

RESEARCH ARTICLE

The Tp0684 (MgIB-2) Lipoprotein of *Treponema pallidum*: A Glucose-Binding Protein with Divergent Topology

Chad A. Brautigam¹, Ranjit K. Deka², Wei Z. Liu², Michael V. Norgard^{2*}

1 Department of Biophysics, The University of Texas Southwestern Medical Center, Dallas, TX, 75390, United States of America, **2** Department of Microbiology, The University of Texas Southwestern Medical Center, Dallas, TX, 75390, United States of America

* Michael.Norgard@UTSouthwestern.edu



CrossMark
click for updates

OPEN ACCESS

Citation: Brautigam CA, Deka RK, Liu WZ, Norgard MV (2016) The Tp0684 (MgIB-2) Lipoprotein of *Treponema pallidum*: A Glucose-Binding Protein with Divergent Topology. PLoS ONE 11(8): e0161022. doi:10.1371/journal.pone.0161022

Editor: Sabato D'Auria, Consiglio Nazionale delle Ricerche, ITALY

Received: May 30, 2016

Accepted: June 24, 2016

Published: August 18, 2016

Copyright: © 2016 Brautigam et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Structure factors and the structural model have been deposited in the Protein Data Bank with accession number 5JX2. All other relevant raw data are included in a Supplemental Information file.

Funding: This work was supported by the National Institutes of Health (grant A1056305, www.nih.gov) to MVN. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Treponema pallidum, the bacterium that causes syphilis, is an obligate human parasite. As such, it must acquire energy, in the form of carbon sources, from the host. There is ample evidence that the principal source of energy for this spirochete is D-glucose acquired from its environment, likely via an ABC transporter. Further, there is genetic evidence of a D-glucose chemotaxis system in *T. pallidum*. Both of these processes may be dependent on a single lipidated chemoreceptor: Tp0684, also called TpMgIB-2 for its sequence homology to MgIB of *Escherichia coli*. To broaden our understanding of this potentially vital protein, we determined a 2.05-Å X-ray crystal structure of a soluble form of the recombinant protein. Like its namesake, TpMgIB-2 adopts a bilobed fold that is similar to that of the ligand-binding proteins (LBPs) of other ABC transporters. However, the protein has an unusual, circularly permuted topology. This feature prompted a series of biophysical studies that examined whether the protein's topological distinctiveness affected its putative chemoreceptor functions. Differential scanning fluorimetry and isothermal titration calorimetry were used to confirm that the protein bound D-glucose in a cleft between its two lobes. Additionally, analytical ultracentrifugation was employed to reveal that D-glucose binding is accompanied by a significant conformational change. TpMgIB-2 thus appears to be fully functional *in vitro*, and given the probable central importance of the protein to *T. pallidum*'s physiology, our results have implications for the viability and pathogenicity of this obligate human pathogen.

Introduction

The sexually transmitted disease syphilis is re-emerging as a global health threat [1–8], motivating continued research on the disease's spirochetal etiologic agent, *Treponema pallidum*. Although this obligate human pathogen is a diderm like Gram-negative bacteria, the cell envelope has unique features. Particularly notable are the scarcity of outer-membrane proteins [9–13] and the fact that periplasmic proteins tend to be membrane-anchored via fatty-acyl groups

Abbreviations: 3-OMe Glc, 3-O-methyl glucose; AUC, analytical ultracentrifugation; DSF, differential scanning calorimetry; EcMglB, the MglB protein from *E. coli*; ITC, isothermal titration calorimetry; LBP, ligand-binding protein; r.m.s.d., root-mean-squared deviations; SV, sedimentation velocity; TpMglB-2, the second MglB homolog from *T. pallidum*.

post-translationally added to their respective processed N-termini [14]. This class of proteolipids, also known as “lipoproteins”, are involved in many vital cellular processes, including nutrient import and flavin homeostasis [15–22]. Because *T. pallidum* cannot be cultivated *in vitro*, a structure-to-function approach of this organism’s cadre of lipoproteins has provided new insights into its biology and, more generally, molecular microbiology [15–29].

The acquisition of nutrients, such as carbon sources, nitrogen sources, and cofactors, is a major undertaking for most microorganisms. For *T. pallidum*, these needs are particularly pressing. The bacterium has a minimized genome [30] of only 1,041 genes; in comparison, *Escherichia coli* strain K-12 has over 4,400 [31]. Thus, *T. pallidum* lacks the machinery for the *de novo* synthesis of many key nutrients, including nucleotides, riboflavin, fatty acids, and amino acids [30]. It must therefore devote considerable resources toward acquiring these nutrients from its human host. In many cases, this is accomplished via multi-protein ABC transporters. In prototypical Gram-negative bacteria, the ligand-binding protein (LBP) components of these transporters exist free in the periplasm, but in *T. pallidum*, they are often lipoproteins that are likely anchored to the outer leaflet of the cytoplasmic membrane. Characterizations of these LBPs have revealed in *T. pallidum* likely transporters for transition metals [16,19], purine nucleosides [15], methionine [17], riboflavin [20], and polyamines [18]. Many others probably are present, but verifications of their probable functions typically are impeded by the need for heterologous hyper-expression in *E. coli* and *in vitro* characterization of recombinant molecules.

All non-photosynthetic organisms must acquire energy sources from their external milieu. For *T. pallidum*, the preferred energy source likely is glucose [32–35], as this monosaccharide is more abundant in human plasma than alternative carbon sources [36], and the spirochete contains all of the necessary enzymes for glycolysis [30]. Further, the organism probably has a means for the conversion of reducing equivalents from glycolysis to electrochemical potentials that can be utilized for further ATP generation [28]. The acquisition of glucose, therefore, is of paramount importance to this human pathogen. Bioinformatic analysis of the open reading frames in the genome of *T. pallidum* reveals two possible lipoproteins that could be the ligand-binding components of a glucose ABC-type transporters. They are Tp0545 and Tp0684 [30,37]. Because of their sequence homologies to *E. coli* MglB, they have been termed TpMglB-1 and TpMglB-2, respectively. In *E. coli*, MglB has a dual function: it serves as the periplasmic ligand-binding protein of an ABC-type transporter, and it is also the chemoreceptor for galactose chemotaxis [38]. Earlier studies on the soluble portion of TpMglB-2 indicated that it binds both glucose and galactose with nearly equal facility [37].

To further elucidate the function(s) of TpMglB-2, we heterologously hyper-expressed the soluble domain of the protein and characterized it structurally and biophysically. The 2.05-Å crystal structure of TpMglB-2 revealed a fold that is similar to other glucose/galactose binding proteins. However, the topology of this fold differed from those of LBP exemplars. Bioinformatic studies showed that this topology is in putative glucose/galactose-binding proteins in other treponemes and is also present in proteins from many species of Gram-positive bacteria. This divergent topology prompted us to perform more finite scrutiny of the binding and hydrodynamic characteristics of the protein. Using isothermal titration calorimetry (ITC) and differential scanning fluorimetry (DSF), we found that the protein exhibited a marked preference for binding to D-glucose over D-galactose, and that no other assayed carbon source bound detectably to the protein. Analytical ultracentrifugation (AUC) studies were consistent with TpMglB-2 being a monomeric protein that changes its conformation upon binding its preferred substrate, D-glucose. When these studies are coupled with genetic considerations, TpMglB-2 appears to be the founding member of a new topological class of LBPs of ABC-type glucose transporters.

Results

The crystal structure of TpMglB-2

Purified, recombinant TpMglB-2 readily formed crystals that diffracted X-rays to a minimum Bragg spacing of 2.05 Å (Table 1). The determination of the structure was accomplished using single-wavelength anomalous diffraction data from a selenomethionyl derivative of the protein (see Materials & Methods). The final model after refinement featured excellent geometry, and 366 of the 387 residues present in the construct were visible in the electron-density maps. The missing residues were mostly in the N-terminal tag, with only two internal residues (L26 and T27) not visible. Note that throughout this report we utilize a numbering scheme based on the mature form of the native TpMglB-2, counting the N-terminal cysteine residue of the mature protein as residue 1.

The overall fold of TpMglB-2 (Fig 1) resembles those of LBPs (sometimes called “Periplasmic Binding Proteins”) that serve as receptors for nutrients and cofactors in bacterial ABC transporters [39,40]. LBPs generally have two structurally similar lobes that are attached by a hinge region. When ligand is absent, there usually is a large, solvent-exposed cleft between the lobes, and the protein displays an open, extended conformation. However, when the cognate ligand binds in the cleft, the two lobes clamp down on it, resulting in a more closed, compact, conformation. This mechanism has been compared to a “Venus fly trap” [41]. In a structural homolog (see below) of TpMglB-2, *E. coli* MglB (EcMglB) [42], these lobes were termed the “N” and “C” lobes because these domains harbored the N- and C-termini of the protein, respectively. Although this rationale fails for TpMglB-2 (both N- and C-termini are found in the same lobe), we choose to retain this nomenclature herein. Hence, the N lobe of TpMglB-2 contains the first visible amino-acid residue, R11, and the C lobe is located on the opposite side of the protein from this terminus (Fig 1A). There are four regions of the macromolecule that cross over between the two domains, forming the hinge region of the protein. The large cleft region is solvent-accessible, and thus this structure of TpMglB-2 appears to be in the “open” conformation.

The N lobe is dominated by a central, twisted β -sheet comprising seven β -strands. All but one of the strands are parallel. This sheet is almost completely surrounded by α -helices, 3_{10} helices, and other parts of the protein that do not adopt a regular secondary structure. The C lobe is organized similarly, but the central β -sheet has only six strands; again, one of them is antiparallel to the others. Decorating one edge of the sheet are two β -strands that have a strong curvature, giving the impression of a partial β -barrel. These strands are not counted as belonging to the central β -sheet, inasmuch as one of their surfaces is completely exposed to solvent.

Two small molecules are found in contact with the C lobe. The first is a calcium ion, which is located at the far end of the lobe, about 30 Å distant from the cleft (Fig 1A). This cation is coordinated by atoms from four side chains of TpMglB-2: D71, N73, D75, and D166 (Fig 1B). Other liganding interactions are provided by the main-chain oxygen atom of I77 and two water molecules. The Ca^{2+} appears to stabilize a loop on the C lobe while simultaneously adhering it to the main body of the domain. It thus likely plays a structural role, ensuring the proper folding and integrity of the C lobe.

The second small molecule is located in the interlobe cleft in association with residues of the C lobe; it is BIS-TRIS, which had been included in both the crystallization medium and the cryoprotection buffer at a concentration of 100 mM (see Materials & Methods). This molecule is stacked on the indole group of W145 (Fig 1C). Hydrogen bonds to hydroxyl groups on the buffer molecule are provided by the side chains of residues D91, R95, N172, and D195, and an additional van der Waals interaction is made by the side chain of H89. Although this site likely overlaps with the ligand-binding site (see below), the fact that TpMglB-2 adopts an open conformation with this small molecule bound suggests that its binding is adventitious.

Table 1. Data collection, phasing, and refinement statistics.

Data Set	TpMgIB-2	SeTpMgIB-2
PDB Accession No.	5JX2	
Data Collection		
Space Group	C222 ₁	C222 ₁
Unit Cell Dimensions (Å)		
a	77.556	76.095
b	130.025	109.927
c	88.250	89.549
α = β = γ (°)	90	90
Resolution (Å)	36.78–2.05 (2.09–2.05) ^a	46.8–2.0 ^b (2.03–2.0)
Completeness (%)	100.0 (100.0)	99.9 (99.4)
Multiplicity	5.6 (5.6)	6.4 (3.6)
Unique Reflections	28,333 (1,373)	25,584 (1,253)
R_{merge}^c	0.075 (0.604)	0.138 (0.550)
$\langle I \rangle / \sigma_I$	19.6 (2.9)	12.7 (2.6)
Wilson B (Å ²)	30.8	26.4
Phasing		
Sites Found	N/A	10
Overall Figure of Merit	N/A	0.296
Automatically Built Residues	N/A	342
Refinement		
Resolution (Å)	36.78–2.05	
No. Non-Solvent Atoms	2,791	
No. Solvent Atoms	191	
Maximum-Likelihood Coordinate Error (Å)	0.20	
Average B-factors		
Non-Solvent (Å ²)	43.95	
Solvent (Å ²)	43.21	
R-values		
R_{work}^d	0.186	
R_{free}^e	0.226	
Ramachandran Statistics		
Outliers (%)	0.3	
Most Favored Region (%)	96.7	
r.m.s. deviations		
Bonds (Å)	0.004	
Angles (°)	0.842	

^aNumbers in the parentheses are reported for the highest-resolution shell of reflections.

^bOnly data to 2.2 Å resolution were used for phasing purposes.

^c $R_{merge} = \frac{\sum_{hkl} \sum_i |I_{h,j} - \langle I_h \rangle|}{\sum_{hkl} \sum_i I_{h,j}}$ where the outer sum (*hkl*) is over the unique reflections and the inner sum (*i*) is over the set of independent observations of each unique reflection.

^d $R_{work} = \frac{\sum_{hkl} ||F_o| - |F_c||}{\sum_{hkl} |F_o|}$, where F_o and F_c are observed and calculated structure factor amplitudes, respectively.

^e R_{free} is calculated using the same formula as R_{work} , but the set *hkl* is a randomly selected subset (5%) of the total structure factors that are never used in refinement.

doi:10.1371/journal.pone.0161022.t001

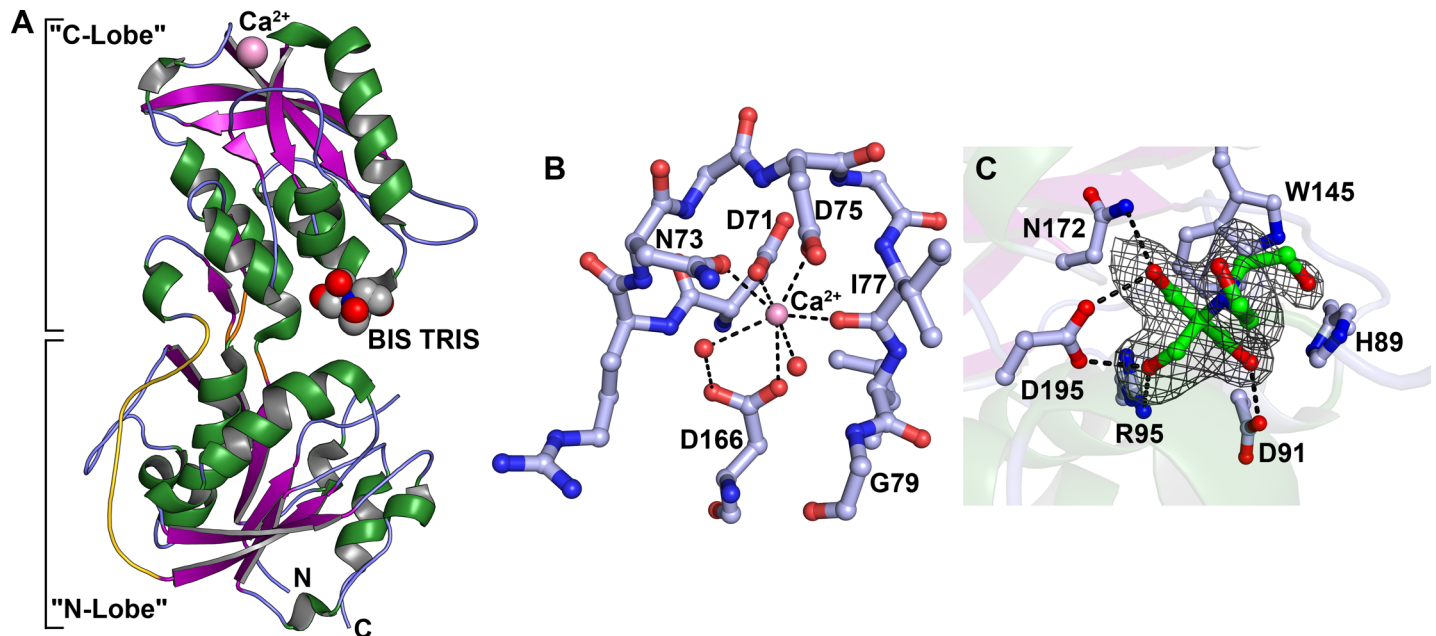


Fig 1. Structural aspects of TpMglB-2. (A) The overall structure. Shown in ribbons representation is the overall crystal structure of TpMglB-2. Helices (either α - or 3_{10}) are shown in green, β -strands are purple arrows, and regions without regular secondary structure are light blue. The three canonical hinge regions are colored orange. The adventitiously bound ligand BIS-TRIS is shown as a group of spheres between the protein's two lobes. Oxygen atoms are colored red, nitrogen atoms blue, and carbon atoms are gray. The Ca^{2+} ion bound to the C Lobe is shown as a pink sphere. The span of the protein that forms the fourth hinge region and also comprises the connecting motif not observed in EcMglB (see text) is colored gold. (B) The Ca^{2+} -binding site. Residues 71–79 are shown, with carbon atoms in light blue, and other atoms colored as described in (A). Water molecules are shown as red spheres. Inner-sphere contacts to the Ca^{2+} ion are depicted as black dashes. (C) The BIS-TRIS binding site. A kicked omit $mF_o - DF_c$ map [43,44] is shown contoured at the $3\text{-}\sigma$ level and superposed on the refined coordinates of BIS-TRIS. Hydrogen bonds between the protein and the buffer molecule are drawn as black dashes. H89 is shown because it is in van der Waals contact with the molecule. Its ϵ N atom is about 3.3 Å from the proximal hydroxyl group on the BIS-TRIS, but this distance coupled with the poor geometry make it unlikely that these two atoms are hydrogen-bonded.

doi:10.1371/journal.pone.0161022.g001

Comparisons of TpMglB-2 to other LBPs

Hidden-Markov-model-based searches [45] of the Protein Data Bank for sequences similar to that of TpMglB-2 revealed a large number of highly probable matches (over 100 matches with a greater than 99.9% probability of being a true positive). Most of these matches were LBPs from ABC transporters, and of these, most were annotated as sugar-binding proteins or putative sugar-binding proteins. Significantly, none of the matches had significant sequence homology to the final 60–100 amino acids of TpMglB-2.

Secondary-structure matching [46] was used to search the Protein Data Bank for structures similar to that of TpMglB-2. Only two matches were located with substantial coverage and root-mean-square deviations (r.m.s.d.'s) on C_α positions of less than 3 Å. These matches were two different structures of *E. coli* MglB (EcMglB), in agreement with predictions that these structures would be similar [47]. In both of the matching structures, the EcMglB was in an “open” conformation. The best match (2.2 Å r.m.s.d. over 276 aligned C_α atoms) was to an apo form of EcMglB [48]. The other matching structure (2.3 Å r.m.s.d. over 277 aligned C_α atoms) was in an open conformation because a non-native ligand, 3-O-methyl glucose (3-OMe Glc), was bound to D-glucose-binding residues in the C lobe, apparently sterically precluding closure of the cleft [49]. Strikingly, when TpMglB-2 and 3-OMe-Glc-bound EcMglB were superposed using the proteins only, the two non-native ligands were in similar positions (Fig 2). Because the 3-OMe Glc and native glucose-binding sites overlapped, this implies that BIS-TRIS occupies the ligand-binding site of TpMglB-2 in the present crystal structure (Fig 1). Furthermore,

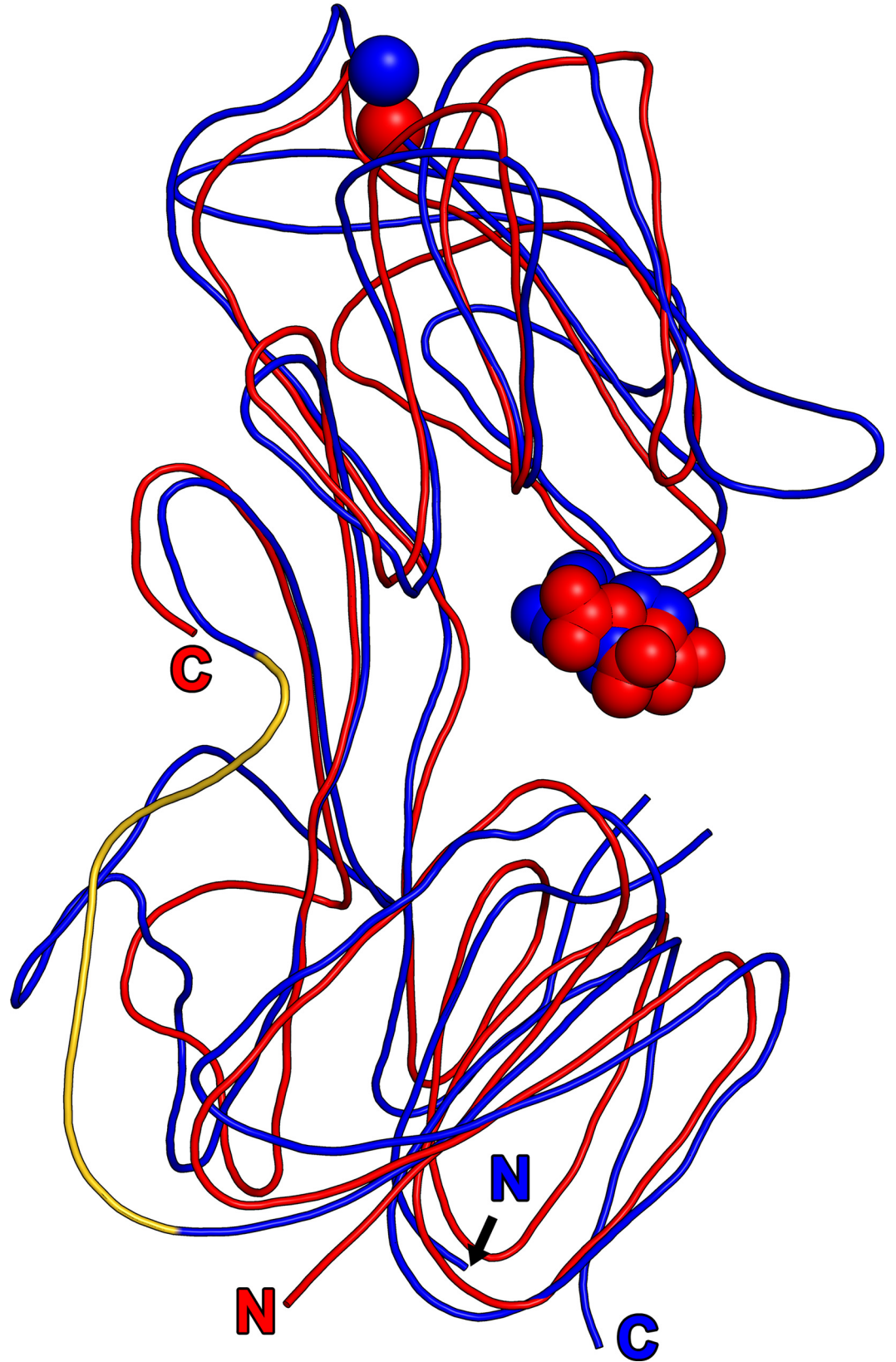


Fig 2. Superposition of TpMglB-2 and EcMglB. For both structures, a smoothed trace through the main chain is drawn. TpMglB-2 is shown in blue, while EcMglB is red. The bound ligands are depicted as spheres and are colored according to their respective protein chains, and the respective N- and C-termini are noted, with the letters color-coded to the respective structures. The gold coloration for the fourth hinge/connection motif from Fig 1 is retained here.

doi:10.1371/journal.pone.0161022.g002

this alignment allowed us to confirm that nearly all of the glucose-binding residues in EcMglB were identical in TpMglB-2 (Table 2). The only changes in the identities of these residues were conservative: D14 and N256 in EcMglB were analogous to N311 and Q215, respectively, in TpMglB2.

Notably, both proteins also possess Ca²⁺-binding sites in analogous positions (Fig 2). In EcMglB, the similarity of the arrangement and identity of amino-acid side chains liganding the calcium has been likened to the EF-loop of EF-hand proteins [42]. This likeness is generally preserved in TpMglB-2, with some differences. In EcMglB, the calcium cation is liganded by seven oxygen atoms from the protein, but in TpMglB-2, only five of the seven ligands emanate from the protein (Fig 1B). The remaining two coordinating atoms are oxygen atoms from water molecules. One of these differences arises from the fact that the coordinating side chain Q142 in EcMglB is not present in TpMglB-2; its equivalent is the non-liganding G79. Another difference is caused by the fact that D166 in TpMglB-2 does not make a bidentate contact to the Ca²⁺, as its EcMglB analog, E205, does. The consequences of these deviations from the usual EF-loop are unclear; no function (other than structural stabilization) has been posited for the Ca²⁺-binding site in EcMglB, as metal binding does not affect sugar binding [50].

Comparisons of the structures of TpMglB-2 and EcMglB also reveal a stark difference in the proteins: their topologies are circularly permuted. This is conveniently illustrated via a comparison of the topologies of the central β-sheets of respective N lobes (Fig 3). If the strands of these sheets are renumbered according to their order in the respective proteins' sequences, the order of the strands (from left to right in Fig 3) in EcMglB is 2-1-3-4-5-6. In two related classification systems for LBPs, the ordering of the first five strands is typical for a Type I [39] or "Cluster B" [40] LBP. However, the strand ordering in TpMglB-2 is 5-4-6-1-2-3. Hence, the first three strands in EcMglB are equivalent to the last three in TpMglB-2. Viewing this arrangement more globally, the first three strands in this lobe in EcMglB are near to the protein's N-terminus (residues 3–9, 34–40, and 62–65), whereas the equivalent strands are close to the C-terminus of TpMglB-2 (residues 300–307, 331–337, and 358–361). The orders of the strands for the respective C lobes are unaffected by the permutation, and they adhere to the canonical 2-1-3-4-5 arrangement for Type I/Cluster B LBPs.

Table 2. Glucose-binding residues from EcMglB and their equivalents in TpMglB-2 and TpMglB-1.

EcMglB	TpMglB-2	TpMglB-1
D14	N311	N18
F16	F313	L20
N91	N19	N93
H152	H89	H154
D154	D91	S156
R158	R95	C160
W183	W145	D185
N211	N172	N212
D236	D195	D241
N256	Q215	N261

doi:10.1371/journal.pone.0161022.t002

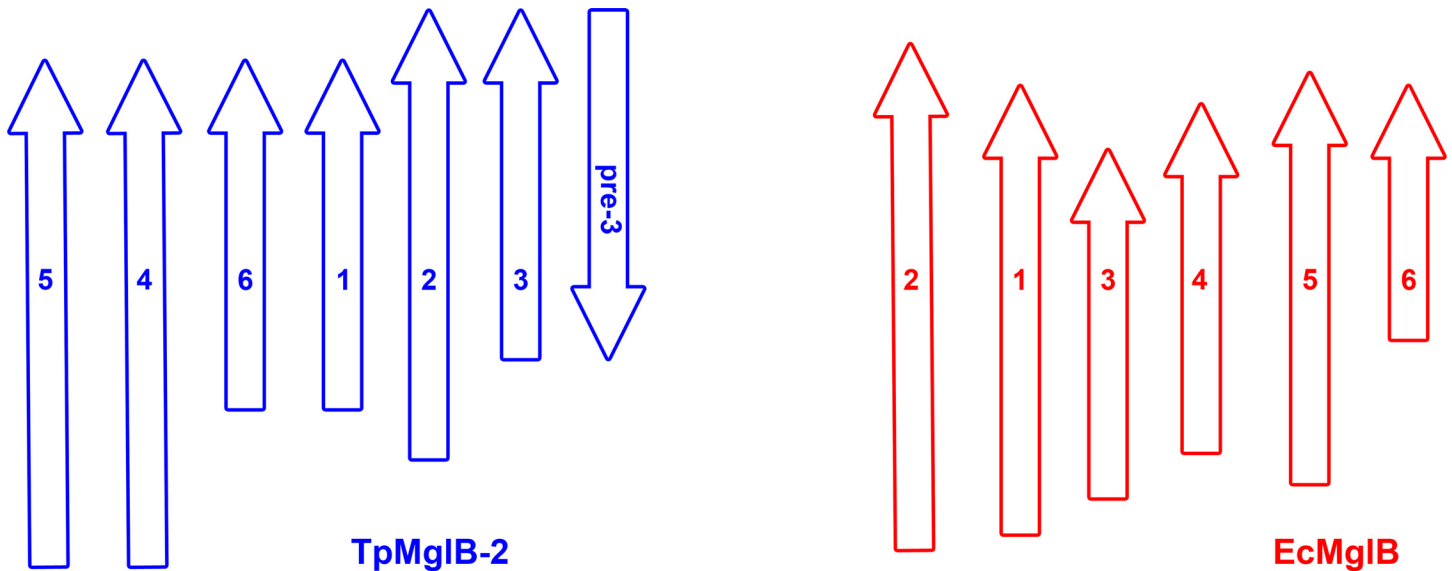


Fig 3. Comparison of the topologies of the central β -sheets in the N-lobes of TpMgIB-2 and EcMgIB. The β -strands for the respective sheets are drawn as arrows, with those from TpMgIB-2 shown in blue and those from EcMgIB in red. The strands are numbered according to the orders in which they appear in the primary structures of their respective proteins. The positions and sizes of the strands are approximately to scale. The antiparallel strand in TpMgIB-2 was not assigned its own number in this scheme because it does not correspond to any strand in EcMgIB. Its designation, “pre-3”, indicates that it immediately precedes strand 3 in the primary structure of TpMgIB-2.

doi:10.1371/journal.pone.0161022.g003

This reordering of the N lobes has additional consequences in the hinge region. The hinges of Type I/Cluster B LBPs are usually made up of three regions of irregular secondary structure representing points where the primary structure crosses over from one lobe to the other. The permutation noted above in TpMgIB-2 necessitates a fourth cross-over; this region of the protein (residues 284–288) is close to the canonical hinge. It leads to an extended structure that spans from this fourth hinge region to the beginning of the permuted strand 4 described above. This fourth hinge and the following extended structure (colored gold in Figs 1 & 2) are not present in EcMgIB.

The discovery of this permutation explains why the sequence searches could not match the final 70–100 residues of TpMgIB-2. That is, this sequence should not be homologous to the C-termini of MgIB homologues. If there is any sequence homology to the C-terminal portion of TpMgIB-2, it should be to the N-termini of EcMgIB and its homologues. An attempt to align the last 100 residues of TpMgIB-2 and to the first 100 residues of EcMgIB yields a poor match with only 13% sequence identity. For comparison, the last 230 amino acids of EcMgIB can be aligned with the first 280 residues of TpMgIB-2, featuring 26% identities.

These structural differences between TpMgIB-2 and EcMgIB prompted us to search sequence databases for proteins with both homologous sequences and similar topologies to that of the treponemal protein. This was accomplished using BLAST [51] and setting as a criterion that the coverage of the match be > 90% and that the N-termini matched well. A large number of matches were discovered. The top 100 matches were pooled, and, after winnowing identical or very similar sequences, a phylogeny was constructed (Fig 4; we included the sequence of EcMgIB and that of the glucose-binding protein from *Salmonella typhimurium* as an outgroup).

The proteins with similar topologies to TpMgIB-2 fall into two main groups: those from organisms of the phylum *Spirochaetes* and those largely from Gram-positive bacteria. While the former are mostly from treponemes, the latter are a mixture of mammalian pathogens and

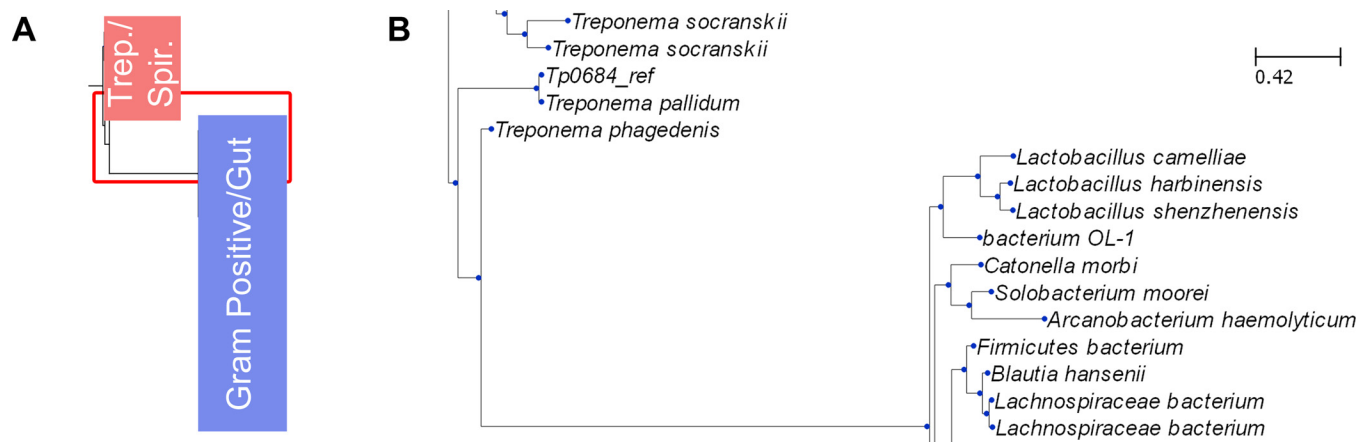


Fig 4. Relationships between proteins containing sequence and topology homology to TpMglB-2. (A) A schematic representation of the full cladogram. There are two main clades, which are shown as colored rectangles. The pink “Trep./Spir.” rectangle is occupied by treponemal and spirochetal organisms, while the “Gram Positive/Gut” blue rectangle is inhabited mainly by gram-positive organisms, many of which are mammalian gut commensal organisms. The red rectangle details the area of the figure that is blown up in (B). The full cladogram is available as a text file in the data supplied in [S1 File](#). (B) A close-up of a branch point in the cladogram. The area of the cladogram highlighted in red in (A) is shown. TpMglB-2 is represented twice here (“Tp0684_ref” and “*T. pallidum*”) because it was included both as a structural template (in.pdb form) and as its native sequence. There is a slight difference between the two because of the 2 missing amino acids in the.pdb file.

doi:10.1371/journal.pone.0161022.g004

commensal bacteria. We also searched for sequence homologues to the gene product of *tp0686*, which is the putative permease for the TpMglB-2’s cognate transporter. The distribution of homologs was similar to that of TpMglB-2, with close homologs appearing in other treponemes and slightly more distant ones found in Gram-positive bacteria. The distribution of these proteins in a disparate group of human parasites, spirochetes, and commensal organisms suggests the possibility that these genes have been shared horizontally among these organisms or their ancestors.

The evident differences between EcMglB and TpMglB-2 raise the question of whether their similarities are the result of convergent, rather than divergent, evolution. The preponderance of the evidence suggests that EcMglB and TpMglB-2 diverged from a common ancestor. The strong sequence conservation in the common parts of the proteins, including the conservation of a Ca²⁺-binding site, serves as ample evidence for this assertion. However, the poor sequence conservation between the permuted portions of the two proteins (and remembering that this part of TpMglB-2 includes a large insertion) may imply that TpMglB-2 and its close homologs acquired this part from a different gene, with evolutionary pressure influencing convergence of structural features.

Functional Studies of TpMglB-2

The unconventional structural organization and the altered Ca²⁺-binding site of TpMglB-2 provoked a question: does such a structurally rearranged LBP function similarly to other LBPs? To address this question, we first examined ligand binding. In a previous study on TpMglB-2, we reported that the protein preferred D-glucose and D-galactose over a small number of other monosaccharide carbon sources [37]. However, we wished to interrogate a larger panel of such sources to investigate whether another ligand would bind to the protein with even higher affinity. We thus employed differential scanning fluorimetry (DSF) as a means to probe for ligand binding [52,53]. This method examines the apparent melting temperature ($T_{m, app}$) of the protein; the “apparent” nomenclature is added in recognition of the fact that the reversibility of the transition was not tested in these experiments. Presumably, in the presence of a specifically

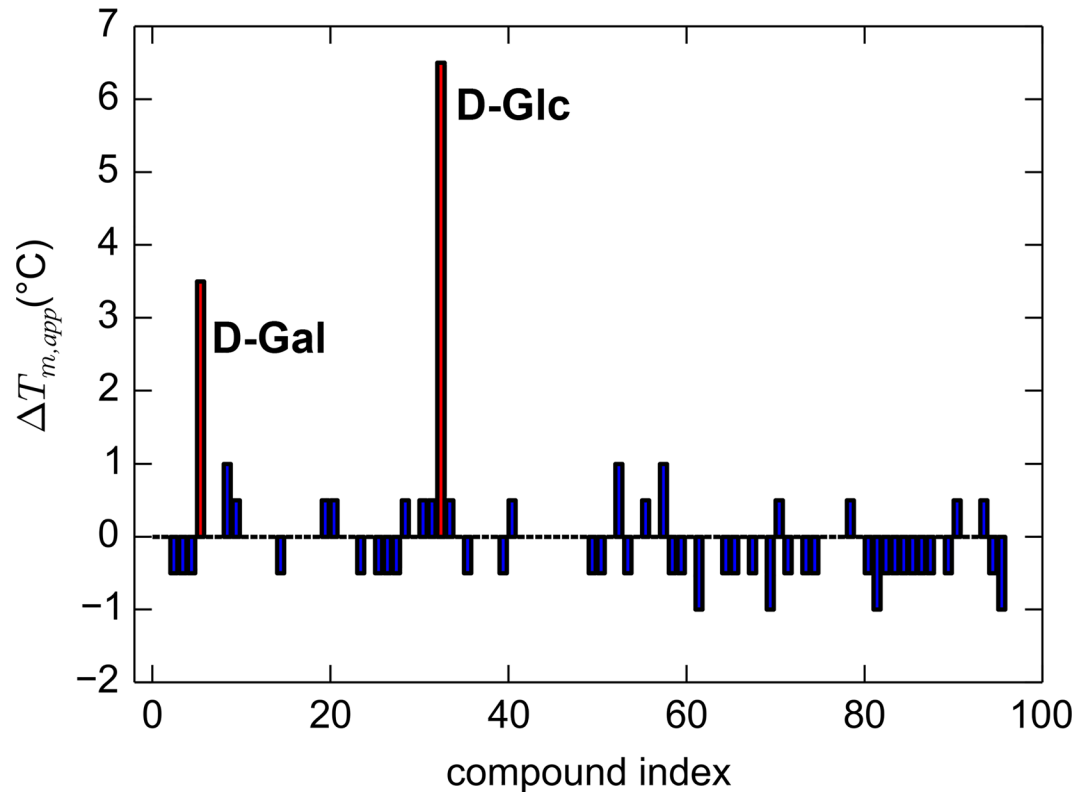


Fig 5. DSF data for a panel of carbon sources in the presence of TpMglB-2. The change in the apparent melting temperature is plotted vs. compound index (ranging from 1 to 96; only results from BIOLOG plate PM1 are shown here). Bars for compounds that caused $T_{m,app}$ shifts of more than 1.5° C from that observed in the presence of the negative control (water, no compound) are colored red and labeled with the compound's identity.

doi:10.1371/journal.pone.0161022.g005

bound ligand, the protein's structure will be stabilized, resulting in a higher $T_{m,app}$. We used this plate-based method to qualitatively assay for the binding of 190 carbon sources. Only D-glucose and D-galactose caused significant shifts ($> 1.5^\circ\text{C}$) in $T_{m,app}$ (Fig 5). Thus, these were the only candidate compounds that were selected for additional study.

ITC was used to quantify the interactions between TpMglB-2 and the subject monosaccharides. Indeed, both of these sugars elicited robust heat signals when titrated into solutions of the protein (Fig 6). D-Mannose and D-ribose, which both apparently failed to bind in the DSF assay, were titrated in control experiments. Corroborating the DSF experiments, they failed to exhibit any binding heats.

Analysis of the binding isotherms resulting from these experiments shows that TpMglB-2 has a pronounced preference for D-glucose over D-galactose in this assay: the measured K_D for D-galactose was 30-fold higher than that of D-glucose (Fig 6; Table 3). Although the data were analyzed using a 1:1 binding model, it was always necessary to refine a parameter [55,56] that compensated for an apparent incompetent fraction of the protein (15–30%). The modest K_D for D-glucose binding to TpMglB-2 (1.1 μM) should not hamper import of the hexose, as its extracellular concentration in *T. pallidum*'s environments is expected to be in the millimolar range [36].

These results contrast with those that we obtained earlier using a similar TpMglB-2 construct [37]. In those experiments, TpMglB-2 bound to both D-glucose and D-galactose with similar, sub-micromolar equilibrium dissociation constants (150 nM for D-glucose, 250 nM for D-galactose). Although the exact reason for this discrepancy is unknown, it is notable that

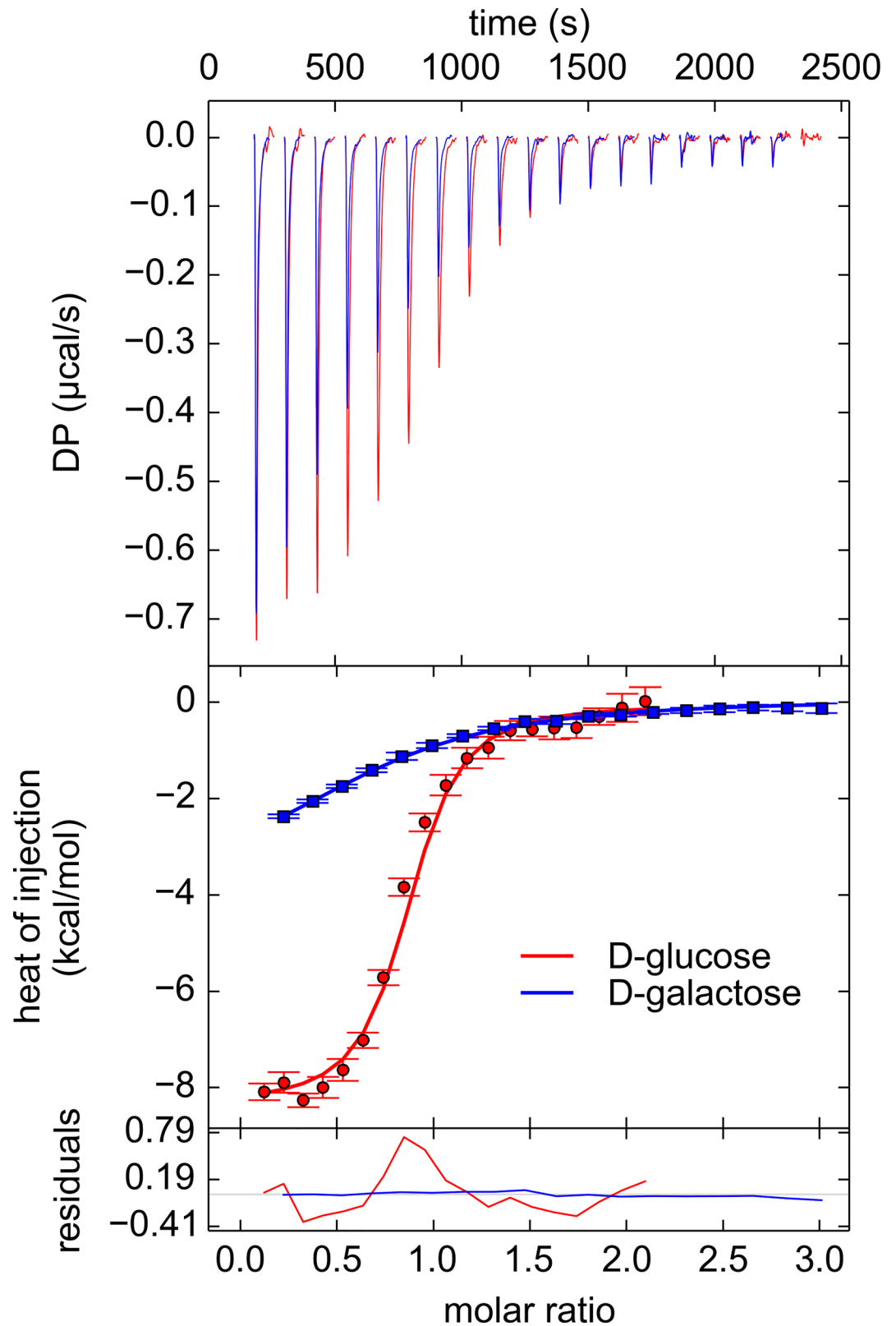


Fig 6. D-glucose and D-galactose binding TpMgIB-2. Baseline-subtracted thermograms reconstructed from the raw data using singular-value decomposition are shown in the top panel. In the middle panel are the integrated heats of injection (markers) with error bars depicted estimated measurement errors [54,55]. The lines are fits to

these data points. The bottom panel shows the residuals between the data and the fit lines. All lines and markers are colored respectively according to the inset legend. Only one of the three analyzed experiments is shown for each monosaccharide.

doi:10.1371/journal.pone.0161022.g006

the present experiments were carried out with fivefold higher NaCl concentrations (20 mM vs. 100 mM). The ionic strength of the present experiments is likely closer to that of the *T. pallidum* periplasm, rendering the measurements presented here more physiologically relevant.

To determine whether the putative binding site—identified by the binding site of BIS-TRIS (Figs 1 & 2) and by homology to the EcMglB binding site (Table 2)—is responsible for binding D-glucose in TpMglB-2, we engineered single-amino-acid changes into the protein and assessed the affinity of the mutant proteins for the hexose (Table 3). None of the mutant proteins (except W145A) displayed an interpretable heat signal for D-glucose binding, strongly suggesting that their D-glucose-binding functions were severely diminished. Analytical ultracentrifugation (AUC) in the sedimentation velocity (SV) mode was performed on the wild-type protein and all of the mutated proteins. All had identical hydrodynamic radii, suggesting that the mutations did not cause large-scale changes in the overall structures of the proteins. The one mutant protein that did display binding heats, W145A, had a tenfold higher K_D for glucose binding than the wild-type protein (Table 3; Fig 7). From these data, we conclude that binding site identified above is very likely to be correct.

The hydrodynamic characteristics of TpMglB-2 gleaned from SV were further scrutinized with respect to the protein’s oligomeric state. We found that this protein was monomeric under our experimental conditions at all concentrations studied (0.6–5 mg/mL) and in the presence and absence of D-glucose. While these data comport with data from other glucose-binding LBPs [49,57], they contrast with reports on other LBPs that can form dimers [58,59].

We also wished to discover whether the topologically distinct TpBglB-2 undergoes a conformational change upon ligand binding; such conformational changes have been observed in other D-glucose-binding LBPs in the crystalline state [48] and in solution [49,57], consistent with the “Venus fly trap” model. All attempts to co-crystallize D-glucose with TpMglB-2 or to introduce it into the crystals described above failed; we do not believe that this failure is due to the circular permutation in the protein disrupting its ability to co-crystallize. Nonetheless, we again chose to characterize the hydrodynamic traits of the protein in the presence and absence of D-glucose. The “closed” forms of LBPs should exhibit a smaller hydrodynamic radius (i.e. larger sedimentation coefficient) because of their more compact conformation. Careful hydrodynamic modeling of liganded (2FVY) and unliganded (2FW0) crystal structures indicated that the TpMglB-2 structural homolog, EcMglB, confirmed this assertion; the protein would evince a 0.080-S higher sedimentation coefficient upon ligand binding. While this change

Table 3. Dissociation constants of the TpMglB-2/sugar interactions.

TpMglB-2 Variant	Sugar	K_D (μ M)
wild-type	D-glucose	1.1 [0.9, 1.4] ^a
wild-type	D-galactose	33 [21, 58]
N19A	D-glucose	NH ^b
W145A	D-glucose	12.7 [10.6, 15.5]
D195A	D-glucose	NH
F313A	D-glucose	NH

^a68.3% confidence intervals are given in brackets.

^bNH: No observable heat of ligand binding.

doi:10.1371/journal.pone.0161022.t003

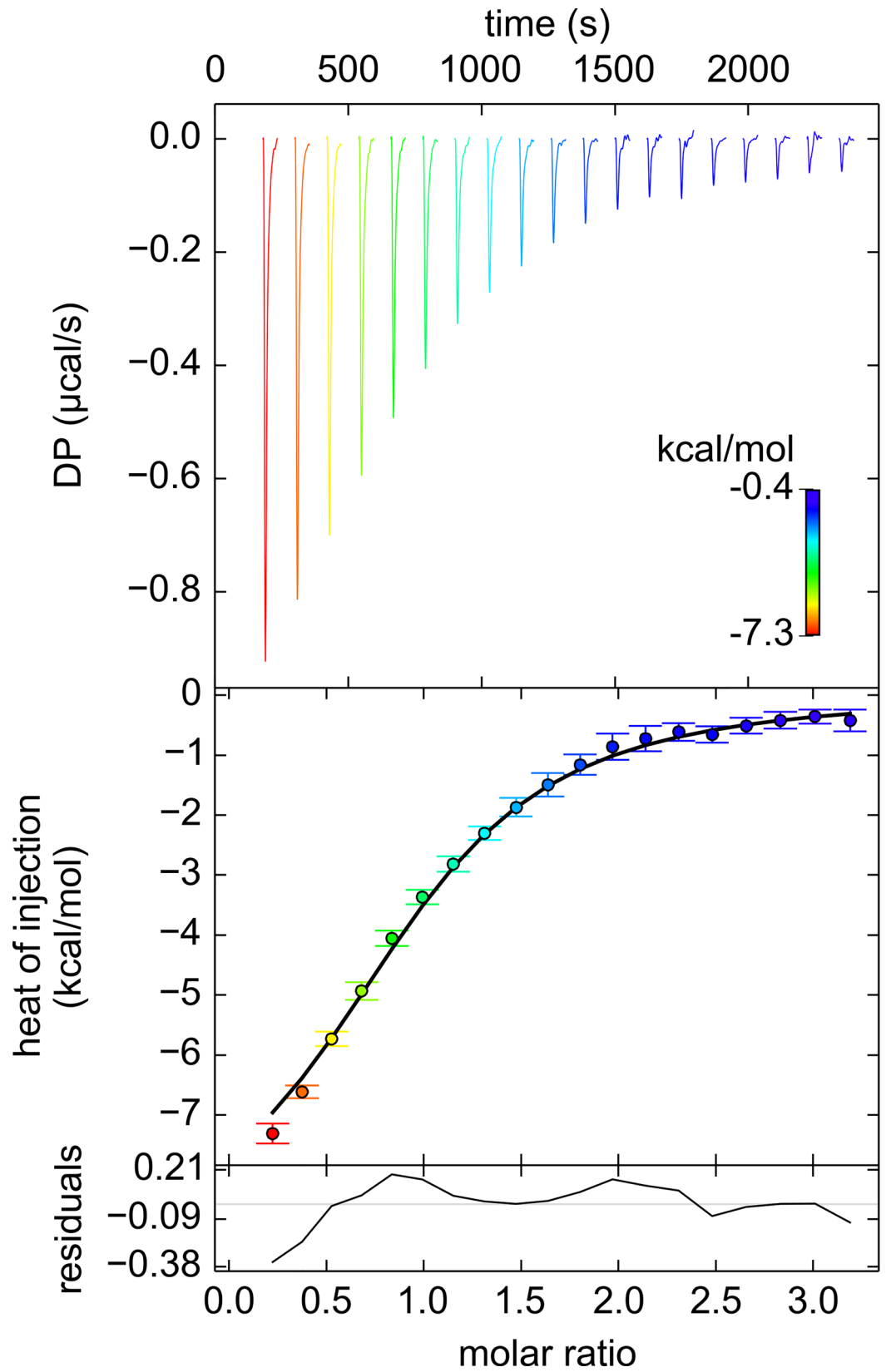


Fig 7. D-glucose binding to the W145A variant of TpMglB-2. The organization of the figure is the same as in Fig 6, but the thermogram peaks and their respective integrated data points are colored according to the color map shown inset in the top panel.

doi:10.1371/journal.pone.0161022.g007

seems quite small (i.e. a 2.5% change in s -value), discerning such differences is well within the capabilities of modern instrumentation and analytical methods [60,61].

Our initial attempt to examine whether this small difference could be detected was conducted using conventional SV. In experiments carried out simultaneously, we sedimented apo-TpMglB-2 and also the protein in the presence of 100 μ M D-glucose. By analyzing the data with a non-interacting discrete species model [62], we could perform a careful analysis using F-statistics [63] that would establish realistic 95% confidence intervals (shown below in brackets) on the refined parameter s . For apo-TpMglB-2, the refined experimental s -value was 3.212 [3.204–3.219] S. The same value of the protein in the presence of D-glucose was 3.292 [3.288–3.297] S. We can estimate the statistical significance of this difference the F-statistic formalism. If we fix the s -value of the apo-protein to that of the liganded protein and repeat the analysis, we arrive at a χ^2 ratio between the two analyses of 1.86. Given the degrees of freedom in this experiment, it is extremely improbable (probability $\approx 10^{-16}$) that there is no difference in the s -values.

Although the previous SV study strongly suggested a conformational change, we sought a more direct comparison of the two sedimentation rates. Minute sedimentation-coefficient differences are readily detected using a method pioneered in H. Schachman's lab called "difference sedimentation velocity" [64]. This method provides a direct, side-by-side comparison of the sedimentation velocity of two macromolecular solutes. In brief, two samples of 5 mg/mL TpMglB-2 were prepared. To one sample, a non-binding sugar, either D-mannose or D-ribose (these sugars evidenced no binding in neither the DSF nor ITC assays), was introduced, whereas an identical concentration of D-glucose was added to the other. The protein sample with the non-binding sugar was placed in the reference sector of a centrifugation centerpiece, and the sample sector of the same centerpiece was filled with the D-glucose-containing protein solution. Centrifugation was initiated, and Rayleigh laser interferometry was used to monitor the sedimentation. Because this mode of data collection reports on refractive-index differences between the two sectors, the difference between the sedimentation in the two sectors is recorded. This signal (for unequal solution-column heights) will appear roughly as a Gaussian-shaped curve inverted about the radial axis (Fig 8A). If the macromolecular solutes in each sector sediment at different rates, the absolute value of the area of (or, more specifically, the normalized first moment of) this curve will increase as sedimentation continues. This trend reports on the fractional sedimentation-coefficient difference between the solutes in the reference and sample sectors, termed $\Delta s/\bar{s}$, i.e. the change in s normalized by the mean s -value. Thus, when a normalized first moment of the interference signal is plotted as shown in Fig 8B, the slope of the fitted line is taken as $\Delta s/\bar{s}$. A nonzero slope is considered to be strong evidence of a sedimentation coefficient difference between the solutes in the two sectors.

As shown in Fig 8B, we observed a $\Delta s/\bar{s}$ of about 0.0177 [0.0163, 0.0191], or about 1.8% (for $\Delta s/\bar{s}$ values, we present the 95% confidence interval for the fitting in brackets) when comparing the sedimentation of TpMglB-2 in the presence of D-mannose/D-glucose. When repeating the experiment using the D-ribose/D-glucose pair, the resulting $\Delta s/\bar{s}$ was nearly identical (0.0180 [0.0163, 0.0197]). Multiplying $\Delta s/\bar{s}$ by the observed s for this same concentration of TpMglB-2, 3.2 S, results in an observed change of approximately 0.06 S. This correlates well with the difference obtained from comparing the s -value of individual experiments and also with the hydrodynamic modeling performed on the structural homolog EcMglB (see above).

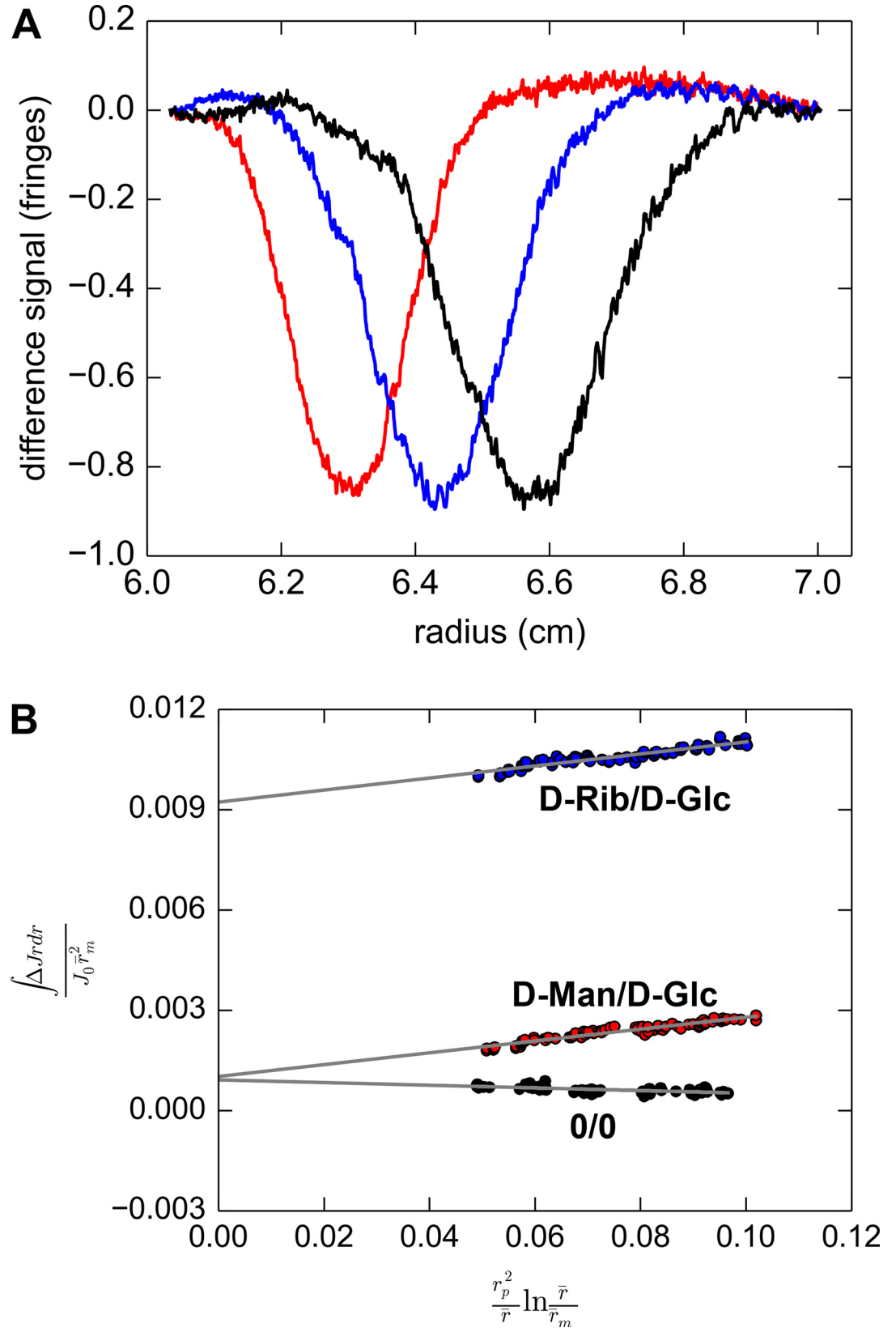


Fig 8. Difference sedimentation velocity of Tp0684 in the presence and absence of sugars. (A) Example difference SV traces. Three traces are shown after noise and baseline subtraction. The traces are colored according to the time of the scan, with red, blue, and black representing 92, 132, and 172 min after the start of centrifugation. The three traces are from the D-mannose/D-glucose experiment. (B) Three difference sedimentation velocity analysis plots. Data similar to those shown in (A) were treated as described in Materials and Methods; each trace resulted in one of the data points shown here, which are colored according to their respective experiments. Each experiment is labeled; “D-Man/D-Glc” stands for D-mannose vs. D-glucose, “D-Rib/D-Glc” is the D-ribose vs. D-glucose experiment, and “0/0” is the negative control in which no sugar was added to either sector. The gray lines show the linear fits to each data set. Eighty-six traces were analyzed per experiment. The D-Rib/D-Glc data are displaced upward because of the larger difference in the reference/sample sector menisci in this experiment (see [Materials & Methods](#)).

doi:10.1371/journal.pone.0161022.g008

An important control experiment in difference SV is to conduct the study with identically prepared solutions in each sector. When applied to TpMglB-2, this approach resulted in a weak negative trend in the first moments, eliciting a $\Delta s/\bar{s}$ of -0.0040 [-0.0052, -0.0028]. The fact that this is nonzero may indicate that this is the inherent error in the experiment or that there is a systematic -0.4% offset in the values obtained using our instrument. A systematic study of the errors associated with this method is required to distinguish between these possibilities.

Another critical consideration in difference SV experiments is the $\Delta s/\bar{s}$ expected solely from the gain in mass resulting from ligand binding. We calculated this (see [Materials & Methods](#)) to be 0.006, or about 0.6%. This value is well outside the 95% confidence intervals for our glucose-containing experiments. This result, coupled with the close correlation between the solution experiments and hydrodynamic modeling, allowed us to conclude that it is very likely that a conformational change occurs when D-glucose binds to apo-TpMglB-2.

Discussion

T. pallidum is thought to be microaerophilic [65], and there is no evidence that the electron-transport chain that would allow canonical aerobic respiration exists in the organism [30]. Further, D-glucose is believed to be the main, if not sole, carbon source for the microorganism [32–35]. Thus, a modified glycolytic pathway could be the only means for energy generation in *T. pallidum*. These facts make the study of the cellular import of D-glucose critical for understanding the lifestyle of the organism. In this report, we found that TpMglB-2, the product of gene *tp0684*, is very likely the D-glucose-binding protein of an ABC transporter dedicated to the import of this vital molecule. This conclusion identifies this protein as a crucial cog in the machinery of energy acquisition for *T. pallidum*. TpMglB-2 has an unconventional topology for an LBP (Fig 3), but retains the hallmarks of ligand binding and ligand-induced conformational change that are common to this family of proteins (Figs 5–8). Further, its marked preference for D-glucose over D-galactose is consistent with the likely paucity of the latter sugar in native (human) *T. pallidum* environments.

The unusual topology (Figs 1–3) of TpMglB-2 is present in putative sugar-binding proteins in other bacteria (Fig 4). The structure detailed herein thus serves as the founding member of a family of proteins related to the “Type I” [39] or “Cluster B” [40] LBPs. We propose that these proteins belong to a sub-class/cluster of these classification systems, given the strong structural homology to canonical members. We suggest that they TpMglB-2-like proteins be referred to as “Type I_c” or “Cluster B_c”, with the subscripted “c” appended to note the circular permutation of this subclass.

TpMglB-2 homologs (in both sequence and topology) are found mostly in other treponemes or Gram-positive bacteria. A noteworthy commonality among most of these organisms is that their respective LBPs are probably tethered. In the treponemes, these proteins are membrane-

tethered in the periplasmic space, and they are probably attached to the outer cell wall in Gram-positive bacteria. These facts raise the intriguing possibility that this divergent topology arose as an adaptation to the necessity of anchoring the proteins to a membrane. No physical or mechanical advantage to this topology is immediately obvious from the structure, but we do note that the N-terminus of EcMglB is near to two sugar-binding amino acids (D14 and F16), whereas the topologically distinct N-terminus of TpMglB-2 is near to only one (N19). Therefore, strain on the N-terminus induced by tethering and Brownian motion could have a more adverse effect on the binding of D-glucose in the *E. coli* protein. Another possibility is that the different topology better satisfies the requirements of the export/lipidation machinery in these organisms.

An important question to resolve regarding the biology of *T. pallidum* is whether TpMglB-2 and its cognate ABC permease/ATPase represent the organism's only means of glucose uptake. Another protein, the product of *tp0545*, also has sequence homology to EcMglB and has been termed "TpMglB-1" [30,37]. The *tp0545* gene does not appear to be in an operon with other ABC transporter components. Hidden-Markov-model searches [45] of the Protein Data Bank reveal a high probability that TpMglB-1 is structurally similar to EcMglB. Efforts to produce a recombinant form of TpMglB-1 to test its sugar-binding capabilities have thus far failed. However, a sequence alignment of this protein with EcMglB and TpMglB-2 demonstrates that certain sugar-binding residues are not conserved (Table 2). This result casts doubt on the ability of TpMglB-1 to serve as a D-glucose-binding protein. Thus, TpMglB-1 may resemble EcMglB only insofar as its structural homology to that protein.

As detailed elsewhere [38], the *E. coli* structural homolog of TpMglB-2, EcMglB, is a multifunctional protein. In addition to its role as the LBP for a D-glucose/D-galactose ABC transporter, it is also the chemoreceptor for a well-studied sugar chemotaxis system. We and others [37,66,67] have noted that all of the essential elements of a homologous chemotaxis apparatus are apparently present in *T. pallidum*. Indeed, motility and chemotaxis are likely major mechanisms for this spirochete to invade tissues distal from the site of initial infection [68,69], thus clinically manifesting as secondary and tertiary syphilis. If TpMglB-2 is also a chemoreceptor for a D-glucose-tactic system, its proper functioning has implications not only for spirochetal survival but also for pathogenicity.

Conclusions

TpMglB-2 adopts a fold that is common for periplasmic LBPs; however, it displays a hitherto unknown topology for this class of proteins. Careful investigation of the binding and physical characteristics of the protein are consistent with the notion that this protein is acting as the ligand-binding element of an ABC transporter for D-glucose. This assertion is buttressed by noting that the gene for this protein, *tp0684*, is clustered with probable transmembrane and ATP-binding elements of the cognate ABC transporter. As no other carbon source has been reliably identified that supports the growth and chemotaxis of *T. pallidum*, this transporter has implications for the survival and pathogenicity of this bacterium.

Materials & Methods

Cloning, overexpression, and protein preparation

To produce a non-lipidated, recombinant derivative of TP0684 in *E. coli*, the DNA fragment encoding amino acid residues 11–378 (cloned without the post-translationally modified N-terminal Cys plus nine other residues comprising the predicted N-terminal transmembrane helix) of TP0684 was PCR amplified from *T. pallidum* genomic DNA by the polymerase incomplete primer extension (PIPE) cloning method using ends-specific primers (PIPE insert). The

expression vector, pSpeedET (DNASU, Tempe, AZ), which encodes an N-terminal TEV-protease cleavable expression and purification hexa-histidine tag (MGSDKIHSHHHHHENLYFQG), was PCR amplified with PIPE-vector primers. The PIPE-insert and PIPE-vector was mixed to anneal the amplified DNA fragments together [70]. *E. coli* HK100 competent cells were transformed with the mixtures (PIPE-vector and insert) and selected for kanamycin resistance on LB agar plates. Cloning junctions/fragments were verified by DNA sequencing. A verified plasmid was then co-transformed with pGroESL (Takara, Shiga, Japan) into *E. coli* BL21 AI (Invitrogen) cells for soluble protein expression. *E. coli* BL21 AI cells were grown at 37° C in LB medium containing 0.1% (w/v) glucose, 40 µg/mL of kanamycin and 30 µg/mL of chloramphenicol until the cell density reached an A600 of 0.5. The culture was then induced for overnight at 16° C with 0.2% (w/v) L-arabinose. The procedures for expression and purification of the recombinant proteins were essentially as previously described [22].

For the production of selenomethionine labeled protein, *tp0684* was recloned into a pProEx HTb vector (Invitrogen) and co-transformed with pGroESL (Takara) into a methionine auxotroph *E. coli* B834 (DE3). The recombinant protein was overproduced and purified as described previously [24].

Site-directed mutagenesis and protein concentration determination

For the construction of the recombinant TpMglB-2 variants, the N19A, W145A, D195A and F313A mutation was individually introduced into the plasmid carrying the wild-type *tp0684* sequence using the QuikChange site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA). The mutation was confirmed by DNA sequencing. The mutant protein was expressed and purified as described above. Protein concentrations were determined in buffer A (20 mM Hepes, 0.1 M NaCl, pH 7.5, 2 mM n-Octyl-β-D-glucopyranoside) using spectrophotometry. Extinction coefficients were calculated using the Protparam tool of ExPASy (www.expasy.org).

Crystallization and cryoprotection

Crystals of TpMglB-2 were obtained by mixing 4 µL of TP0684 (~22 mg/mL in buffer A) with 4 µL of crystallization buffer (0.1 M BIS-TRIS, pH 6.5, 20% (w/v) PEG MME 5,000) and incubating them over 0.5 mL of the reservoir (containing crystallization buffer) for 10 days. The crystals were transferred to the stabilization buffer (SB: 0.1 M BIS-TRIS, pH 6.5, 100 mM NaCl, 20% (w/v) PEG MME 5,000, 5% (v/v) ethylene glycol). After about 5 minutes in SB, they were serially transferred to buffers that were the same as SB except that they had higher concentrations of ethylene glycol. The final concentration of ethylene glycol was 25%. After about 1 min in this solution, the crystals were flash-cooled in liquid nitrogen. Toward solving the phase problem, crystals of a selenomethionyl derivative of TpMglB-2 (SeTpMglB-2) were grown by mixing 4 µL of SeTpMglB-2 (~9 mg/mL in buffer A) with 4 µL of crystallization buffer (0.1 M BIS-TRIS, pH 5.5, 0.1 M ammonium acetate, 17% (w/v) PEG 10,000) and incubating them over 0.5 mL of reservoir solution containing the crystallization buffer for 7 days. The crystals were transferred to the stabilization buffer 2 (SB2: 0.1 M BIS-TRIS, pH 5.5, 0.1 M ammonium acetate, 17% (w/v) PEG 10,000, 5% (v/v) ethylene glycol) and cryoprotected/flash cooled as above.

Data collection, structure determination and refinement

All X-ray diffraction data were acquired at beamline 19-ID of the Structural Biology Center at Argonne National Laboratories. Native TpMglB-2 crystals had the symmetry of space group C222₁ and diffracted X-rays with a d_{\min} spacing of 2.05 Å. The data were reduced and scaled

using HKL3000 [71]. Phase determination was undertaken using single-wavelength anomalous diffraction from crystals of a SeTpMglB-2 (see above). These crystals had the same symmetry as the native crystals and diffracted with a similar d_{\min} spacing, but exhibited significantly divergent unit-cell constants. Diffraction data from the selenomethionine-derivatized crystals were reduced and scaled using HKL3000. The GUI interface in that program was utilized to perform a substructure solution (using SHELXD [72]), site refinement (using MLPHARE [73]), density modification (using DM [74]; no non-crystallographic symmetry is present), and automated model building (using Arp [75]). This strategy resulted in a model that was 93% complete. This model was used as a molecular replacement model to determine phases for the native structure using Phaser [76]. The model was completed and corrected in Coot [77]. After an initial round of simulated annealing, the model was refined in PHENIX [43] using positional and TLS refinement. Riding hydrogen atoms were included in the model. The weights between the chemical and X-ray terms were refined to optimize R_{free} . The final model had excellent statistics and geometry (Table 1), and it featured one molecule of BIS-TRIS and one Ca^{2+} cation. The model and structure factors for TpMglB-2 have been deposited in the Protein Data Bank with accession number 5JX2. Structure figures were rendered using PyMol (Schrödinger LLC).

Isothermal titration calorimetry

TpMglB-2 was dialyzed exhaustively against Assay Buffer (10 mM sodium phosphate, pH 7.4, 100 mM NaCl). The sugars to be titrated were dissolved in the same buffer. All titrations were carried out in an iTC200 calorimeter (Malvern, Malvern, UK). A typical titration scheme featured 19 or 20 2- μL injections of 500 μM of the monosaccharide into 50 μM protein in the ca. 200- μL stirred interaction cell. However, the concentrations were sometimes altered to sample a wider range of concentration space [56]. Three titrations were performed per sugar, and the resulting thermograms were serially integrated using NITPIC v. 1.1.7 [54,78]. All titrations for a given sugar were globally analyzed in ITCsy [55], and confidence intervals were obtained using the error-surface projection method [55,63]. All ITC illustrations were rendered in GUSI [79].

Hydrodynamic modeling

HYDROPRO [80] was used to carry out all hydrodynamic modeling in this study. For apo-EcMglB, the molar mass (M) and partial-specific volume (\bar{v}) were calculated using SEDNTERP [81] and inputted into HYDROPRO. The crystal structure used was 2FW0 [48] denuded of all water molecules and ligands (although this structure is described as “apo”, there is a citrate molecule located in the D-glucose-binding site). For modeling of the liganded, “closed” structure, 2FVY [48] was used. The procedure was the same, but mass (180 g/mol) was added and the weighted-average \bar{v} was calculated to account for the bound ligand.

Analytical ultracentrifugation

All AUC experiments were carried out at 20° C in a Beckman-Coulter Optima XL-I centrifuge (Beckman-Coulter Inc., Indianapolis, IN), and Assay Buffer was used in all cases. Charcoal-filled Epon centerpieces with 1.2-cm path-lengths were housed between two sapphire windows in a standard aluminum housing. After introduction to the centerpiece sectors, all sample cells were placed in an An50-Ti rotor that was subsequently incubated at the experimental temperature under vacuum in the centrifuge for at least 2.5 h prior to the initiation of rotation. The rotor speed was 50,000 rpm for all AUC studies. For normal SV experiments, 400 μL of Assay Buffer was introduced into the reference sector, and 400 μL of sample was placed in the sample

sector. Centrifugation continued for about 16 h, and concentration-profile data were collected using both absorbance and interference optics. Data were initially analyzed using the $c(s)$ method in SEDFIT [82,83]. As warranted, data were also analyzed using a discrete-species model in SEDPHAT [62].

For difference SV studies, a slightly different protocol was followed. First, a 5 mg/mL stock of TpMglB-2 (in Assay Buffer) was made, and stock solutions of the sugars to be examined were also prepared in Assay Buffer. Identical volumes of the sugars were dispensed to separate tubes, followed by the addition of identical volumes of the protein stock solution. The two pairs of sugars studied were D-mannose/D-glucose and D-ribose/D-glucose. In a third experiment (a negative control), no sugars were added, only Assay Buffer. The solution containing the non-binding sugar was placed into the reference sector of a centerpiece, and the solution containing D-glucose was inserted in the sample sector. For the D-mannose/D-glucose and buffer/buffer experiments (performed on different days), the experiment was carried out in a meniscus-matching centerpiece (Spin Analytical, Inc., Berwick, ME). Thus, ca. 10 μ L less of the glucose containing solution was dispensed than that of the reference-sector solution. The solution-column heights were equalized prior to the experiment by 30 min of centrifugation at 9,000 rpm. After equalization, the centrifugation cell was removed and repeatedly inverted to thoroughly mix the sector contents. The equalization step resulted in a small amount of non-binding sugar being transferred to the sample sector. The D-ribose/D-glucose experiments were done in a normal Beckman-Coulter centerpiece, so the column-height equalization step was not possible. A deliberate meniscus mismatch was introduced into this experiment by pipetting ca. 6 μ L less into the sample sector compared to the reference sector, as suggested [64]. The same pre-experiment incubation period and rotor speed as described above were utilized for these experiments. A “fringe-control” experiment in which buffer was placed in the reference sector and an identical concentration of protein (compared to the difference SV experiments) was always run side-by-side with the difference SV experiments. Sedimentation and sedimentation differences were monitored using the interference optics only; scans were obtained every 60 s for almost 17 h.

As presented by the data-acquisition software, interference data have the amount of fringe displacement, J , on the y -axis and the radius from the center of rotation on x -axis (r). In this case, of course, the difference in fringe displacement, ΔJ , was measured. Difference SV data, as mentioned above, take on the appearance of a Gaussian distribution inverted about the x -axis (Fig 8A). The details of the analytical method will be published elsewhere (C.A.B., R.K.D., & M.V.N., in preparation). Briefly, to analyze these data, we first identified by eye which scans displayed a suitable amount of both meniscus depletion and evident plateau regions. Usually a subset of ca. 80 scans was suitable for analysis under this criterion. Next, we determined the time-invariant noise in the scans by examining the last 10 scans (in which no sedimentation was occurring, as solute had been depleted from the analysis range) and calculating the noise assuming no sedimenting species in SEDFIT [84]. This noise was subtracted from the subject difference-data scans. Next, the first moment of the difference curve by first subtracting a baseline value and then finding the area above the curve $\Delta J \cdot r$, in essence performing the integral

$$\int_{r_1}^{r_2} \Delta J r dr, \tag{1}$$

where r_1 and r_2 are dynamically chosen radial values proximal to the meniscus and plateau regions, respectively. These values were normalized by quantity $J_0 \bar{r}_m^2$, where J_0 is the total amount of fringe displacement of the protein (from the “fringe-control” experiment described above) and \bar{r}_m is the average of the two menisci in the difference SV experiment. Thus the

quantity

$$\left| \frac{\int_{r_1}^{r_2} \Delta J r dr}{J_0 \bar{r}_m^2} \right| \quad (2)$$

was tabulated for every scan. Finally these data were plotted as the ordinate for the respective abscissa

$$\frac{r_m^2}{\bar{r}} \ln \frac{\bar{r}}{\bar{r}_m}, \quad (3)$$

where \bar{r} is the radial position of the minimum in the ΔJ curve. These data were fitted to a straight line, and the slope was taken as the quantity $\Delta s/\bar{s}$. The 95% confidence interval of this value was determined using F-statistics and the error-surface projection method, as discussed elsewhere [63]. Estimation of $\Delta s/\bar{s}$ for the addition of ligand without conformational change was accomplished using Eq. 19 of Kirschner & Schachman [64].

Differential scanning fluorimetry

The thermal stability of TpMglB-2 in the presence of various putative ligands was determined in a 96-well PCR-plate (Bio-Rad Laboratories, Inc., Hercules, CA). The assumption was that binding ligands would enhance the thermal stability of the protein. For carbon-source screening, Biolog's Phenotype MicroArray (PM) compounds supplied in 96-well microplates (BIOLOG, Inc., Hayward, CA) were dissolved in 50 μ L of sterile water to obtain a final concentration of around 10–20 mM. Screenings were performed with plates PM1 and PM2A. Each plate contains 95 compounds and a blank (no ligand) control. The complete plate contents are available at the BIOLOG website (www.biolog.com).

Each 20 μ L standard assay mixture in a 96-well PCR-plate contained 10 μ M purified protein and SYPRO Orange (Life Technologies) at 5x concentration in a buffer containing 10 mM phosphate, 100 mM NaCl, pH 7.4. Two μ L of the resuspended BIOLOG compounds were added to each well. The plate was sealed and placed in a BioRad CFX96 real-time PCR detector coupled to a C1000 thermal cycler (BioRad Laboratories, Inc., Hercules, CA), and the fluorescence of the dye was monitored as a function of temperature from 4° C to 95° C. Fluorescence readings were recorded every 0.5° C. The data were transformed to $-dF/dT$ v. T curves, and the abscissas of the minima in these curves were defined as the apparent melting temperatures ($T_{m, app}$). Significant positive shifts in $T_{m, app}$ were taken as positive results and were confirmed using ITC.

Bioinformatics

BLAST [51] was used to identify protein sequences similar TpMglB-2. All hits were visually examined to ensure that full coverage of the TpMglB-2 was achieved. The top 100 hits were assembled, and identical or very similar sequences (i.e. they contained identical sequential cores with differences only on the periphery) were eliminated. In the end, 78 sequences were analyzed, including the outgroup of two sequences (MglB proteins from *E. coli* and *S. typhimurium*). PROMALS-3D [85] was used to align the sequences using the current structure of TpMglB-2 as a structural template. This output was converted to Phylip format using the online resource at insilico.ehu.es/tophylip/. PhyML [86] was applied to establish the maximum-likelihood tree for the protein sequences. The output from this session was visualized by running a short Python script that utilized the toolkit ETE [87] version 3.

Supporting Information

S1 File. Raw Data Presented in this Paper. File S1_File.7z is a compressed folder containing the raw data for the AUC, ITC, and bioinformatics results described in this paper. (7Z)

Acknowledgments

This research was supported by a grant (AI056305) to M.V.N. from the National Institutes of Health. Some results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

Author Contributions

Conceived and designed the experiments: CAB RKD MVN.

Performed the experiments: CAB RKD WZL.

Analyzed the data: CAB RKD.

Wrote the paper: CAB RKD WZL MVN.

References

1. Simms I, Fenton KA, Ashton M, Turner KME, Crawley-Boevey EE, Gorton R, et al. The Re-Emergence of Syphilis in the United Kingdom: The New Epidemic Phases. *Sex Transm Dis.* 2005; 32: 220–226. PMID: [15788919](#)
2. Golden MR, Marra CM, Holmes KK. Update on syphilis: resurgence of an old problem. *J Am Med Assoc.* 2003; 290: 1510–1514.
3. Bremer V, Marcus U, Hamouda O. Syphilis on the rise again in Germany—results from surveillance data for 2011. *Sex Transm Dis.* 2012; 7–11.
4. Velicko I, Unemo M. Recent trends in gonorrhoea and syphilis epidemiology in Sweden: 2007 to 2011. *Control.* 2012; 1–6.
5. Mattei PL, Beachkofsky TM, Gilson RT, Wisco OJ. Syphilis: a reemerging infection. *Am Fam Physician.* 2012; 86: 433–440. PMID: [22963062](#)
6. Kerani RP, Handsfield HH, Stenger MS, Shafii T, Zick E, Brewer D, et al. Rising rates of syphilis in the era of syphilis elimination. *Sex Transm Dis.* 2007; 34: 154–61. PMID: [17179773](#)
7. Cohen MS, Hawkes S, Mabey D. Syphilis returns to China. . .with a vengeance. *Sex Transm Dis.* 2006; 33: 724–725. PMID: [17130806](#)
8. Centers for Disease Control and Prevention. Primary and secondary syphilis- Jefferson County, Alabama, 2002–2007. *Morb Mortal Wkly Rep.* 2009; 58: 463–467.
9. Cox DL, Luthra A, Dunham-Ems S, Desrosiers DC, Salazar JC, Caimano MJ, et al. Surface immunolabeling and consensus computational framework to identify candidate rare outer membrane proteins of *Treponema pallidum*. *Infect Immun.* 2010; 78: 5178–94. doi: [10.1128/IAI.00834-10](#) PMID: [20876295](#)
10. Cox DL, Chang P, McDowall AW, Radolf JD. The outer membrane, not a coat of host proteins, limits antigenicity of virulent *Treponema pallidum*. *Infect Immun.* 1992; 60: 1076–1083. PMID: [1541522](#)
11. Cullen PA, Haake DA, Adler B. Outer membrane proteins of pathogenic spirochetes. *FEMS Microbiol Rev.* 2004; 28: 291–318. PMID: [15449605](#)
12. Desrosiers DC, Anand A, Luthra A, Dunham-Ems SM, LeDoyt M, Cummings MAD, et al. TP0326, a *Treponema pallidum* β -barrel assembly machinery A (BamA) orthologue and rare outer membrane protein. *Mol Microbiol.* 2011; 80: 1496–515. doi: [10.1111/j.1365-2958.2011.07662.x](#) PMID: [21488980](#)
13. Radolf JD, Robinson EJ, Bourell KW, Akins DR, Porcella SF, Weigel LM, et al. Characterization of outer membranes isolated from *Treponema pallidum*, the syphilis spirochete. *Infect Immun.* 1995; 63: 4244–4252. PMID: [7591054](#)

14. Haake DA. Spirochaetal lipoproteins and pathogenesis. *Microbiology*. 2000; 146: 1491–1504. PMID: [10878114](#)
15. Deka RK, Brautigam CA, Yang XF, Blevins JS, Machius M, Tomchick DR, et al. The PnrA (Tp0319; TmpC) lipoprotein represents a new family of bacterial purine nucleoside receptor encoded within an ATP-binding cassette (ABC)-like operon in *Treponema pallidum*. *J Biol Chem*. 2006; 281: 8072–8081. PMID: [16418175](#)
16. Lee Y-H, Deka RK, Norgard M V., Radolf JD, Hasemann CA. *Treponema pallidum* TroA is a periplasmic zinc-binding protein with a helical backbone. *Nat Struct Biol*. 1999; 184: 628–633.
17. Deka RK, Neil L, Hagman KE, Machius M, Tomchick DR, Brautigam CA, et al. Structural evidence that the 32-kilodalton lipoprotein (Tp32) of *Treponema pallidum* is an L-methionine-binding protein. *J Biol Chem*. 2004; 279: 55644–55650. PMID: [15489229](#)
18. Machius M, Brautigam CA, Tomchick DR, Ward P, Otwinowski Z, Blevins JS, et al. Structural and biochemical basis for polyamine binding to the Tp0655 lipoprotein of *Treponema pallidum*: putative role for Tp0655 (TpPotD) as a polyamine receptor. *J Mol Biol*. 2007; 373: 681–694. PMID: [17868688](#)
19. Desrosiers DC, Sun YC, Zaidi AA, Eggers CH, Cox DL, Radolf JD. The general transition metal (Tro) and Zn²⁺ (Znu) transporters in *Treponema pallidum*: analysis of metal specificities and expression profiles. *Mol Microbiol*. 2007; 65: 137–152. PMID: [17581125](#)
20. Deka RK, Brautigam CA, Bidy BA, Liu WZ, Norgard M V. Evidence for an ABC-type riboflavin transporter system in pathogenic spirochetes. *MBio*. 2013; 4: e00615–12. doi: [10.1128/mBio.00615-12](#) PMID: [23404400](#)
21. Deka RK, Brautigam CA, Liu WZ, Tomchick DR, Norgard M V. The TP0796 lipoprotein of *Treponema pallidum* is a bimetal-dependent FAD pyrophosphatase with a potential role in flavin homeostasis. *J Biol Chem*. 2013; 288: 11106–21. doi: [10.1074/jbc.M113.449975](#) PMID: [23447540](#)
22. Deka RK, Brautigam CA, Liu WZ, Tomchick DR, Norgard M V. Molecular insights into the enzymatic diversity of flavin-trafficking protein (Ftp; formerly ApbE) in flavoprotein biogenesis in the bacterial periplasm. *Microbiologypopen*. 2015; 5: 21–38. doi: [10.1002/mbo3.306](#) PMID: [26626129](#)
23. Brautigam CA, Deka RK, Liu WZ, Norgard M V. Insights into the potential function and membrane organization of the TP0435 (Tp17) lipoprotein from *Treponema pallidum* derived from structural and biophysical analyses. *Protein Sci*. 2015; 24: 11–19. doi: [10.1002/pro.2576](#) PMID: [25287511](#)
24. Deka RK, Brautigam CA, Goldberg M, Schuck P, Tomchick DR, Norgard M V. Structural, bioinformatic, and *in vivo* analyses of two *Treponema pallidum* lipoproteins reveal a unique TRAP transporter. *J Mol Biol*. 2012; 416: 678–696. doi: [10.1016/j.jmb.2012.01.015](#) PMID: [22306465](#)
25. Brautigam CA, Deka RK, Schuck P, Tomchick DR, Norgard M V. Structural and thermodynamic characterization of the interaction between two periplasmic *Treponema pallidum* lipoproteins that are components of a TPR-protein-Associated TRAP transporter (TPAT). *J Mol Biol*. 2012; 420: 70–86. doi: [10.1016/j.jmb.2012.04.001](#) PMID: [22504226](#)
26. Deka RK, Machius M, Norgard M V., Tomchick DR. Crystal structure of the 47-kDa lipoprotein of *Treponema pallidum* reveals a novel penicillin-binding protein. *J Biol Chem*. 2002; 277: 41857–41864. PMID: [12196546](#)
27. Brautigam CA, Deka RK, Ouyang Z, Machius M, Knutsen G, Tomchick DR, et al. Biophysical and bioinformatic analyses implicate the *Treponema pallidum* Tp34 lipoprotein (Tp0971) in transition metal homeostasis. *J Bacteriol*. 2012; 194: 6771–6781. doi: [10.1128/JB.01494-12](#) PMID: [23042995](#)
28. Deka RK, Brautigam CA, Liu WZ, Tomchick DR, Norgard M V. Evidence for posttranslational protein flavinylation in the syphilis spirochete *Treponema pallidum*: Structural and biochemical insights from the catalytic core of a periplasmic flavin-trafficking protein. *MBio*. 2015; 6: 1–9.
29. Deka RK, Brautigam CA, Tomson FL, Lumpkins SB, Tomchick DR, Machius M, et al. Crystal structure of the Tp34 (TP0971) lipoprotein of *Treponema pallidum*: implications of its metal-bound state and affinity for human lactoferrin. *J Biol Chem*. 2007; 282: 5944–5958. PMID: [17192261](#)
30. Fraser CM, Norris SJ, Weinstock GM, White O, Sutton GG, Dodson R, et al. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science*. 1998; 281: 375–388. PMID: [9665876](#)
31. Riley M, Abe T, Arnaud MB, Berlyn MKB, Blattner FR, Chaudhuri RR, et al. *Escherichia coli* K-12: A cooperatively developed annotation snapshot—2005. *Nucleic Acids Res*. 2006; 34: 1–9. PMID: [16397293](#)
32. Barbieri JT, Austin FE, Cox CD. Distribution of glucose incorporated into macromolecular material by *Treponema pallidum*. *Infect Immun*. 1981; 31: 1071–1077. PMID: [7014454](#)
33. Barbieri JT, Cox CD. Glucose incorporation by *Treponema pallidum*. *Infect Immun*. 1979; 24: 291–293. PMID: [378845](#)

34. Nichols JC, Baseman JB. Carbon source utilized by virulent *Treponema pallidum*. *Infect Immun*. 1975; 12: 1044–1050. PMID: [1104482](#)
35. Schiller NL, Cox CD. Catabolism of glucose and fatty acids by virulent *Treponema pallidum*. *Infect Immun*. 1977; 16: 60–68. PMID: [326678](#)
36. Ning C, Segal S. Plasma galactose and galactitol concentration in patients with galactose-1-phosphate uridylyltransferase deficiency galactosemia: Determination by gas chromatography/mass spectrometry. *Metabolism*. 2000; 49: 1460–1466. PMID: [11092512](#)
37. Deka RK, Goldberg MS, Hagman KE, Norgard M V. The Tp38 (TpMg1B-2) lipoprotein binds glucose in a manner consistent with receptor function in *Treponema pallidum*. *J Bacteriol*. 2004; 186: 2303–2308. PMID: [15060032](#)
38. Hazelbauer GL, Adler J. Role of the galactose binding protein in chemotaxis of *Escherichia coli* toward galactose. *Nat New Biol*. 1971; 230: 101–104. PMID: [4927373](#)
39. Fukami-Kobayashi K, Tateno Y, Nishikawa K. Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J Mol Biol*. 1999; 286: 279–290. PMID: [9931266](#)
40. Berntsson RP-A, Smits SHJ, Schmitt L, Slotboom D-J, Poolman B. A structural classification of substrate-binding proteins. *FEBS Lett*. 2010; 584: 2606–17. doi: [10.1016/j.febslet.2010.04.043](#) PMID: [20412802](#)
41. Felder CB, Graul RC, Lee AY, Merkle H-P, Sadee W. The venus flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors. *AAPS J*. 1999; 1: 7–26.
42. Vyas NK, Vyas MN, Quioco FA. A novel calcium binding site in the galactose-binding protein of bacterial transport and chemotaxis. *Nature*. 1987; 327: 635–637. PMID: [3600760](#)
43. Adams PD, Afonine P V, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure determination. *Acta Crystallogr Sect D Biol Crystallogr*. 2010; 66: 213–221.
44. Hodel A, Kim S-H, Brunger AT. Model bias in molecular crystal structures. *Acta Crystallogr A*. 1992; 48: 851–859.
45. Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res*. 2005; 33: W244–W248. PMID: [15980461](#)
46. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr Sect D Biol Crystallogr*. 2004; 60: 2256–2268.
47. Becker PS, Akins DR, Radolf JD, Norgard M V. Similarity between the 38-kilodalton lipoprotein of *Treponema pallidum* and the glucose/galactose-binding (MgIB) protein of *Escherichia coli*. *Infect Immun*. 1994; 62: 1381–1391. PMID: [8132345](#)
48. Borrok MJ, Kiessling LL, Forest KT. Conformational changes of glucose/galactose-binding protein illuminated by open, unliganded, and ultra-high-resolution ligand-bound structures. *Protein Sci*. 2007; 16: 1032–1041. PMID: [17473016](#)
49. Borrok MJ, Zhu Y, Forest KT, Kiessling LL. Structure-based design of a periplasmic binding protein antagonist that prevents domain closure. *ACS Chem Biol*. 2009; 4: 447–456. doi: [10.1021/cb900021q](#) PMID: [19348466](#)
50. Vyas MN, Jacobson BL, Quioco FA. The calcium-binding site in the galactose chemoreceptor protein. Crystallographic and metal-binding studies. *J Biol Chem*. 1989; 264: 20817–20821. PMID: [2684986](#)
51. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997; 25: 3389–3402. PMID: [9254694](#)
52. Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ, Wasney GA, et al. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A*. 2006; 103: 15835–40. PMID: [17035505](#)
53. Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc*. 2007; 2: 2212–2221. PMID: [17853878](#)
54. Keller S, Vargas C, Zhao H, Piszczek G, Brautigam CA, Schuck P. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal Chem*. 2012; 84: 5066–73. doi: [10.1021/ac3007522](#) PMID: [22530732](#)
55. Brautigam CA, Zhao H, Vargas C, Keller S, Schuck P. Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nat Protoc*. Nature Publishing Group; 2016; 11: 882–894.
56. Zhao H, Piszczek G, Schuck P. SEDPHAT- a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods*. 2015; 76: 137–148. doi: [10.1016/j.ymeth.2014.11.012](#) PMID: [25477226](#)

57. Shilton BH, Flocco MM, Nilsson M, Mowbray SL. Conformational changes of three periplasmic receptors for bacterial chemotaxis and transport: the maltose-, glucose/galactose- and ribose-binding proteins. *J Mol Biol.* 1996; 264: 350–363. PMID: [8951381](#)
58. Cuneo MJ, Changela A, Miklos AE, Beese LS, Krueger JK, Hellinga HW. Structural analysis of a periplasmic binding protein in the tripartite ATP-independent transporter family reveals a tetrameric assembly that may have a role in ligand transport. *J Biol Chem.* 2008; 283: 32812–32820. doi: [10.1074/jbc.M803595200](#) PMID: [18723845](#)
59. Gonin S, Arnoux P, Pierru B, Lavergne J, Alonso B, Sabaty M, et al. Crystal structures of an extracytoplasmic solute receptor from a TRAP transporter in its open and closed forms reveal a helix-swapped dimer requiring a cation for α -keto acid binding. *BMC Struct Biol.* 2007; 7: 11. PMID: [17362499](#)
60. Errington N, Rowe AJ. Probing conformation and conformational change in proteins is optimally undertaken in relative mode. *Eur Biophys J.* 2003; 32: 511–517. PMID: [12830332](#)
61. Ghirlando R, Balbo A, Piszczek G, Brown PH, Lewis MS, Brautigam CA, et al. Improving the thermal, radial, and temporal accuracy of the analytical ultracentrifuge through external references. *Anal Biochem.* 2013; 440: 81–95. doi: [10.1016/j.ab.2013.05.011](#) PMID: [23711724](#)
62. Schuck P. Sedimentation analysis of noninteracting and self-associating solutes using numerical solutions to the Lamm Equation. *Biophys J.* 1998; 75: 1503–1512. PMID: [9726952](#)
63. Bevington PR, Robinson DK. Data reduction and error analysis for the physical sciences. 2nd ed. Boston, MA: WCB/McGraw-Hill; 1992.
64. Kirschner MW, Schachman HK. Conformational changes in proteins as measured by difference sedimentation studies. I. A technique for measuring small changes in sedimentation coefficient. *Biochemistry.* 1971; 10: 1900–1919. PMID: [5563767](#)
65. Cover WH, Norris SJ, Miller JN. The microaerophilic nature of *Treponema pallidum*: enhanced survival and incorporation of tritiated adenine under microaerobic conditions in the presence or absence of reducing compounds. *Sex Transm Dis.* 1982; 9: 1–8. PMID: [10328016](#)
66. Hagman KE, Porcella SF, Popova TG, Norgard M V. Evidence for a methyl-accepting chemotaxis protein gene (*mcp1*) that encodes a putative sensory transducer in virulent *Treponema pallidum*. *Infect Immun.* 1997; 65: 1701–1709. PMID: [9125550](#)
67. Greene SR, Stamm L V. Molecular characterization of *Treponema pallidum mcp2*, a putative chemotaxis protein gene. *Infect Immun.* 1998; 66: 2999–3002. PMID: [9596781](#)
68. Charon NW, Goldstein SF. Genetics of motility and chemotaxis of a fascinating group of bacteria: the spirochetes. *Annu Rev Genet.* 2002; 36: 47–73. PMID: [12429686](#)
69. Norris SJ, Cox DL, Weinstock GM. Biology of *Treponema pallidum*: correlation of functional activities with genome sequence data. *J Mol Microbiol Biotechnol.* 2001; 3: 37–62. PMID: [11200228](#)
70. Klock HE, Koesema EJ, Knuth MW, Lesley S a. Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins Struct Funct Genet.* 2008; 71: 982–994. PMID: [18004753](#)
71. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr Sect D Biol Crystallogr.* 2006; 62: 859–866.
72. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. *Acta Crystallogr Sect D Biol Crystallogr.* 2002; D58: 1772–1779.
73. Otwinowski Z. Maximum likelihood refinement of heavy atom parameters. *Proc CCP4 Daresbury Study Weekend.* 1991; 80–86.
74. Cowtan K, Main P. Miscellaneous algorithms for density modification. *Acta Crystallogr Sect D Biol Crystallogr.* 1998; 54: 487–493.
75. Morris RJ, Perrakis A, Lamzin VS. ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.* 2003; 374: 229–244. PMID: [14696376](#)
76. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr. International Union of Crystallography;* 2007; 40: 658–674. PMID: [19461840](#)
77. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr.* 2004; 60: 2126–2132.
78. Scheuermann TH, Brautigam CA. High-precision, automated integration of multiple isothermal titration calorimetric thermograms: new features of NITPIC. *Methods.* 2015; 76: 87–98. doi: [10.1016/j.ymeth.2014.11.024](#) PMID: [25524420](#)
79. Brautigam CA. Calculations and publication-quality illustrations for analytical ultracentrifugation data. *Methods Enzymol.* 2015; 562: 109–134. doi: [10.1016/bs.mie.2015.05.001](#) PMID: [26412649](#)

80. de la Torre JG, Huertas ML, Carrasco B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structures. *Biophys J.* 2000; 78: 719–730. PMID: [10653785](#)
81. Laue TM, Shah BD, Ridgeway RM, Pelletier SL. Computer-aided interpretation of analytical sedimentation data for proteins. In: Harding SE, Rowe AJ, Horton JC, editors. *Analytical Ultracentrifugation in Biochemistry and Polymer Science.* Cambridge, UK: The Royal Society of Chemistry; 1992. pp. 90–125.
82. Schuck P. Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys J.* 2000; 78: 1606–1619. PMID: [10692345](#)
83. Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D. Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys J.* 2002; 82: 1096–1111. PMID: [11806949](#)
84. Schuck P, Demeler B. Direct sedimentation analysis of interference optical data in analytical ultracentrifugation. *Biophys J.* 1999; 76: 2288–2296. PMID: [10096923](#)
85. Pei J, Kim B-H, Grishin N V. PROMALS3D: a tool for multiple sequence and structure alignment. *Nucleic Acids Res.* 2008; 36: 2295–2300. doi: [10.1093/nar/gkn072](#) PMID: [18287115](#)
86. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 2010; 59: 307–321. doi: [10.1093/sysbio/syq010](#) PMID: [20525638](#)
87. Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, analysis and visualization of phylogenomic data. *Mol Biol Evol.* 2016; in press.