Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# The porphobilinogen synthase family of metalloenzymes

**Eileen K. Jaffe**

Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA

Correspondence e-mail: ek_jaffe@fccc.edu

Eileen K. Jaffe became interested in the interactions of metals and enzymes as an undergraduate not long after Gunther Eichhorn coined the term inorganic biochemistry. She did her graduate work at the University of Pennsylvania with Mildred Cohn, a pioneer in the use of magnetic resonance techniques for studying enzymes. At PENN, Dr Jaffe discovered and characterized the metal-ion dependence of various kinases for ATP$\alpha$S and ATP$\beta$S, which are the diastereomeric phosphorothioate analogs of ATP. She carried out postdoctoral studies at Harvard University with Jeremy Knowles, who then wore the hat of a suburb enzymologist. Following postgraduate work, Dr Jaffe settled in Philadelphia and currently directs a laboratory at the Institute for Cancer Research, Fox Chase Cancer Center. She chose to focus on porphobilinogen synthase because its sensitivity to lead was so different from the more garden-variety zinc-containing metalloenzymes. Dr Jaffe has addressed the rich and varied bioinorganic chemistry of porphobilinogen synthase with a wide variety of biochemical and biophysical tools. For example, significant mechanistic information has come from the application of $^{13}$C NMR to this almost 300 kDa protein. Dr Jaffe has most recently taken up the design of artificial genes for the purpose of expressing massive quantities of the mammalian and plant proteins for structural and mechanistic studies.



Received 9 August 1999
Accepted 19 November 1999

The porphobilinogen synthase (PBGS) family of enzymes catalyzes the first common step in the biosynthesis of the essential tetrapyrroles such as chlorophyll and porphyrin. Although PBGSs are highly conserved at all four levels of protein structure, there is considerable diversity in the use of divalent cations for the catalytically essential and allosteric roles. Assumptions regarding commonalities among the PBGS proteins coupled with the diversity of usage of metal ions has led to a confused literature. The recent publication of crystal structures for three PBGS proteins coupled with more than 50 individual PBGS sequences allows an evaluation of these assumptions. This topical review focuses on the usage of metals by the PBGS family of proteins. It raises doubt concerning a dogma that there has been an evolutionary shift between $Zn^{II}$ and $Mg^{II}$ at one or more of the divalent metal-binding sites. It also raises the possibility that there may be up to four specific divalent metal ion-binding sites, each serving a unique function that can be alternatively filled by amino acids in some of the PBGSs.

## 1. Introduction

Tetrapyrrole cofactors are essential to archea, bacteria and eukaryotes and comprise essential physiological components such as heme and chlorophyll (see Fig. 1*a*). Most organisms perform *de novo* biosynthesis of at least some of their required tetrapyrroles. The first common step in tetrapyrrole biosynthesis is the asymmetric condensation of two molecules of 5-aminolevulinic acid (ALA) catalyzed by the enzyme porphobilinogen synthase (PBGS; E.C. 4.2.1.24; Shemin & Russell, 1953; Shemin, 1972; Jaffe, 1995; Shoolingin-Jordan, 1998). This reaction is illustrated and described in Fig. 1(*b*). The substrate that becomes the acetyl-containing half of porphobilinogen is called A-side ALA; the propionyl-containing half of porphobilinogen derives from P-side ALA. The enzyme-catalyzed reaction involves multiple bond-breaking and bond-making events, the order of which has yet to be determined. PBGS is large; it is generally a 280–300 kDa protein composed of eight identical subunits. The literature on this enzyme extends back to the early 1950s and includes some of David Shemin's landmark metabolic studies using stable isotopes to follow the steps of heme biosynthesis (Shemin, 1979). Nearly 50 years has not been sufficient to sort out the mysteries of the PBGS protein and the reaction it catalyzes. In fact, documented experimental observations, particularly with regard to the metalloprotein nature of PBGS, continue to be interpreted in a non-uniform fashion that is confusing. The attempted focus of this work is to bring some clarity to the existing PBGS literature and to extend recent X-ray crystallographic results to the general family of PBGS proteins.

## 2. What's in a name?

One simple factor confounding the literature is that PBGS is also referred to by its pre-IUPAC designation, δ-aminolevulinic acid dehydratase (ALAD). The latter name is found in the literature under more than a dozen different spellings, misspellings or abbreviations. The automatic text readers of abstract services have read PBGS as PEGS, thus transfiguring a complex metalloprotein into a common polymer. Table 1 lists the various published designations for the enzyme and is intended to assist anyone who wishes to tackle the PBGS literature.

**Table 1**
The published names and abbreviations used for porphobilinogen synthase (E.C. 4.2.1.24).

| |
|---|
| Porphobilinogen synthase (PBGS and mis-indexed as PEGS) |
| Porphobilinogen synthetase (an incorrect name, as ATP is not utilized) |
| δ-Aminolevulinate dehydratase; δ-aminolevulinic acid dehydratase (δ is also seen as 'delta') |
| δ-Aminolaevulinic acid dehydratase; δ-aminolaevulinate dehydratase |
| 5-Aminolevulinate dehydratase; 5-aminolevulinic acid dehydratase |
| 5-Aminolaevulinic acid dehydratase; 5-aminolaevulinate dehydratase |
| Abbreviations: ALAD, ALADH, δ-ALAD |
| 5-Levulinic acid dehydratase (an incorrect name) |
| γ-Aminolevulinic acid dehydratase (an incorrect name) |
| δ-Aminolevulinic acid dehydrase (an incorrect name) |
| δ-Aminolevulinic acid dehydrogenase (an incorrect name) |

## 3. How many metal ions does it take to make a pyrrole?

A second and less trivial confounding factor is the diversity of metal-ion utilization by the PBGS family of proteins. One general dogma is that the PBGS enzymes fall into two main classes: those that require $Zn^{II}$ and those that require $Mg^{II}$. The assumption is often made that the evolution of the $Mg^{II}$-requiring PBGSs to or from the $Zn^{II}$-requiring PBGSs can be explained by considering a 'metal-binding region of the protein sequence' identified with the elucidation of the first PBGS sequence (Wetmur *et al.*, 1986) and illustrated in Fig. 2. In general, for any one protein, this sequence is either rich in cysteine and histidine residues, which are good ligands for zinc, or it is rich in aspartic acid residues, which are good ligands for magnesium. These amino acids are assumed to form the binding residues for a catalytically essential metal ion, either $Zn^{II}$ or $Mg^{II}$. These two metal ions have been assumed to have evolved to replace each other in function. For the $Zn^{II}$-requiring PBGS, it is assumed that the exquisitely sensitive inhibition by the environmental toxin $Pb^{II}$ is caused by a substitution of $Pb^{II}$ for $Zn^{II}$ bound to the cysteines of this region of protein sequence (Warren *et al.*, 1998). One goal of this review is to critically evaluate the dogmas and assumptions described in this paragraph with the aid of recently published X-ray crystal structures.

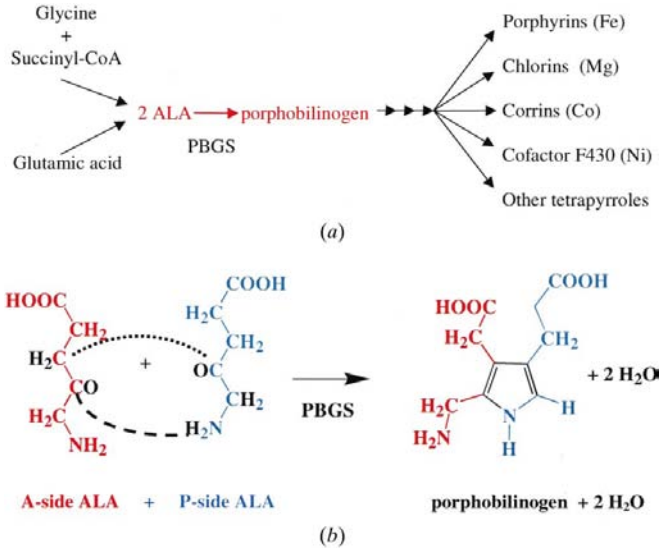



**Figure 1**
(*a*) Tetrapyrrole biosynthesis is an ancient and essential pathway. The first common intermediate is 5-aminolevulinic acid (ALA). One molecule of ALA condenses with a second one to form the monopyrrole porphobilinogen in a reaction catalyzed by the metalloenzyme porphobilinogen synthase (PBGS). The PBGS-catalyzed reaction is the first common step in the biosynthesis of the tetrapyrrole cofactors. (*b*) The PBGS-catalyzed reaction involves the formation of one carbon–carbon bond (dotted line) and one carbon–nitrogen bond (dashed line), as well as the loss of four protons and two O atoms (shown in black). The only known reaction intermediate is a Schiff base formed between an active-site lysine and C4 of the ALA molecule whose fate is the propionyl-containing half of porphobilinogen (P-side ALA, shown in turquoise). For mammalian PBGS, the $Zn^{II}$ is implicated in the binding and reactivity of the ALA molecule whose fate is the acetyl-containing half of porphobilinogen (A-side ALA, shown in red). Little else is known about the order of chemical events or the active-site residues that mediate the chemical transformations.

Over the past two decades, biochemical, kinetic and metal-binding data from several different PBGS suggested the existence of a variety of different roles for divalent metal ions in PBGS structure and function. Based on these studies, the PBGSs have been sorted into types based on the numbers and kinds of metal ions utilized. Table 2 sorts the PBGSs by their usage of divalent metal ions and also indicates whether or not the activities respond to monovalent metal ions. Table 2 is based partly on experimental observations and is partly predictive. The unproven (predictive) components are italicized. A prevailing assumption in the synthesis of Table 2 is that the PBGS proteins are more alike than they are different from each other. This assumption comes in part from the high degree of protein-sequence conservation illustrated in Fig. 3.

Table 2 also includes an indication of whether or not the activity of a given PBGS can be stimulated by monovalent cations such as $K^{I}$. Little is known about the specific role of monovalent cations, their location in the protein or the extent

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mammalian PBGS | **C** | D | V | **C** | L | **C** | P | Y | T | S | **H** | G | **H** | **C** | G |
| Plant PBGS | T | **D** | V | A | L | **D** | P | Y | S | S | **D** | G | H | **D** | G |

**Figure 2**
Portions of the sequences of human and pea PBGS are used to illustrate a metal-binding region that has been credited for an evolution of the $Zn^{II}$-requiring PBGSs (sequences rich in cysteine and histidine shaded grey) to or from the $Mg^{II}$-requiring PBGSs (sequences rich in aspartic acids shown as white on black).

**Table 2**
Sorting of the PBGS proteins by metal-ion utilization.

| | | Metal ions required for full activity per homooctamer† | | | | |
|---|---|---|---|---|---|---|
| Type | Species | Site A | Site B | Site C | Total divalents | Monovalent requirement |
| I | *Homo sapiens* | 4 Zn | 4 Zn | Absent | 8 | No |
| ? | *S. cerevisiae* | *8 Zn* | *8 Zn* | Absent | 16 | No |
| II | *E. coli* | *4 Zn* | *4 Zn* | 8 Mg | 16 | No |
| III | *P. sativum* | *4 Mg* | *4 Mg* | *8 Mg* | 16 | No |
| IV | *B. japonicum* | 4 Mg | | 8 Mg | 12 | Yes |
| V | *P. aeruginosa* | Absent | Absent | 4 Mg | 4 | Yes |

† Unproven hypothetical distributions are listed in italics.

to which they are important to the PBGSs of various phyla. For PBGS from *Bradyrhizobium japonicum*, the role of the monovalent ion is to downshift an essential p$K_a$ by about one pH unit (from 7.5 to 6.5), so that at neutral pH the protein is far more active upon addition of monovalent cations (Petrovich *et al.*, 1996). The monovalent cation effect was first seen as a seriously confounding buffer dependence in pH-rate profiles as has commonly been the case (Woehl & Dunn, 1995). The location of the monovalent ion(s) is not apparent from existing crystal structures and, therefore, this work focuses predominantly on the divalent metal-ion sites.

The nomenclature used for the metal-ion sites is an additional complicating factor in the PBGS literature. The divalent metal-ion sites listed in Table 2 follow the nomenclature introduced on the basis of an EXAFS analysis of a $Zn^{II}$-containing PBGS (Dent *et al.*, 1990). The term ZnA was used to describe a pentacoordinate $Zn^{II}$ with one cysteine and several other oxygen- and nitrogen-containing ligands. The designation ZnB was used in the EXAFS study to describe a putative tetrahedral site with all sulfur-containing ligands. In Table 2, these are called metal site A and metal site B, respectively. Metal site C refers to an allosteric site whose nomenclature was added when such a site became apparent from studies of *Escherichia coli* PBGS (Mitchell & Jaffe, 1993; Jaffe *et al.*, 1995). Naming a potential 'catalytic' metal site as A and an 'allosteric site' as C is yet another unfortunate historical consequence. In an effort to circumvent a potential confusion of metal site A with A-side ALA (see Fig. 1), a simultaneous report on *E. coli* PBGS suggested the use of $\alpha$ and $\beta$ in place of A and B (Spencer & Jordan, 1993). The intention was for $\alpha$ to correspond to A and for $\beta$ to correspond to B. Unfortunately, the metal site referred to as $\alpha$ corresponds to the one originally identified as B (where the $Zn^{II}$ is rich in sulfur ligands) and the metal site referred to as $\beta$ most often corresponds to the one that was identified as C. This further confounds the literature and attempts will be made herein to clarify the confusion.

In general, the literature which uses the $\alpha$ and $\beta$ nomenclature considers that there are only *two* divalent metal-binding sites in the family of PBGS proteins. Discordantly, the literature using the A, B, C nomenclature considers that there are at least *three* divalent metal-binding sites in the family of PBGS proteins. Since the available X-ray crystal structures prove the existence of the three different divalent metal ion-binding sites that correspond to ZnA, ZnB and MgC, this work will use the latter nomenclature to describe the PBGS family of proteins. There are chemical modification data that implicate ZnA in the binding and reactivity of A-side ALA (Jaffe & Hanes, 1986; Jaffe *et al.*, 1992) and similar data that place ZnB in the vicinity of P-side ALA (Jaffe *et al.*, 1994). The MgC site is strictly allosteric and was predicted to lie at the subunit interface (Jaffe *et al.*, 1995). The crystallographic data support some but not all of these predictions. In particular, the crystal structures place ZnB nearer to the presumed A-side ALA-binding site.

## 4. The structure of the PBGS protein

Prior to description of the divalent metal ion-binding sites seen in PBGS, it is essential to lay out the general structure of the protein. As shown in Fig. 3, the core sequence of about 330 amino acids is highly conserved among 53 sequences found in the sequence databases. Furthermore, extensive PSI-BLAST (Altschul *et al.*, 1997) searches bring up only PBGS proteins, thus defining PBGS as a protein family not obviously related to any other proteins. Prior to the determination of any PBGS crystal structures, a secondary-structure prediction showed alternating helices and sheets that were predictive of a TIM-like protein (Rost & Sander, 1994; Jaffe, 1995). The predicted structure was a good match for those subsequently experimentally determined. All the published structures show an octamer with extensive inter-subunit interactions (Erskine *et al.*, 1997; Erskine, Newbold *et al.*, 1999; Erskine, Norton *et al.*, 1999; Frankenberg, Erskine *et al.*, 1999). Each monomer contains an $\alpha\beta$-barrel domain where known active-site residues are located. Each monomer also contains an N-terminal arm which is involved in extensive interactions with neighboring subunits. The arm wraps around the barrel of one neighbor with whom there are barrel–barrel contacts. With respect to a different neighbor, the arm of one monomer sits at the base of the barrel of a second monomer. There is a mobile portion of

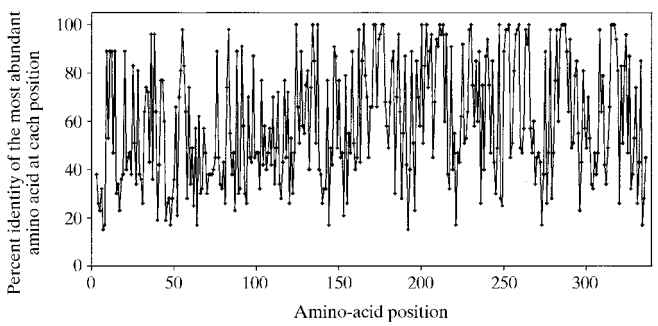



**Figure 3**
The PBGS sequence is highly conserved. The figure is based on 53 PBGS sequences found in the databases. For any one position in the 330 amino-acid PBGS core, the figure illustrates the numbers of sequences for which this position is like the predominant amino acid found at that site.

the sequence (~11 amino acids) which is seen in some structures to be disordered and in some structures to be an ordered lid which can isolate the active site from bulk solvent. Fig. 4 uses the *Saccharomyces cerevisiae* (yeast) protein to illustrate the general tertiary and quaternary structure of PBGS. Similar interactions would be seen if illustrated with the structures of the *Pseudomonas aeruginosa* or *E. coli* enzymes. The programs *RasMol*2 (copyright Roger Sayle, 1992–1999) and *MOLSCRIPT* (Kraulis, 1991) were used to prepare the structure figures.

## 5. PBGS crystal structures provide evidence for three divalent metal ion-binding sites

Most of the structural analysis presented herein is based on the crystal structures of the PBGS from yeast (Erskine *et al.*, 1997) and from *P. aeruginosa* for which coordinates have been made available (Frankenberg, Erskine *et al.*, 1999). The respective PDB entries are 1aw5, which is at 2.3 Å resolution, and 1b4k, which is a 1.67 Å resolution structure. Yeast PBGS is one of the $Zn^{II}$-requiring enzymes (Senior *et al.*, 1996) and

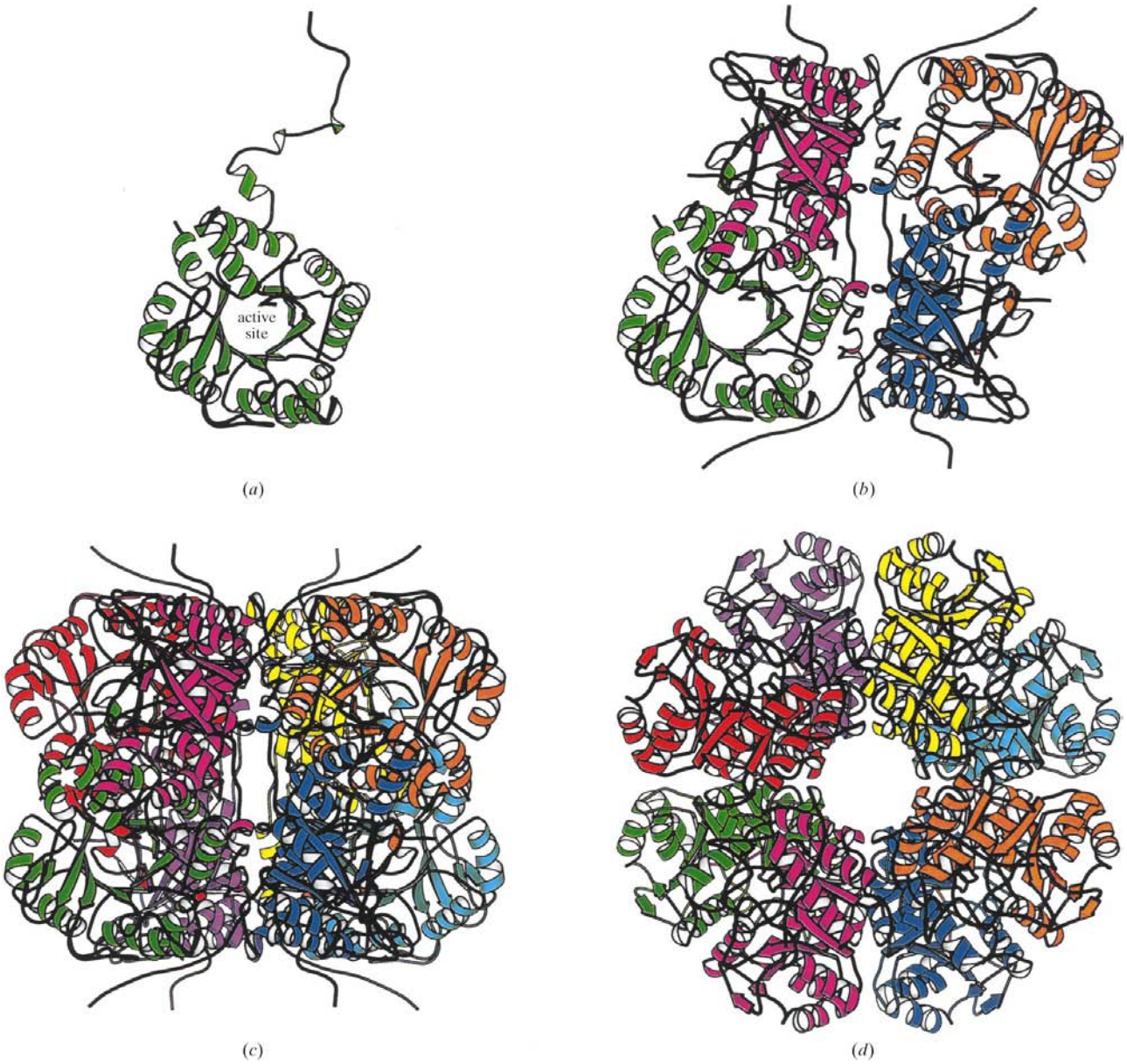



**Figure 4**
Each PBGS monomer is a TIM-like barrel with a long N-terminal arm as illustrated in (*a*) for yeast PBGS (law5). This monomer is shown in green in all figures and all monomers are missing electron density from 11 amino acids that are poised to act as an active-site lid. The lid, when ordered, sits atop the active site and isolates the center of the barrel from bulk solvent. (*b*) A cutaway half of the octamer, from which it can be seen that the arm regions play an extensive role in holding together the octamer. For instance, the arm of the magenta monomer wraps around the barrel of the green monomer and lies against the base of the barrel of the dark blue monomer. Through these arm interactions it is possible to imagine subunit–subunit communication through two different kinds of dimers (magenta–green or magenta–dark blue). The octamer is shown in (*c*), keeping the orientation of the green monomer constant. (*d*) A 90° vertical rotation of the octamer relative to (*c*).

the structure reveals two different kinds of $Zn^{II}$ sites, as illustrated in Fig. 5. *P. aeruginosa* PBGS activity responds to $Mg^{II}$ and its structure reveals one kind of $Mg^{II}$ site, which is also illustrated in Fig. 5 and is distant from either of the $Zn^{II}$ sites of the yeast protein. Thus, it is established that the PBGS family of proteins contains at least three different divalent metal-ion sites, as proposed previously (Jaffe, 1993; Mitchell & Jaffe, 1993). However, there are many aspects of the current three metal-ion model, presented in Table 2, which are not established. It remains to be established whether or not any PBGS exists that uses all three metal-ion sites in the same protein. It also remains to be established whether or not any of the PBGSs use $Mg^{II}$ in a catalytic role (*e.g.* in place of the $Zn^{II}$ seen in the yeast protein) or whether any PBGS uses $Zn^{II}$ in an allosteric role (*e.g.* in place of the $Mg^{II}$ seen in *P. aeruginosa* PBGS). The analysis presented herein addresses these questions. Finally, it remains to be established if three different metal sites are sufficient to describe the variety of divalent metal-ion positions that function in the PBGS-catalyzed reaction.

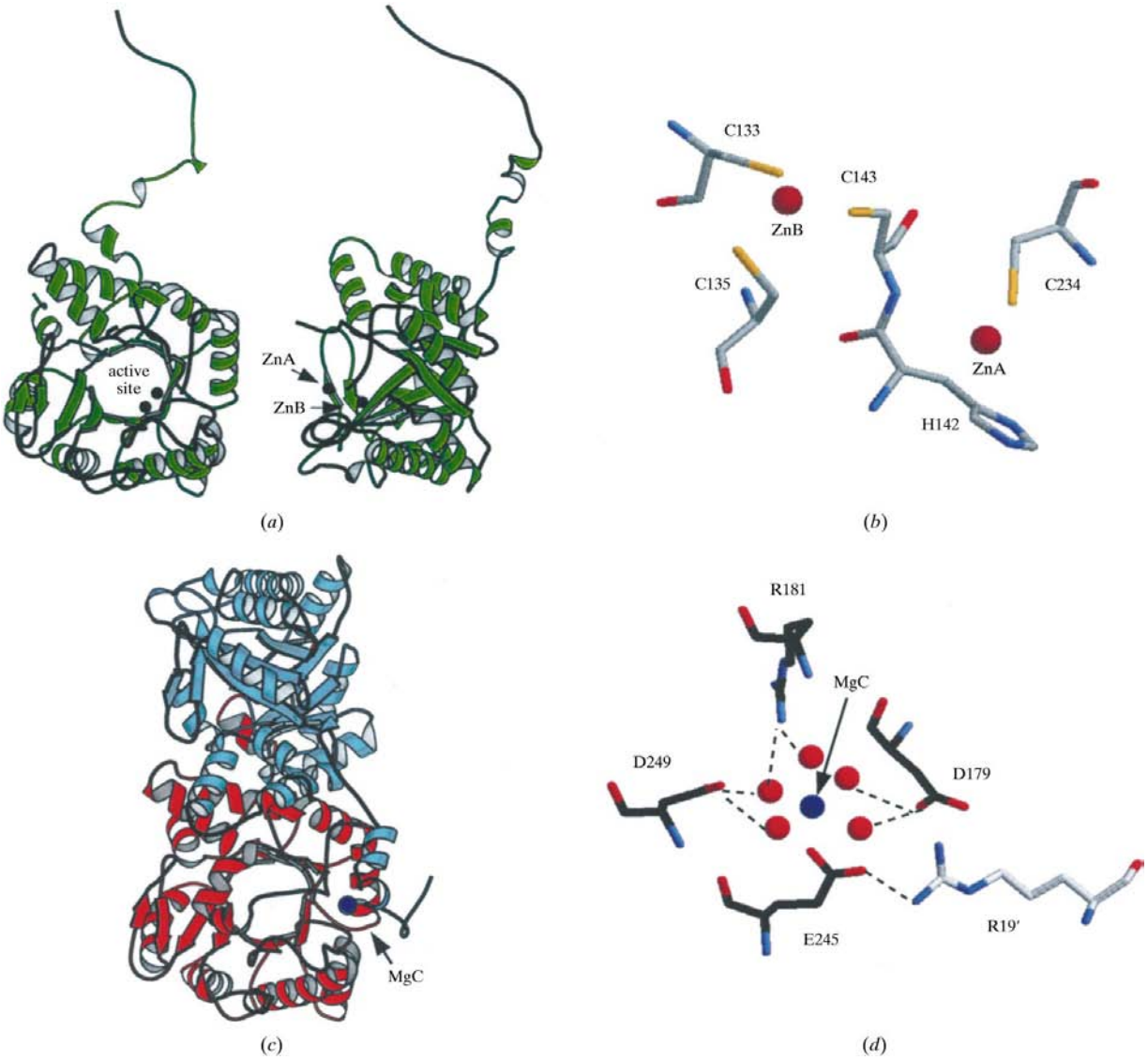



**Figure 5**
The published crystal structures prove the existence of three different types of divalent metal ion-binding sites. The yeast PBGS structure reveals two different $Zn^{II}$ environments in the vicinity of the active site as illustrated in (*a*). Two orthogonal orientations of the monomer are shown at a 90° rotation; the black balls depict the $Zn^{II}$. (*b*) The specific $Zn^{II}$ ligands. The $Zn^{II}$ are labeled ZnA and ZnB following the ZnA and ZnB nomenclature of an early EXAFS study (Dent *et al.*, 1990). In this structure, the active-site lid (residues 220–233) is disordered and one of the $Zn^{II}$ ligands is the penultimate ordered residue (Cys234). The other ligands all derive from the metal-binding region illustrated in Fig. 2. The *P. aeruginosa* PBGS reveals one kind of $Mg^{II}$-binding site that is shown in (*c*) to be distant from the $Zn^{II}$ sites of yeast PBGS. (*c*) A PBGS dimer; the $Mg^{II}$ (shown in dark blue) lies at an interface of the barrel region of the red monomer and the arm region of the cyan monomer. (*d*) The $Mg^{II}$ ligands which consist of one glutamic acid moiety (E345) and five water molecules, shown in red. These first coordination sphere ligands then contain extensive hydrogen bonds to other residues, D249, R181, D179 and R19′, the latter of which comes from the arm of the other monomer. The carbon atoms in black in (*d*) correspond to amino acids from the red subunit of (*c*). The carbon atoms in light grey correspond to the cyan subunit of (*c*). It is important to note that yeast PBGS does not contain the $Mg^{II}$ seen in *P. aeruginosa* PBGS, which in turn does not contain the $Zn^{II}$ seen in yeast PBGS (see Table 2).

PBGS is the only biological catalyst for which it has been proposed that a catalytically essential $Zn^{II}$ function evolved from or to a catalytically essential $Mg^{II}$ function. For quite some time this author resisted accepting this proposal on the basis of the inorganic chemistry of $Mg^{II}$ *versus* $Zn^{II}$ (Jaffe, 1993, 1995). The former is an alkaline-earth metal and the latter is a transition element. In fact, $Mg^{II}$ and $Zn^{II}$ are similar in size and charge. Neither is typically involved in redox chemistry and both can function as Lewis acids. Their coordination chemistries, however, are quite different (Glusker *et al.*, 1996). With the exception of chlorins, $Mg^{II}$ in proteins is normally situated in an octahedral coordination environment surrounded by six oxygen-containing (hard) ligands. In contrast, $Zn^{II}$ can be found in proteins with a variety of coordination geometries and ligands containing sulfur, nitrogen or oxygen (Lippard & Berg, 1994; Glusker *et al.*, 1996). In general, $Zn^{II}$ ions involved in catalysis contain predominantly nitrogen and oxygen ligands and their role in catalysis is often accompanied by a change in the coordination geometry (Vallee & Auld, 1995). More recent data show that $Zn^{II}$ can activate cysteine ligands to participate in catalysis (Myers *et al.*, 1994; Matthews & Goulding, 1997; Hightower & Fierke, 1999). Most importantly, $Zn^{II}$ has the special property of being able to dramatically depress the p$K_a$ of a directly bound hydroxyl or alcohol by up to 8 pH units (Kimura *et al.*, 1999). With regard to a $Mg^{II}$ to $Zn^{II}$ switch in PBGS based on the metal-binding region illustrated in Fig. 2, it is now established that some PBGSs have an absolute requirement for $Mg^{II}$ instead of $Zn^{II}$ (Petrovich *et al.*, 1996; Senior *et al.*, 1996) and that substrate alters the relative affinity of enzyme-bound $Mg^{II}$ for two different $Mg^{II}$-binding sites (Petrovich & Jaffe, 1997). Therefore, it becomes attractive (even to this author) to claim a direct replacement of function based on the sequence illustrated in Fig. 2. However appealing this proposal, it remains to be proven because $Mg^{II}$ has yet to be observed at this site. Mutagenesis studies have placed the human sequence of this region into the pea protein and demonstrated the introduction of a $Zn^{II}$ requirement into a protein that retains an allosteric response to $Mg^{II}$; similar work with *B. japonicum* PBGS also resulted in a $Zn^{II}$ requirement (Chauhan & O'Brian, 1995). However, it has yet to be established that insertion of the pea sequence of this region into the human protein will introduce a $Mg^{II}$ requirement.

In summary, Fig. 5(*b*) shows two $Zn^{II}$ seen in the yeast PBGS crystal structure. ZnA is coordinated to one histidine (H142) and one cysteine ligand (C234). The histidine corresponds to the second histidine of the metal-binding region shown in Fig. 2. The cysteine is not from this region but corresponds to one previously identified as a ZnA ligand from chemical modification studies (Jaffe *et al.*, 1992). The cysteine is adjacent to the disordered lid region of the protein (residues 220–233) that might provide additional $Zn^{II}$ ligands. Furthermore, the C4 carbonyl and C5 amino groups of A-side ALA have been proposed to provide two more ligands to ZnA (Jaffe, 1995). The ZnA site of yeast PBGS was observed to be at a low occupancy, consistent with the lack of substrate in the active site (Erskine *et al.*, 1997). The yeast crystal structure also shows ZnB, at higher occupancy, bound to three cysteines (C133, C135, C143) which all derive from the sequence illustrated in Fig. 2. The cysteines are arranged on the three corners of a tetrahedron, with the fourth corner vacant as if it might also interact with substrate. There are no ordered water molecules within 4.0 Å of ZnB. The most N-terminal fourth cysteine of this region of sequence, once proposed as a ZnB ligand (Jaffe, 1993), is quite distant from the metal-binding site.

The total stoichiometry of $Zn^{II}$ bound in the yeast PBGS crystals was not reported, though ZnA is cited as present at low occupancy (Erskine *et al.*, 1997). Therefore, it remains to be proven whether ZnA and ZnB can bind simultaneously to any one subunit. It had previously been suggested that the half-site reactivity of mammalian PBGS might be based on the mutually exclusive binding of either ZnA or ZnB to residues in the metal-binding region, where each $Zn^{II}$ is bound to alternate PBGS subunits (Dent *et al.*, 1990; Jaffe *et al.*, 1992). The crystal structure implies that this is not the case. Fig. 5(*b*) shows ZnA and ZnB bound to adjacent amino acids (H142 and C143), which provides a nice model for communication (cooperativity) between these two sites if indeed two $Zn^{II}$ ions can bind to the same subunit. The relative occupancy of the ZnA and ZnB sites may be altered in the presence of substrate or product, as was seen for the $Mg^{II}$ sites of *B. japonicum* PBGS (Petrovich & Jaffe, 1997).

The MgC seen in the *P. aeruginosa* PBGS structure (Fig. 5*d*) is not coordinated through the region of sequence shown in Fig. 2 (Frankenberg, Erskine *et al.*, 1999). In Fig. 5(*d*), it is shown that MgC binds five water molecules and E245. This E245 first coordination-sphere ligand derives from a region of sequence originally proposed to provide the MgC ligands for *E. coli* PBGS (Jaffe, 1993; Mitchell & Jaffe, 1993). It is important to stress that the MgC seen in *P. aeruginosa* PBGS is structurally and functionally distinct from either of the two $Zn^{II}$ seen in yeast PBGS. The MgC seen in *P. aeruginosa* PBGS lies at subunit interfaces, as had also been predicted based on studies of *E. coli* PBGS (Jaffe *et al.*, 1995). One additional prediction was that the PBGS octamer would contain four active sites located at the subunit interfaces (Jaffe *et al.*, 1995). Clearly, the active sites are not at subunit interfaces.

## 6. The structure of *P. aeruginosa* PBGS shows subunit asymmetry

Over the decades, evidence has been presented supporting the notion that each PBGS octamer contains only four functional active sites (Cheh & Neilands, 1973; Shemin, 1976; Bevan *et al.*, 1980; Seehra *et al.*, 1981; Seehra & Jordan, 1981; Lingner & Kleinschmidt, 1983; Jaffe & Hanes, 1986; Jaffe *et al.*, 1992). Another body of literature supports the notion that each PBGS octamer contains eight equivalent and functional active sites (Gibbs & Jordan, 1986; Spencer & Jordan, 1994; Cheung *et al.*, 1997; Erskine *et al.*, 1997). In some cases, these studies are reported on PBGS from the same organism (Spencer & Jordan, 1994; Mitchell *et al.*, 1995). The experimental basis for

the differing results, predominantly divided by the Atlantic Ocean (Evanston and Philadelphia *versus* Southampton and London), remains unclear.

The structure of the yeast PBGS (illustrated in Fig. 4) is that of a symmetric octamer, in part because the monomer was seen in the asymmetric unit of the crystallographic data. In contrast, the *P. aeruginosa* PBGS crystallographic data obtained at higher resolution show an asymmetric dimer as the fundamental structural unit (Frankenberg, Erskine *et al.*, 1999). In this case, the sequence-identical subunits do not have precisely the same structure. This dimer is illustrated in Fig. 6(*a*) and described in the figure legend. The barrel region of monomer *A* is seen associated with the $Mg^{II}$ and its active-site lid is ordered. The barrel domain of monomer *B* is not associated with $Mg^{II}$ and its active-site lid is disordered. Based on thermal factors, the arm of monomer *A* is part of the structural unit containing the barrel of monomer *B* and *vice versa*. The structural differences between monomers *A* and *B* are subtle, yet the analysis of Heinz shows a beautiful relationship between the presence of the $Mg^{II}$ ion and the closing of the active-site lid. The coordination of R181 as a second-sphere ligand to $Mg^{II}$ frees up D139, allowing the latter to hydrogen bond with K229, thus stabilizing the closed configuration of the lid. This is reproduced in Fig. 6(*b*).

## 7. The methodology used for a comparison of the metal-binding sites of yeast and *P. aeruginosa* PBGS

The two crystal structures provide an interesting background for investigating the validity of some assumptions laid forth earlier about the locations and roles of the metal ions of PBGS, since each is lacking the metal ions seen in the other. For comparative purposes, the program *MINAREA* (Falicov & Cohen, 1996) was used to align each of the monomers *A* and *B* of *P. aeruginosa* PBGS with the yeast PBGS monomer; the metal ions, active-site ligands and water molecules were included in the superimposition files. The program *RasMol*2 was then used to view and analyze these superimposed structures for the purpose of comparing the environments around the three observed divalent metal ions ZnA, ZnB, and MgC. To compare and group the 53 PBGS sequences in the regions of the observed metal ligands and make figures with only those regions of sequence, the sequence data were processed manually using the program Microsoft *EXCEL*. The three sites are presented below as MeA, MeB and MeC.

## 8. The MeA site: within the active site, but no place for $Mg^{II}$

Fig. 7(*a*) shows a superposition of the ZnA site of yeast PBGS and its two ligands with the identical amino acids of monomer *A* of *P. aeruginosa* PBGS. Only the backbone portions of H142 and its cognate Q138 from *P. aeruginosa* PBGS superimpose. The amide oxygen atom of the side chain of Q138 spatially conflicts with the position of ZnA. The location of T231 of *P. aeruginosa* PBGS is distant to that of its cognate C234, which serves as a ZnA ligand. Although the crystal-

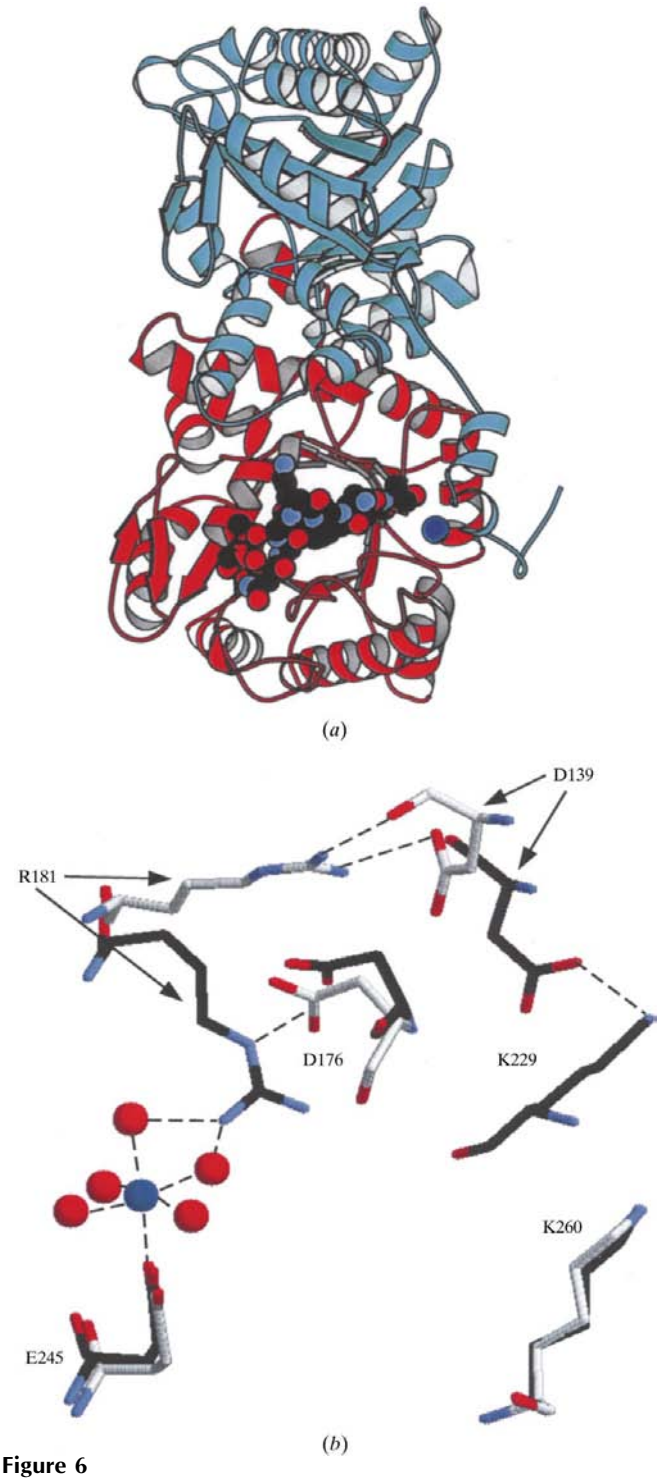
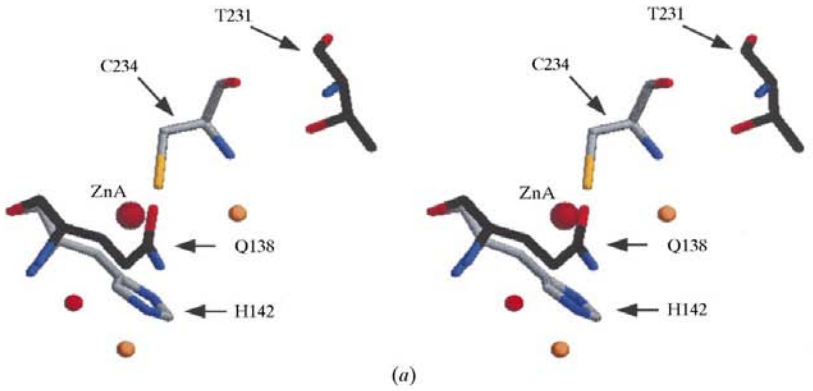
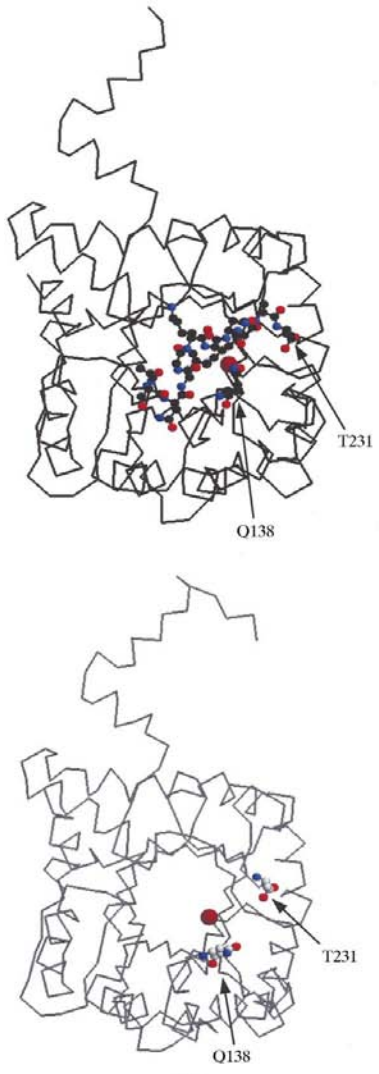



**Figure 6**
The dimer of *P. aeruginosa* PBGS (*a*) shows an asymmetry. Monomer *A*, shown in red, has generally lower thermal factors, the barrel region is associated with a $Mg^{II}$ ion (blue) and contains well defined electron density for the active-site lid (shown as space-fill with black carbon atoms and cpk-colored oxygen and nitrogen atoms). Monomer *B*, shown in cyan, has generally higher thermal factors, the barrel region is not associated with a $Mg^{II}$ and the active-site lid is disordered. (*b*) The relationship between the presence of $Mg^{II}$ (shown in blue) and the ordering of the active-site lid. In the absence of $Mg^{II}$ (carbon atoms in light grey), R181 is hydrogen bonded to D139 and there are no interactions with amino acids from the lid. In the presence of $Mg^{II}$ (carbon atoms in black), R181 is hydrogen bonded to two water ligands (shown in red) in the $Mg^{II}$ first coordination sphere and D139 is free to rotate and hydrogen bond to K229, which is part of the mobile lid.




(*a*)

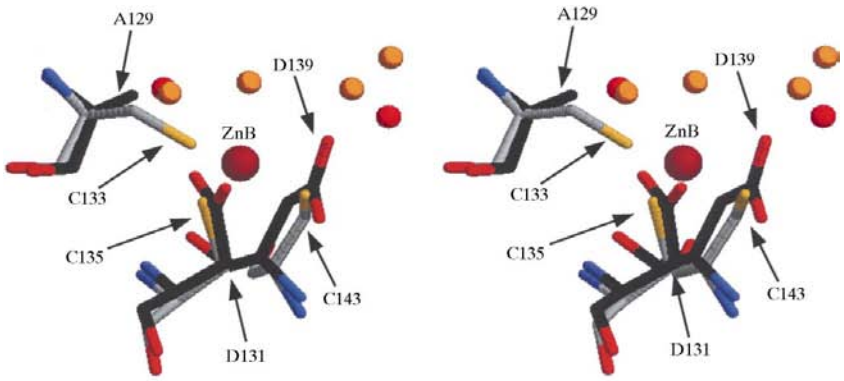



(*b*)

| *P. aeruginosa* numbering | 138 | 214 | 221 | 224 | 228 | 229 | 231 |
|---|---|---|---|---|---|---|---|
| Yeast numbering | 142 | 219 | 226 | 229 | 231 | 232 | 234 |
| *Mus & Rattus* | HG H C... F | RDAAQS | S PA | F G..... | D | R R | C YQ |
| *Homo* | HG H C... F | RDAAKS | S PA | F G..... | D | R R | C YQ |
| *Candida* | HG H C... F | RDAACS | S PS | N G..... | D | R K | C YQ |
| ***Saccharomyces*** | **HG H C... F** | **RDAACS** | **A PS** | **N G.....** | **D** | **R K** | **C YQ** |
| *Schizosaccharomyces* | HG H C... F | RAAANG | A PK | F G..... | D | R S | C YQ |
| *Chlamydia* | HG H D... F | RDALSS | H VT | S G..... | D | K | K Q YQ |
| *Chlamydia* | SG H D... F | RDALSS | H LQ | S G..... | D | K | R T YQ |
| Six plant species | DG H D... F | REALDS | N PR | F G..... | D | K | K T YQ |
| *Physcomitrella* | DG H D... S | REALDS | N PR | F G..... | D | K | K T YQ |
| *Chlamydomonas* | DG H D... F | RDALAS | A PK | P GQAHRRIPP | N | K | K T YQ |
| Two *Neisseria* species | HG Q D... F | RDAVGS | S GN | L G.....KA | D | K | K T YQ |
| *Bordella* | HG Q D... F | RDAVGS | A SN | L G.....KS | D | K | M A YQ |
| *Shewanella* | HG Q D... F | RDAVGS | A GN | L K.....GG | N | K | H S YQ |
| *Vibrio* | HG Q D... F | RDAVGS | S AN | L K.....GG | N | K | K N YQ |
| *Actinobacillus* | HG Q D... F | RDAVGS | S GN | L K.....GG | D | K | K T YQ |
| *Yersinia* | HG Q D... F | RDAIGS | S SN | L K.....GG | D | K | K T YQ |
| ***Pseudomonas*** | **HG Q D... F** | **RDAVGS** | **A SN** | **L G.....KG** | **N** | **K** | **A T YQ** |
| *Rhodobacter* | NG H D... F | RDAVGA | S SA | L K.....G | D | K | K T YQ |
| *Rhodobacter* | NG H D... F | RDAVGA | S GA | L K.....G | D | K | K T YQ |
| *Bradyrhizobium* | HG H D... F | RDAIGS | A KT | L T.....G | D | K | R T YQ |
| *Rhodopseudomonas* | HG H D... F | RDAIGS | A KA | L T.....G | D | K | R T YQ |
| *Caulobacter* | HG H D... Y | RDAIGS | A KL | S A.....GQG | D | K | K T XQ |
| *Rickettsia* | SG H D... F | RDAVQS | N KK | N Y.....L | D | K | S S YQ |
| Two *Mycobacterium* species | HG H C... F | REAVSS | S LS | .....G | D | R | R T YQ |
| *Mycobacterium* | HG H C... F | REAVSC | S LS | .....G | D | R | R T YQ |
| *Streptomyces* | HG H C... F | REAVGS | S LK | .....G | D | R | K T YQ |
| *Propionibacterium* | HG H C... F | REAVDS | Q LK | .....G | D | R | R A YQ |
| *Escherichia* | HG H C... F | REAAGS | A LK | .....G | D | R | K S YQ |
| *Salmonella* | HG H C... F | REAAGT | A LK | .....G | D | R | K T YQ |
| *Campylobacter* | HG H C... F | RDVAES | A PS | Y.....G | D | R | K S YQ |
| *Helicobacter* | HG H C... F | RDVANS | A PS | F.....G | D | R | K S YQ |
| *Methanobacterium* | HG H C... F | RGAVSS | A PA | F.....G | D | R | R S YQ |
| *Methanothermus* | HG H C... F | RSAVYS | S PA | F.....G | D | R | S T YQ |
| *Archaeoglobus* | HG H C... F | RDAAES | G FK | F.....G | D | R | R G YQ |
| *Methanococcus* | HG H C... F | REAAES | A PK | F.....G | D | R | K S YQ |
| *Clostridium* | HG H C... F | REAANS | A PQ | F.....G | D | R | K T YQ |
| Two *Chlorobium* species | FG H D... F | RDALHS | A PQ | F.....G | D | K | S T YQ |
| *Anacystis & Synechococcus* | HG H C... F | RDAAES | T PQ | F.....G | D | R | R T YQ |
| *Synechocystis* | HG H C... F | RDAADS | S PQ | F.....G | D | R | R T YQ |
| *Aquifex* | HG H C... F | REAAES | A PA | F.....G | D | R | R T YQ |
| *Bacillus* | HG H C... F | RDAANS | T PQ | F.....G | D | R | K T YQ |
| *Staphylococcus* | HG H C... F | RDAADS | A PS | F.....G | D | R | K T YQ |
| *Deinococcus* | HG H C... F | REAAGS | A PS | K.....G | D | R | A S YQ |

(*c*)

**Figure 7**
(*a*) A stereoview superposition of the ZnA binding site of yeast PBGS (carbon atoms in grey) with the corresponding region of monomer *A* of *P. aeruginosa* PBGS (carbon atoms in black). ZnA is dark red, water from the yeast structure is red, water from the *P. aeruginosa* structure is orange and all other atoms are colored cpk. (*b*) The ZnA of yeast PBGS in the context of the two monomers of *P. aeruginosa* PBGS. Residues T231 and Q138 are shown in ball and stick form and are labeled; the ordered active-site lid is shown in ball and stick form. Monomer *A* is above and shows the comparable position of ZnA as enclosed in the active site. Monomer *B* is below and shows the comparable location of ZnA as open to solvent owing to the disorder of the lid. The configuration of Q138 is seen to change with the apparent closing of the lid. (*c*) A distillation of the various sequences present in 53 available PBGS sequences in the region of the ZnA binding site. The sequence illustration includes the regions of sequence which are disordered in the yeast PBGS structure and in monomer *B* of the *P. aeruginosa* PBGS structure. Residues like the yeast PBGS are shown on grey, the analogous residues like *P. aeruginosa* PBGS are shown as white on black and residues like those of the active-site lid of *P. aeruginosa* PBGS are shown in olive green. The Asp residue that might serve as a ligand to MeA is shown in red.




(*a*)

| *P. aeruginosa* numbering | | | 129 | 131 | | | | | | | 139 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Yeast numbering | | | 133 | 135 | | | | | | | 143 | |
| **Three** mammalian species | C | D V | **C** | L | **C** | P | Y | T | S | H | **C** | G |
| **Two** yeast species | C | D V | **C** | L | **C** | E | Y | T | S | H | **C** | G |
| *Schizosaccharomyces* | C | D V | **C** | L | **C** | E | Y | T | D | H | **C** | G |
| *Chlamydia* | S | D I | **A** | L | **D** | P | Y | T | T | H | **D** | G |
| *Chlamydia* | S | D I | **A** | L | **D** | P | F | T | T | S | **D** | G |
| **Six** plant species | T | D V | **A** | L | **D** | P | Y | S | S | D | **D** | G |
| *Spinacia* | T | D V | **A** | L | **D** | P | Y | Y | Y | D | **D** | G |
| *Chlamydomonas* | T | D V | **A** | L | **D** | P | Y | N | S | D | **D** | G |
| **Two** *Neisseria* species | T | D V | **A** | L | **D** | P | Y | T | V | H | **D** | Q |
| *Bordella* | T | D V | **A** | L | **D** | P | Y | T | S | H | **D** | Q |
| *Shewanella, Vibrio, Yersinia* | T | D V | **A** | L | **D** | P | F | T | T | H | **D** | Q |
| *Actinobacillus* | T | D V | **A** | L | **D** | P | Y | T | V | H | **D** | Q |
| *Pseudomonas* | T | D V | **A** | L | **D** | P | F | T | T | H | **D** | Q |
| *Rhodobacter* | T | D I | **A** | L | **D** | P | Y | N | A | N | **D** | H |
| *Rhodobacter* | T | D V | **A** | L | **D** | P | Y | N | A | N | **D** | H |
| *Bradyrhizobium* | C | D V | **A** | L | **D** | P | F | T | S | H | **D** | G |
| *Rhodopseudomonas* | C | D V | **A** | L | **D** | P | Y | T | S | H | **D** | G |
| *Caulobacter* | C | D V | **A** | L | **D** | P | F | T | D | H | **D** | G |
| *Rickettsia* | C | D V | **A** | L | **D** | P | Y | T | I | S | **D** | G |
| **Three** *Mycobacterium* species | A | D T | **C** | L | **D** | E | F | T | D | H | **C** | G |
| *Streptomyces* | S | D L | **C** | L | **D** | E | F | T | D | H | **C** | G |
| *Propionibacterium* | A | D V | **C** | L | **D** | E | F | T | D | H | **C** | G |
| *Escherichia & Salmonella* | S | D T | **C** | F | **C** | E | Y | T | S | H | **C** | G |
| *Campylobacter* | S | D L | **C** | F | **C** | E | Y | T | D | H | **C** | G |
| *Helicobacter* | A | D L | **C** | F | **C** | E | Y | T | D | H | **C** | G |
| *Methanobacterium* | T | D V | **C** | L | **C** | Q | Y | T | S | H | **C** | G |
| *Methanothermus* | T | D V | **C** | L | **C** | Q | Y | T | E | H | **C** | G |
| *Archaeoglobus* | T | D V | **C** | L | **C** | E | Y | T | T | H | **C** | G |
| *Methanococcus* | A | D C | **C** | L | **C** | E | Y | T | S | H | **C** | G |
| *Clostridum* | T | D V | **C** | M | **C** | E | Y | T | S | H | **C** | G |
| **Two** *Chlorobium* species | T | D V | **A** | L | **D** | P | F | T | P | F | **D** | G |
| *Anacystis & Synechococcus* | A | D T | **C** | L | **C** | E | Y | T | P | H | **C** | G |
| *Synechocystis* | V | D T | **C** | L | **C** | E | Y | T | N | H | **C** | G |
| *Aquifex* | T | D V | **C** | F | **C** | E | Y | T | T | H | **C** | G |
| *Bacillus & Staphylococcus* | A | D T | **C** | L | **C** | E | Y | T | D | H | **C** | G |
| *Deinococcus* | A | D T | **C** | L | **C** | E | Y | T | S | H | **C** | G |

(*b*)

**Figure 8**
(*a*) A stereoview superposition of the ZnB binding site of yeast PBGS (carbon atoms in grey) with the corresponding region of monomer A of *P. aeruginosa* PBGS (carbon atoms in black). ZnA is dark red, water from the yeast structure is red, water from the *P. aeruginosa* structure is orange and all other atoms are colored cpk. (*b*) A distillation of the various sequences present in 53 available PBGS sequences in the region of the ZnB binding site. Residues like the yeast PBGS are shown on grey; the analogous residues like *P. aeruginosa* PBGS are shown as white on black. Uniformly conserved residues are in bold.

solvent. The lid residues that are near the ZnA site are L224, D228 and K229; these do not form a metal-binding site. However, alternate lid residues on other PBGS may form a reasonable binding site for an MgA. From this analysis, ZnA is seen poised to be enclosed in the active site, which argues for a functional significance. However, it is 12 Å from the Schiff base and 7 Å from the nearest oxygen atom of an ordered sulfate molecule seen in the *P. aeruginosa* PBGS active site. The sulfate was interpreted to sit at the site of A-side ALA and this is in general agreement with the predicted location of A-side ALA based on *E. coli* PBGS (Erskine, Norton *et al.*, 1999). Therefore, the currently available crystal structures for PBGS do not support the notion that A-side ALA acts as a bidentate ligand to ZnA (Jaffe, 1995). It remains possible that motion within the active site could bring these two moieties together.

The PBGS sequence in the regions of amino acids seen as ZnA ligands and the amino acids that comprise the active site lid are presented in Fig. 7(*c*). Identical sequences (at these residues only) are grouped together such that all 53 available sequences can be presented. There is little information in these sequences to support the existence of a MgA, unless one considers residues from the loop. In that case, a comparable MgA site might exist for some PBGS that contain an aspartic acid in the third position prior to C234. This suggests that the D*xx*C/S motif might form the nucleation site for the binding of MeA. It has previously been shown that the HGH motif included in Fig. 2 does not form a $Zn^{II}$ nucleation site for *E. coli* PBGS (Mitchell *et al.*, 1995).

## 9. The MeB site: is it the metal or the amino acids that matter?

Fig. 8(*a*) shows a superposition of the ZnB site seen in yeast PBGS with the comparable location of monomer *A* of *P. aeruginosa* PBGS. In this case, the corresponding amino acids are nearly isosteric with each other. In the two cases where an aspartic acid of *P. aeruginosa* PBGS replaces a cysteine ligand of yeast PBGS, the sulfur atom of the cysteine lies approximately at the same site as the carboxyl carbon atom of these aspartates. In the case of the cysteine to alanine transition, the alanine falls right on top of the cysteine. These residues fall in the metal-binding region first illustrated in Fig. 2 and *P. aeru-*

*ginosa* PBGS contains each of the aspartates predicted to be ligands to a putative MgB. The lack of $Mg^{II}$ at the active site of *P. aeruginosa* PBGS raises the question as to whether *any* of the PBGSs contain $Mg^{II}$ in this site.

Fig. 8(*b*) shows this metal-binding region of sequence for all 53 PBGS sequences. Although the sequences of the neighboring residues vary, the three residues that bind ZnB fall into only three different constellations. Each PBGS is either like yeast (three cysteines in these positions), like *P. aeruginosa* PBGS (one alanine and two aspartic acid residues in these positions) or is a hybrid of the two, with two cysteines and one aspartic acid. The relatively invariant nature of these sequences and the proximity of the amino acids to the active site argues for their functional significance. However, the lack of an MgB in *P. aeruginosa* PBGS (Frankenberg, Jahn *et al.*, 1999) suggests that the significant functional switch may be a $Zn^{II}$-bound cysteine for a metal-free aspartic acid rather than the previously presumed $Zn^{II}$ to $Mg^{II}$ switch.

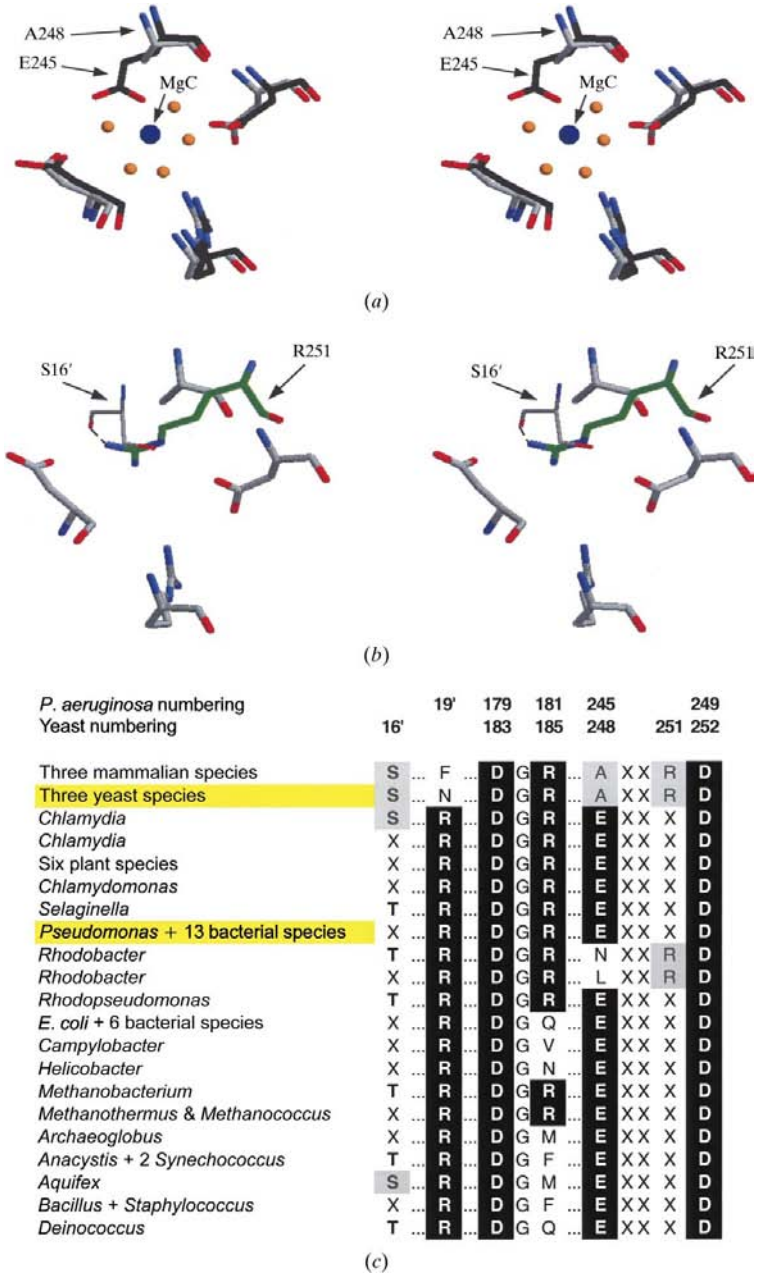



**Figure 9**
(*a*) A stereoview of the MgC site, seen in monomer *A* of *P. aeruginosa* PBGS (carbon atoms colored black) superimposed on the homologous amino acids of yeast PBGS (carbon atoms colored grey). The water molecules in the first coordination sphere of the $Mg^{II}$ are orange. (*b*) The homologous location of the positively charged R251 (carbon atoms colored green) which fills the void left by the lack of MgC. S16′, shown in narrow sticks, forms a hydrogen bond to R251 and derives from the arm of a second monomer. (*c*) A distillation of the sequence around the MgC binding site. Residues like yeast are on grey and residues like *P. aeruginosa* are shown as white on black.

## 10. The MeC site: not a universal determinant for asymmetry

Fig. 9(*a*) shows the superposition of the $Mg^{II}$ binding site of *P. aeruginosa* PBGS with the comparable location of yeast PBGS. In this case, nearly all the amino acids are identical and isosteric. The only amino-acid change is E245 of *P. aeruginosa* PBGS to A242 of yeast PBGS, thus omitting the single first coordination sphere amino-acid ligand to the $Mg^{II}$. Fig. 9(*b*) includes R251 of yeast PBGS which is seen to take the place of the positive charge of the $Mg^{II}$ of *P. aeruginosa* PBGS. In this case, the linkage to the arm of the adjacent monomer is accomplished by a hydrogen bond between the guanidinium group of R251 and S16′, which derives from the arm of the other subunit. Thus, it becomes clear why the octameric structure of mammalian and yeast PBGS is more stable than that of those PBGSs which contain MgC. In the latter case, it has been shown that the specific activity is protein-concentration dependent (the octamer dissociates to smaller less active units at low protein concentration) and EDTA promotes the dissociation of the octamer during native gel electrophoresis (Jaffe *et al.*, 1995; Petrovich *et al.*, 1996; Frankenberg, Jahn *et al.*, 1999).

Fig. 9(*c*) distills the 53 known sequences of this region into a smaller basis set in order to look at the binding of MgC. All but the mammalian, yeast and *Rhodobacter* PBGSs contain a glutamate analogous to E245 of *P. aeruginosa* PBGS and each of these sequences also contains the R19′ that forms the intersubunit bridge (see Fig. 5*d*). Two second-sphere MgC ligands are universally conserved. However, R181, which is important to the Mg switch illustrated in Fig. 6(*b*), is not present in all PBGSs containing E245, but it is present in yeast and mammalian PBGSs. From this analysis it appears that most PBGSs contain an MgC binding site at their subunit interfaces and many of these may utilize the Mg-switch mechanism seen in *P. aeruginosa* PBGS. However, an analogous switch is not obvious for mammalian PBGS, which shows strong evidence for half-site reactivity.

## 11. The PBGS reaction mechanism

In 1968, a Schiff-base intermediate was identified between one ALA substrate molecule and an active-site lysine residue (Nandi & Shemin, 1968). It was reasoned that the Schiff base formed to A-side ALA and an attractive mechanism was proposed that lived on in biochemistry texts for two decades. Ten years later, the Schiff base was reassigned to the P-side ALA and an equally attractive mechanism was proposed, which had some difficulty displacing the former mechanism from texts (Jordan & Seehra, 1980). During this time, the stereospecificity of proton removal from the C5 of P-side ALA was also determined (Chaudhry & Jordan, 1976), but none of the proposed mechanisms acknowledged a potential role for metal ions. By the mid 1980s, evidence supporting a role for $Zn^{II}$ in the binding and reactivity of A-side ALA was presented (Jaffe & Hanes, 1986), but no new mechanism was proposed as there was little conclusive data on the order of bond formation between the two ALA molecules. This uncertainty remains (Erskine, Newbold *et al.*, 1999). The late 1980s saw the publication of a series of $^{13}C$ and $^{15}N$ NMR studies that defined the stereochemistry and protonation states of the Schiff-base intermediate (Jaffe & Markham, 1987, 1988; Jaffe *et al.*, 1990). These studies support a mechanism where inter-subunit C—N bond formation (a second Schiff base) precedes C—C bond formation (aldol condensation) (Jaffe, 1995). Thus, a crystal structure was much anticipated with regard to the information that might be provided regarding the PBGS catalytic reaction mechanism.

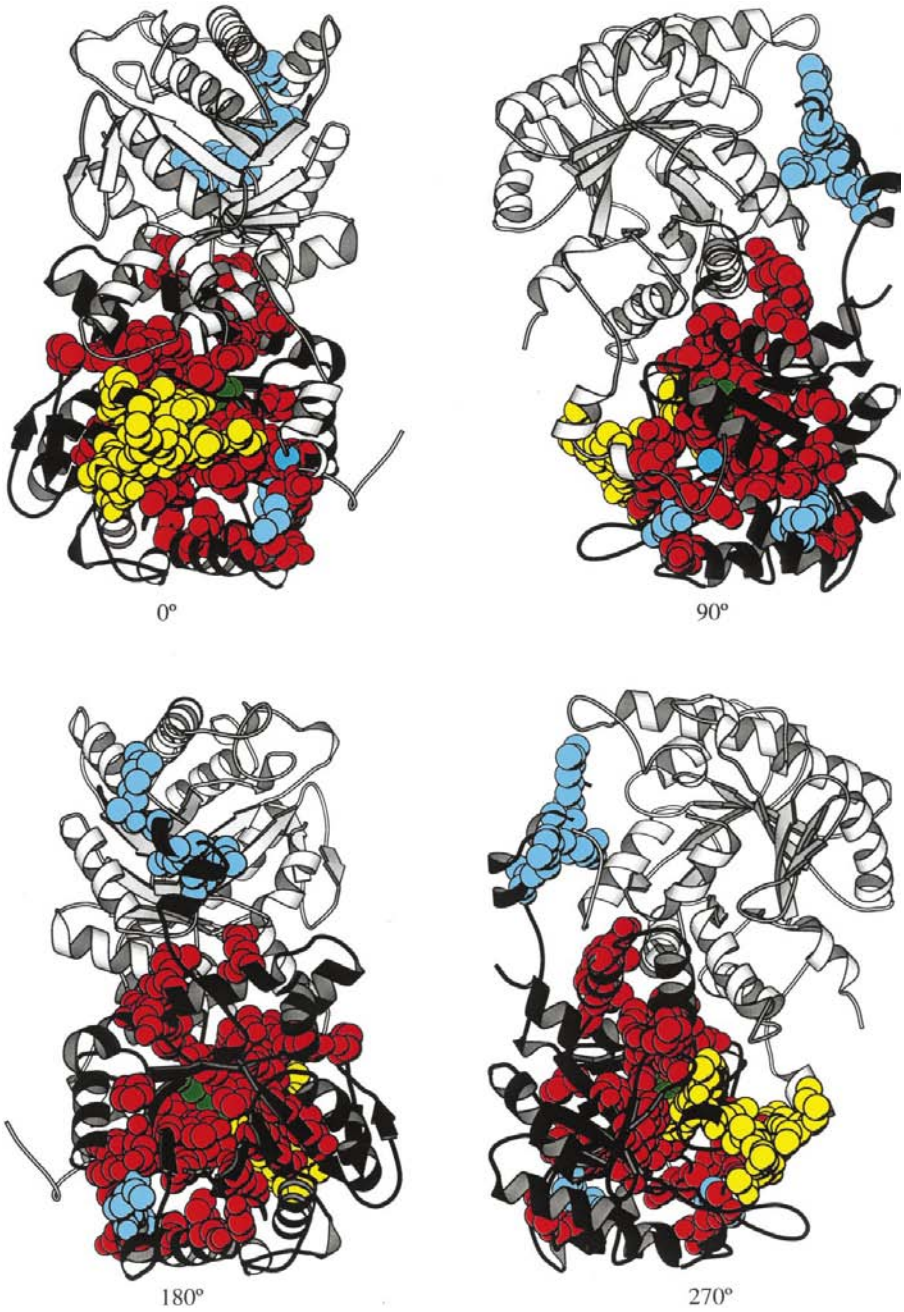



**Figure 10**
Monomer *A* of *P. aeruginosa* PBGS is used to illustrate the amino acids that are conserved in at least 52 of the 53 sequences (shown in red; set Y = F, D = E, K = R, S = T). Those residues that are conserved in all but the yeast and mammalian sequences are shown in cyan. The Schiff-base-forming lysine is shown in green, MgC is shown in blue and the region that is disordered in monomer *B* is shown in yellow. Monomer *B* is a light-grey ribbon diagram. The four views are each rotated 90° to the left relative to the preceding view.

In the PBGS-catalyzed reaction, the important roles the enzyme must play are to bind the two ALA molecules, preferentially through parts of the molecule that do not undergo hybridization changes, and to facilitate the chemical transmogrification from substrate to product (see Fig. 1*b*). At the least, one would expect the enzyme to provide specific bases for the abstraction of the protons. An excellent description of this problem has been presented based on an analysis of the crystal structure of *E. coli* PBGS (Erskine, Norton *et al.*, 1999). In that case, the proton-abstraction steps were assigned to the Schiff-base-forming lysine, an invariant lysine that lies adjacent in the crystal structure and a hydroxyl group bound to $Zn^{II}$ that is not present in *P. aeruginosa* PBGS, but is presented here as ZnB of the yeast PBGS structure. Clearly, of the two $Zn^{II}$ seen in yeast PBGS, ZnB is closer to the putative location of A-side ALA, which is between the ZnB site and the P-side Schiff base. Nevertheless, the analysis presented above

places ZnA in the active site. Finally, the recently established lack of metals in the *P. aeruginosa* PBGS active site requires that whatever the roles of ZnA and ZnB, there must be amino acids (or monovalent cations) that can play these roles in PBGSs like that of *P. aeruginosa*.

One general assumption for the PBGS mechanism is that some catalytically essential residues will be conserved. Clearly the metal ions and their ligands are not conserved. Fig. 10 illustrates the *P. aeruginosa* PBGS dimer using monomer *A* to show the highly conserved residues of PBGS. The uniformly conserved residues form a band that extends from the barrel-arm interface (where MgC is seen) through the active site and up to the other side. Some conserved residues are also seen at the barrel–barrel interface of the hugging dimer. Those residues that are conserved in all but the animal and yeast PBGSs occur predominantly at the barrel–arm interface of the hugging dimer. From this view, it appears that subunit interactions are important to catalytic activity, consistent with limited evidence suggesting that the dimer is the minimal functional unit of activity (Stafforini *et al.*, 1980). It is interesting to note that the mobile active-site lid, often cited as 'highly conserved' (Warren *et al.*, 1998; Frankenberg, Erskine *et al.*, 1999), contains few uniformly conserved residues. K229 of *P. aeruginosa* PBGS, which is either a lysine or an arginine, is one of the conserved few; K229 is the residue that directly responds to the presence of MgC (see also Figs. 6*b*, 7*b* and 7*c*). K229 brings a third lysine into the active site and may be the key to setting up the environment for catalysis.

## 12. So, how many metal ions does it really take to make a pyrrole?

The dogma that that PBGS can be divided into two main groups based on metal-ion usage is clearly an oversimplification. The assumption that the sequence region illustrated in Figs. 2 and 8(*b*) explains a functional substitution of $Zn^{II}$ for $Mg^{II}$ is not supported by the analysis presented here, which raises some doubt as to whether there are any PBGSs that use $Mg^{II}$ in place of either ZnA or ZnB. It is clear is that the MgC site exists in some PBGSs at subunit interfaces and that MgC can communicate with the active site, thus giving a structural basis to the allosteric effect (Frankenberg, Erskine *et al.*, 1999). It is also clear that the existence of a specific arginine–serine pair can functionally substitute for the role of MgC in stabilizing the quaternary structure of PBGS (see Fig. 9*b*). It is similarly possible that glutamic acids functionally substitute for the $Zn^{II}$-bound cysteines seen in Fig. 8. It is not clear which amino acids might substitute for the ZnA shown in Fig. 7, though this is where monovalent metal ions may come in. Table 2 suggests that the monovalent effect is present when MeA is absent. In any case, if there is no MgA and no MgB, then what are the roles and/or binding sites of the $Mg^{II}$ that are not MgC? The total stoichiometry of metal-ion binding presented in Table 2 is well documented (Jaffe *et al.*, 1992; Petrovich *et al.*, 1996; Senior *et al.*, 1996; Cheung *et al.*, 1997; Frankenberg, Jahn *et al.*, 1999). Could there be a fourth divalent metal ion binding site?

**Table 3**
What if there are four different divalent metal ion sites in PBGS?

| | | Metal ions required for full activity per homooctamer | | | | | |
|---|---|---|---|---|---|---|---|
| Type | Species | ZnA | ZnB | MgC | MgD | Total divalents | Monovalent requirement |
| I | *H. sapiens* | 4 | 4 | 0 | 0 | 8 | No |
| ? | *S. cerevisiae* | *4* | *4* | 0 | 8 | 16 | No |
| II | *E. coli* | *4 or 0* | *4 or 8* | 8 or 8 | 0 or 0 | 16 | No |
| III | *P. sativum* | *0* | *0* | *8* | *8* | 16 | No |
| IV | *B. japonicum* | *0* | *0* | *8* | *4* | 12 | Yes |
| V | *P. aeruginosa* | 0 | 0 | 4 | 0 | 4 | Yes |

The published biochemical data on yeast PBGS may provide information about a potential fourth site. It is well established that yeast PBGS, *E. coli* PBGS and *P. sativum* (pea) PBGS can all bind 16 divalent metal ions (Senior *et al.*, 1996; Cheung *et al.*, 1997). Yeast and *E. coli* PBGSs can each bind 16 $Zn^{II}$ or eight $Zn^{II}$ plus eight $Mg^{II}$. Pea PBGS can only bind 16 $Mg^{II}$. A comparative analysis of the former two proteins made the assumption that their binding identical numbers and types of metal ions meant that they bound the same configuration of metal ions (Senior *et al.*, 1996). However, the crystal structures of these two proteins show this assumption to be unsubstantiated. The *E. coli* PBGS sequence shows the MgC binding site and the *E. coli* crystal structure, which was obtained on the all-$Zn^{II}$ form of the protein, contains a $Zn^{II}$ in the MgC site (Erskine, Norton *et al.*, 1999). MgC cannot bind to yeast PBGS because its binding is prohibited by the presence of Arg251 (see Fig. 9). Therefore, the eight $Mg^{II}$ seen to bind to yeast PBGS must be different from the eight $Mg^{II}$ seen to bind to *E. coli* PBGS. This is functionally substantiated by the fact that the $Mg^{II}$ of *E. coli* PBGS is an allosteric activator, whereas the $Mg^{II}$ of yeast PBGS has no kinetic effect (Senior *et al.*, 1996). One possible interpretation is presented in Table 2, where yeast PBGS is considered as containing eight ZnA and eight ZnB. However, there are no data from any PBGSs to suggest that $Mg^{II}$ can bind to the ZnB site with its three cysteine ligands; therefore, in yeast PBGS the eight $Zn^{II}$ and eight $Mg^{II}$ are unlikely to be eight ZnA and eight MgB. Since the ZnA site for yeast PBGS contains the same ligands as mammalian PBGS and there is no data suggesting that $Mg^{II}$ can bind to mammalian PBGS, it is unlikely that the yeast PBGS containing equal numbers of $Zn^{II}$ and $Mg^{II}$ can correspond to eight MgA and eight ZnB. Hence, this suggests that the eight $Mg^{II}$ which can bind to yeast PBGS might bind to a fourth divalent metal ion binding site.

The possibility of a fourth divalent metal ion binding site is consistent with the doubts raised in this work as to whether *any* PBGSs contain $Mg^{II}$ in the MeA or MeB sites. Since it is well documented that pea PBGS can bind 16 $Mg^{II}$ (Cheung *et al.*, 1997), eight of these are assigned to the MgC sites and perhaps the others are at the fourth site. Table 3 presents a preliminary idea of how one might categorize the PBGSs on the basis of four specially spatially and functionally distinct metal-binding sites. Although complex, Table 3 is consistent

with the available data. One attractive aspect of Table 3 is that it does not infer a functional substitution of $Zn^{II}$ for $Mg^{II}$ at any one site.

The history of our knowledge of the metal ions in PBGS is instructive. In 1972, there were data suggesting that PBGS used $Zn^{II}$ or $Mg^{II}$ and might also respond to monovalent metal ions. By 1980, there were data suggesting that the eight $Zn^{II}$ of mammalian PBGS fell into two different classes (Bevan *et al.*, 1980). The purification and characterization of *E. coli* PBGS in the early 1990s yielded data suggesting the three divalent metal-ion sites described in this work (Mitchell & Jaffe, 1993) and the turn of the millennium suggests that there might even be more than three.

## 13. Lead poisoning may be a ZnA phenomenon

The history and sociology of lead poisoning in humans as it relates to PBGS is a fascinating topic that has recently seen an outstanding review (Warren *et al.*, 1998). With regard to $Pb^{II}$ inhibition of PBGS, the general dogma within the medical and biochemical communities is that $Pb^{II}$ inhibition of the $Zn^{II}$-utilizing PBGS is caused by binding of $Pb^{II}$ to the cysteine-rich site. Crystallographic data show that such binding occurs and also shows that $Pb^{II}$ can bind to these residues in more than one way (Warren *et al.*, 1998). The uncertainty lies in extending this binding observation to functional inhibition. For instance, it has been shown that the competitive occupancy of $Pb^{II}$ and $Zn^{II}$ in mammalian PBGS is a function of the presence or absence of substrate (Jaffe *et al.*, 1991). Furthermore, the disproportionation of the available $Pb^{II}$ and $Zn^{II}$ between metal-binding sites is a slow process following addition of substrate to the reaction mixture. Thus, the preferential binding of $Pb^{II}$ to the cysteine-rich ZnB site seen in the absence of ALA does not necessarily mean that catalytic inhibition occurs when $Pb^{II}$ is bound to the ZnB site.

We and others have attempted to use the $Zn^{II}$-requiring microbial PBGS as a model for lead poisoning, with a variety of results. An excellent example of this is the crystallographic observation of $Pb^{II}$ bound to yeast PBGS (Warren *et al.*, 1998) and functional studies of $Pb^{II}$ inhibition of that protein currently ongoing (Godwin *et al.*, 1999). Another example is our unpublished observation that *E. coli* PBGS is not nearly as sensitive to inhibition by $Pb^{II}$ as is the mammalian protein. The crystal structure of *E. coli* PBGS suggests that it does not contain ZnA. Perhaps $Pb^{II}$ inhibition is only limited to those PBGSs that contain ZnA.

## 14. Conclusions

The PBGS family of proteins contains considerable diversity in metal-ion usage. The current analysis suggests that the long-held notion that PBGSs use either $Zn^{II}$ or $Mg^{II}$ does not necessarily correspond to a functional/spatial replacement of these two ions. The diversity in metal-ion usage corresponds to differences in active-site structures that might translate to species-specific PBGS inhibitors.

This work was supported by NIH grants ES03654 (EKJ), CA06927 (Institute for Cancer Research) and an appropriation from the Commonwealth of Pennsylvania. Its contents are solely the opinion of the author and do not necessarily represent the official views of the National Institute for Environmental Health Sciences or of the National Cancer Institute. Dr Michael Sauder is acknowledged for preparing the coordinate files containing the superpositions of the yeast and *P. aeruginosa* PBGS structures. James Tisdale is acknowledged for writing a computer program that allowed the author to construct Fig. 3 from the vast array of PBGS sequence data. Drs Nicole Frankenberg, Dieter Jahn and Dirk Heinz are acknowledged for their generous contribution of the X-ray crystal structure coordinates for the *P. aeruginosa* PBGS.

## References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**(17), 3389–3402.
Bevan, D. R., Bodlaender, P. & Shemin, D. (1980). *J. Biol. Chem.* **255**(5), 2030–2035.
Chaudhry, A. G. & Jordan, P. M. (1976). *Biochem. Soc. Trans.* **4**(4), 760–761.
Chauhan, S. & O'Brian, M. R. (1995). *J. Biol. Chem.* **270**(34), 19823–19827.
Cheh, A. & Neilands, J. B. (1973). *Biochem. Biophys. Res. Commun.* **55**(4), 1060–1063.
Cheung, K. M., Spencer, P., Timko, M. P. & Shoolingin-Jordan, P. M. (1997). *Biochemistry*, **36**(5), 1148–1156.
Dent, A. J., Beyersmann, D., Block, C. & Hasnain, S. S. (1990). *Biochemistry*, **29**(34), 7822–7828.
Erskine, P. T., Newbold, R., Roper, J., Coker, A., Warren, M. J., Shoolingin-Jordan, P. M., Wood, S. P. & Cooper, J. B. (1999). *Protein Sci.* **8**(6), 1250–1256.
Erskine, P. T., Norton, E., Cooper, J. B., Lambert, R., Coker, A., Lewis, G., Spencer, P., Sarwar, M., Wood, S. P., Warren, M. J. & Shoolingin-Jordan, P. M. (1999). *Biochemistry*, **38**(14), 4266–4276.
Erskine, P. T., Senior, N., Awan, S., Lambert, R., Lewis, G., Tickle, I. J., Sarwar, M., Spencer, P., Thomas, P., Warren, M. J., Shoolingin-Jordan, P. M., Wood, S. P. & Cooper, J. B. (1997). *Nature Struct. Biol.* **4**(12), 1025–1031.
Falicov, A. & Cohen, F. E. (1996). *J. Mol. Biol.* **258**(5), 871–892.
Frankenberg, N., Erskine, P. T., Cooper, J. B., Shoolingin-Jordan, P. M., Jahn, D. & Heinz, D. W. (1999). *J. Mol. Biol.* **289**(3), 591–602.
Frankenberg, N., Jahn, D. & Jaffe, E. K. (1999). *Biochemistry*, **38**(42), 13976–13982.
Gibbs, P. N. & Jordan, P. M. (1986). *Biochem. J.* **236**(2), 447–451.
Glusker, J. P., Katz, A. K. & Bock, C. W. (1996). *Acta Cryst.* A**52**, C133.
Godwin, H. A., Ghering, A. B., Wang, J. & Padia, S. (1999). *J. Biol. Inorg. Chem.* **74**(1–4), 141.
Hightower, K. E. & Fierke, C. A. (1999). *Curr. Opin. Chem. Biol.* **3**(2), 176–181.
Jaffe, E. K. (1993). *Comments Inorg. Chem.* **15**, 67–93.
Jaffe, E. K. (1995). *J. Bioenerg. Biomembr.* **27**(2), 169–179.
Jaffe, E. K., Abrams, W. R., Kaempfen, H. X. & Harris, K. A. Jr (1992). *Biochemistry*, **31**(7), 2113–2123.
Jaffe, E. K., Ali, S., Mitchell, L. W., Taylor, K. M., Volin, M. & Markham, G. D. (1995). *Biochemistry*, **34**(1), 244–251.
Jaffe, E. K., Bagla, S. & Michini, P. A. (1991). *Biol. Trace Elem. Res.* **28**(3), 223–231.
Jaffe, E. K. & Hanes, D. (1986). *J. Biol. Chem.* **261**(20), 9348–9353.

Jaffe, E. K. & Markham, G. D. (1987). *Biochemistry*, **26**(14), 4258–4264.
Jaffe, E. K. & Markham, G. D. (1988). *Biochemistry*, **27**(12), 4475–4481.
Jaffe, E. K., Markham, G. D. & Rajagopalan, J. S. (1990). *Biochemistry*, **29**(36), 8345–8350.
Jaffe, E. K., Volin, M., Myers, C. B. & Abrams, W. R. (1994). *Biochemistry*, **33**(38), 11554–11562.
Jordan, P. M. & Seehra, J. S. (1980). *FEBS Lett.* **114**(2), 283–286.
Kimura, E., Gotoh, T., Kioke, T. & Shiro, M. (1999). *J. Am. Chem. Soc.* **121**, 1267–1274.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Lingner, B. & Kleinschmidt, T. (1983). *Z. Naturforsch. C*, **38**(11/12), 1059–1061.
Lippard, S. J. & Berg, J. M. (1994). *Principles of Bioinorganic Chemistry*. Mill Valley, California: University Science Books.
Matthews, R. G. & Goulding, C. W. (1997). *Curr. Opin. Chem. Biol.* **1**(3), 332–339.
Mitchell, L. W. & Jaffe, E. K. (1993). *Arch. Biochem. Biophys.* **300**(1), 169–177.
Mitchell, L. W., Volin, M. & Jaffe, E. K. (1995). *J. Biol. Chem.* **270**(41), 24054–24059.
Myers, L. C., Cushing, T. D., Wagner, G. & Verdine, G. L. (1994). *Chem. Biol.* **1**(2), 91–97.
Nandi, D. L. & Shemin, D. (1968). *J. Biol. Chem.* **243**(6), 1236–1242.
Petrovich, R. M. & Jaffe, E. K. (1997). *Biochemistry*, **36**(43), 13421–13427.
Petrovich, R. M., Litwin, S. & Jaffe, E. K. (1996). *J. Biol. Chem.* **271**(15), 8692–8699.
Rost, B. & Sander, C. (1994). *Proteins*, **19**(1), 55–72.
Seehra, J. S., Gore, M. G., Chaudhry, A. G. & Jordan, P. M. (1981). *Eur. J. Biochem.* **114**(2), 263–269.
Seehra, J. S. & Jordan, P. M. (1981). *Eur. J. Biochem.* **113**(3), 435–446.
Senior, N. M., Brocklehurst, K., Cooper, J. B., Wood, S. P., Erskine, P., Shoolingin-Jordan, P. M., Thomas, P. G. & Warren, M. J. (1996). *Biochem. J.* **320**, 401–412.
Shemin, D. (1972). *The Enzymes*, edited by P. Boyer, Vol. VII, pp. 232–237. New York: Academic Press.
Shemin, D. (1976). *Philos. Trans. R. Soc. London Ser. B, Biol. Sci.* **273**(924), 109–115.
Shemin, D. (1979). *Ann. NY Acad. Sci.* **325**, 210–227.
Shemin, D. & Russell, C. S. (1953). *J. Am. Chem. Soc.* **75**, 4873.
Shoolingin-Jordan, P. M. (1998). *Biochem. Soc. Trans.* **26**(3), 326–336.
Spencer, P. & Jordan, P. M. (1993). *Biochem. J.* **290**, 279–287.
Spencer, P. & Jordan, P. M. (1994). *Ciba Found. Symp.* **180**, 50–64.
Stafforini, D. M., Polo, C. F., Stella, A. M., Wider de Xifra, E. & Batlle, A. M. (1980). *Int. J. Biochem.* **12**(5/6), 757–760.
Vallee, B. L. & Auld, D. S. (1995). *Experientia Suppl.* **73**, 259–277.
Warren, M. J., Cooper, J. B., Wood, S. P. & Shoolingin-Jordan, P. M. (1998). *Trends Biochem. Sci.* **23**(6): 217–221.
Wetmur, J. G., Bishop, D. F., Cantelmo, C. & Desnick, R. J. (1986). *Proc. Natl Acad. Sci. USA*, **83**(20), 7703–7707.
Woehl, E. U. & Dunn, M. F. (1995). *Biochemistry*, **34**(29), 9466–9476.