Journal of Bioenergetics and Biomembranes, Vol. 27, No. 2, 1995

Porphobilinogen Synthase, The First Source of Heme's Asymmetry

Eileen K. Jaffe[1]

Received January 18, 1995; accepted February 1, 1995

Porphobilinogen is the monopyrrole precursor of all biological tetrapyrroles. The biosynthesis of porphobilinogen involves the asymmetric condensation of two molecules of 5-aminolevulinate and is carried out by the enzyme porphobilinogen synthase (PBGS), also known as 5-aminolevulinate dehydratase. This review documents what is known about the mechanism of the PBGS-catalyzed reaction. The metal ion constitutents of PBGS are of particular interest because PBGS is a primary target for the environmental toxin lead. Mammalian PBGS contains two zinc ions at each active site. Bacterial and plant PBGS use a third metal ion, magnesium, as an allosteric activator. In addition, some bacterial and plant PBGS may use magnesium in place of one or both of the zinc ions of mammalian PBGS. These phylogenetic variations in metal ion usage are described along with a proposed rationale for the evolutionary divergence in metal ion usage. Finally, I describe what is known about the structure of PBGS, an enzyme which has as yet eluded crystal structure determination.

KEY WORDS: Porphobilinogen synthase; 5-aminolevulinate dehydratase; enzyme mechanisms; zinc metalloenzyme, magnesium proteins.

INTRODUCTION

Porphobilinogen synthase (PBGS, also known as 5-aminolevulinate dehydratase, E.C. 4.2.1.24) is a metalloprotein that catalyzes the formation of porphobilinogen (PBG) from two molecules of 5-aminolevulinic acid (ALA) (Shemin and Russell, 1953). The PBGS-catalyzed reaction is the first common step in the biosynthesis of all tetrapyrroles and is essential to all life. Sequence comparisons illustrated in Fig. 1 show PBGS to be a highly conserved protein which suggests that the fundamental catalytic mechanism is probably the same for all organisms. Nevertheless, there are variations among phyla in the use of metal ions for catalytic and regulatory function. This monograph reviews what is known about the PBGS-catalyzed reaction mechanism with emphasis on data obtained in our laboratory. A general three-metal ion model is discussed along with a proposed rationale for the observed evolution of metal ion usage by PBGS. Lead inhibition of PBGS and its use as a diagnostic tool for lead poisoning is also described. Finally, we present what is known about the structure of PBGS.

[1] Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111.

THE PBGS-CATALYZED REACTION MECHANISM

The PBGS-catalyzed reaction is an asymmetric condensation of two identical substrate molecules which appears to be carried out by a homo-dimeric active site. For example, each homo-octamer of mammalian PBGS contains four active sites. Figure 2 illustrates the PBGS-catalyzed reaction and differentiates the two substrate molecules denoted as A-side ALA and P-side ALA. The two identical substrate molecules are distinct when bound to the enzyme, and their chemical fates are different. This ancient, essentially irreversible reaction involves the formation or cleavage of at least eight chemical bonds.

0145-479X/95/0400-0169$07.50/0

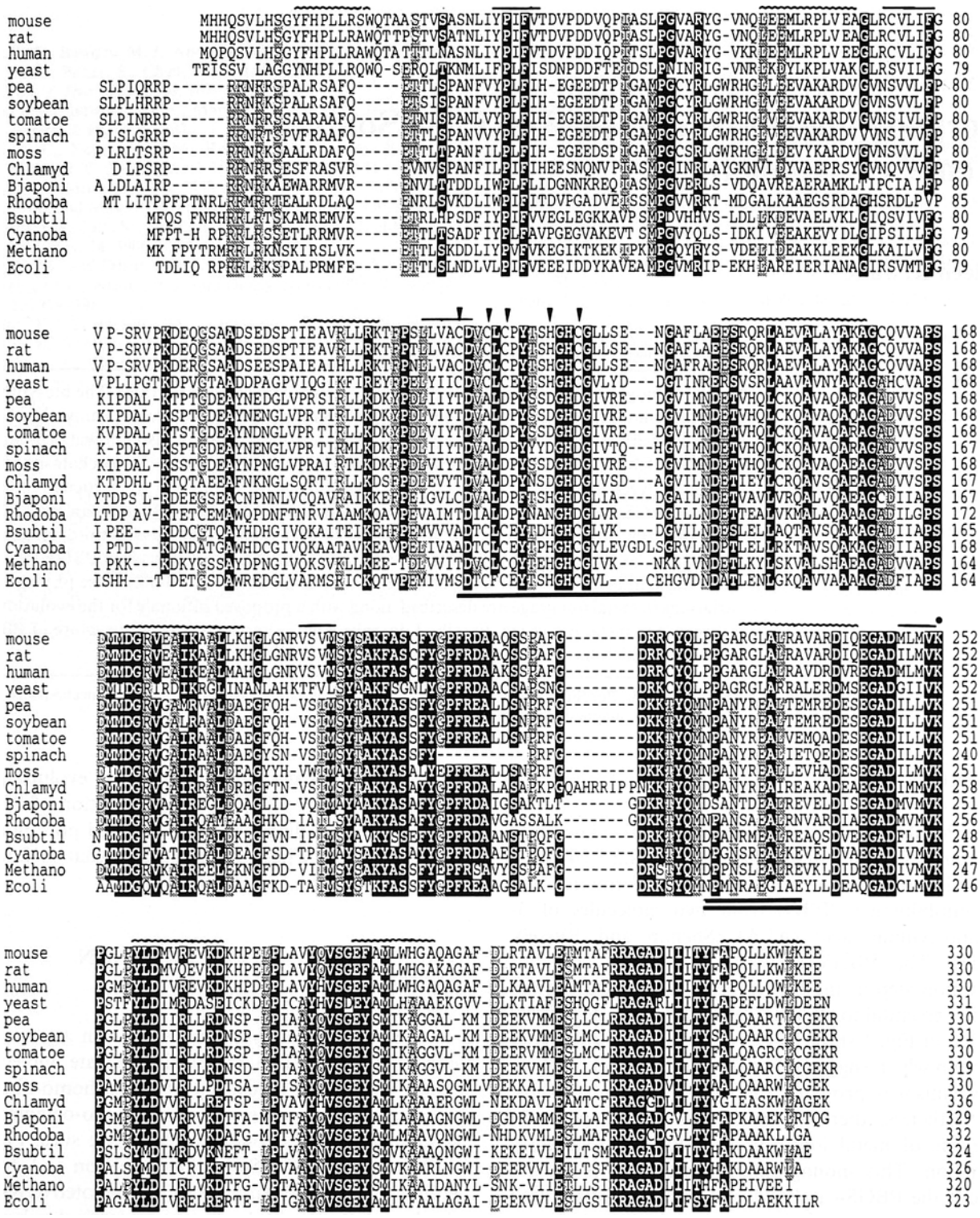
mouse
rat
human
yeast
pea
soybean
tomatoe
spinach
moss
Chlamyd
Bjaponi
Rhodoba
Bsubtil
Cyanoba
Methano
Ecoli

Fig. 1

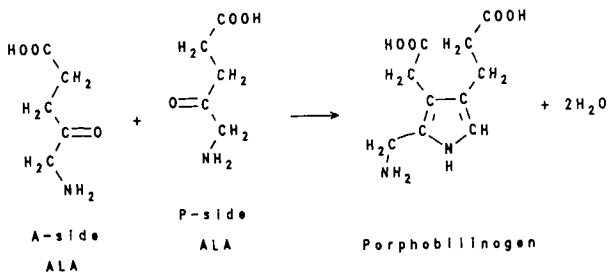

Fig. 2. The PBGS-catalyzed condensation of two molecules of 5-aminolevulinic acid (ALA). A-side ALA forms the acetyl half of the product porphobilinogen; its amino moiety remains intact. P-side ALA forms the propionyl half of porphobilinogen; its amino moiety becomes part of the pyrrole ring.

Consequently one can imagine a variety of chemical mechanisms for the PBGS-catalyzed reaction. The literature contains several fundamentally different mechanistic schemes (Nandi and Shemin, 1968; Batlle and Stella, 1978; Jordan and Seehra, 1980; Gibbs and Jordan, 1986; Jordan, 1990, 1991). However, the experimentally determined facts are limited. Some of the variation in proposed mechanisms arises from an inherent difficulty in determining whether reactions occurring with substrate analogs are at the A-side ALA binding site or the P-side ALA binding site.

An active-site model and abbreviated mechanistic scheme for mammalian PBGS is presented in Fig. 3. Figure 3A contains our model for the ternary ES_2Zn_2 complex (where S = ALA). Each active site binds two ALA and two Zn(II), each in one of two distinct environments. It has been established that P-side ALA can bind first and forms a Schiff base to Lys252 of mammalian PBGS (Nandi, 1978; Gibbs and Jordan, 1986). The formation of this complex does not require Zn(II) (Jaffe and Hanes, 1986). Because the substrates are chemically identical, it is not possible to establish whether P-side ALA *must* bind first. A-side ALA binding determines the apparent K_m for ALA as the affinity for P-side ALA is much higher. Further on in the reaction the pro-R proton of the C_5 of P-side ALA is stereospecifically removed (Chaudhry and Jordan, 1976), but the structure of the intermediate from which it is removed has not been established. The experimental data supporting the details illustrated in Fig. 3 are described below.

The Schiff Base Structure

^{13}C and ^{15}N NMR have been used to decipher the tautomeric structure, stereochemistry, and protonation state of the P-side Schiff base intermediate as illustrated in Fig. 3A (Jaffe and Markham, 1987, 1988; Jaffe *et al.*, 1990). Structure elucidation of the P-side Schiff base intermediate was facilitated by a methyl methanethiosulfonate (MMTS)-modified form of mammalian PBGS which contains no Zn(II) and is arrested after formation of the intermediate (Jaffe *et al.*, 1984). Table I contains the chemical shift data used to determine the P-side Schiff base structure. Because C_5 of P-side ALA loses a proton during PBG formation, the C_4–C_5 enamine is an attractive structural possibility. Nevertheless, all ^{13}C NMR data support the imine structure of E stereochemistry; ^{15}N NMR data indicate deprotonation of the amino group. These studies established the precedent for observing ^{13}C- and ^{15}N-labelled ligands bound to a protein as large as 280,000 dalton.

Fig. 1. The ~330 amino acid core of the PBGS sequences. This multiple sequence alignment was prepared using the programs MALIGNED and MALFORM (Clark, 1992) on a multiple sequence file prepared using PILEUP within the GCG package of sequence analysis programs (Gribskov and Devereux, 1991). In the figure, the following amino acids are set as equivalent, R = K, D = E, Q = N, S = T, and Y = F. The amino acids are shaded by percent conservation (≥87% = white on black, ≥75% = shaded). All PBGS are highly homologous throughout the ~330 amino acids with the highest variations seen in the fifty most N-terminal residues. The singly underlined region is referred to as the putative metal binding region of PBGS. It appears to contribute to the binding of Zn_A, Zn_B, and/or Mg_B. Those amino acids marked with an arrow relate to the switch between Zn(II) and Mg(II). The doubly underlined region has been proposed to bind Mg_C. The active site lysine is marked with a dot. From top to bottom, the species are mouse, rat, human, yeast, pea, soybean, tomato, spinach, moss, *Chlamydomonus*, *B. japonicum* (a nitrogen-fixing symbiot), *R. capsulatus*, *B. subtilis*, *A. nidulans* (a cyanobacter), *M. sociabilis* (a methanobacter), and *E. coli*. For some sequences the N-terminal chloroplast targeting region is omitted. Currently there are 16 different DNA sequences for PBGS's as well as considerable protein sequence information from peptides of both bovine and *E. coli* PBGS (Lingner and Kleinschmidt, 1983; Bishop *et al.*, 1986; Wetmur *et al.*, 1986; Myers *et al.*, 1987; Echelard *et al.*, 1988; Bishop *et al.*, 1989; Li *et al.*, 1989; Boese *et al.*, 1991; Hansson *et al.*, 1991; Bröckl *et al.*, 1992; Jaffe *et al.*, 1992; Schaumburg *et al.*, 1992; Chauhan and O'Brian, 1993; Markham *et al.*, 1993; Sollbach and Schneider-Poetsch, 1993; Beale, 1994; Jones *et al.*, 1994; Kaczor *et al.*, 1994; Polking *et al.*, 1994; Indest and Biel, 1994). The bovine PBGS sequence is omitted from Fig. 1 because it is incomplete. Sufficient sequence information is available to predict the secondary-structure elements of PBGS with some certainty. The data illustrated in Fig. 1 were submitted to EMBL for secondary-structure prediction (Rost and Sander, 1993, 1994; Rost *et al.*, 1994). The regions marked with a wavy line above the sequence are predicted to be α-helices. The regions marked with a bar above the sequence are predicted to be beta sheets.

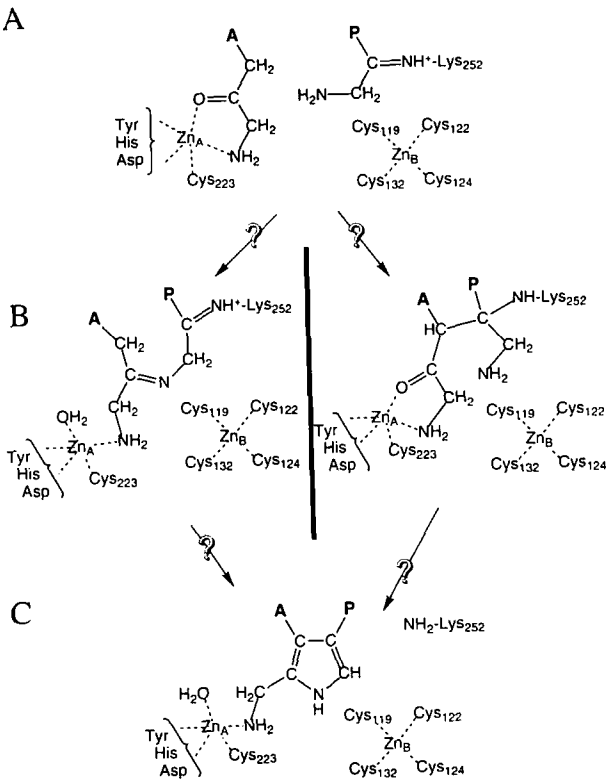

Fig. 3. The PBGS reaction mechanism illustrating two plausible pathways. Part A illustrates a transient ES_2Zn_2 complex and is specific for bovine PBGS. Data supporting this model are extensive and are described in the test. Part B illustrates two possible structures for the first addition product. Part C illustrates the enzyme–product complex. The A and P abbreviations refer to acetyl and propionyl moieties.

The Zn(II) Complex of A-Side ALA

Studies of MMTS-modified bovine PBGS also determined that a catalytically essential Zn(II) must bind before A-side ALA (Jaffe and Hanes, 1986). The current terminology refers to this Zn(II) as Zn_A (Dent *et al.*, 1990); four Zn_A are required for catalysis (Bevan *et al.*, 1980; Jaffe *et al.*, 1984). The K_d for Zn_A is much less than 0.1 μM and too tight to quantify by atomic absorption spectroscopy (Jaffe *et al.*, 1992). Each Zn_A contains five ligands, only one of which is a cysteine (Dent *et al.*, 1990). Mammalian PBGS also contains four Zn_B which have a lower affinity than Zn_A, 5 μM. The Zn_B are each tetrahedrally coordinated to four cysteine ligands and are discussed in more detail below (Dent *et al.*, 1990; Jaffe *et al.*, 1992).

5-Chlorolevulinate (5-CLA), a reactive mimic of ALA, has been used to modify bovine PBGS, and this modification identified a potential role for Zn_A (Jaffe *et al.*, 1992). Prior to our work, it had been established that 5-CLA inactivates bovine PBGS and that modification at four sites per octamer (one per active site) affords near-total inactivation (Seehra and Jordan, 1981). We demonstrated that 5-CLA-modified bovine PBGS can bind four substrate molecules as the P-side Schiff Base (Jaffe *et al.*, 1992). Thus, the preferential site of 5-CLA inactivation was shown to be the A-side ALA binding site of bovine PBGS. 5-CLA modifies Cys223 on one-half the subunits. This modification results in a loss of the Zn_A binding sites, implicating Cys223 as the one-cysteine ligand to Zn_A. Lacking Zn_A, 5-CLA-modified bovine PBGS is unable to catalyze any bond formation between ALA bound as the P-side Schiff base intermediate and an A-side levulinic acid moiety derived from 5-CLA. On this basis Zn_A was proposed to play an essential role in formation of the first bond between the two substrate molecules. Zn_A may polarize the carbonyl of A-side ALA to promote attack by the P-side ALA amino group (leading to Fig. 3B, left) or it may act as a Lewis acid promoting enolization at C_3 of A-side ALA (leading to Fig. 3B, right). A-side ALA is proposed to be a bidentate ligand to Zn_A, a structure that mimics the X-ray crystal structure of glycylglycine bound to the Zn(II) of carboxypeptidase (Christianson and Lipscomb, 1986). A monodentate complex of Zn(II) and A-side ALA has also been proposed (Spencer and Jordan, 1994).

Table I. ^{13}C NMR Chemical Shift Data Supporting the P-Side Schiff Base Structure Illustrated in Fig. 3A

	Expected chemical shift changes for four possible structures				
^{13}C-Labelled substrate	Enamine C_3–C_4	Enamine C_4–C_5	Imine (Z)	Imine (E)	Observed chemical shift changes
[4-^{13}C]ALA	−40 ± 10	−40 ± 10	−40 ± 10	−40 ± 10	−39.4
[5-^{13}C]ALA	6 ± 4	100 ± 20	−12	⩽ −3	−2.7
[3-^{13}C]ALA	100 ± 20	6 ± 4	⩽ −3	−12	−12.5

Placement of Zn(II) Near the P-Side Schiff Base

5-Chloro-[1,4-^{13}C]levulinate ([1,4-^{13}C]CLA) modification of *E. coli* PBGS provides the data to place Zn_B at the enzyme active site (Jaffe *et al.*, 1994). *E. coli* PBGS contains Ser217 in the position analogous to Cys223, the 5-CLA-modified Zn_A ligand of bovine PBGS. Consequently, [1,4-^{13}C]CLA modification of *E. coli* PBGS differs from that of the mammalian enzyme. In fact, [1,4-^{13}C]CLA modification of *E. coli* PBGS occurs preferentially at the P-side ALA binding pocket, although the modification is heterogeneous. For one of four observed forms of [1,4-^{13}C]-CLA modified *E. coli* PBGS, an analog of the P-side Schiff base was found. Proteolysis and peptide sequencing demonstrate that 5-CLA modifies *E. coli* PBGS at either Cys 119 or Cys129 on any one peptide. Cys119 and Cys129 are part of a four-cysteine cluster implicated in binding Zn_B (see arrows in Fig. 1). Thus, Zn_B must be near the C_5 of P-side ALA as illustrated in Fig. 3A. Unlike *E. coli* PBGS, mammalian PBGS requires only Zn_A for maximal catalytic activity in the presence of 2-mercaptoethanol. Thus, the four Zn_B have previously been called the noncatalytic or structural Zn(II). However, in the absence of 2-mercaptoethanol and exogenous Zn(II), mammalian PBGS is prone to disulfide formation, Zn(II) loss, and a concomitant reversible inactivation (Jaffe *et al.*, 1984). It has been suggested that one role for Zn(II) in PBGS is to prevent disulfide formation between catalytically essential SH groups, and this may be the role of Zn_B (Batlle and Stella, 1978; Tsukamoto *et al.*, 1979).

Placement of the 4-cysteinyl Zn_B at the active site suggests unprecedented chemistry. 4-Cysteinyl Zn(II) are typically not catalytic (e.g., Jaffe, 1993). However, the C_5 of P-side ALA loses the pro-R proton at some time during catalysis and this proton is near Zn_B (see Fig. 3A). I propose that the cysteines serve to accept the proton and shuttle it to water. In the absence of Zn_B this process would result in disulfide formation; thus, 2-mercaptoethanol can serve the role of Zn_B in bovine PBGS, making Zn_B appear nonessential. In contrast, the Zn_B of *E. coli* PBGS appears less accessible to solvent, 2-mercaptoethanol cannot fill its role, and Zn_B appears to be essential. PBGS from some species, particularly plants, appear to contain an analogous Mg_B coordinated to aspartates which are also predicted to be essential to catalysis (see Fig. 1 and below).

What Comes First, C–C or C–N Bond Formation?

Figure 3B illustrates uncertainty about the structure of the first bi-substrate intermediate, its mechanism of formation, and its journey to product. The NMR-determined deprotonation of the amino group of the P-side Schiff base intermediate would facilitate formation of a second Schiff base between this amino group and C_4 of the A-side ALA (see Fig. 3B, left side). This general reaction scheme, without the Zn(II), had previously been suggested (Seehra and Jordan, 1981). An alternative first addition product is illustrated in Fig. 3B (right side). Though no evidence proves or refutes its existence, the C–C addition product is preferred by several research groups on the basis of stereoelectronic arguments (e.g., Fabiano and Golding, 1991; Spencer and Jordan, 1994). Both addition products have precedents in synthetic pyrrole chemistry.

The Enzyme–Product Complex

Figure 3C depicts a model of the enzyme–product complex. The ^{15}N and ^{13}C spectra of enzyme-bound product established that at neutral pH the product porphobilinogen is bound in such a way that its free amino group (derived from A-side ALA) is unprotonated. This is surprising as the solution pK_a of this amino group is 10–11 (Evans *et al.*, 1985; Jaffe and Markham, 1988). The finding that Zn_A is required for the binding of A-side ALA as well as of product leads to a model where enzyme-bound PBG is coordinated to Zn_A through the unprotonated free amino group (Jaffe *et al.*, 1990).

Can One Predict the Remaining Zn(II) Ligands of Mammalian PBGS?

Based on extended X-ray absorption fine structure (EXAFS) data, Fig. 3 includes Asp, Tyr, and His as the remaining Zn_A ligands. In the absence of substrate or product, all three may be ligands. In the presence of A-side ALA or PBG, perhaps only two are ligands. I analyzed the Zn(II) ligations patterns of 24 known X-ray and NMR determined structures for proteins which contain either a catalytic or noncatalytic Zn(II) (roughly analogous to Zn_A and Zn_B) in an attempt to predict which specific amino acids serve as the Zn_A and Zn_B ligands (Jaffe, 1993). Catalytic Zn(II) were found to be either tetrahedral or pentacoordinate. The most common catalytic Zn(II) ligand is histidine, and the ligand preference series appears to

be H > E = D > C = Y = K = backbone carbonyl. However, the data set is still small and dominated by carboxypeptidase-like enzymes. It is possible that any nitrogen, oxygen, or sulfur can be a catalytic Zn(II) ligand, though more than two sulfur ligands is unprecedented. In all cases, except alcohol dehydrogenase, two of the catalytic Zn(II) ligands are separated by three or fewer intervening amino acids. On the assumption that all PBGS contain Zn_A, and that the Zn_A ligands are predominantly conserved amino acids, the PBGS sequences were searched for near-neighboring conserved Asp, Tyr, and His residues. Two possibilities were found, Asp220 and Tyr224 or Tyr257 and Asp259. His131 is the only conserved histidine and thus is implicated as a Zn_A ligand. Cys223 was previously described as Zn_A ligand. Because each homo-octameric PBGS contains only four Zn_A, the three or four protein-derived ligands need not all be on the same subunit.

Noncatalytic Zn(II) sites were found to all contain a tetrahedral coordination of at least two, and often four, cysteine ligands. Two of the cysteines are always separated by fewer than four intervening amino acids and all noncatalytic Zn(II) ligands derive from a continuous sequence of up to 35 amino acids (Jaffe, 1993). By analogy, Cys119, Cys122, Cys124, and Cys132 of mammalian PBGS are probable ligands to Zn_B (see Fig. 1). Further evidence that these cysteines are zinc ligands came from sequencing peptides modified with a nitroxide spin-labelled derivative of MMTS (Markham *et al.*, 1993) because MMTS modification releases all Zn(II).

The Role of Histidine in the PBGS-Catalyzed Reaction

The double dehydration catalyzed by PBGS is likely to involve acid-base chemistry at the active site. Since chemical modification of histidine has been shown to inactivate PBGS from various phyla, a supposition is that a histidine participates in the acid-base chemistry of the enzyme (Fukuda *et al.*, 1998, 1990; Maralihalli *et al.*, 1985; Tsukamoto *et al.*, 1975). His128 (numbered according to *E. coli* PBGS) is the only totally conserved histidine in the known PBGS sequences and His126 is conserved in all but plant and *Rhodobacter capsulatus* PBGS (see Fig. 1). Preliminary results on the *E. coli* PBGS mutants H128A, H126A, and H126,128A show that with sufficient Zn(II) and Mg(II), the catalytic activity of these proteins varies only marginally from wild type (Mitchell, Volin, and Jaffe, 1995). These results suggest that neither His126 nor His128 are essential for catalytic activity. However, reduced affinities for Zn(II) are consistent with one or both of these histidines serving as a Zn_A ligand.

Unanswered questions

There remains much uncertainty about the details of the PBGS reaction mechanism as indicated by the large question marks in Fig. 3. We and others have yet to mutate an essential amino acid, though it is presumed that the lysine through which the P-side Schiff base is formed is essential. Perhaps few other amino acids are required. Perhaps the metal ions do much of the work while the protein merely provides the scaffolding to bind substrate. Neier has suggested that the PBGS-catalyzed reaction, beyond the Schiff base formation, may be spontaneous even in the absence of metal ions (Neier, 1995). Further mutagenesis experiments may support this intriguing possibility.

VARIATIONS IN METAL ION USAGE BY PBGS AND THE DEVELOPMENT OF A THREE-METAL ION MODEL

The two Zn(II) active site model depicted in Fig. 3 for mammalian PBGS does not accurately represent PBGS from all phyla. Most PBGS have been described as using Zn(II) or Mg(II) (Shemin, 1972), but interpretation of the data on the role of metal ions in the PBGS-catalyzed reaction has been limited by the assumption that there was a *single* role for the metal (or a single type of metal ion). Now that the metal ion usage of both mammalian and bacterial PBGS have been characterized and many PBGS sequences are known, it is easier to reevaluate the literature in terms of a three-metal ion model first invoked for *E. coli* PBGS (Mitchell and Jaffe, 1993). The three-metal ion model is summarized in Table II.

Table II. A Three-Metal Ion Model for Homo-Octameric PBGS

Species	Site A	Site B	Site C	Total metal ions
Mammalian and yeast	4 Zn	4 Zn		8
E. coli, *Cyanobacter*, and most other bacteria	4 Zn	4 Zn	8 Mg	16
Bradyrhizobium japonicum	4 Zn	4 Mg	8 Mg	16
Rhodobacter, algae, and plants	? 4 Mg ?	4 Mg	8 Mg	16

The Two Metals of Mammalian PBGS

The two metal ions of mammalian PBGS, a homo-octamer of 35 kDa subunits containing eight Zn(II) per octamer (Tsukamoto *et al.*, 1979; Bevan *et al.*, 1980; Jaffe *et al.*, 1984), have already been described as four Zn_A and four Zn_B. No third metal has been found for mammalian PBGS. In the absence of Zn(II), mammalian PBGS is catalytically inactive, but its octameric structure remains intact (Jaffe *et al.*, 1984). The proof of two classes of Zn(II) came from an EXAFS analysis of the Zn_4 vs. Zn_8 forms of bovine PBGS (Dent *et al.*, 1990). The coordination chemistry and roles of Zn_A and Zn_B are described above.

The Three Metal Ions of *E. coli* PBGS

Characterization of *E. coli* PBGS revealed the third type of metal ion. *E. coli* PBGS and mammalian PBGS share some characteristics but differ in others (Mitchell and Jaffe, 1993). The similarities include a requirement for Zn(II), inhibition by 1,10-phenanthroline, exceptional thermal stability, and a requirement for free sulfhydryl(s) for activity as shown by MMTS inactivation. One striking similarity between *E. coli* and bovine PBGS is the ^{13}C NMR spectrum of enzyme-bound [3,5-^{13}C]PBG; the chemical shifts of bound product are identical for both proteins and quite different from free PBG. One significant difference between *E. coli* and mammalian PBGS is the metal ion stoichiometry. Proton-induced X-ray emission analysis showed a Zn(II) stoichiometry of two per subunit in *E. coli* PBGS, in contrast to the one Zn(II) per subunit in mammalian PBGS (Mitchell and Jaffe, 1993). Complementary studies using atomic absorption analysis show that the second equivalent of Zn(II) in *E. coli* PBGS is readily replaced by Mg(II) (Spencer and Jordan, 1993). Without Zn(II), octameric *E. coli* PBGS shows no activity, regardless of whether Mg(II) is present, but the Zn_8Mg_8 enzyme is far more active than the Zn_{16} species. Dialysis of the Zn_{16} species against EDTA shows the rapid loss of all but four Zn/octamer (Spencer and Jordan, 1993). These four Zn, which are lost much more slowly, are herein interpreted to be analogous to the Zn_A of mammalian PBGS.[2] Preliminary EXAFS analysis of the Zn_8Mg_8 form of *E. coli* PBGS reveal spectra similar to mammalian PBGS containing a total of eight Zn(II), four Zn_A, and four Zn_B (unpublished results). Therefore, the eight Mg(II) of *E. coli* PBGS bind *in addition to* Zn_A and Zn_B and are denoted Mg_C. The observation that addition of Mg(II) affords a stimulation of *E. coli* PBGS activity is reminiscent of a well-documented effect of Mg(II) on plant PBGS (e.g., Liedgens *et al.*, 1983); this effect is not seen for bovine PBGS and suggests that Mg_C is common only to plant and bacterial PBGS. It is noteworthy that other workers established that *E. coli* PGBS binds eight Zn(II) and eight Mg(II) and interpreted the use of Zn(II) and Mg(II) by *E. coli* PBGS as analogous to the Zn_A and Zn_B of mammalian PBGS, not necessarily respectively (Spencer and Jordan, 1993). A study using Co(II) as an optical probe was interpreted to indicate that *E. coli* PBGS uses eight catalytic Mg_α and eight noncatalytic Zn_β, where the α and β are analogous to the A and B of Dent *et al.* (1990) (Spencer and Jordan, 1994).

The Three-Metal Ion Model

The aforementioned data and protein sequences of PBGS led to the three-metal ion model for PBGS (see Table II). I propose that mammalian PBGS uses only Zn_A and Zn_B; *E. coli*, *Cyanobacter*, and most other bacterial PBGS use Zn_A, Zn_B, and Mg_C; and *Rhodobacter capsulatus*, algae, and plant PBGS use Mg(II) in both the B and C sites. It is not yet clear if the plantlike PBGS use Zn(II) or Mg(II) in the A site. At least one species, *Bradyrhizobium japonicum*, is proposed to have a Zn_A, Mg_B, Mg_C usage (see below).

The Metal Binding Regions of PBGS

Elucidation of the human PBGS sequence revealed a "cysteine- and histidine-rich ... zinc finger-like" domain (Wetmur *et al.*, 1986). This putative metal binding region is underlined in Fig. 1. In this region it is the nonconserved amino acids which are intriguing (see arrows). We propose that this region of mammalian PBGS contributes ligands to both Zn_A and Zn_B in a *mutually exclusive fashion* (Jaffe *et al.*, 1992). The four cysteines are putative ligands to Zn_B, and at least one histidine, tyrosine, or aspartic acid residue is a ligand to Zn_A as suggested by the EXAFS data (Dent *et al.*, 1990; Jaffe, 1993). Elucidation of the first sequence of a plant PBGS revealed that the cysteine residues of the putative metal binding region were not present (Boese *et al.*, 1991). Instead, this region of plant PBGS is rich in aspartic acid residues. Since it had been previously demonstrated that

[2] Spencer and Jordan (1993) interpret their data differently. The reader is referred to the original figures.

the activity of plant PBGS was stimulated by the addition of Mg(II) (Liedgens *et al.*, 1983), this sequence was proposed to govern whether a PBGS required Zn(II) *or* Mg(II) for activity. However, the fact that *E. coli* PBGS, which contains the cysteine-rich region, requires Zn(II) *and* is stimulated by Mg(II) indicated that the stimulatory Mg(II), now known to be allosteric and denoted Mg_C, must bind somewhere else in the sequence (Mitchell and Jaffe, 1993). The underlined region of Fig. 1 is proposed to govern whether a particular PBGS uses Zn_B or Mg_B (see Table II). *Bradyrhizobium japonicum* PBGS has an unusual sequence in the putative metal binding region (Chauhan and O'Brian, 1993). It contains most of the putative ligands to Zn_A, only one cysteine, most of the putative ligands to Mg_B, as well as the distant proposed Mg_C site. This unusual PBGS is predicted to be a four Zn_A/four Mg_B/eight Mg_C species, a hypothesis we are actively testing.

To localize the Mg_C region, we looked for the sequence regions of *E. coli* PBGS which were similar to plant PBGS (which has the Mg(II) responsive element) and different from yeast and mammalian PBGS which do not respond to Mg(II). We presume that the Mg_C site will contain at least two neighboring oxygen-containing amino acid side chains. On this basis, the region of sequence predicted to bind Mg_C is doubly underlined in Fig. 1. It lies between positions 221 and 233 in the *E. coli* sequence. Recently, pea PBGS was used to confirm that there is an allosteric Mg_C, common to plant and *E. coli* PBGS, which binds at a site outside the putative metal binding region singly underlined in Fig. 1 (J. S. Ventura and M. P. Timko, personal communication). Cassette mutagenesis was used to prepare a hybrid of pea PBGS containing the putative metal binding region of human PBGS. The hybrid protein has a Zn(II) requirement as well as an allosteric Mg(II). The hybrid protein is similar in sequence and characteristics to *E. coli* PBGS. The exact location of the Mg_C binding site remains unproven.

The Allosteric Mg_C

In *E. coli* PBGS, the nonessential Mg_C can be characterized independently of the other metal ions because it is the only Mg(II) present (Jaffe *et al.*, 1995). Mg_C lowers the K_m for ALA, increases the specific activity, and decreases the saturating concentration of the required Zn(II) ($Zn_A + Zn_B$). The K_d for Mg_C is ~0.1 mM in 0.1 M phosphate buffer, and the site is probably saturated under physiologic conditions. Mn(II) is a good substitute for Mg_C and has been used as an EPR-active probe of the Mg_C binding site. Mn(II) binds at a stoichiometry of eight ions per enzyme octamer. The X- and Q-band EPR spectra reflect a single type of binding site with rhombic symmetry, and are consistent with oxygen and/or nitrogen ligands. The addition of unlabelled or [1-^{13}C]-labelled ALA does not significantly affect the Mn(II) EPR spectra. Thus, the Mg_C binding sites are distant from each other and also distant from the active sites.

Mg_C also has a profound effect on the quaternary structure of the protein. *E. coli* PBGS octamers dissociate into hexamers, tetramers, and dimers during native gel electrophoresis. Preincubation of the protein in EDTA potentiates the dissociation while preincubation of the protein in Mg(II) or Mn(II) and/or ALA hinders electrophoretic dissociation. Some PBGS have evolved away from Mg_C without compromising the integrity of the octamer (see Table II).

THE EVOLUTION OF METAL ION USE BY PBGS

Why is it that the PBGS of eukaryotic photosynthetic organisms lack the looser binding Zn_B present in animal, fungal, and most bacterial PBGS? The answer may derive from the thermodynamic challenge of surrounding a Mg(II) with nitrogen ligands, as is seen in chlorophyll. Insertion of Mg(II) into protoporphyrin IX is the first committed step in chlorophyll biosynthesis. The thermodynamic equilibrium between Zn-protoporphyrin IX and Mg-protoporphyrin IX favors the Zn(II) complex by about a factor of 30,000 (Baum and Plane, 1966). Insertion of Zn(II) into protoporphyrin IX is spontaneous. The insertion of Mg(II) into protoporphyrin IX, catalyzed by Mg-chelatase, has recently been shown to require the cleavage of ATP (Walker and Weinstein, 1994). It can be argued that these thermodynamic factors require that free Zn(II) be omitted from the chloroplast. In plants, all chlorophyll biosynthesis steps from glutamate through protoporphyrin IX occur in the chloroplast (Smith, 1988). The Zn_B site of PBGS ($K_d = 5\,\mu M$) requires free Zn(II) in the micromolar range. The absence of μM free Zn(II) may have driven the evolution of PBGS to use Mg_B as Mg(II) is abundant in chloroplasts. The most recent data indicating the presence of a Mg_B site

in PBGS from *Rhodobacter capsulatus* suggests that this organism may also have very little free Zn(II).

What about Zn_A vs. Mg_A? Regardless of the mechanism, during catalysis Zn_A probably undergoes ligand exchange. Ligand exchange at Zn(II) is often associated with its flexible coordination sphere and a change in coordination number. Mg(II) is ill-prepared to take on this role, no matter what the ligands are (Bock *et al.*, 1995). Consequently, we find it hard to envisage Mg(II) playing the role of Zn_A. Lead inhibition data on PBGS from both plants and animals (Hampp *et al.*, 1974) caused us to propose that all PBGS contain and utilize a Zn_A. Binding of Zn_A to mammalian PBGS is much tighter than 0.1 μM and might be tighter still for a plant protein. It is notable that plant PBGS has not yet been purified in large enough quantities to determine the intrinsic metal ion content. Although preliminary kinetic investigations show no apparent requirement for Zn(II) or inhibition by 1,10-phenanthroline (Boese *et al.*, 1991), there are examples in the literature of enzyme-bound Zn(II) which is coordinated so tightly that it takes a week of dialysis against 1,10-phenanthroline to prepare apo-protein (Liu *et al.*, 1993).

Sequence analysis suggests Mg_C to be common to all bacterial and photosynthetic organisms. The reasons for this remain elusive.

LEAD INHIBITION OF PBGS

The predominant importance of PBGS to human health is the inhibition of catalytic activity by trace levels of the environmental pollutant lead (e.g., Hernberg and Nikkanen, 1970). Enzyme inhibition causes increased levels of the substrate ALA which is a reactive molecule and a structural analog of γ-aminobutyric acid. This structural similarity may account for the neurological abnormalities associated with low level lead poisoning (retardation in children, neurosis in adults). In mammalian PBGS lead inhibition is a direct consequence of substitution of lead for the catalytic Zn(II) (Jaffe *et al.*, 1991). Most recently the mammalian PBGS protein was shown to perform a second function as an accessory protein in the ATP-dependent ubiquitin protein degradation pathway (Guo *et al.*, 1994), and similar experiments with *E. coli* PBGS show that this activity is shared by the bacterial protein (Etlinger, J., personal communication). This function, however, is not inhibited by lead.

Measurement of blood Pb(II)-inhibited PBGS is the most sensitive indicator of low-level Pb intoxication, but problems with the assay have diminished its use. Pb(II) has been identified as a slow-binding inhibitor of PBGS, whose activity can change up to sixfold during an hour-long clinical assay of Pb-contaminated blood (Jaffe *et al.*, 1991). Furthermore, Pb(II)-inhibited enzyme activity is profoundly affected by the presence of serum proteins which can act to buffer Zn(II), or by DTT which buffers both Zn(II) and Pb(II) (Gnonlonfoun *et al.*, 1991; Masuoka *et al.*, 1993). However, when assay components are scrupulously controlled and the enzyme is allowed to reach a steady-state rate, kinetic data on purified PBGS show that Pb(II) is a competitive inhibitor relative to Zn(II). Because Zn_A alone is required for mammalian PBGS activity, it follows that Pb(II) inhibits by replacing Zn_A. It is noted, however, that Pb(II) can bind at the Zn_B site without inhibiting mammalian PBGS (Dent *et al.*, 1990).

In an effort to determine where Pb(II) interferes with catalysis, we attempted to use ^{13}C NMR to observe partial reaction intermediates from [4-^{13}C]ALA bound to Pb(II)-PBGS. This proved impossible. The data indicate that neither A-side ALA nor product have any significant affinity for the Pb(II) form of the enzyme. This is consistent with Pb(II) replacing Zn_A which is required for substrate and product binding (see Fig. 3). A small amount of residual activity in a stoichiometric mixture of labelled ALA and Pb-PBGS results in NMR spectra of free product (unpublished data).

THE STRUCTURE OF PBGS

PBGS presents a unique and challenging structural puzzle. Each active site (equivalent to a homodimer) binds two ALA molecules whose chemical fates are distinctly different. For mammalian, yeast, and some bacterial PBGS, each active site also binds two Zn(II) whose coordination geometries are distinctly different (Zn_A and Zn_B). How can the same amino acid sequence (e.g., monomer) bind both Zn(II) and ALA, each in one of two very different ways? We and others have proposed that the binding of the four Zn_A causes a conformational change in the dimer which leads to the necessary asymmetry (Dent *et al.*, 1990; Jaffe *et al.*, 1992). Alternatively, the arrangement of the monomers could dictate formation of asymmetric dimers.

The enzyme aldolase may provide a clue to the symmetry of PBGS. Gurba *et al.* (1972) first cited similarities between the chemistry of the fructose-1,6-bisphosphate aldolase-catalyzed reaction and the PBGS-catalyzed reaction. The aldolases, which have substantial sequence homology, exist as either zinc metalloenzymes or as nonmetalloenzymes. The nonmetalloaldolases use a Schiff base intermediate analogous to P-side ALA forming a Schiff base linkage. The metalloaldolases use Zn(II); this is analogous to A-side ALA directly coordinating to Zn_A. X-ray crystal structures of both forms of aldolase show that the lysine which is involved in the Schiff base of the nonmetalloaldolase is homologous to a lysine which is a second-sphere ligand to Zn(II) in the metalloaldolase (Hester *et al.*, 1991; Sygusch, J., personal communication). On the basis of this precedent, it is possible that half of the PBGS subunits use the active-site lysine in a Schiff base linkage to P-side ALA while the other half use that lysine as a ligand to Zn_A. In both cases the lysine must be deprotonated. X-ray crystal structure determination and/or mutations to the lysine can address this question.

PBGS has long shunned attempts to determine its X-ray crystal structure, despite the relative ease of obtaining crystals (Shemin, 1976). Published electron micrographs show the 280,000 Dalton octameric mammalian protein to resemble a cube made up of eight spheres, one at each corner (Wu *et al.*, 1974). Small angle X-ray diffraction analysis suggests a structure which more closely resembles a cube made up of four parallel cylinders (Pilz *et al.*, 1988). Each cylinder is interpreted to be a dimer with high electron density in the center, suggesting that the active sites, with their required metal ions, are at the intersection of the monomers. We have used the four Zn_A/four Zn_B/eight Mg_C model of *E. coli* PBGS, where the Mg_C are at the interfaces of the dimers but not integral to an individual dimer to propose a quaternary structure model with four-fold symmetry (Jaffe *et al.*, 1995). These models, though intriguing, lack the molecular detail necessary to understand the relationships between the structure and function of PBGS.

ACKNOWLEDGMENT

This work was supported by NIH ES03654 (EKJ), CA06927 (ICR), and CA09035 (FCCC), and an appropriation from the Commonwealth of Pennsylvania. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute.

REFERENCES

Battle, A. M., and Stella, A. M. (1978). *Int. J. Biochem.* **9**, 861–864.
Baum, S. J., and Plane, R. A. (1966). *J. Am. Chem. Soc.* **88**, 910–913.
Beale, S. (1994). The sequence of *Chlamydomonus* PBGS, submitted to GenBank 1/95, Accession Number U19876.
Bevan, D. R., Bodlaender, P., and Shemin, D. (1980). *J. Biol. Chem.* **255**, 2030–2035.
Bishop, T. R., Frelin, L. P., and Boyer, S. H. (1986). *Nucleic Acids Res.* **14**, 0115.
Bishop, T. R., Hodes, Z. I., Frelin, L. P., and Boyer, S. H. (1989) *Nucleic Acids Res.* **17**, 1775.
Bock, C., Katz, A., and Glusker, J. P. (1995). *J. Am. Chem. Soc.*, in press.
Boese, Q. F., Spano, A. J., Li, J. M., and Timko, M. P. (1991). *J. Biol. Chem.* **266**, 17060–17066.
Bröckl, G., Berchtold, M., Behr, M., and Konig, H. (1992) *Gene* **119**, 151–152.
Chaudhry, A. G., and Jordan, P. M. (1976). *Biochem. Soc. Trans.* **4**, 760–761.
Chauhan, S., and O'Brian, M. R. (1993). *J. Bacteriol.* **175**, 7222–7227.
Christianson, D. W., and Lipscomb, W. N. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 7568–7572.
Clark, S. P. (1992). *CABIOS* **8**, 535–583.
Dent, A. J., Beyersmann, D., Block, C., and Hasnain, S. S. (1990). *Biochemistry* **29**, 7822–7828.
Echelard, Y., Dymetryszyn, J., Drolet, M., and Sasarman, A. (1988). *Mol. Gen. Genet.* **214**, 503–508.
Evans, J. N. S., Fagerness, P. E., Mackenzie, N. E., and Scott, A. I. (1985). *Magn. Reson. Chem.* **23**, 939–944.
Fabiano, E., and Goldin, B. T. (1991). *J. Chem. Soc. Perkin Trans.* **1**, 3371–3375.
Fukuda, H., Paredes, S. R., and Batlle, A. M. (1988). *Comp. Biochem. Physiol. B: Comp. Biochem.* **91**, 285–291.
Fukuda, H., Sopena de Kracoff, Y. E., Inigo, L. E., Paredes, S. R., Ferramola de Sancovich, A. M., Sancovich, H. A., and Batlle, A. M. (1990). *J. Enzyme Inhib.* **3**, 295–302.
Gibbs, P. N., and Jordan, P. M. (1986). *Biochem. J.* **236**, 447–451.
Gnonlonfoun, N., Filella, M., and Berthon, G. (1991). *J. Inorg. Biochem.* **42**,207–215.
Gribskov, M., and Devereux, J. (1991). *Sequence Analysis Primer*, Stockton Press, New York.
Guo, G. G., Gu, M., and Etlinger, J. D. (1994). *J. Biol. Chem.* **269**, 12399–12402.
Gurba, P. E., Sennett, R. E., and Kobes, R. D. (1972). *Arch. Biochem. Biophys.* **150**, 130–136.
Hampp, R., Kriebitzsch, C., and Ziegler, H. (1974). *Naturwissenschaften* **61**, 504–505.
Hansson, M., Rutberg, L., Schroder, I., and Hederstedt, L. (1991). *J. Bacteriol.* **173**, 2590–2599.
Hernberg, S., and Nikkanen, J. (1970). *The Lancet*, January 10, p. 63–64.
Hester, G., Brenner-Holzach, O., Rossi, F. A., Struck-Donatz, M., Winterhalter, K. H., Smit, J. D., and Piontek, K. (1991). *FEBS Lett.* **292**, 237–242.
Indest, K., and Biel, A. J. (1994). Genbank Accession Number RCU14593.
Jaffe, E. K. (1993). *Comments Inorg. Chem.* **15**, 67–93.
Jaffe, E. K., and Hanes, D. (1986). *J. Biol. Chem.* **261**, 9348–9353.

Jaffe, E. K., and Markham, G. D. (1987). [published erratum appears in *Biochemistry* **26**, 8030, 1987]. *Biochemistry* **26**, 4258–4264.
Jaffe, E. K., and Markham, G. D. (1988). *Biochemistry* **27**, 4475–4481.
Jaffe, E. K., Salowe, S. P., Chen, N. T., and DeHaven, P. A. (1984) *J. Biol. Chem.* **259**, 5032–5036.
Jaffe, E. K., Markham, G. D., and Rajagopalan, J. S. (1990). *Biochemistry* **29**, 8345–8350.
Jaffe, E. K., Bagla, S., and Michini, P. A. (1991). *Biol. Trace Element Res.* **28**, 223–231.
Jaffe, E. K., Abrams, W. R., Kaempfen, K. X., and Harris, K. A. (1992). *Bioichemistry* **31**, 2113–2123.
Jaffe, E. K., Volin, M., Myers, C. B., and Abrams, W. R. (1994). *Biochemistry* **33**, 11554–11562.
Jaffe, E. K., Ali, S., Mitchell, L. W., Taylor, K. M., Volin, M., and Markham, G. D. (1995). *Biochemistry* **34**, 244–251.
Jones, M. C., Jenkins, J. M., Smith, A. G., and Howe, C. J. (1994). *Plant Mol. Biol.* **24**, 435–448.
Jordan, P. (1990). In *Biosynthesis of Heme and Chlorophylls* (Dailey, H. A., ed.), McGraw-Hill, New York, pp. 55–121.
Jordan, P. M. (1991). In *New Comprehensive Biochemistry*, Vol. 19 (Neuberger, A., and Van Deenan, L. L. M., eds.), Elsevier, Amsterdam.
Jordan, P. M., and Seehra, J. S. (1980). *FEBS Lett.* **114**, 283–286.
Kaczor, C. M., Smith, M. W., Sangwan, I., and O'Brian, M. R. (1994). *Plant Physiol.* **104**, 1411–1417.
Li, J. M., Russell, C. S., and Cosloy, S. D. (1989). *Gene* **75**, 177–184.
Liedgens, W., Lutz, C., and Schneider, H. A. (1983). *Eur. J. Biochem.* **135**, 75–79.
Lingner, B., and Kleinschmidt, T. (1983). *Biosciences* **38**, 1059–1061.
Liu, J., Lin, S., Blochet, J. E., Pegolet, M., and Lapoint, J. (1993). *Biochemistry* **32**, 11390–11396.
Maralihalli, G. B., Rao, S. R., and Bhagwat, A. S. (1985). *Phytochemistry* **24**, 2533–2536.
Markham, G. D., Myers, C. B., Harris, K. A., Jr., Volin, M., and Jaffe, E. K. (1993). *Protein Sci.* **2**, 71–79.
Masuoka, J., Hegenauer, J., Van Dyke, B. R., and Saltman, P. (1993). *J. Biol. Chem.* **268**, 21533–21537.
Mitchell, L. W., Volin, M., and Jaffe, E. K. (1995) *J. Biol. Chem.*, submitted for publication.
Mitchell, L. W., and Jaffe, E. K. (1993). *Arch. Biochem. Biophys.* **300**, 169–177.
Myers, A. M., Crivellone, M. D., Koerner, T. J., and Tzagoloff, A. (1987). *J. Biol. Chem.* **262**, 16822–16829.
Nandi, D. L. (1978). *Biosciences* **33**, 799–802.
Nandi, D. L., and Shemin, D. (1968). *J. Biol. Chem.* **243**, 1236–1242.
Neier, R. (1996). In *Advances in Nitrogen Heterocycles*, Vol. 2 (Christopher, J. Moody, ed.), JAI Press, Greenwich, Connecticut.
Pilz, I., Schwarz, E., Vuga, M., and Beyersmann, D. (1988). *Biol. Chem. Hoppe-Seyler* **269**, 1099–1103.
Polking, G. F., Hannapel, D. J., and Gladon, R. J. (1994). Submission to GenBank 3/94, Accession Number L31367.
Rost, B., and Sander, C. (1993). *J. Mol. Biol.* **232**, 584–599.
Rost, B., and Sander, C. (1994). *Proteins* **19**, 55–72.
Rost, B., Sander, C., and Schneider, R. (1994). *CABIOS* **10**, 53–60.
Schaumburg, A., Schneider-Poetsch, H. J. A. W., and Eckerskorn, C. (1992). *Z. Naturforsch. Teil C J. Biosci.* **47**, 77–84.
Seehra, J. S., and Jordan, P. M. (1981). *Eur. J. Biochem.* **113**, 435–446.
Shemin, D. (1972). *The Enzymes*, 3rd edn. (Boyer, P. D., ed.), Academic Press, New York, pp. 323–337.
Shemin, D. (1976). *J. Biochem.* **79**, 37P–38P.
Shemin, D., and Russell, C. S. (1953). *J. Am. Chem. Soc.* **75**, 4873.
Smith, A. G. (1988). *Biochem. J.* **249**, 423–428.
Sollbach, M., and Schneider-Poetsch, H. J. A. W. (1993). Submission to GenBank 9/93, Accession Number X75043.
Spencer, P., and Jordan, P. M. (1993). *Biochem. J.* **290**, 279–287.
Spencer, P., and Jordan, P. M. (1994). *Biochem. J.* **300**, 373–381.
Tsukamoto, I., Yoshinaga, T., and Sano, S. (1975). *Biochem. Biophys. Res. Commun.* **67**, 294–300.
Tsukamoto, I., Yoshinaga, T., and Sano, S. (1979). *Biochim. Biophys. Acta* **570**, 167–178.
Walker, C. J., and Weinstein, J. D. (1994). *Biochem. J.* **299**, 277–284.
Wetmur, J. G., Bishop, D. F., Cantelmo, C., and Desnick, R. J. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 7703–7707.
Wu, W. H., Shemin, D., Richards, K. E., and Williams, R. C. (1974). *Proc. Natl. Acad. Sci. USA* **71**, 1767–1770.