Structure and proposed mechanism of α-glycerophosphate oxidase from *Mycoplasma pneumoniae*

Callia K. Elkhal,[1#] *Kelsey Kean,*[1#] *Derek Parsonage,*[2] *Somchart Maenpuen,*[3] *Pimchai Chaiyen,*[4] *Al Claiborne,*[2,*] and *P. Andrew Karplus*[1*]

[1] Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331

[2] Center for Structural Biology, Wake Forest School of Medicine, Winston-Salem, North Carolina 27157

[3] Department of Biochemistry, Faculty of Science, Burapha University, Chonburi, Thailand

[4] Department of Biochemistry and Center of Excellence in Protein Structure and Function, Faculty of Science, Mahidol University, Bangkok, Thailand

[*] To whom correspondence should be addressed: e-mail: karplusp@science.oregonstate.edu, alc@csb.wfu.edu

[#] These two authors contributed equally to this work.

Contact information for editorial correspondence: Andrew Karplus; 2011 ALS Bldg, Dept. of Biochemistry, Oregon State University, Corvallis, OR 97331; ph: 541-737-3200; Fax: 541-737-0481; email: karplusp@science.oregonstate.edu

Running Title: α-glycerophosphate oxidase structure and mechanism

Database: Structural data are available in PDB database under the accession number(s) **4x9m** (oxidized) and **4x9n** (reduced).

Abbreviations: Glp – ʟ-α-glycerophosphate; DHAP – dihydroxyacetone phosphate; GlpO - ʟ-α-glycerophosphate oxidase; DH - ʟ-α-glycerophosphate dehydrogenase; *Mp – Mycoplasma pneumoniae*; *Bp – Bordetella pertussis*; *Sp – Streptococcus pneumonia*; *Ssp – Streptococcus* sp.

Enzyme Commission number: ʟ-α-glycerophosphate oxidase – EC 1.1.3.21

**Keywords:** flavoenzyme, drug design, protein evolution, GlpA, hydride transfer

ABSTRACT

The formation of hydrogen peroxide (H_2O_2) by the FAD-dependent α-glycerophosphate oxidase (GlpO), is important for the pathogenesis of *Streptococcus pneumoniae* and *Mycoplasma pneumoniae*. The structurally known GlpO from *Streptococcus* sp. (*Ssp*GlpO) is similar to the pneumococcal protein (*Sp*GlpO) and provides a guide for drug design against that target. However, *M. pneumoniae* GlpO (*Mp*GlpO), having <20% sequence identity with structurally known GlpOs, appears to represent a second type of GlpO we designate as Type II GlpOs. Here, the recombinant His-tagged *Mp*GlpO structure is described at ~2.5 Å resolution, solved by molecular replacement using as a search model the *Bordetella pertussis* protein 3253 (*Bp*3253) a protein of unknown function solved by structural genomics efforts. Recombinant *Mp*GlpO is an active oxidase with a turnover number of ~580 min^{-1} while *Bp*3253 showed no GlpO activity. No substantial differences exist between the oxidized and dithionite-reduced *Mp*GlpO structures. Although, no liganded structures were determined, a comparison with the tartrate-bound *Bp*3253 structure and consideration of residue conservation patterns guided the construction of a model for α-glycerophosphate (Glp) recognition and turnover by *Mp*GlpO. The predicted binding mode also appears relevant for the type I GlpOs (such as *Ssp*GlpO) despite differences in substrate recognition residues, and it implicates a histidine conserved in type I and II Glp oxidases and dehydrogenases as the catalytic acid/base. This work provides a solid foundation for guiding further studies of the mitochondrial Glp dehydrogenases as well as for continued studies of *M. pneumoniae* and *S. pneumoniae* glycerol metabolism and the development of novel therapeutics targeting *Mp*GlpO and *Sp*GlpO.

INTRODUCTION

Mycoplasma pneumoniae is a human respiratory tract pathogen that causes 40% or more of community-acquired pneumonias [1], with the typical syndrome being tracheobronchitis in children. This pathogen initiates colonization of the host airway mucosal epithelium *via* a special attachment organelle [2], which also provides an important gliding function. Among other distinctive features of *M. pneumoniae* are a small genome, the lack of a rigid cell wall, and limited metabolic capabilities. As in the Gram-positive streptococci from which the mycoplasmas diverged ca. 600 million years ago [1], the tricarboxylic acid cycle, electron transport chain, and respiratory cytochromes are absent. However, glycerol metabolism appears to be an important pathogenicity factor for *M. pneumoniae* [3-6] and thus the enzymes involved provide potential drug targets for combating respiratory infectious diseases.

Of particular interest in this context is the *M. pneumoniae glpD* gene (MPN051) that is annotated as encoding a glycerol-3-phosphate dehydrogenase [4], and based on its sequence can be identified as a member of the D-amino acid oxidase (DAAO) superfamily [7] of FAD-dependent enzymes combining a 'glutathione-reductase-2' type FAD-binding domain and an antiparallel β-sheet based substrate-binding domain. Despite the annotation as a dehydrogenase, the encoded enzyme has been shown to be a constitutively expressed cytosolic FAD-dependent α-glycerophosphate oxidase [4], using O_2 as the final electron acceptor and producing dihydroxyacetone phosphate (DHAP) and H_2O_2:
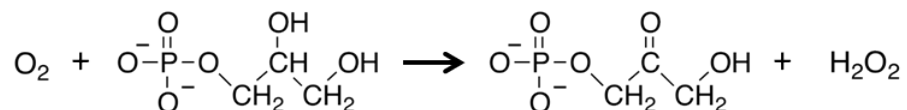
We therefore refer here to the encoded protein as *M. pneumoniae* GlpO (*Mp*GlpO), rather than GlpD. The gene is tightly linked with that for glycerol kinase (*glpK*), and together these enzymes can catalyze the ATP-dependent conversion of glycerol to the glycolytic intermediate dihydroxyacetone phosphate [4]. The peroxide produced by *Mp*GlpO has been shown to be crucial for pathogenicity, and the ortholog from the animal pathogen, *Mycoplasma mycoides* subsp. *mycoides* SC, has also been implicated as a primary virulence factor [3]. Similarly, Mahdi *et al.* [8] have shown that *Streptococcus pneumoniae* GlpO (*Sp*GlpO) is responsible for H_2O_2-mediated cytotoxicity against human brain microvascular endothelial cells and for promoting pneumococcal meningitis.

Given the clinical significance of *M. pneumoniae* and *S. pneumoniae* and the importance of *Mp*GlpO and *Sp*GlpO in cytotoxicity and virulence, structural studies of these enzymes would provide a valuable foundation for drug design. A good understanding of the *Sp*GlpO structure is in hand because it is quite similar (~62% sequence identity) to the structurally known *Streptococcus* sp. GlpO (*Ssp*GlpO) [9]. However, insufficient information exists for modeling the 43 kD *Mp*GlpO enzyme, as it is surprisingly divergent from the higher molecular weight (~65 kD) streptococcal enzymes: it is not only missing a C-terminal α-helical domain, but it has only ~20% sequence identity with *Ssp*GlpO and is actually more sequence similar to other DAAO superfamily enzymes, such as glycine oxidase, than it is to *Ssp*GlpO. To provide a foundation for understanding catalysis and guiding drug design against *Mp*GlpO, we report here its crystal structure. As part of this work, we also carried out some functional characterizations of the protein encoded by *Bordetella pertussis* gene 3253 (*Bp*3253), which is a related protein of unknown function solved by the NorthEast Structural Genomics group (PDB entry 3DME, deposited 2008). It became of interest because it is the structurally known protein most similar to *Mp*GlpO (27% sequence identity).

RESULTS and DISCUSSION

Expression and biochemical properties of MpGlpO and the B. pertussis protein Bp*3253*. Previously published studies of *Mp*GlpO were carried out with an N-terminal His-tagged protein derived from *M. pneumoniae* M129 [4]. In the present study, the pET28a vector introduces a 34-residue His-tag in-frame with the codon-optimized synthetic gene, and this tag was present in all studies. The characterization of *B. pertussis Bp*3253 was carried out using the same N-terminally His-tagged construct used in the crystal structure determination of PDB entry 3DME. Both recombinant *Mp*GlpO and *Bp*3253 were expressed and purified (yielding ~5 and ~60 mg/L culture respectively), and the visible absorption spectra for the purified proteins are characteristic of properly folded flavoenzymes (Figure 1). The native molecular weight for His-tagged *Mp*GlpO as determined by gel filtration, at 41.7 kD, is reasonably close to the value of 46.3 kD calculated for His-tagged *Mp*GlpO indicating that the recombinant protein is a monomer in solution.

Although the oxidase activity of *Mp*GlpO was documented previously [4], we assayed the recombinant His-tagged *Mp*GlpO purified here and found a specific activity of 12.8 U/mg at 25

˚C with 200 mM D/L-Glp and ca. 260 μM O_2 (the saturating concentration in water at 25 ℃). This is equivalent to a k_{cat} of 9.7 s^{-1} (i.e. 580 min^{-1}). This level of activity is reasonable for an oxidase even though it is somewhat lower than was seen for native and recombinant GlpOs from *E. casseliflavus* (*Ecass*GlpO) [10] and *Streptococcus* sp. (*Ssp*GlpO) [11] which gave k_{cat} values of 70-90 s^{-1} at 25 ℃. When *Bp*3253 was tested using the same assay, it showed no ability to catalyze the Glp-dependent formation of H_2O_2.

Many flavoprotein oxidases show a pronounced reactivity with sulfite to form a reversible adduct at the flavin N5-position [12]. The *Ecass*GlpO and *Ssp*GlpO giving similar values of K_d = 0.82 mM and 1.5 mM [10, 11], and we here performed similar sulfite titrations with freshly prepared *Mp*GlpO and *Bp*3253 (Figure 1A). For *Mp*GlpO, the measured K_d was 3.0 mM, but flavin absorbance was only partially bleached with about half of the *Mp*GlpO flavin being refractory to sulfite adduct formation, even after an incubation of 30 min at 48 mM sulfite. A parallel experiment with the *Bp*3253 protein (Figure 1B) shows a monophasic sulfite titration giving essentially complete bleaching of the flavin absorbance spectrum with a K_d of 0.42 mM, about seven-fold more favorable than that for *Mp*GlpO. We do not have any good proposals at this time for why only roughly half of the *Mp*GlpO reacts with sulfite.

Structure determination of oxidized and reduced Mp*GlpO*. Crystals obtained of recombinant His-tagged *Mp*GlpO yielded diffraction data to ~2.5 Å resolution. Attempts to solve the structure by molecular replacement with a structure of *Ssp*GlpOΔ (PDB entry 2RGH) as the search model were unsuccessful, which was not surprising given the low ~20% sequence identity between *Mp*GlpO and *Ssp*GlpO. A PDB database query for better search models, led to the identification of the *B. pertussis* protein *Bp*3253 (PDB entry 3DME; 27% identity) as the known structure with the highest sequence similarity. Using this as a search model, molecular replacement was successful and led to ~2.5 Å resolution structures with acceptable statistics (Table 1) for both oxidized and dithionite-reduced *Mp*GlpO.

The refined models contain a single *Mp*GlpO chain in the asymmetric unit including residues 1-384, the FAD cofactor, ordered water sites, and a Ni^{2+} atom. The $2F_o$-F_c map for the FAD cofactor and some of its environment illustrates the quality of electron density in well-ordered regions of the protein (Figure 2). Missing from the models and presumed mobile are all 34 residues of the N-terminal His-tag and two side-chains; also, a few residues are modeled with

alternate conformations (see Experimental Procedures). The Ni^{2+} ion, identified using an X-ray fluorescence scan and presumably introduced during Ni-affinity chromatography, is situated at a crystallographic three-fold packing interface where it is coordinated by three His59-side chains. As the gel filtration result described above shows that the recombinant His-tagged *Mp*GlpO is a monomer, this three-fold interface and a two-fold crystal packing interaction involving strand β15 and burying ca. 1600 $Å^2$ of surface area are not physiologically relevant. So the nickel, possibly picked up during purification, can be considered a fortuitous crystallization aid.

Overall structure. As expected for a DAAO superfamily member, the *Mp*GlpO chain is organized into an FAD-binding domain with a predominantly parallel, six-stranded β-sheet and a substrate-binding domain with a core antiparallel, eight-stranded β-sheet (Figure 3). The two domains are discontinuous, with the FAD-binding domain including residues 1-87, 149-219, and 330-364, and the substrate-binding domain formed by residues 86-148 and 227-323.

A search using the DALI server [13] to identify structural homologs of *Mp*GlpO, confirms that there are no highly similar known structures. The most similar structure is that of the molecular replacement search model *Bp*3253 (PDB code 3DME; Z-score=44, 26% sequence identity and 2.0 Å rmsd over 352 residues), and a bit less similar is a glycine oxidase from *Bacillus subtilis* (PDB code 1NG4; Z-score=42; 20% sequence identity and 1.9 Å rmsd over 340 residues). An overlay with these proteins shows the high similarity extends throughout the protein chains (Figure 4A). Remarkably, the enzymes with more similarity in function, *Ssp*GlpO (PDB codes 2RGO) and *E. coli* GlpD (PDB code 2QCU), are structurally less similar, giving *Z*-scores of only 28 and 36, respectively (with higher rmsd values of 2.9 and 3.7 Å), and showing only ~17% sequence identity over the ~350 residues that are in the two domains common to the proteins. Visually, the overlay of *Mp*GlpO with *Ssp*GlpO and *E. coli* GlpD shows their greater divergence (Figure 4B). A structure-based sequence alignment of these proteins (Figure 4C) provides further details of the comparisons, and also is useful for tracking the similarities and differences in active site residues that will be discussed in the active site section.

That *Mp*GlpO is more similar to other DAAO superfamily enzymes than it is to *Ssp*GlpO led us to consider that whether GlpO activity may have independently evolved twice in the DAAO superfamily. To explore this possibility further we generated a relatedness tree of structurally known proteins similar to *Mp*GlpO (Figure 5). In this tree, despite the DALI scores, the *Mp*GlpO

and *Ssp*GlpO and *E. coli* GlpD actually do come out as being more closely-related to each other than to other functionally characterized DAAO superfamily members. The *Ssp*GlpO-like enzymes that cluster together include both GlpOs and the mitochondrial/bacterial GlpD dehydrogenases [14]. They represent a very widely distributed group that we are designating as 'Type I GlpO/DH' enzymes. In contrast, the *Mp*GlpO-like enzymes, which we are designating 'Type II GlpO/DH' enzymes, are more narrowly distributed, being found only bacteria of the class mollicutes (including *M. pneumoniae*) and the closely related low G+C Gram-positive anaerobic bacteria of the class Erysipelotrichia [15].

Interestingly, the Type II GlpO/DHs from anaerobic bacteria appear to all have an additional ca. 85 residue C-terminal domain that is not closely related to any known structure but includes two segments with conserved pairs of Cys residues (CxCE and CQxGFC) and has some similarity (e-value 4×10^{-10}) with the Pfam family of bacterioferritin-associated ferredoxin-like [2Fe-2S] domains (Pfam04324). The one biochemically studied Type II GlpO/DH containing such a C-terminal domain is the *E. coli glpA* gene product that is expressed under anaerobic conditions (replacing GlpD); and consistent with the assignment of the additional domain having an iron-sulfur center is the report that *E. coli* GlpA binds both FAD and non-heme iron [16]. Our proposal based on these observations is that these Type II GlpO/DHs with the additional C-terminal domain are all dehydrogenases for which the Type II GlpO/DH module converts Glp to DHAP, and the C-terminal domain serves as a conduit to receive electrons from the flavin and pass them on to a further (anaerobic) acceptor.

Flavin binding and active site. General features of FAD binding to DAAO superfamily enzymes have been well-described [7], so we will not detail those here but will focus in on the flavin and the substrate binding site. In *Mp*GlpO, the bound flavin is slightly twisted with a 3_{10}-helix involving residues Thr42 - Asn46 closely covering its *si* face, and Ser47 – Val49 interacting with the N5, O4, and N3 atoms of the flavin (Figure 2). Optimizing the hydrogen-bond between flavin-N3 and Val49-O appears to be a cause of the flavin twist. Above the flavin (on the *si* side), there is open space above and in front of atom N5, creating a cavity lined by His51, Arg320, and Ser348 (Figure 2). Interestingly, we do not see any substantive differences between the active sites of oxidized and reduced *Mp*GlpO structures at this resolution, even though the pale color of the reduced *Mp*GlpO crystals provided visual evidence that they truly

were reduced. Given the lack of differences, in the remaining discussion of the active site features, we will focus solely on the somewhat higher resolution oxidized structure.

Unfortunately, attempts to obtain an *Mp*GlpO structure with a substrate or substrate analog bound were unsuccessful (see methods), as was also true for our earlier structural work on *Ssp*GlpO [9]. However, the protein structure that is the most similar to *Mp*GlpO, that of *Bp*3253, fortuitously has a ligand – L-tartrate – bound in its active site. Although the function of *Bp*3253 is not yet known, it binds tartrate similarly to how glycine oxidase [17, 18] and DAAO [19] bind their substrate analogs (overlays not shown), implying that the *Bp*3253:tartrate complex is an informative one. Specifically, one tartrate α-carbon (i.e. alpha to one of the carboxylates) is placed just 3.7 Å from flavin N5 and with excellent geometry for hydride transfer (Figure 6).

An overlay of *Mp*GlpO with *Bp*3253 (Figure 6) shows that the flavins and most nearby peptide backbone segments align well. In the *Bp*3253 complex there are just six side chains (and no backbone atoms) making van der Waals or hydrogen-bond contacts with the tartrate. Starting with the key Arg sitting above the flavin C7 and C8 methyl groups, the *Bp*3253 residues, with their *Mp*GlpO equivalents in parentheses, are Arg316(Arg320), Pro272(Pro274), His259(Ile261), Tyr248(Phe250), His52(His51), and Ser350(Ser348). Hydrogen bonds connect Arg316 with the C1-carboxylate, His259 via a water molecule with the C3 hydroxyl, and Tyr248, His52, and Ser350 with the C4-carboxylate (Figure 6). Four of these positions are perfectly conserved in *Mp*GlpO, both in identity and potential placement, recognizing that Arg320 in *Mp*GlpO could easily shift to match the position seen for the equivalent *Bp*3253 residue. The conserved placement of Ser350(Ser348) projecting over the flavin depends on a conserved *cis*-peptide bond between the Ser and Pro351(Pro349). Notable differences in the active sites are the Tyr248->Phe replacement in *Mp*GlpO and the His259->Ile261 replacement that is compounded by a large difference in the path of the β12-β13 loop, with the backbone of *Mp*GlpO Gly259 occupying the space filled by the His259 side chain in *Bp*3253.

Regarding this β12-β13 loop difference, a key question is why the path of *Mp*GlpO residues 256-261 is different. Are they adopting an arbitrary conformation in this crystal form or a conformation that reliably represents the chain path of the Type II GlpO/DHs? Three evidences suggest it is representative rather than arbitrary: first, the *Mp*GlpO path is similar to the paths seen in broader superfamily members such as glycine oxidase meaning that it is the *Bp*3253 path that is unusual; second, the distinct *Bp*3253 path appears to be directly related to it having a

proline (Pro249) at the end of strand β12 that disrupts the normal β-sheet hydrogen bonding (Figure 6) and is not present in *Mp*GlpO (Figure 4C); and third, the loop is well ordered and not involved in crystal contacts.

Modeling Glp binding to type II GlpO/DHs. Given the reliability of the β12-β13 loop conformation, we can now consider how *Mp*GlpO binds substrate (Figure 7A). Compared with *Bp*3253, the presence of Phe250 and the Gly259 methylene make the pocket above the pyrimidyl portion of the flavin less polar, leading us to speculate that this region would recognize the C1-end of Glp and the region near Arg320 would accommodate the negatively charged phosphoryl group. As to what could provide additional recognition of the phosphoryl, the side chains of Arg230, Lys258 and Lys347 are nearby and appear adjustable (B-factors at the side-chain tips being ~20 $Å^2$ higher than at the backbone). These side chains are also well conserved among the 70 sequences having >40% sequence identity with *Mp*GlpO (i.e. BLAST e-value <10^{-80}): Lys258 is fully conserved; Lys347 is conserved as either a Lys or a Gln; and Arg230 is Arg or Lys in all but four sequences, but in these it is a Leu implying a lesser importance. Given these insights, we created a model for docked Glp guided by the constraints that the C2-hydrogen be oriented for transfer to the flavin-N5, that either the C1- or C2-hydroxyl replaces the water binding to His51 and Ser348 (see Figure 2), and the phosphate interacts with Arg320. The binding mode obtained (Figure 7A) fortuitously places His51 ideally for serving as a catalytic base that could deprotonate the C2-hydroxyl in the forward reaction.

Extrapolation of Glp binding to type I GlpO/DHs. Interestingly, this proposed mode of substrate binding differs from those proposed previously for the type I GlpO/DHs *Ssp*GlpO [9] and *Ec*GlpD [14], neither of which were very satisfactory. For *Ssp*GlpO, we had proposed that the substrate would be oriented the other way, with the 3'-phosphoryl moiety near the equivalent of *Mp*GlpO His51 and the C1-hydroxyl interacting with the equivalent of *Mp*GlpO Arg320 [9]. We had also noted that the equivalent of His51 seemed the only potential base, but how this could be achieved was not clear. For *Ec*GlpD [14], a phosphate bound in the native crystal structure interacted with Arg317 (equivalent of *Mp*GlpO Arg320), Arg54 and Tyr55, but ligand soaks were taken to indicate binding modes that placed the substrate C2-atom far from flavin-N5.

Given the unsatisfactory nature of those proposals, we explored whether the mode of binding we predict for *Mp*GlpO could also work for the Type I enzymes. An overlay of *Ec*GlpD onto

*Mp*GlpO with its docked Glp shows a remarkable compatibility, with the predicted position of the phosphoryl of Glp matching closely with the experimentally observed phosphate of *Ec*GlpD (Figure 7B). The ligand fits reasonably into the *Ec*GlpD active site, and it can be seen that the functionality of key *Mp*GlpO residues that are not conserved in *Ec*GlpD appears to be fulfilled by substitutions of residues often from different parts of the chain: with Lys354 replacing Ser348 in interacting with the C2-hydroxyl, Arg54 and Tyr55 replacing Lys258 and Lys347 in interacting with the phosphoryl, and Phe257 replacing Phe250 in providing a non-polar environment for the C1-methylene (Figure 7B). These functional substitutions are most easily seen in schematic drawings of the interactions (Figure 8). We emphasize that whereas the general features associated with this rough placement of Glp are plausible and have explanatory power, the details are not reliably defined because the side chain and backbone positions of protein groups are expected to shift during ligand binding from the positions they adopt in the unliganded *Mp*GlpO structure used to guide the modeling.

Catalytic mechanism and outlook. Given the predicted mode of Glp binding to the *Mp*GlpO (Figure 8A) and the *SSp*GlpO (Figure 8B) active sites, how the electrons flow during catalysis and how the forward reaction is enhanced by base catalysis become readily apparent. As indicated by the arrows in the Figure 8A schematic, in *Mp*GlpO the substrate is well-aligned for His51 to deprotonate the C2-hydroxyl, promoting its electrons to fold in to form a carbonyl, and facilitating loss of the C2-hydrogen as a hydride that can attack the flavin N5; the ensuing shifting of flavin electrons would lead to the formation of the reduced flavin N1-anionic form, with the negative charge at the N1/O2 locus stabilized by three hydrogen bonds donated by the Leu351 and Thr352 backbone amides and the Thr352 hydroxyl. The geometries of the interactions are also stereoelectronically reasonable, in that the His51 interaction with the C2-hydroxyl has a roughly "anti" orientation with respect to the hydride leaving group. The Ser47 hydroxyl – supported by its interactions with Tyr234 – is well-placed to help stabilize the protonated N5 of the reduced flavin through minor shifts in its position in association with flavin reduction. A set of equivalent interactions exist in the *SSp*GlpO active site (Figure 8B).

A full kinetic analysis of *Mp*GlpO has recently been completed (Maenpuen et al; submitted for publication in the same issue), and our structural results allow us to propose explanations for some of the observations made in that study. One such area is the redox potential difference

between the *Mp*GlpO flavin, which at -167 mV is much lower than those seen for type I GlpOs (e.g. *Enterococcus casseliflavus* GlpO at -118 mV; [10]). While many subtle factors can influence redox potential, we note that one very clear difference in the flavin electrostatic environment consistent with this shift is that the Type I GlpO/DH enzymes conserve an active site lysine (equivalent to Lys354 in *Ec*GlpD; Figure 8B, 7B, 4C) having its amino group just at van der Waals distance above the flavin N1/O2 locus, whereas the type II enzymes have a conserved neutral serine side chain in that place (equivalent to Ser348 in *Mp*GlpO). The additional local positive charge in the type I enzymes would make them more easily reduced. As was pointed out by Maenpuen et al, the difference in redox potential may be the reason that *Mp*GlpO can catalyze the reverse reaction (i.e. DHAP oxidation of the reduced flavin) whereas those type I enzymes tested cannot.

Another very interesting result of the kinetics study was the observation of two enzyme populations, partitioning as 70:30 with 70% of the oxidized enzyme reacting rapidly with Glp, and 70% of the reduced enzyme reacting slowly with DHAP in the reverse reaction, but reacting rapidly with O_2. Based on the structure, a plausible single explanation for all of these observations is that the two populations are defined by the protonation state of the catalytic acid/base His-51, with 70% being deprotonated and 30% protonated at the pH of 7 at which the studies were done. For the reaction with Glp, the deprotonated form of His51 is required for abstracting the C2-hydroxyl proton, so 70% of the enzyme would react rapidly. The slower population would reasonably be limited by the rate of deprotonation of His51 in the ligand bound form. For the reaction of the reduced flavin with DHAP, the deprotonated 70% of the enzyme reacts poorly (in bimolecular limited fashion at 56 M^{-1} s^{-1}), while the protonated 30% reacts rapidly enough that the k_{off} of the Glp produced (at 6 s^{-1}) is rate limiting. In terms of the O_2 reactivity of the reduced enzyme, it is reasonable that the His51-protonated/charged and the His51-deprotonated/neutral active sites would react differently and our proposal implies that the deprotonated 70% reacts rapidly with oxygen (at ~600 s^{-1}) while the protonated 30% reacts more slowly (at ~100 s^{-1}). This explanation predicts that the enzyme will be more active above pH=7, and a set of assays looking at the enzymes pH dependence at a single high substrate concentration shows the pH optimum is near 8. More extensive studies of k_{cat} and K_M as a function of pH are now planned to test these ideas and better define the enzyme mechanism.

The proposed binding mode of Glp also implies the importance of the negatively charged phosphoryl group for substrate recognition, as both the Type I and Type II enzymes have 3 to 4 positively charged side chains involved in its recognition (Figure 8). The importance of the phosphoryl is consistent with the binding studies reported by Maenpuen et al (submitted) showing that *Mp*GlpO effectively binds the negatively charged glyceraldehyde-3-phosphate, lactate, and malate, but not the neutral glycerol. While much remains to be done, we see the most influential contributions of these analyses are two-fold: first, is the recognition of the Type I and Type II GlpO/DH enzymes as distinct variants with striking active site differences that were not predicted from sequence alignments; and second, is the plausible concrete predicted mode for substrate recognition and catalysis that is relevant not only for bacterial GlpO enzymes, but also for the widespread mitochondrial GlpD dehydrogenases, and that will guide mutational studies that test the proposed mechanism and dissect the roles of specific active site residues. Also, the distinction between the Type I and Type II GlpO/DH active sites raises the encouraging possibility that it will be possible to design inhibitors that may selectively block activity of bacterial Type II enzymes such as *Mp*GlpO while not inhibiting the mitochondrial GlpD dehydrogenase of the host.

EXPERIMENTAL PROCEDURES

Expression and Purification of Mp*GlpO and* Bp*3253.* The codon-optimized MPN051 gene encoding *Mp*GlpO was synthesized by GenScript (Piscataway, NJ, USA) and subcloned into the expression plasmid pET28a (Novagen, Darmstadt, Germany). *Mp*GlpO was expressed with an N-terminal His-tag in *E. coli* B834(DE3) cells using autoinduction medium at 28ºC. All steps of purification were conducted at 4ºC. Harvested cells were resuspended in 50 mM potassium phosphate, pH 7.0, containing 200 mM NaCl, 20 mM imidazole, 0.5 mM 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF) protease inhibitor and 10 % glycerol. Cells were disrupted using an Avestin EmulsiFlex-C5 homogenizer, and centrifuged (27,000 g for 60 min). The clarified extract was loaded onto a 25 mL Co Sepharose High Performance column (GE Healthcare, Piscataway, NJ, USA), and *Mp*GlpO protein was eluted with 0.5 M imidazole after washing with 20 mm imidazole. The pooled yellow fractions were dialyzed overnight against 50 mM potassium phosphate, pH 7.0, containing 0.5 mM EDTA. FAD (0.25 mM) was added to the enzyme and incubated for 45 min on ice before loading on a 75-mL SP-Sepharose HP column

(GE Healthcare). The column was washed with 50 mM potassium phosphate, pH 7.0, 0.5 mM EDTA, 100 mM NaCl before *Mp*GlpO was eluted with a 100 mM to 1 M NaCl gradient. Fractions were analyzed by SDS-PAGE, pooled, buffer-exchanged into 50 mM potassium phosphate, pH 7.0, 0.5 mM EDTA and concentrated to 10 mg/mL before freezing in aliquots at -80ºC.

The pET21_NESG clone for expression of His-tagged *Bp*3253 (corresponding to PDB entry 3DME) was purchased from the DNASU Plasmid Repository, and its expression and purification followed the protocol described above for *Mp*GlpO, with the following modifications. Recombinant *E. coli* B834(DE3) cells were grown in TYP medium, prior to induction with 0.5 mM IPTG and overnight protein expression at 16 °C. The cells were broken in a solution containing 50 mM Tris-Cl, pH 8.0 (4 °C). Nucleic acids were removed by adding 2% (w/v) streptomycin sulfate and centrifugation. Crude extract was loaded to the Co column in 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 20 mM imidazole, the column was washed with more buffer, and pure *Bp*3253 was eluted with 500 mM imidazole. *Bp*3253 protein was concentrated to 10 mg/ml in an Amicon ultrafiltration cell with a YM30 membrane and for freezing was buffer-exchanged into 50 mM potassium phosphate, pH 7.0, containing 0.5 mM EDTA.

Biochemical characterizations. Extinction coefficients for His-tagged *Mp*GlpO and *Bp*3253 were determined by standard methods, using an Agilent model 8453 diode-array spectrophotometer. The specific activities of the two proteins were measured as described earlier for *Ssp*GlpO [11], using the standard spectrophotometric assay with a Cary 50 spectrophotometer (Varian) thermostatted at 25ºC. The native M_r for recombinant His-tagged *Mp*GlpO was determined by gel filtration at 25ºC, in a pH 7 phosphate buffer with 150 mM NaCl, using a Sephadex G-200 medium HR 10/30 column calibrated with six standard proteins covering the range 15,600 to 440,000. Titrations of both MpGlpO and Bp3253 proteins with sulfite followed established protocols [10, 11].

Crystallization and Data Collection. Thawed aliquots of wild-type His-tagged *Mp*GlpO were subjected to a variety of crystal screens at 4 ˚C and the most promising crystal leads grew using a reservoir of 2.68 M NaCl, 3.35% v/v isopropanol, and 0.1 M HEPES pH 7.5 (Wizard Precipitant Synergy Screen Block 1 condition B7). Optimization at 4 ºC using hanging-drops led to yellow, trigonal-pyramidal crystals measuring approximately 0.3 x 0.3 x 0.3 mm^3 growing within one week using drops made from 1 μL protein at 5 mg/ml plus 2 μL of a reservoir solution as above

but containing 2% (v/v) isopropanol. For data collection, crystals in an artificial mother liquor (AML) of 3 M NaCl in 0.1 M Bis-Tris, pH 7.0 were placed for 3 minutes in AML with 15% glycerol as a cryoprotectant before being flash-frozen by plunging into liquid nitrogen. Data sets to first 2.50 Å and then 2.40 Å resolution were collected at beamline 5.0.3 at the Advanced Light Source synchrotron. For analyses of additional forms of *Mp*GlpO, crystals were soaked in either 10 mM dithionite, 10 mM L-Glp, 10 mM L-tartrate, 10 mM 2-phosphoglycerate, or 10 mM phosphoenolpyruvate (PEP) in AML for 1 h prior to freezing and data collection. Two experiments with dithionite-soaked crystals gave a complete merged data set to 2.50 Å. Data sets at between 2.5 and 3.0 Å resolution were collected for the ligand soaks.

Mp*GlpO Phasing and Structure Refinement.* Diffraction data for oxidized *Mp*GlpO were processed with iMOSFLM [20] and with SCALA in the CCP4 suite [21, 22]. *Mp*GlpO crystallizes in space group *P*23, and there is one molecule per asymmetric unit. Attempts to solve the structure by molecular replacement using AMoRe with the *Ssp*GlpO coordinates were not successful. However, using chain A of PDB entry 3DME – annotated as the "conserved exported protein (BP3253) from *Bordetella pertussis*" – as a search model gave a solution using the 2.50 Å data set and AutoMR [23]. Manual rebuilding of the initial model was carried out in COOT [24], and this model was refined with the PHENIX software suite [25]. Six rounds of simulated annealing and minimization refinement gave a partially refined model with R/R_{free} of 32%/45%, respectively. At this point, the improved 2.40 Å resolution *Mp*GlpO data were used, and six rounds of refinement with BUSTER [26] led to an R_{free} of 28%. Water molecules were added, as indicated by both electron density peak height and hydrogen-bonding interactions, and refinement continued with REFMAC [27]. A nickel ion (see "Results") was introduced late in the process and was confirmed by X-ray fluorescence, and further manual modeling and refinement led to the final oxidized *Mp*GlpO structure with R/R_{free} 15.8%/16.4%. The model does not include the Lys79 and Lys298 side chains beyond -CB, as these surface residues have little-or-no side chain density. Alternate side chain and/or backbone conformations are included for Gln40, His244, and the Trp375-Asn376-Gly377 backbone.

The 2.50 Å refinement of the dithionite-reduced GlpO structure began from the oxidized structure. Rounds of manual modeling and REFMAC refinement gave a model with only minor changes, and with R/R_{free} values of 16.4%/22.6%. Of the soaks with Glp, 2-phosphoglycerate,

PEP, and tartrate, no substantive interpretable density differences were observed in the vicinity of the isoalloxazine and so these structural analyses were not pursued further.

AUTHOR CONTRIBUTIONS

CKE Planned and performed experiments and analyzed data; KK Analyzed data and wrote the paper; DP Planned and performed experiments, contributed reagents, and analyzed data; SM and PC Planned and performed experiments and analyzed data; AC and PAK Planned experiments, analyzed data and wrote the paper.

REFERENCES

1. Waites, KB & Talkington, DF (2004) Mycoplasma pneumoniae and its role as a human pathogen. *Clin Microbiol Rev* **17**, 697-728.

2. Prince, OA, Krunkosky, TM & Krause, DC (2014) In vitro spatial and temporal analysis of Mycoplasma pneumoniae colonization of human airway epithelium. *Infect Immun* **82**, 579-86.

3. Pilo, P, Vilei, EM, Peterhans, E, Bonvin-Klotz, L, Stoffel, MH, Dobbelaere, D & Frey, J (2005) A metabolic enzyme as a primary virulence factor of Mycoplasma mycoides subsp. mycoides small colony. *J Bacteriol* **187**, 6824-31.

4. Hames, C, Halbedel, S, Hoppert, M, Frey, J & Stulke, J (2009) Glycerol metabolism is important for cytotoxicity of Mycoplasma pneumoniae. *J Bacteriol* **191**, 747-53.

5. Schmidl, SR, Otto, A, Lluch-Senar, M, Pinol, J, Busse, J, Becher, D & Stulke, J (2011) A trigger enzyme in Mycoplasma pneumoniae: impact of the glycerophosphodiesterase GlpQ on virulence and gene expression. *PLoS Pathog* **7**, e1002263.

6. Grosshennig, S, Schmidl, SR, Schmeisky, G, Busse, J & Stulke, J (2013) Implication of glycerol and phospholipid transporters in Mycoplasma pneumoniae growth and virulence. *Infect Immun* **81**, 896-904.

7. Dym, O & Eisenberg, D (2001) Sequence-structure analysis of FAD-containing proteins. *Protein Sci* **10**, 1712-28.

8. Mahdi, LK, Wang, H, Van der Hoek, MB, Paton, JC & Ogunniyi, AD (2012) Identification of a novel pneumococcal vaccine antigen preferentially expressed during meningitis in mice. *J Clin Invest* **122**, 2208-20.

9. Colussi, T, Parsonage, D, Boles, W, Matsuoka, T, Mallett, TC, Karplus, PA & Claiborne, A (2008) Structure of alpha-glycerophosphate oxidase from Streptococcus sp.: a template for the mitochondrial alpha-glycerophosphate dehydrogenase. *Biochemistry* **47**, 965-77.

10. Parsonage, D, Luba, J, Mallett, TC & Claiborne, A (1998) The soluble alpha-glycerophosphate oxidase from Enterococcus casseliflavus. Sequence homology with the membrane-associated dehydrogenase and kinetic analysis of the recombinant enzyme. *J Biol Chem* **273**, 23812-22.

11. Charrier, V, Luba, J, Parsonage, D & Claiborne, A (2000) Limited proteolysis as a structural probe of the soluble alpha-glycerophosphate oxidase from Streptococcus sp. *Biochemistry* **39**, 5035-44.

12. Massey, V, Muller, F, Feldberg, R, Schuman, M, Sullivan, P A, Howell, LG, Mayhew, SG, Matthews, RG & Foust, GP (1969) The reactivity of flavoproteins with sulfite: Possible relevance to the problem of oxygen reactivity. *J Biol Chem* **244**, 3999-4006.

13. Holm, L & Rosenstrom, P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-9.

14. Yeh, JI, Chinte, U & Du, S (2008) Structure of glycerol-3-phosphate dehydrogenase, an essential monotopic membrane enzyme involved in respiration and metabolism. *Proc Natl Acad Sci USA* **105**, 3280-5.

15. Davis, JJ, Xia, F, Overbeek, RA & Olsen, GJ (2013) Genomes of the class Erysipelotrichia clarify the firmicute origin of the class Mollicutes. *Int J Syst Evol Microbiol* **63**, 2727-41.

16. Schryvers, A & Weiner, JH (1981) The anaerobic sn-glycerol-3-phosphate dehydrogenase of Escherichia coli. Purification and characterization. *J Biol Chem* **256**, 9959-65.

17. Mortl, M, Diederichs, K, Welte, W, Molla, G, Motteran, L, Andriolo, G, Pilone, MS & Pollegioni, L (2004) Structure-function correlation in glycine oxidase from Bacillus subtilis. *J Biol Chem* **279**, 29718-27.

18. Settembre, EC, Dorrestein, PC, Park, JH, Augustine, AM, Begley, TP & Ealick, SE (2003) Structural and mechanistic studies on ThiO, a glycine oxidase essential for thiamin biosynthesis in Bacillus subtilis. *Biochemistry* **42**, 2971-81.

19. Umhau, S, Pollegioni, L, Molla, G, Diederichs, K, Welte, W, Pilone, MS & Ghisla, S (2000) The x-ray structure of D-amino acid oxidase at very high resolution identifies the chemical mechanism of flavin-dependent substrate dehydrogenation. *Proc Natl Acad Sci USA* **97**, 12463-8.

20. Battye, TG, Kontogiannis, L, Johnson, O, Powell, HR & Leslie, AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271-81.

21. Winn, MD, Ballard, CC, Cowtan, KD, Dodson, EJ, Emsley, P, Evans, PR, Keegan, RM, Krissinel, EB, Leslie, AG, McCoy, A, McNicholas, SJ, Murshudov, GN, Pannu, NS,

Potterton, EA, Powell, HR, Read, RJ, Vagin, A & Wilson, KS (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42.

22. Evans, PR (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72-82.

23. McCoy, AJ, Grosse-Kunstleve, RW, Adams, PD, Winn, MD, Storoni, LC & Read, RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674.

24. Emsley, P, Lohkamp, B, Scott, WG & Cowtan, K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501.

25. Adams, PD, Afonine, PV, Bunkoczi, G, Chen, VB, Davis, IW, Echols, N, Headd, JJ, Hung, LW, Kapral, GJ, Grosse-Kunstleve, RW, McCoy, AJ, Moriarty, NW, Oeffner, R, Read, RJ, Richardson, DC, Richardson, JS, Terwilliger, TC & Zwart, PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21.

26. Bricogne, G, Blanc, E, Brandl, M, Flensburg, C, Keller, P, Paciorek, W, Roversi, P, Sharff, A, Smart, OS, Vonrhein, C & Womack, TO (2011) BUSTER version 2.11.2 in Global Phasing Ltd., Cambridge, UK.

27. Murshudov, GN, Vagin, AA & Dodson, EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55.

28. Guindon, S, Dufayard, JF, Lefort, V, Anisimova, M, Hordijk, W & Gascuel, O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**, 307-21.

Scheme 1
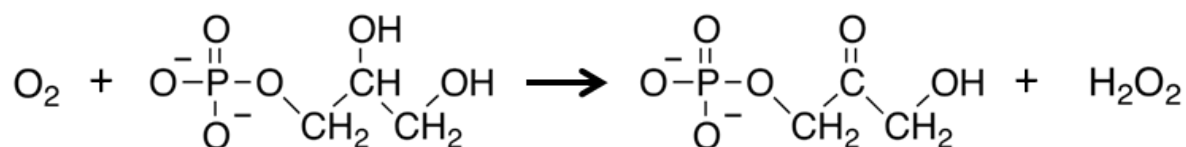
Table 1. Data Collection and Refinement Statistics for *Mp*GlpO structures. Numbers in parentheses represent data for the high-resolution shell.

	oxidized	reduced
PDB entry code	**4X9M**	**4X9N**
Data quality statistics		
Wavelength (Å)	0.9765	0.9765
Space group	*P*23	*P*23
Cell dimensions, $a = b = c$ (Å)	111.59	111.61
Resolution range (Å)	50-2.4 (2.53-2.4)	60-2.5 (2.64-2.5)
Reflections	430,576	836,296
Unique reflections	18,401	16,368
Completeness (%)	100 (99)	100 (100)
Multiplicity	23.4 (23.6)[a]	51.1 (28.6)
R_{pim}	0.022 (0.19)	0.072 (0.28)
R_{meas}	0.109 (0.90)	0.533 (1.51)
I/σ	22.7 (4.4)	19.9 (3.0)
Refinement statistics		
R_{work} (%)	14.8 (21.9)	14.7 (23.5)
R_{free} (%)	20.4 (25.3)	21.1 (34.3)
No. amino acid residues	384	384
No. solvent atoms	200	200
No. non-hydrogen atoms	3301	3313
$\langle B \rangle$ protein (Å^2)	35	39

FIGURE LEGENDS

Figure 1. Sulfite titrations of *Mp*GlpO and *Bp*3253. **A.** MpGlpO [35.4 μM in 0.8 mL of 50 mM potassium phosphate, 0.5 mM EDTA, pH 7.0] was titrated aerobically with a 1 M sodium sulfite solution. The spectra shown correspond to the addition of 0 (black), 1.43 (blue), 4.27 (green), 11.3 (orange) and 47.6 mM (red) total sulfite. The inset shows the absorbance at 450 nm as a function of added sulfite. **B.** The titration of *Bp*3253 (36.2 μM in 1 mL) was carried out as described in (A), with 25 and 100 mM solutions of sodium sulfite. The spectra shown are with 0 (black), 0.15 (blue), 0.50 (green), 0.80 (orange) and 2.0 mM (red) total sulfite. The inset shows the absorbance changes at 449 nm as sulfite is added.

Figure 2. Electron density quality for the flavin and nearby side chains. Stereoview of the final $2F_o$-F_c electron density map (cyan; contoured at $2.0\rho_{rms}$) for the *Mp*GlpO flavin (yellow carbons) and the adjacent protein atoms (off-white carbons) and one water (red sphere). Hydrogen bonds (dashed lines) to the flavin and the water are indicated. Residues are labeled.

Figure 3. Tertiary structure of the *Mp*GlpO monomer. Stereo ribbon diagram of the *Mp*GlpO monomer showing the FAD (sticks with yellow carbons) and labeling the secondary structural elements. The FAD-binding domain is at the bottom and the substrate-binding domain is at the top.

Figure 4. Comparisons of the *Mp*GlpO structure and sequence with select homologs. **A.** The flavin region in an overlay of *Mp*GlpO (off-white protein and yellow FAD) on *Bp*3253 (violet; PDB entry 3DME) and glycine oxidase (salmon; PDB entry 1RYI; [17]). For this view the Figure 3 molecule was rotated 180˚ around a vertical axis (i.e. this view is from the back of that image). Note the similar paths of the loop in front of the flavin, in *Mp*GlpO containing Ser 348. **B.** Same as (A) but overlaying *Mp*GlpO (colored as in A) with a form of *Ssp*GlpO missing a 50 residue segment (blue; PDB entry 2RGH; [9]) and *Ec*GlpD (cyan; PDB entry 2QCU; [14]). Note the different paths of the loop in front of the flavin of the type I enzymes compared with *Mp*GlpO. **C.** Structure-based sequence alignment of the five enzymes shown in panels (A) and (B). Conserved residues (*) and residues involved in β-strands (yellow), α-helices (cyan) and 3_{10}-helices (blue) are indicated. Also highlighted are residues discussed as important for in flavin binding (green boxes; and including the *cis*-Pro in *Mp*GlpO and *Bp*3253) and substrate binding

(red boxes). Dots above the *Mp*GlpO sequence mark its every 10th residue, and at the end of each line is a numbered residue for each of the sequences.

Figure 5. Relatedness tree of structurally known DAAO superfamily members most similar to *Mp*GlpO. A DALI [13] search in November 2014 using the PDB90 database option and oxidized *Mp*GlpO as the search model, provided a gap-removed alignment for hits with Z-scores higher than 20. These were used to generate a tree with PhyML [28]. Branches are labeled with individual PDB entry names and known enzyme types are indicated.

Figure 6. Comparison of the *Bp*3253 tartrate binding pocket and *Mp*GlpO. Stereoview shown of the *Bp*3253-tartrate complex (violet carbons) overlaid on the equivalent region of *Mp*GlpO (semi-transparent off-white carbons for protein and yellow for flavin). Also indicated are H-bonds (grey dashed lines for *Mp*GlpO, violet dashed lines for *Bp*3253-tartrate complex) and the close approach of the tartrate C3 atom to the flavin N5 (green thick dotted line). Select residues in *Bp*3253-tartrate complex (violet) and *Mp*GlpO (grey) are identified. H259 in the *Bp*3253-tartrate complex and I261 in *Mp*GlpO are equivalent residues, but their labels are placed near their respective side chains.

Figure 7. A predicted *Mp*GlpO-Glp complex and its comparison with *Ec*GlpD. **A.** Stereoview of the *Mp*GlpO active site with a roughly positioned Glp-bound (off-white carbons for protein, yellow for flavin, and green for Glp). H-bonds (grey dashed lines) and the close approach of the modeled Glp C2 atom to the flavin N5 (green thick dotted line) are also shown. Criteria used for placing the Glp are described in the text. **B.** Stereoview of the *Mp*GlpO (colored as in A, but semitransparent and without H-bonds shown) overlaid on the equivalent region of *Ec*GlpD (cyan carbons) with its bound inorganic phosphate (cyan phosphorus) and H-bonds (cyan dashed lines). Select residues in *Ec*GlpD (cyan) and *Mp*GlpO (grey) are identified.

Figure 8. Schematic drawings of residues involved in substrate binding and catalysis in *Mp*GlpO and *Ec*GlpD. **A.** The *Mp*GlpO active site atoms roughly in the plane of the flavin are shown smaller and with thinner bonds and hydrogen bonds (dashed), and atoms in front of the flavin are shown with thicker bonds and hydrogen bonds (dashed). Residues shown interacting with the flavin and the Glp are in Figure 4C highlighted in red and green boxes, respectively,. Curved

arrows indicate the proposed flow of electrons during the reductive half-reaction. **B.** The same except for showing the *Ec*GlpD active site as a representative of a Type I GlpO.
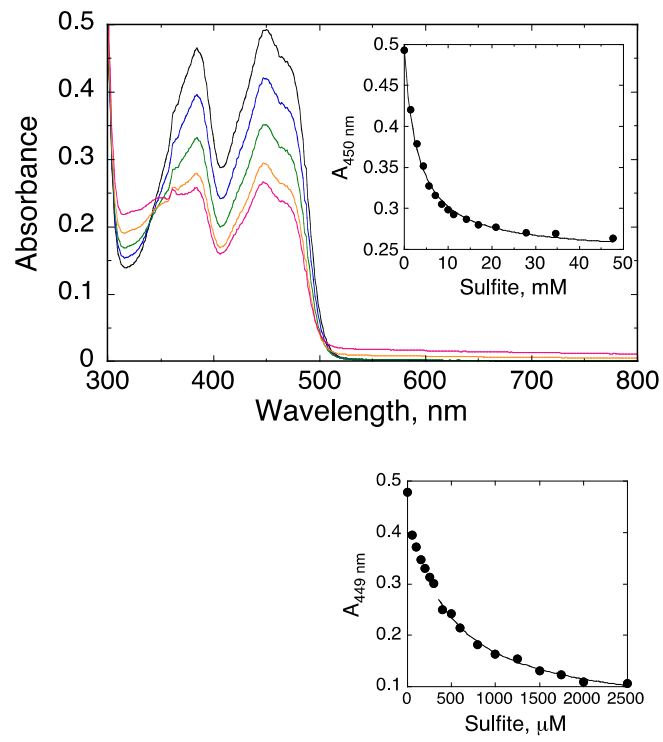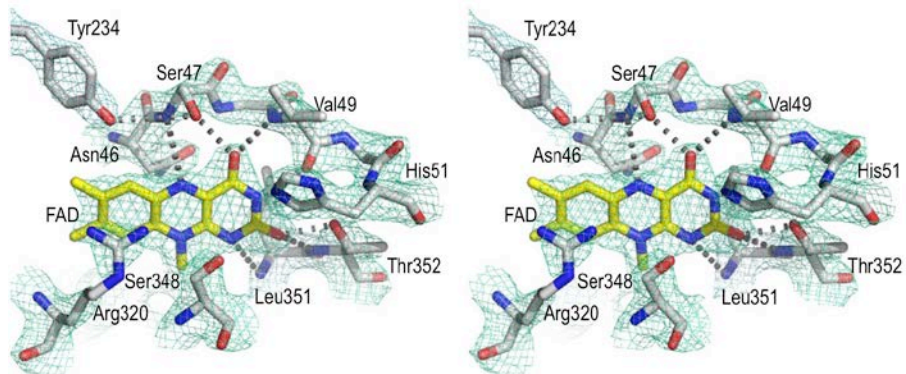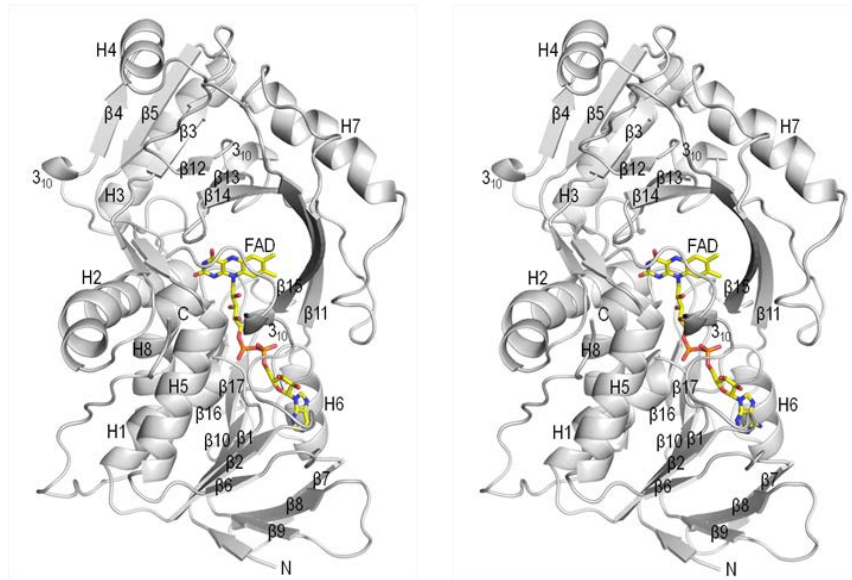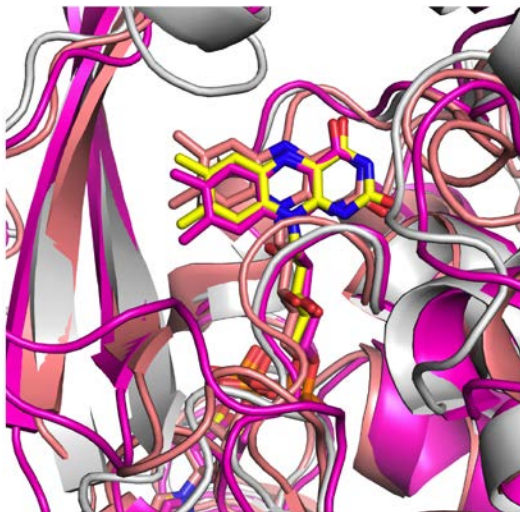
Figure 1

Figure 2

Figure 3

A

Figure 4

A

B

C

β β1   α1   β2   α2

```
MpGlpO  ----------------METRDVLIVGGGVIGCATAYELSQYKLKVTLVEKHHYLAQETSHANSGVIHTGIDPNP-HKLTAKYNILGKKLWLNTYFKR------ 80
Bp3253  ---------------msTDIDCIVIGAGVVGLAIARALAAGGHEVLVAEAAEGIGTGTSSRNSEVIHAGIYYPA-DSLKARLCVRGKHLLYEYCAAR------ 81
SsGlpO  CMFSNKTRQDSIQKMQQELDLLIIGGGITGAGVAVQAAASGIKTGLIEMQDFAEG-TSSRSTKLVHGGLRYLKTFDVEVVADTVGERAVVQGIAP------ 95
EcGlpD  ----------METKDLIVIGGGINGAGIAADAAGRGLSVLMLEAQDLACA-TSSASSKLIHGGLRYLEHYEFRLVSEALAEREVLLKMAP------- 79
BsGlyOx  ---------------MKRHYEAVVIGGGIIGSAIAYYLAKENKNTALFESGTMGGR-TTSAAAGMLG-AHACEERDAFFDFAMHSQRLYKGLGEELYALSGV 86
             *   *  *      *           *           *      *                    *
```

β β3   α3   β4 α4   β5 β   α5   β6 β7

```
LGFPRQKIRTLIVAF-----NEMEREQLEVLKQRGIANQIN--LEDIQMLSKEETLKLEPYVNPEIVAGLKIEGSWAIDPVLASKCLALAAQQNKVQICTNTEVTNISKQV 184
-GVPHQRLGKLIVAT-----SDAEASQLDSIARRAGANGV--DDLQHIDGAAARRLEPALH---CTAALVSPSTGIVDSHALMLAYQGDAESDGAQLVFHTPLIAGRVRP 180
-HIPKPDPMLLPIYEDEGATTFNMFSVKVAMDLYDKLANVTGTKYENYTLTPEEVLEREPFLKKEGLKGAGVYLDFRNNDARLVIDNIKKAAEDGAYLVSKMKAVGFLYEG 205
--IAFPMRFRLPHR-----PHLRPAWMIRIGLFMYDHLGKRTSLPGSTGLRF---GANSVLKPEIKRGFEYSD--CWVDGDARLVLANAQMVVRKGGEVLTRTRATSARRE- 178
DIRQHNGGMFKLAFS-----EEDVLQLRQMD--------DLDSVSWYSKEEVLEKEPYASGDIFGASFIQDDVHVEPYFVCKAYVKAAKMLGAEIFEHTPVLHVERD- 180
```

β8   β9 β10 3₁₀ α6   β11   β12   β13   β14

```
DGTYLVWT---NNETTPSFKVKKIIDAAGHYADYLAHLA----KADDFEQTTRRGQYVVVTNQGELHLNSMVFMVPTIH-GKGVISPMLDGFLVGPTALDGVDKEATRYI 288
EGGFELDF----GGAEPMTLSCRVLINAAGLHAPGLARRIEGIPRDSIPPEYLCKGSYFTLAG--RAPFSRLIYPVPQHA-GLGVHLTLDLGGQAKFGPDTEWIA-TEDYTLD 285
DQIVGVKARDLLTDEVIEIKAKLVINTSGPWVDKVRNLNFT--RPVSPKMRPTKGILIKGSHIVVP---RVHTQKQAYILQNE-DKRIVFAIPR-ENKTYFGTTDTDYQGDFTDPKV 313
NGLWIVEAEDIDTGKKYSWQARGLVNATGPWVKQFFDDGMH--LPSSPYGIRLIKGSHIVVP---RVHTQKQAYILQNE-DKRIVFVIPWMDEFSIIGTTDVEYKGDPKAVKI 284
GEALFIKT------PSGDVWANHVVVASGWSGMFFKQLGL-----NNAFLPVKGECLSVW--NDDIPLTKTLYH-------DHCYIVPRKSGRLVVGATMKPGDW---SETP 270
                                                                          *
```

α7   β15   β16   β17   α8

```
TKDAPCMLTKIGKHMVPSLNIN--NALISFAGSRPIDKA-------TNDFIIRVAHN---DPDFVILGGMKSPGLTAAPAIVREREAVRLLNWKLTKKPNWNGKYNLPWI 384
PRRADVFYAAVRSYWPALP---DGALAPGYTGIRPKISGPHEPAAD----FAIAGPASHGV-AGLVNLYGIESPGLTASLAIAEETLARLAa 369
TQEDVDYLLDVINHRYPEANITLADIEASWAGLRPLLIgnsgspstisrGSSLER----EPDGLLTLSGG---KITDYRKMAEGALRLIRQLLKEEYGIETKEIDSKKY 465...
EESEINYLLNVYNTHFKK-QLSRDDIVWTYSGVRPLCDDESDSPQAITRDYTLDIHDENGKAPLLSVFGG---KLTTYRKLAEHALEKLTPYYQGIGPAWTKESVLPGG 389...
DLGGGLESVMKKAKTMLPAIQNM--KVDRFWAGLRPGTKD--------GKPYIGRHPE---DSRILFAAGHFRNGILAPATGALISDLIMNKEVNQDWLHAFRIDRKeavni 369
```

Figure 5

fructosamine oxidase

3C4N

3DJE

piperideine-2-carboxylate oxidase

D-Arg DH

3HZL

L-Pro DH   3AXB

3NYE

type II Glp oxidase/DH

2UZZ   N-methyl-Trp oxidase

4X9M

3DME

2GB0
1ZOV   sarcosine oxidase

2RGO

3AD7

3DA1

1Y56   L-Pro DH

type I Glp oxidase/DH

4PAB

2QCU

1NG3

1PJ6   glycine oxidase

3PVC

cmnm5U34-tRNA amine oxidase

3PS9

1C0P

3WGT   2E49

3ATQ

D-amino acid oxidase

3OZ2

geranylgeranyl reductase

25

Figure 6

Figure 7

A

B

Figure 8

A

B