spectroscopy and several chemical modification approaches (see Aisen and Brown,1975b; Chasteen, 1977). Candidate ligands included up to three phenolic groups (Warner and Weber, 1953), two histidine imidazole groups (Aasa *et al.*,1963) and possibly a tryptophan indole ring (Tomimsatsu *et al.*,1973). Though the identity of the binding ligands is now conclusive, residues removed or even distant from the metal may yet prove to be critical for normal iron binding. For example a Gly65Arg variant of serum transferrin has a characteristic yellow absorption maximum, likely due to an alteration of the electronic environment near the Asp63 ligand (Evans and Madden,1984). The third tyrosine residue implicated in several earlier studies may prove to be Tyr83 (hLF numbering system), as this residue is invariant in both lobes of all TF species and is situated within 6Å of the metal (Baker *et al.*,1987) Residues

involved in fixing the anion would also be crucial for normal iron binding, as evidenced by chemical modification studies in which an essential arginyl residue participates at the metal binding site (Rogers *et al.*,1977), a feature confirmed by crystal structure determinations. Nuclear magnetic resonance spectroscopic studies on whole molecule transferrin species (Markley,1975; Alsaadi *et al.*,1987; Woodworth,1986) and half molecules (Valcour and Woodworth,1987) have shown histidyl residues to be directly affected by iron binding.

The presence of a suitable anion is an absolute requirement for specific iron binding (Bates and Schlabach,1975). The binding of metal and anion is cooperative, as neither is bound in the absence of the other (Price and Gibbons,1972; Aisen *et al.*,1973), though many small anions can be substituted for carbonate in the complex. These include oxalate, malonate, EDTA, glycinate, thioglycolate and nitrilotriacetate (Aisen *et al.*,1967; Woodworth *et al.*,1975, Schlabach and Bates,1975). These are all molecules that share the structural characteristic of having at least one carboxylate function and a second functional group in close proximity. These functionalities likely act as metal ligands, reflected in the finding that the ligand fields and iron-binding affinities vary with the bound ligand (Pinkowitz and Aisen,1972). In addition, the synergistic anion also functions to shield the ferric ion from local positive charge within the binding site. As shown in the hLF crystal structure (Anderson *et al.*,1987), the carbonate dianion is in close proximity to the invariant Arg124 side chain and additional positive charge is provided by a dipole effect from the N-terminus of nearby α-helix 121-137. NMR studies have also shown anion binding likely precedes metal binding (Woodworth,1975).

Proteolytic digestion of several species of serum transferrin and lactoferrin releases a small (20 kDa) iron-binding fragment which includes domain II of the protein as well as portions of the bridging strands (Evans and Madden, 1984; Legrand *et al.*,1984). These polypeptides do not contain the Asp ligand provided by domain I, and display yellow-shifted absorption maxima and reduced iron-binding affinities. The

presence of a synergistic anion is required for iron binding and binding to these fragments is specific and reversible (Legrand *et al.*,1990).

A plausible model for iron binding suggests that the initial complex includes the anion, metal ion and ligands contributed from domain II and the bridging strands. Hinge-like movement of the two domains closes the lobe as the Asp ligand coordinates, thereby locking the metal in a tight complex (Baker *et al.*,1987). Several mechanisms may effect iron release, including protonation of the side-chain ligands, direct attack on the anion, chelation *via* ligand exchange and reduction of the ferric ion (Aisen and Listowsky,1980)

v) The Two Lobes of Transferrin

The preparation of 35 kDa iron-binding fragments from OTF by limited proteolysis demonstrated conclusively the presence of two structurally distinct lobes in transferrin molecules (Williams 1974,1975). This finding was later confirmed by low-resolution x-ray diffraction data which showed two lobes of similar size and structure, best approximated by two ellipsoids of revolution inclined at 30° to each other (Gorinsky *et al.*,1979). The independence of metal binding by the two lobes was demonstrated conclusively by the introduction of the urea-polyacrylamide electrophoresis procedure of Makey and Seal (1976) which allowed the resolution of mono-ferric forms of transferrin molecules which have slightly different electrophoretic mobilities.

A large number of studies have sought to determine whether the two lobes of TF differ in functional properties. An initial impetus for this work was the reported finding of Warner and Weber (1953) of strong cooperativity between the two metal-binding sites ($K_2=100K_1$), though this finding was later refuted by the demonstration of distinct mono-ferric forms of the molecule (Aisen and Listowsky,1980). However many investigators have now reported slight but real differences between the N and C lobes (Brock,1985a). EPR studies of Cr^{+3}- (Aisen *et al.*,1969) and VO^{+2}-loaded TF (Cannon

and Chasteen,1975) showed clear differences between the two sites as have Mossbauer and optical spectroscopic measurements (Aisen *et al.*,1973). The pH-dependence of iron binding also differs. The C lobe is more susceptible to proton attack with a five-fold lower affinity at pH 7.4, but a 33-fold lower affinity at pH 6.7 (Lestas,1976; Princiotto and Zapolski,1975). In addition, the two lobes differ in their ability to bind large ionic radii metals such as Nd^{+3} and Pr^{+3} (Luk,1971) and various anions, such as oxalate (Ainscough *et al.*,1983). Structural studies have shown differences between the two lobes near the distal regions of the iron-binding clefts which may affect the accessibility of the two sites for small molecules, while the presence of an additional disulfide bond in the C lobe likely accounts for the greater stability and higher iron affinity of that lobe (Baker *et al.*,1987).

The production of a "half-molecule" of transferrin, corresponding to the N lobe of the intact molecule was first achieved by Williams (1974), who capitalized on the finding that partially saturated OTF selectively loads the N lobe site, leaving the vacant C lobe in a loose conformation which is susceptible to proteolytic digestion. The resulting N lobe fragment (OTF/2N) maintains the ability to bind metals and was also reported to donate iron to reticulocytes. The isolation of the corresponding C lobe (OTF/2C) was achieved by tryptic digestion of iron-saturated OTF under mildly acidic conditions, whereupon the N lobe is selectively relieved of iron (Williams,1975). The OTF/2C lobe was also shown to bind metals and donate iron to reticulocytes and also carries the carbohydrate moiety. Isolation of the corresponding half molecules of human serum TF (Evans and Williams,1978; Lineback-Zins and Brew,1980; Zak and Aisen,1985) and bovine serum TF (Brock *et al.*,1978) followed; however, individual bovine half molecule species were incapable of donating iron to rabbit reticulocytes, whereas diferric bovine TF could donate iron. A similar result was later reported by Brown-Mason and Woodworth (1984) who, in contrast to previous studies, failed to observe donation of iron to reticulocytes by OTF half molecules, suggesting that each lobe of the OTF

molecule contains a distinct recognition region necessary for receptor binding and subsequent iron donation.

4. Nucleotide Sequence Analysis

Investigations carried out over the past ten years on the gene and cDNA sequences of transferrin family members has proven valuable in identifying mechanisms by which gene families evolve. Comparison of coding sequences, exon organization and chromosomal localization have provided finer resolution of the lineage of the transferrin gene family and have confirmed that both intragenic duplication and gene amplification events have occurred during the evolution of modern transferrins. The availability of DNA probes has also augmented studies of the expression patterns of transferrins, the results of which have been recently reviewed (Bowman *et al.*,1988, Funk and MacGillivray,1990). The following section relates some of the important findings of studies on the molecular biology of transferrins, with particular emphasis on the evolution of this gene family.

i) Chicken Ovotransferrin

The chicken OTF gene was the first transferrin species investigated at the nucleotide level. Southern blot analysis of chicken genomic DNA demonstrated that serum TF and OTF are the product of a single gene consisting of at least 17 exons separated by at least 16 introns (Cochet *et al.*,1979). This observation has now been confirmed by the determination of the complete nucleotide sequence of the ovotransferrin gene (Jeltsch and Chambon,1979), and provides direct evidence that the internal homology between the amino- and carboxy-terminal domains arose as the result of the duplication of an ancestral iron binding protein, as had been suggested previously for human serum transferrin (MacGillivray and Brew,1975). The suggested scheme for this event is identical to that proposed for the evolution of the human serum

transferrin gene (Park *et al.*,1985), involving the intragenic duplication of exons 2-9 from the 9 exon ancestral gene followed by the gain of exon 11 in the second domain (see Figure 11). Several conserved histidyl and arginyl residues in human transferrin have been proposed to contribute directly to the iron binding sites, and their occurrence has been postulated to result from an even more ancient quadruplication event, such that a primordial iron binding structure originally duplicated to yield the first lobe of transferrin, then duplicated again to produce the present structure (MacGillivray *et al.*,1975). When the coding sequences for these fourfold homology blocks in the chicken ovotransferrin cDNA are compared, the sequence identity between any pairs varies from 37 to 67%, a value far higher than would be expected at random, considering that the proposed first duplication in this scheme would necessarily be dated well prior to 500 million years b.p. (Bowman *et al.*,1988) (see Figure 12). However these residues do not occupy "equivalent" positions in the domains of a given lobe (Baker *et al.*,1987) and thus this finding cannot be substantiated through structural studies.

The promoter region of the ovotransferrin gene displays a typical eukaryotic TATA box at position -31 (Cochet *et al.*,1979). Chicken ovotransferrin mRNA is 2376 nucleotides in size and contains a base-paired loop (nucleotides 2-29) of unknown function at the 5' end. This is followed by a short region complementary to 18S rRNA. A polyadenylation site is found at the distal end of the 182 bp 3'untranslated region. Comparison of the gene and cDNA sequences has revealed several sequence dissimilarities resulting in 3 predicted amino acid differences. These are the most likely causes of the known polymorphisms found in ovotransferrins (Williams *et al.*,1982).

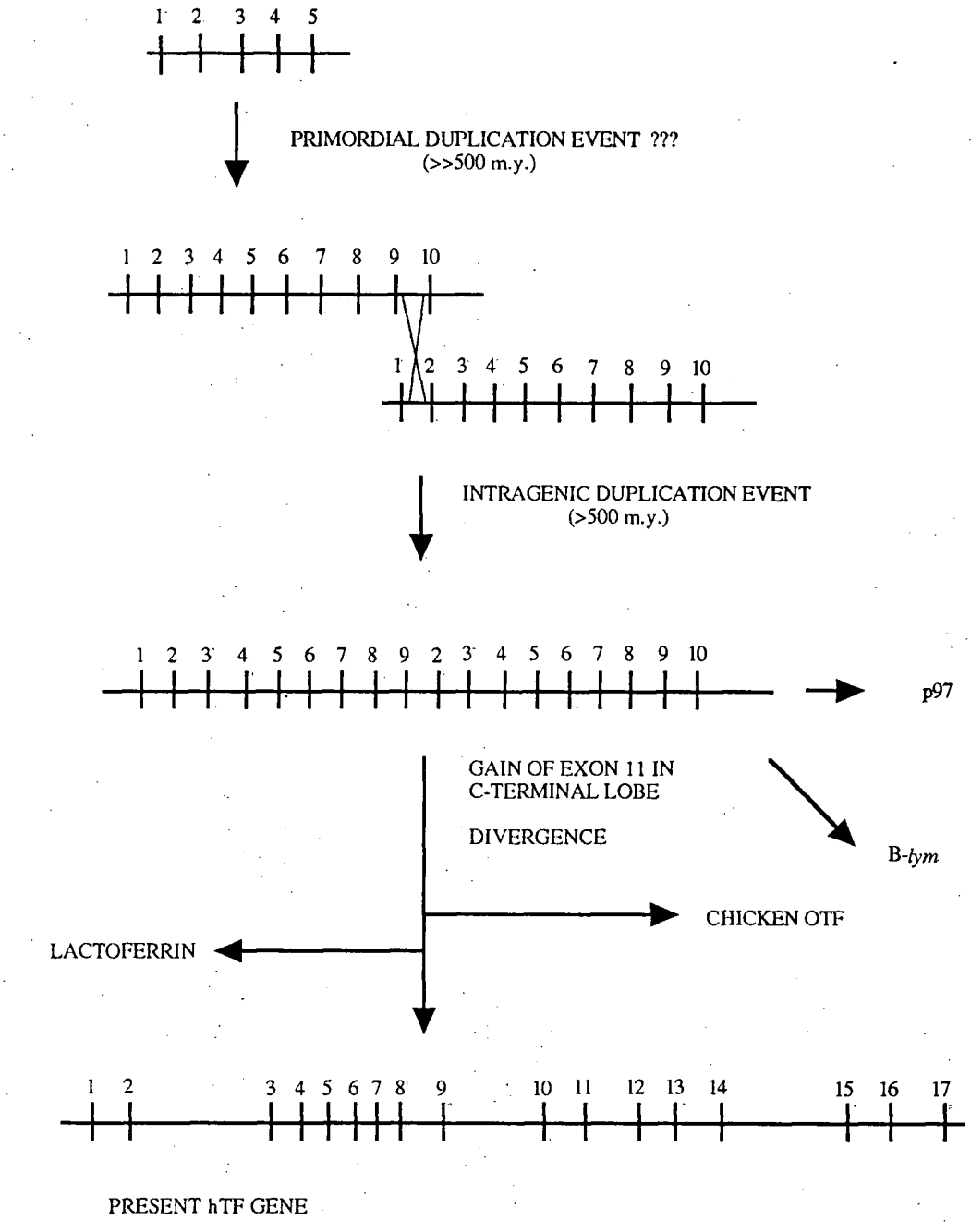

1 2 3 4 5
PRIMORDIAL DUPLICATION EVENT ???
(>>500 m.y.)
1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10
INTRAGENIC DUPLICATION EVENT
(>500 m.y.)
1 2 3 4 5 6 7 8 9 2 3 4 5 6 7 8 9 10
p97
GAIN OF EXON 11 IN
C-TERMINAL LOBE
DIVERGENCE
B-lym
CHICKEN OTF
LACTOFERRIN
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
PRESENT hTF GENE

Figure 11: Possible Scheme for the Evolution of the Present Human Serum Transferrin Gene (based on Park et al.,1985).

Certainly one, and possibly two intragenic duplication events have occurred to give rise to present day transferrin genes. The presence of a "half-sized" transferrin-like protein in prochordates (Bowman *et al.*,1988) suggests the precursor iron-binding molecule was present prior to the divergence of this lineage from the vertebrates (500 m.y.). Sequence comparison of the two domains of serum transferrin provides weak evidence that this single lobe molecule may itself be the result of an even earlier duplication event (see Figure 12). Following the assemblage of a two-lobed transferrin gene, this progenitor sequence gained an additional exon within the C-terminal lobe, followed by several gene amplification events which gave rise to the modern suite of transferrin-related gene sequences.

---

ii) Human Serum Transferrin and Lactoferrin

The overall organization of the human gene is the same as that seen in the chicken gene (Park *et al.*,1985: Breathnach *et al.*,1978). When compared to their ovotransferrin counterparts, the 17 exons show very similar sizes, and the intron/exon splicing pattern is identical for the homologous sequences in the two domains of the human transferrin sequence, providing further evidence for the gene duplication event. The human gene spans 33.5 kbp. Virtually all of the size difference compared to the ovotransferrin gene is because the introns are larger in the human gene. The human transferrin mRNA is 2.3 kbp in size, including 54 bp of 5' untranslated sequence and 170bp of 3' untranslated sequence (Williams *et al.*,1982: Schaeffer *et al.*,1987). The 5' flanking region of the human transferrin gene contains a near-canonical CAAT transcription factor (CTF) recognition sequence (Jones *et al.*,1987) at position -187 relative to the transcriptional start site. This region binds purified CTF protein (Brunel *et al.*,1988). Other consensus regulatory sequences have been identified in the 5' flanking region of the human transferrin gene, but their functional significance is unclear (Lucero *et al.*,1986).

The cloning of a partial cDNA containing approximately 40% of the coding sequence for human lactoferrin has been reported (Rado *et al.*,1987), and it shows almost complete agreement with the previously reported protein sequence.

A

I	116	HisThrGlyLeuGlyArgSerAlaGlyTrpAsnIleProIleGlyThrLeuLeu	133
II	250	HisAlaValValAlaArgAspAspAsnLysValGluAspIleTrpSerPheLeu	267
III	455	HisThrAlaValGlyArgThrAlaGlyTrpValIleProMetGly---LeuIle	471
IV	592	HisAlaValValValArgProGluLysAlaAsnLys---IleArgAspLeuLeu	608

I	479	CACACGGGGCUGGGCAGGUCUGCGGGCUGGAACAUCCCCAUUGGGACACUCCUC	532
II	881	CACGCCGUCGUGGCUCGGGAUGACAACAAGGUUGAAGAUAUCUGGAGCUUCCUC	934
III	1496	CACACCGCUGUGGGGAGGACUGCUGGCUGGGUCAUCCCCAUGGGC---UUGAUU	1546
IV	1907	CACGCUGUGGUCGUGCGCCCGGAGAAAGCAAACAAA---AUCCGUGAUCUGCUG	1957

B

COMPARISON OF THE FOUR REGIONS

	I/II	I/III	I/IV	II/III	II/IV	III/IV
LENGTH IN NUCLEOTIDES	54/54	54/51	54/51	54/51	54/51	51/51
NUMBER OF CODON DELETIONS	-/-	-/1	-/1	-/1	-/1	1/1
NUMBER OF IDENTICAL NUCLEOTIDES	24	36	24	24	25	20
% NUCLEOTIDE IDENTITY	44	67	44	44	46	37

Figure 12: Sequence Comparison of the Four Domains of Ovotransferrin (from Jeltsch and Chambon, 1981)

Panel A: The amino acid and nucleotide sequences of Domains I-IV of chicken OTF have been aligned to demonstrate the putative four-fold homology of transferrin molecules. Single codon gaps in Domains III and IV were introduced in order to maximize the homologies. Similar homologies have been described for human serum TF (MacGillivray *et al.*, 1977).

Panel B: Summary of the nucleotide sequence identities from Panel A. Nucleotides which occur in at least three of the Domains are scored as identical.

The genomic arrangement of the transferrin gene and related gene species reflects their common evolutionary ancestry, as the human genes for transferrin, lactoferrin, p97 and the transferrin receptor are all found on the same chromosome.

Hybridization of a labeled human transferrin cDNA to metaphase chromosome spreads localized the human transferrin gene to the 3q21-25 region of chromosome 3 (Yang *et al.*,1984). The long arm of chromosome 3 also contains the genes for the transferrin receptor (Goodfellow *et al.*,1982), p97 (Plowman *et al.*,1983) and lactoferrin (Teng *et al.*,1987) (see Funk and MacGillivray,1990).

These results clearly indicate a common ancestry for members of the transferrin gene family. Presumably, the gene for a primordial iron-binding protein has been duplicated several times. Initially, a gene duplication gave rise to the present day transferrin molecule consisting of two homologous domains each with an iron-binding site. Subsequent gene duplications and divergence then gave rise to the other members of the transferrin gene family such as lactoferrin and p97 (see Figure 11).

5. Molecular Structure of Transferrin and Lactoferrin

The recently published crystallographic structures of iron-saturated human lactoferrin (Anderson *et al.*,1987) and rabbit serum transferrin (Bailey *et al.*,1988) have corroborated and clarified many of the previous physiochemical and structural studies of transferrin molecules. As expected, the two available crystal structures clearly define two distinct iron-binding lobes in each molecule connected by a short helical linker sequence, and have allowed absolute determination of the iron ligands at each site. The initial description the hLF structure was subsequently used to aid in the interpretation of a 3.3Å resolution map for the rabbit TF structure and though the complete primary sequence for the rabbit serum TF molecule is currently unavailable, it has been possible to extract a detailed structural analysis of serum Tf as well.

i) Human Lactoferrin

The human lactoferrin crystallographic structure provided conclusive evidence that the high degree of primary sequence identity between the two halves of the coding

sequence is conserved in the overall three dimensional arrangement of the protein. Each lobe contains a single iron binding site with four amino acid side chains (Asp61, Tyr93, Tyr191 and His252 in the N lobe; Asp407, Tyr447, Tyr540 and His609 in the C lobe) contributing ligands to the metal (Figure 13). Further refinement of the structure has conclusively identified a bidentate carbonate ion filling the remaining two coordination positions (Anderson *et al.*,1989).

The two lobes of the molecule are arranged in an approximately antiparallel configuration, such that one lobe can be superimposed following a rotation of 180° and translation of 25Å (see Figure 14 for the equivalent organization of rabbit serum TF). Each lobe in turn is composed of two structural domains (termed N1,NII, CI and CII after the nomenclature proposed by Anderson *et al.*,1987)(see Figure 14). In examining the structural organization of the N lobe (Figure 15), the alpha-carbon backbone begins in domain NI (residues 1-90) as a series of β strands with connecting loops and helices, passes into domain NII via a strand composed of residues 78-100, continues as domain NII (residues 91-251), returns to domains NI via a second connecting strand (244-257) and completes the lobe in domain NI (residues 252-320) A final short helix (residues 320-331) runs across the lobe towards domain NII and leads into the connecting helix which joins the lobes. Minor differences in the topology of the N and C lobes are restricted to the external loops and do not affect the overall folding pattern. The N lobe contains 6 disulfide bridges, and these are maintained in equivalent positions in the C lobe suggesting that this pattern arose prior to the gene duplication event that gave rise to modern transferrins. In addition the C lobe contains four additional cystine residues, two of which are involved in folding the terminal residues back over the surface of the C lobe. None of the disulfides span between the two lobes, nor do they bridge across the two domains except at or near the strands which connect domains. Carbohydrate attachment sites have been identified at Asn137 and Asn490, though the density map does not allow for accurate structural definition of the glycan moiety. The overall

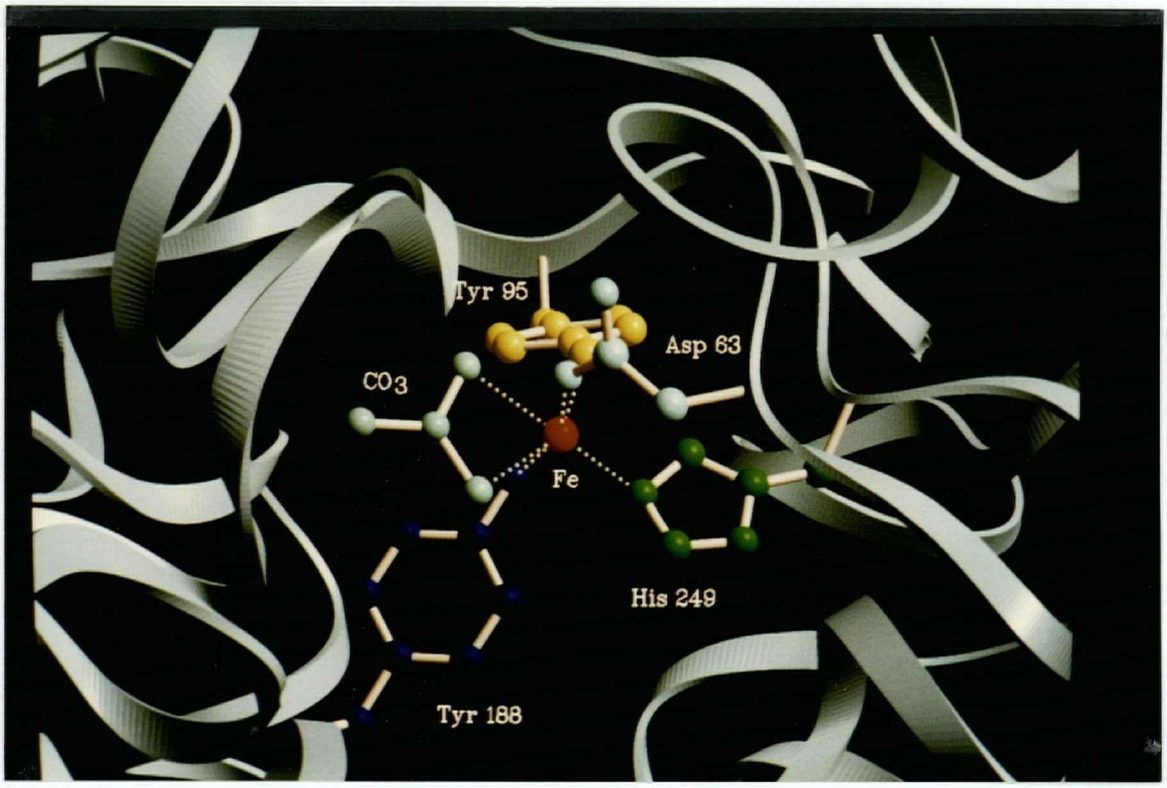


Figure 13: Metal Ligands of Transferrin

The four side-chain ligands and carbonate dianion of the N-terminal lobe of rabbit serum transferrin are shown as ball-and-stick representations, with the surrounding peptide backbone represented as ribbon strand. Tyr188 is contributed from Domain II, Tyr95 and His249 are from the bridging strands which connect the two domains, and Asp63 originates from Domain I.

---

configuration of each lobe therefore is hinge-like, involving two domains which are restricted in conformation by multiple cystines and are connected by two flexible strands.

ii) Rabbit Serum Transferrin

The general organization of rabbit serum transferrin is the same as that of human lactoferrin. The alpha-C chain follows the same path as that described for hLF with

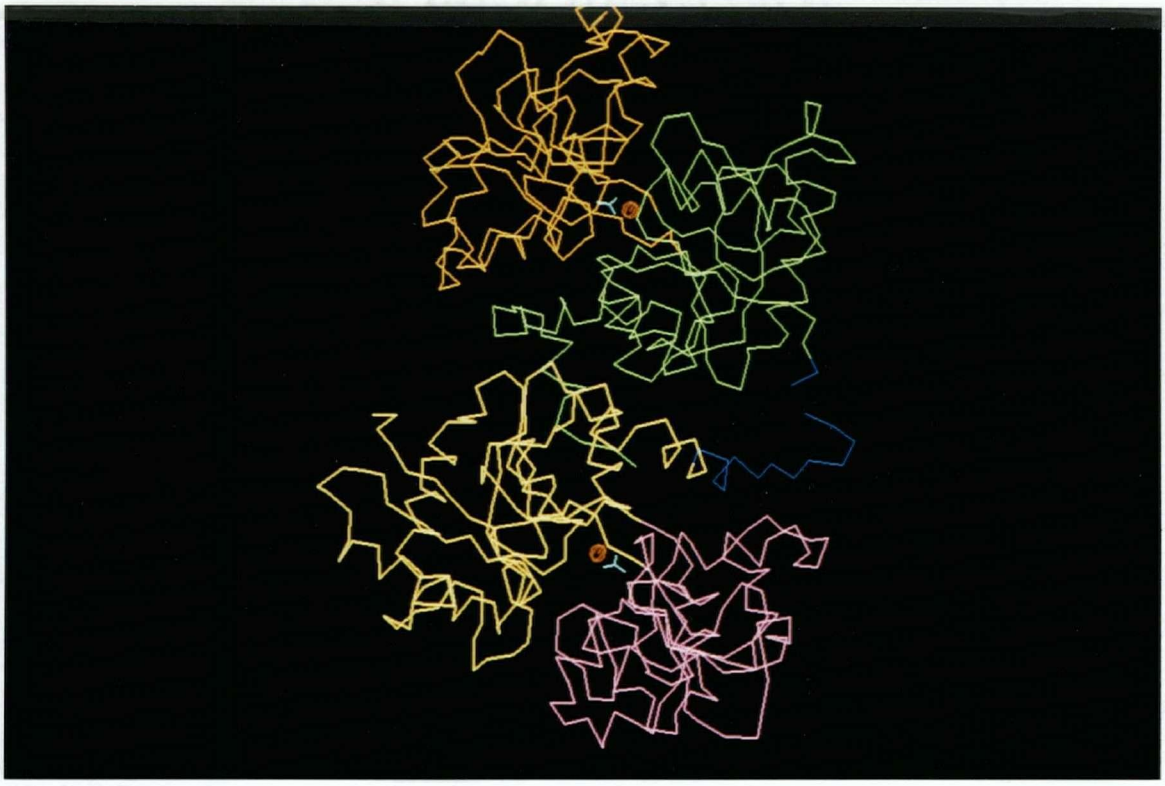


Figure 14: Overall Molecular Organization of Rabbit Serum Transferrin

The four domains of rabbit serum TF are demonstrated in this α-carbon representation. Domains NI (yellow) and NII (pink) of the N-lobe are connected with Domains CI (green) and CII (orange) of the C-lobe by a linker peptide (dark blue). The iron atoms and accompanying bicarbonate anions are shown in red and sky blue respectively. The two lobes of the molecule are related across a screw axis of 167° and 25Å translation.

---

domain II of the N lobe being composed of residues 97-245; domain NI is composed of residues which flank domain NII in the primary sequence. A total of nineteen cystine residues are found in serum TF, sixteen of which are located in the same relative position in hLF. Of the three additional disulfides, one is a short-range bridge found in the N lobe, while the other two are located at the ends of the interlobe connecting peptide. The presence of additional cystine residues in this linker peptide may confer conformational restrictions upon the serum TF molecule, which is reflected in the rotation required to superimpose the two lobes of the molecule (167° rotation, 23Å

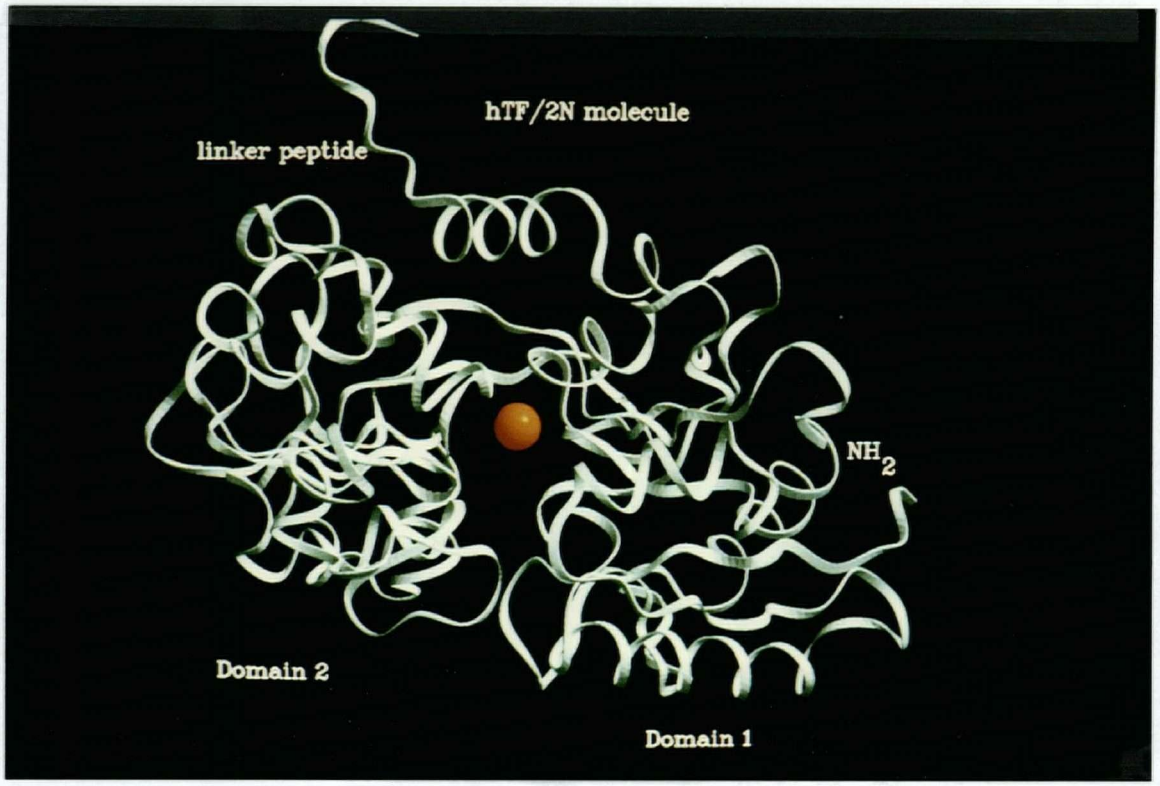


Figure 15: Structural Organization of the N-terminal Lobe of Rabbit Serum Transferrin

In this ribbon representation of the N lobe of transferrin, the two domains (NI and NII) of the molecule are shown to be composed of β-sheet structures joined by connecting loops and helices. In turn, the two domains are themselves connected by two bridging strands. In the full two-lobed molecule, the linker peptide joins the two lobes across the NI and CI domains. The binding of iron (red) involves a hinge-like movement of the two domains about the connecting strands thus trapping the metal deep within the protein interior.

---

translation for serum TF *versus* 180° rotation and 25Å translation for hLF). Whether these differences are of functional significance is as yet undetermined.

Rabbit serum TF reportedly possesses either one or two carbohydrate moieties (Leger *et al.*,1978; Strickland *et al.*,1979). The crystal structure of the molecule can be interpreted in terms of at least one glycan structure at or near positions 491-495 in domain CII; however, the structure lacks precise definition.

The availability of gene sequences for both hen ovoTF (Cochet *et al.*,1979) and human serum TF (Park *et al.*,1985; Schaeffer *et al.*,1987) has shown a high degree of conservation in the exon organization of TF genes from species distant in evolution. Though information on the organization of the rabbit serum TF gene is presently unavailable, an attempt has been made to correlate the exon organization of the human TF gene with the presence of discrete structural elements within the rabbit crystallographic model (Bailey *et al.*,1988) The N-terminal signal sequence is almost completely encoded by a single exon (Exon-1) while the remaining exons consist of seven homologous pairs common to both lobes, and two exons unique to the C lobe. Each amino acid ligand of the metal binding site is contributed by a separate exon, an unusual feature for eukaryotic metal binding proteins in which ligands are frequently clustered in discrete domains. Exon-2 of the N lobe (Exon-9 in the C lobe) displays high identity with single exons in both the chicken and human B*lym* oncogenes. Conserved residues in this region of the B*lym* transforming proteins are also those conserved between different transferrin species. This portion of the TF molecule forms a discrete supersecondary structural element which is remote from the iron binding site and therefore may be required for functions unrelated to metal binding (see Figure 16). Since the homologous regions in the transforming proteins constitute the majority of the primary sequence (56 of 65 residues), exon-2 may therefore retain a function in the TF molecule, such as interaction with a receptor, which is independent of iron binding.

6. Iron Binding Site and Structural Considerations

The elucidation of the hLF crystal structure unequivocally identified the protein ligands involved in iron binding and in doing so clarified the interpretation of data from previous physiochemical studies. As the protein ligands in both the hLF and rabbit

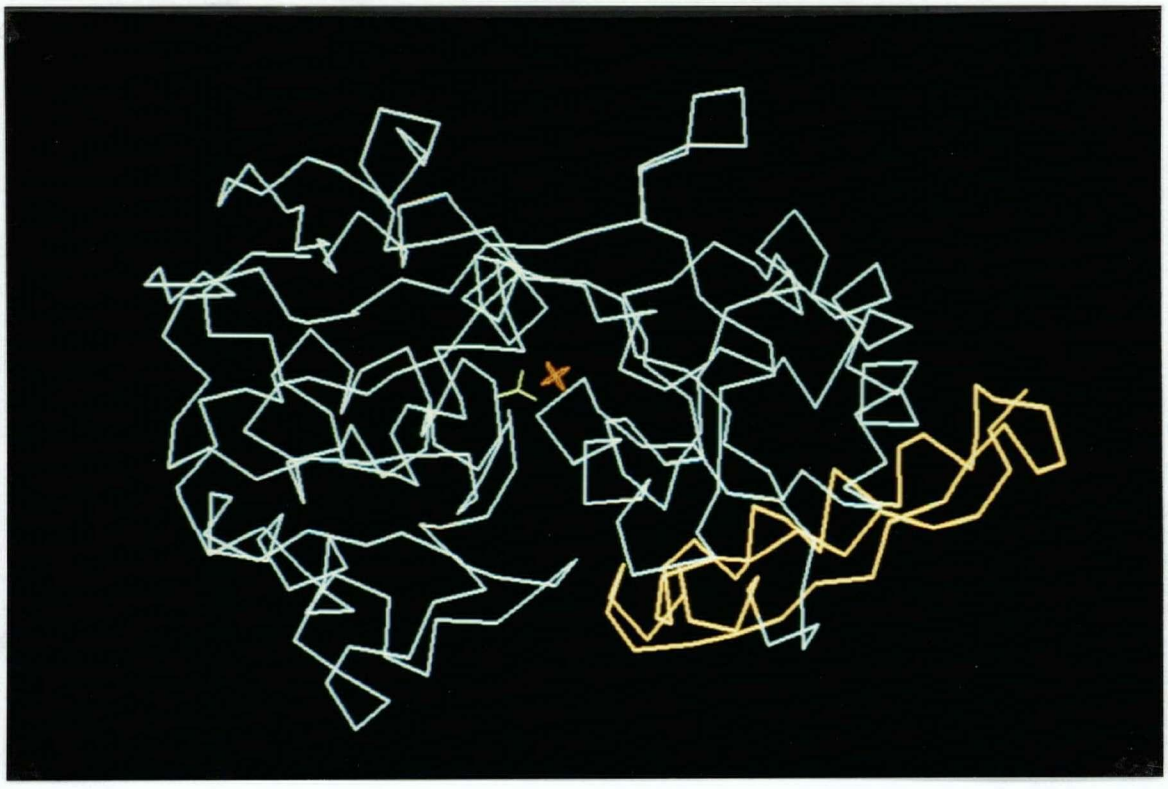


Figure 16: Region of the N Lobe Encoded by Exon 2.

Exon 2 of the human transferrin gene encodes a distinct super-secondary structure (yellow) in the mature molecule which is remote from the iron-binding site (red) (Bailey *et al.*, 1988). This region shows high identity with both the human and chicken B*lym* transforming proteins (Diamond *et al.*, 1984; Goubin *et al.*, 1983). In lactoferrin, Domain NI of the N-terminal lobe is the primary recognition site for the lymphocytic lactoferrin receptor (Rochard *et al.*, 1989), raising the possibility that this interaction may involve the exon 2 domain.

---

serum TF molecules are identical and occur in equivalent positions in the two molecules, the following discussion applies equally to both.

In the N lobe of the hLF molecule, Asp61, Tyr93, Tyr191 and His252 provide ligands to the iron through their respective side-chain functionalities (equivalent positions in the human and rabbit serum TF molecules are Asp63, Tyr95, Tyr188 and His249) while the C lobe ligands are contributed by the equivalent residues in that half of the molecule. Asp61 is contributed by a loop in Domain NI. The presence of a

carboxylate ligand was not anticipated from previous chemical studies. The replacement of this residue with serine in the equivalent position in the C-terminal lobe of melanoTF suggests this lobe of the molecule may not bind iron or would bind iron in a manner different to other transferrin lobes. Domain NII provides the Tyr191 ligand from the N-terminus of a helical segment and the invariant Arg121 residue from the N-terminus of another helical segment. The remaining two ligands (Tyr93 and His252) originate from the two domain-connecting strands.

The final two coordination positions of the iron atom are occupied by a bidentate carbonate ion in the refined version of the hLF structure. The counterion occupies the coordination positions closest to the opening of the cleft between the two domains, and as such is considered to bind the "open" side of the iron atom. This also places the carbonate atom in a position to interact directly with the invariant Arg121 as had been anticipated from modification studies (Rogers *et al.*,1978) and may also interact with the dipole moment at the N-terminus of an adjacent helix. Several other amino acid side chains identified within the iron-binding cleft have been implicated as being necessary for iron and counterion binding. These include the invariant Tyr83 residue, likely the "third Tyr ligand" identified in chemical studies (Aisen and Listowsky,1980; Brock,1985b; Chasteen,1983), and Gly63, which has been substituted by Arg in the C lobe of a naturally occurring variant of human serum TF whose visible absorption spectrum is significantly shifted (Evans *et al.*,1982). As well, Lys206 of serum TF approaches within 6Å of the iron center and has been implicated by labelling experiments to undergo a change in reactivity upon iron binding (Shewale and Brew,1982). The equivalent residues in the N lobe of hLF and C lobe of OTF have been replaced by arginine and glutamine respectively, though the functional implications of these substitutions has yet to be demonstrated.

The isolation of iron-binding fragments from the N lobe which correspond to Domain NII has been reported for several transferrin molecules (Evans and

Madden,1984; Legrand *et al.*,1984). As these fragments lack the Asp61 ligand, a plausible sequence for ligand binding might involve initial binding by the carbonate dianion and the other three protein ligands in a structurally open conformation after which the two domains would be brought within close proximity by pivoting about the flexible hinge formed by the interdomain strands (Baker *et al.*,1987). Coordination of the final position by Asp61 would complete the binding and lock the molecule into a tight configuration, in agreement with studies which show a significant compaction of the molecule upon iron binding.

In order to understand better the structural requirements for iron-binding by TF, a study was undertaken to express recombinant molecules for the purpose of employing the techniques of site-directed mutagenesis. Previous investigators have expressed the chicken OTF gene in transgenic mice (McKnight *et al.*,1983) and production of a portion of rat TF fused to β-galactosidase has also been reported (Aldred *et al.*,1984). The study reported in this dissertation was designed to achieve a high level of protein expression in a biological system that would allow ready recovery of the recombinant product. Our initial efforts have been to produce a recombinant N lobe half molecule of human serum TF, though we are now extending our analysis in attempting to express both the C lobe half molecule and a full length TF species. Successful completion of these goals will allow us to dissect separate iron binding sites, assess the requirements for TF receptor binding and investigate individual and collective properties of the two lobes.

II. Experimental Procedures and Materials

A. General Materials and Supplies

The large fragment (Klenow) of *E. coli* DNA polymerase I, T4 polynucleotide kinase, T4 DNA ligase, and all restriction enzymes were purchased either from Pharmacia-PL Biochemicals or Bethesda Research Laboratories. Deoxy- and dideoxyribonucleotidetriphosphates were purchased from Pharmacia-PL Biochemicals. α-[^{32}P]dATP and γ-[^{32}P]ATP were from Amersham and New England Nuclear. Hen egg white lysozyme, deoxyribonuclease I, ribonuclease A and goat anti-human transferrin antiserum were from Sigma. Polyethylene glycol 4000 (PEG) was from BDH Chemicals. Ferriprotoporphyrin IX dimethylester was obtained from Porphyrin Products, Logan, UT. $Ru(NH_3)_6Cl_3$ (Alfa) was purified as described by Pladziewicz *et al.* (1973). All other chemicals were reagent grade or better.

The three-dimensional coordinates for the crystal structure (1.9Å resolution) of the trypsin-solubilized form of bovine microsomal cytochrome b_5 were kindly provided by Dr. F.S. Mathews (Washington University, St. Louis, MO). Coordinates for rabbit serum TF (3.3Å resolution) were provided by Dr. Peter Lindley (Birkbeck College, University of London, London, U.K.). All imaging was performed with the MMS (molecular modelling system) package (S. Dempsey, University of Cailfornia, S.D.) in the Department of Biochemistry at the University of British Columbia.

B. Cytochrome b_5 Studies

1. *Gene construction*

The four oligonucleotides used to assemble the gene cassette encoding the lipase-solubilized form of bovine liver microsomal cytochrome b_5 (designated B1-B4) were synthesized on an Applied Biosystems 380A DNA synthesizer by Mr. Tom Atkinson (University of British Columbia). Post-synthesis purification was achieved by

electrophoresis through an 8% denaturing polyacrylamide gel (7M urea), followed by reverse phase liquid chromatography of the appropriately sized product through Sep-Pak C18 cartridges (Waters Associates)(Atkinson & Smith, 1984). Concentrations were determined by absorbance at 260 nm. (1 A_{260}= 33μg/ml of single stranded DNA).

One microgram of each oligonucleotide was phosphorylated at the 5' end using ATP and T4 polynucleotide kinase (Chaconas and van de Sande, 1980). Second strand fill-in was performed using the Klenow fragment in the presence of all four deoxyribonucleotides and the resulting blunt-ended products were then cloned into *Sma*I-cut M13mp18 (Messing,1983). The fidelity of the DNA synthesis step was confirmed by sequence analysis using the chain termination method on single stranded template DNA (Sanger *et al.*,1977). A *Bam*HI/*Hha*I fragment corresponding to the sequence covered by oligonucleotides B1 and B2 , and a *Hha*I/*Eco*RI fragment covering the B3 and B4 sequence were then excised from the replicative form of the appropriate clones and ligated into *Bam*HI/*Eco*RI cut pUC 19 (Messing,1983). $CaCl_2$ treated *E. coli* JM83 bacteria were transformed with the ligation products and plasmid DNA was then prepared from select clones by the alkaline lysis procedure for restriction digest analysis (Maniatis *et al.*,1982). The complete nucleotide sequence of the final construction was obtained by recloning a *Bam*HI/*Eco*RI fragment back into both M13mp18 and M13mp19 followed by dideoxy sequence analysis of both strands of the gene cassette.

2. *Host Selection*

The pUC19 plasmid encoding the lipase-solubilized form of bovine hepatic cytochrome b_5 (designated pLpb_5) was transformed into $CaCl_2$ treated *E. coli* host strains and grown overnight at 37° with shaking in 2 mL of YT broth (Maniatis *et al.*,1982) supplemented with ampicillin to 50 μg/mL. One hundred microlitres of each culture was subsequently used to inoculate parallel 20 mL cultures of the same medium supplemented additionally with isopropylthiogalactoside (IPTG) to 1mM and grown for

18 hrs. Cell densities were estimated by measuring the optical density of diluted samples at 600 nm. Equivalent volumes of cell culture were pelleted by centrifugation, resuspended in NaDodSO$_4$-PAGE loading buffer, boiled for 5 minutes and then resolved on a 15% NaDodSO$_4$ polyacrylamide gel (Laemli, 1970). The gel was then fixed, stained with Coomasie Brilliant Blue R250 (Merck) and scanned using an LKB Ultroscan XL laser densitometer.

3. *Induction Studies*

Parallel 50 mL YT broth samples were inoculated with one mL of overnight culture of RR1 and JM 83 strains of *E. coli* and grown with shaking at 37°. RR1 cultures were induced with the addition of IPTG to 1 mM and 1 mL samples were then removed from the cultures at various time points for analysis by NaDodSO$_4$-PAGE.

4. *Protein Purification*

Transformed JM 83 bacteria containing pLpb_5 were grown for 16 hours at 37° with shaking in YT medium (4L) containing 50 μg/mL ampicillin. This culture was used to inoculate 50 L of the same medium in a Bilthoven fermentation apparatus. This culture was grown with vigorous aeration until the O.D.$_{600}$ reached 2.5, at which time the air supply was reduced to a minimum. The culture was maintained under these conditions for a further 20 hours. This latter step improved the yield of holo-cytochrome b_5, presumably by permitting heme synthesis to accommodate the large load of apocytochrome b_5 (S.G. Sligar, University of Illinois; personal communication). Bacteria were harvested with a CEPA continuous flow centrifuge, and the resulting cell paste was resuspended in 1L of lysis buffer containing 3 mg/ml lysozyme as described (von Bodman *et al.*, 1986). Following a one hour incubation at 37°, cell lysis was completed by a single pass through a French press (1200 p.s.i.), and the cell debris was cleared by centrifugation at 8600 x g for 30 minutes at 4^0. KCl and polyethyleneglycol

(PEG 4000) (final concentrations 0.4 M and 6% respectively) were added to the supernatant solution. This solution was stirred for 15 minutes and then centrifuged at 8600 x g for 30 minutes at 4°. The resulting supernatant solution was brought to 50% $(NH_4)_2SO_4$ saturation, stirred at 4° for 1 hr, and then centrifuged at 8600 x g for 30 minutes at 4°. The supernatant solution was dialyzed exhaustively against de-ionized water until the conductivity equalled that of sodium phosphate buffer (20 mM, pH 7.2). The dialysate was clarified by centrifugation at 8600 x g , loaded onto a column (2.5 x 12.5 cm) of DE-52 cellulose (Whatman), washed with one column volume of 50 mM sodium phosphate (pH 7.2), and eluted with 150 mM sodium phosphate buffer (pH 7.2). Red coloured fractions were concentrated by ultrafiltration (Amicon YM-5 and Centriprep-10), exchanged into 20 mM sodium phosphate (pH 7.2), and loaded onto a column (5 x 150 cm) of Sephadex G75 superfine resin (Pharmacia). The column was developed with 20 mM sodium phosphate buffer (pH 7.2), and the $A_{412.5}/A_{280}$ ratio of individual fractions was determined. Fractions with an $A_{412.5}/A_{280}$ ratio of greater than 5.0 were concentrated by ultrafiltration and exchanged into 20 mM triethanolamine (pH 7.3). Further purification was achieved by chromatography on an HR 10/10 Mono-Q FPLC column (Pharmacia) in 20 mM triethanolamine (pH 7.3) with a linear gradient of 0.2 to 0.29 M NaCl. The S64A mutant form of the bovine cytochrome was not stable when chromatographed on Mono-Q FPLC. Therefore, concentrated fractions from the first Sephadex G75 column were re-chromatographed on a second smaller column (2.5 x 100 cm) of Sephadex G75. For all forms of the cytochrome, fractions with an $A_{412.5}/A_{280}$ ratio of >5.9 were concentrated by ultracentrifugation and stored in liquid nitrogen.

5. *Dimethylesterheme IX Substitution*

Replacement of the native protoheme IX group with dimethylesterheme IX (DME) was performed as described previously (Reid *et al.*, 1984). Briefly, purified

cytochrome b_5 was exchanged into de-ionized water, concentrated to 10 mg/mL by ultracentrifugation, chilled on ice and acidified to pH 1.5 with HCl. All subsequent manipulations were performed at 4°. This solution was extracted with an equal volume of methylethylketone four times and then dialyzed consecutively against:

1. 50 mg/L $NaHCO_3$, 320 mg/L EDTA
2. 50 mg/L $NaHCO_3$
3. 20 mM sodium phosphate pH 7.2
4. 50 mg/L $NaHCO_3$.

DME was dissolved in dimethylformamide and added to the dialyzed apocytochrome dropwise. The reconstituted DME-cytochrome was separated from the free DME by chromatography on a column (100 X 2.5 cm) of Sephadex G75 (Pharmacia).

6. *Tryptic Peptide Mapping*

Apocytochrome b_5 was prepared by the method of Teale (1959; Reid *et al.*, 1984). One milligram of the apoprotein was digested in 0.5 mL of 0.2 M NH_4HCO_3 with 12.5 μl of a stock solution of L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK)-trypsin (Worthington, Cooper Diagnostics)(1.0 mg/mL TPCK-trypsin in 1 mM HCl) at 37° for 6 hrs. An additional 12.5 μl of the TPCK-trypsin stock was then added, and the hydrolysis continued for a further 18 hours. After the addition of 100 μl of 1M HCl, the sample was lyophilized, and the hydrolysis procedure was repeated. The final lyophilized pellet was dissolved in 1.0 mL of 0.05% trifluoroacetic acid (TFA). The remainder of the tryptic mapping procedure was performed by Marcia Mauk (Dept. of Biochemistry, University of British Columbia). Reverse phase HPLC tryptic peptide maps were developed from 100 μl samples of this digest with a Varian Model 5060 HPLC system equipped with a Vista 401 data station and fitted with an Alltech C18 reverse phase column (4.6 x 250 mm). The gradient (all solutions in 0.05% TFA) was linear, 0-60% acetonitrile, over 135 minutes at a flow rate of 1.0 mL/min.

7. *Amino Acid Sequence Determination*

The amino-terminal sequence of recombinant lipase solubilized cytochrome b_5 and the total sequences of selected tryptic peptides obtained from HPLC maps of the authentic trypsin-solubilized bovine liver cytochrome b_5 were determined on 1 nmol samples of each. Vapor phase amino terminal sequence analysis of tryptic peptides was performed with an Applied Biosystems 477A Protein Sequencer by the Protein Microchemistry Centre at the University of Victoria. Amino-terminal sequence determination of the recombinant cytochrome was performed by Prof. Ian Clarke-Lewis at the Biomedical Research Centre at the University of British Columbia.

8. *Limited Tryptic Hydrolysis*

Tryptic hydrolysis to convert the recombinant lipase-solubilized form of cytochrome b_5 to the trypsin-solubilized form was performed by the method of Strittmatter and Ozols (1966). Briefly, purified protein was exchanged into 20 mM sodium phosphate (pH 7.5), TPCK-trypsin was added to a final concentration of 50 μg/mL and the digestion allowed to proceed for two hours at 25°. Tryptic hydrolysates were chromatographed on a Mono-Q column (HR5/5)(Pharmacia) in 20 mM triethanolamine (pH 7.3) using a linear gradient of 0.19 to 0.32 M NaCl.

9. *Site-Directed Mutagenesis*

Modifications to the gene encoding the lipase-solubilized form of bovine cytochrome b_5 were accomplished with three separate techniques. The correction of codons for positions 11,13 and 57 of the recombinant Lpb_5, the substitution at position 64 and the production of a gene encoding the trypsin-solubilized form of bovine cytochrome b_5 was accomplished by the method of Zoller and Smith (1983) with single stranded M13mp18 template DNA purified from RZ1032 strain *E. coli*, using the dut-,

ung- selection procedure (Kunkel, 1985). The alteration of the lipase-solubilized b_5 gene cassette to encode the erythrocytic form of human cytochrome b_5 was accomplished using the Polymerase Chain Reaction (PCR) based procedure of Nelson and Long (1989). The production of the E11Q,Q13E and N57D forms of bovine lipase-solubilized b_5 was achieved by splicing *Bam*HI/*Acc*I and *Acc*I/*Eco*RI restriction fragments from the triple mutant and corrected wild-type gene cassettes.

10. *Spectroelectrochemistry*

Potentiometric titrations were performed with an optically transparent thin layer electrode (OTTLE) described previously (Reid *et al.*,1982,; Reid 1984) with $Ru(NH_3)_6Cl_3$ used as a mediator. Solution potentials were measured against the saturated calomel electrode (Radiometer #K4112)(SCE) and converted to the hydrogen scale as described by Dutton (1978). Spectra were processed and Nernst plots obtained using OLIS software (On-Line-Instrument-Systems, Jefferson, GA). All reported midpoint potentials represent at least two independent titrations.

C. Human Serum Transferrin Studies

The majority of the studies on the expression of human serum transferrin were done in collaboration and consultation with Drs. Robert C. Woodworth and Anne B. Mason (University of Vermont).

1. *Isolation of Human Serum Transferrin cDNA*

The initial isolation of a full-length cDNA for human serum transferrin was performed by Mrs. Susan Lejay. A human liver cDNA library constructed in the *E. coli* expression vector pKT-218 (Prochownik *et al.*,1983) was kindly provided by Dr. Stuart Orkin (Harvard University). This cDNA library was screened by using a synthetic oligonucleotide coding for the amino-terminal eight amino acids of serum hTF as a

hybridization probe. This oligonucleotide corresponded to nucleotides 88-111 of the human serum TF cDNA sequence reported by Yang *et al.* (1984). The oligonucleotide was end-labelled with T4 polynucleotide kinase and γ-[^{32}P]ATP (Chaconas & van de Sande,1980). Replica nitrocellulose filters lifted from approximately 10^5 colonies (Schleicher and Schuell) were screened under high stringency conditions (final washes in 1X SSC, 1% $NaDodSO_4$ at 60°). Restriction endonuclease mapping of positive clones and DNA sequence analysis of positive clones was performed using standard procedures with pUC19 and M13mp19 vectors, respectively (Maniatas *et al.*,1982; Messing,1983; Sanger *et al.*,1977).

2. *Expression Vector Assembly*

A *Pst*I/*Hha*I restriction fragment spanning the coding region for the signal sequence and N-terminal lobe of human TF was sub-cloned into M13mp18. Oligonucleotides were synthesized and purified as described for the cytochrome b_5 methods above. Subsequent alteration of this sequence was accomplished by using a mutagenesis kit purchased from Amersham and the procedures described by Taylor *et al.* (1985). Unique *Hin*dIII and *Kpn*I endonuclease recognition sites and two translational stop codons were engineered to allow the excision of coding sequence for the mature N-terminal lobe of TF. An altered version of the expression vector pKK223-3 (Brosius and Holy,1984), termed ΔpKK223-3 and oligonucleotides encoding the *E. coli* alkaline phosphatase signal sequence were kindly provided by Dr. Deborah Cleveland (University of British Columbia). The expression vector pNUT (Palmiter *et al.*,1987) was kindly provided by Dr. Richard Palmiter (Howard Hughes Medical Institute, University of Washington). A *Hpa*II/*Hin*dIII restriction fragment corresponding to the signal peptide and N-terminal lobe of human TF was made blunt using the Klenow fragment and deoxyribonucleotides and then cloned into the large *Sma*I fragment of pNUT. All subsequent manipulations of DNA were performed using standard methodologies

(Maniatis *et al.*,1982).

3. *DNA Transfection and Eukaryotic Cell Culture*

Baby hamster kidney (BHK) cells were a gift of Dr. Richard Palmiter (Howard Hughes Medical Institute, University of Washington). Super-coiled plasmid DNA was purified from *E. coli* JM105 by two successive centrifugation steps through cesium chloride gradients. BHK cells were grown in Dulbecco's modified essential medium (DMEM) (Gibco) containing 10% fetal calf serum (Gibco) and 50 U/ml penicillin G and 37 U/mL streptomycin (both from Sigma) in a 37° humidified incubator under 5% CO_2. For the transfection procedures, approximately 10^7 cells were seeded per 90-mm culture dish and 12 hours later were transfected with 10 μg of plasmid DNA using the calcium phosphate co-precipitation procedure as described by Searle *et al.* (1985). After 24 h, the medium was changed to medium consisting of a 1:1 (v:v) mixture of DMEM and DMEM harvested from semi-confluent BHK cultures (conditioned DMEM), with a final concentration of 10% fetal calf serum and 500 μM methotrexate (MTX). Stocks of selected cells were stored in liquid nitrogen using DMEM containing 20% fetal calf serum and 10% dimethylsulfoxide (DMSO) (Sigma, tissue culture grade) as a preservative. Large-scale roller cultures were initiated by seeding approximately 5 x 10^7 cells into each 850 cm^2 roller bottle (Corning) containing 100 mL of medium consisting of a 1:1 (v:v) mixture of DMEM and Hamm's F-12 nutrient mixture, with a final concentration of fetal calf serum of 5% and supplemented with 500 μM MTX. Cultures were induced at 80% confluency by the addition of $ZnSO_4$ to a final concentration of 0.02 mM. The medium was harvested at 48 h intervals and replaced with 100 mL of the same medium.

4. *Immumoprecipitation and Western Blotting*

Immunoprecipitation of cell culture medium and cell lysates was performed by the method of van Oost *et al.* (1986). Briefly, medium was removed from Zn-induced cultures, and the cell layer rinsed twice with phosphate buffered saline (PBS). The cell layer was solubilized with a detergent solution containing 0.15 M NaCl, 20 mM Tris-HCl (pH 8.0), 0.5% sodium deoxycholate (NadChol), 1% Nonidet-P40 and 0.1 % $NaDodSO_4$. The cell debris was removed by centrifugation at 8000 x g at ambient temperature. Two mL of a detergent solution containing 0.5% NadChol, 0.1% $NaDodSO_4$, and 1% Nonidet-P40 was added to each 10 mL of cell culture medium harvested. Both cell culture medium and cell lysates were treated with pre-immune mouse serum (1000-fold diluted) and 1 mL of formalin-fixed *Staphylococcus aureus* cells (10% v:v)(Bethesda Research Laboratories) and then cleared by centrifugation at 8000 x g. Goat anti-hTF serum (Sigma) was then added (250-fold dilution), left overnight at 4° and immunoconjugates were then recovered with the addition of *S. aureus* cells. Precipitated material was resolved by electrophoresis on 12% polyacrylamide gels in the presence of $NaDodSO_4$ (Laemmli, 1970), followed by electroblotting onto nitrocellulose membranes (Schleicher & Schuell). Blots were incubated in PBS containing 0.1 mg/mL gelatin, then treated with goat anti-hTf antiserum (250-fold dilution), and finally developed with an alkaline phophatase conjugated, rabbit anti-goat IgG antibody (Promega) according to the supplier's instructions.

5. *Protein Purification*

Harvested culture medium was made 0.01% in phenylmethylsulfonylfluoride (PMSF)(Sigma), and sufficient ferricnitrilotriacetic acid ($Fe^{III}(NTA)_2$) was added to saturate all transferrin in the medium. After stirring for 30 min, the solution was dialyzed for 24 h versus cold running tap water and then for 12 h versus several changes of cold de-ionized water. Concentrated Tris-HCl buffer, pH 8.4, was added to a final

concentration of 5 mM. The preparation was centrifuged at 6000 x g to remove insoluble debris, and then loaded onto a column (2.5 x 80 cm) of DEAE-Sephacel (Pharmacia) equilibrated with 10 mM Tris-HCl buffer, pH 8.4. The column was eluted with a linear gradient of NaCl (0-0.3 M) in the same buffer. Fractions showing pink colour were analyzed by $NaDodSO_4$-PAGE and fractions containing the recombinant protein were pooled and concentrated by ultrafiltration over an Amicon PM-10 membrane. The protein was then chromatographed on a column (2.5 x 100 cm) of Sephadex G-75 Superfine (Pharmacia) equilibrated with 100 mM ammonium bicarbonate and fractions were analyzed by $NaDodSO_4$-PAGE. Fractions still containing full-length bovine transferrin and the recombinant protein were re-chromatographed on the same G-75 column. The recombinant protein was purified to homogeneity by FPLC on a column (1 x 10 cm) of Polyanion SI (Pharmacia) using a linear gradient of NaCl (0-0.3M) in 50 mM Tris-HCl, pH 8.0, over a period of 1 h at a flow rate of 1 mL/min. Fractions of 1 mL were collected.

$NaDodSO_4$-PAGE was performed on 5-12% gradient gels, and urea-PAGE was performed according to a modification (Brown-Mason & Woodworth,1984) of the Makey & Seal procedure (1976). Iron was removed from the holoprotein by incubation in a buffer containing 1mM NTA, 1mM EDTA, and 0.5M sodium acetate, pH 4.9. For iron-binding titration studies, the apoprotein was made 10 mM in $NaHCO_3$ and titrated with $Fe^{III}(NTA)_2$ while monitoring the absorbance at 465 nm.

6. *Amino-Terminal Sequence Analysis*

The amino-terminal sequences of both the major and minor forms of the recombinant hTF/2N were determined on an Applied Biosystems 470A protein sequencer in the Given Analytical Facility at the University of Vermont.

7. *hTF/2C-pNUT Assembly*

The initial assembly of the hTF/2C coding sequence was accomplished by Dr. Janet Lineback (Florida International University). A construction encoding a hTF/2C recombinant species was produced using a PCR procedure to effect the splicing of the natural signal peptide sequence to the C lobe coding sequence. The oligonucleotide primer defining the 5' end of the clone encoded the natural signal sequence of the hTF cDNA and residues 334-341 of the C-terminal lobe and produced a *Sma*I site at the 5' terminus of the amplified product, while the 3' oligonucleotide spanned a region of the 3' non-coding sequence and introduced an additional *Sma*I site. The initial pKT218-hTF cDNA vector was first cut with *Eco*RI to remove almost the entire coding sequence for the N-terminal domain, after which the remaining coding sequence served as a template for 25 rounds of PCR amplification with *Thermophillus aquaticus* (*Taq*) DNA polymerase (AmpliTaq; Perkin-Elmer) using reaction conditions recommended by the supplier. The amplified product was cut with *Sma*I, cloned into *Sma*I-cut pNUT vector and introduced into BHK cells as before.

In order to determine the fidelity of the PCR amplification procedure, the *Sma*I-cut PCR amplification product was also cloned into the *Sma*I site of the vector BlueScript KS+ (Stratagene) and subject to ExoIII deletion analysis (Heinikoff, 1984). Clones from individual time points were analyzed by restriction digest and then infected with helper phage K107 to produce single stranded template for dideoxy sequence analysis.

III Results

A. Cytochrome b_5 Studies

1. *Gene Design and Synthesis*

The assembly strategy for the gene cassette encoding the lipase-solubilized bovine microsomal cytochrome b_5 protein is shown in Figure 17. The oligonucleotide pairs (B1/B2 and B3/B4) were designed to overlap sufficiently so as to be stable at the extension temperature (37°). An autoradiogram of the extension of end-labelled primer pairs is shown in Figure 18. Both primer pairs could be extended to yield appropriately sized products of 170 bp., with yields approaching 50% based on comparison of band intensities of the extension products and remaining unextended oligonucleotides. The sequence of the oligonucleotides was screened prior to synthesis to ensure that the regions of overlap between pairs would constitute a unique priming structure. Nevertheless, the extension reaction with oligonucleotides B3 and B4 produced a substantial amount of a shorter product of approximately 140 bp. As this reaction product was seen only when oligonucleotide B4 was labelled, it must represent an alternate priming event in which the 3' end of B4 hybridizes to an internal sequence in B3. Such illegitimate priming may restrict the number of oligonucleotides that could be used simultaneously in fill-in synthesis procedures. In the present study, no full-length product was observed in reactions where all four oligonucleotides were annealed and extended (results not shown). A second limitation of this strategy is the fidelity of the polymerase activity. The first clone sequenced during assembly contained a deletion of a single thymidylate residue, and resulted in the introduction of an ochre codon at position 29 of the translation product. The use of alternate polymerases with greater processivity, such as that of T7 bacteriophage (Tabor and Richardson,1987), may increase the accuracy of second strand synthesis.

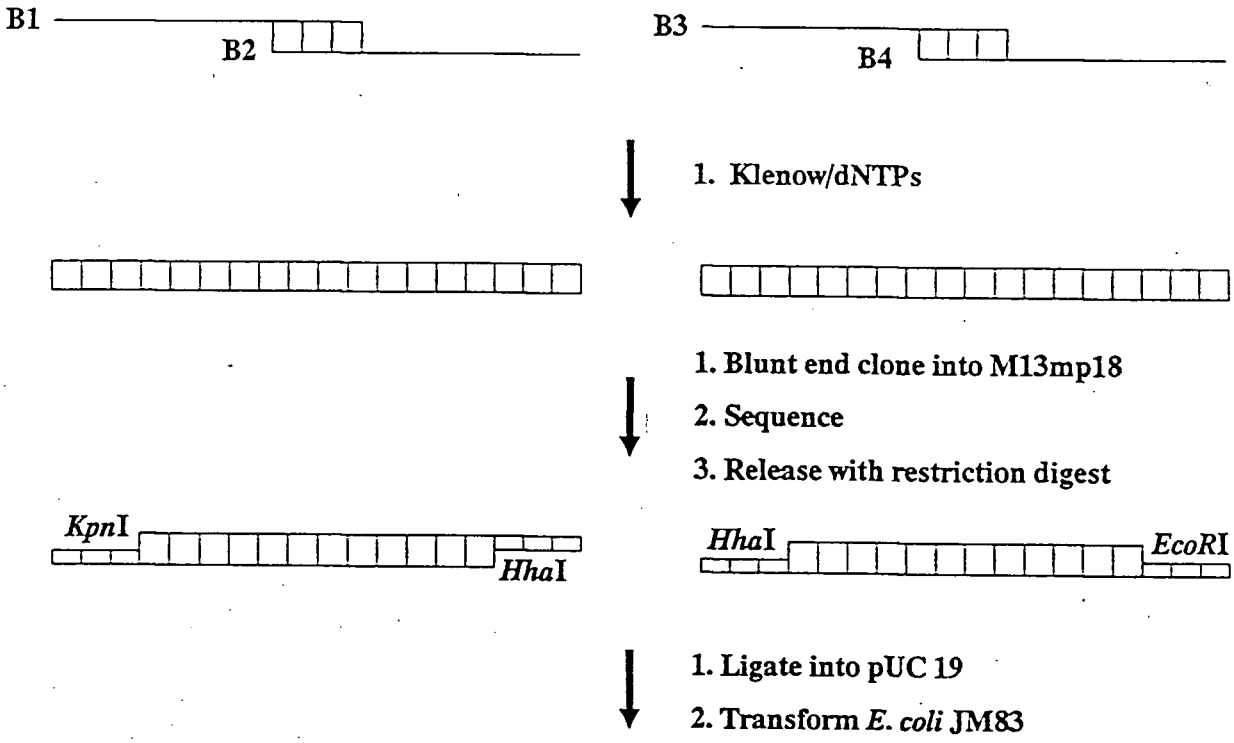


Figure 17: Assembly Strategy for the Cytochrome b_5 Gene Cassette

Two pairs of overlapping oligonucleotides (B1/B2 and B3/B4) were annealed, then filled-in using Pol I (Klenow). The resulting blunt fragments were ligated into M13mp18, sequenced, then released with the indicated restriction endonuclease digestions. These fragments were then cloned directionally into pUC19 and transformed into *E. coli* JM83.

Internal *Kpn*I, *Hha*I and *Eco*RI restriction sites were included in the design of the gene cassette to allow the directed assembly of fill-in reaction products into pUC19. Direct cleavage of the extension products proved to be inefficient, even though

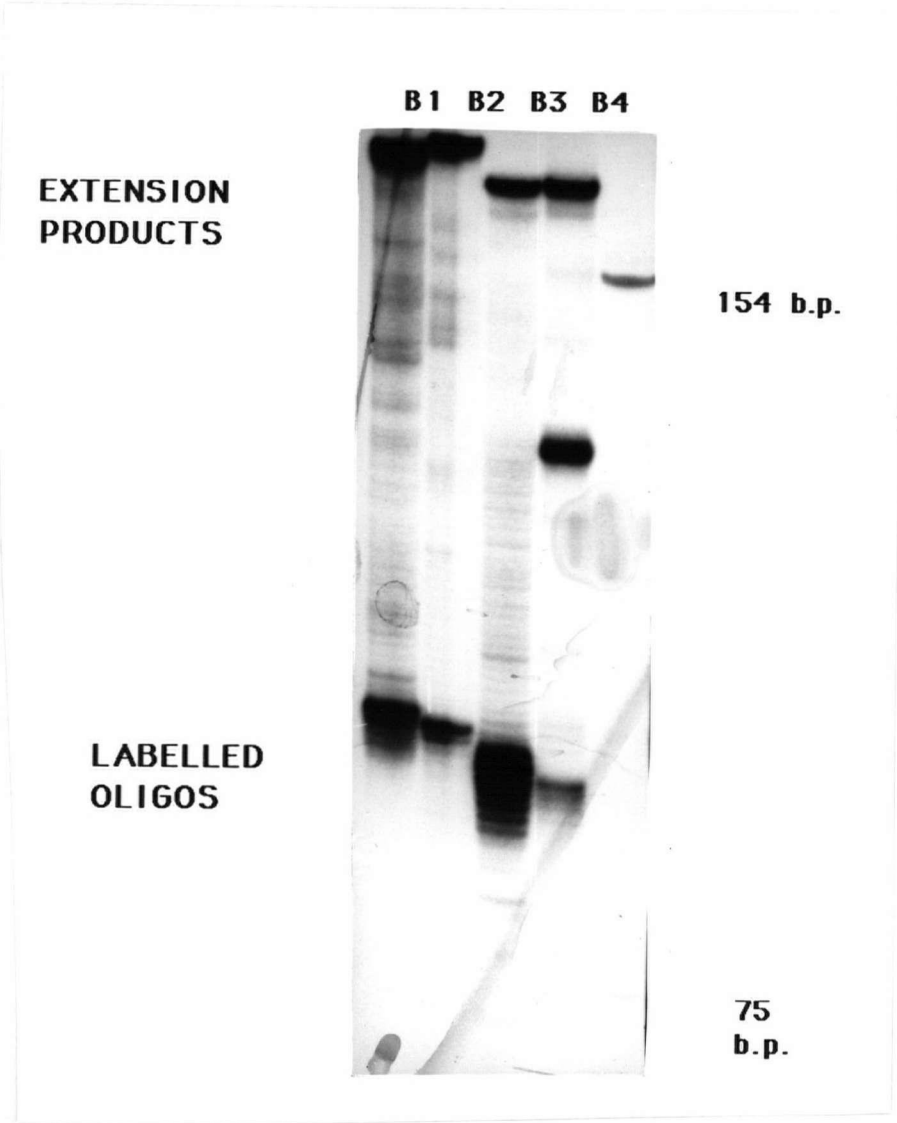


Figure 18: Second Strand Fill-in Reaction of B1-B4 Oligonucleotides

Four separate oligonucleotide extension reactions were performed using paired oligonucleotides, one of which was end-labelled with ^{32}P. Reaction products were electrophoresed through a denaturing polyacrylamide gel (10% polyacrylamide, 8M urea). The indicated size markers are ^{32}P-labelled *Hin*fI digestion products of pBR322.

---

additional bases had been allowed at the presumptive termini of the extension products to encourage endonuclease recognition. Thus the extension products were cloned without modification into the *Sma*I site of M13mp18 for dideoxy sequence analysis and were

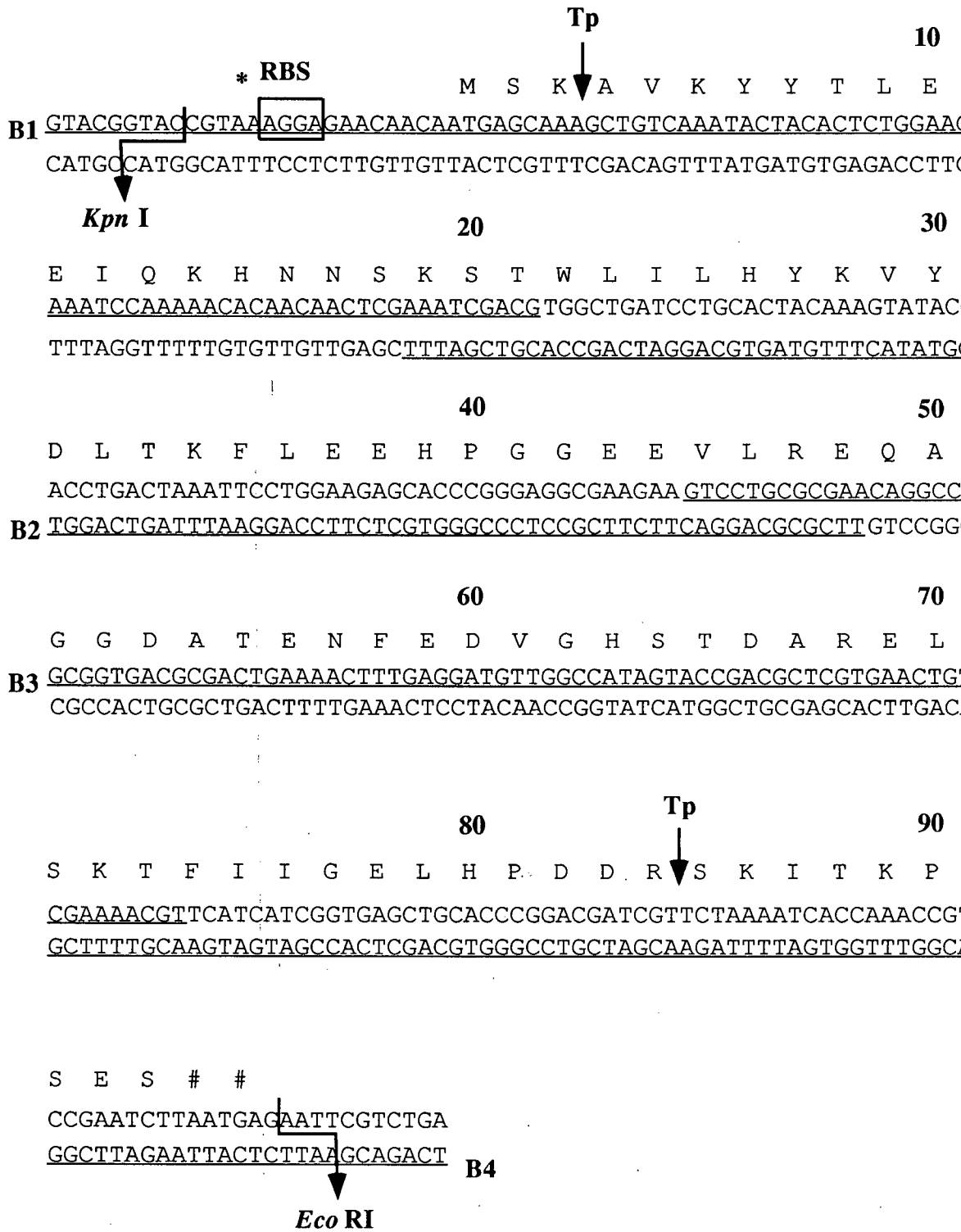

Tp
10
* RBS
M S K A V K Y Y T L E
B1 GTACGGTACCGTAAAGGAGAACAACAATGAGCAAAGCTGTCAAATACTACACTCTGGAAG
CATGCCATGGCATTTCCTCTTGTTGTTACTCGTTTCGACAGTTTATGATGTGAGACCTTC
Kpn I
20
30
E I Q K H N N S K S T W L I L H Y K V Y
AAATCCAAAAACACAACAACTCGAAATCGACGTGGCTGATCCTGCACTACAAAGTATACG
TTTAGGTTTTTGTGTTGTTGAGCTTTAGCTGCACCGACTAGGACGTGATGTTTCATATGC
40
50
D L T K F L E E H P G G E E V L R E Q A
ACCTGACTAAATTCCTGGAAGAGCACCCGGGAGGCGAAGAAGTCCTGCGCGAACAGGCCG
B2 TGGACTGATTTAAGGACCTTCTCGTGGGCCCTCCGCTTCTTCAGGACGCGCTTGTCCGGC
60
70
G G D A T E N F E D V G H S T D A R E L
B3 GCGGTGACGCGACTGAAAACTTTGAGGATGTTGGCCATAGTACCGACGCTCGTGAACTGT
CGCCACTGCGCTGACTTTTGAAACTCCTACAACCGGTATCATGGCTGCGAGCACTTGACA
Tp
80
90
S K T F I I G E L H P D D R S K I T K P
CGAAAACGTTCATCATCGGTGAGCTGCACCCGGACGATCGTTCTAAAATCACCAAACCGT
GCTTTTGCAAGTAGTAGCCACTCGACGTGGGCCTGCTAGCAAGATTTTAGTGGTTTGGCA
S E S # #
CCGAATCTTAATGAGAATTCGTCTGA
GGCTTAGAATTACTCTTAAGCAGACT
B4
Eco RI

Figure 19: Nucleotide Sequence of the Lipase-Solubilized, Bovine Microsomal Cytochrome b_5 Gene Cassette.

The oligonucleotides used in the synthesis of the gene cassette (B1-B4) are underlined. Restriction endonuclease recognition sequences used in the final assembly step are indicated by bent arrows while the sites of trypsin cleavage which yield the heme-binding core domain are shown by straight arrows. Termination codons for the lac Z (*) and cytochrome b_5 (##) translation products are indicated while the ribosome binding site for cytochrome b_5 initiation is boxed.

---

then efficiently released from the replicative form and assembled into pUC 19 (see Figure 19).

The gene cassette utilizes a transcriptional fusion strategy with the lacZ transcript from pUC and includes an in-frame stop codon for the lacZ translation product that is followed immediately by a consensus ribosome binding (Shine-Delgarno) site (Figure 20). The spacer region between the ribosome binding site and initiator methionine was modelled on the *Pseudomonas* cytochrome P-450_{cam} sequence, which has been used successfully in the expression of a synthetic gene for rat liver cytochrome b_5 in *E. coli* (von Bodman *et al.*,1986). This region was modified to conform to consensus base rules for this region (Sherer *et al.*,1980; Stormo *et al.*,1982) by shortening the distance between the ribosome binding site and initiator methionine to 9 bp. Codons were chosen to reflect the *E. coli* bias (Maruyama *et al.*,1986).

4. *Expression Studies*

Transformation of *E. coli* JM83 with the pUC construction containing the gene for the lipase-solubilized cytochrome b_5 (pLpb_5) resulted in high levels of the recombinant protein as demonstrated by the bright pink colour of bacterial pellets. As a first approach towards optimizing expression, several different strains of *E. coli* were transformed, and the level of recombinant protein in each was estimated by NaDodSO$_4$-PAGE analysis. A representative trial is shown in Figure 21. As judged by laser

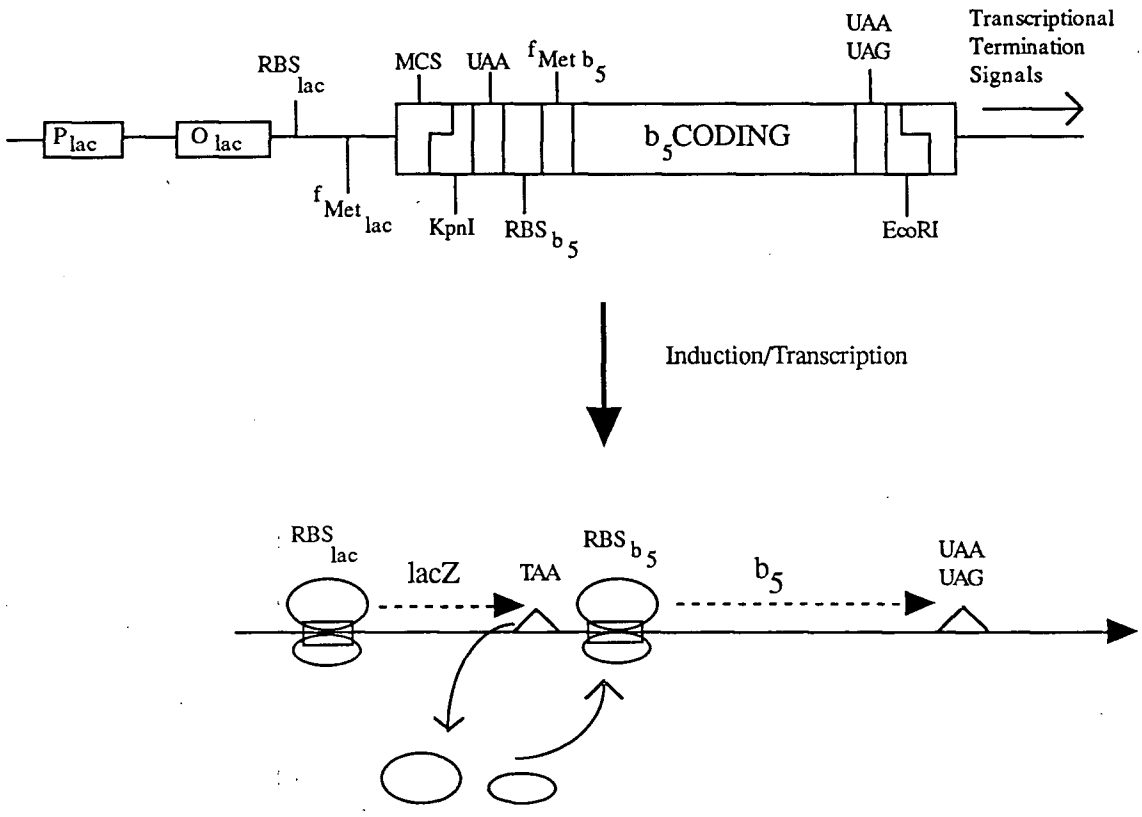


Figure 20: Transcriptional Fusion Strategy for the Expression of Recombinant Cytochrome b5.

The cytochrome b_5 gene cassette (large block) includes an in-frame translational termination codon for the lac Z product (UAA). Induction of the lac promoter of pUC19 (P_{lac}) results in a polycistronic message, thus the information for cytochrome b_5 expression is "transcriptionally fused" to the lac operon. Upon translation, the lac Z product is terminated, and the translation machinery re-initiates at the adjacent ribosome binding site (RBSb_5) after which the cytochrome b_5 coding information is translated.

scanning densitometry, the recombinant cytochrome b_5 accounts for up to 15% of the total stain density (Figure 20, sample RR1).

To optimize the expression of recombinant protein further, an induction time course experiment was performed. Several cultures of transformed *E. coli* RR1 were induced at varying times after which samples were removed for optical density

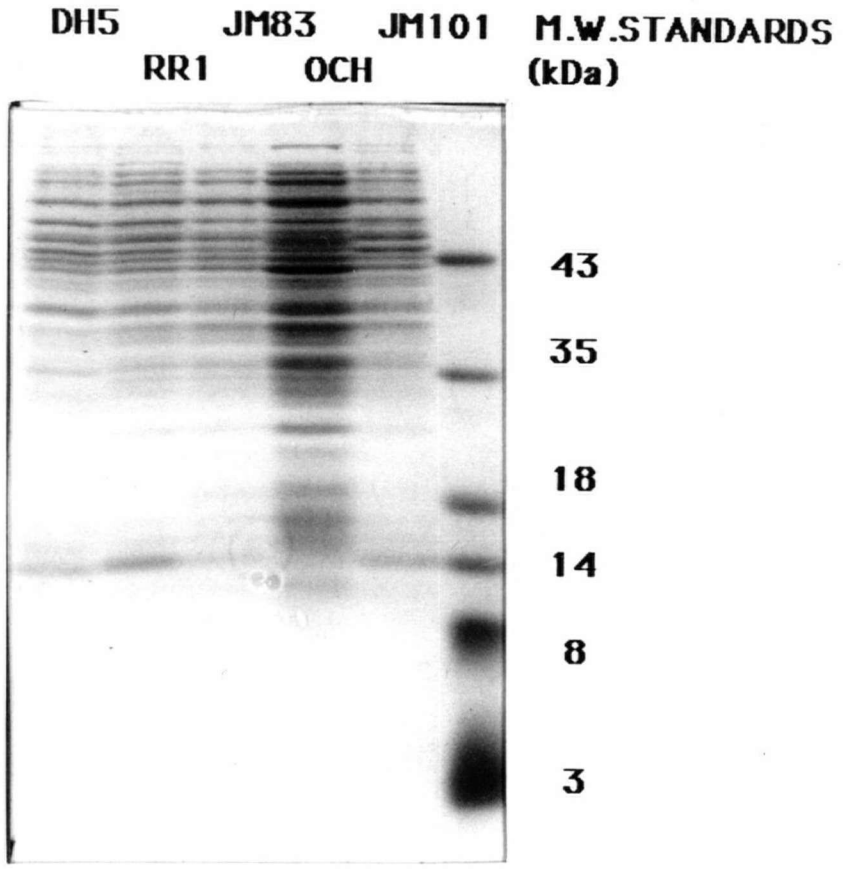


Figure 21: E. coli **Host Selection for Cytochrome b_5 Expression**

Plasmid pLpb_5, encoding the lipase-solubilized form of cytochrome b_5 was used to transform four different strains of *E. coli*. DH5, RR1 and JM101 strains were grown in YT medium supplemented with 1 mM IPTG, while JM83 strains were grown in YT medium alone. Based on $O.D._{600nm}$ readings, equivalent amounts of bacteria from 16 h. cultures were lysed by boiling and the resulting supernatants were analyzed by $NaDodSO_4$-PAGE (15% polyacrylamide). Also included was an overloaded amount of JM83 culture transformed with an ochre mutant of pLpb_5. Cytochrome b_5 runs anomalously large in these gel systems (M_r 14 kDa *versus* an empirically estimated weight of 10,590 Da for apo-Lpb_5).

---

determination and $NaDodSO_4$-PAGE analysis. As shown in Figure 22, the total amount of recombinant protein was highest when IPTG induction was initiated early in the culture (compare lanes A and J), and the maximal yield of all cultures was observed only

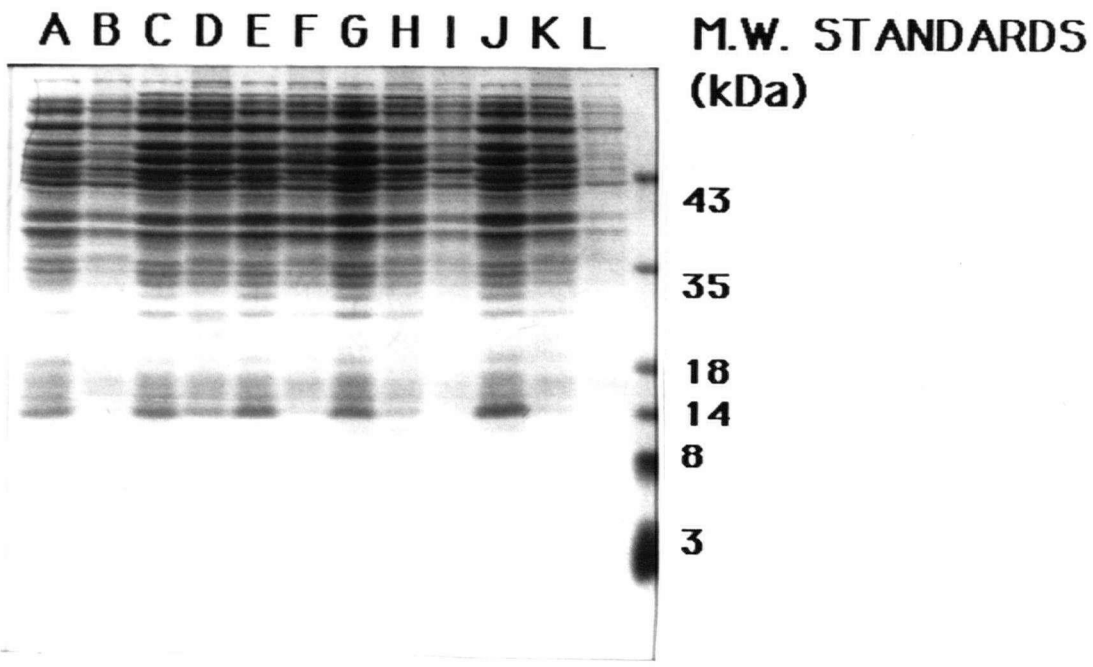


Figure 22: IPTG Induction Time Course of Lpb_5 Expression

Parallel cultures of pLpb_5-transformed *E. coli* RR1 and JM83 were grown in YT medium. RR1 cultures were induced with IPTG (1 mM final concentration) at various culture densities. Samples were harvested at subsequent time points for analysis by NaDodSO_4-PAGE (15% polyacrylamide).

Lane	Host	O.D.$_{600}$ at Induction	O.D.$_{600}$ at Harvest
A	JM 83	-	2.5
B	JM 83	-	1.5
C	RR1	2.0	2.5
D	RR1	2.0	2.1
E	RR1	1.5	2.5
F	RR1	1.5	1.9
G	RR1	1.0	2.5
H	RR1	1.0	2.2
I	RR1	1.0	1.3
J	RR1	0.5	2.5
K	RR1	0.5	1.6
L	RR1	0.5	0.9

after an 18 hour incubation. This observation demonstrates that the bacteria continue to produce recombinant protein well after entering stationary phase ($O.D._{600}$=2.4).

Fifty liter fermentation runs with JM83 routinely yield up to 500 mg of purified protein. However this yield represents only about 10% of the expected yield based on the gel analysis described above, if it is assumed that 15% of the wet weight of *E. coli* is protein (Lehninger,1970). As this discrepancy may reflect an inability of the bacterial host to supply sufficient heme for the cytochrome load, experiments were undertaken in which exogenous heme was added to culture lysates in an attempt to saturate any apoprotein which might remain in a relatively native form. No significant increase in protein yield was observed in these experiments, nor in other studies in which porphyrin precursors such as glycine and δ-aminolevulenic acid were added to the cultures. Alternatively, this discrepancy might suggest that only a fraction of the recombinant protein is in a form competent to bind heme, whether as a result of intracellular compartmentalization or protein degradation. Anomalous staining of cytochrome b_5 in the gel analysis could also explain this finding.

5. *Amino-terminal Sequence Analysis*

Gas phase protein sequencing of the amino-terminal region of the recombinant cytochrome showed complete agreement with the predicted sequence from DNA analysis (Table I). The absence of a peptide containing formyl-methionine in the tryptic peptide maps (see section IIIA7. Tryptic Peptide Mapping) indicates complete processing of the amino-terminus as reported for recombinant rat liver cytochrome b_5 (von Bodman *et al.*,1986).

6. *FPLC Analysis of Authentic and Recombinant Cytochromes* b_5

Our efforts to characterize fully the recombinant cytochrome b_5 included performing limited tryptic digests on the holoprotein in order to convert the lipase-

	Recombinant Cytochrome b_5		Tryptic Peptide T2		Tryptic Peptide T7			
	uncorrected		authentic		authentic		uncorrected	
Cycle	residue	yield*	residue	yield*	residue	yield*	residue	yield*
1	Ser	309	Tyr	818	Glu	963	Glu	989
2	Lys	156	Tyr	897	Gln	751	Gln	724
3	Ala	240	Thr	306	Ala	786	Ala	901
4	Val	809	Leu	572	Gly	612	Gly	655
5	Lys	180	Glu	545	Gly	725	Gly	875
6	Tyr	664	Glu	585	Asp	504	Asp	568
7	Tyr	544	Ile	299	Ala	510	Ala	599
8	Thr	380	Gln	434	Thr	242	Thr	252
9	Leu	145	Lys	188	Glu	375	Glu	473
10	Glu	135			Asp	312	Asn	396
11	Gln	583			Phe	324	Phe	414
12	Ile	359			Glu	275	Glu	370
13					Asp	238	Asp	277
14					Val	183	Val	239
15					Gly	152	Gly	239
16					His	78	His	106
17					Ser	54	Ser	54
18					Thr	54	Thr	55
19					Asp	83	Asp	102
20					Ala	61	Ala	78
21					Arg	10	Arg	15

*pmol

Table I: Amino-Terminal Sequence Analysis of Recombinant Lpb$_5$.

Gas phase sequence analysis was performed on the initial recombinant Lpb_5 (Q11,E13,D57), on tryptic digestion peptide T2 from the authentic tryptic core protein from bovine liver microsomes (corresponding to residues 6-14, using the Lpb_5 recombinant numbering system) and on peptide T7 (residues 52-76) from both the recombinant Lpb_5 and authentic tryptic core proteins.

---

solubilized form to the corresponding trypsin-solubilized form as reported for the authentic bovine protein (Ozols and Strittmatter,1966). This process releases peptides from both the N- and C-termini (see Figure 19) and therefore allowed direct comparison of recombinant and authentic molecules of identical chain length. FPLC analysis of the resulting digestion products revealed three species (data not shown), none of which exhibited elution behavior identical to that of the authentic trypsin-solubilized cytochrome b_5. This profile persisted even when the heme group was removed and subsequently replaced with exogenous hemin, indicating that the altered tryptic susceptibility was not the result of defective protein folding. The major digestion product had a prolonged retention time relative to that of the authentic sample (compare panels A and D, Figure 23), suggesting that it is more anionic in nature.

7. *Tryptic Peptide Mapping*

The initial peptide map of the recombinant cytochrome indicated that all of the predicted tryptic fragments were present. Within the resolution and error of the mapping process, the HPLC retention times of these peptides appeared to match those of the authentic trypsin-solubilized cytochrome. Additional peptides in the profile of the recombinant sample were predicted to result from the amino- and carboxy-terminal extensions of the lipase-solubilized recombinant form *versus* the trypsin-solubilized bovine liver cytochrome b_5. Amino acid composition analysis of these additional peptides confirmed this expectation.

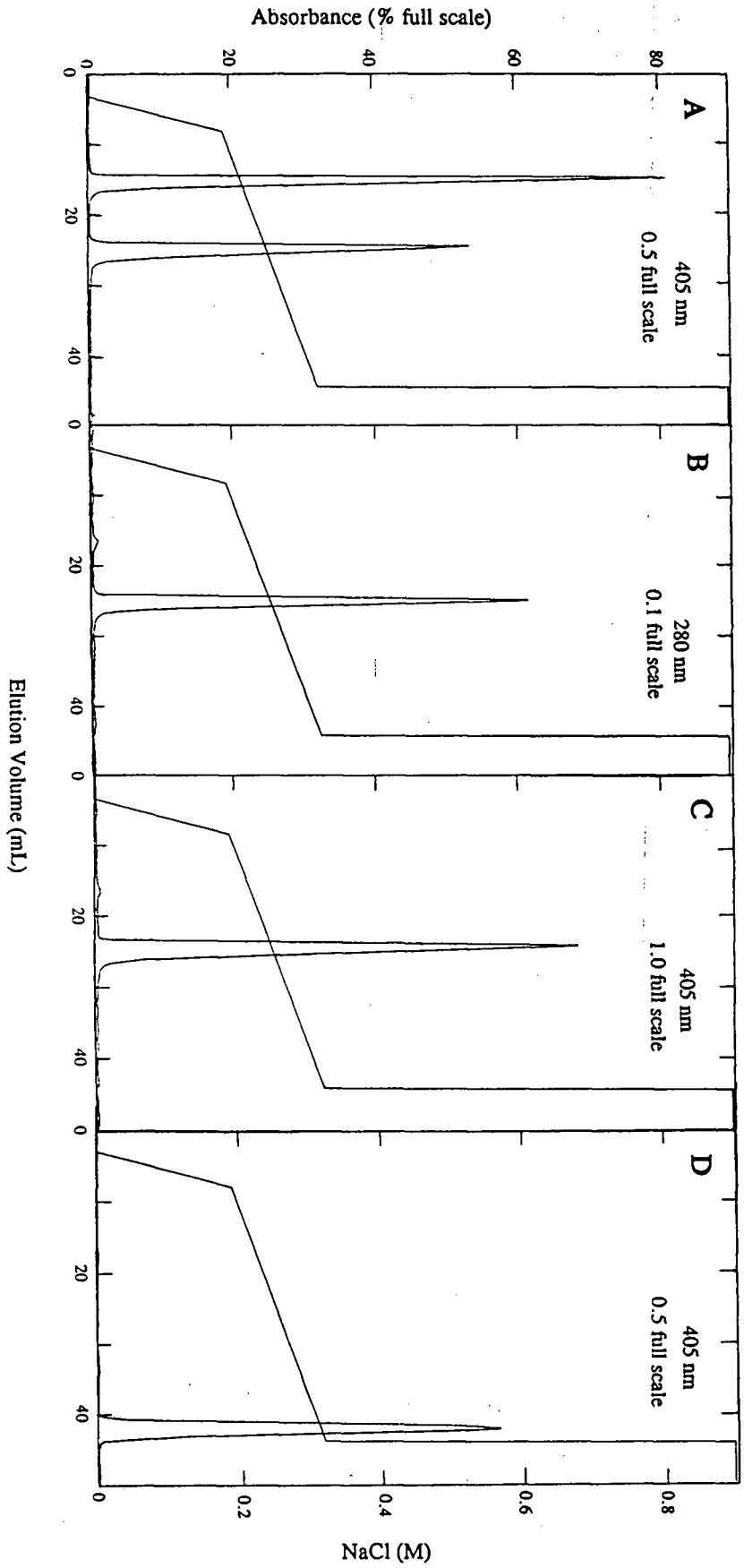

Absorbance (% full scale)
A
405 nm
0.5 full scale
B
280 nm
0.1 full scale
C
405 nm
1.0 full scale
D
405 nm
0.5 full scale
Elution Volume (mL)
NaCl (M)

Figure 23: FPLC Elution Profiles of Recombinant and Authentic Microsomal Cytochrome b_5

Limited tryptic digestion was performed on the corrected (E11,Q13,N57) and initial forms (Q11,E13,D57) of recombinant Lpb_5 (see IIB8, Experimental Procedures and Materials).

Panel A: Corrected recombinant Lpb_5 (V_e 15 mL) and tryptic core of authentic bovine microsomal cytochrome b_5 (V_e 25 mL).

Panel B: Tryptic core of corrected recombinant Lpb_5 (280 nm).

Panel C: Tryptic core of corrected recombinant Lpb_5 and the authentic bovine microsomal cytochrome b_5 (405 nm).

Panel D: Major product from the limited tryptic digestion of initial recombinant Lpb_5 (Q11,E13,D57).

Comparisons of the electrochemical (*vide infra*), chromatographic, cytochrome *c*-binding and methemoglobin-binding behavior (M.R. Mauk, W.D. Funk, unpublished results) and NMR properties (Burch *et al.*,1988) of the recombinant *versus* authentic cytochromes showed additional functional differences between the two proteins. Circular dichroism (CD) spectroscopy of the recombinant protein showed no significant difference relative to the authentic trypsin-solubilized protein (data not shown), suggesting that no gross structural re-organization of the molecule had occurred. However multiple peptide mapping experiments of both the recombinant and authentic cytochromes demonstrated slight but reproducible differences in the retention times of peptides T2 and T7 from the two proteins (Table II). Gas phase sequence analysis of these peptides indicated that the peptides from authentic trypsin-solubilized cytochrome b_5 were in concordance with the predicted amino acid sequence from cDNA analysis (Christiano and Steggles,1989), while the peptides from the recombinant protein agreed with the sequence obtained by Ozols and Strittmatter (1969), upon which the gene cassette sequence had been targetted (Table I). The two sequences differ in the amidation status of three residues; the corrected assignments are: Asn57, Glu11, and Gln13 (numbering of Ozols and Strittmatter,1969)

Peptide	Residues	Retention Time (min) Authentic[a]	Recombinant[b]	Difference
	5-6	---	6.2	---
T1	7-9	18.2	18.4	0.2
T2	10-18	71.4	72.8	1.4
T3	19-23	16.5	16.4	0.1
T4	24-32	109.0	109.1	0.1
T5	33-38	55.0	54.8	0.2
T6	39-51	86.4	86.2	0.2
T7	52-76	63.9	66.0	2.1
T8	73-76	28.4	28.5	0.1
T9	77-88	97.7	97.7	0.0
	89-90	---	6.2	---
	91-97	---	36.6	---

[a]Average of 5 peptide maps.
[b]Average of 11 peptide maps.

Table II: HPLC Retention Times of Tryptic Peptides from Recombinant Bovine Microsomal Lipase-solubilized Cytochrome b_5 (Q11,E13,D57) and Authentic Bovine Microsomal Trypsin-solubilized Cytochrome b_5

Peptides generated from the complete tryptic digestion of recombinant and authentic cytochrome b_5 samples were analyzed by HPLC. The mean retention times for each fragment are presented, along with the differences in retention times between molecular species. The authentic trypsin-solubilized cytochrome b_5 is shorter at both the amino- and carboxy-terminus compared to the recombinant lipase-solubilized form.

8. *Primary Sequence Correction*

The amidation status of residues 11, 13 and 57 of the bovine liver cytochrome b_5 was confirmed independently by sequence analysis of the corresponding cDNA by Christiano and Steggles (1989). Two oligonucleotides spanning the sequence for residues 11 and 13 and for residue 57 were synthesized and used to effect the necessary replacements by site-directed mutagenesis. The resulting wild-type (corrected) lipase-solubilized cytochrome still displayed a different FPLC retention time than that exhibited by the authentic trypsin-solubilized protein (Figure 23, panel A). Significantly, the trypsin digestion product of the corrected protein eluted as a single peak (panel B) that eluted with a retention time identical to that of the authentic trypsin-solubilized protein (panel C).

The complex elution profile of the triple mutant trypsin hydrolysis reaction suggested that the termini of the molecule were less susceptible to proteolytic hydrolysis by trypsin than wild-type lipase b_5. To investigate this possibility, restriction fragments from the gene cassettes for the corrected wild-type and triple mutant cytochromes were joined to produce molecules mutated either at position 57 (D57) or 11 and 13 (Q11,E13). The resulting molecules were subjected to limited trypsin digestion and chromatographed as before. Surprisingly, both the D57 and Q11,E13 mutant proteins yielded a single reaction product (Figure 24, panel B) suggesting the behavior of the triple mutant towards limited tryptic digestion must result from a subtle interplay between the two ends of the molecule. The FPLC elution profiles of the lipase form of the D57 mutant shows that this mutation alone is sufficient to account for the increased anionic character of the triple mutant (compare Figure 24, panel C *versus* Figure 23, panel D). Interestingly, the lipase form of the Q11,E13 mutant appears to have a slightly altered FPLC retention time from the corrected wild-type molecule (compare Figure 24, panel

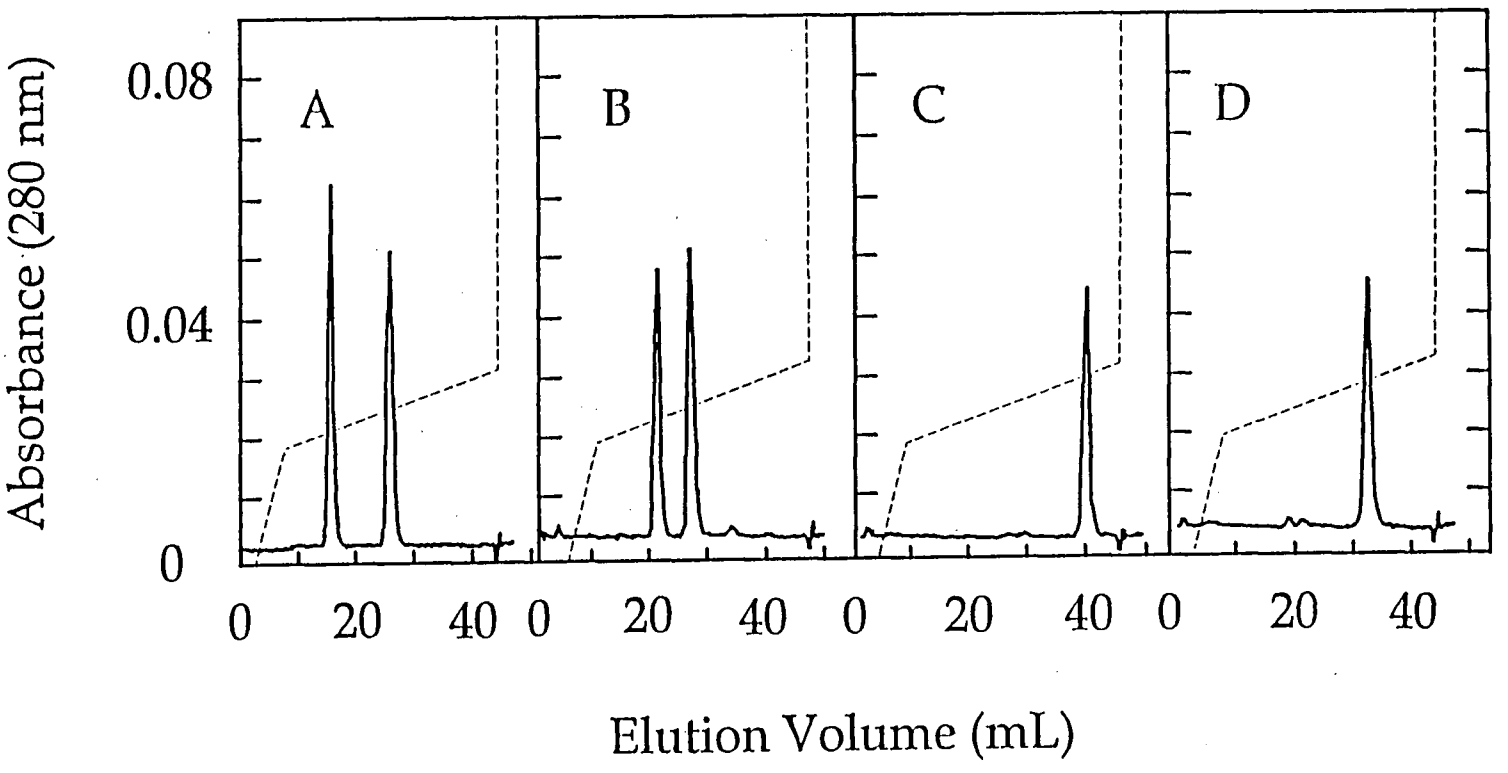

Absorbance (280 nm)
0.08
0.04
0
A
B
C
D
0
20
40
Elution Volume (mL)

Figure 24: FPLC Elution Profile of the Limited Digestion Products of Recombinant Lpb$_5$ Species (Q11,E13 and D57)

Limited tryptic digestion was performed on two species of recombinant Lp*b*$_5$, Q11,E13 and D57 as described in Experimental Procedures and Materials, B8.

Panel A: Corrected recombinant Lp*b*$_5$ (V_e 15 mL) and the tryptic core of authentic bovine microsomal cytochrome b_5 (V_e 25 mL).

Panel B Recombinant Lp*b*$_5$ species prior to digestion: Q11,E13 (V_e 22 mL) and D57 (V_e 28 mL).

Panel C: Recombinant L*p*b$_5$ D57 following limited tryptic digestion.

Panel D: Recombinant Lp*b*$_5$ Q11,E13 following limited tryptic digestion.

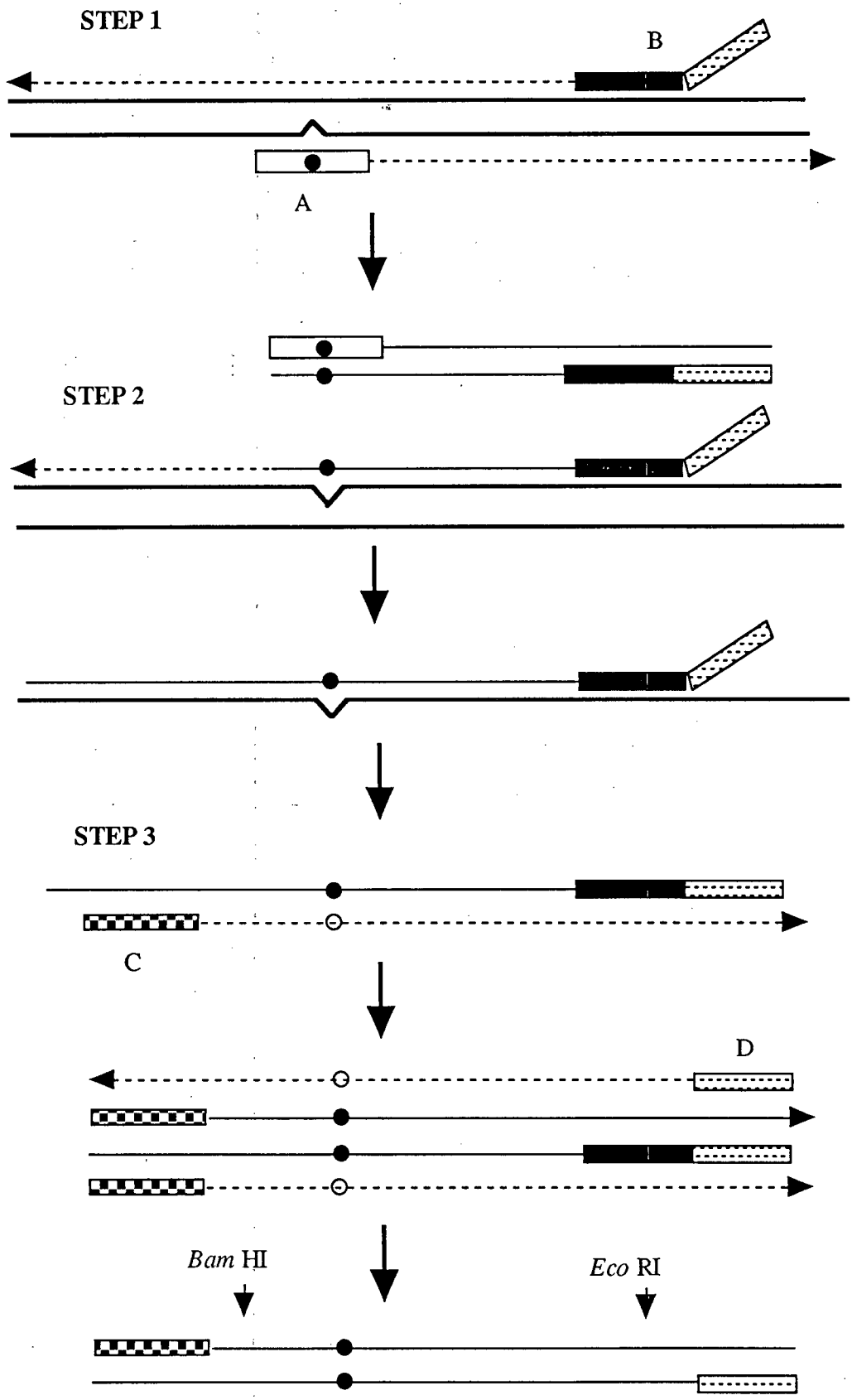
D *versus* Figure 23, panel B). Following conversion to the smaller tryptic core, the Q11,E13 mutant elutes at a later point in the gradient (V_e 33 mL, Figure 24, panel D) than does the authentic tryptic core (V_e 25 mL, Figure 24, panel B).

9. *Constructions for the Expression of Trypsin-Solubilized Bovine Hepatic and Human and Erythrocytic Cytochrome b_5*

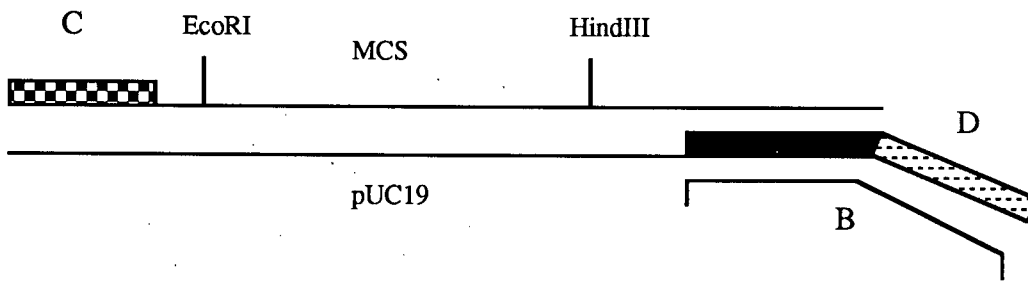
Conversion of the original pLp*b*$_5$ plasmid to encode the smaller trypsin-solubilized form of the same protein was achieved by using site-directed mutagenesis to first delete the coding sequence of Ser1 and Lys2 at the 5' end of the coding sequence and then to introduce a new translational stop codon immediately following the codon for Arg84. The resulting protein was recovered with similar yields as the lipase-solubilized form, exhibited identical FPLC elution characteristics with the authentic bovine-liver isolated protein (data not shown) and has a midpoint potential of -3 mV *versus* SHE (P.D. Barker, personal communication).

The conversion of pLp*b*$_5$ to encode the human erythrocytic form of cytochrome b_5 was accomplished using the PCR-directed mutagenesis procedure of Nelson and Long (1988). As illustrated in Figure 25, the process involves two separate PCR amplification procedures, primed by two pairs of oligonucleotides. Step 1 amplifies the sequence

A



B



PCR Mutagenesis Oligonucleotides

B 5'-P-GGA GTA CTA GTA ACC CTG GCC CCA GTC ACG ACG TTG TAA A-OH-3' (40 MER) (- STRAND)

C 5'-P-CAG GAA ACA GCT ATG ACC AT-OH-3' (20 MER) (+ STRAND)

D 5'-P-GGA GTA CTA GTA ACC CTG GC-OH-3' (20 MER) (- STRAND)

Figure 25: PCR Mutagenesis Strategy

Panel A: The method of Nelson and Long (1989) was used to generate the human erythrocytic form of cytochrome b_5 from the pLpb_5 gene cassette. Step 1 involves the amplification of a segment defined by the mutagenizing oligonucleotide A and a flanking oligonucleotide B. At its 3' end, Oligo B matches the multiple cloning site (MCS) flanking sequence of pUC vectors but its 5' end consists of a unique sequence. In Step 2, the product of Step 1 is used to prime a single cycle of polymerization which extends beyond the opposite flanking region. As the 3' end of the sense strand of the Step 1 product is not complementary to the template, only the mutant strand will prime polymerization. Oligonucleotides C and D are added to the reaction in Step 3 and amplification of the intervening sequence is completed. The product of Step 3 can be digested with appropriate restriction enzymes for subsequent re-cloning.

Panel B: The sequences of the flanking oligonucleotides (B,C, and D) and their relative orientations at the pUC 19 multiple cloning site (MCS) are shown. As these oligonucleotides lie outside of the MCS, they can be used with either orientation of pUC vectors as well as with the M13mp series of single strand-producing vectors.

enclosed by the mutagenizing oligonucleotide and a second oligonucleotide hybridized to the cloning site flanking region. The double-stranded product of Step 1 is then used to prime a second round of PCR amplification spanning the remainder of the coding sequence (see Figure 25 for details).

This PCR-based mutagenesis procedure offers several advantages over standard M13-based procedures. First, the amplification strategy assures that only the mutant will be produced in significant amounts, thereby eliminating the sometimes tedious process of screening multiple samples as is normally the case with M13-based mutagenesis. In addition, PCR mutagenesis is performed on double stranded template and thus eliminates both the need for isolating single-stranded template and the inherent shuttling of modified coding sequences back into the expression plasmid. The coding sequence of the resulting mutants was readily verified by double-stranded dideoxy sequence analysis (Gatermann *et al.*,1988) using both forward and reverse orientation universal sequencing primers. One recurrent problem with this PCR procedure has been the inadvertent introduction of additional nucleotide substitutions during the amplification procedure. These errors were found predominantly at the nucleotide pair immediately 5' to the non-priming (5') end of the mutagenizing oligonucleotide, and almost certainly result from the propensity of *Taq* polymerase to add an additional residue to the 3' ends of amplified products (Gelfman and White,1990). The clones resulting from the PCR process are predominantly of the desired sequence. It may be possible to eliminate such mutants introduced by polymerase errors through the addition of a blunting procedure following Step 1 using an exonuclease activity such as that provided by the Klenow fragment.

The resulting plasmid encoding human erythrocytic cytochrome b_5 (phErb_5) failed to produce any recombinant protein as detected by NaDodSO$_4$-PAGE in the bacterial host strain *E. coli* JM83. The cause of this poor expression may be different from that responsible for the poor yield observed with several of the site-specific

mutations performed on pLpb_5; these mutants fail to incorporate heme yet show equivalent amounts of recombinant protein when compared to the wild-type standard on NaDodSO$_4$-PAGE (data not shown). As the promoter regions and transcription patterns of these two constructions should be identical, these results suggest that the human erythrocytic construction is translated with extremely low efficiency.

10. *Spectroelectrochemical Studies*

Potentiometric titrations were performed on both the corrected and triple mutant wild-type cytochromes and on the S64A mutant in both the corrected and triple mutant backgrounds. As seen in the representative spectra of Figure 26, the recombinant cytochromes displayed electrochemically reversible behavior as had been observed for the cytochrome purified from liver (Reid *et al.*,1982). Such Optically Transparent Thin Layer Electrode (OTTLE) potentiometric measurements rely on the presence of a small molecule mediator to transfer electrons from a thin surface working electrode to the protein (see Appendix I for a discussion of the electrochemical methods used in this investigation). This oxidation-reduction equilibrium is monitored by measuring the electronic absorption of the solution at a wavelength where the reduced and oxidized forms of the protein exhibit different extinction coefficients. The resulting data are analyzed in terms of a Nernst plot by fitting to the Nernst equation:

$$E_{app} = E_{1/2} + RT/nF \log ([Red]/[Ox])$$

In general, all of the cytochrome species examined achieved electrochemical equilibrium at a given potential within 30 minutes. The electronic absorption spectra of the S64A mutant appeared to match those of the wild-type protein closely. Isosbestic points for the wild-type protein were observed at 566, 545, 535, 514, 438, 416 and 353 nm and at 567, 545, 535, 514, 438, 416 and 351 nm for the S64A mutant, indicating that this substitution did not significantly affect the heme chromophore. The corresponding Nernst plots shown in the inset of Figure 26 were calculated from absorption readings

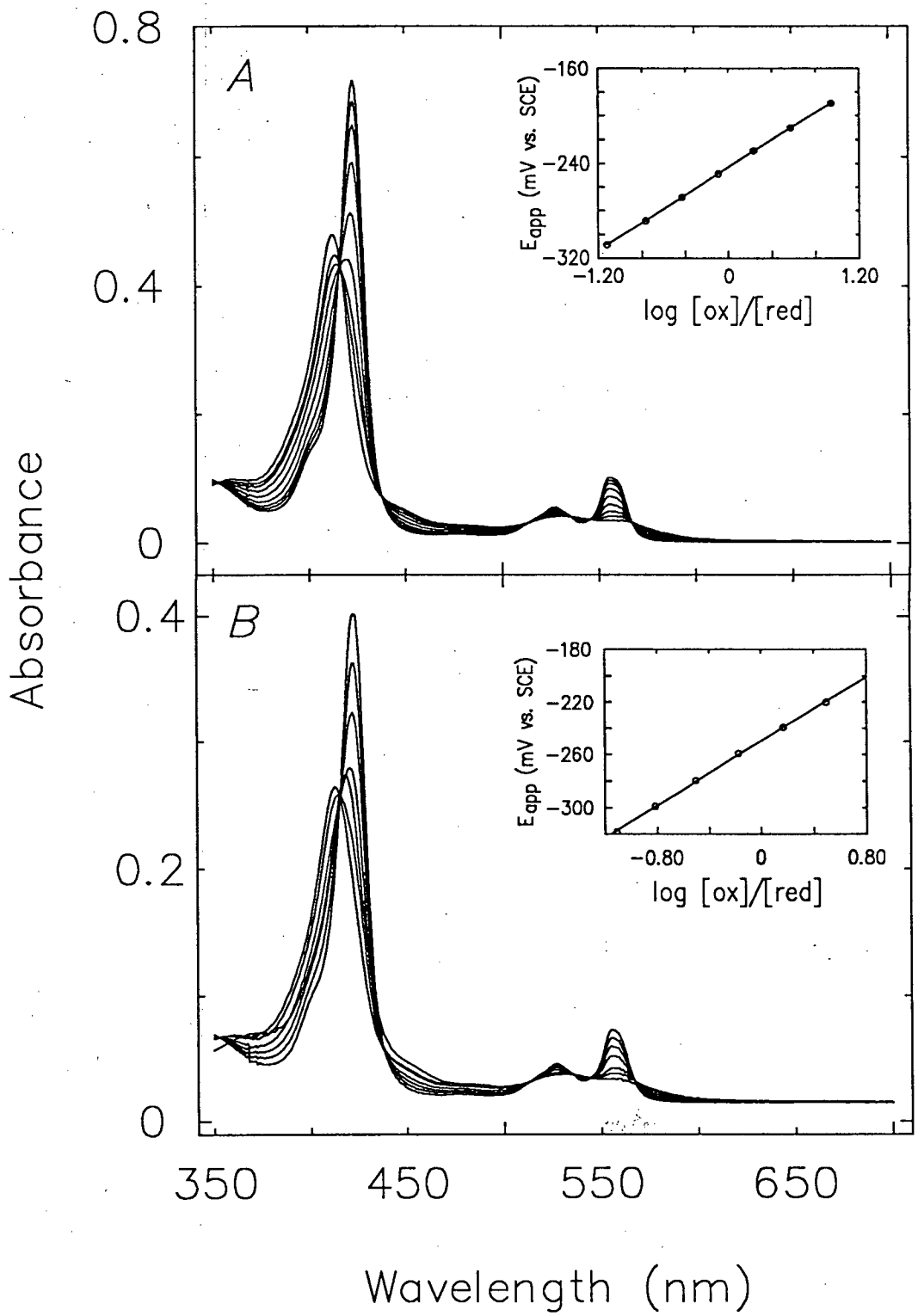

A
B
Absorbance
Wavelength (nm)
0.8
0.4
0
0.4
0.2
0
350
450
550
650
E_{app} (mV vs. SCE)
−160
−240
−320
−1.20
0
1.20
log [ox]/[red]
−180
−220
−260
−300
−0.80
0
0.80
log [ox]/[red]

Figure 26: Thin Layer Spectra of Recombinant Wild-type and S64A Forms of Lpb_5

Thin layer spectra were obtained using an OTTLE cell apparatus mounted in a Cary 219 spectrophotometer. The applied voltage ranged from -320 to +50 mV *versus* SCE. The absorbance values at 423.5 nm were used to generate Nernst plots which are presented in the inserts. Cytochrome b_5 (120 μM), $Ru(NH_3)_6^{2+3+}$ (12μM), 25°C, pH 7.0 (phosphate), μ= 0.1 M.

A: Wild-type recombinant bovine microsomal Lpb_5.

B: S64A variant of recombinant bovine microsomal Lpb_5.

---

taken at 423.5 nm. The resulting midpoint potentials are listed in Table III.

The midpoint potential of the corrected wild-type molecule (3 mV) is virtually identical to the value reported for the trypsin-solubilized protein isolated from liver (Reid *et al.*,1982). The S64A mutant exhibits a slightly lower potential (-4 mV). This result is somewhat surprising result in that the removal of the hydroxyl group at this position was predicted to destabilize the propionate carboxyl charge, stabilize the reduced form of the cytochrome, and thereby increase the reduction potential of the protein (see Figure 7). The midpoint potentials of both the corrected wild-type and S64A species were shifted upwards by 67 mV when reconstituted with DME heme as had been observed for the authentic liver protein (Reid *et al.*,1984). The S64A substitution in the triple mutant background could only be recovered from the bacterial lysate with the addition of exogenous hemin and quickly lost heme when chromatographed over MonoQ FPLC columns. Attempts to reconstitute this species with DME-heme were unsuccessful.

Cyclic voltammetry was used as an alternate method for obtaining the midpoint potentials of the D57 and Q11,E13 mutant species. This procedure cycles the protein through a range of applied potentials and relies on measurement of the current produced during reduction and oxidation to determine the midpoint potential. At low potential, the protein is essentially fully reduced. As the potential is increased, the resulting

Protein	Prosthetic Group	E 1/2 (mV *vs.* SHE)[a]	Nernst Slope (mV)
wild-type (E11,Q13,N57) (correct sequence)	protoheme IX	3	57.9
	DME-protoheme IX	70	61.2
S64A (E11,Q13,N57)	protoheme IX	-4	59.0
	DME-protoheme IX	63	56.9
wild-type (Q11,E13,D57) (incorrect sequence)	protoheme IX	-14	59.0
S64A (Q11,E13,D57)	protoheme IX	-19	62.0
wild-type (Q11,E13,N57)[b]	protoheme IX	3	
wild-type (E11,Q13,D57)[b]	protoheme IX	-15	

[a] 25 C, 0.1M sodium phosphate, pH 7.0
[b] Determined by using cyclic voltammetry

Table III: Midpoint Reduction Potentials of Recombinant Cytochrome b_5 Species.

Midpoint potential were determined for several forms of recombinant Lpb_5 and selected dimethylester heme (DME) replacement species. Reduction potentials were estimated either from Nernst plots of thin-layer spectra measurements or by using cyclic voltametric techniques (see also Appendix I).

current first increases with rising potential, reaches a maximum, called the cathodic peak and then subsides. An anodic peak is obtained upon reversing the scan and for reversible (Nernstian) reactions, the midpoint potential lies midway between the cathodic and anodic peaks, which are themselves separated by 57/n mV (n is the number of

electrons in the half reaction). However, for many real systems, such as those involving proteins, the rate of electron transfer is slow compared to the rate of diffusion and the reaction is referred to as being quasi-reversible. In such cases the peak currents are separated by a larger potential (see Appendix I for additional detail on cyclic voltammetric methods).

The D57 mutant displays the same midpoint potential as the triple mutant (Table III), indicating that this substitution is solely responsible for the decreased potential of the triple mutant. In concordance with this result, the Q11,E13 mutant yielded the same midpoint potential as the corrected wild-type protein (+3 mV).

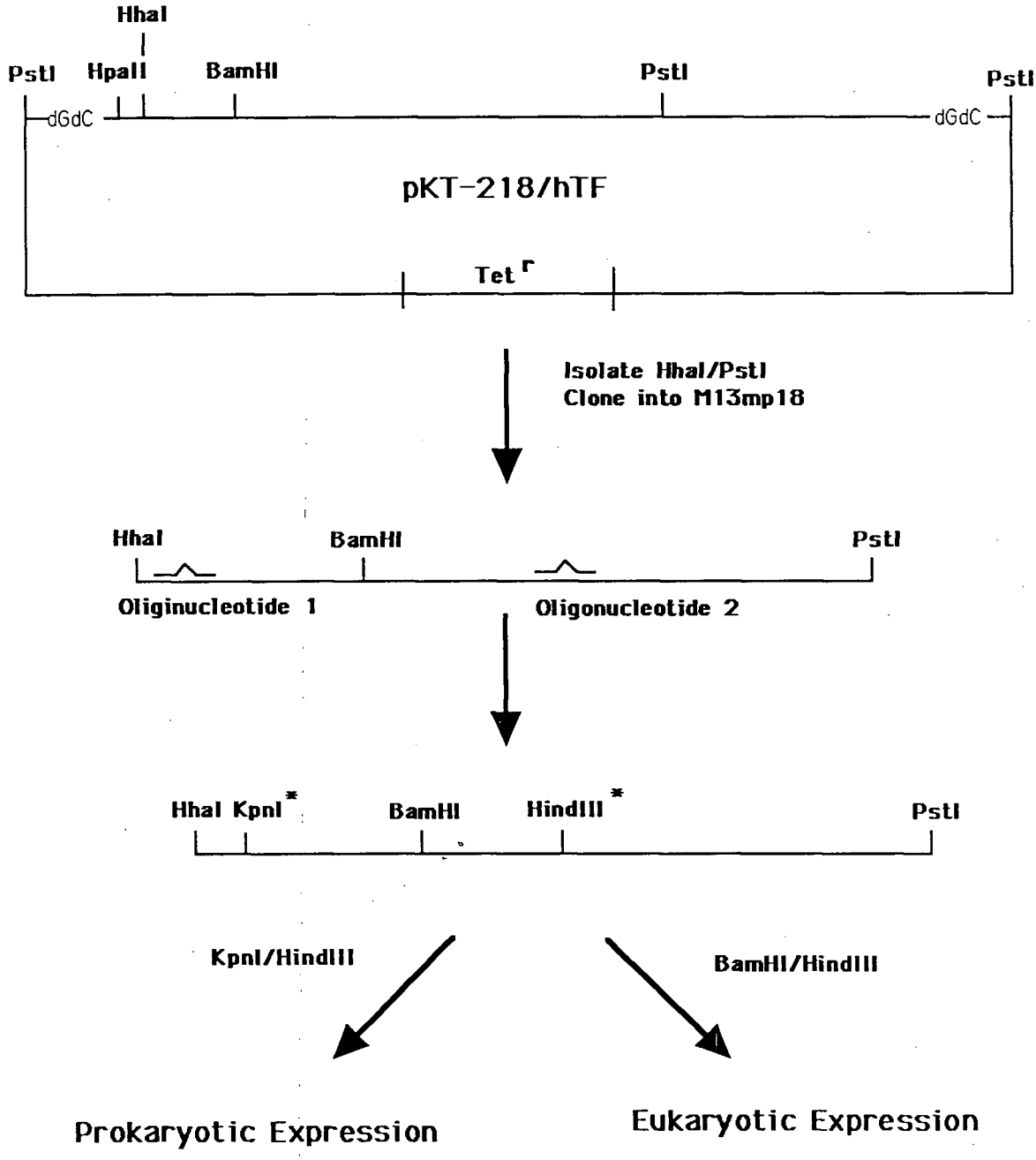
B. Transferrin Studies

1. *cDNA Isolation and Prokaryotic Expression Constructions*

Approximately 100,000 colonies of a human liver cDNA library (Prochownik *et al.*,1983) were screened by using a 24-base oligonucleotide corresponding to the 5' sequence of the human TF cDNA as a hybridization probe. A single positive clone was isolated whose plasmid (pKT-218/hTF) displayed the identical restriction endonuclease mapping pattern as the full-length cDNA isolated by Yang *et al.* from the same library (1984). Dideoxy sequence analysis of this clone showed that the 5' and 3' termini were identical to those of the full-length cDNA from Yang *et al.* and all subsequent sequence analysis performed during this study were also in complete agreement.

A *Hha*I/*Pst*I fragment from pKT218-hTF was subcloned into M13mp18 to produce single-stranded template for use in site-specific mutagenesis. Two translational stop codons and a unique *Hin*dIII recognition were introduced to the linker peptide coding sequence and a second oligonucleotide was used to produce a *Kpn*I recognition site at the signal peptide cleavage site coding sequence (Figure 27). The predicted translation product from this sequence terminates following Asp337. These

A


HhaI
PstI
HpaII
BamHI
PstI
PstI
dGdC
dGdC
pKT-218/hTF
Tet r
Isolate HhaI/PstI
Clone into M13mp18
HhaI
BamHI
PstI
Oliginucleotide 1
Oligonucleotide 2
HhaI KpnI *
BamHI
HindIII *
PstI
KpnI/HindIII
BamHI/HindIII
Prokaryotic Expression
Eukaryotic Expression

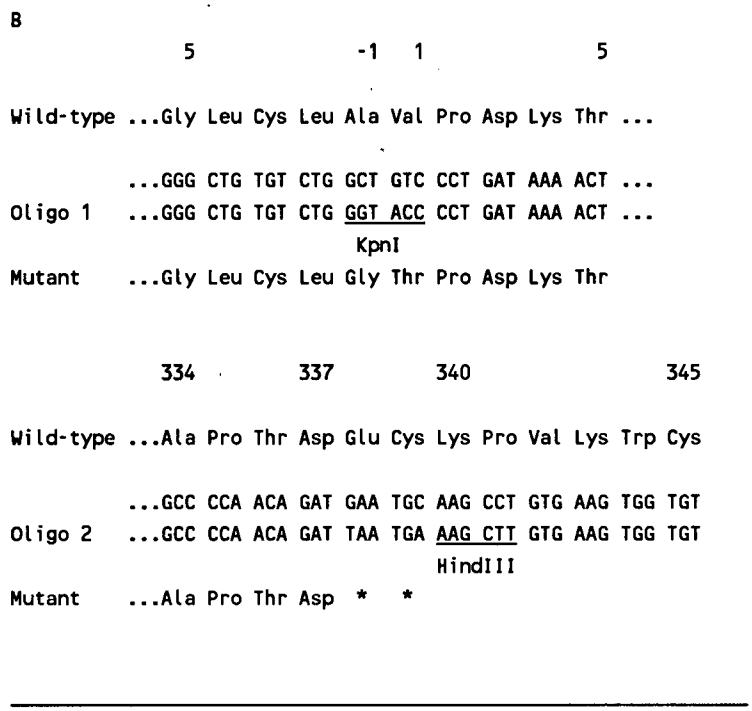


Figure 27: In vitro **Mutagenesis Strategy for the Assembly of hTF/2N Coding Sequence**

Panel A: A portion of the human serum transferrin cDNA containing the entire N lobe coding sequence was subcloned from pKT-218/hTF into M13mp18. A single round of mutagenesis was performed using two separate oligonucleotides (1 and 2) after which clones were selected for the presence of two new restriction endonuclease recognition sites, *Kpn*I and *Hin*dIII. A *Kpn*I/*Hin*dIII fragment was released for use in prokaryotic expression studies, while a *Bam*HI/*Hin*dIII fragment was used in the assembly of the pNUT-based expression system.

Panel B: The nucleotide and amino acid substitutions effected by oligonucleotides 1 and 2 from Panel A are presented. Oligonucleotide 1 creates a unique *Kpn*I recognition sequence at the signal cleavage site of hTF. This substitution results in the replacement of residues -1 and +1 in the amino acid sequence and allows the separation of the mature N lobe sequence from the signal peptide sequence. Oligonucleotide 2 introduces a unique *Hin*dIII recognition sequence and two translational termination codons immediately 3' to the codon for Asp337.

---

manipulations allowed the isolation of a *Kpn*I/*Hin*dIII restriction fragment which lacks the signal peptide sequence and encodes the amino-terminal lobe of transferrin. This fragment was then spliced to an oligonucleotide duplex which encodes the *E. coli*

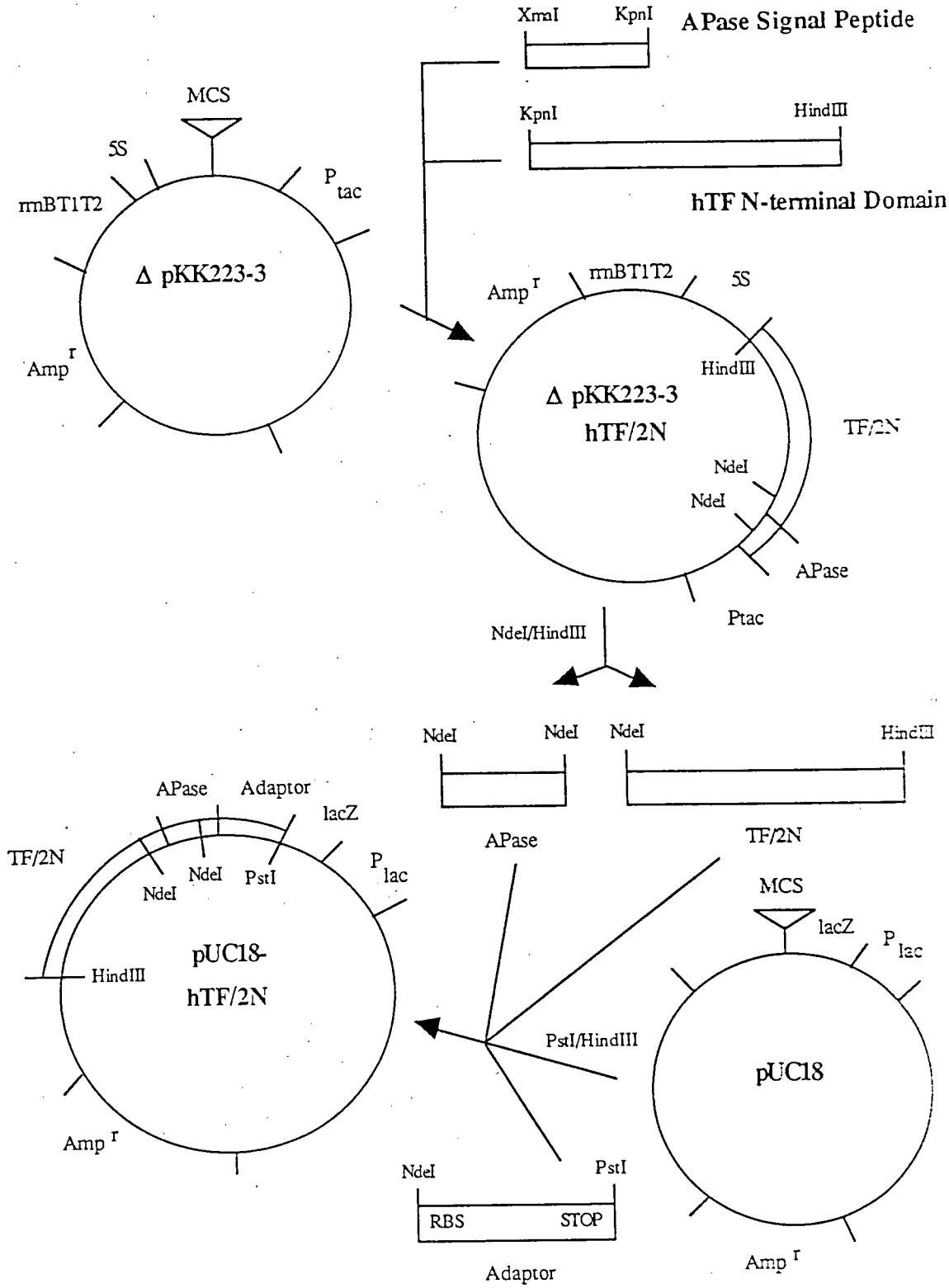

XmaI
KpnI
APase Signal Peptide
KpnI
HindIII
hTF N-terminal Domain
MCS
5S
rrnBT1T2
P tac
Δ pKK223-3
Amp r
Amp r
rrnBT1T2
5S
HindIII
Δ pKK223-3
hTF/2N
TF/2N
NdeI
NdeI
APase
Ptac
NdeI/HindIII
NdeI
NdeI
NdeI
HindIII
APase
Adaptor
lacZ
TF/2N
P lac
NdeI
NdeI
PstI
HindIII
pUC18-
hTF/2N
Amp r
TF/2N
MCS
lacZ
P lac
PstI/HindIII
pUC18
NdeI
PstI
RBS
STOP
Adaptor
Amp r

Figure 28: Vectors for the Expression of Recombinant hTF/2N in Bacteria

Two separate expression vectors were assembled for the expression of hTF/2N in bacterial systems. Both utilized a synthetic adaptor encoding the *E. coli* alkaline phosphatase (APase) signal sequence fused to the coding sequence for the mature form of hTF/2N. The expression vector ΔpKK223-3 is a modified version of pKK223-3 in which the sequence separating the ribosome binding site and initiator methionine has been shortened to 11 b.p. (D.R. Cleveland, University of British Columbia). pKK223-3 utilizes the hybrid tac promoter element (P_{tac}) to drive transcription of sequences inserted into the multiple cloning site (MCS). Transcriptional termination is assured by the presence of the *E. coli* ribosomal RNA operon which includes a GC-rich inverted repeat termination signal (rrnT1T2).
A second expression system utilizing the same transcriptional fusion strategy described for the cytochrome b_5 studies (Figure 20) was assembled into the parental vector pUC18. A short adaptor oligonucleotide encoding an in-frame translational stop codon for the lac Z gene product and an adjacent ribosome binding site (RBS) was ligated with the appropriate restriction fragments encoding hTF/2N from ΔpKK223-3/hTF/2N to yield pUC18-hTF/2N.

---

alkaline phosphatase signal peptide and was subsequently introduced into the expression vector ΔpKK223-3 (Figure 28). ΔpKK223-3 (D.R. Cleveland, University of British Columbia) differs from the parental vector pKK223-3 (Brosius and Holy,1984) in that the spacer sequence between the ribosome binding site and initiator methionine codon has been shortened to 11 bases. The same alkaline phosphatase signal peptide had been previously shown to target recombinant bovine pancreatic trypsin inhibitor to the periplasmic space (Marks *et al.*,1986) and we had hoped that passage of the transferrin half molecule through the bacterial membrane would encourage proper folding of the protein. However NaDodSO$_4$-PAGE and Western analysis of lysates from transformed bacteria failed to detect recombinant protein (results not shown).

As a second attempt at eukaryotic expression, a transcriptional fusion strategy similar to that described in the cytochrome b_5 studies was employed for the transferrin half molecule. As illustrated in Figure 28, an adaptor oligonucleotide duplex was assembled which encoded an in-frame translational stop codon for the lac Z product of pUC18 and an adjacent ribosome binding site. This was then joined with the alkaline phosphatase-TF/2N coding sequence and introduced into pUC 18. Again, recombinant protein was undetectable in bacterial lysates.

2. *Eukaryotic Expression Construction*

The expression vector pNUT (Palmiter et al.1987) contains a mouse metallothionein I/human growth hormone gene fusion that has been shown to direct the expression of high levels of human growth hormone in transgenic mice (Palmiter *et al.*,1983). Important components of this expression vector include the heavy metal-inducible metallothionein promoter, pUC18 sequences to allow replication in *E. coli*, and a dihydrofolate reductase (DHFR) cDNA driven by the SV40 early promoter and terminated by sequences from the 3' end of the hepatitis B virus (Figure 29). The DHFR cDNA (Simonsen and Levinson, 1983) encodes a mutant form of the protein which has a 270-fold lower affinity for the competitive inhibitor methotrexate (MTX). This allows for the immediate selection of transfected cells with high concentrations of MTX (0.5 mM), thereby abrogating the need for a recipient cell line that is deficient in DHFR.

The assembly of the expression vector pNUT-hTF/2N is illustrated in Figure 29. A *Bam*HI/*Hin*dIII restriction fragment was first isolated from the bacterial vector ΔpKK223-3-hTF/2N and then ligated with a *Bam*HI/*Hpa*II fragment from the 5' end of the full length cDNA vector pKT-218/hTF into M13mp18. This allowed the isolation of an *Xba*I/*Hin*dIII fragment which was subsequently made blunt and inserted into the large *Sma*I fragment of pNUT. The *Xba*I/*Hin*dIII fragment is free of the dG/dC tail found in the cDNA vector and encodes the natural signal sequence of hTF. Transcriptional termination and polyadenylation signals are retained from the 3' end of the hGH gene in the final pNUT-hTF/2N construction.

Transfection of BHK cells with pNUT-hTF/2N allows for rapid selection of cells with high concentrations of MTX. Ribonucleic acid was isolated from pNUT and pNUT-hTF/2N transfected cells, electrophoresed through an agarose gel in the presence of formaldehyde (Maniatis *et al.*,1982) and then transferred to nitrocellulose. Using an

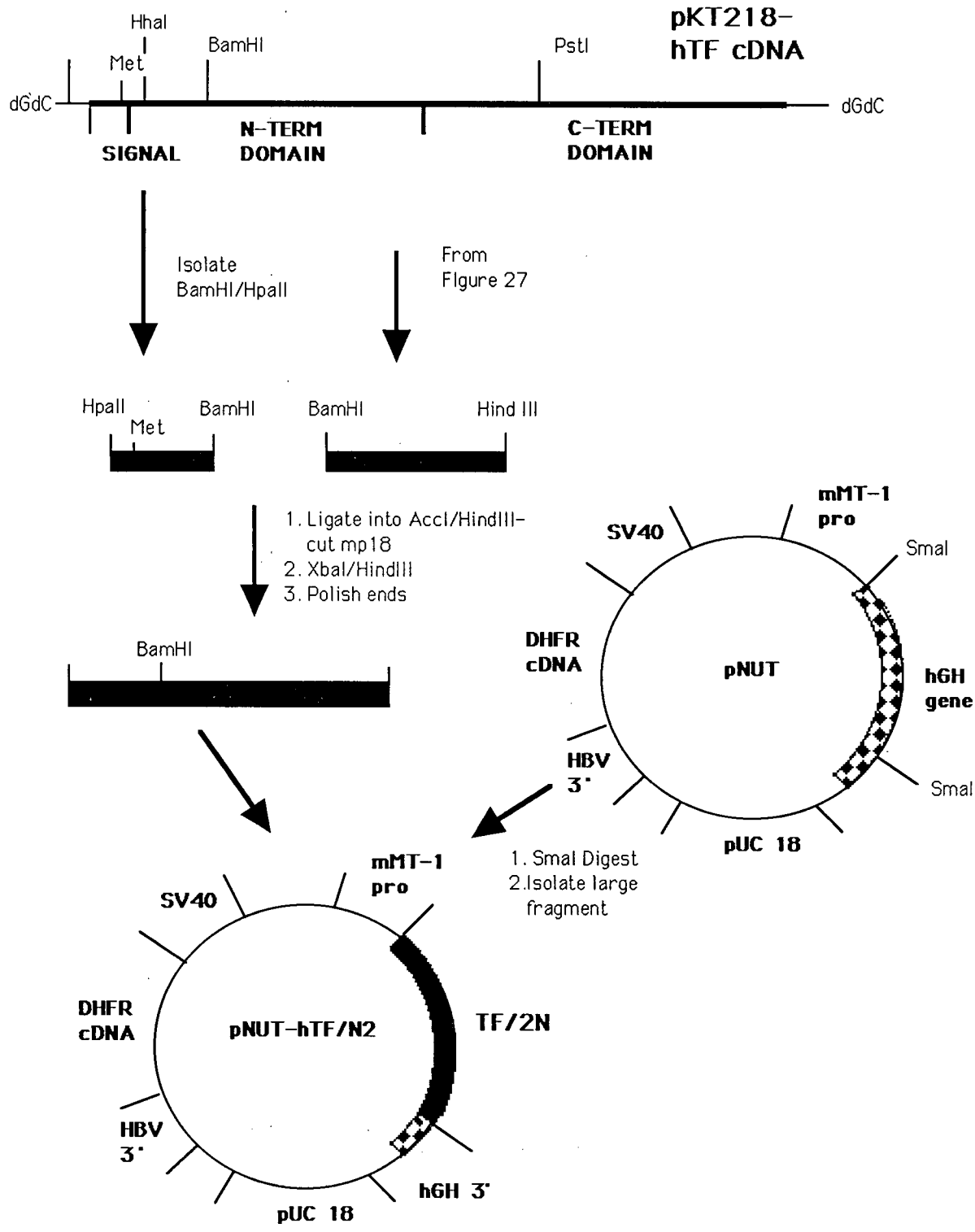

pKT218-
hTF cDNA
HhaI
Met
BamHI
PstI
dGdC
dGdC
SIGNAL
N-TERM
DOMAIN
C-TERM
DOMAIN
Isolate
BamHI/HpaII
From
Figure 27
HpaII
Met
BamHI
BamHI
Hind III
1. Ligate into AccI/HindIII-
cut mp18
2. XbaI/HindIII
3. Polish ends
BamHI
mMT-1
pro
SV40
SmaI
DHFR
cDNA
pNUT
hGH
gene
HBV
3'
SmaI
pUC 18
1. SmaI Digest
2.Isolate large
fragment
mMT-1
pro
SV40
DHFR
cDNA
pNUT-hTF/N2
TF/2N
HBV
3'
hGH 3'
pUC 18

Figure 29: pNUT-hTF/2N Expression Vector Assembly

The coding sequence for hTF/2N was assembled from a *Hpa*II/*Bam*HI restriction fragment from the 5' end of the hTF cDNA and a *Bam*HI/*Hin*dIII fragment from the initial mutagenesis procedures described in Figure 27. This coding sequence was released by restriction digestion and the resulting fragment was made blunt and inserted into the large *Sma*I fragment of pNUT. Plasmid pNUT includes the promoter element of the mouse metallothionein-1 gene (mMT-1) to allow metal inducible transcription of downstream sequences and the *Sma*I digestion fragment retains the 3' transcriptional termination and polyadenylation signals from the human growth hormone gene (hGH). A cDNA encoding a mutant form of the dihydrofolate reductase gene (DHFR) is driven by the SV40 early promoter (SV40) to provide a selectable element for eukaryotic cell expression, while bacterial replication is maintained by the pUC18 sequences.

---

oligonucleotide to the 3' end of the hGH gene as a hybridization probe, an inducible mRNA of approximately 1.4 kbp was detected in the pNUT-hTF/2N transfected cells, in good agreement with the predicted size of the hTF/2N-hGH mRNA (data not shown).

Western blot analysis was performed on both the medium and solubilized cell pellet from three cell lines. Samples were treated successively with goat anti-hTF antiserum and *S. aureus* cells and the resulting precipitate was then electrophoresed through a 12% polyacrylamide gel and transferred to nitrocellulose. The blot was incubated with the same goat anti-hTF antiserum and finally with alkaline phosphatase-conjugated anti-goat immunoglobulin, then developed with the appropriate substrates. As seen in Figure 30, medium and cell lysates from BHK cells and pNUT-transfected cell lines showed only the expected goat immunoglobulin bands (M_r 25 kDa and 50 kDa) and a small amount of cross reactive material common to all samples, while the pNUT-hTF/2N material showed the presence of a strong band at Mr 37 kDa, in excellent agreement with the calculated molecular weight of recombinant hTF/2N based on amino acid sequence (37833 Da).

3. *Isolation and Characterization of Recombinant hTF/2N*

Large-scale culture of the hTF/2N cell line was performed using roller culture bottles. Radioimmunoassays of the culture medium indicated an initial concentration of

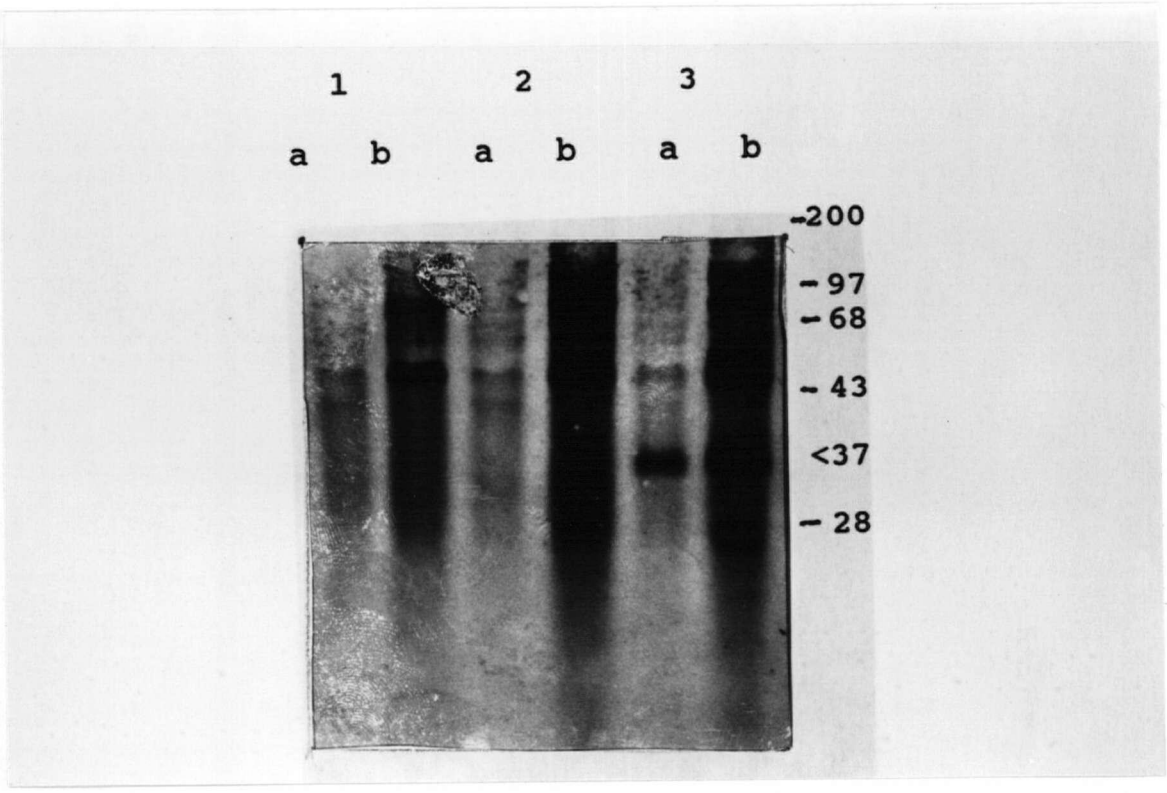


Figure 30: Western Blot Analysis of Immunoprecipitates from Transfected Eukaryotic Cell Cultures

Samples of cell lysates (a) or medium (b) from zinc-induced cell cultures were immunoprecipitated with anti-hTF antiserum. Samples of the precipitated material were resolved by using NaDodSO$_4$-PAGE, transferred to nitrocellulose membrane, and developed consecutively with anti-hTF antiserum and an alkaline phosphatase-conjugated anti-IgG detection system. The pNUT and hTF/2N-pNUT cell lines were selected in 500 μM MTX and all cell culture utilized DMEM medium supplemented with 10% fetal calf serum.

Lane 1: BHK cells
Lane 2: pNUT-transfected BHK cells
Lane 3: pNUT-hTF/2N-transfected BHK cells

The positions of molecular size markers (x10^{-3}Da) are indicated to the right as well as the position of the major immunoprecipitated protein of M_r 37000 (<37).

10-15 μg/mL, however continued culturing of the cells in roller bottles has resulted in levels exceeding 80 μg/mL (R.C. Woodworth and Anne B. Mason, University of

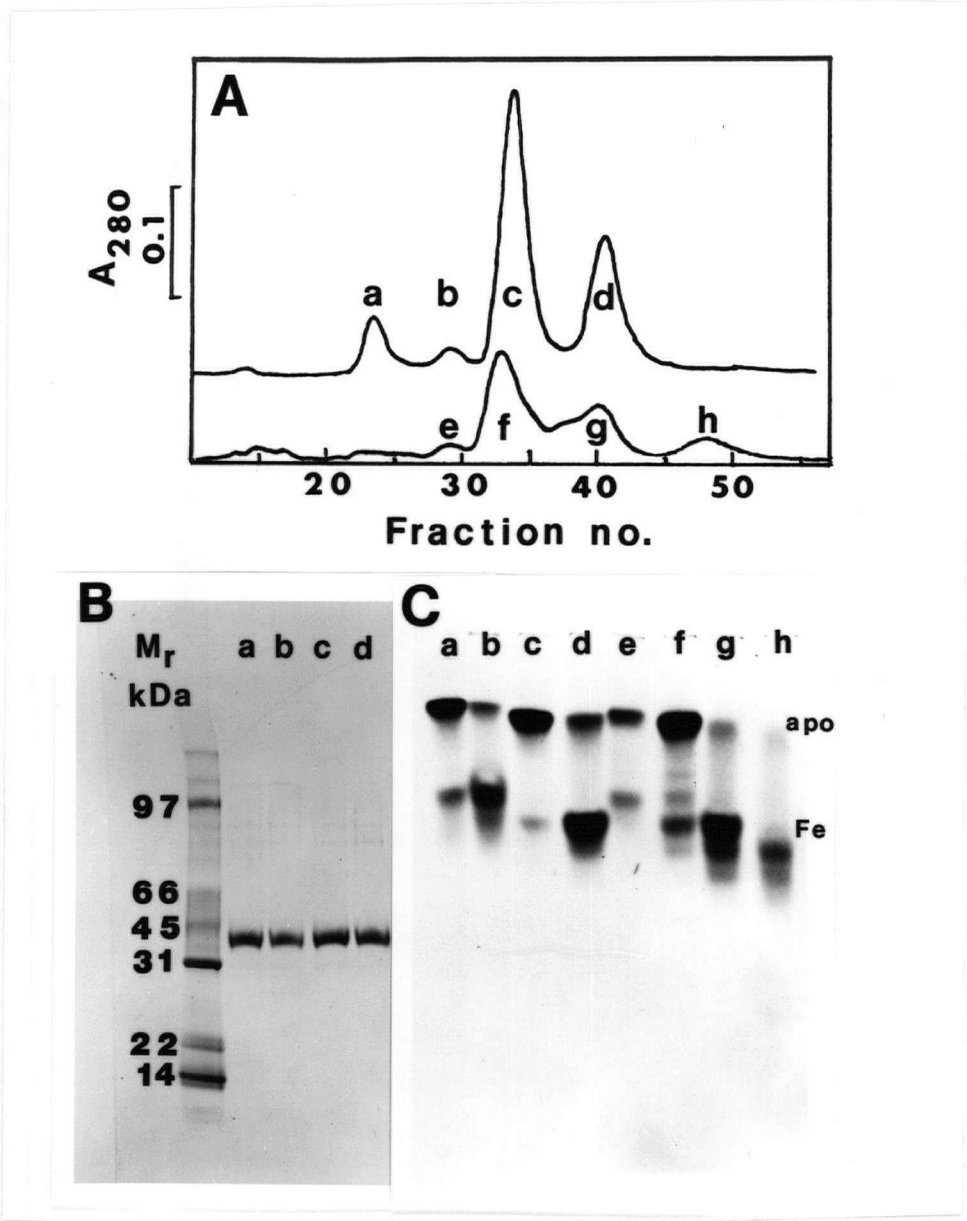


Figure 31: Isolation and Analysis of Recombinant hTF/2N

Panel A: The FPLC (Polyanion SI) elution profiles of recombinant hTF/2N and proteolytically-derived hTF/2N are shown in the upper and lower tracings respectively. Four major peaks are seen for both preparations, indicative of apo and holo forms of two distinct molecular species.

Panel B: The four major components (a-d) from the polyanion SI elution of recombinant hTF/2n seen in Panel A have been analyzed by $NaDodSO_4$-PAGE (5-12% polyacrylamide gradient). The size of molecular size standards is indicated ($x10^{-3}$ kDa).

Panel C: The major components of recombinant hTF/2N (a-d) and proteolytically-derived hTF/2N (e-h) were resolved by urea-PAGE. The relative position of the apo and holo forms of hTF/2N are indicated.

Protein	Amino Acid Sequence	Ref
human serum transferrin	V-P-D-K-T-V-R-W-C-A-V-S-	MacGillivray et al.(1983)
recombinant hTF/2N (major)[a]	V-P-D-K-T-V-R-W-X-A-V-S	Funk et al. (1990a)
recombinant hTF/2N (minor)[b]	V-P-D-K-T-V-	Funk et al. (1990a)

[a] Twelve cycles were analyzed; no residue was identified at cycle 9, however cysteine residues were not modified
[b] Six cycles were analyzed

Table IV: Amino-Terminal Sequence of Human Transferrin and of the Recombinant hTF/2N Half-Molecule

Approximately 200 pmol of the major and minor forms of recombinant hTF/2N were sequenced as described in section IIC6 of Experimental Procedures and Materials.

Vermont, personal communication). The reason for this increase is not clear, though it may conceivably result from an increased cell density in long term cultures as BHK cells are only partially inhibited by cell-cell contact.

Recombinant hTF/2N was purified by a three-step procedure with 80% recovery achieved routinely. The final Polyanion SI elution showed four distinct peaks (a-d, Figure 31, panel A), representing the apo- and holo- forms of both a major (>95%) and minor (<5%) species. NaDodSO$_4$-PAGE analysis of the peaks a-d showed the major and minor forms to be monodisperse and of equal relative molecular size (Figure 31, panel B) and amino terminal sequence analysis showed identical sequences for both forms (Table IV).

The visible absorption spectrum of iron-saturated recombinant hTF/2N and human serum transferrin is shown in Figure 32. Absorption maxima for both human serum TF and recombinant hTF/2N were observed at 465 nm. Spectral ratios ($A_{280/465}$)

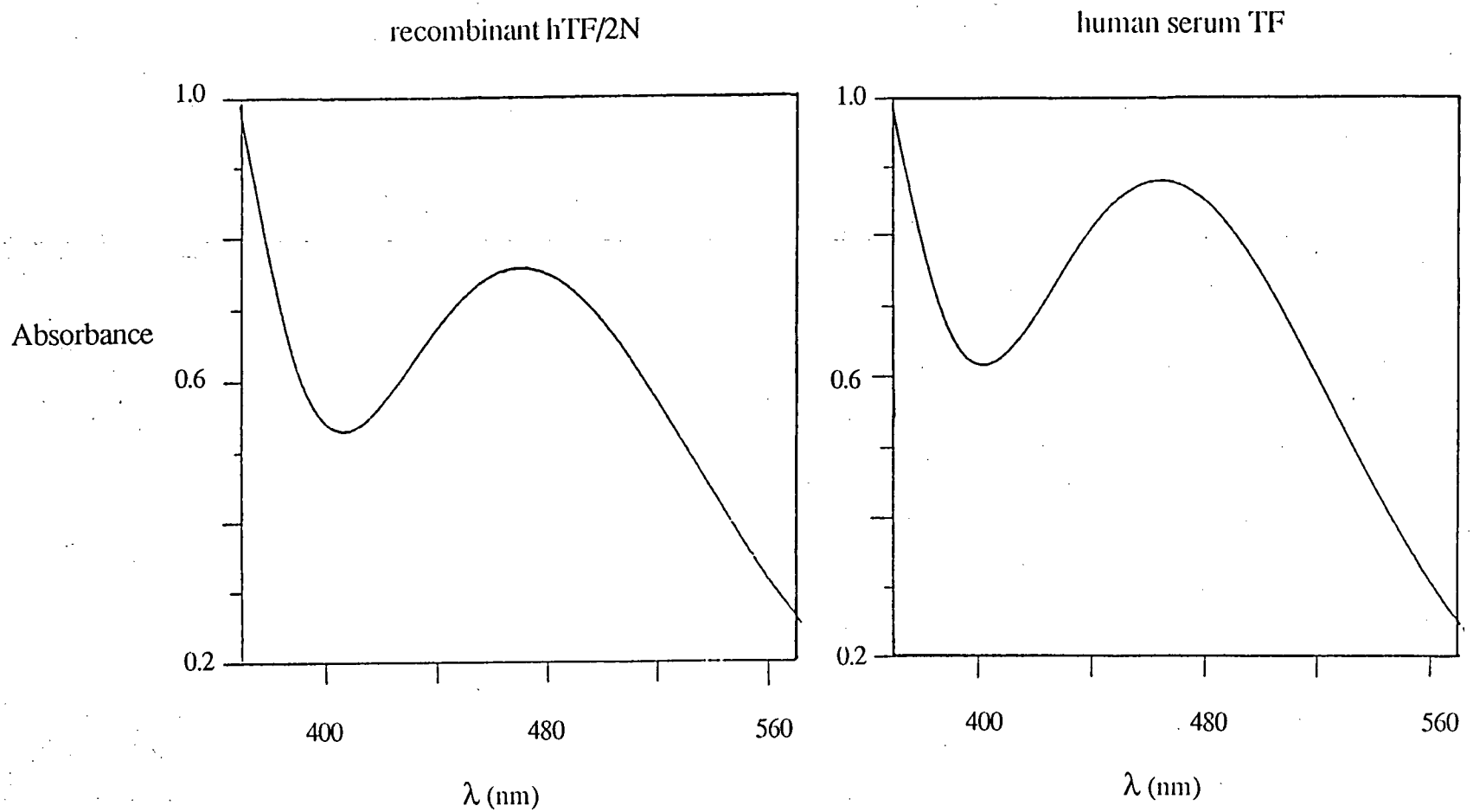

recombinant hTF/2N
human serum TF
Absorbance
1.0
0.6
0.2
400
480
560
λ (nm)

Figure 32: Visible Absorption Spectra of Recombinant hTF/2N and Human Serum Transferrin

Absorption spectra of iron-loaded preparations of recombinant hTF/2N (0.36 mM) and hTF (0.40 mM) were recorded using a Cary 219 spectrophotometer (all protein in 10 mM $NaHCO_3$).

---

for hTF/2N (21) and human serum transferrin (23.5) compared favourably. Titration of 3.68 A_{280} units of apoprotein with $Fe^{III}(NTA)_2$ (Figure 33) showed a slope corresponding to an $E_{465,mM}$ = 2.1, yielding $E_{280,mM}$ = 38.8, both reasonable values for the transferrin half molecule (Lineback-Zins and Brew,1980).

Urea-PAGE analysis of the four peak components (a-d) from FPLC is displayed in Figure 31, panel C. The apo forms of transferrin molecules run with greater mobility than iron-bound forms. The presence of both apo and holoprotein in these fractions is likely due to the limited resolution of the FPLC procedure and the presence of contaminating iron during the subsequent work up. In general, peaks a and c represent the apo form of the minor and major species respectively, while peaks b and d are the corresponding iron-bound forms. In comparison to the proteolytically-derived half molecules, the recombinant molecules showed better monodispersity.

4. *hTF/2C Studies*

The production of a gene construction encoding the C lobe of hTF fused to the natural hTF signal sequence is illustrated in Figure 34. PCR amplification using the described oligonucleotides produced only a single band by agarose gel analysis whose size (1.1 kbp) agreed with the predicted length of the target product. Dideoxy sequence analysis of the termini of this product showed it to indeed code for the desired product, however when this sequence was cloned into the expression vector pNUT and introduced into BHK cells, immunoreactive product was not detected either by immunoprecipitation or radioimmunoassay.

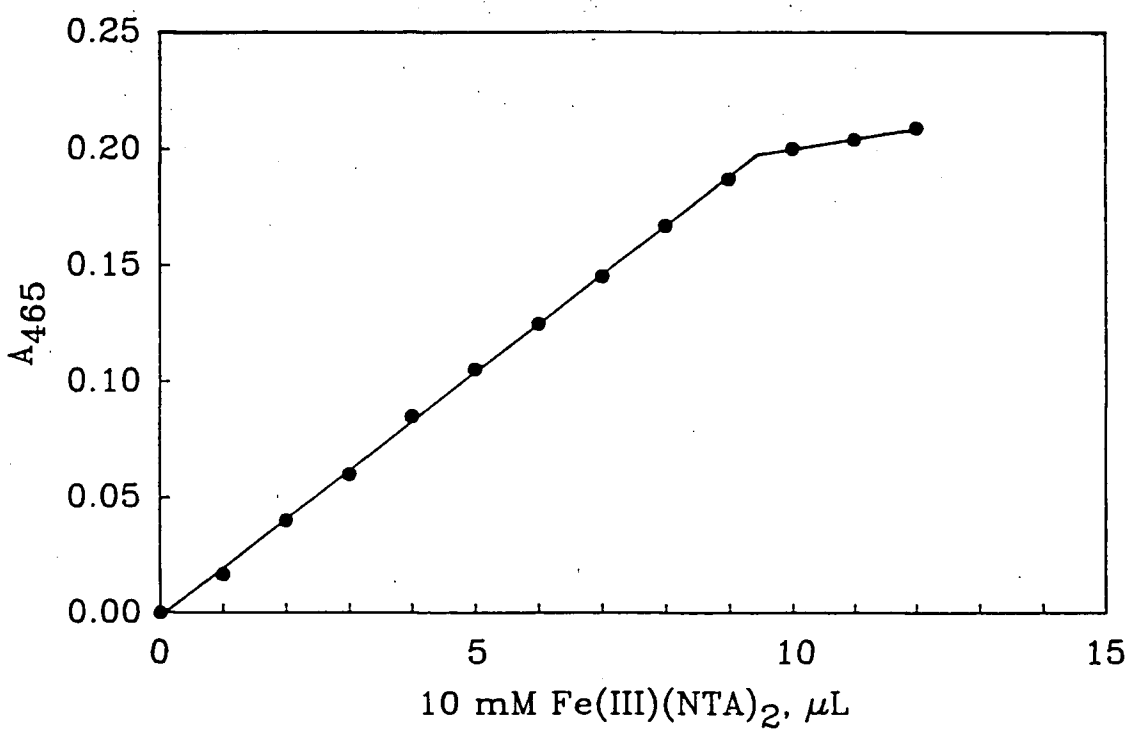
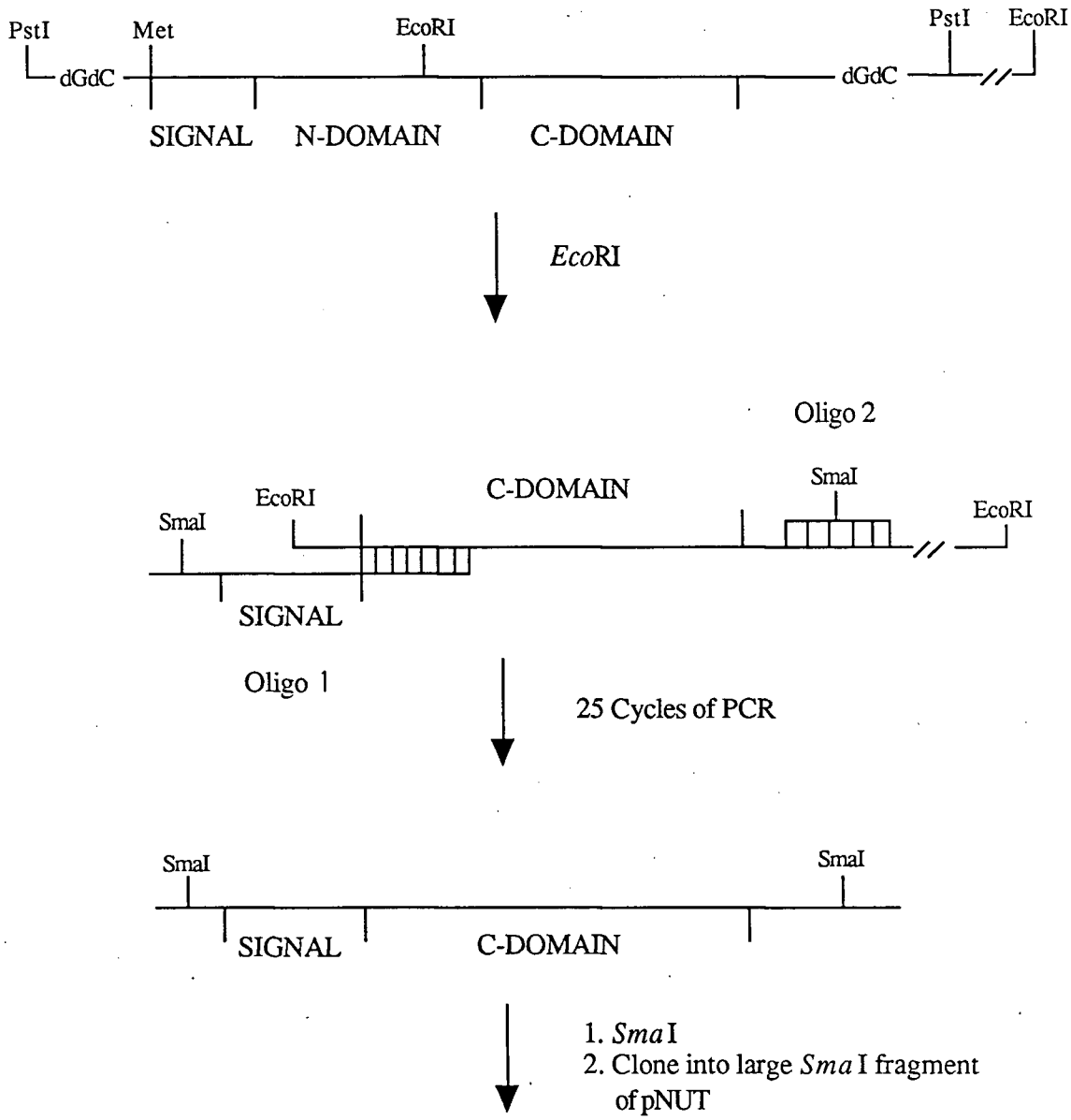


Figure 33: Iron Titration of the Apo form of Recombinant hTF/2N

$Fe^{III}(NTA)_2$ (10 mM) was used to titrate 3.68 A_{280} units of the major form of recombinant hTF/2N in 10 mM $NaHCO_3$ (1 mL volume). Visible spectra were recorded 5-10 minutes after each addition.

To characterize the hTF/2C PCR product fully, sequence analysis was performed using the ExoIII deletion protocol (Heinikoff,1984). Two unintentional errors were found in the coding sequence, resulting in the replacement of Glu507 by Gly and Lys657 by Glu. Sequence analysis of the 5' splice site of the hTF/2C-pNUT construction showed that the initiator methionine had not been lost during the assembly of the vector.


PstI
Met
EcoRI
PstI
EcoRI
dGdC
dGdC
SIGNAL
N-DOMAIN
C-DOMAIN
EcoRI
Oligo 2
SmaI
EcoRI
C-DOMAIN
SmaI
EcoRI
SIGNAL
Oligo 1
25 Cycles of PCR
SmaI
SmaI
SIGNAL
C-DOMAIN
1. SmaI
2. Clone into large SmaI fragment
of pNUT

B

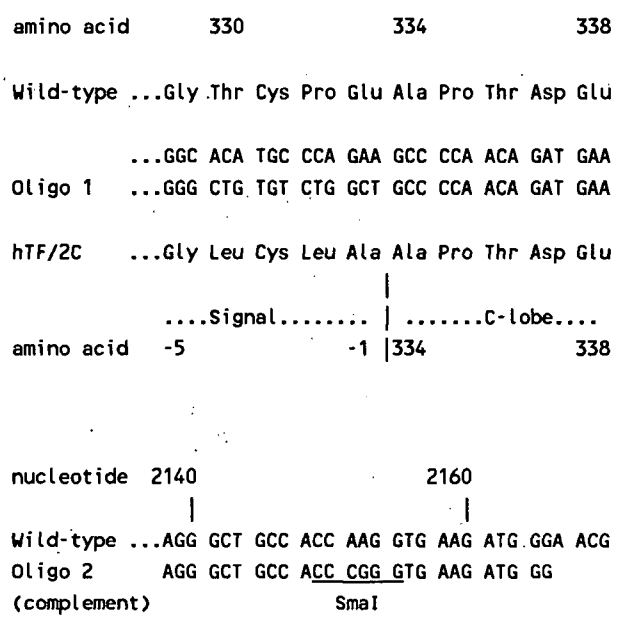


Figure 34: Assembly Strategy for Recombinant hTF/2C Coding Sequence

Panel A: Two separate oligonucleotides were used as PCR primers to create the hTF/2C coding sequence. An *Eco*RI restriction fragment including the entire C lobe coding sequence was used as a template for 25 rounds of PCR amplification. Oliogonucleotide 1 includes a *Sma*I recognition site and the natural hTF signal sequence at its 5'end and matches the coding sequence for amino acids 334 -341 of hTF at its 3' end. Oligonucleotide 2 matches sequence in the 3' nontranslated region of the hTF cDNA and introduces a second *Sma*I recognition sequence at this site.

Panel B: The nucleotide and amino acid substitutions effected by Oligonucleotides 1 and 2 are presented. Oligonucleotide 1 splices the natural signal sequence from the hTF cDNA to the C lobe coding sequence starting at position 334 which is positioned in the linker peptide domain which normally joins the two lobes of hTF. Oligonucleotide 2 introduces a *Sma*I recognition site downstream of the normal translation termination site (nucleotides 2125-2127 of the hTF cDNA).

IV. Discussion

A. Prokaryotic Expression Studies

Bacterial expression systems have been used routinely in studies where the production of large quantities of recombinant protein is desired. As pioneered by Sligar and co-workers in their studies on cytochrome b_5 (von Bodman *et al.*,1986) and myoglobin (Springer and Sligar,1987), the use of synthetic genes has been shown to constitute a powerful alternative to natural cDNA or gene sequences in heterologous expression systems.

Many independent factors must be considered in the design and assembly of prokaryotic expression vectors. Typically these systems are based on high copy number vectors which incorporate strong, inducible transcriptional promoter elements such as P_{lac} of pUC vectors or the hybrid P_{tac} promoter (De Boer *et al.*, 1983) of pKK223-3. The tight coupling of RNA transcription, translation and degradation in bacteria makes it difficult to quantify transcription rates precisely, however strong promoters such as those mentioned above likely provide high levels of transcription of even eukaryotic sequences. Rather, the achievement of high level protein expression hinges on the efficiency of sequence translation and the modification, folding, processing and targeting of the polypeptide chain.

Message translation in prokaryotes begins with the binding of the 16S subunit of the 30S ribosomal subunit to consensus binding sites adjacent to initiator methionine codons (Shine and Delgarno,1975). Consensus sequences covering the ribosome binding site (RBS) and initiator methionine have been compiled (Scherer *et al.*,1980; Stormo *et al.*,1982) and have been summarized as follows:

Rule 1: The typical sequence register is AGGN6-9ATG

Rule 2: G should not occur at position -3 (see Figure 35)

Rule 3: Less than two G residues should occur between positions -1 and -7.

Rule 4: Positions +5 and +10 should be either A or T

A comparison of the sequence regions between the ribosome binding site and the initiator methionine codon for the five prokaryotic expression vectors tested is presented in Figure 35. The vector ΔpKK223-3 hTF/2N violates Rules 1 and 3; however, the same parent vector has been shown to drive expression of human prothrombin (D.R. Cleveland, University of British Columbia, personal communication). Our studies did not allow us to determine quantitatively whether the sequence near the translation start site actually affected the translation efficiency of the recombinant proteins, although the complete absence of antigenically active product from either ΔpKK223-3-hTF/2N or pUC18-hTF/2N suggests that transcripts from these constructions were translated with poor efficiency, or recombinant hTF/2N may simply be highly unstable and is thus quickly degraded. The translation of eukaryotic DNA sequences in bacteria could also be affected by the codon bias of the heterologous gene, though the significance of this factor cannot be readily evaluated.

The complete lack of recombinant protein observed for the human erythrocytic cytochrome b_5 expression plasmid is indicative of the sensitivity of bacterial systems to poorly resolved factors. The amino acid sequence of the core structure of human erythrocytic cytochrome b_5 differs at only three positions compared to the bovine lipase-solubilized hepatic cytochrome b_5, specifically asparagine 17, tyrosine 27 and methionine 70, though the carboxy and amino terminal regions differ substantially (see Figure 3). Although any of these substitutions could readily account for the inability of the protein to fail to incorporate heme, the complete absence of apoprotein in $NaDodSO_4$-PAGE analysis demonstrates that the protein is translated with poor efficiency. Studies by Gheysen *et al.* (1982) examining the expression of the SV40 t-antigen in a bacterial expression system have shown that translation efficiency can vary over 2 orders of magnitude depending upon the nucleotide sequences surrounding the

Vector Construction	-7 -1+1 5 10
pUC19-bovLpb5	AGGA AACAACA ATG AGC AAA GCT GTC
pUC19-humErb5	AGGA AACAACA ATG GCC GAA CAA AGC
pUC19-bovTpb5	AGGA ACCAACA ATG GCT CTC AAA TAC
ΔpKK223-3-hTF/2N	AGGA AAACAGGGGCAT ATG AAA CAA AGC ACT
pUC18-hTF/2N	AGGA AACACAAT ATG AAA CAA AGC ACT

Figure 35: Comparison of the Upstream Sequences from Various Prokaryotic Expression Constructions

Sequences between the ribosome binding site (underlined) and the initiator methionine are shown for the bacterial expression constructions used in this study. The A residue of the initiator methionine is designated +1.

ribosome binding site and initiator methionine. These differences are attributed to hypothetical RNA secondary structures which trap either the ribosome binding site, initiator codon or both in double-stranded regions and presumably affect the translation efficiency of such transcripts. The use of *in vitro* mutagenesis to randomize the sequence surrounding the initiator codon has been demonstrated to increase the translation efficiency of bacterial expression systems substantially.

The use of overlapping oligonucleotides in the cytochrome b_5 gene assembly process provides an economical alternative to the complete synthesis of both DNA strands. The ready availability of oligonucleotides of extended length (100 bases or more) has now made it feasible to assemble genes of considerable size, and the use of enzymatic second strand synthesis almost halves the material costs involved in gene synthesis. The region of overlap need only be sufficient to form a stable hybrid at

ambient temperature. The results of this study clearly demonstrate the need for thorough screening of the target sequences for potential illegitimate priming events.

As illustrated by the present study, the fidelity of the target amino acid sequence is critical to the use of synthetic genes in the production of recombinant proteins. Unfortunately, the available amino acid sequence for bovine cytochrome b_5 was reported from chemical sequencing studies which relied on extensive acid treatment of derivatized peptides, which often results in deamidation of the sample. Had we not had the authentic protein isolated from liver for comparison, there would not have been reason to suspect the behavior of the initial recombinant protein. Only the report of the full-length cDNA sequence for the protein eliminated the necessity of re-sequencing the entire hepatic cytochrome b_5 protein.

The incorporation of heme into the recombinant cytochrome molecules appears to be the limiting factor in determining overall yields of the holoprotein. The S64A substitution in the triple mutant background sequence provides one example of the limited ability of some mutants to incorporate heme. The holo form of this particular mutant could only be recovered following prolonged incubation with exogenous heme, even though NaDodSO$_4$-PAGE analysis showed the same total amount of protein as the high yield wild-type species. Interestingly, the S64A substitution in the corrected wild-type background could be isolated without the need for exogenous hemin, albeit total yields were reduced from that of the wild-type (60 mg for S64A versus 500 mg for w.t.). Similar low-yield mutants have been observed for mutants in which surface histidyl residues were substituted, even though these replacements were not predicted to affect the heme-binding crevice directly (M. Mauk, W. Funk, unpublished work). Nuclear magnetic resonance studies on the apoprotein form of rat liver cytochrome b_5 have demonstrated the presence of a stable core structure which centers on a cluster of residues about Trp 22 (Moore and Lecomte,1990) . Presumably, the apoprotein must maintain its own structural stability prior to the insertion of heme *in vivo* (Shawver *et*

al.,1984). Consequently any substitutions that affect the stability of this conformation would be expected to limit the yields of recombinant protein. The rate limiting reaction in heme synthesis in virtually all organisms studied is δ-aminolevulenic acid (ALA) synthesis (Lascelles *et al.*,1964). Attempts at improving holoprotein yields by supplementing cultures with exogenous ALA were not successful.

B. Analysis of the Q11,E13,D57 Triple-Mutant of Cytochrome b_5

The incorrect sequence of the original recombinant cytochrome may be regarded as a serendipitous mutagenesis experiment. The altered electrochemical and conformational characteristics of the triple mutant have been substantiated by the solution of a 1.9Å resolution crystal structure for the protein (Funk *et al.*,1990b) which shows only slight deviations from the authentic wild-type bovine cytochrome b_5 model (Figure 4).

The substitution of Asn57 by an aspartyl residue brings the γ-carbon of a negatively-charged carboxylate within 11.3Å of the heme iron atom (Figure 36). As demonstrated by the midpoint potential determination of the N57D mutant, this replacement is solely responsible for the 17 mV decrease in reduction potential in the triple mutant. The β-carbon of Asp 57 lies within 3.3Å of the methylene carbon atom of the heme 4-vinyl group. Presumably the carboxylate charge can stabilize the oxidized form of the protein through an electronic inductive effect *via* conjugation of the 4-vinyl group to the π-system of the porphyrin ring. This interpretation may however be oversimplified in that the dominant heme orientation in the triple mutant is present at only 60%, compared to 80% (Mathews,1980) or 90% (Walker *et al.*,1988; Reid *et al.*, 1987) reported for the authentic wild-type protein. As the dominant and alternate isomers display different reduction potentials (calculated as +0.8 mV and -26.2 mV respectively) (Walker *et al*,1988), the observed drop in potential for the Asp57

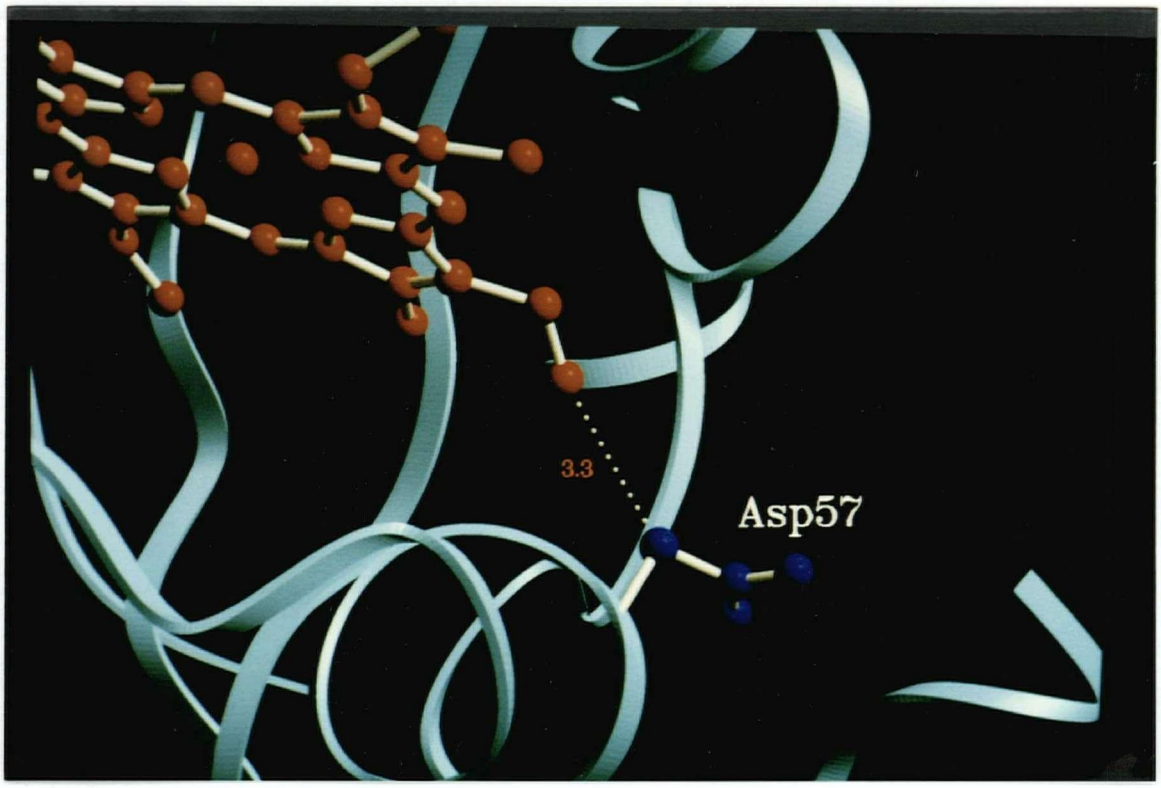


Figure 36: Relative Positioning of Asp57 in the Crystal Structure of Recombinant Lpb$_5$

In the crystal structure of the triple mutant form of recombinant Lpb_5 (Q11,E13,D57), the β-carbon of Asp 57 is located within 3.3Å of the methylene carbon of heme vinyl 4 (dominant heme orientation). The introduction of a negative charge at this position decreases the reduction potential of the heme center by 17 mV, likely due to an inductive electronic effect transmitted to the iron center via heme vinyl 4.

---

replacement may also reflect a difference in conformational preference of the heme. Regardless, the sensitivity of the protein to substitutions at position 57 has been confirmed by studies in which Met, Phe and Cys replacements have been introduced (P. Barker, W.D. Funk, A.G. Mauk, work in progress). Most dramatically, the heme prosthetic group of the N57C substituted protein undergoes a spontaneous conversion to form a chlorin derivative, almost certainly as a result of the participation of the cysteine thiol in the reduction of a heme pyrrole ring.

The replacements at positions 11 and 13 (Q and E respectively) are illustrated in Figure 37. These substitutions do not affect the net charge of the molecule, nor do they affect the reduction potential of the heme center. Perturbations of the hydrogen bonding pattern of this region of the molecule were observed in the crystallographic structure nonetheless. In the authentic wild-type protein, the Glu 11 carboxylate interacts with both Lys 14 and His 15, while substitution of glutamine at this position appears to eliminate such electrostatic interactions. In addition, the Gln 11 side chain forms hydrogen bonds to the side chain hydroxyl and amide nitrogen of Thr 8 which are not observed in the authentic protein. Although the N- and C-termini of the protein are not well ordered in the triple mutant crystal structure, the subtle structural differences at position 11 and 13 may affect the conformation at the termini and thus account for the lowered susceptibility of the triple mutant to the action of trypsin and for the altered FPLC elution characteristics of the Q11,E13 mutant following conversion to the trypsin-solubilized form.

C. S64A Substitution and DME Replacement Studies

The reduction potential of the heme center in a given cytochrome is governed primarily by the nature of the axial ligands. However cytochromes that employ identical ligand sets can vary by more than 300 mV, as seen in the *c*-type family (Lemberg and Barrett,1973). Moore and Williams (1977) have summarized additional factors which govern the potentials of such heme centers as follows:

1. Dielectric constant of the heme pocket
2. Electrostatic charge on the axial ligand; higher negative charge yields a lower redox potential
3. π-acceptor power of the ligand; greater acceptor power yields a higher redox potential
4. Spin-state of the redox interconversion

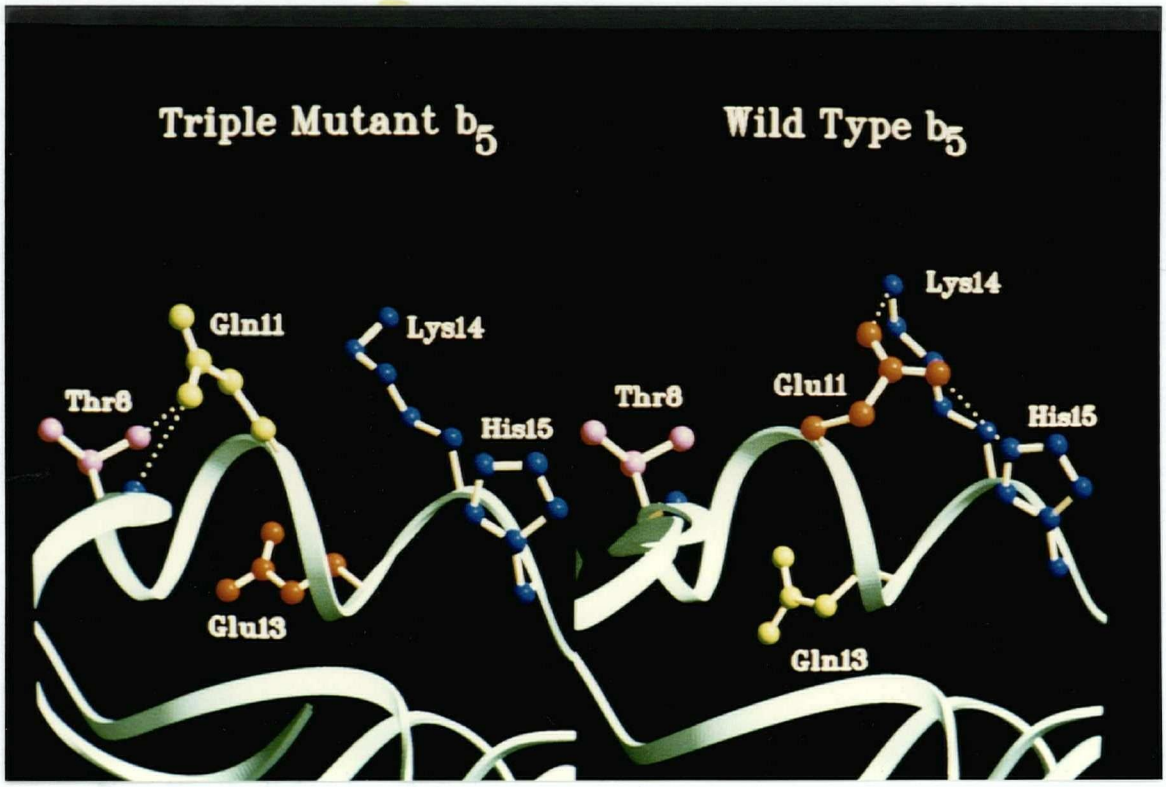


Figure 37: Comparison of the Crystal Structures of Recombinant and Authentic Cytochrome b_5 at Positions 11 and 13

In the triple mutant form of recombinant Lpb_5 (Q11,E13,D57), the side chain of Gln11 forms hydrogen bonds with the side chain hydroxyl of Thr8 and a main chain amide. In the crystal structure of the authentic trypsin-solubilized form of bovine liver microsomal cytochrome b_5 (Mathews *et al.*, 1971b), the carboxylate oxygens of Glu11 form hydrogen bonds with the side chains of Lys14 and His 15.

---

5. Steric factors favoring oxidized *versus* reduced form

6. Special electrostatic interactions

In comparing the variants of recombinant cytochrome b_5, only the first and last factors on this list would likely have bearing, provided of course that the substitution does not alter significantly the architecture of the protein. The dielectric constant of the heme pocket will vary according to the microenvironment of specific residues and must

therefore be considered to be heterogeneous. In general the heme crevice maintains a low dielectric constant and thus electrostatic interactions between the heme and protein can have a substantial affect on the reduction potential, as illustrated by the position 57 substitutions described in this study.

The crystal structure of bovine hepatic cytochrome b_5 clearly shows that heme propionate 7 forms hydrogen bonds with the both the side chain hydroxyl group and main chain amide of Ser 64 and causes the propionate to rotate towards the protein surface (Mathews *et al.*,1971b; Mathews *et al.*,1972) (see Figure 5). Analysis of the crystal structure of the reduced form of the protein, ferrocytochrome b_5, demonstrated that reduction did not significantly displace any of the protein residues, though the binding of a cation by heme propionate 7 was observed in the reduced structure (Argos and Mathews,1975). These authors suggested a model for the red-ox process in which the carboxylate of heme propionate 7 serves to stabilize the formal positive charge of the heme in the oxidized state. Upon reduction, the same carboxylate is freed to bind a cation from solution (see Figure 7). The work of Reid *et al.* (1984) in which substitution of dimethylesterified heme (DME) into the cytochrome raised the reduction potential by 64 mV is consistent with this model. As predicted by the model, removal of the negative charge on the propionates would de-stabilize the ferric form of the protein and consequently raise the reduction potential. Disruption of the hydrogen bonding pattern to heme propionate 7 would also be predicted to raise the reduction potential by destabilizing the carboxylate charge. However the potential of the S64A mutant was actually lower than that of the wild-type by several millivolts, indicating that this substitution slightly increased the stability of the ferric state. These results suggest that the hydrogen bond between heme propionate 7 and the main chain amide is sufficient to stabilize ferricytochrome b_5 and that the removal of the second hydrogen bond by the alanine substitution may allow the carboxylate group to re-orient to a position in closer proximity to the heme center that increases Coulombic stabilization of

the Fe^{III} center. Attempts at crystallizing the S64A mutant protein are currently in progress to test this hypothesis.

The esterification of propionate carboxyls in DME-substituted species results in an identical increase in reduction potential (67 mV) for both the recombinant wild-type and S64A cytochromes. These findings are consistent with the view that the regulation of heme centers by peripheral carboxylates is primarily a function of the carboxylate charge and its relative position with respect to the iron center.

D. Expression of hTF/2N in Eukaryotic Cells

The highly convoluted structure of transferrin and the large number of disulfide bridges in the molecule may have been major impediments to the expression of hTF/2N in bacteria. Our attempt to mimic partially the natural processing of transferrin by targeting the protein for bacterial membrane transport *via* an attached alkaline phosphatase signal sequence was not successful, forcing us to explore eukaryotic expression systems.

The mutant dihydrofolate reductase activity of pNUT represents a significant advantage over other selection systems in that transfected cells can be immediately selected with high levels of MTX, abrogating the need for a dihydrofolate reductase deficient cell line or tedious selection amplification procedures. The relatively short generation time of BHK cells (approximately 12 hours, mid-log phase) and their general hardiness make them ideal for large-scale cultivation.

Several attempts have been made to increase further the product yield of the pNUT-hTF/2N BHK cell line. Attempts to achieve "superinduction" of the metallothionein promoter by using low serum content medium in the presence of actinomycin D (Walden *et al.*,1987) proved unsuccessful, likely due to the decreased viability of the treated cells in low serum (results not shown). Clonal isolation of individual cells from the primary cell line has yielded a line producing almost twice the

level of hTF/2N as judged by radioimmunoassay while serial selection of this same overproducing line to 5 mM MTX failed to show any additional increase in product level (B.K-C. Chow, A.B. Mason, R.C.Woodworth, unpublished results). Most significantly, the continued maintenance of roller cultures over an extended period (up to two months) causes a large increase in the levels of recombinant product, achieving a peak concentration of over 80 mg/L (R.C. Woodworth, A.B. Mason, unpublished results). Product levels of this magnitude allow us to recover sufficient product for experiments such as NMR spectroscopy and crystallization studies which require large amounts of protein.

The recombinant product occurs in both a major and minor form which share identical N-termini and molecular size but differ in their relative elution times on Polyanion SI and in their mobilities in urea-PAGE. Combined, these results suggest the minor form does not result from proteolytic degradation of the recombinant protein, nor does it represent a species with an improperly cleaved signal peptide. The decreased mobility of the minor form on urea-PAGE gels suggests that it maintains a tighter conformation than either the major form or thermolysin-produced hTF/2N molecule. Recombinant hTF/2N recovered from cultures maintained with serum substitutes requires fewer chromatographic steps in the purification process, and the resulting product is composed of less than 1% of the minor species, suggesting strongly that the minor component is derived from the major form during the isolation process (R.C. Woodworth, personal communication).

The spectroscopic and iron-binding behavior of the recombinant hTF/2N match closely those of diferric human serum transferrin and the thermolysin-derived hTF/2N (Lineback-Zins and Brew,1980; Zak *et al.*,1983). Subsequent characterization of the isoelectric point, carbohydrate status and ^{1}H nuclear magnetic resonance spectrum of the recombinant hTF/2N molecule are all in good agreement with results from the proteolytically-derived half molecule (R.C. Woodworth, A.B. Mason,unpublished results;

Funk *et al.*,1990a).

E. hTF/2C Construction

The complete absence of recombinant product from the hTF/2C-pNUT construction again precludes thorough analysis of the system although several factors could conceivably affect the efficiency of expression. This construction shares the same 5' untranslated region, initiator codon region and signal peptide sequence with hTF/2N-pNUT plasmid and thus at least the initiation of translation for the two constructions should, to a first approximation, be similar. The predicted amino acid sequence at the site of signal peptide cleavage for the recombinant hTF/2C compared favourably to consensus rules developed for such cleavage sites in eukaryotic proteins (von Heijne,1983), although the failure of this processing step would account for the absence of antigenically-recognized material. Alternatively, the *de novo* folding of the C-terminal lobe of hTF may require the presence of the N-terminal lobe which would be absent in the recombinant system. In support of this hypothesis, non-covalent interaction between N- and C-terminal half molecules of human serum TF have been shown to stabilize both lobes against denaturation by urea (Ikeda *et al.*,1985).

The presence of two inadvertent substitutions in the nucleotide sequence of the hTF/2C construction were most likely introduced during the PCR amplification procedure as a result of replication errors by the *Taq* polymerase. This is a common observation in PCR amplification procedures and is attributed to the lower fidelity of the thermostable enzyme compared to typical *E. coli* DNA polymerase activities. Conceivably the observed substitutions (E507G and K657E corresponding to GAA-GGA and AAG-AGG codon substitutions respectively) could destabilize the protein and thus prevent the accumulation and detection of recombinant product, although neither of these substitutions involve conserved residues seen in other transferrin species (Anderson *et al.*,1989). The correction of these sequence errors by site-directed mutagenesis or by

re-assembly of other PCR-generated clones will ultimately resolve this possibility and is currently underway. As a second alternative, the expression of a full-length recombinant hTF molecule has now been accomplished (W.D. Funk, D.K. Banfield, R.T.A. MacGillivray, unpublished results). Conceivably, a convenient protease recognition sequence could be engineered within the linker peptide separating the two lobes of the molecule to allow the subsequent recovery of both half molecules.

F. Additional Projects and Work in Progress

The successful expression of recombinant forms of cytochrome b_5 and human transferrin have led to the initiation of several projects designed to probe structural and functional properties of these two molecules. In one ongoing study, surface histidyl residues of bovine cytochrome b_5 (residues 15, 26 and 80) have been substituted in order to produce molecules in which only a single surface histidyl residue remains. This leaves a unique target residue for chemical modifications such as the binding of $Ru(NH_3)_6Cl_3$, and as such eliminates many of the side-reaction products, such as doubly and triply modified proteins, which complicate studies on the wild-type species. These replacements are also valuable for studying electron transfer from the heme iron to bound copper complexes which interact with surface histidyl residues (Reid *et al.*,1987).

Recombinant cytochrome b_5 has also been used in probing molecular interactions with cytochrome c (Burch *et al.*,1988;1990). The availability of complete NMR signal assignments for both cytochrome b_5 (Guiles et al.,1990) and cytochrome *c* (Wand et al.,1989), together with recombinant expression systems for both molecules (Funk *et al.*,1990b-cytochrome b_5: Pielak *et al.*,1985-yeast iso-1-cytochrome *c*), will allow manipulations of either or both binding partners.

Completed substitutions at position 57 of cytochrome b_5 include Met, Phe, and Cys (P.D. Barker, W.D. Funk, A.G. Mauk, unpublished results). The sensitivity of the electrochemical potential to substitutions at this position has been described in this

report and subsequently confirmed by these additional replacements. Initial characterization of the Cys57 variant introduces a thiol group which appears to be directly responsible for the conversion of protoporphyrin IX in the protein to a chlorin-type porphyrin, characterized by the presence of a reduced pyrrole ring. Studies are presently underway to define precisely the role of the thiol at position 57 in this reduction reaction, as initial results suggest the absence of a free thiol in the protein following conversion.

The study of mutant forms of hTF/2N has been initiated with the completion and initial characterization of a D63S variant of the molecule (R.C. Woodworth, W.D. Funk, A.B. Mason, R.T.A. MacGillivray; unpublished results). This substitution mimics a ligand replacement at this site seen in the C lobe of melanoTF and was created to test the relative importance of this ligand in the binding of metals. In addition to directly coordinating the metal, Asp 63 also participates in a hydrogen bonding pattern which links this residue with Domain NII of the molecule and thus may be essential in maintaining the closed conformation of holoTF (E.N. Baker, personal communication). Initial characterization of D63S-hTF/2N shows that it binds iron with a significantly decreased affinity *versus* wild-type and has a yellow-shifted absorption maximum characteristic of the yellow 20 kDa Domain II fragment from duck TF (Evans and Madden,1984). Additional substitution studies currently in progress include a replacement of the same Asp 63 ligand by Cys with the goal of possibly altering the metal-binding specificity of the molecule, and the production of K206Q and E207H variants to examine the role of these residues in stabilizing the bound anion/metal complex as has been suggested from a comparison of the primary sequences of transferrin and lactoferrin (Anderson *et al.*,1989).

V. REFERENCES

Aasa R., Malmstrom B.G., Saltman P. and Vanngard T. (1963) *Biochim. Biophys. Acta* **75**, 203-222.

Abe K., Kimura S., Kizawa R., Anan F.K. and Sugita Y. (1985) *J. Biochem. (Tokyo)* **97**, 1659-1668.

Adrian G.S., Korinek B.W., Bowman B.H. and Yang F. (1986) *Gene* **490**, 167-172.

Ainscough E.W., Brodie A.M., MacLachlan S.J. and Ritchie V.S. (1983) *J. Inorg. Biochem.* **18**, 103-109.

Aisen P., Aasa R., Malmstrom B. and Vanngard T. (1967) *J. Biol. Chem.* **242**, 2484-2493.

Aisen P., Aasa R. and Redfield A. (1969) *J. Biol. Chem.* **244**, 4628-4633.

Aisen P. and Brown E.B. (1975a) *Prog. Hematol.* **9**, 31-53.

Aisen P. and Brown E.B. (1975b) *Prog. Hematol.* **14**, 25-56.

Aisen P., Koenig S.H., Schillinger W.E., Scheinberg D.H., Mann K.G. and Fish W. (1970) *Nature* **226**, 859-861.

Aisen P. and Liebman A. (1972) *Biochim. Biophys. Acta* **257**, 314-319.

Aisen P. and Liebman A. (1973) *Biochim. Biophys. Acta* **304**, 797-804.

Aisen P. and Liebman A. (1978) *In "Transport by Proteins"* (G. Blanes, H. Sund, eds.) Walter de Gruyter, Berlin.

Aisen P., Liebman A., Pinkowitz R.A. and Pollock S. (1973) *Biochemistry* **12**, 3679-3684.

Aisen P., Liebman A. and Sia C.L. (1972) *Biochemistry* **11**, 3461-3464.

Aisen P. and Listowsky I. (1980) *Ann. Rev. Biochem.* **49**, 357-393.

Aldred A.R., Howlett G.J. and Schreiber G. (1984) *Biochim. Biophys. Res. Commun.* **122**, 960-965.

Alderton G., Word W.H. and Fevold H.L. (1946) *Arch. Biochem.* **11**, 9-13.

Alsaadi B.M., Williams R.J.P. and Woodworth R.C. (1987) *J. Inorg. Biochem.* **15**, 1-12.

Anderson B.F., Baker H.M., Norris G.E., Rice D.W. and Baker E.N. (1989) *J. Mol. Biol.* **209**, 711-734.

Anderson B.F., Baker H.M., Dodson E.J., Norris G.E., Rumball S.V., Waters J.M. and Baker E.N. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1769-1773.

Argos P. and Mathews F.S. (1975) *J. Biol. Chem.* **250**, 747-757.

Arnold R.R., Cole M.F. and McGhee J.R. (1977) *Science*, **1917**, 263-266.

Ashwell G. and Morell A.G. (1974) *Adv. Enzymol.* **41**, 99-128.

Atkinson T. and Smith M. (1989) *In "Oligonucleotide Synthesis: A Practical Approach"* (Gait M.J., ed) IRL Oxford.

Azari P.R. and Feeney R.E. (1958) *J. Biol. Chem.* **232**, 293-299.

Azari P.R. and Feeney R.E. (1961) *Arch. Biochim. Biophys.* **92**, 44-52.

Bach J.S., Dixon M. and Zerfas L.G. (1946) *Biochem J.* **40**, 229-239.

Bagby S., Barker P.D., Di Gleria H., Hill A.O. and Lowe V.J. (1988) *Biochem. Soc. Trans.* **16**, 958-959.

Bailey S., Evans R.W., Garratt R.C., Gorinsky B., Hasnain S., Horsburgh C., Jhoti H., Lindley P.F., Mydin A., Sarra R. and Watson J.L. (1988) *Biochemistry* **27**, 5804-5812.

Baker E.N., Rumball S.V. and Anderson B.F. (1987) *TIBS* **12**, 350-353.

Baker E.N., Anderson B.F., Baker H.M., Haridas M., Norris G.E. and Rumball S.V. (1989) X-ray Structural Studies on Human Lactoferrin in Both the Iron- and Apo-Forms. *IXth International Conference on Proteins of Iron Transport and Storage*, Brisbane, Australia.

Bates G.W. and Schlabach M.R.C. (1975) *J. Biol. Chem.* **250**, 2177-2181.

Bezkorovainy A. (1966) *Biochim. Biophys. Acta* **127**, 535-537.

Bezkorovainy A. and Grohlich D. (1967) *Biochim. Biophys. Acta* **310**, 365-375.

Bezkorovainy A., Rafelson M.E. and Likhite V. (1963) *Arch. Biochim. Biophys.* **103**, 371-378.

Bilimoria M.H. and Kamin H.C. (1973) *Ann N.Y. Acad. Sci.* **212**, 428-448.

Blumberg B.S. and Warren L. (1961) *Biochim. Biophys. Acta* **50**, 90-96.

Boulik M.D. (1961) *Clin. Chim. Acta* **6**, 493-497.

Bowman B.H., Yang F. and Adrian G.S. (1988) *Adv. Genetics* **25**, 1-36.

Breathnach D., Benoist C., O'Hare K., Gannon F. and Chambon P. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4853-4860.

Brock J.H. (1985a) *Top. Mol. Struct. Biol.* **7**, 183-262.

Brock J.H. (1985b) *In Metalloproteins (Part 2)* (Harrison P.M., ed.) MacMillan, New York.

Brock J.H., Arzabe F.R., Richardson N.E. and Deverson E.V. (1978) *Biochem J.* **171**, 73-82.

Brosius J., and Holy A. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6929-6936.

Brown J.P., Hewick R.N., Hellstrom I., Hellstrom K.E., Doolittle R.F. and Drezer W.J. (1982) *Nature* **296**, 171-173.

Brown J.P., Nishiyama K., Hellstrom K.E. and Hellstrom I. (1981) *J. Immunol.* **127**, 539-546.

Brown J.P., Wright P.W., Hart C.E., Woodbury R.G., Hellstrom K.E. and Hellstrom I. (1980) *J. Biol. Chem.* **255**, 4980-4983.

Brown-Mason A. and Woodworth R.C. (1984) *J. Biol. Chem.* **259**, 1866-1873.

Brunel F., Ochoa A., Schaeffer E., Boissier F., Guillou Y., Cereghini S., Cohen G.N. and Zakin M.M. (1988) *J. Biol. Chem.* **263**, 10180-10189.

Burch A.M., Rigby S.E.J., Funk W.D., MacGillivray R.T.A., Mauk A.G., Mauk M.R. and Moore G.R. (1988) *Biochem. Soc. Trans.* **16**, 844-845.

Burch A.M., Rigby S.E.J., Funk W.D., MacGillivray R.T.A., Mauk M.R., Mauk A.G. and Moore G.R. (1990) *Science* **247**, 831-833.

Cannon J.C. and Chasteen N.D. (1975) *Biochemistry* **14**, 4573-4577.

Casey J.L., Hentze H.W., Koeller D.M., Caughman S.W., Reuault T.A., Klausner R.D. and Harford J.B. (1988) *Science* **240**, 924-927.

Chaconas G. and van de Sande J.H. (1980) *Methods Enzymol.* **65**, 75-85.

Chance B. and Williams G.R. (1954) *J. Biol. Chem.* **248**, 4947-4955.

Charlwood P.A. (1963) *Biochem. J.* **88**, 394-398.

Chasteen N.D. (1977) *Coord. Chem. Rev.* **22**, 1-36.

Chasteen N.D. (1983) *Trend Biochem. Sci.* (Pers. ed.) **8**, 272-275.

Ciehanover A., Schwartz A.L. and Lodish H.F. (1983) *Cell* **32**, 267-276.

Cochet M., Gannon F., Hen R., Maroteaux E., Perrin F. and Chambon P. (1979) *Nature* **282**, 567-574.

Cohen H.J. and Fridovich I. (1971) *J. Biol. Chem.* **246**, 359-366.

Cohn E.J. (1947) *Experimentia* **3**, 125-136.

Cohn E.J., Strong L.E., Hughes W.L., Mulford D.J., Ashworth J.N., Melin M. and Taylor H.L. (1946) *J. Am. Chem. Soc.* **68**, 459-475.

Cristiano R.J. and Steggles A.W. (1989) *Nucleic Acids Res.* **17**, 799.

Daruish N., Fischer C.W. and Steggles A.W. (1988) *Prot. Seg. Data Analysis* **1**, 351-353.

DeBoer H.A., Comstock L.J. and Vasser M. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 21-25.

Diamond A., Devine J.M. and Cooper G.M. (1984) *Science* **225**, 516-519.

Doolittle R.F. (1984) Evolution of the Vertebrate Plasma Proteins. *In "The Plasma Proteins: Structure, Function and Genetic Control"* (F.W. Putnam, ed.) 2nd Ed., **Vol. 4**, pp 317-359. Academic Press, New York.

Doolittle R.F. and Feng Y. (1987) *C.S.H. Symp. Quant. Biol.* **52**, 869-874.

Dorland L., Haverkamp J., Schut B.L., Vliegenthart J.F.G., Spik G., Strecker G., Fournet B. and Montreuil J. (1977) *FEBS Lett.* **77**, 15-20.

Douglas R.H. and Hultquist D.E. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3118-3122.

Druyan R., DeBernhard B. and Robinowitz M.C. (1968) *J. Biol. Chem.* **244**, 5874-5878.

Dutton P.L. (1978) *Methods Enzymol.* **54**, 411-435.

Egami F., Hahashi M., Sato R. and Mori T. (1953) *J. Biochem. (Tokyo)* **40**, 527-534.

Ehrenberg A. and Laurell C-.B. (1955) *Acta Chem. Sci.* **9**, 68-72.

Enoch H.G. and Strittmatter P. (1979) *J. Biol. Chem.* **254**, 6483-6488.

Evans R.W. and Madden A.D. (1984) *Biochem. Soc. Trans.* **12**, 661-662.

Evans R.W. and Williams J.C. (1978) *Biochem J.* **173**, 543-552.

Evans R.W., Williams J. and Moreton K. (1982) *Biochem J.* **210**, 19-26.

Fleischer S.F., Fleischer B., Azzi A. and Chance B. (1971) *Biochim. Biophys. Acta* **225**, 194-200.

Fleming P.J., Dailey H.A., Corcoran D. and Strittmatter P. (1978) *J. Biol. Chem.* **253**, 5369-5372.

Fontes G. and Thivolle L. (1925) *Compt. Rend. Soc. Biol.* **93**, 687-689.

Fukushima K. and Sato R.C. (1972) *J. Biochem. (Tokyo)* **71**, 447-461.

Funk W.D. and MacGillivray R.T.A. (1990) Molecular Biology of Transferrin *In "Proteins of Iron Transport and Storage"* (P. Ponka, R.C. Woodworth, eds.) C.R.C. Press Reviews, Boca Raton, Fl.

Funk W.D., MacGillivray R.T.A., Mason A.B., Brown S.A. and Woodworth R.C. (1990a) *Biochemistry* **29**, 1654-1660.

Funk W.D., Lo T.P., Mauk M.R., Brayer G.D., MacGillivray R.T.A. and Mauk A.G. (1990b) *Biochemistry* **29**, 5500-5508.

Gaber B.P. and Aisen P. (1970) *Biochim. Biophys. Acta* **221**, 228-233.

Gatermann K.B., Rosenberg G.H. and Kaufer N.F. (1988) *Biotechniques* **6**, 63-64.

Galfand D.H. and White T.J. (1990) *In "PCR Protocols: A Guide to Methods and Applications"* (Innis M.A., Gelfand D.H. and White T.J., eds.) Academic Press, San Diego.

George S.K., Najera L., Sandoval R.P., Countryman C., Davis R.W. and Ihler G.M. (1989) *J. Bacteriol* **171**, 4569-4576.

Gheysen D., Iserentent D., Derom C. and Fiers, W. (1982) *Gene* **17**, 55-63.

Giblett E.R. (1968) *Bibl. Haematol. (Basel)* **29**, 10-20.

Giblett E.R., Hickman C.G. and Smithies O. (1959) *Nature* **183**, 1589-1560.

Gilbert W. (1978) *Nature* **271**, 50.

Gitlow S.E. and Beyers M.R. (1952) *J. Lab. Clin. Med.* **39**, 337-346.

Goodfellow P.N., Banting G., Sutherland R., Greaves M., Soloman E. and Povey S. (1982) *Somat. Cell Genet.* **8**, 197-193.

Gorinsky B., Horsburgh C., Lindley P.F., Moss D.S., Parkar M. and Watson J.L. (1979) *Nature* **281**, 157-158.

Goubin G., Goldman G.S., Luce J., Weiman D.E. and Cooper G.M. (1983) *Nature* **302**, 114-119.

Graham I. and Williams J. (1975) *Biochem J.* **145**, 263-279.

Guiard B. and Lederer F. (1979) *J. Molec. Biol.* **135**, 639-650.

Guiles R.D., Altman J., Kuntz I.D. and Waskell L. (1990) *Biochemistry* **29**, 1276-1289.

Harding C., Heuser J. and Stahl P. (1983) *J. Cell Biol.* **97**, 329-335.

Harris W.R. (1983) *Biochemistry* **22**, 3920-3926.

Hebbert D. and Morgan E.H. (1985) *Blood* **65**, 758-764.

Hegesh E., Hegesh J. and Kaftory A. (1986) *New England J. Med.* **314**, 757-761.

Heimburger W., Heide K., Haupt H. and Schulze H.E. (1964) *Clin. Chim. Acta* **10**, 293-307.

Heinikoff S. (1984) *Gene* **28**, 351-359.

Holloway P.W. and Katz J.T. (1972) *Biochemistry* **11**, 3689-3696.

Holmberg C.G. and Laurell C-B., (1945b) *Acta Physiol. Scand.* **10**, 307-312.

Holmberg C.G. and Laurell C-B., (1947) *Acta Chem. Scand.* **1**, 944-950.

Hopkins C.R. (1983) *Cell* **35**, 321-328.

Hopkins L.L. and Schwartz K. (1964) *Biochim. Biophys. Acta* **90**, 484-491.

Hopkins L.L. and Trowbridge I.S. (1983) *J. Cell Biol.* **97**, 508-511.

Huntley T.E. and Strittmatter P. (1972) *J. Biol. Chem.* **247**, 4641-4647.

Iacopetta B.J. and Morgan E.H. (1983) *J. Biol. Chem.* **258**, 9108-9114.

Iacopetta B.J., Rothenberger S. and Kuhn L.C. (1988) *Cell* **54**, 485-495.

Ichikawa Y. and Yamano T. (1970) *Biochim. Biophys. Res. Commun.* **40**, 297-305.

Ikeda H., Nabuchi Y., Nakazato K., Tanaka Y. and Satake K. (1985) *FEBS Lett.* **182**, 305-310.

Imai Y. and Sato R.C. (1977) *Biochim. Biophys. Res. Commun.* **75**, 420-426.

Inman J.K., Coryell F.C., McCall K.B., Sgouris J.T. and Anderson H.D. (1961) *Vox. Sang.* **6**, 34-52.

Ito A. (1980) *J. Biochem (Tokyo)* **87**, 63-71.

Ito A. and Sato R. (1968) *J. Biol. Chem.* **243**, 4922-4923.

Jacq C. and Lederer F. (1974) *Eur J. Biochem.* **41**, 311-318.

Jamieson G.A., Jett M. and DeBernardo S. (1971) *J. Biol. Chem.* **246**, 3686-3694.

Jansson I. and Schenkman J.B. (1977) *Archs. Biochim. Biophys.* **178**, 89-107.

Jeltsch J.M. and Chambon P. (1981) *Eur. J. Biochem.* **120**, 291-295.

Johansson B. (1958) *Nature* **181**, 996-997.

Johansson B. (1960) *Acta Chem. Scand.* **14**, 510-512.

Jarnum S. and Lassen N.A. (1961) *Scand. J. Clin. Lab. Invest.* **13**, 357-359.

Jolles J., Mazurier J., Metz-Boutigue M.H., Spik G., Montreuil J. and Jolles P. (1976) *FEBS Lett.* **69**, 27-31.

Jones K.A., Kadonaga T.T., Rosenfeld P.J., Kelly T.J. and Tijan R. (1983) *Cell* **48**, 79-89.

Kadlubar F.F. and Ziegler D.M. (1974) *Arch. Biochim. Biophys.* **169**, 83-92.

Kajihara T. and Hagihara B. (1968) *J. Biochem (Tokyo)* **63**, 453-461.

Karlsson K-A., Pascher I., Samuelsson B.E., Finne J., Krusivs T. and Rauvala H. (1978) *FEBS Lett.* **94**, 413-417.

Keilin D. (1925) *Proc. R. Soc. Lond.* **B98**, 312-339.

Keilin D. and Hartree E.F. (1940) *Proc. R. Soc. Lond.* **B129**, 277-306.

Keller R.M. and Wuthrich K. (1980) *Biochim. Biophys Acta* **621**, 204-210.

Keller W. and Pennell R.B. (1959) *J. Lab Clin. Med.* **53**, 638-645.

Keyes S.R., Alfano J.A., Jannson I. and Cinti D.L. (1979) *J. Biol. Chem.* **254**, 7778-7784.

Kimura S., Abe K. and Sugita Y. (1984) *FEBS Lett.* **169**, 143-146.

Koechlin B.A. (1952) *J. Am. Chem. Soc.* **74**, 2649-2653.

Kuhn L.C., Ponka P. and Schulman H.M. (1990) Iron-Transferrin Requirements and Transferrin Receptor Expression in Proliferating Cells *In "Proteins of Iron Transport and Storage"* (P. Ponka and R.C. Woodworth, eds.) C.R.C. Press Reviews, Boca Raton, Fl.

Kuhn L.C., McClelland A. and Ruddle F.H. (1984) *Cell* **37**, 95-103.

Kunkel T.A. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488-492.

Laemmli N.K. (1970) *Nature* **227**, 680-685.

LaMar G.N., Budd D.L., Vislio D.B., Smith K.M. and Langry K. (1978) *Proc. Natl Acad. Sci. U.S.A.* **75**, 5755-5762.

LaMar G.N., Burns P.D., Jackson J.T., Smith K.M., Langry K.C. and Strittmatter P. (1981) *J. Biol. Chem.* **256**, 6075-6083.

Lascelles J. (1964) *Tetrapyrole biosynthesis and regulation.* W.A. Benjamin, New York.

Laurell C.B. (1960) Metal Binding Plasma Proteins *In "The Plasma Proteins"* (F.W. Putnam, ed.) p. 349. Academic Press, New York.

Le K.H.D. and Lederer F. (1983) *EMBO J.* **2**, 1909-1914.

Lederer F., Ghrir R., Guiard B., Cortial S. and Ito A. (1983) *Eur. J. Biochem.* **132**, 95-102.

Lee D.C., McKnight G.S. and Palmiter R.D. (1978) *J. Biol. Chem.* **253**, 3494-3503.

Leger D., Tordera V., Spik G., Dorland L., Haverkamp J. and Vleigenthart J.F.G. (1978) *FEBS Lett.* **93**, 255-259.

Legrand D., Mazurier J., Colavizza D., Montreuil J. and Spik G. (1990) *Biochem. J.* **266**, 575-581.

Legrand D., Mazurier J., Metz-Boutigue M.H., Jolles J., Jolles P., Montreuil J. and Spik G. (1984) *Biochim. Biophys. Acta* **787**, 90-96.

Lehninger A.L. (1970) *Biochemistry*, Worth Publishers, Inc., New York.

Lemberg R. and Barrett J. (1973) *"Cytochromes"*, Academic Press, New York.

Lestas A.N. (1976) *Br. J. Haematol* **32**, 341-350.

Leto T.L., Roseman M.A. and Holloway P. (1980) *Biochemistry* **19**, 1911-1916.

Lineback-Zins J. and Brew K. (1980) *J. Biol. Chem.* **255**, 708-714.

Livingston D.J., McLachlan S.J., LaMar G.N. and Brown W.D. (1985) *J. Biol. Chem.* **260**, 15699-15706.

Lu A.Y.H., West S.B., Vore M., Ryan D. and Levin W. (1974) *J. Biol. Chem.* **249**, 6701-6709.

Lucero M.A., Schaeffer E., Cohen G.N. and Zakin M.M. (1986) *Nucleic Acids Res.* **14**, 8692.

Luk C-K., (1971) *Biochemistry* **10**, 2838-2843.

Lum J.B., Infante A.J., Makker D.M., Yang F. and Bowman B.H. (1986) *J. Clin. Invest.* **77**, 841-849.

MacGillivray R.T.A. (1975) Ph.D. Dissertation, University of Miami.

MacGillivray R.T.A. and Brew K. (1975) *Science* **190**, 1306-1309.

MacGillivray R.T.A., Mendez E. and Brew K. (1977) *In "Proteins of Iron Metabolism"* (E.B. Brown, P. Aisen, J. Fielding, R.R. Crichton, eds.) Grove and Stratton, New York.

MacGillivray R.T.A., Mendez E., Shewale J.G., Sinha S.K., Lineback-Zins J. and Brew K. (1983) *J. Biol. Chem.* **258**, 3543-3553.

Makey D.G. and Seal U.S. (1976) *Biochim. Biophys. Acta* **453**, 250-256.

Maniatis F., Fritsch E.F. and Sambrook J. (1982) *"Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Mann K.G., Fish W., Cox A. and Tanford C. (1970) *Biochemistry* **9**, 1348-1353.

Markley J.L. (1975) *Acc. Chem. Res.* **8**, 70-75.

Marks C.B., Vasser M., Ng P., Henzen W. and Anderson S. (1986) *J. Biol. Chem.* **261**, 7115-7118.

Marsh J.B. and James A.T. (1962) *Biochim. Biophys. Acta* **60**, 320-328.

Martin E.M. and Morton R.K. (1955) *Nature* **176**, 113-114.

Martin A.W., Huebers E., Webb J. and Finch C.A. (1984) *Blood* **64**, 1047-1052.

Maruyama T., Gojobor A.S. and Ikemura T. (1986) *Nucleic Acids Res.* **14**, r151-r197.

Masson P.L. and Heremans J.F. (1966) *Protides Biol. Fluids* **14**, 115-121.

Masson P.L., Heremans J.F. and Schonne E. (1969) *J. Exp. Med.* **130**, 643-649.

Masson P.L., Carbonara A.O. and Heremans J.F. (1965) *Biochim. Biophys. Acta* **107**, 485-500.

Masson P.L., Heremans J.F. and Dive C. (1966) *Clin. Chem. Acta* **14**, 735-739.

Mathews F.S. (1980) *Biochim. Biophys. Acta* **662**, 375-379.

Mathews F.S. (1985) *Prog. Biophys. Molec. Biol.* **45**, 1-56.

Mathews F.S. and Czerwinski E.W. (1976) Cytochrome b_5 and Cytochrome b_5 Reductase from a Chemical and X-Ray Diffraction Viewpoint *In "The Enzymes of Biological Membranes"*, **Vol. 4** (A. Martinosi ed.) Plenum Publishing, New York.

Mathews F.S., Argos P. and Levine M. (1971a) *Cold Spring Harbor Symp. Quant. Biol.* **36**, 387-394.

Mathews F.S., Levine M. and Argos P. (1971b) *Nature* **223**, 1519.

Mathews F.S., Levine M. and Argos P. (1972) *J. Mol. Biol.* **64**, 449-458.

Mauk M.R., Reid L.S. and Mauk A.G. (1982) *Biochemistry* **21**, 1843-1851.

McClelland A., Kuhn L.C. and Rodde F.H. (1984) *Cell* **39**, 267-275.

McKnight G.S., Hammer K.E., Kuenzel E.A. and Brinster R.L. (1983) *Cell* **34**, 335-341.

McKnight G.S., Lee D.C., Hemmaplardh D., Finch C.A. and Palmiter R.D. (1980a) *J. Biol. Chem.* **255**, 144-147.

McKnight G.S., Lee D.C. and Palmiter R.D. (1980b) *J. Biol. Chem.* **255**, 148-152.

McKnight G.S. and Palmiter R.D. (1979) *J. Biol. Chem.* **254**, 9050-9058.

Meek J. and Adamson E.D. (1985) *J. Embryol Exp. Morphol* **86**, 205-212.

Messing J. (1983) *Meth. Enzymol.* **101**, 20-28.

Metz-Boutigue M.H., Jolles J., Mazurier J., Schoentgen R., Legrade D., Spik G., Montreuil J. and Jolles P. (1984) *Eur. J. Biochem.* **145**, 659-676.

Meyers E.W. and Miller W. (1988) *Cabios* **4**, 11-17.

Montreuil J., Tonnelat J. and Mullet S. (1960) *Biochim. Biophys. Acta.* **45**, 413-421.

Moore C.D. and Lecomte J.T.J. (1990) *Biochemistry* **29**, 1984-1989.

Moore G.R. and Williams R.J.P. (1977) *FEBS Lett.* **79**, 229-232.

Morgan E.H. (1983) Synthesis and Secretion of Transferrin *in "Plasma Protein Secretions by the Liver"* (Glaumann H., Peters T. Jr. and Ledman C., eds.) Academic Press, London.

Morgan E.H., Marsagalia G., Giblett E.K. and Finch C.A. (1967) *J. Lab. Clin. Med.* **69**, 370-381.

Mullner E.W., Neupert B. and Kuhn L.C. (1989) *Cell* **58**, 373-381.

Nelson R.M. and Long G.L. (1989) *Analyt. Biochem.* **180**, 147-151.

Newman R., Scheider C., Sutherland R., Vodinelich C. and Greaves M. (1982) *Trends Biochem. Sci.* **8**, 397-1982.

Nisimoto Y., Takevchi F. and Shibata Y. (1977) *J. Biochem. (Tokyo)* **82**, 1257-1266.

Oka T. and Schimke R.T. (1969) *J. Biol. Chem.* **43**, 123-127,

Okada Y., Frey A., Guenther T., Vesch F., Sabitini D.D. and Kreibich G. (1982) *Eur. J. Biochem* **122**, 393-402.

Okada Y., Frey A., Guenther T., Vesch F., Sabitini D.D. and Kreibich G. (1982) *Eur. J. Biochem* **122**, 393-402.

Omura T., Siekevitz P. and Palade G.E. (1967) *J. Biol. Chem.* **242**, 2389-2396.

Oncley J.L., Scatchard G. and Brown A. (1947) *J. Phys. Colloid Chem.* **51**, 184-198.

Oshino N. (1982) Cytochrome b5 and its physiological significance *In "Int. Encyclopedia of Pharm. and Ther."* (Schenkman T. and Kupfer D. eds.), Pergamon Press.

Oshino N. (1978) *Pharmac. Ther. A.* **2**, 477-515.

Ozols J. (1989) *Biochim. Biophys. Acta* **947**, 121-130.

Ozols J. and Heinemann F.S. (1977) *J. Biol. Chem.* **252**, 8549-8553.

Ozols J. and Heinemann F.S. (1982) *Biochim. Biophys. Acta* **704**, 163-173.

Ozols J. and Strittmatter P. (1969) *J. Biol. Chem.* **244**, 6617-6618.

Palmiter R.D., Norstedt G., Gelings R.E., Hammer R.E. and Brinster R.L. (1983) *Science* **222**, 809-814.

Palmiter R.D., Behringer R.R., Quaife C.J., Maxwell F., Maxwell I.H. and Brinster R.L. (1987) *Cell* **50**, 435-443.

Palmour R.M. and Sutton H.E. (1971) *Biochemistry* **10**, 4020-4032.

Paltauf F. and Holasek A. (1973) *J. Biol. Chem.* **248**, 1609-1615.

Pappenheimer A.M. Jr. and Williams C.M. (1954) *J. Biol. Chem.* **209**, 915-929.

Park I., Schaeffer E., Sidoli A., Baralle F.E., Cohen G.N. and Zukin M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3149-3157.

Parker W.C. and Bearn A.G. (1962) *J. Exptl. Med.* **115**, 83-105.

Passon P.G., Reed D.W. and Hultquist D.E. (1972) *Biochim. Biophys. Acta* **275**, 51-61.

Phillips M.E. and Thorbecke G.J. (1966) *Int. Arch. Allergy Appl. Immunol.* **29**, 553-557.

Pielak G.J., Mauk A.G. and Smith M. (1985) *Nature* **313**, 152-154.

Pinkowitz R.A. and Aisen P. (1972) *J. Biol. Chem.* **237**, 7830-7834.

Pladziewicz J.R., Meyer T.J., Broomhead J.A. and Taubelt I. (1973) *Inorg. Chem.* **12**, 639.

Plowman G.D., Brown J.P., Enns. C.A., Schroder J., Nikinana B., Sussman H.H. and Hellstrom I. (1983) *Nature* **305**, 70-72.

Polis B.D. and Schmukler W. (1953) *J. Biol. Chem.* **201**, 475-500.

Poulik M.D. (1957) *Nature* **180**, 1477-1479.

Poulos T.L. and Mauk A.G. (1983) *J. Biol. Chem.* **258**, 7369-7373.

Price E.M. and Gibbons J. (1972) *Biochim. Biophys. Res. Comm.* **46**, 646-651.

Princiotto J.V. and Zapolski E.J. (1975) *Nature* **255**, 87-88.

Prochownik E.V., Markham A.F. and Orkin S.H. (1983) *J. Biol. Chem.* **258**, 8389-8394.

Pugh E.L. and Kutes M. (1977) *J. Biol. Chem.* **252**, 68-73.

Putnam F.W. (1975) Transferrin *In "The Plasma Proteins"* (F.W. Putnam, eds.) **Vol. 1**, 2nd Ed., Academic Press, New York.

Rachubinski D., Verman P.S. and Bergeron, J.J.M. (1980) *J. Cell Biol.* **84**, 705-716.

Rado T.A., Wei X. and Benz, E.J. (1987) *Blood* **70**, 989-993.

Raw I. and Mahler, H.R. (1959) *J. Biol. Chem* **234**, 1867-1873.

Reddy V.V.R., Kupfer D. and Caspi E. (1977) *J. Biol. Chem.* **252**, 2797-2801.

Reid L.S. (1984) Ph.D. Dissertation, University of British Columbia.

Reid L.S., Mauk M.R. and Mauk A.G. (1984) *J. Am. Chem. Soc.* **106**, 2182-2185.

Reid L.S., Taniguchi V.T., Gray H.B. and Mauk A.G. (1982) *J. Am. Chem. Soc.* **104**, 7516-7519.

Reid L.S., Gray H.B., Dalvit C., Wright P.E. and Saltman P. (1987) *Biochemistry* **26**, 7102-7107.

Remacle J., Fowler S., Amar-Costesec A. and Berthet J. (1976) *J. Cell Biol.* **71**, 551-564.

Rochard E., Legrand D., Mazurier J., Montruil J. and Spik G. (1989) *FEBS Lett.* **255**, 201-204.

Rodgers K.K., Pachapsky T.C. and Sligar S.G. (1988) *Science* **240**, 1657-1659.

Rogers J. (1985) *Nature* **315**, 458-459.

Rogers M.J. and Strittmatter P.C. (1973) *J. Biol. Chem.* **248**, 800-806.

Rogers M.J. and Strittmatter P.C. (1975) *J.* Biol. Chem. **250**, 5713-5718.

Rogers T.B., Feeney R.E. and Meares C.F. (1977) *J. Biol. Chem.* **252**, 8108-8112.

Rogers T.B., Boerrsen T. and Feeney R.E. (1978) *Biochem. J.* **17**, 1105-1109.

Roop W.E. and Putnam F.W. (1967) *J. Biol. Chem.* **242**, 2507-2513.

Rose T.M., Plowman G.D., Teplow D.B., Dreyer W.J., Hellstrom K.E. and Brown J.P. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1261-1265.

Rosseneu-Moutreff M.Y., Soetewey F., Lamote K. and Peeters H. (1971) *Biopolymers* **10**, 1039-1048.

Rossman M.G. and Argos P.C. (1975) *J. Biol. Chem.* **250**, 7525-7533.

Rouault T.A., Hentze M.W., Caughman S.W., Harford J.B. and Klausner R.D. (1988) *Science* **241**, 1207-1211.

Salemme F.R. (1976) *J. Mol. Biol.* **102**, 563-568.

Sanger F., Nicklen S. and Koulson A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467.

Schade A.L. and Caroline L. (1944) *Science* **100**, 14-15.

Schade A.L. and Caroline L. (1946) *Science* **104**, 340-341.

Schade A.L., Reinhard R. and Levy H. (1949) *Arch. Biochem.* **20**, 170-172.

Schaeffer E., Lucero M.A., Jelsch J.M., Py M.C., Levin M.J., Cohen G.W. and Zakin M.M. (1987) *Gene* **56**, 109-116.

Schafer D.A. and Hultquist D.E. (1983) *Biochem. Biophys. Res. Commun.* **115**, 807-813.

Scherer G.R.E., Walkinshaw M.D., Arnot S and Moore D.J. (1980) *Nucleic Acids Res.* **8**, 3893-3907.

Schultze H.E. and Heremans J.F. (1960) *Molecular Biology of Human Proteins*, Elsevier, Amsterdam.

Schultze H.E. and Schwick G. (1957) *Behring Werk-Mitt* **33**, 11-17.

Searle P.F., Stuart G.W. and Palmiter R.D. (1985) *Mol. Cell. Biol.* **5**, 1480-1484.

Shawver L.K., Siedel S.L., Krieter P.A. and Shires T.K. (1984) *Biochem. J.* **217**, 623-632.

Shewale J.G. and Brew K. (1982) *J. Biol. Chem.* **257**, 9406-9415.

Shine J. and Delgarno L. (1975) *Nature* **254**, 34-38.

Simonsen C.C. and Levinson A.D. (1983) *Proc. Natl Acad. Sci. U.S.A.* **80**, 2495-2499.

Slaughter S.R. and Hultquist D.E. (1979) *J. Cell. Biol.* **83**, 231-239.

Slaughter S.R., Williams C.H. Jr. and Hultquist D.E. (1982) *Biochim. Biophys. Acta* **705**, 228-237.

Spatz L. and Strittmatter P. (1971) *Proc. Natl. Acad. Sci. U.S.A.* **68**, 1042-1046.

Spatz L. and Strittmatter P. (1973) *J. Biol. Chem.* **248**, 793-796.

Springer B.A. and Sligar S.G. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8961-8965.

Spik G., Bayard B., Fournet B., Strecker G., Bouquelet S. and Montreuil J. (1975) *FEBS Lett.* **50**, 296-299.

Starkenstein S. and Harvalik Z. (1933) *Arch. Exptl. Pathol. Pharmakol.* **172**, 75-92.

Stayton P.S., Poulos T.L. and Sligar S.G. (1989) *Biochemistry* **28**, 8201-8205.

Stellwagen E. (1978) *Nature* **275**, 73-74.

Stoffel W. (1961) *Biochem. Biophys. Res. Commun.* **6**, 270-273.

Stoneheurner E., Williams J.B. and Millet F. (1979) *Biochemistry* **18**, 5422-5427.

Stormo G.D., Schneider T.D. and Gold L.M. (1982) *Nucleic Acids Res.* **10**, 2971-2995.

Strickland D.K., Hamilton J.W. and Hudson B.G. (1979) *Biochemistry* **18**, 2549-2554.

Strittmatter C.F. and Ball E.G. (1954) *J. Cell Comp. Physiol.* **43**, 57-78.

Strittmatter C.F. and Velick S.F. (1956) *J. Biol. Chem.* **221**, 253-264.

Strittmatter P. and Ozols J. (1966) *J. Biol. Chem.* **241**, 4787-4792.

Struck D.K., Sivta P.B., Lane M.D. and Lennarz W.J. (1978) *J. Biol. Chem.* **253**, 5332-5337.

Sugiyama T., Miki N. and Yamuno T. (1979) *Biochem. Biophys. Res. Commun.* **28**, 73-78.

Surgenor D.M., Koechlin B.A. and Strong E.C. (1949) *J. Clin Invest.* **28**, 73-78.

Sutton M.R. and Brew K. (1974) *FEBS Lett.* **40**, 146-148.

Tabor S. and Richardson C.C.C. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4767-4771.

Takagaki Y., Radharkrishman R., Wirtz K.W.A. and Khorana H.G. (1983) *J. Biol. Chem.* **258**, 9136-9142.

Tan A.T. and Woodworth R.C. (1969) *Biochemistry* **8**, 3711-3718.

Taylor J.W., Ott J. and Eckstein F. (1985) *Nucleic Acids Res.* **13**, 8749-8764.

Teale F.W.J. (1959) *Biochim. Biophys. Acta* **35**, 543.

Teng C.T., Pentecost B.T., Marshall A., Bowman B.H., Laltey P.A. and Naylor S.L. (1987) *Somat. Cel Mol. Genet.* **13**, 689-695.

Thorbecke G.J., Liem H., Knight S. Cox K. and Muller-Eberhard V. (1973) *J. Clin. Invest.* **52**, 725-731.

Tomimatsu Y., Kint S. and Scherer J.R. (1973) *Biochem. Biophys. Res. Commun.* **54**, 1067-1074.

Ulmer D.B. (1969) *Biochim. Biophys. Acta* **181**, 305-311.

Valcour A.A. and Woodworth R.C. (1987) *Biochemistry* **26**, 3115-3120.

van Bockxmeer F.M. and Morgan E.H. (1979) *Biochim. Biophys. Acta* **584**, 76-83.

van Oost B.A., Edgell C.J.S., Hay C.W. and MacGillivray R.T.A. (1986) *Biochem. Cell Biol.* **64**, 699-705.

Veitch N.C., Concar D.W., Williams R.J.P., and Whitford P. (1988) *FEBS Lett.* **238**, 49-55.

von Bodman S.B., Schuler M.A., Jollie D.R. and Sligar S.G. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9443-9447.

von Heijne (1983) *J. Mol. Biol.* **159**, 537-541.

Wainio W.W. (1982) Chemistry and Biology of Mammalian b-Type Cytochromes In *"Int. Encyclopedia of Pharm. and Ther."* (Schenkman T. and Kupfer D., eds.) Pergamon Press.

Wainio W.W. (1978) *Pharmacol. Ther. A.* **2**, 359-372.

Walden W.E., Daniels-McQueen S., Smith L.L. and Thach R.E. (1987) *Gene* **61**, 317-327.

Walker F.A., Emrick D., Rivera J.E., Hanquet B.J. and Buttlaire D.H.C. (1988) *J. Am. Chem. Soc.* **110**, 6234-6240.

Wand A.J., DiStefano D.L., Feng Y., Roder H. and Englander S.W. (1989) *Biochemistry* **28**, 49-55.

Warner R.C. and Weber I. (1953) *J. Am. Chem. Soc.* **75**, 5094-5101.

Watts C. (1985) *J. Cell Biol.* **100**, 633-637.

West S.B., Levin W., Ryan D., Vore M. and Lu A.Y.H. (1974) *Biochem. Biophys. Res. Commun.* **58**, 516-522.

Williams J. (1975) *Biochem. J.* **149**, 237-244.

Williams J. (1974) *Biochem J.* **141**, 745-752.

Williams J. (1962) *Biochem. J.* **83**, 355.

Williams J., Elleman T.C., Kingston B., Wilkins A.G. and Kuhn K.A. (1982) *Eur. J. Biochem.* **122**, 297-403.

Woodbury R.G., Brown J.P., Yeh M.U., Hellstrom I. and Hellstrom K.E. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77, 2183-2187.

Woodworth R.C. (1986) *J. Inorg. Biochem.* **28**, 245-252.

Woodworth R.C., Morallee K.G. and Williams R.J.P. (1970) *Biochemistry* **9**, 839-842.

Woodworth R.C., Virkaitis L.M., Woodburg R.G. and Fava R. (1975) *In "Proteins of Iron Storage and Transport in Biochemistry and Medicine"* (R.R. Crichton ed.) North-Holland, Amsterdam.

Yang F., Lum J.B. and Bowman B.H. (1985) Molecular Genetics of Human Transferrin. *In "Protides of the Biological Fluids"* (H. Peeters, ed.) pp. 31-34, Pergamon, Oxford.

Yang F., Lum J.B., McGill J.R., Moore C.M., Naylor S.L., Van Bragt P.H., Baldwin W.D. and Bownam B.H. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2752-2756.

Yeh Y., Iwai S. and Feeney R.E. (1979) *Biochemistry* **18**, 882-889.

Yoo M. and Steggles A.W. (1988) *Biochem. Biophys. Res. Commun.* **156**, 576-580.

Yoo M. and Steggles A.W. (1989) *Biochem. Biophys. Res. Commun.* **163**, 18-24.

Yoshikawa H. (1951) *J. Biochem.* (Tokyo) **388**, 1-5.

Young S.P. and Aisen P. (1981) *Hepatology* **1**, 114-118.

Zak O. and Aisen P. (1985) *Biochim. Biophys. Acta* **829**, 348-353.

Zak O., Leibman A. and Aisen P. (1983) *Biochim. Biophys. Acta* **742**, 490-495.

Zhang H. and Somerville C. (1988) *Arch. Biochem. Biophys.* **264**, 343-347.

Zoller M.J. and Smith M. (1983) *Meth. Enzymol.* **100**, 458-500.

Appendix I: Measurement of the Reduction Potential of Biological Molecules

Most protein molecules do not exchange electrons with an unmodified electrode efficiently. In this report, two methods have been used to overcome this phenomenon and determine the reduction potentials of various recombinant cytochrome b_5 species. A brief summary of the technical and theoretical aspects of spectroelectrochemical and cyclic voltammetric methods is given below.

1. *Spectroelectrochemistry*

The transfer of electrons to an oxidized species can be described by the general reaction:

$$Ox + ne^- \underset{k_r}{\overset{k_f}{\rightleftharpoons}} Red \quad (1)$$

The distribution of reduced and oxidized species in such a reaction can be manipulated by adjusting the applied voltage to the system and is governed by the standard reaction potential (E^O) which is expressed in terms of the Nernst equation:

$$E = E^O + \frac{RT}{nF} \ln \frac{[Ox]}{[Red]} \quad (2)$$

Typically, optically transparent thin layer electrodes (OTTLE) consist of a thin minigrid, working electrode which is mounted across the width of a small cuvette. This working electrode covers the majority of the optical cross section of the cell and is of sufficient porosity to be transparent to transmitted light (*ca.* 60%). Application of a potential to the cell from an external source establishes the potential of the solution and the position of the red/ox equilibrium of the solution. The ratio of oxidized to reduced protein is determined at an appropriate wavelength.

Potential measurements can only be made relative to a reference potential. For the sake of experimental convenience, the Saturated Calomel Electrode (SCE) is a useful reference:

$Hg/Hg_2Cl_2/KCl$ (saturated, aq.) E= +0.244 V *vs.* SHE @ 25°C

Thus measurements made against the SCE are conveniently converted to the Saturated Hydrogen Electrode (SHE) scale by adding the factor 244.4 mV (25°C).

Electron transfer between the working electrode and the protein is achieved by the inclusion of a soluble small molecule mediator to the protein solution. $Ru(NH_3)_6Cl_3$ has been shown to undergo red/ox reactions with cytochrome b_5 and gold surfaces and thus serves to shuttle electrons between the working electrode surface and the cytochrome (Reid *et al.*, *J. Am. Chem. Soc.* **104**, 7516 (1982)). However, care must be taken with such a system to ensure that the solution has attained true electrochemical equilibration between each change in applied potential and the subsequent absorption spectrum determination. Solutions of relatively high ionic strength accelerate equilibration.

2. *Cyclic Voltammetry*

Cyclic voltammetry involves the detection of current during the linear scanning of applied potential between two limits. Several electronic effects are observed at the surface of an active electrode and must be considered when interpreting cyclic voltammograms.

The overall rate of heterogeneous electron transfer will depend primarily on the rate of electron transfer at the surface and on the rate of bulk transfer of electroactive species to and from the electrode surface. As proteins do not interact directly with unmodified electrodes, the electrode surface is usually functionalized with a coating of an organic promoter (see Armstrong *et al.*, *Quart. Rev. Biop.* **18**, 261 (1986)). In the present study, a small thiol-containing peptide (KCTCCA) from proteolytically-digested metallothionein was used to modify a gold electrode as has been previously described (Bagby *et al.*, *Biochem. Soc. Trans.* **16**, 958 (1988)).

For samples in which the rate of electron transfer is fast compared to the rate of bulk transfer, the reaction rate will be limited by rate of mass transfer from the electrode surface. Systems with this characteristic are electrochemically reversible (Nernstian), and it is implicit that species at the electrode surface are always at an equilibrium position governed by the electrode potential. At the other limit, samples in which the electron transfer rate is slow compared to the bulk transfer rate are referred to as irreversible. The electrochemical behaviour of proteins at modified electrodes is often characterized by rates between these two limits. Such quasi-reversible systems display Nernstian behaviour at a low potential sweep rate, but become increasingly irreversible at higher sweep rates.

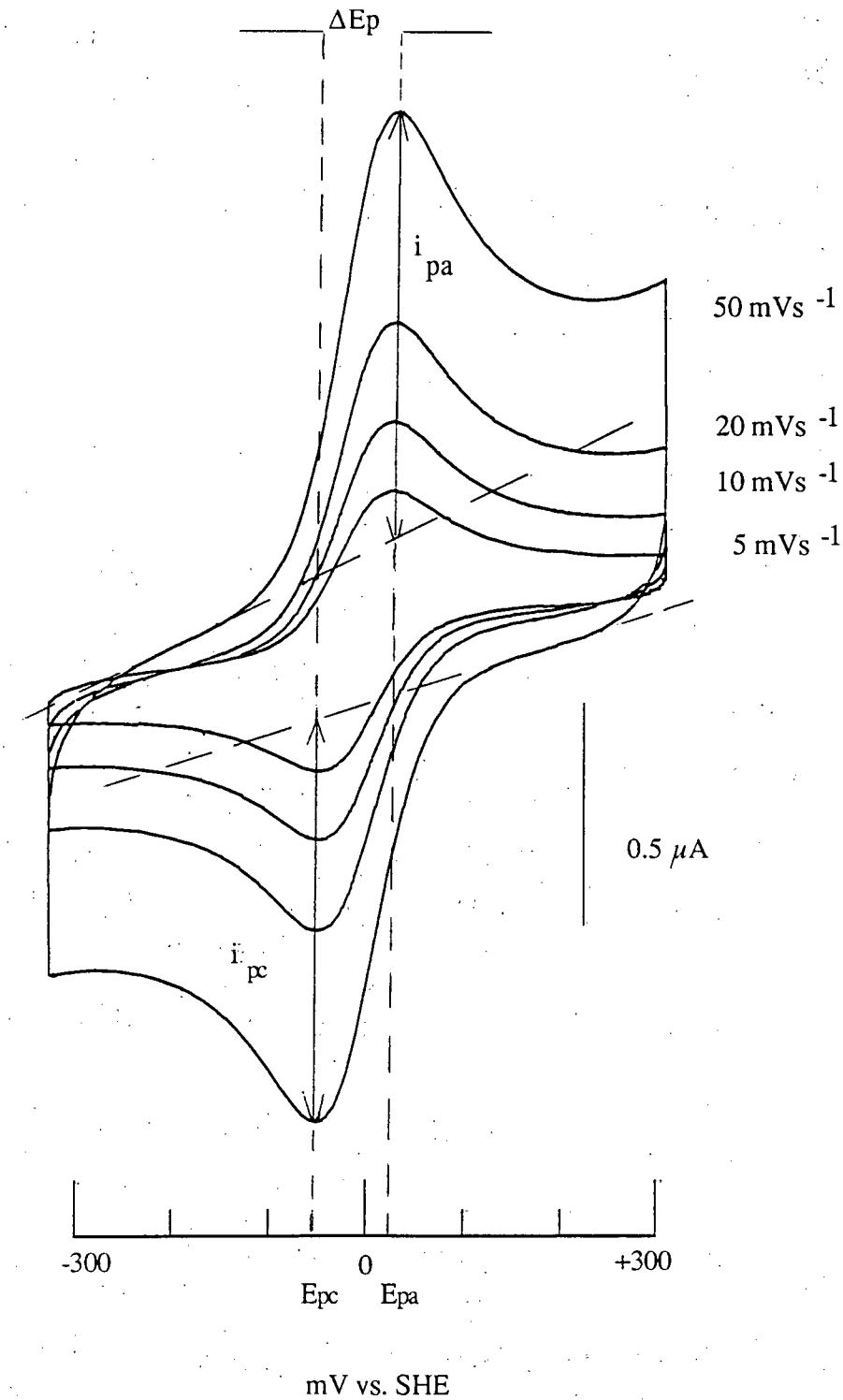
As the potential is raised from a state of full oxidation, reduction occurs at the electrode surface and the resulting faradaic current increases. Initially, substrate reduction increases with the rising potential as seen in Figure IA. As the process continues, a layer of reduced product forms at the electrode surface which can only be relieved by the diffusion-controlled removal of product. A cathodic peak current (i_{pc}) results from the competition between increasing potential and the developing accumulation layer. Upon sweep reversal, a corresponding anodic peak current (i_{pa}) is observed for analogous reasons involving an increasing oxidizing potential competing with the developing accumulation layer of oxidized products. An estimation of i_{pc} and i_{pa} can be made by extrapolating the background (or capacitive) current as shown in Figure IA and the corresponding peak potentials (E_{pc} and E_{pa}) can be assigned. The half-wave potential ($E_{1/2}$) is found midway between E_{pc} and E_{pa} and is analogous to the midpoint or reduction potential, provided the kinetics of electron transfer are identical for both the reduced and the oxidized species.

An estimation of the reversibility of the system can be made by examining the peak current (i_{pc}) as a function of sweep rate (ν). The Cottrell equation describes the current i, at time t after a potential step as:

$$i(t)= nFAC_0^{bulk}(\pi D_0 a)^{1/2}\nu(at) \quad \text{with } a=(nF\nu t)/RT \quad (3)$$

where n is the number of electrons involved, F is Faraday's constant, A is the area of the electrode, C_0^{bulk} is the bulk concentration of oxidized species and D_0 is the diffusional coefficient of the oxidized species. From (3) it is apparent that the current, i, is proportional to $\nu^{1/2}$, provided the red/ox process is limited by planar diffusion. Thus for fully reversible processes, a plot of the peak current, i_{pc} *versus* $\nu^{1/2}$, yields a constant. For quasi-reversible couples, such as most protein systems, the reversible limit is achieved at low sweep rates, but at higher rates (>0.2 Vs^{-1}), the rate of electron transfer becomes limiting as is the case for the cytochromes b_5 studied here (Figure IB).

A


ΔEp
i pa
50 mVs -1
20 mVs -1
10 mVs -1
5 mVs -1
0.5 μA
i pc
-300
0
+300
Epc
Epa
mV vs. SHE

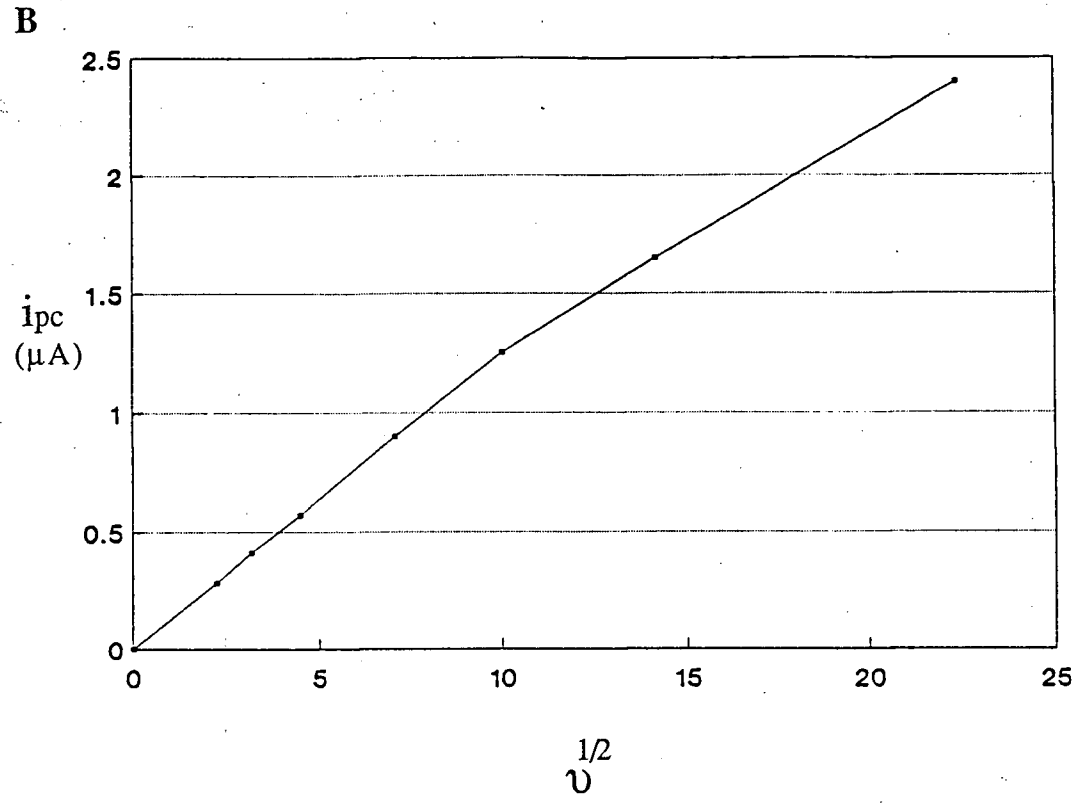


Figure I: Cyclic Voltamogram of the D57 variant of recombinant Lpb$_5$

Panel A: Cyclic voltammetry was performed on a solution of D57 Lpb_5 (100μM) using a peptide-modified gold electrode as described previously (Bagby *et al.*,1988). The corresponding values for the peak currents (i_{pc} and i_{pa}), peak potentials (E_{pc} and E_{pa}) and the peak separation (ΔE_p) were calculated from the sweep at 50 mVs^{-1}.

Panel B: A plot of the peak current *versus* $\nu^{1/2}$ shows the quasi-reversible characteristics of many protein couples. At low sweep rate, the relationship is linear, obeying the Cottrell relationship (Equation 3). At higher sweep rates, the rate of electron transfer becomes limiting, and linearity is lost.