

The structure of sedoheptulose-7-phosphate isomerase from *Burkholderia pseudomallei* reveals a zinc binding site at the heart of the active site

Nicholas J. Harmer¹

From: ¹School of Biosciences, University of Exeter, Exeter, U.K.

Author for correspondence:

Nicholas Harmer,
School of Biosciences,
Geoffrey Pope Building,
Stocker Road,
Exeter EX4 4QD,
U.K.

E-mail: N.J.Harmer@exeter.ac.uk

Tel: +44 (0) 1392 725197

Fax: +44 (0) 1392 263489

Abstract

Heptoses are found in the surface polysaccharides of most bacteria, contributing to structures that are essential for virulence and antibiotic resistance. Consequently, the biosynthetic enzymes for these sugars are attractive targets for novel antibiotics. The best characterized biosynthetic enzyme is GmhA, which catalyzes the conversion of sedoheptulose-7-phosphate into D-*glycero*-D-*manno*-heptopyranose-7-phosphate, the first step in the biosynthesis of heptose. Here, the structure of GmhA from *Burkholderia pseudomallei* is reported. This enzyme contains a zinc ion at the heart of its active site: this ion stabilizes the active, closed form of the enzyme, and presents co-ordinating side chains as a potential acid and base to drive catalysis. A complex with the product demonstrates that the enzyme retains activity in the crystal, and so suggests that the closed conformation is catalytically relevant, and is an excellent target for the development of therapeutics. A revised mechanism for the action of GmhA is postulated on the basis of this structure and the activity of *B. pseudomallei* GmhA mutants.

Keywords:

Burkholderia pseudomallei, capsule biosynthesis, GmhA, zinc binding, isomerase.

Abbreviations used:

2O6dDHep, 2-*O*-acetyl-6-deoxy- β -D-*manno*-heptopyranose; 6dDHep, 6-deoxy-D-*manno*-heptopyranose; CPS, capsular polysaccharide; CPS-I, type I capsular polysaccharide; DHep, L-*glycero*- α -D-*manno*-heptopyranose; EXAFS, extended X-ray absorption fine structure; LPS, lipopolysaccharide; M7P, D-*glycero*-D-*manno*-heptopyranose-7-phosphate; S7P, sedoheptulose-7-phosphate.

Introduction

The bacterium *Burkholderia pseudomallei* is a major form of community acquired infection in Thailand and Northern Australia.¹⁻³ In acute cases, mortality ranges from 20-50 %, depending on the treatment regimen.⁴ *B. pseudomallei* is also endemic over much of the tropics,^{1,5,6} and likely causes undiagnosed disease over much of South-East Asia. This bacterium shows innate resistance to most antibiotics,⁷ has no available vaccine, and is extremely persistent, with infections developing up to 63 years after exposure.⁸ Both *B. pseudomallei*, and the related bacterium *B. mallei*, are considered to be potential biowarfare agents,⁹ and consequently novel treatments are being urgently sought.

One key potential source of vaccines is the cell surface polysaccharides. Bacteria maintain a range of polysaccharides as a protective barrier against a variety of external insults, including hydrophobic molecules, divalent cations, and antibiotic macromolecules.¹⁰ Furthermore, in animal infections, polysaccharides interact with both innate and adaptive arms of the immune system.^{11,12} The principal surface polysaccharides of Gram-negative bacteria are lipopolysaccharides (LPS), whose core is usually essential for outer membrane integrity and viability,^{10,12} and capsular polysaccharides (CPS; Figure 1a). Many organisms maintain multiple CPS clusters, with each CPS providing access to a range of environmental niches, hosts, or infection stages.¹³

In *B. pseudomallei*, both the best expressed of the organism's four capsules¹⁴ (CPS type I, CPS-I; Figure 1b) and LPS have been proposed as potential vaccine candidates. CPS-I is one of the best validated virulence factors of the bacterium: immunization with capsule provides protection against future

infection;^{15,16} acapsular mutants are strongly attenuated;^{14,17,18} and the biosynthetic genes are absent in related, less pathogenic species.¹⁴ In addition to providing a general defence against immune challenge, *B. pseudomallei* CPS-I actively inhibits complement C3b, providing a defense against innate immunity.¹⁹ CPS-I is thus an attractive target for vaccine and small molecule inhibitor development.

The CPS-I structure is extremely simple, comprising a linear polymer of 2-*O*-acetyl-6-deoxy- β -D-*manno*-heptopyranose (2O6dDHep; Figure 1b, c).^{14,20} Heptoses are rarely found in higher eukaryotes, but are common in bacterial polysaccharides. In *E. coli* LPS, L-*glycero*- α -D-*manno*-heptopyranose (DHep), forms part of the “inner core”, which is essential for virulence. In other pathogens, heptoses are used in a variety of polysaccharides: 6-deoxy-D-*manno*-heptopyranose (6dDHep) forms part of the LPS O-antigen in both *Yersinia pseudotuberculosis*²¹ and *Campylobacter jejuni*.²² Heptose biosynthesis follows one of two divergent pathways in different bacteria, providing different nucleotide-linked sugars according to the needs of the bacterium.^{10,23} However, the initial stages of biosynthesis of DHep from the pentose phosphate pathway intermediate sedoheptulose-7-phosphate (S7P) are common to both pathways. The wide distribution of these enzymes, their essential nature for virulence, and the lack of equivalents in higher eukaryotes make such enzymes excellent targets for the development of novel, broad spectrum antibiotics.²⁴

A route to the biosynthesis of 2O6dDHep in *B. pseudomallei* has been proposed (Figure 1d; Cuccui *et al.*, manuscript in preparation). A key enzyme in this biosynthetic pathway is GmhA: this is the first enzyme in the pathway, and converts sedoheptulose-7-phosphate into D-*glycero*-D-*manno*-heptopyranose-7-phosphate (M7P). The structures of GmhA from *Escherichia coli*, *Pseudomonas aeruginosa*, *C. jejuni* and *Vibrio cholerae* have recently been determined (Table 1).^{25,26} These structures suggested that GmhA adopts two significant conformations, representing substrate and product binding conformers; and a plausible mechanism was proposed, consistent with all the crystallographic and mutational data and known mechanisms of homologous enzymes. However, whilst the proposed mechanism is a first step towards developing new interventions based on GmhA, some key details were not revealed. It is not clear

which conformation is relevant to catalysis, or whether catalysis occurs during a transition between conformations; and no role is apparent for several residues that are in environments often associated with a catalytic role (Figure 2).²⁷

The *B. pseudomallei* GmhA ortholog retains all of the key conserved amino acids seen in GmhA from these species, whilst having a slightly lower sequence identity than other extant structures (39 % identity to *E. coli* ortholog; others show > 40 % identity). It therefore represents an attractive target for ascertaining the role of these amino acids whose role is unaccounted for, and the two conformations, in GmhA activity. Furthermore, it should highlight whether GmhA might similarly represent a good target for antibiotic development against *B. pseudomallei*. I therefore determined the structure of GmhA from this organism. Surprisingly, the solved structure contains a metal ion in the core of the active site: careful investigation demonstrated that this metal ion is zinc, and suggests that it contributes towards the full activity of the enzyme. The side chains required for binding of this ion are completely conserved across a wide range of bacterial species, suggesting that zinc binding may be common to some other bacterial GmhA enzymes. The solved structure indicates that the “closed” conformation of GmhA is likely to be the catalytically relevant one, and that the previously proposed mechanism for GmhA is unlikely to be correct. The structure suggests an alternative mechanism for GmhA, consistent with all previously published data. This work reveals a novel metal binding site, unprecedented in the SIS enzyme class, and highlights a conformation that should represent an excellent target for the development of effective broad spectrum inhibitors of GmhA.

RESULTS

Structural analysis of His-tagged *B. pseudomallei* GmhA reveals a novel cation binding site

Crystals of 6-His tagged GmhA grew in space group *P1* (Crystal 1; Table 2), and the structure was solved by molecular replacement to GmhA from *C. jejuni* (PDB ID: **1TK9**).²⁵ Eight GmhA protomers were

observed in the asymmetric unit: these appeared to form into two tetramers. This is consistent with the previously observed structures of GmhA from *E. coli*, *P. aeruginosa*, *V. cholerae*, and *C. jejuni*;^{25,26} and with the behavior of the protein in size exclusion chromatography, where it eluted with a Stoke's radius consistent with a tetramer (data not shown). The final model was refined to R and R_{free} values of 18.9 and 24.5 % respectively (Table 2).

Previously solved structures of GmhA from other organisms have suggested that GmhA crystallizes in one of two conformers, corresponding to substrate and product bound forms (Table 1). Comparison of GmhA from *B. pseudomallei* with these structures showed that this enzyme strongly resembles the “closed” structure, observed in two of the extant GmhA structures, including the one known product bound structure (Figure 3a). Moreover, the active site of the *B. pseudomallei* structure shows a clear evidence for an ion. There is an atom present that has electron density consistent with a first row transition metal, which is tetrahedrally co-ordinated (Figure 3b, c). The atom was therefore modelled as zinc, which proved to be an excellent fit to the electron density.

Crystallization of untagged GmhA and determination of metal binding

Since this crystal suggested that GmhA is a metalloenzyme, but did not unambiguously identify the native metal, protein with the His-tag removed was crystallized. The initial crystal form had some crystal contacts involving the TEV cleavage site, so removing this was expected to alter the crystal form.

Untagged GmhA crystallized readily in a range of conditions: all of these crystals were in the space group $P2_12_12_1$, with similar cell dimensions. To determine the identity of the metal in the active site, an extended X-ray absorption fine structure (EXAFS) scan was performed on crystal 2, following collection of a high resolution dataset to 1.62 Å. The emission peaks from this scan were unambiguously characteristic for zinc. Following this, a standard fluorescence scan at the zinc anomalous peak wavelength was performed on crystal 3. This showed a clear anomalous peak (Figure 4a). Data were collected for anomalous phasing for crystal 3, and these data were sufficient to locate the zinc atoms (Figure 4b), determine initial phases, and build an initial model for GmhA. This initial model was then refined against the data from crystal 2.

This model was refined to R and R_{free} values of 14.6 and 17.9 % respectively (Table 2).

Identification of ligands in the crystal structure

The overall architecture of the structure in this crystal form is strikingly similar to that of the tagged protein (not shown). The structures of crystal 1 and crystal 2 show a main chain RMSD of 0.47 Å. Crystal 2 clearly showed a molecule of the product in each active site (Figure 5a). The location of the product molecule strongly resembles that observed for the product bound form of *P. aeruginosa* GmhA, with very similar interactions with side chains (Figure 6).

GmhA mutants confirm that predicted active site residues are required for full activity

As these structural data all suggested that the GmhA from *B. pseudomallei* is a metalloenzyme, a range of mutants were designed to test whether the enzyme has similar activity to the *E. coli* enzyme,²⁶ and to provide evidence in support of zinc binding. Consequently, the mutations H64Q, E68Q, and Q175E mutants were made in residues ligating the zinc (Figure 5b); D98N and T124A mutants in ligand binding residues; and D61A, in a residue drawing protons away from H64 (Figure 5c). With the exception of the D61A, which was not tested, these mutations all had displayed reduced or no activity in *E. coli*. The kinetic parameters for the wild-type enzyme and these mutants (Table 3) demonstrate that GmhA from *B. pseudomallei* shows a similar activity to the *E. coli* enzyme. The E68Q, D98N, T124A and Q175E mutants are all inactive, whilst H64Q and D61A mutants retain some activity. This mirrors the behavior of *E. coli* GmhA, where equivalent mutations showed similar reductions in activity (Table 3). Addition of chelating agents (e.g. EDTA) did not significantly affect the rate of the reaction (not shown).

DISCUSSION

***B. pseudomallei* GmhA is a metalloenzyme**

The two structures of GmhA from *B. pseudomallei* presented here clearly demonstrate that the enzyme is

a metalloenzyme, as both structures show a metal at the heart of the active site (Figures 3 and 5). The most convincing proof for this is that the structure could be solved from the anomalous signal of the four zinc atoms located alone (Figure 4b). It is likely that the observed zinc is natively bound to the active site, and has been retained from the cell: no zinc was added to any of the purification or crystallization buffers used. Zinc is a common enzyme cofactor, with over 10 % of the well characterized enzyme mechanisms in the MACiE database²⁸ including a zinc ion.

Comparison of the *B. pseudomallei* structure with those from other species demonstrates that this enzyme strongly favours the “closed” conformation.²⁶ This conformation is observed in all three tetramers in the two crystal forms: the conformation is likely to be driven by the metal binding, as a similar constellation of the metal binding side chains is observed in the two structures that also adopt this conformation (Figure 7). As the presence of the metal is clearly important for activity in the *B. pseudomallei* enzyme, it is unlikely to be able to access the “open” conformation at all, as this would imply loss of the bound metal. The metal appears to be tightly bound, as zinc binding persisted despite prolonged exposure to imidazole in purification, and several chelating agents failed to reduce the rate of the reaction.

Implications for other orthologs

The structure of GmhA from *B. pseudomallei* strongly suggests that the “closed” conformation, observed also in structures of *P. aeruginosa* and *V. cholerae* GmhA,^{25,26} is most likely to be the catalytically relevant conformation. The structure of the untagged protein presented here shows the product of GmhA, D-glycero- α -D-manno-heptopyranose-7-phosphate: this indicates that the enzyme is capable of turning over the substrate whilst in the “closed” conformation. Furthermore, crystal 2 was harvested five minutes after soaking with sedoheptulose-7-phosphate, consistent with a rapid turnover in the crystal. These observations suggest that the “closed” conformation is sufficient for substrate binding, and for all of the catalytic steps required to form the product in this enzyme. The “closed” conformation maintains a largely solvent exposed active site, so rapid ingress and egress of substrate and product should be possible. Given the high level of sequence identity between GmhA orthologs (Figure 2), this is likely to be the case for

GmhA from all species. The open conformation appears to be suitable for substrate binding,²⁶ and may be used by some orthologs as a means of facilitating more rapid product release. The relative propensity of different orthologs to access this state may reflect differences in evolved turnover rate; *B. pseudomallei* represents an extreme, but perhaps not unique, case where the enzyme is locked into the “closed” conformation. Indeed, all of the previously determined structures used protein that had been purified in the presence of reducing agents, chelating agents, or both, and so it is possible that zinc may have been leached from some of these proteins prior to crystallization. These observations will be extremely important for efforts to rationally design inhibitors against GmhA: the “closed”, catalytic, conformation, of the enzyme offers an excellent target for obtaining a strong inhibitor that will mimic the substrate, product or transition state.

Implications for the mechanism of GmhA

Taylor *et al.*²⁶ proposed a mechanism for the action of GmhA involving an enediol-switch, with E68 and H183 (their *E. coli* equivalents) acting as an acid and base respectively to promote isomerization of GmhA (Figure 8a). Based on the structures presented here, this mechanism appears unlikely for *B. pseudomallei* GmhA, as H183 is buried by the zinc ion in the active site. In contrast, E68 and Q175 bind to the zinc in a manner that strikingly presents these side chains as a potential acid and base (Figure 5d). Mutation of either of these residues to the conjugate amide or acid is sufficient to reduce the enzyme rate at least 80 fold: as these mutations would not be expected to prevent zinc binding, this implies that the zinc ion serves principally to orientate these side chains. I therefore propose an amended mechanism, with E68 and Q175 acting as the base and acid respectively to promote this reaction (Figure 8b). Such a mechanism is consistent with the presence of zinc in the active site of some GmhA orthologs, as zinc ions in other enzymes commonly act to increase the acidity or nucleophilicity of co-ordinating side chains; and the co-ordination of the acid and base in a enediol switch mechanism by zinc has been previously described.²⁹

The presence of a zinc ion in the active site, and the proposed mechanism, are also consistent with mutant

data produced by Taylor *et al.*²⁶ and in this work: the most crucial side chains for enzyme activity appear to be E68, H183 and Q175. Mutation of H183 would be expected to have significant consequences for zinc binding in this structure (as it is the most buried zinc ligand), and in both the non-metal bound structures, where a hydrogen bond between H183 and H64 stabilizes the “closed” conformation (Figure 7). Furthermore, the D61A mutant, which should reduce the affinity of H64 for the zinc, shows a considerably reduced activity, whilst H64Q, which should still be capable of providing a ligand to zinc, shows only a moderate loss in activity (Table 3). These observations explain the roles of the conserved residues of D61, H64, E68, Q175, E176, H178 and H183 across bacterial species (Figure 2): four of these residues are required to co-ordinate the zinc ion (H64, E68, Q175 and H183), whilst D61 and H178 position and remove electrons from H64 (Figure 5c), and E176 performs a similar role with H183. Taylor *et al.*²⁶ proposed that cyclization of the heptose following isomerization would occur outside the enzyme. However, in crystal 2, which was soaked for 5 minutes with S7P, the cyclized product is unambiguously observed in all four active sites. This rapid turnover to the cyclized sugar, and the apparent preference for binding this form in the active site argue that GmhA also catalyzes cyclization. Based on the structures presented here, and on the structure of the complex of *P. aeruginosa* GmhA with M7P, I propose that D98 acts as a proton shuffle, providing both acid and base functions, to promote the cyclization of the sugar (Figure 8c). The crystal structures show that O1 of M7P is located suggestively close to the main chain amide of R72 (Figure 9): it is likely that, in the linear form, O1 will form a hydrogen bond with this amide. This interaction drives the conformation of the sugar such that the α -anomer of M7P will be preferred in the product. This is in agreement with the proposed pathway for the biosynthesis of the *B. pseudomallei* capsule, and suggests that anomer selection is made by GmhA in its catalytic steps, rather than by selection from a racemate by a later enzyme.

Conclusions and future prospects

The data presented in this paper suggest a novel metal binding site in GmhA, a key enzyme for the production of heptoses in bacteria. The discovery of this metal binding site has significant implications

for the mechanism of the enzyme: the revised mechanism explains the striking conservation of many side chains, including those quite distant from the catalytic centre, throughout the bacterial kingdom; and the structure clearly shows two side chains poised to act as a catalytic acid and base for the reaction. This work suggests that efforts to design inhibitors of GmhA as broad-spectrum antibiotics might be focused on the “closed” conformation, and should take account of the possibility of zinc binding across a range of species. The development of such specific inhibitors will provide important details as to the precise mechanism of the enzyme.

MATERIALS AND METHODS

Preparation of GmhA, WcbL and WcbN

Full length sedoheptulose-7-phosphate isomerase (GmhA; Figure 1), D,D-heptose-7-phosphate 1-kinase (WcbL) and D,D-heptose-1,7-bisphosphate 7-phosphatase (WcbN) from *B. pseudomallei* strain K96423 (genomic DNA a gift of R. Titball, University of Exeter) were cloned into pNIC28-Bsa4 (gift of O. Gileadi, SGC Oxford). Constructs were transformed into Rosetta (DE3) cells (Merck), and grown in ZYM-5052 media³⁰ supplemented with 100 µg/mL kanamycin and 20 µg/mL chloramphenicol. Cells were grown at 37°C until OD₆₀₀ was 0.5, and then at 20°C for 16 hours. Harvested cells were resuspended in 20 mM Tris-HCl pH 8.0, 0.5 M NaCl (Buffer A), and lysed using a Soniprep 150 sonicator (MSE). Clarified lysate was purified using a nickel-agarose column (Bioline). Briefly, the loaded protein was washed with Buffer A supplemented with 20 mM imidazole-HCl pH 8.0, and eluted with Buffer A supplemented with 250 mM imidazole-HCl pH 8.0 (WcbL was eluted with Buffer A supplemented with 250 mM imidazole-citrate pH 8.0). GmhA and WcbL were then loaded onto a Superdex 200 16/60 hr column (GE Healthcare), and eluted isocratically with 10 mM Hepes pH 7.0, 0.5 M NaCl. GmhA for enzyme assays was concentrated to 2.25 mg/mL using a Vivaspin centrifugal concentrator (Vivascience), and stored at -20°C with 20 % (v/v) glycerol added: WcbL and WcbN were diluted to 0.9 and 3.4 mg/mL

respectively, with 20 % (v/v) glycerol added, and stored at -20°C. For preparation of tag-free GmhA, approximately 50 mg of nickel affinity purified protein was incubated at 4°C for 36 hours with 700 µg of TEV protease (S219V mutant, purified using methods available at www.mcl1.ncifcrf.gov/waugh_tech/protocols/pur_histev.pdf). The sample was then purified by size-exclusion chromatography as described above, and extensively loaded over nickel-agarose to remove residual undigested GmhA and TEV.

GmhA Mutagenesis

Site directed mutants of GmhA were generated using the QuikChange II XL mutagenesis kit (Agilent). Mutations were verified by DNA sequence analysis (MWG, London).

Structure Determination of GmhA

All crystals were grown using the microbatch method, and were prepared using an Oryx6 crystallization robot (Douglas Instruments). Tagged *B. pseudomallei* GmhA at 10 mg/mL was mixed with an equal volume of crystallization solutions (see Table 4 for details of crystallization and soaking conditions), and grown at 20°C. Prior to flash-freezing in liquid nitrogen, crystals were soaked for 30-60 seconds in a cryoprotectant solution. Tag-removed GmhA at 10 mg/mL was mixed with an equal volume of crystallization solutions, and crystals were grown at 4°C. Prior to flash freezing, crystals were soaked in cryoprotectant solutions containing proposed substrates (Table 4) for 5 (crystal 2) or 70 (crystal 3) minutes. All substrate was converted to product in both crystals. All X-ray diffraction data sets were collected at 100 K at beamline I03 of the Diamond synchrotron. Single wavelength datasets were collected for all crystals as detailed in Table 2. To determine the identity of the binding metal, an EXAFS scan was performed using station I03. An energy of 20 keV was used for the EXAFS scan. Data for crystal 1 were processed using MOSFLM version 7.0.2³¹ and SCALA;³² data for crystals 2 and 3 were processed using iMOSFLM version 1.0.3 and SCALA. Initial phases for crystal 1 were determined by

molecular replacement using the structure of one monomer from *C. jejuni* GmhA (PDB ID: **1TK9**) as a model. Molecular replacement was performed using PHASER.³³ Eight protomers were observed in the asymmetric unit. Single wavelength anomalous data were collected for crystal 3: using SHELX C/D/E³⁴ and the HKL2MAP pipeline,³⁵ precisely four zinc atom sites were located. Phasing statistics were calculated using SHARP.³⁶ An initial structure was determined using ARP/wARP version 7.0.³⁷ Model building and refinement of the structures was performed using Coot version 0.5.2³⁸ and PHENIX version 1.5-2.³⁹ TLS groups were generated using TLSMD.⁴⁰ Structures were validated using PHENIX, Coot and MOLPROBITY.⁴¹ Structural images were prepared using the PyMOL molecular graphics system (DeLano Scientific).

Sedoheptulose 7-Phosphate (S7P) Synthesis

S7P for enzyme assays was purchased from Carbosynth. For crystallization, S7P was synthesized using an irreversible activity of transketolase. *E. coli* transketolase was the gift of J. Littlechild, Exeter. Purified transketolase was incubated at 0.15 mg/mL in 50 mM Hepes pH 7.5, 1 mM EDTA pH 8.0, 6 mM 2-mercaptoethanol, 2 mM thiamine pyrophosphate, 7.5 mM CaCl₂ for 20 min at room temperature. One volume each of 40 mM ribose-5-phosphate and 80 mM β-hydroxypruvate were added to four volumes of this mixture, and incubated at 37°C overnight. The mixture was then incubated at 70°C for 10 min to eliminate the enzyme activity. Greater than 95 % conversion of ribose-5-phosphate to sedoheptulose 7-phosphate was confirmed by HPLC.

GmhA Enzyme Activity Assay

GmhA activity was monitored by coupling product formation to WcbL and WcbN and monitoring P_i release, in an analogous manner to the methods of DeLeon *et al.*⁴² Briefly, a reaction mixture was formed consisting of 20 mM HEPES, pH 8.0, 10 mM MgCl₂, 10 mM KCl, 200 μM ATP, 9 μg/mL WcbL, 34 μg/mL WcbN, and 36-90 μg/mL GmhA (final concentrations) in a total volume of 50 μL. Reactions were

initiated with 50 μL of S7P for final concentrations ranging from 0 to 1.5 mM. Samples were incubated at 37°C for 10 min, and the reaction was terminated by heating to 70°C for 5 minutes. The phosphate concentration was determined by the addition of 25 μL of 5.6 % (w/v) ammonium molybdate, 0.006 % (w/v) malachite green in 6 M HCl to 90 μL of reaction product. The absorbance at 660 nm was then determined using a Spectra MR plate reader (Dynex). The rate of GmhA was determined by comparison to a phosphate standard curve ranging from 0-200 μM phosphate, and to negative controls containing no GmhA. All reactions were performed in triplicate, and a total of eight substrate concentrations were tested for each sample. The kinetic parameters of each GmhA preparation were determined by fitting the data to the Michaelis-Menten equation, using SPSS version 16 (SPSS Inc.). Activity figures are given per GmhA holoenzyme. A detection limit of 0.006 s^{-1} was determined for this assay.

Accession numbers

The co-ordinates and structure factors for the structures described in this paper have been submitted to the Protein Data Bank with accession numbers [2X3Y](#) and [2XBL](#) for crystals 1 and 2 respectively. Experimental data for crystal 3 have been submitted with the structure for crystal 2.

Acknowledgements

The author thanks the staff at Diamond beamlines I02 and I03, especially James Sandy, for their assistance in collecting data; Misha Isupov, Paul James, Chris Sayer and Vahid Saneii (University of Exeter) for assistance in data collection; Pavel Afonine (Lawrence Berkeley National Laboratory, CA) for advice on refinement; Benjamin Svendsen (University of Exeter) for assistance in preparing mutants; and Clive Butler (University of Exeter) for critical reading of the manuscript. This work was funded by the University of Exeter.

REFERENCES

1. Hsueh, P.-R., Teng, L.-J., Lee, L.-N., Yu, C.-J., Yang, P.-C., Ho, S.-W. & Luh, K.-T. (2001). Melioidosis: an emerging infection in Taiwan? *Emerg. Infect. Dis.* **7**, 428-433.
2. Stone, R. (2007). Racing to defuse a bacterial time bomb. *Science* **317**, 1021-1023.
3. Wiersinga, W. J., van der Poll, T., White, N. J., Day, N. P. & Peacock, S. J. (2006). Melioidosis: insights into the pathogenicity of *Burkholderia pseudomallei*. *Nat. Rev. Microbiol.* **4**, 272-282.
4. White, N. (2003). Melioidosis. *Lancet* **361**, 1715-1722.
5. John, T., Jesudason, M., Lalitha, M., Ganesh, A., Mohandas, V., Cherian, T., Mathai, D. & Chandy, M. (1996). Melioidosis in India: the tip of the iceberg. *Ind. J. Med. Res.* **103**, 62-65.
6. Lazar Adler, N. R., Govan, B., Cullinane, M., Harper, M., Adler, B. & Boyce, J. D. (2009). The molecular and cellular basis of pathogenesis in melioidosis: how does *Burkholderia pseudomallei* cause disease? *FEMS Microbiol. Rev.* **33**, 1079-1099.
7. Jenney, A. W., Lum, G., Fisher, D. A. & Currie, B. J. (2001). Antibiotic susceptibility of *Burkholderia pseudomallei* from tropical northern Australia and implications for therapy of melioidosis. *Int. J. Antimicrob. Agents* **17**, 109-113.
8. Ngauy, V., Lemeshev, Y., Sadkowski, L. & Crawford, G. (2005). Cutaneous melioidosis in a man who was taken as a prisoner of war by the Japanese during World War II. *J. Clin. Microbiol.* **43**, 970-972.
9. Rotz, L. D., Khan, A. S., Lillibridge, S. R., Ostroff, S. M. & Hughes, J. M. (2002). Public health assessment of potential biological terror agents. *Emerg. Infect. Dis.* **8**, 225-230.
10. Valvano, M. A., Messner, P. & Kosma, P. (2002). Novel pathways for biosynthesis of nucleotide-activated *glycero-manno*-heptose precursors of bacterial glycoproteins and cell surface polysaccharides. *Microbiology* **148**, 1979-1989.
11. Caroff, M. & Karibian, D. (2003). Structure of bacterial lipopolysaccharides. *Carbohydr. Res.* **338**, 2431-2447.
12. Raetz, C., Reynolds, C., Trent, M. & Bishop, R. (2007). Lipid A modification systems in Gram-negative bacteria. *Ann. Rev. Biochem.* **76**, 295-329.
13. Whitfield, C. (2006). Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Ann. Rev. Biochem.* **75**, 39-68.
14. Reckseidler, S. L., DeShazer, D., Sokol, P. A. & Woods, D. E. (2001). Detection of bacterial virulence genes by subtractive hybridization: identification of capsular polysaccharide of *Burkholderia pseudomallei* as a major virulence determinant. *Infect. Immun.* **69**, 34-44.
15. Atkins, T., Prior, R., Mack, K., Russell, P., Nelson, M., Prior, J., Ellis, J., Oyston, P. C., Dougan, G. & Titball, R. W. (2002). Characterisation of an acapsular mutant of *Burkholderia pseudomallei* identified by signature tagged mutagenesis. *J. Med. Microbiol.* **51**, 539-547.
16. Nelson, M., Prior, J., Lever, M. S., Jones, H. E., Atkins, T. P. & Titball, R. W. (2004). Evaluation of lipopolysaccharide and capsular polysaccharide as subunit vaccines against experimental melioidosis. *J. Med. Microbiol.* **53**, 1177-1182.
17. Cuccui, J., Easton, A., Chu, K., Bancroft, G., Oyston, P., Titball, R. & Wren, B. (2007). Development of signature-tagged mutagenesis in *Burkholderia pseudomallei* to identify genes important in survival and pathogenesis. *Infect. Immun.* **75**, 1186-1195.
18. Warawa, J., Long, D., Rosenke, R., Gardner, D. & Gherardini, C. (2009). A role for the *Burkholderia pseudomallei* capsular polysaccharide encoded by the *wcb* operon in acute

- disseminated melioidosis. *Infect. Immun.* **77**, 5252-5261.
19. Reckseidler-Zenteno, S. L., DeVinney, R. & Woods, D. E. (2005). The capsular polysaccharide of *Burkholderia pseudomallei* contributes to survival in serum by reducing complement factor C3b deposition. *Infect. Immun.* **73**, 1106-1115.
 20. DeShazer, D., Waag, D. M., Fritz, D. L. & Woods, D. E. (2001). Identification of a *Burkholderia mallei* polysaccharide gene cluster by subtractive hybridization and demonstration that the encoded capsule is an essential virulence determinant. *Microb. Pathogen.* **30**, 253-269.
 21. Pacinelli, E., Wang, L. & Reeves, P. R. (2002). Relationship of *Yersinia pseudotuberculosis* O Antigens IA, IIA, and IVB: the IIA gene cluster was derived from that of IVB. *Infect. Immun.* **70**, 3271-3276.
 22. St. Michael, F., Szymanski, C., Li, J., Chan, K., Khieu, N., Larocque, S., Wakarchuk, W., Brisson, J. & Monteiro, M. (2002). The structures of the lipooligosaccharide and capsule polysaccharide of *Campylobacter jejuni* genome sequenced strain NCTC 11168. *Eur. J. Biochem.* **269**, 5119-5136.
 23. Butty, F., Aucoin, M., Morrison, L., Ho, N., Shaw, G. & Creuzenet, C. (2009). Elucidating the formation of 6-deoxyheptose: biochemical characterisation of the GDP-D-glycero-D-manno-heptose C6 dehydratase, DmhA, and its associated C4 reductase, DmhB. *Biochemistry* **48**, 7764-7775.
 24. Taylor, P. L. & Wright, G. D. (2008). Novel approaches to discovery of antibacterial agents. *Anim. Health Res. Rev.* **9**, 237-246.
 25. Seetharaman, J., Kanagalaghatta, R. R., Solorzano, V., Kniewel, R., Lima, C. D., Bonanno, J. B., Burley, S. K. & Swaminathan, S. (2006). Crystal structures of two putative phosphoheptose isomerases. *Proteins: Struct., Funct. Bioinf.* **63**, 1092-1096.
 26. Taylor, P. L., Blakely, K. M., De Leon, G. P., Walker, J. R., McArthur, F., Evdokimova, E., Zhang, K., Valvano, M. A., Wright, G. D. & Junop, M. S. (2008). Structure and function of GmhA (sedoheptulose 7-phosphate isomerase): a critical enzyme for lipopolysaccharide biosynthesis and a target for antibiotic adjuvants. *J. Biol. Chem.* **283**, 2835-2845.
 27. Overington, J., Johnson, M. S., Sali, A. & Blundell, T. (1990). Tertiary structural constraints on protein evolutionary diversity: templates, key residues and structure prediction. *Proc. Roy. Soc. B* **241**, 132-145.
 28. Holliday, G. L., Almonacid, D. E., Bartlett, G. J., O'Boyle, N. M., Torrance, J. W., Murray-Rust, P., Mitchell, J. B. O. & Thornton, J. M. (2007). MACiE (Mechanism, Annotation and Classification in Enzymes): novel tools for searching catalytic mechanisms. *Nucleic Acids Res.* **35**, D515-D520.
 29. Cameron, A. D., Ridderström, M., Olin, B., Kavarana, M. J., Creighton, D. J. & Mannervik, B. (1999). Reaction mechanism of glyoxalase I explored by an X-ray crystallographic analysis of the human enzyme in complex with a transition state analogue. *Biochemistry* **38**, 13480-13490.
 30. Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression Purif.* **41**, 207-234.
 31. Leslie, A. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 ESF-EAMCB Newsl. Protein Crystallogr.* **26**.
 32. Evans, P. R. (2005). Scaling and assessment of data quality. *Acta Crystallogr. D* **D62**, 72-82.

33. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674.
34. Sheldrick, G. M. (2008). A short history of SHELX. *Acta Crystallogr. A* **A64**, 112-122.
35. Pape, T. & Schneider, T. R. (2004). HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843-844.
36. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2008). Automated structure solution with autoSHARP. *Meth. Mol. Biol.* **364**, 215-230.
37. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171-1179.
38. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D* **D66**, 486-501.
39. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **D58**, 1948-1954.
40. Painter, J. & Merritt, E. A. (2006). *TLSMD* web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.* **39**, 109-111.
41. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall III, B., Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375-W383.
42. De Leon, G. P., Elowe, N. H., Koteva, K. P., Valvano, M. A. & Wright, G. D. (2006). An in vitro screen of bacterial lipopolysaccharide biosynthetic enzymes identifies an inhibitor of ADP-heptose biosynthesis. *Chem. Biol.* **13**, 437-441.
43. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.
44. Mizuguchi, K., Deane, C. M., Blundell, T., Johnson, M. S. & Overington, J. P. (1998). JOY: protein sequence-structure representation and analysis. *Bioinformatics* **14**, 617-623.

Figure 1: Polysaccharides in *Burkholderia pseudomallei*. **(a)** Schematic of bacterial polysaccharides. These be broadly categorised as the stress bearing, periplasmic, peptidoglycan; the lipopolysaccharides (LPS), which constitute the outer leaflet of the outer membrane; and the capsular polysaccharides (CPS), which form a protective barrier around the cell (after ¹³). **(b)** Structure of the repeating unit of *B. pseudomallei* CPS-I, -3)-2-*O*-acetyl-6-deoxy-D-glycero- β -D-manno-heptopyranose-(1-; the capsule consists of long linear polymers of this linked to an unknown lipid. **(c)** Structure of D-glycero-D-manno-heptopyranose (DHep), a sugar used by many bacteria in LPS or CPS. **(d)** Proposed pathway for the biosynthesis of the *B. pseudomallei* CPS-I precursor, GDP-2-*O*-acetyl-6-deoxy- α -D-manno-heptopyranose (Cuccui *et al.*, in preparation). The sites acted on by the next enzyme in the pathway are indicated at each step by an arrowhead.

Figure 2: Sequence comparison of known GmhA sequences and structures. Sequences chosen to include all crystallized proteins, and to represent a wide range of bacterial taxa, were aligned using CLUSTALX version 2.0.⁴³ Image prepared using JOY,⁴⁴ with sequences of known structures annotated to indicate their environment (see key, bottom): residues shown with more prominent annotation are more likely to be in interesting environments, and conserved side chains of a type often associated with catalysis in such environments are indicated as explained below. Absolutely conserved residues are indicated with an asterisk below the residue. Side chains previously implicated in catalysis are indicated with a black arrow. Other zinc binding side chains proposed by this work are indicated with a gray arrow. Highly conserved side chains with unusual environments, located near to the active site, with no proposed role in catalysis, are indicated with a black arrowhead. Other residues selected for mutation are indicated with a gray arrowhead. Highly conserved residues with suggestive environments, but a clear role in providing structural stability are indicated with dashed black arrows. Sequences taken from the following species: *Burkholderia pseudomallei* K96243 (2XBL), *Vibrio cholerae* (1X94), *Pseudomonas aeruginosa* (1X92), *Campylobacter jejuni* (1TK9) *Escherichia coli* (2I22), *Alphaproteobacterium* HTCC2255, *Aquifex aeolicans* VF5, *Methanococcus aeolicus* Nankai-3, *Bacteriodes* sp. 1_1_6, and *Desulfovibrio desulfuricans* subsp. *desulfuricans* strain G20.

Figure 3: Structure of 6*His-tagged GmhA from *B. pseudomallei*. **(a)** *B. pseudomallei* GmhA adopts a “closed” conformation: GmhA from *B. pseudomallei* (crystal 1, left, orange) is compared with the “open” conformation structure of *E. coli* GmhA (PDB ID: 2I22, middle, cyan) and the “closed” conformation of *P. aeruginosa* GmhA (PDB ID: 1X92, right, magenta²⁶). All proteins are shown in cartoon form. Note the disordered helices (arrowhead) and change in the angle of helices connecting the two dimers (arrow) at the centre of the “open” conformation, leading to a translocation of the lower dimer by 6 Å. **(b)** Crystal 1 contains a zinc ion. Zinc is shown as a magenta sphere. The three protein chains that form the active site are colored orange, pink and cyan respectively. Zinc co-ordinating residues are shown as sticks, with the nearest water as a red sphere. In green is a simulated annealing omit map ($F_o - F_c$) with zinc and co-ordinating side chains removed, contoured at 3σ . Difference density at the zinc position is still visible to 9.5σ **(c)** This structure shows each ion clearly tetrahedrally co-ordinated with ligands approximately equally spaced (pink dashed lines), with the closest waters too distant to form an addition ligand (blue dashed line). Distances are shown in Å. Structural images here and in other figures were prepared using the PyMOL molecular graphics system (DeLano Scientific).

Figure 4: GmhA from *B. pseudomallei* contains a zinc ion in the active site. **(a)** Fluorescence scan of crystal 3 at the zinc peak shows a clear absorption edge. **(b)** SAD data for crystal 3 allows zinc ions to be located. Shown is an anomalous difference map, contoured at 10σ , for the final refined structure of crystal 2 refined once in PHENIX against the data from crystal 3. The upper and lower panels show two representative zinc ions are clearly located.

Figure 5: Ligand binding to *B. pseudomallei* GmhA. **(a)** Crystal 2 shows a molecule of M7P in each

active site. Protein is shown in cartoon form, with side chains within 3.5 Å of M7P shown as lines. The three protein chains that constitute the active site are colored yellow, pink and cyan. Zinc ion is shown as a purple sphere, and M7P as white sticks. In green is a simulated annealing omit map ($F_o - F_c$) with M7P removed, contoured at 3σ . All ligand atoms are clearly resolved in the density. **(b, c)** Rationale for selection of mutants. **(b)** The active site of crystal 2 is shown, with side chains selected for mutation shown as sticks. Coloring is as above. **(c)** Residue D61 helps co-ordinate a water filled cavity (waters shown as red spheres) between two active sites, and acts through water to draw a proton from H64. **(d)** E68 and Q175 are poised to act as proton donors and acceptors to drive catalysis. Active site of crystal 2 is shown, with zinc binding residues and M7P shown as sticks, colored as above. This conformation presents one oxygen of E68 and the nitrogen of Q175 (arrowheads) in ideal locations for performing an acid-base catalysis of the isomerization of S7P.

Figure 6: Active site interactions with M7P. The interactions of the active sites of *B. pseudomallei* **(a)** and *P. aeruginosa* **(b)** GmhA in complex with M7P were compared using MOE (Chemical Computing Group). E68, D98 and Q175 are indicated on each image by black, blue and red arrows, respectively. Other side chains mutated in this work are indicated by a green arrow.

Figure 7: Comparison of the active sites of “closed” conformation GmhA. The histidine residues H64 and H183 of *B. pseudomallei* GmhA crystal 3 (left, yellow/pink/cyan) show a subtly different conformation to their equivalents in *P. aeruginosa* (middle, magenta) or *C. jejuni* (right, slate). This presents the *B. pseudomallei* side chains to form an interaction with the zinc ion (purple sphere), and extends the nitrogen-nitrogen distance to 3.4 Å, precluding a hydrogen bond between H64 and H183. *B. pseudomallei* GmhA H64, E68, Q175, H183 and their equivalents in other structures are shown as sticks, with main chain shown as cartoon.

Figure 8: Proposed mechanism for GmhA. **(a)** The mechanism previously proposed for GmhA²⁶ implicated H180 (numbering is given for *E. coli* ortholog for residues marked *; equivalent *B. pseudomallei* residues are E68 and H183 respectively) as the catalytic acid for the isomerization of S7P. **(b)** Mechanism proposed by this work shows Q175 as a catalytic acid. In the middle step, Q175 forms an imine, stabilized by the presence of zinc. **(c)** Proposed mechanism for the cyclization to form M7P. D98 is proposed to act as both acid and base, with a protonated form stabilized by a hydrogen bond to S71. The preference for the α -anomer appears to be driven by a hydrogen bond between the carbonyl oxygen and the main chain of R72 in the linear form (see Figure 9).

Figure 9: Proposed mechanism for sugar cyclization to form M7P. A representative active site of crystal 2 is shown. Key side chains and ligand are shown as sticks, with the protein chains colored yellow, pink and cyan, and the ligand white. D98 is poised to act as a proton donor/acceptor for M7P O1 and O5 respectively (dashed orange lines), with S71 making a hydrogen bond (dashed pink line) to stabilize the protonated form of D98. In the linear form of the sugar, O1 would make a hydrogen bond (dashed blue line) with the main chain amide of R72, driving the cyclized sugar to the α -anomer.