Mycobacterial OtsA structures unveil substrate preference mechanism and allosteric regulation by 2-oxoglutarate and 2-phosphoglycerate

Vítor Mendes[1*], Marta Acebrón-García-de-Eulate[1], Nupur Verma[1], Michal Blaszczyk[1], Márcio V. B. Dias[2,3] and Tom L. Blundell[1*]

1 - Department of Biochemistry, University of Cambridge, Cambridge CB21GA, UK.

2 - Department of Microbiology, Institute of Biomedical Science, University of São Paulo, São Paulo, Brazil.

3 – Department of Chemistry, University of Warwick, Coventry CV4 7EQ, UK.

*Corresponding authors

Tom L. Blundell

Address: Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, CB2 1GA Cambridge, UK.

Phone: +441223333628

Email: tom@cryst.bioc.cam.ac.uk

Vitor Mendes

Address: Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, CB2 1GA Cambridge, UK.

Phone: +441223766028

Email: vgm23@cam.ac.uk

Abstract

Trehalose is an essential disaccharide for mycobacteria and a key constituent of several cell wall glycolipids with fundamental roles in pathogenesis. Mycobacteria possess two pathways for trehalose biosynthesis. However, only the OtsAB pathway was found to be essential in *M. tuberculosis*, with marked growth and virulence defects of OtsA mutants and strict essentiality of OtsB2. Herein, we report the first mycobacterial OtsA structures from *M. thermoresistibile* in both apo and ligand-bound forms. Structural information reveals three key residues in the mechanism of substrate preference that were further confirmed by site-directed mutagenesis. Additionally, we identify 2-oxoglutarate and 2-phosphoglycerate as allosteric regulators of OtsA. The structural analysis in this work strongly contributed to define the mechanisms for feedback inhibition, show different conformational states of the enzyme and map a new allosteric site.

Importance

Mycobacterial infections are a significant source of mortality worldwide, causing millions deaths annually. Trehalose is a multipurpose disaccharide that plays a fundamental structural role in these organisms as a component of mycolic acids, a molecular hallmark of the cell envelope of mycobacteria. Here we describe the first mycobacterial OtsA structures. We show mechanisms of substrate preference and show that OtsA is regulated allosterically by 2-oxoglutarate and 2-phosphoglycerate at an interfacial site. These results identify a new allosteric site and provide insight on the regulation of trehalose synthesis through the OtsAB pathway in mycobacteria.

Introduction

Trehalose is a non-reducing disaccharide, formed by α-(1-1) linked glucoses, with a wide distribution in nature and present in all three domains of life (1). This remarkably widespread sugar performs multiple roles in a wide variety of organisms and it can also fulfil different roles within the same organism. Trehalose has been considered a compatible solute, conferring protection to proteins, DNA, membranes and whole cells from thermal shock, osmotic shock, freezing, ionizing radiation, oxidative stress and desiccation (1-9). This disaccharide, which can further function as a carbon and energy reserve molecule (1, 10-12), was also recently related to the pathogenicity of *Pseudomonas aeruginosa* in plants (13). It additionally plays fundamental signaling roles in plants, in the phosphorylated form (trehalose-6-phosphate), where it regulates sucrose metabolism and flowering (1, 14-16), and in yeast where it regulates gluconeogenesis and glycolysis (17, 18). Trehalose was further shown to be an autophagy inducer, both in plants and in mammals, with potential biotechnological and clinical implications (19, 20). In mycobacteria, trehalose is also an essential component of mycolic acids and other cell wall glycolipids, which are major protagonists in *Mycobacterium tuberculosis* pathogenesis (1), the causative agent of the widespread infectious disease tuberculosis. In these organisms, trehalose was further identified as a key signaling molecule of cell-envelope stress, playing a role as an activator of the iniBAC operon, which is induced when mycobacteria are exposed to the first line drug isoniazid (21).

All mycobacterial species, with a few exceptions, possess two pathways to synthesize trehalose, the OtsAB and the TreYZ pathway (1). The OtsAB pathway is the most widely distributed trehalose biosynthesis pathway, present in bacteria, archaea and eukaryotes (22, 23). In this pathway, which is conserved and essential in mycobacteria, trehalose is

synthesized in a two-step process involving OtsA and OtsB2 enzymes. *M. tuberculosis* mutants showed that the OtsAB pathway was not only the dominant route for trehalose biosynthesis in this pathogen, but also required for growth both *in vitro* and in a mouse infection model, with marked growth and virulence defects of OtsA mutants (24) and strict essentiality of OtsB2 due to the toxic effect of trehalose-6-phosphate (T6P) accumulation (24, 25).

OtsA, a glycosyltransferase that belongs to the GT20 family of the CAZY classification (www.cazy.org), uses the α anomer of glucose-6-phosphate (G6P) as acceptor and nucleoside diphosphate glucose (NDP-glucose) as the donor to synthesize T6P, with net retention of the anomeric configuration of the donor substrate (26, 27). Binding of the substrates occurs sequentially in a "bi-bi mechanism" with NDP-glucose binding first and G6P second (26, 28). Interestingly, OtsAs from different organisms show different donor preferences as shown by kinetic studies and X-ray crystal structures, but the reasons behind these preferences are poorly understood (27, 29-33). Recently, different roles were identified for OtsA beyond its enzymatic activity. OtsA was reported to act as an osmotic stress sensor and morphogenetic protein that can regulate the switch to myceloid growth in *Arthrobacter sp.* strain A3, a pleomorphic soil dwelling actinobacteria (34).

In this work we purified, crystallized and solved the structure of *Mycobacterium thermoresistibile* OtsA (MtrOtsA). To gain further insight into substrate preference of mycobacterial OtsAs we have obtained structures with substrates that define the mechanisms of ADP-glucose preference. Structure-guided point mutations of key residues in the active site were performed and characterization of the mutants showed that three mutations are enough to change the donor substrate preference to UDP-glucose. Further structures with OtsA product and pathway product were obtained showing how this enzyme is feedback inhibited by trehalose. Importantly with these structures, we have

also identified a new allosteric site and allosteric regulators of this enzyme that link glycolytic and TCA cycle metabolites to the regulation of trehalose synthesis.

Results

Overall structure

The structure of MtrOtsA in its apo form was solved by molecular replacement, using the *Escherichia coli* OtsA structure (PDB code: 1UQU) as the search model (Figure S1). Data collection and refinement statistics are summarized in (Table S1). MtrOtsA is composed of two Rossmann-fold domains with a deep catalytic site at their interface, in an arrangement typical of GT-B glycosyltransferases as described previously for other organisms (27, 29, 30). The apo protein crystallized in the $I4_122$ space group, with one protomer per asymmetric unit, and diffracted to ~1.8 Å resolution. The N-terminal domain is formed by a core of 7 parallel β-strands flanked on both sides by an antiparallel β-strand and surrounded by 8 α-helixes, one of which is composed of the final C-terminal residues (Figure 1A). The C-terminal domain contains 6 parallel β-strands associated with 9 α-helixes with the last one undergoing a kink and extending to the N-terminal domain, characteristic of the GT-B fold glycosyltransferases (Figure 1A). Analysis of the B-factor distribution shows that the N-terminal domain has the highest values for atomic temperature factors, suggesting that this domain is more dynamic, which is consistent with the large movements of α-1 during catalytic activity.

MtrOtsA forms a tetramer in solution, as previously observed for *M. tuberculosis* OtsA. This tetrameric form is also observed in all the crystal structures reported herein (Figure 1B). The amino acids involved in the tetramer interfaces are not conserved beyond mycobacteria and related species (Figure S2), suggesting that OtsA might have a different

molecular assembly in different species. Indeed in *E. coli* both tetrameric and dimeric forms were reported (27) while in *Streptomyces venezuelae* and *Aspergillus fumigatus* only the dimeric form was observed (29, 30). However, in mycobacteria and closely related organisms the tetramer interfaces are highly conserved, suggesting a tetrameric assembly of OtsA in all of these organisms (Figure S2).

Catalytic site

OtsA, is a glycosyltransferase that uses the α anomer of G6P as acceptor and NDP-glucose as the donor to synthesize T6P. The MtrOtsA catalytic site is located between the two Rossmann-fold domains in a large and deep cavity. Structures with donor substrates were determined by soaking the apo form crystals with ADP-glucose and GDP-glucose (Figure 2). The donor substrates interact primarily with the C-terminal domain through the side chains of the highly conserved Arg286, Lys291, Asp385, Glu393 and with the side chain of Arg365 (Figure 2), which is conserved in all mycobacteria and closely related species but less conserved outside this group (Figure S2). Backbone interactions with the absolutely conserved residues Gly386, Met387, Asn388 and Leu389 amine groups are also observed. Similar interactions were observed before for *E. coli* OtsA (27, 32). While the adenine moiety interacts with Val363 amino group, Leu319 carbonyl and a coordinated water (Figure 2A), the guanine moiety forms interactions with Val 363 only (Figure 2B). N-terminal domain interactions with the donor substrate are only observed for His168 when the active site is in an open conformation (Figure 2). To obtain a structure with the acceptor substrate, we co-crystallized OtsA in the presence of 5 mM ADP and G6P. As G6P binds to the protein in the presence of a donor substrate or its nucleotide, the active site adopts to a closed conformation (Figure 3A). The OtsA:ADP:G6P ternary complex crystallized in the $P6_222$ space group, with one

151 protomer per asymmetric unit, and the crystals diffracted up to ~1.7 Å resolution. In this
152 closed conformation new interactions with the donor substrate are formed with Gly39
153 peptide-NH function (Figure 3). Due to the absolute conserved nature of Gly39 among
154 all functional OtsAs and related enzymes (27) and the contact it forms with the glucose-
155 bound phosphate, this residue is highly likely to be mechanistically involved in the
156 catalytic activity (27).
157 The acceptor interactions recapitulate what was observed before for *E. coli* (27) with the
158 substrate interacting with the highly conserved residues Arg18, Tyr90, Asp144 and
159 Gln146 of the N-terminal domain and with only a single residue of the C-terminal domain,
160 Arg324 (Figure 3B). The residues involved in acceptor substrate interaction are highly
161 conserved across both bacteria and yeast (Figure S2). As described for *E. coli* OtsA (27),
162 the OtsA:ADP:G6P ternary structure shows the catalytic site in a closed conformation
163 that substantially differs from the apo form, with α-helix 1 and the Arg35-Gly39 loop
164 region moving up to ~12 Å (Figure 3A).
165
166 **Properties of *M. thermoresistibile* OtsA.**
167 *Mtr*OtsA uses ADP-glucose, UDP-glucose and GDP-glucose as glucose donors with a
168 preference for ADP-glucose (Table 1) and G6P as the only acceptor, in accordance to
169 what was reported before for *M. tuberculosis* OtsA (33, 35). The kinetic parameters for
170 *Mtr*OtsA were obtained and are reported in (Table 1) and (Figure S3) with K_m^{app} values
171 for preferred donor substrate (ADP-Glucose) of 0.25 ± 0.02 mM and for the acceptor
172 (G6P) of 3.3 ± 0.1 mM. The K_m^{app} for GDP-glucose of 0.29 ± 0.02 mM was in the same
173 range as the one obtained for ADP-glucose, however the turnover was 5 fold lower (Table
174 1). For UDP-glucose, the enzyme showed a ~7 fold higher K_m^{app} of 1.7 ± 0.1 mM than
175 the one obtained for ADP-glucose (Table 1).

The preference of ADP-glucose over the other glucose donors was further confirmed by isothermal titration calorimetry (ITC). The binding affinity of substrate donors could only be determined for ADP-glucose in the tested conditions with an observable K_d of 27.17 ± 2.66 μM (Figure S4). Other glucose donors analysed (GDP-glucose and UDP-glucose) had no observable heat of binding and the same was observed for G6P. The lack of heat of binding for GDP- and UDP-glucose is most likely due to their reduced affinity. For G6P case, it reflects a necessity of previous binding of the donor substrate.

Donor substrate preference is largely mediated by three residues

OtsAs from different organisms have been shown to have different donor substrate preferences, while being capable of using several nucleotide donors (30, 33, 36, 37), which is reflected in lower conservation of the donor substrate interacting residues (Figure S2). Preference for ADP-glucose as the donor substrate in *Mtr*OtsA is conferred by interactions with the deeply buried adenine moiety (Figure 4A). The carbonyl groups of Leu319 and Arg361 interact with the primary amine, and the amide group of Val363 interacts with N1 of the adenine moiety (Figure 4A). A highly coordinated water interacts with the primary amine of the adenine moiety and the carbonyl groups of Leu359 and Arg361. The guanine moiety of GDP-glucose does not occupy the same deeply buried pocket as the adenine because its primary amine group would sterically clash with Val363 carbonyl group (Figure 4B). It thus binds to OtsA more weakly than ADP-glucose explaining the substrate preference and consequently the lack of observable heat of binding in the tested ITC conditions. A binary complex structure with OtsA and UDP-glucose was also obtained but the electron density was only observed for glucose and the two phosphates (not shown) indicating a reduced preference for this substrate as confirmed by enzymatic and biophysical data. Comparing the OtsA structures of *M.*

thermoresistibile and *E. coli*, we hypothesised that the ADP preference was mediated by the substitution of an isoleucine (Ile295 in *E. coli*) for a leucine (Leu319 in *M. thermoresistibile*). The presence of the leucine in *M. thermoresistibile* allows the primary amine of the adenine moiety to occupy a buried position, interacting with the carbonyl group of Leu319 (Figure 4C) that no other nucleotide activated donor can occupy. However, given the differences between the two enzymes it is likely that other residues could also play a role. To improve the selection of residues to mutate we employed a computational approach using a mCSM-lig (38), a software developed by our group that predicts the effects of mutations on binding affinity. The software predicted several that residues within 4.5 Å of the adenine moiety have a destabilizing effect on the interactions between *M. thermoresisitibile* OtsA and ADP-glucose if mutated to the *E. coli* OtsA equivalent (Table S2). Combining this information with structural analysis three mutations predicted to be destabilizing were selected. These mutations are Val363Phe, Leu319Ile and Glu367Leu, the latter being a long distance mutation that we hypothesised would have a strong effect on ligand binding.

To confirm our hypothesis, we generated several *Mtr*OtsA mutants (Leu319Ile, Val363Phe, Leu319Ile-Glu367Leu and a triple mutant) and obtained kinetic data for these mutants (Table 1). As predicted the mutant Leu319Ile showed a ~3 fold increase in the K_m^{app} for ADP-glucose whereas, K_m^{app} was similar GDP-glucose and UDP-glucose when compared to the wild type. Even though GDP-glucose now showed the best K_m^{app} among the three different donors, the catalytic efficiency and turnover rate was still higher for ADP-glucose (Table 1). A complete reversal of the donor substrate preference for UDP-glucose was obtained only with a combination of the three mutations (Table 1). The triple mutant showed a ~5 fold increase in ADP-glucose K_m^{app} and a reduction for the UDP-glucose K_m^{app} by more than 3 fold (Table 1). The catalytic efficiency was also completely

reversed for ADP- and UDP-glucose between the wild-type and the triple mutant (Table 1). Although Leu319Ile showed the largest contribution for ADP-glucose K_m^{app} increase, Val363Phe and Glu367Leu were also determinant for the reversal to UDP-Glucose preference (Table 1). This can be explained by the fact that the phenylalanine substitution at position 363 forces the loop between β-14 and α-14 to move towards the nucleotide binding site (Figure 4C), thus establishing stronger interactions with a pyrimidine nucleotide and clashing with purine ones. This effect is also observed for both ADP- and GDP-glucose, with a ~2.5 fold increase in K_m^{app} when compared to the wild type. The leucine substitution at position 367 further helps in this move since the hydrogen bonds between the side chains of Arg266 and Glu367 are no longer present and α-14 is repelled from Arg266 due to the Leu side chain (Figure 4C).

Feedback inhibition

We soaked both trehalose and T6P into apo OtsA crystals and solved structures with both ligands (Figures 4D and 4E). In both structures OtsA active site assumes an open conformation and two compounds superpose almost perfectly at active site, recapitulating all of the interactions for the donor glucose and also interacting with His199 and Arg286, the latter a residue that interacts with the phosphates of the donor substrate (Figures 4D and 4E). The structure with T6P further shows the phosphate group of the product occupying a similar position to the phosphate group of G6P, interacting with the side chains of Tyr90, Gln146, Arg324 and Tyr145 a residue that does not interact with G6P (Figure 4E). Interestingly the buffer present in the crystallization condition N-Cyclohexyl-2-aminoethanesulfonic acid (CHES) is observed in both trehalose and trehalose-6-phosphate bound structures, occupying the site of the nucleotide donor with

the sulphate group binding in a phosphate site and interacting with hydroxyl groups from both glucose units of trehalose (Figures 4D and 4E).

After obtaining these structures, it was expected to observe some degree of feedback inhibition of OtsA by trehalose given that under physiological conditions, free trehalose is highly abundant inside mycobacteria cells reaching up to 15% of the total organic material in actively growing *M. smegmatis* cells (39), while T6P could not be detected in *M. tuberculosis* cell extracts (25). Trehalose was found to inhibit the enzyme with an IC_{50} of ~24 mM (Figure S5). These results suggest that under physiological conditions OtsA can be regulated by trehalose, the final product of the pathway which is highly abundant inside mycobacterial cells (39).

Identification of an allosteric site in *M. thermoresistible* OtsA

Trehalose is a multipurpose molecule that is abundant in mycobacteria. It can function not only as an energy reserve and as structural component of cell wall glycolipids, but can also lead to the synthesis of glycogen through the TreS-Mak-GlgE pathway and be directly synthesised from glycogen from the TreX-TreY-TreZ pathway. It is therefore likely that trehalose synthesis through the OtsA-B pathway is under strong regulation.

In all the crystallization conditions that contained CHES, we could observe this compound in the crystal structures occupying a pocket formed by the contact interface of 2 OtsA protomers in the tetrameric assembly (Figure 5). The CHES sulphate group directly interacts with the side chains of Arg384 of protomer A and Arg213 of protomer B, forming hydrogen bonds with both, but also hydrogen-π interactions with Phe410 of protomer A (Figure 5). The residues composing this site are completely conserved in mycobacterial and closely related species that are likely to harbour OtsA tetramers but not on others species known to have other oligomeric forms (Figure S2).

Arg384 sits in a loop that forms extensive contacts with the donor substrate through Asp385, Gly386, Met387 and Leu389 (Figure 2). Ligands interacting with the side chain of Arg384 could therefore have an impact on the activity of the enzyme making this site a prime candidate for allosteric regulation.

While CHES had no discernible impact on OtsA activity in the tested concentrations, it was reported before that F6P acted as an allosteric regulator of *M. tuberculosis* OtsA (35), and the same was reported for yeast OtsA-B complex (40). We have tested *Mtr*OtsA in the same conditions as those reported by Diez and colleagues (35), and could not detect any effect of F6P for *Mtr*OtsA. Although we did not observe any effect of F6P on enzyme activity for *Mtr*OtsAs, we nevertheless attempted to soak this compound in both apo form and OtsA:ADP:G6P ternary complex crystals, and also co-crystallize it in the presence and absence of ADP and G6P, all in CHES free conditions. None of these conditions provided a structure in which F6P occupied an allosteric site. We could however observe F6P when co-crystallized together with ADP, bound to the acceptor substrate site, with OtsA presenting the active site in a closed conformation (Figure S6).

OtsA was found to interact with a multitude of proteins in a large-scale proteomics study (41), including enolase that catalyses the penultimate step of glycolysis. In our search for an allosteric regulator we, therefore, decided to test several glycolytic metabolites glucose-1-phosphate (G1P), fructose-1,6-biphosphate (F16BP), 3-phosphoglycerate (3PG), the enolase substrate 2-phosphoglyceric acid (2PG) and product phosphoenolpyruvate (PEP), but also the master metabolic regulators cAMP and 2-oxoglutarate (2OG) (42), and assess their impact in OtsA activity.

From the several compounds tested only 2OG and 2PG showed clear inhibition of OtsA activity with IC_{50} of 1.8 mM and 2.3 mM respectively (Table 2 and Figure 5C). To assess whether they could be acting allosterically on *Mtr*OtsA and binding to the site identified

by CHES we performed two mutations on the arginines that interacted with CHES (Arg213Glu and Arg384Glu) and tested the two mutants activity in the presence of 2OG and 2PG. The results show that the Arg213Glu mutation abolished the strong response of the protein to 2OG while there was a ~5 fold increase in 2PG IC_{50} to 12 mM. For the Arg384Glu mutant 2OG becomes unexpectedly becomes a strong activator with an EC_{50} of 0.7 mM while 2PG has no effect (Table 2).

Discussion

Trehalose, an essential disaccharide in mycobacteria, is one of the critical components of the cell wall. The OtsA-B pathway is an essential source of trehalose for these organisms with knockouts of this pathway being growth defective or non-viable (24, 25). Here we reveal the first mycobacterial OtsA structure. The overall structure of *M. thermoresistibile* OtsA is similar to the reported structures of *E. coli*, *S. venezuelae* and *Paraburkholderia xenovorans* OtsAs despite relatively low sequence identity (30-35%). Remarkably, OtsAs have also a highly similar fold to the pseudo-glycosyltransferases VldE and ValL involved in validamycin A synthesis, a potent antifungal agent (43, 44).

Nevertheless, there are substantial differences in the oligomeric organization of bacterial OtsAs, with tetrameric, dimeric and protomeric forms reported. While mycobacteria and related species present conservation of interfaces showing that the observed tetrameric form for *Mtr*OtsA is present in all these organisms, beyond this group such conservation is absent, which is consistent with other oligomeric states (Figure S2).

The structures obtained in this work allow us to explain how mycobacterial OtsA prefers ADP-glucose as a substrate. A cavity found to accommodate the binding of the adenine moiety is surrounded with Leu319 and Arg361, which are critical for substrate recognition. It is further mediated by Val363 carbonyl group that does not allow GDP-

glucose to occupy the same position and Glu367. Mutating these three residues to the *E. coli* OtsA equivalents completely changed the preference to UDP-glucose, the preferred *E. coli* OtsA donor substrate. Both Leu319 and Glu367 are highly conserved in all mycobacterial and closely related species analysed but not beyond this group, while Val363 can be replaced by other small hydrophobic amino acids such as alanine and isoleucine in other mycobacterial species and closely related species (Figure S2). However, outside this group Val363 is replaced by bulkier residues such as tyrosine and phenylalanine (Figure S2). These differences are suggestive of alternative donor substrate preferences for organisms outside the sub-order *Corynebacterineae*.

Trehalose is not only abundant inside the mycobacterial cell but also involved in a cycle that links it to the synthesis and degradation of glycogen (45). We have shown that OtsA is feedback inhibited by trehalose but not by T6P even though the structures obtained show that both compounds form extensive interactions with OtsA. However, given that these two structures were obtained by soaking apo form crystals with high concentrations of trehalose and T6P, and that CHES is also observed at the active site interacting with both compounds, it is possible that these two structures do not entirely reflect their natural interactions with the enzyme. The toxic accumulation of T6P in OtsB2 knockout mutants (25) can be explained by these results since OtsA shows very low sensitivity to T6P. Furthermore trehalose levels are reduced in the mutant (25) further contributing to an increase in OtsA activity leading to higher T6P production.

Regulation of OtsA by phosphorylation and methylation has been observed for yeast trehalose-6-phosphate synthase complex (46, 47) and phosphorylation of *E. coli* OtsA has also been reported (48) in two residues close to the active site. Even though only one of the residues is conserved in mycobacteria (Ser323 for *Mtr*OtsA), phosphorylation of this residue would most likely lead to inactivation of the enzyme due to the location close to

the acceptor binding site (Figure S7). Interestingly this residue is highly conserved in all sequences analysed with a single exception where it is substituted by a threonine (Figure S2), suggesting that phosphorylation of this residue might be a common regulatory mechanism of OtsA activity.

Allosteric regulation was reported for both *M. tuberculosis* OtsA and yeast trehalose phosphate synthase complex, mediated by F6P (35, 40). The structure obtained with F6P shows this compound occupying the donor site, in the presence of ADP, with the enzyme adopting a closed conformation. Nevertheless, OtsA could only use G6P as an acceptor and F6P showed no effect in the enzyme activity, indicating that the presence of F6P in the acceptor site was forced by the high concentration used for co-crystallization (5mM). The observed differences in regulation by F6P between mycobacterial enzymes are difficult to reconcile since the allosteric site is highly conserved in mycobacterial OtsAs and closely related species (Figure S2).

*Mtr*OtsA is allosteric regulated by 2PG, the substrate of enolase, a glycolytic enzyme, and 2OG a master metabolic regulator that sits at the interface of carbon and nitrogen metabolism. The observed effect is within relevant physiological concentrations for both compounds (49, 50).

The effect of 2PG on OtsA and the reported interaction with enolase suggests an interplay between these two enzymes that would be interesting to explore further on the enolase side. Enolase is a multifunctional protein involved in a variety of cellular processes, beyond its enzymatic activity, that include response to oxidative and thermal stress (51). Given that trehalose is a known chemical chaperone and compatible solute (1), the association between enolase and OtsA, through the effect of 2PG in OtsA activity and the reported interaction between the two enzymes, points to further regulation to stress response.

2OG is a molecule that sits at the interface of carbon and nitrogen metabolism and that has been shown to regulate many different pathways (42). The role of trehalose as an energy reserve molecule (1) and its relationship with synthesis and degradation of glycogen (52) points to a mechanism in which 2OG influences the synthesis of these two compounds through regulation of OtsA activity that deserves to be further explored.

We have shown that OtsA can be allosteric inhibited, however a single mutation at the allosteric site changed the behaviour of 2OG, but not 2PG, from an inhibitor to an activator. This hints at the possible existence of allosteric activators of OtsA yet to be discovered.

The conservation of allosteric site residues and oligomeric assembly in the sub-order *Corynebacterineae* but not outside, suggests that allosteric regulation of OtsA through this site might be limited to this group of organisms. The results of this work are a significant step forward in understanding the regulation of trehalose synthesis in mycobacteria the structural reasons behind substrate preference, and provide new important insight into this enzyme.

Methods

Bacterial strains and cloning.

M. thermoresistibile (DSM 44167) *otsA* gene was amplified from chromosomal DNA obtained from the Deutsch Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ, Braunschweig, Germany). Primers were designed based on the sequence available in NCBI database and the gene was cloned between the BamHI and HindIII sites in pET28a vector (Novagen), modified with an N-terminal 6xHis-SUMO tag. The resulting plasmid was confirmed by DNA sequencing and transformed into *E. coli*

BL21(DE3) strain (Invitrogen). Six *Mtr*OtsA mutants (Arg213Glu, Leu319Ile, Val363Phe, Arg384Glu, Leu319Ile-Glu367Leu and Leu319Ile-Val363Phe-Glu367Leu) were also constructed by site-specific mutagenesis, sequenced and transformed into *E. coli* BL21(DE3) strain. Primers used in this work are listed in table S3.

Recombinant expression and protein purification

Transformed *E. coli* BL21(DE3) cells were grown to mid-exponential growth phase (OD_{610} = 0.6) in LB media (Invitrogen) containing 30 mg L^{-1} kanamycin at 37 ºC. Isopropyl β–D-1-thiogalactopyranoside (IPTG) was then added at a final concentration of 0.5 mM to induce gene expression and the temperature was lowered to 18 ºC. Cells were harvested 18 h-20 h later by centrifugation and re-suspended in 20 mM TRIS pH 7.5, 500 mM NaCl and 20 mM Imidazole with protease inhibitor tablets (Roche), DNAseI and 5 mM $MgCl_2$. Cells were lysed by sonication and cell lysate was centrifuged at 27000 g for 30 min to remove cell debris.

Recombinant *Mtr*OtsA was purified with a HiTrap IMAC Sepharose FF column (GE-Healthcare), equilibrated with 20 mM TRIS pH 7.5, 500 mM NaCl and 20 mM Imidazole. Elution was performed in the same buffer, but with 500 mM Imidazole. Imidazole was removed with a desalting column and SUMO tag was cleaved overnight at 4 ºC by adding Ulp1 protease at 1:100 ratio in 20 mM TRIS pH 7.5, 500 mM NaCl. SUMO tag, Ulp1 protease and uncleaved SUMO-OtsA were removed with a HiTrap IMAC Sepharose FF column (GE-Healthcare), equilibrated with 20 mM TRIS pH 7.5, 500 mM NaCl and 20 mM imidazole. Flow through containing OtsA was collected, concentrated and loaded to a Superdex 200 column equilibrated with 20 mM TRIS pH 7.5, 500 mM NaCl. Fraction purity was determined by SDS-page and purest fractions were pooled concentrated to ~10

mg.ml^{-1} in 20 mM TRIS pH 7.5 350 mM NaCl, flash frozen in liquid nitrogen and stored at -80 ºC. The same purification protocol was used for all *Mtr*OtsA mutants.

Crystallization and data collection

*Mtr*OtsA crystallization screens and optimization were performed at 18 ºC using sitting-drop vapour diffusion method. 300 nl of pure OtsA at 10 mg.ml^{-1} was mixed in 1:1 ratio with well solution using a Phoenix robot (Art Robbins). Initial conditions were obtained in classics suite crystallization screen (Qiagen), solution 36. Crystals obtained in this condition diffracted only up to 3 Å. Therefore, further optimization was performed using the additive screen HT (Hampton Research) and ethylene glycol was found to be the best additive. The final optimized condition consisted of 0.7 M sodium potassium tartrate, 0.1 M CHES pH 10 and 10% v/v ethylene glycol. Crystals appeared after 4 days in this condition. To obtain ligand-bound structures soaking was performed in the optimized condition using the hanging drop vapour diffusion method as follows: 1 µl of protein storage buffer containing 5 mM of ligand was mixed with 1 µl of reservoir solution and drops were left to equilibrate against 500 µl of reservoir solution for 3 days. Crystals were then transferred to the pre-equilibrated drops and incubated for 24h. A cryogenic solution was prepared by adding ethylene glycol up to 25% v/v to mother liquor. Crystals were briefly transferred to this solution, flash frozen in liquid nitrogen and stored for data collection. To obtain an ADP-G6P-OtsA ternary complex co-crystallization with 5mM ADP and G6P was performed instead, as all attempts to soak G6P alone or G6P in the presence of ADP have failed. Crystals were obtained in the Wizard Classic I&II screen (Rigaku), solution E10 and were flash frozen in liquid nitrogen after a brief soak in a solution containing mother liquor and 25% ethylene glycol. The same condition was used to obtain ADP:F6P:OtsA complex.

All data sets were collected at stations I04, I04-1 and I24 at Diamond Light Source (Oxford, UK). Data collection and refinement statistics are summarized in (Table S1).

Structure solution and refinement

Diffraction data were processed and reduced using MOSFLM (53) and Aimless (54) from the CCP4 suite (55), or autoPROC from Global Phasing Limited (56). The apo form crystalized in $I4_122$ space group with one protomer per asymmetric unit. ADP-G6P-OtsA ternary complex crystallized in $P6_222$ space group again with one protomer per asymmetric unit.

Initial phases were determined with PHASER (57) from PHENIX software package (58) using the structure of *E. coli* OtsA (PDB entry 1UQU) (32) as a search model. Model building was done with Coot (59) and refinement was performed in PHENIX (58). Structure validation was performed using Coot and PHENIX tools (58, 59). All the figures were prepared with Pymol (http://www.pymol.org).

Multiple sequence alignment

The multiple sequence alignment was performed with T-Coffee (60) using the OtsA sequences of *Mycobacterium thermoresistibile*, *Mycobacterium tuberculosis*, *Mycobacterium leprae*, *Mycobacterium avium*, *Mycobacterium abscessus*, *Nocardia farcinica*, *Arthrobacter alpinus*, *Salmonella typhimurium*, *Escherichia coli*, *Candida albicans*, *Paraburkholderia xenovorans* and *Streptomyces venezuelae*.

Prediction of mutations in ligand affinity

To provide an insight on which residues to mutate, we used mCSM-lig a software that predicts the effect of mutations in ligand affinity (38) on the X-ray crystal structure of OtsA with ADP-glucose (5JIO).

Enzymatic Assays

Formation of T6P was assessed by a continuous colorimetric assay that followed the release of NDP by measuring the oxidation of NADH at 340nm, in the presence of pyruvate kinase and lactate dehydrogenase. All reagents used were obtained from Sigma-Aldrich. The enzymatic reactions (200 µl) were performed at 37 ºC and contained 50 mM TRIS pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$, 50 mM KCl, 0.3 mM NADH, 2.5 mM phosphoenolpyruvate (PEP), 10 units of pyruvate kinase/lactate dehydrogenase, 0.1 µM enzyme and varying concentrations of G6P and NDP-glucose to determine kinetic parameters.

To assess the specificity of OtsA for G6P, several compounds (F6P, Mannose-6-phosphate, glucosamine-6-phosphate and G1P) were tested for their capacity to act as acceptor substrates, in the conditions above, using ADP-glucose as the donor substrate. The effect of trehalose and T6P as feedback inhibitors, and G1P, F6P, F16BP, CHES, 2PG, 3PG, cAMP and 2OG as allosteric regulators, was tested in the conditions described above but with fixed concentrations of G6P and ADP-glucose (both 0.3 mM). The same conditions were used to test possible allosteric regulators for wild type OtsA and mutants Arg213Glu and Arg384Glu. All experiments were performed in triplicate in a PheraStar plate reader (BMG Labtech) and data was analysed with Prism 5 (Graphpad Software).

To assess the inhibitory effect of ADP and allosteric effect of phosphoenolpyruvate an end-point assay was performed instead. Reactions (100 µl) containing 50 mM TRIS pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$, 50 mM KCl, 4 µM *Mtr*OtsA, 0.3 mM ADP-glucose,

0.3 mM G6P and varying concentrations of ADP and PEP were incubated at 37 ºC and stopped at different time points with 10 µl of 1 M HCl, incubated for 1 minute and neutralized with NaOH. The reactions were diluted with 100 µl acetonitrile and run at 25 ºC on a Waters I-class UHPLC with a PDA detector (258 nm) using a ACQUITY UHPLC BEH Amide column (2.1 x 100 mm, 1.7 µm). Gradient elution (delivered at 0.13 mL/min) was employed using 80/20 acrilamide/water with 0.1 % NH_2OH (A) and water with 0.1 % NH_2OH (B) which started at 90 % A and decreased linearly to 55 % A over 10 min.

Isothermal Titration Calorimetry

Binding interaction between OtsA and ligands was characterized at 25 ºC, using a Microcal ITC200 titration calorimeter (Microcal). OtsA concentration of 60 µM was used for all titrations. Ligands (1 mM) were injected in 2ul aliquots. Titration data was recorded in 20mM TRIS pH 7.5, 500mM NaCl. Data was analysed by fitting a simple single-site model using Origin software (Microcal).

Accession numbers

The coordinates and structure factors presented in this work have been deposited in the Protein Data Bank under accession codes 5JIJ, 5K41, 5K42, 5JIO, 5L3K, 5K5C and 5K44.

Acknowledgements

This work was funded by Bill & Melinda Gates Foundation (subcontract by the Foundation for the National Institutes of Health - NIH) (OPP1158806). VM would like to acknowledge Fundação para a Ciência e a Tecnologia for a postdoctoral fellowship (SFRH/BPD/79531/2011).

References

1. Nobre A, Alarico S, Maranha A, Mendes V, Empadinhas N. 2014. The molecular biology of mycobacterial trehalose in the quest for advanced tuberculosis therapies. Microbiology 160:1547-70.
2. Iturriaga G, Suarez R, Nova-Franco B. 2009. Trehalose metabolism: from osmoprotection to signaling. Int J Mol Sci 10:3793-810.
3. Webb KM, DiRuggiero J. 2012. Role of Mn2+ and compatible solutes in the radiation resistance of thermophilic bacteria and archaea. Archaea 2012:845756.
4. Singer MA, Lindquist S. 1998. Thermotolerance in Saccharomyces cerevisiae: the Yin and Yang of trehalose. Trends Biotechnol 16:460-8.
5. Lu H, Zhu Z, Dong L, Jia X, Sun X, Yan L, Chai Y, Jiang Y, Cao Y. 2011. Lack of trehalose accelerates H2O2-induced Candida albicans apoptosis through regulating Ca2+ signaling pathway and caspase activity. PLoS ONE 6:e15808.
6. Lee J, Lin EW, Lau UY, Hedrick JL, Bat E, Maynard HD. 2013. Trehalose Glycopolymers as Excipients for Protein Stabilization. Biomacromolecules 14:2561-2569.
7. Kandror O, DeLeon A, Goldberg AL. 2002. Trehalose synthesis is induced upon exposure of Escherichia coli to cold and is essential for viability at low temperatures. Proceedings of the National Academy of Sciences of the United States of America 99:9727-9732.
8. Cardoso FS, Castro RF, Borges N, Santos H. 2007. Biochemical and genetic characterization of the pathways for trehalose metabolism in Propionibacterium freudenreichii, and their role in stress response. Microbiology 153:270-80.
9. Tapia H, Koshland DE. 2014. Trehalose is a versatile and long-lived chaperone for desiccation tolerance. Curr Biol 24:2758-66.
10. Thevelein JM. 1984. Regulation of trehalose mobilization in fungi. Microbiol Rev 48:42-59.
11. Elbein AD. 1974. The metabolism of alpha,alpha-trehalose. Adv Carbohydr Chem Biochem 30:227-56.
12. Elbein AD, Pan YT, Pastuszak I, Carroll D. 2003. New insights on trehalose: a multifunctional molecule. Glycobiology 13:17R-27R.
13. Djonovic S, Urbach JM, Drenkard E, Bush J, Feinbaum R, Ausubel JL, Traficante D, Risech M, Kocks C, Fischbach MA, Priebe GP, Ausubel FM. 2013. Trehalose biosynthesis promotes Pseudomonas aeruginosa pathogenicity in plants. PLoS Pathog 9:e1003217.
14. Wahl V, Ponnu J, Schlereth A, Arrivault S, Langenecker T, Franke A, Feil R, Lunn JE, Stitt M, Schmid M. 2013. Regulation of flowering by trehalose-6-phosphate signaling in Arabidopsis thaliana. Science 339:704-7.
15. O'Hara LE, Paul MJ, Wingler A. 2013. How do sugars regulate plant growth and development? New insight into the role of trehalose-6-phosphate. Mol Plant 6:261-74.
16. Paul MJ, Primavesi LF, Jhurreea D, Zhang Y. 2008. Trehalose metabolism and signaling. Annu Rev Plant Biol 59:417-41.
17. Fraenkel D, Nielsen J. 2016. Trehalose-6-phosphate synthase and stabilization of yeast glycolysis. FEMS Yeast Res 16:fov100.
18. Deroover S, Ghillebert R, Broeckx T, Winderickx J, Rolland F. 2016. Trehalose-6-phosphate synthesis controls yeast gluconeogenesis downstream and independent of SNF1. FEMS Yeast Res 16.

19. Williams B, Njaci I, Moghaddam L, Long H, Dickman MB, Zhang X, Mundree S. 2015. Trehalose Accumulation Triggers Autophagy during Plant Desiccation. PLoS Genet 11:e1005705.
20. Rubinsztein DC, Bento CF, Deretic V. 2015. Therapeutic targeting of autophagy in neurodegenerative and infectious diseases. J Exp Med 212:979-90.
21. Boot M, van Winden VJC, Sparrius M, van de Weerd R, Speer A, Ummels R, Rustad T, Sherman DR, Bitter W. 2017. Cell envelope stress in mycobacteria is regulated by the novel signal transduction ATPase IniR in response to trehalose. PLoS Genet 13:e1007131.
22. Avonce N, Mendoza-Vargas A, Morett E, Iturriaga G. 2006. Insights on the evolution of trehalose biosynthesis. BMC Evol Biol 6:109.
23. Zaparty M, Hagemann A, Brasen C, Hensel R, Lupas AN, Brinkmann H, Siebers B. 2013. The first prokaryotic trehalose synthase complex identified in the hyperthermophilic crenarchaeon Thermoproteus tenax. PLoS ONE 8:e61354.
24. Murphy HN, Stewart GR, Mischenko VV, Apt AS, Harris R, McAlister MS, Driscoll PC, Young DB, Robertson BD. 2005. The OtsAB pathway is essential for trehalose biosynthesis in Mycobacterium tuberculosis. J Biol Chem 280:14524-9.
25. Korte J, Alber M, Trujillo CM, Syson K, Koliwer-Brandl H, Deenen R, Kohrer K, DeJesus MA, Hartman T, Jacobs WR, Jr., Bornemann S, Ioerger TR, Ehrt S, Kalscheuer R. 2016. Trehalose-6-Phosphate-Mediated Toxicity Determines Essentiality of OtsB2 in Mycobacterium tuberculosis In Vitro and in Mice. PLoS Pathog 12:e1006043.
26. Errey JC, Lee SS, Gibson RP, Martinez Fleites C, Barry CS, Jung PM, O'Sullivan AC, Davis BG, Davies GJ. 2010. Mechanistic insight into enzymatic glycosyl transfer with retention of configuration through analysis of glycomimetic inhibitors. Angew Chem Int Ed Engl 49:1234-7.
27. Gibson RP, Turkenburg JP, Charnock SJ, Lloyd R, Davies GJ. 2002. Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. Chem Biol 9:1337-46.
28. Lee SS, Hong SY, Errey JC, Izumi A, Davies GJ, Davis BG. 2011. Mechanistic evidence for a front-side, SNi-type reaction in a retaining glycosyltransferase. Nat Chem Biol 7:631-8.
29. Miao Y, Tenor JL, Toffaletti DL, Maskarinec SA, Liu J, Lee RE, Perfect JR, Brennan RG. 2017. Structural and In Vivo Studies on Trehalose-6-Phosphate Synthase from Pathogenic Fungi Provide Insights into Its Catalytic Mechanism, Biological Necessity, and Potential for Novel Antifungal Drug Design. MBio 8.
30. Asencion Diez MD, Miah F, Stevenson CE, Lawson DM, Iglesias AA, Bornemann S. 2017. The Production and Utilization of GDP-glucose in the Biosynthesis of Trehalose 6-Phosphate by Streptomyces venezuelae. J Biol Chem 292:945-954.
31. Jiang Y, Chen XM, Liu YJ, Li YT, Zhang HH, Dyson P, Sheng HM, An LZ. 2010. The catalytic efficiency of trehalose-6-phosphate synthase is effected by the N-loop at low temperatures. Arch Microbiol 192:937-43.
32. Gibson RP, Tarling CA, Roberts S, Withers SG, Davies GJ. 2004. The donor subsite of trehalose-6-phosphate synthase: binary complexes with UDP-glucose and UDP-2-deoxy-2-fluoro-glucose at 2 A resolution. J Biol Chem 279:1950-5.
33. Pan YT, Carroll JD, Elbein AD. 2002. Trehalose-phosphate synthase of Mycobacterium tuberculosis. Cloning, expression and properties of the recombinant enzyme. Eur J Biochem 269:6091-100.
34. Chen X, An L, Fan X, Ju F, Zhang B, Sun H, Xiao J, Hu W, Qu T, Guan L, Tang S, Chen T, Liu G, Dyson P. 2017. A trehalose biosynthetic enzyme doubles as an osmotic stress sensor to regulate bacterial morphogenesis. PLoS Genet 13:e1007062.
35. Asencion Diez MD, Demonte AM, Syson K, Arias DG, Gorelik A, Guerrero SA, Bornemann S, Iglesias AA. 2015. Allosteric regulation of the partitioning of glucose-1-phosphate

between glycogen and trehalose biosynthesis in Mycobacterium tuberculosis. Biochim Biophys Acta 1850:13-21.
36. Killick KA. 1979. Trehalose 6-phosphate synthase from Dictyostelium discoideum: partial purification and characterization of the enzyme from young sorocarps. Arch Biochem Biophys 196:121-33.
37. Silva Z, Alarico S, da Costa MS. 2005. Trehalose biosynthesis in Thermus thermophilus RQ-1: biochemical properties of the trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase. Extremophiles 9:29-36.
38. Pires DE, Blundell TL, Ascher DB. 2016. mCSM-lig: quantifying the effects of mutations on protein-small molecule affinity in genetic disease and emergence of drug resistance. Sci Rep 6:29575.
39. Shleeva MO, Trutneva KA, Demina GR, Zinin AI, Sorokoumova GM, Laptinskaya PK, Shumkova ES, Kaprelyants AS. 2017. Free Trehalose Accumulation in Dormant Mycobacterium smegmatis Cells and Its Breakdown in Early Resuscitation Phase. Front Microbiol 8:524.
40. Londesborough J, Vuorio OE. 1993. Purification of trehalose synthase from baker's yeast. Its temperature-dependent activation by fructose 6-phosphate and inhibition by phosphate. Eur J Biochem 216:841-8.
41. Hu P, Janga SC, Babu M, Diaz-Mejia JJ, Butland G, Yang W, Pogoutse O, Guo X, Phanse S, Wong P, Chandran S, Christopoulos C, Nazarians-Armavil A, Nasseri NK, Musso G, Ali M, Nazemof N, Eroukova V, Golshani A, Paccanaro A, Greenblatt JF, Moreno-Hagelsieb G, Emili A. 2009. Global functional atlas of Escherichia coli encompassing previously uncharacterized proteins. PLoS Biol 7:e96.
42. Huergo LF, Dixon R. 2015. The Emergence of 2-Oxoglutarate as a Master Regulator Metabolite. Microbiol Mol Biol Rev 79:419-35.
43. Zheng L, Zhou X, Zhang H, Ji X, Li L, Huang L, Bai L. 2012. Structural and functional analysis of validoxylamine A 7'-phosphate synthase ValL involved in validamycin A biosynthesis. PLoS ONE 7:e32033.
44. Cavalier MC, Yim YS, Asamizu S, Neau D, Almabruk KH, Mahmud T, Lee YH. 2012. Mechanistic Insights into Validoxylamine A 7 '-Phosphate Synthesis by VldE Using the Structure of the Entire Product Complex. PLoS ONE 7.
45. Chandra G, Chater KF, Bornemann S. 2011. Unexpected and widespread connections between bacterial glycogen and trehalose metabolism. Microbiology 157:1565-72.
46. Trevisol ET, Panek AD, De Mesquita JF, Eleutherio EC. 2014. Regulation of the yeast trehalose-synthase complex by cyclic AMP-dependent phosphorylation. Biochim Biophys Acta 1840:1646-50.
47. Sengupta S, Banerjee S, Lahiri S, Dutta T, Dhar TK, Ghosh AK. 2014. Purification, characterization, sequencing and molecular cloning of a novel cysteine methyltransferase that regulates trehalose-6-phosphate synthase from Saccharomyces cerevisiae. Biochim Biophys Acta 1840:1861-71.
48. Potel CM, Lin MH, Heck AJR, Lemeer S. 2018. Widespread bacterial protein histidine phosphorylation revealed by mass spectrometry-based proteomics. Nat Methods 15:187-190.
49. Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. Nat Chem Biol 5:593-9.
50. Albe KR, Butler MH, Wright BE. 1990. Cellular concentrations of enzymes and their substrates. J Theor Biol 143:163-95.
51. Sharma S, Jadli M, Singh A, Arora K, Malhotra P. 2014. A secretory multifunctional serine protease, DegP of Plasmodium falciparum, plays an important role in thermo-oxidative stress, parasite growth and development. FEBS J 281:1679-99.

52. Bornemann S. 2016. alpha-Glucan biosynthesis and the GlgE pathway in Mycobacterium tuberculosis. Biochem Soc Trans 44:68-73.
53. Leslie AGW, Powell HR. 2007. Processing diffraction data with MOSFLM. Evolving Methods for Macromolecular Crystallography 245:41-51.
54. Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? Acta Crystallographica Section D-Biological Crystallography 69:1204-1214.
55. 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.
56. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, Bricogne G. 2011. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr D Biol Crystallogr 67:293-302.
57. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. J Appl Crystallogr 40:658-674.
58. 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.
59. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486-501.
60. Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang JM, Taly JF, Notredame C. 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Res 39:W13-7.
61. Jubb HC, Higueruelo AP, Ochoa-Montano B, Pitt WR, Ascher DB, Blundell TL. 2017. Arpeggio: A Web Server for Calculating and Visualising Interatomic Interactions in Protein Structures. J Mol Biol 429:365-371.

Figures

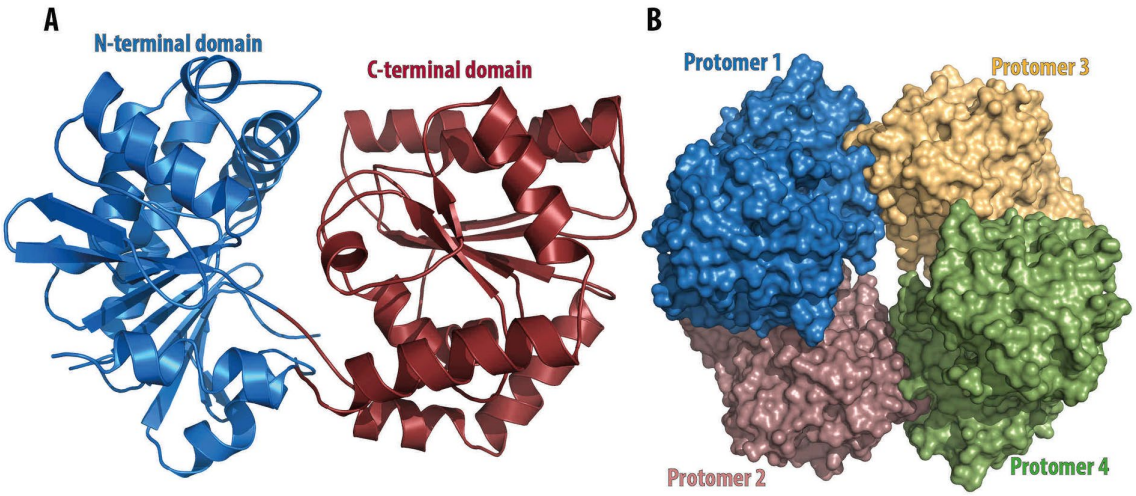


Figure 1: (A) Representation of the overall structure of *M. thermoresistibile* OtsA. The N-terminal domain consists of residues 1-247 and 462-486, the C-terminal domain of residues 248-461. View of *M. thermoresistibile* OtsA tetramer (B).

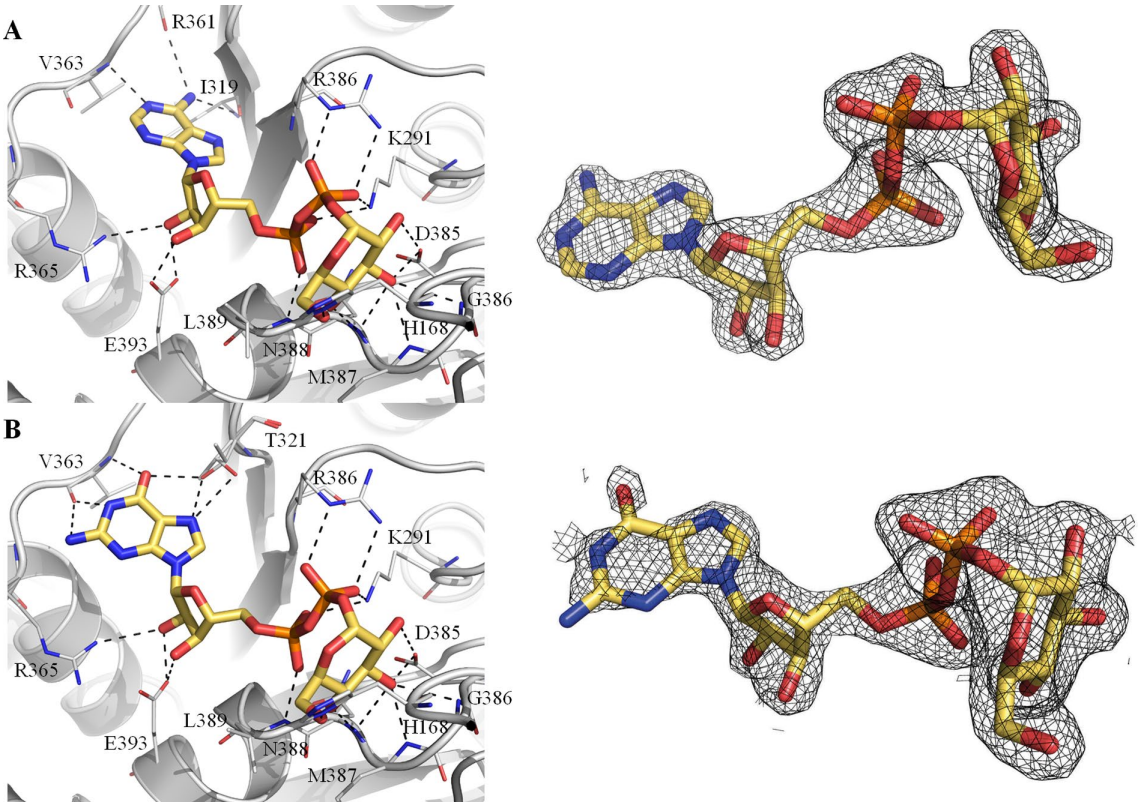
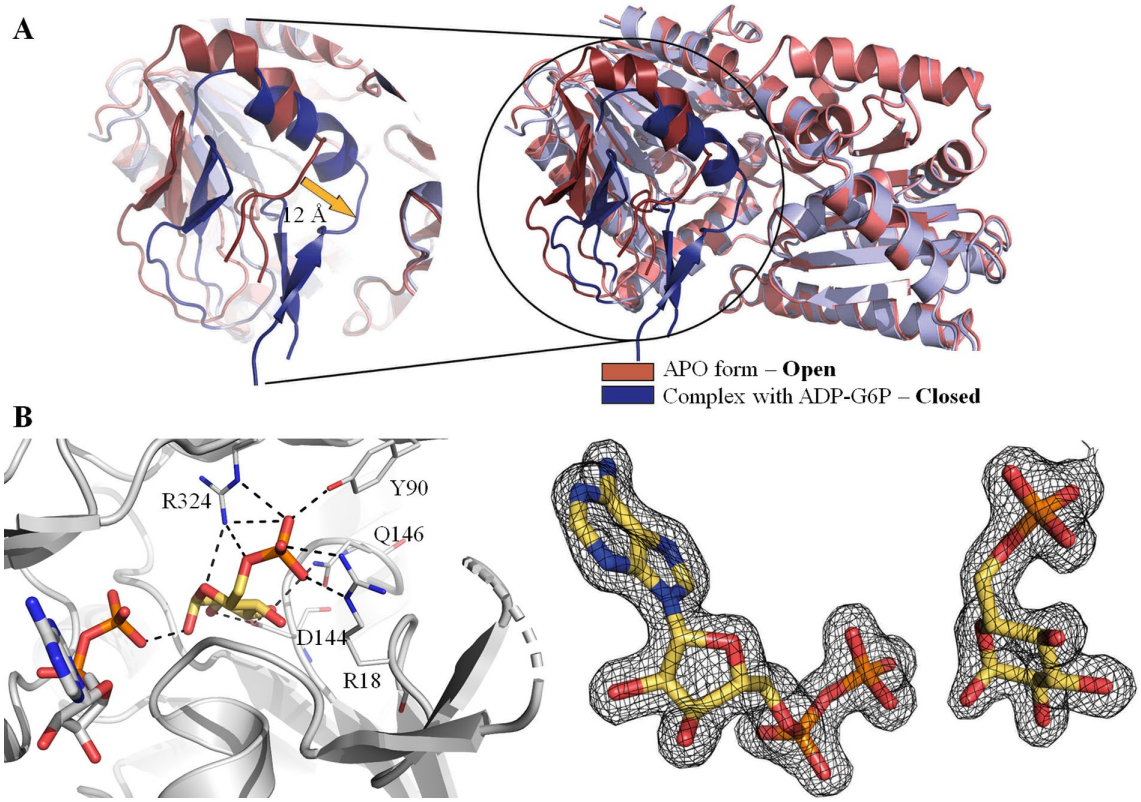


Figure 2: Detailed view of the active site of *M. thermoresistibile* OtsA with ADP-glucose (A) and with GDP-glucose bound (B) with Fo-Fc "omit maps" shown contoured at 1.5 σ. Black dashed lines represent hydrogen bonds.



724

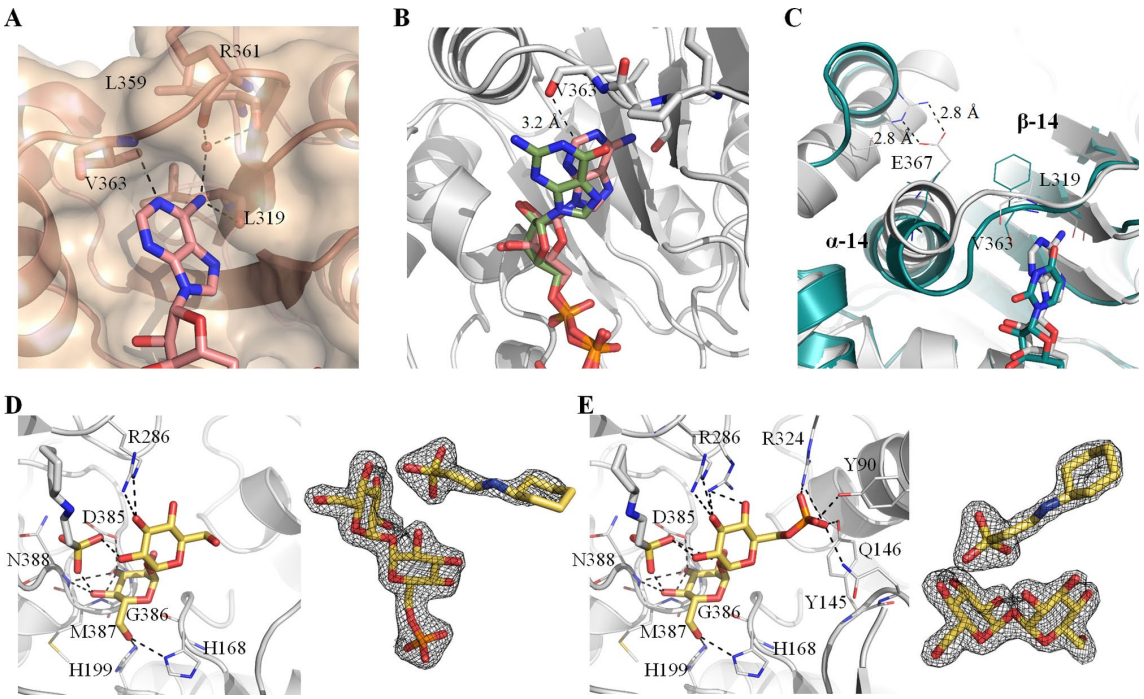
725 **Figure 3:** View of OtsA in an open (salmon and red) and closed conformation (violet and

726 blue) and superposition of the two conformations. (B) Representation of the active site of

727 *M. thermoresistibile* OtsA with ADP (white) and G6P (yellow) bound with Fo-Fc "omit

728 map" shown contoured at 1.5 σ. Black dashed lines represent hydrogen bonds.

729



730

731 **Figure 4:** View of the binding site of the adenine moiety of ADP-glucose. A water that
732 mediates the adenine moiety interaction with OtsA is represented as a red sphere (A).
733 Superposition of ADP-glucose and GDP-glucose structures (B). Superposition of
734 *Mtr*OtsA structure with ADP-glucose bound and *E. coli* OtsA structure with UDP-glucose
735 bound (PDB code: 1UQU) (C). Alpha helix 14 (α-14) and beta strand 14 (β-14) are
736 shown. Detailed view of the active site of *M. thermoresistibile* OtsA with trehalose (D)
737 and with T6P bound (E) with CHES visible in both structures. Fo-Fc "omit maps" are
738 shown contoured at 1.5 σ for trehalose, T6P and CHES. Black dashed lines represent
739 hydrogen bonds.

740

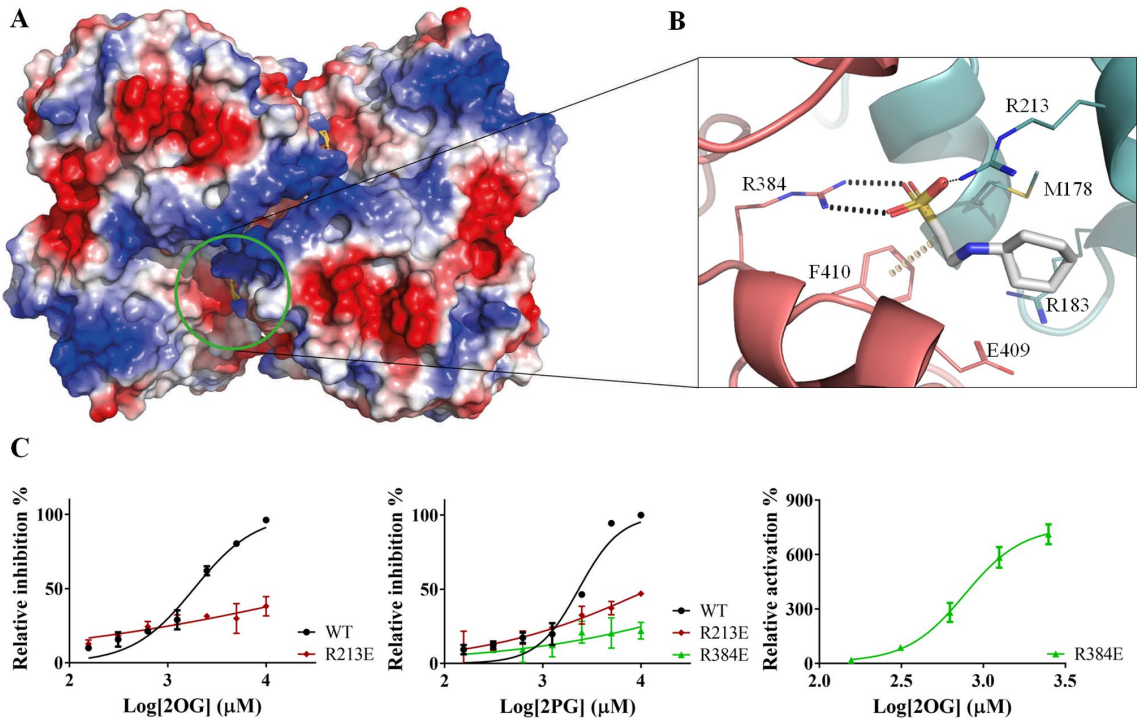


Figure 5: View of the allosteric site of *M. thermoresistibile* OtsA with CHES bound with protein surface electrostatic potential shown (A). Detailed view of the CHES binding site and its interactions with OtsA (B). The interactions were calculated with Arpeggio (61) using the apo structure (5JIJ). Black dots represent hydrogen bonds and yellow disks a carbon-π interaction. The two protomers are colored differently. (C) Activity profiles of *M. thermoresistibile* wild type OtsA and the allosteric site mutants Arg213Glu and Arg384Glu in the presence of the allosteric effectors 2OG and 2PG.

Tables

Table 1: Kinetic parameters of *M. thermoresistibile* OtsA.

Substrate	K_m^{app} (mM)	k_{cat}^{app} (s^{-1})	k_{cat}/K_m ($s^{-1}.mM^{-1}$)

WT	ADP-Glucose	0.25 ± 0.02	26 ± 1	104 ± 13
	GDP-Glucose	0.29 ± 0.02	5.1 ± 0.2	18 ± 2
	UDP-Glucose	1.7 ± 0.1	36 ± 1	21 ± 2
	Glucose-6P	3.3 ± 0.1	26 ± 1	7.9 ± 0.5
L319I	ADP-Glucose	0.71 ± 0.18	8.4 ± 0.7	12 ± 2
	GDP-Glucose	0.35 ± 0.04	3.1 ± 0.1	8.9 ± 1.4
	UDP-Glucose	1.4 ± 0.1	8.9 ± 0.4	6.4 ± 0.7
	Glucose-6P	3.3 ± 0.2	6.4 ± 0.2	1.9 ± 0.2
V363F	ADP-Glucose	0.60 ± 0.03	11 ± 1	18 ± 2
	GDP-Glucose	0.74 ± 0.08	14 ± 1	19 ± 3
	UDP-Glucose	0.68 ± 0.06	13 ± 1	19 ± 3
	Glucose-6P	3.1 ± 0.2	19 ± 1	6.1 ± 0.7
L319I, E367L	ADP-Glucose	0.63 ± 0.04	14 ± 1	22 ± 3
	GDP-Glucose	0.70 ± 0.05	9.2 ± 0.4	13 ± 2
	UDP-Glucose	0.63 ± 0.03	15 ± 1	24 ± 3
	Glucose-6P	3.7 ± 0.2	21 ± 1	5.7 ± 0.5
L319I, V363F, E367L	ADP-Glucose	1.3 ± 0.1	19 ± 1	15 ± 2
	GDP-Glucose	0.90 ± 0.06	16 ± 1	18 ± 2
	UDP-Glucose	0.52 ± 0.05	22 ± 1	42 ± 6
	Glucose-6P	3.2 ± 0.2	17 ± 1	5.3 ± 0.6
R213E	ADP-Glucose	0.31 ± 0.01	56 ± 2	180 ± 13
	Glucose-6P	3.3 ± 0.3	33 ± 2	10 ± 1
R384E	ADP-Glucose	1.9 ± 0.2	3.2 ± 0.1	1.7 ± 0.2
	Glucose-6P	2.9 ± 0.2	2.1 ± 0.2	0.77 ± 0.05

Table 2: Effect of allosteric regulators in *M. thermoresistibile* OtsA activity. IC_{50} and EC_{50} values are in (mM) and 95% confidence intervals are given in brackets.

	WT	**R213E**	**R384E**
2OG	IC_{50} = 1.8	IC_{50} = 73	EC_{50} = 0.8
	(1.6-2.1)	(18-302)	(0.6-0.9)
	n_H = 1.4	n_H = 0.26	n_H = 2.5
	(1.1-1.7)	(0.16-0.35)	(1.4-3.5)
2PG	IC_{50} = 2.3	IC_{50} = 12	IC_{50} < 100
	(2.0-2.7)	(6.3-24)	
	n_H = 2.0	n_H = 0.51	
	(1.4-2.7)	(0.40-0.63)	

n_H – Hill coefficient

Supplemental material

Figure S1: Alignment of the X-ray crystal structures of *E. coli* OtsA (PDB code: 1UQU) in cyan and *M. thermoresistibile* OtsA Apo form in orange (RMSD = 2.996).

Figure S2: Sequence comparison of OtsA from *Mycobacterium thermoresistibile*, *Mycobacterium tuberculosis*, *Mycobacterium leprae*, *Mycobacterium avium*, *Mycobacterium abscessus*, *Nocardia farcinica*, *Arthrobacter alpinus*, *Salmonella typhimurium*, *Escherichia coli*, *Candida albicans*, *Paraburkholderia xenovorans* and *Streptomyces venezuelae*. Residues that contact with the substrates are highlighted with blue circles (donor site) and green circles (acceptor site). The allosteric site residues are marked with crosses and tetramer interfaces are highlight in red. The tetramer interfaces are highly conserved only in the several mycobacteria and *N. farcinica* and less *A. alpinus*. The remaining non-actinobacterial species show very little conservation of the interfaces. The same is observed for the allosteric site. Acceptor site residues are conserved throughout. Donor site residues are less conserved indicating the known differences in substrate preference.

Figure S3: *Mtr*OtsA wt and mutants kinetics. G6P concentration was fixed at 10 mM for all NDP-glucose. ADP-glucose concentration was kept at 2.5 mM for G6P.

Figure S4: ITC trace with *Mtr*OtsA for ADP-glucose binding

Figure S5: Feedback inhibition by trehalose.

Figure S6: (A) View of the active site of *M. thermoresistibile* OtsA complexed with ADP (white) and F6P (yellow), with F6P occupying the acceptor site. Fo-Fc "Omit" map for the two ligands is shown contoured at 1.5 σ (B). Black dashed lines represent hydrogen bonds.

Figure S7: View of the active site of *M. thermoresistibile* OtsA complexed with ADP and G6P (white). Distances between Ser323 hydroxyl group and other atoms interacting with the phosphate group of G6P are shown. Phosphorylation of Ser323 would place the phosphate group in a similar position to the one occupied by G6P.

Table S1: X-ray crystallography data collection and final refinement statistics

Table S2: mCSM-lig predicitons

Table S3: Primers used in this work

Supplemental information

Mycobacterial OtsA structures unveil substrate preference mechanism and allosteric regulation by 2-oxoglutarate and 2-phosphoglycerate.

Vitor Mendes[1*], Marta Acebrón-García-de-Eulate[1], Nupur Verma[1], Michal Blaszczyk[1], Márcio V. B. Dias[2,3] and Tom L. Blundell[1*]

1 - Department of Biochemistry, University of Cambridge, Cambridge CB21GA, UK.

2 - Department of Microbiology, Institute of Biomedical Science, University of São Paulo, São Paulo, Brazil.

3 – Department of Chemistry, University of Warwick, Coventry CV4 7EQ, UK.

*Corresponding authors

Tom L. Blundell

Address: Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, CB2 1GA Cambridge, UK.

Phone: +441223333628

Email: tom@cryst.bioc.cam.ac.uk

Vitor Mendes

Address: Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, CB2 1GA Cambridge, UK.

Phone: +441223766028

Email: vgm23@cam.ac.uk

Results

Table S1: X-ray crystallography data collection and final refinement statistics

Ligand#	**APO**	**ADP-glucose**	**GDP-glucose**	**ADP G6P**	**ADP-F6P**	**Trehalose**	**T6P**
PDB ID	5JIJ	5K41	5K42	5JIO	5L3K	5K5C	5K44
Data collection*							
Space group	$I4_122$	$I4_122$	$I4_122$	$P6_222$	$P6_2$	$I4_122$	$I4_122$
Cell parameters:							
a [Å]	126.86	127.41	127.31	105.36	216.96	127.32	127.23
b [Å]	126.86	127.41	127.31	105.36	216.96	127.32	127.23
c [Å]	207.19	205.62	206.78	158.87	159.84	207.38	207.61
α/β/γ [˚]	90/90/90	90/90/90	90/90/90	90/90/120	90/90/120	90/90/90	90/90/90
Resolution range [Å]	60.66 – 1.82 (1.86 – 1.82)	102.81 – 1.97 (2.20 – 1.97)	108.41 – 1.92 (2.15 – 1.92)	91.24– 1.71 (1.87 – 1.71)	159.84 – 2.31 (2.43 – 2.31)	103.69 – 1.85 (2.02 – 1.85)	103.80 – 1.93 (2.15 – 1.93)
No. of observations							
total	833006 (50109)	753000 (216413)	842990 (231847)	1030194 (193800)	1927301 (266711)	923936 (184758)	817022 (229455)
unique	75576 (4427)	59763 (16773)	64773 (18176)	56816 (13323)	187701 (27326)	72774 (17204)	64496 (18114)
R_{merge}	0.116 (0.716)	0.154 (0.669)	0.078 (0.958)	0.096 (0.790)	0.268 (1.138)	0.076 (0.762)	0.093 (0.544)
I/σ(I)	9.5 (2.0)	9.5 (2.8)	23.1 (3.3)	19.1 (3.3)	7.0 (2.1)	20.1 (3.0)	15.6 (3.5)
Completeness [%]	100.0 (100.0)	100.0 (99.9)	100.0 (99.9)	100.0 (100.0)	99.9 (99.9)	100.0 (100.0)	100.0 (100.0)
Multiplicity	11.0 (11.3)	12.6 (12.9)	13.0 (12.8)	18.1 (14.5)	10.3 (9.8)	12.7 (10.7)	12.7 (12.7)
Refinement							
Refinement program	PHENIX	PHENIX	PHENIX	PHENIX	PHENIX	PHENIX	PHENIX
Resolution [Å]	60.66 – 1.82	90.09 – 1.97	45.01 – 1.92	91.24 – 1.71	121.74 – 2.31	41.58 – 1.85	40.23 – 1.93
No. reflections	74139	59634	64577	56756	187654	72668	64406
R_{work}/R_{free} [%]	16.5/18.6	16.5/19.1	15.6/18.4	15.2/17.9	15.3/21.3	15.1/17.5	14.9/17.4
RMS deviations							
Bonds [Å]	0.007	0.007	0.007	0.007	0.009	0.007	0.008
Angles [˚]	1.009	1.042	1.045	1.092	1.048	1.028	1.030
Ramachandran							
Favoured [%]	97	97	97	97	95	97	96
Outliers [%]	0	0	0	0	0	0	0
Average B-factor [Å^2]							
macromolecule	40.3	44.8	44.9	31.2	32.6	40.2	40.7
ligands	49.3	43.0	66.0	25.8	30.9	45.5	43.0
solvent	43.7	49.2	50.0	39.8	30.8	47.6	48.0

* Parameters shown in brackets are for the highest resolution shell

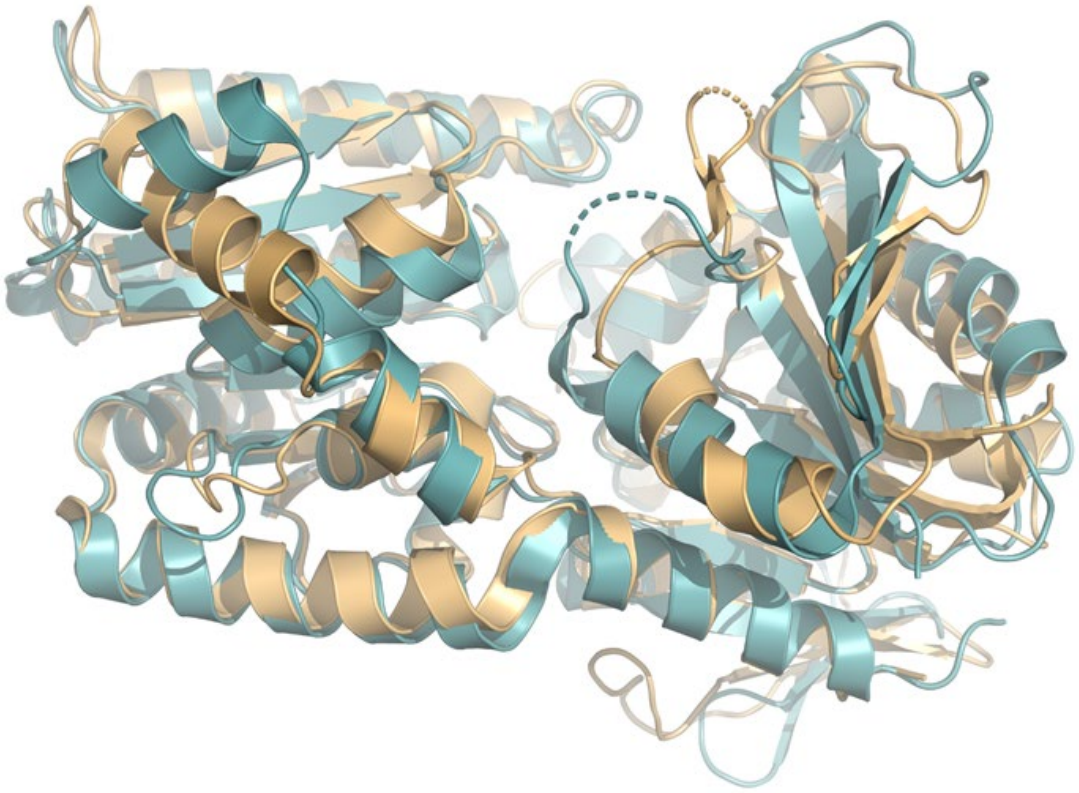
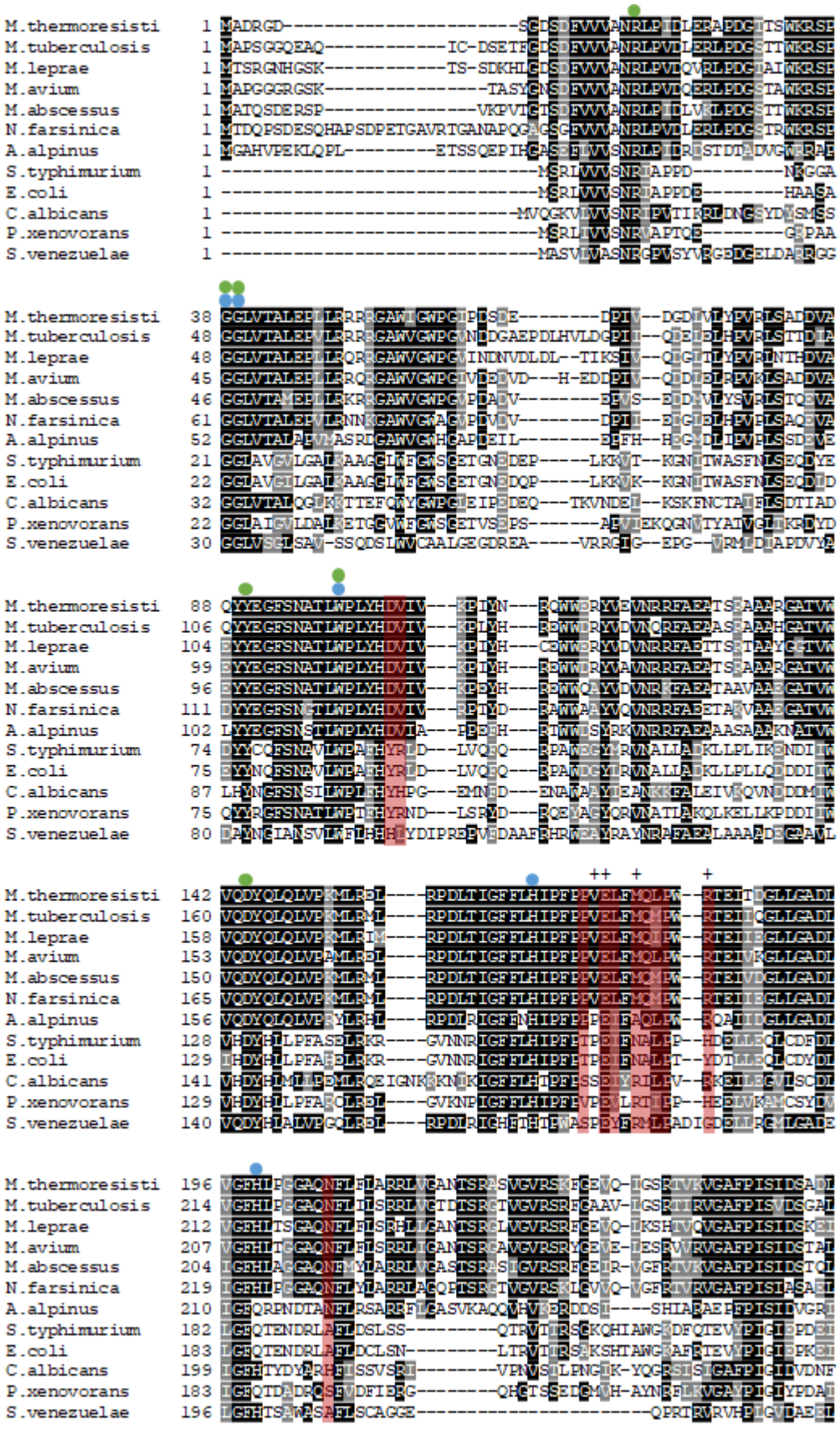


Figure S1: Alignment of the X-ray crystal structures of *E. coli* OtsA (PDB code: 1UQU) in cyan and *M. thermoresistibile* OtsA Apo form in orange (RMSD = 2.996).


M.thermoresisti
M.tuberculosis
M.leprae
M.avium
M.abscessus
N.farsinica
A.alpinus
S.typhimurium
E.coli
C.albicans
P.xenovorans
S.venezuelae

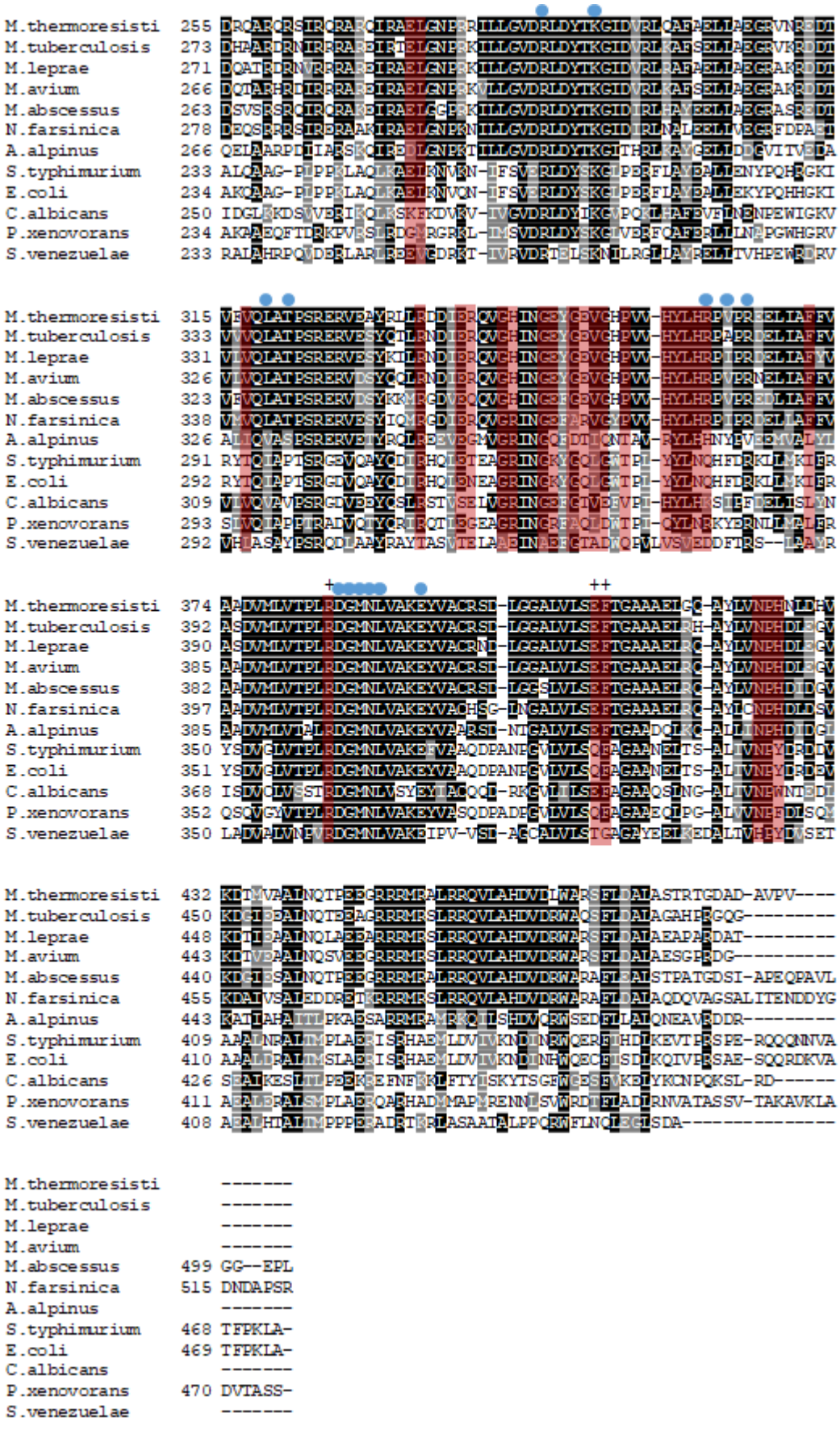

M.thermoresisti
M.tuberculosis
M.leprae
M.avium
M.abscessus
N.farsinica
A.alpinus
S.typhimurium
E.coli
C.albicans
P.xenovorans
S.venezuelae

Figure S2: Sequence comparison of OtsA from *Mycobacterium thermoresistibile*, *Mycobacterium tuberculosis*, *Mycobacterium leprae*, *Mycobacterium avium*, *Mycobacterium abscessus*, *Nocardia farcinica*, *Arthrobacter alpinus*, *Salmonella typhimurium*, *Escherichia coli*, *Candida albicans*, *Paraburkholderia xenovorans* and *Streptomyces venezuelae*. Residues that contact with the substrates are highlighted with blue circles (donor site) and green circles (acceptor site). The allosteric site residues are marked with crosses and tetramer interfaces are highlight in red. The tetramer interfaces are highly conserved only in the several mycobacteria and *N. farcinica* and less *A. alpinus*. The remaining non-actinobacterial species show very little conservation of the interfaces. The same is observed for the allosteric site. Acceptor site residues are conserved throughout. Donor site residues are less conserved indicating the known differences in substrate preference. The multiple sequence alignment was performed with T-Coffee (1).

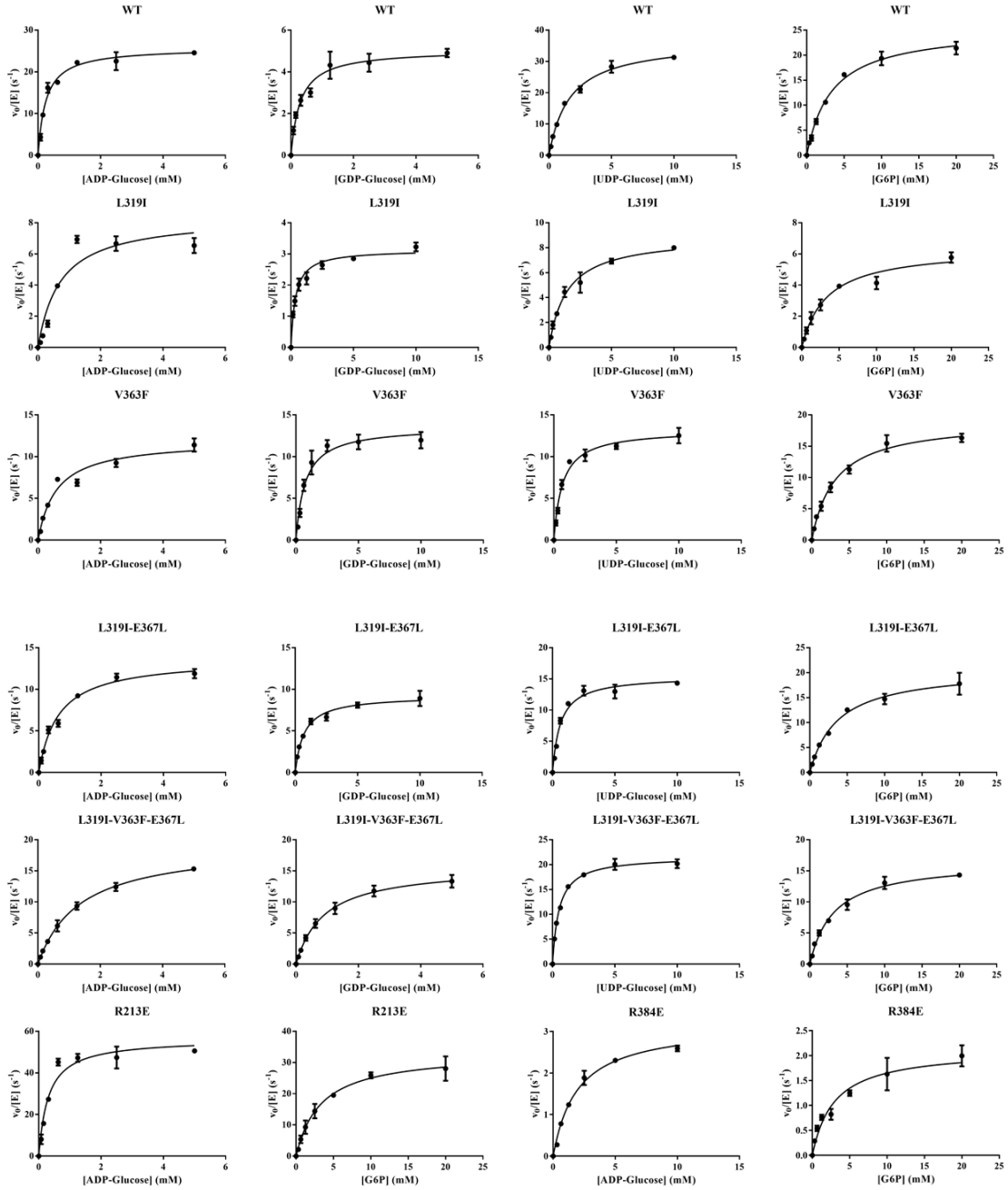


Figure S3: *Mtr*OtsA wt and mutants kinetics. G6P concentration was fixed at 10 mM for all NDP-glucose. ADP-glucose concentration was kept at 2.5 mM for G6P.

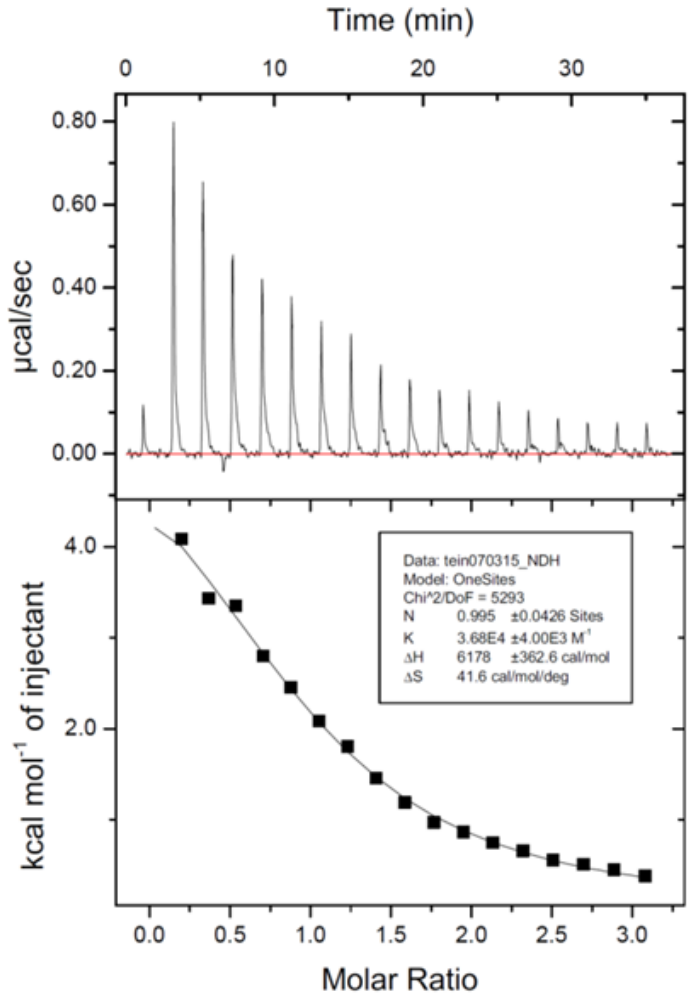


Figure S4: ITC trace with *Mtr*OtsA for ADP-glucose binding

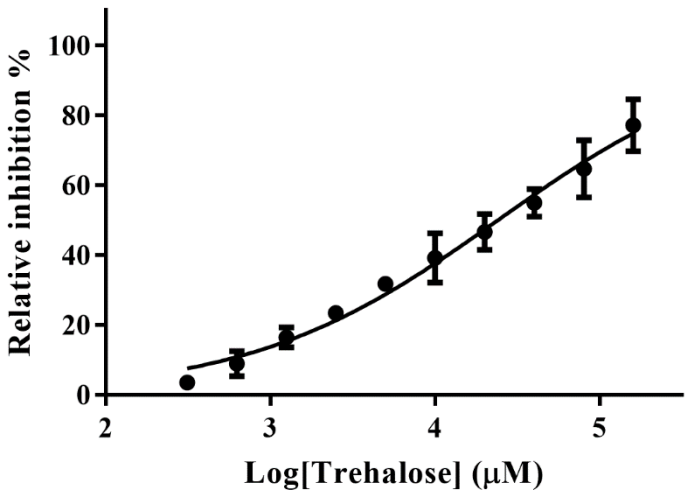
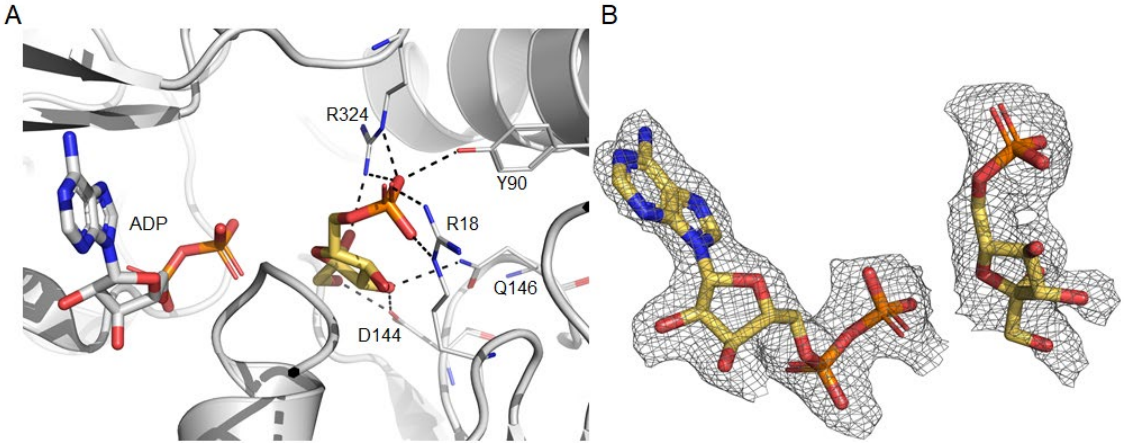
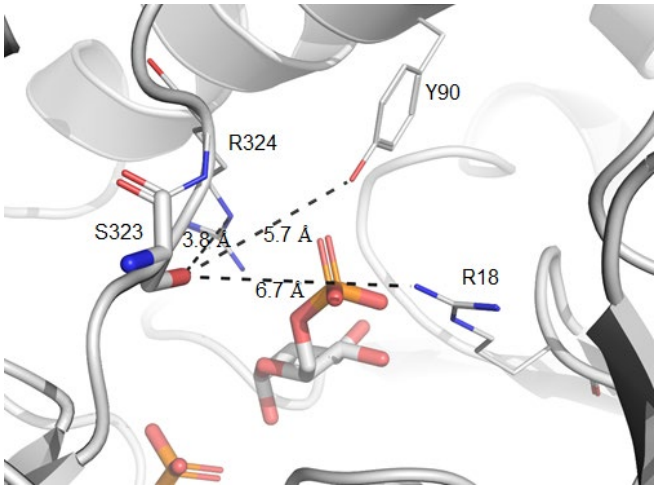


Figure S5: Feedback inhibition by trehalose.



Supplementary Figure S6: (A) View of the active site of *M. thermoresistibile* OtsA complexed with ADP (white) and F6P (yellow), with F6P occupying the acceptor site. Fo-Fc "Omit" map for the two ligands is shown contoured at 1.5 σ (B). Black dashed lines represent hydrogen bonds.



Supplementary Figure S7: View of the active site of *M. thermoresistibile* OtsA complexed with ADP and G6P (white). Distances between Ser323 hydroxyl group and other atoms interacting with the phosphate group of G6P are shown. Phosphorylation of Ser323 would place the phosphate group in a similar position to the one occupied by G6P.

Table S2: mCSM-lig predicitons

Residues	Log change in affinity
***V363F**	-1.4
D285E	-1.3
T321P	-1.3
***L319I**	-1.1
***E367L**	-0.8
P363H	-0.6

*Residues selected for mutation. All residues except E367 are within 4.5 Å of the adenine moiety of ADP-glucose.

Table S3: Primers used in this work

*otsA*pET28SUMO_F	ATAGGATCCATGGCTGACCGGGGCGACTC
*otsA*pET28SUMO_R	ATTAAGCTTTCACACCGGAACCGCGTCGG
L319I_F	GACACCGTCTTCGTCCAGATCGCCACCCCCAGCCGCGAG
L319I_R	CTCGCGGCTGGGGGTGGCGATCTGGACGAAGACGGTGTC
V363F_F	CCTGCACCGGCCGTTTCCGCGTGAGGAAC
V363F_R	GTTCCTCACGCGGAAACGGCCGGTGCAGG
V363F_E367L_F	CCTGCACCGGCCGTTTCCGCGTGAGCTAC
V363F_E367L_R	GTAGCTCACGCGGAAACGGCCGGTGCAGG
R213E_F	CTTCCTGTTCCTGGCGCGAGAGCTGGTGGGCGCCAACAC
R213E_R	GTGTTGGCGCCCACCAGCTCTCGCGCCAGGAACAGGAAG
R384E_F	CATGCTGGTCACCCCGCTGGAGGACGGGATGAACCTGGT
R384E_R	ACCAGGTTCATCCCGTCCTCCAGCGGGGTGACCAGCATG

References

1. Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang JM, Taly JF, Notredame C. 2011. T-Coffee: a web server for the multiple sequence alignment of

protein and RNA sequences using structural information and homology extension. Nucleic Acids Res 39:W13-7.