Published in final edited form as:
J Mol Biol. 2012 March 9; 416(5): 678–696. doi:10.1016/j.jmb.2012.01.015.

Structural, bioinformatic, and *in vivo* analyses of two *Treponema pallidum* lipoproteins reveal a unique TRAP transporter

Ranjit K. Deka[a,*], **Chad A. Brautigam**[b,*], **Martin Goldberg**[a], **Peter Schuck**[c], **Diana R. Tomchick**[b], and **Michael V. Norgard**[a,§]

[a]Department of Microbiology, The University of Texas, Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, TX 75390, USA

[b]Department of Biochemistry, The University of Texas, Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, TX 75390, USA

[c]Dynamics of Molecular Assembly Section, Laboratory of Cellular Imaging and Macromolecular Biophysics, NIBIB, National Institutes of Health, Bethesda, MD, USA

Abstract

Treponema pallidum, the bacterial agent of syphilis, is predicted to encode one tripartite ATP-independent periplasmic transporter (TRAP-T). TRAP-Ts typically employ a periplasmic substrate-binding protein (SBP) to deliver the cognate ligand to the transmembrane symporter. Herein, we demonstrate that the genes encoding the putative TRAP-T components from *T. pallidum*, *tp0957* (the SBP) and *tp0958* (the symporter) are in an operon with an uncharacterized third gene, *tp0956*. We determined the crystal structure of recombinant Tp0956; the protein is trimeric and perforated by a pore. Part of Tp0956 forms an assembly similar to those of "tetratricopeptide repeat" (TPR) motifs. The crystal structure of recombinant Tp0957 was also determined; like the SBPs of other TRAP-Ts, there are two lobes separated by a cleft. In these other SBPs, the cleft binds a negatively charged ligand. However, the cleft of Tp0957 has a strikingly hydrophobic chemical composition, indicating that its ligand may be substantially different and likely hydrophobic. Analytical ultracentrifugation of the recombinant versions of Tp0956 and Tp0957 established that these proteins associate avidly. This unprecedented interaction was confirmed for the native molecules using *in vivo* cross-linking experiments. Finally, bioinformatic analyses suggested that this transporter exemplifies a new subfamily of TPR-protein associated TRAP transporters (TPATs) that require the action of a TPR-containing accessory protein for the periplasmic transport of a potentially hydrophobic ligand(s).

Keywords

TRAP transporter; syphilis; *Treponema pallidum*; TPR motif; protein interactions

§To whom correspondence should be addressed: Phone: 214-633-0015, Fax: 214-648-5905, Michael.Norgard@UTSouthwestern.edu.
*These authors contributed equally to this work.

Accession Numbers

Coordinates and structure factors for the TatT and $TatP_T$ structures have been deposited in the Protein Data Bank; the accession numbers are 3U64 and 3U65, respectively.

Introduction

Treponema pallidum, the etiologic agent of syphilis, is an obligate human pathogen that metabolically relies heavily on its human host [1]. The spirochete encodes only 1,067 genes (cf. the genome of *Escherichia coli* K-12, which has over 4,000 genes). *T. pallidum* lacks genes of the biosynthetic pathways for fatty acids, nucleotides, most of the amino acids, and many other vital metabolites. Natural selection has thus honed *T. pallidum* to be an organism with an extremely parsimonious genome, likely at the cost of its ability to be free-living. From a more practical perspective, the need for *T. pallidum* to acquire so many of its vital nutrients from its human host likely has accounted for the historic inability to continuously cultivate the organism *in vitro* [2]. Indeed, much of our lack of understanding of this enigmatic human pathogen derives directly from this major investigative obstacle.

Bacteria use many means to import needed metabolites from their extracellular milieu into their cytoplasms. The primary ABC transporters and secondary symporters are examples [3;4]. ABC transporters utilize a substrate-binding protein (SBP) in the periplasm to secure the ligand and deliver it to a transmembrane permease, which uses the energy from ATP hydrolysis to transport the ligand against a concentration gradient into the cytoplasm. Canonical symporters do not require the services of an SBP; rather, a periplasmic ligand binds directly to the transmembrane symporter. An abundant ion (usually either H^+ or Na^+) also binds, and the favorable energetics of co-transporting the ion are used to allow delivery of the ligand to the cytoplasm, again typically against a concentration gradient.

Recently, a different family of solute transporters has been described— the tripartite ATP-independent periplasmic transporters (TRAP-Ts) [5]. The TRAP-Ts incorporate features of both the primary and secondary transporters. The *dctPQM* operon of *Rhodobacter capsulatus* is the prototypical TRAP-T; it is responsible for the import of C_4-dicarboxylates [6]. Like the ABC transporters, the DctPQM system and all other TRAP-Ts employ an SBP (called “P” herein) to capture the ligand [7] and deliver it to the transmembrane parts of the transporter (called “Q” and “M” herein) [6]. ATP hydrolysis does not energize these transporters. Rather, like secondary transporters, the energy stored in an electrochemical gradient across the plasma membrane facilitates ligand transport [8]. The molecular details of how the transport is achieved are not yet fully elucidated, though studies on the SiaPQM TRAP-T indicate the requirement for the P component, an electric gradient, and the co-transport of at least two Na^+ ions per ligand transported [8]. The ligands for TRAP-Ts are typically acidic small molecules [9]. A related, but distinct, family of transporters that import tricarboxylic acids has also been described [10].

The genome of *T. pallidum* is predicted to encode a single TRAP-T [1; 5], apparently organized as an operon. The genes are *tp0957* (the P component) and *tp0958* (putative fused Q and M components). A third gene within this putative operon is *tp0956*, which belongs to the Cluster of Orthologous Groups (COG) 5660[11], a family of putative integral membrane proteins. This protein is also flagged as a member of the RskA superfamily[11]; in *Mycobacterium tuberculosis*, this protein binds to sigma factor K, repressing its activity[12]. Alternatively, because it is similar in size to the P protein, it had been speculated that Tp0956 might serve as a second SBP for this TRAP-T [5]. Having a second SBP would likely have important functional consequences, given the very limited genome of *T. pallidum*.

In this study, we first set out to characterize this putative second SBP, Tp0956. However, *T. pallidum* is not cultivatable *in vitro*, thereby precluding genetic manipulation of the organism. Thus, we took advantage of a structure-to-function approach for this protein that has been successful for other *T. pallidum* proteins of unknown functions [13;14;15;16]. First, reverse transcription and amplification of RNA from rabbit-tissue-derived *T. pallidum*

confirmed that *tp0956*, *tp0957*, and *tp0958* constitute a single transcriptional unit. We then determined the crystal structure of Tp0956. Notably, we found that the protein has no structural homology to known P proteins, making it unlikely that it serves as an additional SBP for this putative TRAP-T. The protein is a mostly α-helical irregular cylinder with a pore running through its center. It features substructures similar to tetratricopeptide repeat (TPR) motifs, which are hallmarks of proteins that are involved in protein-protein interactions[17;18]. Because of the unusual nature of this TRAP operon, we also pursued and obtained a crystal structure for Tp0957, the putative P component. The overall fold of Tp0957 is similar to P components from other TRAP-Ts, but there are notable differences, particularly in the ligand-binding clefts of the proteins; that of Tp0957 is more hydrophobic than other structurally characterized P proteins. Although there is some evidence of a small-molecule ligand bound to the protein, Tp0957 adopts a conformation more similar to an unliganded TRAP-T SBP. The existence of TPR motifs in Tp0956 led us to seek its binding partner(s). Using analytical ultracentrifugation (AUC), we found that trimeric Tp0956 interacts with Tp0957. This observation was confirmed *in vivo* with formaldehyde cross-linking. Finally, we showed that proteins homologous to Tp0956 are found in other TRAP-T operons. Thus, the *T. pallidum* TRAP-T may serve as a paradigm for a hitherto unknown class of TRAP-Ts (which we call TPATs; see below) that apparently require a TPR-containing accessory protein.

Results

Analyses of *tp0956*, *tp0957*, and *tp0958*

Genomic analysis of *T. pallidum* suggested that *tp0956*, *tp0957* and *tp0958* are oriented in the same direction and thus may be cotranscribed[1] (Fig. 1). Furthermore, bacterial genes belonging to a transporter often are cotranscribed, and bioinformatic analyses predicted *tp0957* and *tp0958* as components of a TRAP-T. Indeed, the TRAP-T operon (SiaPQM) from *Haemophilus influenzae* is organized similarly[19] (Fig. 1). Therefore, we performed RT-PCR to examine the transcriptional linkage of these genes. As shown in Fig. 1, the three-gene cluster from *T. pallidum* is transcriptionally linked. One potential caveat is that the RT-PCR product observed between the *tp0954* and *tp0956* pair (Fig. 1; Table S1) may be due to the presence of non-coding mRNA that is transcribed in the opposite direction of this operon.

In accord with its putative function as a permease, Tp0958 (a likely fusion of the M and Q components) is quite hydrophobic (average hydropathy = 1.084) and is predicted to be an all-α-helical transmembrane protein. Signal sequences at their N-termini identify both Tp0956 and Tp0957 as putative lipoproteins[20] (see Table S2). As such, their N-terminal signal sequences would be removed, leaving a tri-acylated cysteine residue at their respective N-termini that would anchor the proteins to the periplasmic side of either the inner or outer membranes of the bacterium.

A conserved-domain database (CDD) search[11] using the sequence of Tp0956 places it in COG5660, a family of integral membrane proteins. This fact raises the question: is the Tp0956 an integral membrane protein or a lipoprotein? We purified recombinant forms of both Tp0956 and Tp0957 almost identically (see Materials & Methods). In these proteins, the N-terminal cysteines had been removed. This expedient was designed to prevent the acylation reaction mentioned above, and should render the proteins water soluble. Indeed, they could be purified with only a low concentration (2 mM; well below the critical micelle concentration) of the non-ionic detergent 2-octyl-glucopyranoside (β-OG). We often include this additive in preparations of recombinantly solubilized treponemal lipoproteins[13;14;15;16], as it appears to improve their stability in solution. Additionally, b-OG can have a positive effect on protein crystallization[21]. The water solubility and probable lipidation of both

proteins suggest that they exist in the periplasm, not in the cell membrane, as implied for Tp0956. The CDD search also placed Tp0956 in the RskA superfamily; as noted above, RskA is an anti-sigma factor in *M. tuberculosis*. It seems unlikely that Tp0956 could play a role in transcription, given its probable periplasmic localization.

The crystal structure of Tp0956

In view of the possibility that Tp0956 may be a second SBP for a TRAP-T, we set out to determine the three-dimensional structure of Tp0956 (Fig. 2; Table 1). The crystal structure of Tp0956 was refined using X-ray diffraction data to 2.3 Å resolution. The model (Figs. 2A & B) has 13 α-helices, one short 3_{10} helix, and no β-strands. The structure is very different from the bilobed structures (see below) of known P components of TRAP-Ts, making it unlikely that Tp0956 fulfills this proposed role. Residues 31-301 (in the numbering of the mature, processed form of the protein) are visible in electron-density maps; the first thirty residues of this protein construct are presumed present but were disordered in the crystals. The α-helices, designated “A” through “M”, have an antiparallel arrangement that forms an irregular ring (Figs. 2A-C). The ring surrounds a pore, which perforates the entire protein (Fig. 2C). Hereafter, we refer to the “N,” “C,” and “lateral” faces of Tp0956, which are defined in Fig. 2B. The “N” and “C” faces are so called because they respectively house the observed N-terminus and C-terminus of Tp0956.

Characteristics of the pore

Tp0956’s pore runs from the N to the C faces of the protein, a total of ~43 Å (Fig. 3A). It has an inner portion (~20 Å in length; average diameter ~5.8 Å; minimum diameter = 4.8 Å; see Fig. S1) and two vestibules (the “N” and “C” vestibules, depending on their respective proximity to the N and C faces). Some soluble protein enzymes have pores that are related to their respective functions; the archeal 20 S proteasome is a multimeric assembly that has narrow (~13 Å) annuli[22], and the DNA topoisomerase of *E. coli* has a large (~28 Å) hole [23]. However, there is no evidence of an enzymatic activity for Tp0956. Although Tp0956 is a soluble protein (see above), the dimensions of its channel invite comparisons to membrane-spanning proteins. For example, the minimum pore diameters of aquaporin (AQP1), glyceroporins (GlpF), and an ion channel (KcsA) range from 2.6 to 3.7 Å [24;25;26], and those of bacterial porin outer-membrane proteins can be as small as 5 Å [27]. The pore of Tp0956 accommodates water molecules and a sulfate ion (from the precipitation buffer). There is a hydrophobic antechamber (HA) in the N vestibule (Fig 3C); it measures approximately 10 × 5 Å. The identities of the side chains lining the HA are shown in Table S3. Electrostatic calculations show (Fig. 2D) that the C vestibule is mostly negatively charged, the N vestibule is mostly positive, and the pore is mixed. No particular class of amino acid dominates the lining of the pore.

Comparisons to other structures

DALI [28] was used to find two structural homologs of Tp0956. They are DrR162B from *Deinococcus radiodurans* (PDB ID: 3GW4, no associated publication; Fig. 4A) and TTC0263 (PDB ID: 2PL2) from *Thermus thermophilus* strain HB27[29]. Their structural homologies (both r.m.s.d.’s are 1.8 Å) apply to roughly half (148 and 149, respectively) of the C_α atoms in Tp0956. Neither homologous protein forms a pore, nor do they appear to be involved in a TRAP-T operon.

These structural comparisons reveal that Tp0956 contains four helical hairpins that are structurally similar to “tetratricopeptide repeat” (TPR) motifs. Both DrR162B and TTC0263 contain these motifs, which occur as tandem, 34-amino-acid repeats in the proteins that harbor them. The TPR-like portions of Tp0956 encompass the α-helical pairs αC/αD, αG/αH, αI/αJ, and αK/αL (Fig. 4B); thus, these motifs form part of the pore and also part of the

lateral surface of the protein. To our knowledge, Tp0956 is the only known protein having TPR-like motifs that form part of a pore.

TPR motifs exist in many proteins[17]. Their folds comprise an antiparallel α-helical hairpin containing two α-helices and a short turn that connects them. These motifs are most often found as tandem repeats of three or more in proteins, where they stack to form a superhelical arrangement that features a concave and a convex surface. The concave surface of the TPR-like superhelix in Tp0956 forms part of the protein's pore (Fig. 4C). The superhelical stack is often "capped" by a C-terminal α-helix with an unrelated sequence; helix αM serves this function in Tp0956. The tandem arrays of TPR motifs mostly function as protein-interaction domains [17] in glucocorticoid receptors [30] and in the anaphase-promoting complex[31]. To date, the transporter association of Tp0956 (and others identified below) is unique among TPR-motif-bearing proteins.

The TPR motifs in Tp0956 are cryptic. That is, they failed to be detected by a hidden-Markhov algorithm designed to reveal their presence[32], and they do not adhere to the canonical 34-residue length. Nevertheless, the structural homology between the TPRs of DrR162B and Tp0956 is pronounced (Fig. 4A). Sequence analysis of Tp0956 in its TPR-like regions reveals that the protein does exhibit TPR-like sequence signatures (Fig. S2), but the interhelical region is longer than usual in all cases except one. These cryptic TPRs of Tp0956 are termed "cTPR motifs" (below).

The quaternary structure of Tp0956

Although there is only one copy of Tp0956 in the asymmetric unit, applying crystallographic symmetry to the protein reveals the presence of multiple potential oligomeric forms. There are two identical dodecamers in the unit cell, and these comprise 8 identical trimers (Fig. 5). Physicochemical characterization of the interfaces in the crystal structure using PISA[33] demonstrated that the most stable oligomer present in the crystals is the crystallographic trimer shown in Fig. 5. To assess the oligomeric state of Tp0956 in solution, we subjected it to sedimentation velocity (SV) analytical ultracentrifugation. As seen in Fig. 6A, there are two dominant forms of the protein under our solution conditions. The virtually concentration-independent sedimentation coefficients and the diffusional spreads of the clearly separate boundaries denote the presence of slowly interconverting species with apparent molar masses of 98,000 and 380,000 g/mol, within error consistent with trimeric and dodecameric forms of Tp0956 (theoretical values of 101,028 and 404,112 g/mol, respectively). There is evidence of monomeric Tp0956 (Fig. 6A), but only at a low concentration (≤ 1.6 μM). The dodecameric form only becomes prevalent at high concentrations (>20 μM). In the presence of a low concentration (2 mM; well below the critical micelle concentration) of β-OG, the trimeric form is predominant, and the tendency to form dodecamers is substantially reduced (Fig. S3). It therefore appears that the detergent disrupts the relatively weak interactions that drive dodecamer formation, but it cannot disrupt the formation of the trimer at this concentration. We therefore conclude that the trimeric form of Tp0956 is the most relevant in solution. The cTPR motifs are arrayed on the outside of this assembly (Fig. 5). Importantly, all of the membrane-anchoring N-termini (i.e. the "N" faces) are located on the same face of this trimer, i.e. facing away from the viewer in Fig. 5. This datum confirms the topological feasibility of this assembly. In the dodecamer (not shown), the N faces all point to the center of the "ball" of Tp09056s, a very unlikely arrangement for a membrane-tethered oligomer.

The crystal structure of Tp0957

The existence of a cTPR-motif-containing protein in the operon of this putative treponemal TRAP-T suggested that this transporter may act differently than other TRAP-Ts. Further, no

crystallographically characterized P components had a sequence more than 25% identical to the putative treponemal P, Tp0957. To explore the impact these differences have on the structure of the P component, we determined the crystal structure of Tp0957. X-ray diffraction data to 1.4 Å were used for the determination and model refinement (Table 1). The protein is bilobed (Figs. 7A & 7B); here, we call the two lobes the "Lobe I" (residues 9-136 and 226-263) and "Lobe II" (residues 137-225 and 264-320). The protein chain crosses from one lobe to the other at three points, the most prominent being a 40-residue α-helix (α11) that forms a backbone-like feature in the protein. Lobe I has a central, five-stranded β-sheet that is surrounded by helices. Four of the strands in the β-sheet are parallel. Like Lobe I, Lobe II also has a central β-sheet in which only one strand is antiparallel; however, this sheet has six strands. This second β-sheet is also surrounded by helices. The protein has no outstanding surface electrostatic features (Fig. 7C). Between the lobes is an opening (or "cleft") in the protein (Figs. 7C & 8A).

There are two copies of Tp0957 in the asymmetric unit of these crystals. These molecules have very similar structures; the r.m.s.d. of all comparable C_α atoms is 0.5 Å. The contents of this asymmetric unit raise the question of the oligomeric status of Tp0957. Most characterized P proteins are monomeric, but two show evidence of oligomerization [34;35]. We used SV experiments to establish Tp0957's oligomeric status in solution, and we found that Tp0957 is monomeric under our solution conditions (Fig. 6B). Further, a physicochemical analysis [33] of the interactions of the monomers found none that would lead to the formation of oligomers.

The fold of Tp0957 is typical of a P component of a TRAP-T. This group of SBPs belongs to superfamily 7 of the extracellular transport proteins [36]. Thus far, the structures of P components suggest that the proteins use a "Venus-fly trap" motion to capture the ligand in a cleft between the two lobes[9]. This mechanism has been convincingly shown for SiaP, which is part of a TRAP-T for sialic acid [19]. That protein has been crystallized in both an "open" (unliganded) form and a "closed" (liganded) form [37;38]. The open-to-closed transition may be characterized [37] as a ~30° rigid-body rotation of Lobe I closer to Lobe II about three hinge points (Fig. 8B). The structure of Tp0957 is much closer to the "open" form of SiaP (Fig. 8C); according to DaliLite [39], compared to Tp0957, 296 comparable C_α positions of "open" SiaP have an r.m.s.d. of 2.8 Å, while 297 C_αs of "closed" SiaP have an r.m.s.d. of 3.6 Å.

The unique cleft of Tp0957

As mentioned above, Tp0957 has a cleft between the two lobes. This opening leads to three different cavities (Figs. 7C & 8A). We call them Cavity A (19 Å deep by 5 Å wide), Cavity B (11 × 7 Å), and Cavity C (10 × 4 Å). The residues whose side chains line these cavities are shown in Table S3. They are predominantly hydrophobic. While Cavities B & C have analogs in other P structures, Cavity A does not.

In comparison to other P proteins, the ligand-binding pocket of Tp0957 is more hydrophobic. Indeed, almost all of the residues making polar contacts to ligands in other P proteins are hydrophobic at the equivalent positions in Tp0957. Notably, an arginine known to contact a carboxylate moiety in the ligands of structurally characterized P proteins[9;40;41] (cf. R147 of SiaP) is not present in Tp0957 (it is an alanine), and no compensatory positively charged residue is observed near to this position. The lack of polar side chains in this area implies that ligand recognition by Tp0957 is based on shape complementarity alone; matching patterns of opposite charge or hydrogen bond donors/acceptors on the ligand and protein are not likely to be binding determinants. This finding could imply a lack of specificity for ligands.

Further, the shapes of the clefts of "open" SiaP and Tp0957 differ significantly. In its open form, the cleft of SiaP is deep and spans the protein (Fig. 8B, left), whereas the cleft of Tp0957 resembles a much smaller pocket leading to the three aforementioned cavities (Figs. 7C and 8A). In comparing the two clefts, that of Tp0957 appears to have been partially covered with protein residues from a loop between β6 and α7 (residues 161-164), forming Cavity B. For Tp0957 to undergo the rigid-body rotation observed in SiaP, this loop would have to be significantly rearranged. These data suggest that the cognate ligand of Tp0957 differs in its chemical character from the small organic acid ligands of most P proteins and that the ligand-capture mechanism of Tp0957 may deviate from those of other P and SBP proteins.

There is no clear evidence for a bound ligand in the electron-density maps for Tp0957. There is a small patch of density proximal to the indole ring of W24 (Fig.8D), as if the unidentified chemical moiety were stacked on the ring. This density does not closely correspond to the structures of any buffer, salt, precipitant, or cryoprotectant present. It may represent a small molecule that copurified with the protein or a contaminant in the precipitation or cryoprotection solutions. Despite this ambiguity, the position of the density feature roughly corresponds to that adopted by the ligands in other P-protein structures. Besides the side chain of W24, the feature is proximal to the side chains of residues I17, P19, A136, L163, and I228, making it likely that the moiety bound here is hydrophobic.

The interaction of Tp0956 and Tp0957

Given that cTPR motifs exist inTp0956, we set out to identify its binding partner. Our first attempt was with Tp0957, as the transcription of the two proteins is linked (Fig. 1). SV experiments were thus used to query a possible interaction between Tp0956 and Tp0957 in solution. We studied a mixture of the proteins in which the Tp0956 trimer (0.7 μM) was saturated with a 24-fold molar excess of Tp0957. As is evident in Fig. 6C, all of the Tp0956 was in a complex that sedimented faster than Tp0956 alone. The 9.3-S complex has an estimated molar mass of 217,000 g/mol, consistent with a single trimer of Tp0956 interacting with three monomers of Tp0957, forming a 1:3 complex (theoretical molar mass = 214,560 g/mol). A mixture of Tp0956 and Tp0957 also formed a co-eluting complex in size-exclusion chromatography experiments (not shown). As a negative control to demonstrate specificity, we also sedimented Tp0956 with the treponemal lipoprotein Tp0655, the SBP of an ABC-type polyamine transporter [16], and there was no evidence of an interaction between the two (Fig. S4). Further, TDE 1020, the *T. denticola* homolog of Tp0957, does not interact well with Tp0956, demonstrating species specificity in this interaction (Fig. S4).

We then used SV to examine in more detail the strength of the interaction between Tp0956 and Tp0957, exploiting the exquisite sensitivity of the time-average sedimentation velocity of the reaction boundary to the complex size and the fractional dissociation of its components. We analyzed data from mixtures of Tp0956 and Tp0957 at a range of concentrations with a model directly fitted to the SV data (Figs. 5D & S6) given the reaction kinetics, sedimentation coefficients, and three macroscopic equilibrium association constants associated with an A + B + B + B ↔ AB + B + B ↔ ABB + B ↔ ABBB reaction scheme (see Materials & Methods; here, A ≡ Tp0956 trimer and B ≡ Tp0957). Because all of the binding sites on trimeric Tp0956 are identical, the model includes the K_d of the first binding event ($K_d(1)$) and cooperativity factors (α-factors) for adding the second (α_2) and third (α_3) molecules of B (an $\alpha > 1$ suggests positive cooperativity and *vice versa*). Excellent global fits to the raw data were achieved (Fig. 6D & S6), demonstrating that the first binding event is strong ($K_d(1) = K_A(1)^{-1} = 60$ [50, 80] nM; the numbers in square brackets indicate the 68.3% error interval), and that a slight positive cooperativity may be associated with the second event ($\alpha_2 = 1.5$ [1, 3]). However, there is a small, but statistically significant (on a

99.3% confidence interval) negative cooperativity (α_3 = 0.34 [0.23, 0.51]) for binding the third Tp0957 to the Tp0956 trimer.

In vivo evidence for interaction between native Tp0956 and Tp0957

The strong, specific interaction between the two recombinant proteins *in vitro* prompted us to demonstrate an association between native Tp0956 and Tp0957 *in vivo*. We therefore used *in vivo* formaldehyde cross-linking of *T. pallidum* to preserve putative Tp0956-Tp0957 complexes. Formaldehyde is a short cross-linker that covalently links primary amines within ~2 Å of each other (reflecting protein-protein interactions as they occur *in vivo*). The link is reversible and used to preserve *in vivo* protein-protein complexes. In this study, we chemically cross-linked treponemes with 1% formaldehyde. Lysed cells were resolved in duplicate on two identical SDS polyacrylamide gels, electroblotted, and then individually probed with either an anti-Tp0956 or an anti-Tp0957 antibody (Fig. 9). As a negative control, immunodetection was also performed using pre-immune serum collected from the same rats later immunized with antigens (data not shown). In the presence of cross-linker, both blots featured a strong band at a position consistent with a 250 kDa protein or complex (Lanes 2). The molecules in this band were recognized by both the Tp0956 and Tp0957 antibodies, providing evidence for the occurrence of an *in vivo* complex. The blots were then stripped and re-probed with the negative control antibody raised against Tp47 (the most abundant periplasmic lipoprotein in *T. pallidum*). Nonspecific cross-linked intermediates of native Tp47 were not observed with either native Tp0956 or Tp0957 (data not shown). The broader profiles of cross-linked species may be caused by inefficient dissociation of the complex in the presence of SDS. As shown in Fig. 9, the large complexes formed upon formaldehyde-treatment were no longer detected after boiling; only Tp0956 and Tp0957 monomers were observed (Lanes 3). For a number of reasons, including the unequal electrotransfer of small *versus* large complex proteins, the signals on the blot do not give an accurate representation of the amount of target proteins (especially in the case of cross-linked complexes of ~250 kDa). Nevertheless, this result strongly supports the conclusion that the large entities (Fig. 9, Lanes 2) are reversible complexes of Tp0956 and Tp0957. Whereas the majority of Tp0956 protein is found at the apparent molecular mass of ~250 kDa, Tp0956-specific bands corresponding to molecular masses of 35, 75 and 120 kDa arose, indicating partial dissociation of native complexes (Tp0956-Tp0957) into Tp0956 trimer, dimer and monomer in the presence of SDS upon incubation at 65°C. The smear observed in the cross-linked sample developed using anti-Tp0957 antibody also warrants comment. Some of the broader bands might correspond to Tp0957 cross-linked to as yet unknown physiological partner(s) or possibly to its operonic partner, Tp0958 (however, we do not possess Tp0958 antibodies to substantiate this second possibility). A control experiment was conducted in which treponemes were incubated under cross-linking conditions without the addition of formaldehyde. In this control, only uncomplexed Tp0956 and Tp0957 were observed (Lanes 1). Based on these results, we contend that there is a bona fide interaction between native Tp0956 and Tp0957 *in vivo*.

Tp0956 and Tp0957 homologs in other microorganisms

A position-specific iterated BLAST search [42] of genomic databases for proteins similar to Tp0956 returned only 38 proteins that may reliably be considered homologous (see Table 2). For comparison, there are 756 P-protein homologs currently collected in a database of known TRAP system proteins [43]. Homologous proteins are clearly present in the genus *Treponema*, but are also found in other spirochetes, as well as in β-, γ-, and δ-proteobacteria. No eukaryotic homologs were found. The treponemes are the only human pathogens that were found to harbor Tp0956 homologs. One Tp0956-homolog-containing organism, *Bacteriovorax marinus*, parasitizes other bacteria [44;45]. The remaining proteins were from

diverse free-living oligotrophic bacteria. Remarkably, roughly one third of the 35 species harboring Tp0956-like proteins appear to metabolize saturated or aromatic hydrocarbons.

We examined the genome maps for all organisms whose genome contains a Tp0956 homolog. Strikingly, 92% of the Tp0956 homologs appear to be components of TRAP-Ts (Table 2). Of those that do not, the Tp0956-like gene is directly adjacent to a putative P protein of a TRAP system 67% of the time; i.e., the Q and M components are missing, perhaps encoded elsewhere in the genome. These correlations strongly suggest that the Tp0956 homologs are carrying out an essential function for their respective TRAP-T. That function probably involves binding to the P component. Indeed, a preliminary study shows that the Tp0956- and Tp0957-like proteins (Tde1020 and Tde1021, respectively) from *T. denticola* interact (Fig. S5). Given the structural and functional characterizations delineated above, the Tp0956 homologs likely perform an auxiliary function in the transport of the target ligand. We name these cTPR-containing proteins the “T components”.

Do the TRAP-Ts with T components have other features that distinguish them from canonical TRAP systems? To begin to answer this question, we compared the sequences of 218 P proteins, including all of those that neighbor T components. The resulting phylogenetic tree (Figs. 10 & S7) demonstrates that the Ps associated with T-containing TRAP-Ts (P_Ts) form a distinct clade; i.e. they are evolutionarily more related to each other than to other Ps (Fig. 10). Further, as noted above, most P proteins have an arginine residue at the position equivalent to A159 of Tp0957 [9]. The positively charged guanidinium moiety of this residue is critical for the binding of the carboxylate moiety of the small carboxylic acids that are normally bound by P proteins. This arginine is found in only 3 of the 36 P_Ts that we have characterized to date. The two most likely explanations for this phenomenon are (1) the P_Ts bind to chemically different compounds or (2) they bind to chemically similar compounds in a manner that differs from most P proteins. Regardless, we conclude that the transport systems harboring a T component are a distinct subfamily of TRAP-Ts that we call the TPATs (TPR-protein-associated TRAP transporters). We suggest that the operon take on the prefix “*tat*”, standing for “TPR-protein-associated transporter”. Thus the treponemal operon would be called *tatTPQ/M*.

Discussion

This investigation has uncovered a new subfamily of TRAP-Ts, the TPATs (Fig. 10). The existence of a strong, specific interaction *in vitro* (Fig. 6C, see above) and *in vivo* (Fig. 9) between the T and P_T components suggests that the mechanism of these transporters apparently is assisted by this association. In ABC and TRAP transporters, obviously protein-protein interactions between the SBP and the transmembrane component(s) are functionally vital [3;8;46;47]. In *Halomonas elongata*, the cytoplasmic, tetrameric, ATP-binding universal stress protein TeaD of the *teaABCD* modulates the activity of an ectoine TRAP transporter (TeaA/TeaB/TeaC), presumably by affecting translation of its components or the activity of the permease[48]. However, the association of an SBP with another periplasmic protein is unprecedented.

What purpose is served by the binding of the T and P_T components of the TPATs? The full answer to this question is not known at this time. It is possible that T may have a role in delivering a ligand to P_T or in receiving a ligand from P_T. Interestingly, if the ligand does interact with T, this association may represent an additional layer of specificity for the TPAT; i.e., ligands must have chemical characteristics that are compatible with association with both the T and P_T components. Additionally, if T interacts with the ligand, it implies that the protein may be acting as a chaperone to escort the ligand through a periplasmic environment that is otherwise hostile to its stability or solubility. Recently, a similar protein-

mediated mechanism for the transport of hydrophobic compounds was proposed for an ABC-type transporter [49].

What is the ligand for this TPAT system in *T. pallidum*? Three lines of evidence suggest that the ligand is hydrophobic: (1) the existence of the hydrophobic antechamber (HA) of the T component (Fig. 3C, Table S3) (2) the hydrophobic nature of the P_T component's cleft and associated cavities (Fig. 8A; Table S3), and (3) the preponderance of T-containing organisms that utilize hydrophobic compounds (Table 2). While these oligotrophs can metabolize essentially insoluble petroleum hydrocarbons, treponemes probably acquired or retained the TPATs during their evolution to import required hydrophobic ligands that they encounter in humans.

This study represents only the beginning of the characterization of the TPATs. It will be necessary to discern the ligands of these transporters; such research currently remains infeasible for *T. pallidum*, owing to the organism's need to be cultivated in rabbit testicles. The nature of the ligand may explain the need for the T component and may also provide new seminal clues concerning what "nutrients" are needed ultimately to achieve the *in vitro* culture of *T. pallidum*. Also, knowledge of whether the ligand interacts with and passes through the T component's pore is imperative for a full understanding of the system. Finally, additional biophysical, biochemical, and biological studies of the interactions between the T, P_T, Q, and M (or Q/M) components are warranted to better comprehend the mechanism(s) of this new subfamily of bacterial transporters. Further examinations of this system might be better accomplished in a different bacterium, such as *T. denticola*, which can replicate in culture. Additionally, information regarding ligand variability in the TPAT P components should be available from a more comprehensive study of the TPAT proteins from diverse bacterial species (Table 2).

Materials and Methods

Protein expression and purification

To produce a non-lipidated, recombinant derivative of Tp0957 in *Escherichia coli*, the DNA fragment encoding amino acid residues 7-328 (cloned without the post-translationally modified N-terminal Cys plus five other hydrophobic residues) of Tp0957 was PCR amplified using ends-specific primers. Note that the recombinant versions of both Tp0956 and Tp0957 are numbered in this report such that the processed, modified cysteine residue is Residue 1. The PCR product was subcloned into the pGEM-T Easy vector (Promega) and transformed into *E. coli* XL1-Blue cells (Stratagene). The plasmids isolated from colonies that tested positive by restriction digestion were verified by DNA sequencing. The inserted fragment was excised by digestion with NotI and BamHI and then ligated into the corresponding restriction sites of the expression vector pIVEX2.4d vector (5 Prime). The resultant construct encoded a fusion protein with a His_6-tag at its N-terminus. Ligation mixtures were transformed into *E. coli* XL1-Blue cells, and plasmids isolated from colonies were verified by DNA sequencing. A verified plasmid was then cotransformed with pGroESL (TAKARA) 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, 100 μg/mL of ampicillin and 30 μg/mL of chloramphenicol until the cell density reached an A_{600} of 0.5. The culture was then induced for 3 h with 0.2% (w/v) L-arabinose. Cells derived from one liter of culture were harvested by centrifugation and lysed at room temperature with gentle rocking for 30 min using 50 mL of B-PER II (Thermo Scientific). The resulting suspension was centrifuged at 25,000 x g for 15 min to remove cell debris. Tp0957 was isolated from the supernatant by affinity chromatography using Ni-NTA Agarose (Qiagen). The protein was then subjected to size-exclusion chromatography (SEC)

using a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) equilibrated with Buffer A (20 mM Hepes, 0.1 M NaCl, pH 7.5, 2 mM *n*-octyl β-D-glucoside (β-OG)). Peak fractions were analyzed by SDS-PAGE. Fractions containing purified rTp0957 were pooled and stored at 4° C in buffer A. After two weeks, mass spectrometry and N-terminal sequencing determined that the tag of rTp0957 was specifically cleaved, leaving two residual amino acid residues (Gly-Arg) at the N-terminus followed by residues 7-328 of Tp0957. The mixture was further purified by SEC using buffer A. This purified, cleaved Tp0957 was used for co-crystallization with Tp0956.

For the production of a non-lipidated, recombinant derivative of Tp0956 in *E. coli*, a DNA fragment encoding residues 2-302 (cloned without the post-translationally modified N-terminal Cys) was amplified by PCR from treponemal genomic DNA using primer pairs encoding the predicted 5′-and 3′-termini. The PCR product was subcloned into BsaI/XbaI digested pE-SUMOpro3-based bacterial expression vector (LifeSensors) in frame with an N-terminal His-SUMO tag. The construct was confirmed by DNA sequencing. The recombinant protein was overexpressed in *E. coli* BL21(DE3) cells. The selenomethionine-substituted protein was overexpressed in the *E. coli* methionine auxotroph B834(DE3). Protein expression was induced using an overnight express autoinduction system 2 (Novagen). Cells harvested from one litre of induced culture were lysed at room temperature for 30 min using 50 ml of B-PER II (Thermo Scientific) with gentle rocking. Lysed cells were centrifuged at 25,000 x g for 15 min to remove cell debris and soluble recombinant Tp0956 was immobilized on a 4 mL Ni^{2+} affinity column (Qiagen). The immobilized protein was then washed with 10 column volumes of 20 mM Tris-HCl, 20 mM NaCl, 20 mM imidazole, pH 8.5 (buffer B), followed by 10 column volumes of 20 mM Tris-HCl, 1 M NaCl, pH 8.5 and 5 column volumes of buffer B. The protein was eluted with 3 column volumes of 20 mM Tris-HCl, 200 mM NaCl, 200 mM imidazole, pH 8.5. The eluted protein was then subjected size-exclusion chromatography (SEC) using a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) equilibrated with Buffer A (20 mM Hepes, 0.1 M NaCl, pH 7.5, 2 mM β -OG). The SEC was performed at 4°C using a FPLC system (GE Healthcare). The His-SUMO tag was removed by the SUMO-specific protease 2 (LifeSensors), and the digestion mixture was further purified by SEC as above. The selenomethionine variant of Tp0956 was purified in the presence of 2 mM DTT.

Protein concentrations were determined spectrophotometrically from their extinction coefficients calculated using the ProtParam utility of ExPASy (http://us.expasy.org).

RNA isolation and RT-PCR

Treponemal RNA extraction and RT-PCR methods were described previously [13;14]. The multi-gene operon was examined by RT-PCR using RNA isolated from *T. pallidum* that had been extracted from rabbit tissue. Intergenic regions were amplified to verify that the genes are cotranscribed in one polycistronic mRNA. cDNA and cDNA negative control were used as templates. The latter ensured that no chromosomal DNA was carried out over to cDNA preparation. The PCR amplification was carried out with GoTaq DNA polymerase (Promega) with a standard protocol and the primers listed in Table S1.

Analytical Ultracentrifugation

All AUC experiments were performed in an Optima XL-I centrifuge (Beckman-Coulter, Fullerton, CA). The samples, which had been incubated overnight at 4° C, were placed in centrifugation cells equipped with dual-sectored Epon centerpieces and sapphire windows inserted in an An50-Ti rotor. The samples were centrifuged at speeds between 45,000 and 50,000 rpm at 20° C after they had equilibrated for several hours. Both absorbance and interference optics were used to acquire the concentration profiles.

The *c*(*s*) analyses were carried out using SEDFIT and SEDPHAT [50]. All sedimentation coefficients were converted to $s_{20,w}$-values. These distributions typically were regularized using the maximum-entropy approach with a confidence level of 0.7. Alternatively, to evaluate the binding equilibria, Tp0956:Tp0957 mixture data were globally analyzed [51;52] by direct modeling with a set of Lamm equations describing the coupled sedimentation/diffusion/reaction process for the three-site scheme with macroscopic equilibrium constants:

$$K_A(1) = \frac{[\mathrm{AB}]}{[\mathrm{A}][\mathrm{B}]}$$

$$K_A(2) = \frac{[\mathrm{ABB}]}{[\mathrm{AB}][\mathrm{B}]}$$

$$K_A(3) = \frac{[\mathrm{ABBB}]}{[\mathrm{ABB}][\mathrm{B}]}$$

where A≡ Tp0956 trimer and B≡Tp0957 monomer. Because the binding sites for B on A are identical, statistical factors govern that $K_A(2)/K_A(1) = 1/3$, and $K_A(3)/K_A(1) = 1/9$ in the absence of cooperativity. Thus, the equilibrium was treated as a system of equations with

$$K_A(2) = 1/3\alpha_2 K_A(1) \quad \text{and}$$

$$K_A(3) = 1/9\alpha_3 K_A(1)$$

In the nonlinear regression, in addition to $K_A(1)$, the factors α_2 and α_3 were allowed to refine to account for cooperativity. The kinetic component of all of the reactions was fixed such that the k_{off} for reaction 1 was 10^{-3} s^{-1}, and no kinetic cooperativity was taken into account. We used HYDROPRO [53] to estimate the sedimentation coefficients of the complexes using the appropriately edited complex crystal structures [54]. Because the program overestimated this value for Tp0956, we compensated by lowering the estimated complex sedimentation coefficients by about 3%. These values were also fixed during the analyses. Thus, the only refined quantities were the positions of the menisci, the exact concentrations of soluble Tp0956 and Tp0957 present in the centrifugation cells, and the association constants parameters $\log(K_A(1))$, $\log(1/3\cdot\alpha_2)$, and $\log(1/9\cdot\alpha_3)$.

In vivo *formaldehyde cross-linking*

Protein cross-linking was performed according to the method of Skare *et al.* [55]. Approximately 1×10^8 freshly harvested *T. pallidum* (prepared as described above for RNA extraction) in phosphate-buffered saline (PBS) was incubated with 1% (v/v) methanol-free formaldehyde (Thermo Scientific) for 2 h at 25°C with gentle rocking on a nutating mixer. To stop the cross-linking reaction, 1.25 M glycine was added to a final concentration of 125 mM (5 min at 25°C). A control reaction was performed similarly, but without adding formaldehyde. Both cross-linked and uncross-linked treponemes were harvested by centrifugation at 16,000 x *g* and rinsed twice with ice-cold PBS to remove excess unreacted cross-linkers, and frozen at −70°C until needed for analysis.

Antibodies, SDS PAGE, electroblotting and immunodetection

Antisera recognizing Tp0956 and Tp0957 were obtained from rats by injecting purified recombinant antigens. Antibodies from the antisera were then affinity purified using an antibody purification kit (Thermo Scientific). The specificity of the purified antibodies was documented against recombinant test and control antigens (not shown).

For SDS PAGE analysis, ~1×10^8 treponemes were lysed in 50 μl of 2X SDS Laemmli sample buffer containing 500 mM β-mercaptoethanol. Cross-linked samples were either incubated at 65°C for 5 min (to maintain the cross-links) or boiled for 20 min (to break the chemical cross-links) prior to SDS PAGE analysis on a 4-15% (w/v) polyacrylamide TGX precast gel (BioRad). The control sample was boiled for 20 min. Identical numbers of cells (~5×10^7 per lane or 25 μl per lane) were loaded within each experiment. After gel electrophoresis, the proteins were electrotransferred onto a nitrocellulose membrane and probed with antigen-specific antibodies using a standard Western blot protocol. The primary antibodies were detected using anti-goat IgG-HRP antibodies (Jackson ImmunoResearch) and visualized by an enhanced chemiluminescent reaction system (Thermo Scientific) and exposure to a Fuji luminescent image analyzer LAS-3000.

Crystallization and X-ray diffraction data collection

All crystals were obtained using the hanging-drop vapor diffusion method. Crystals of selenomethionine-substituted Tp0956 were grown at 20° C from solutions containing 4 μL protein (40 mg/mL in buffer A containing 2 mM dithiothreitol) and 4 μL of reservoir solution (30% (v/v) PEG-400, 0.1 M Bicine, pH 9.0, 0.2 M $MgCl_2$) that were equilibrated over 0.5 mL of reservoir solution. Cryoprotection was performed by transferring the crystals first to a stabilization solution of 30% PEG-400, 0.1 M Bicine, pH 9.0, 0.2 M $MgCl_2$, 0.1 M NaCl, then serially transferring them to similar solutions of increasing [PEG-400]; the final [PEG-400] was 40%. The crystals were flash-cooled in liquid nitrogen. Native crystals of Tp0956 were obtained and treated similarly, except as noted below. A 4-μL drop of the protein solution (13 mg/mL in buffer A) was mixed with 4 μL of a different reservoir solution (1.26 M ammonium sulfate, 0.1 M CHES, pH 9.5, 0.2 M NaCl). The stabilization solution was 1.36 M ammonium sulfate, 5% (v/v) ethylene glycol, 0.1 M CHES, pH 9.0, and 0.3 M NaCl, and the final cryostabilization solution was 1.36 M ammonium sulfate, 15% (v/v) ethylene glycol, 0.1 M CHES, pH 9.0, and 0.3 M NaCl. All X-ray diffraction data in this report (Table 1) were obtained at 100 K at beamline 19-ID at the Structural Biology Center at the Advanced Photon Source in Argonne National Laboratories. Tp0956 crystals exhibited the symmetry of space group I23 and contained one molecule of Tp0956 per asymmetric unit. Data were indexed, integrated and scaled using the HKL-3000 program package [56]. Data from selenomethionine-substituted crystals were indexed in the lower symmetry space group C2.

Tp0957 was crystallized at 20°C. Drops comprised 4 μL protein solution (12 mg/mL in Buffer A) and 4 μL of reservoir solution (0.2 M KSCN, 20% (w/v) polyethylene glycol 3350). Parallelepiped-shaped crystals appeared within four days and varied widely in size. The crystal used for data collection was approximately 0.1 × 0.1 × 0.25 mm. Crystals were stabilized by transferring them to a solution consisting of 0.1 M Hepes pH 7.5, 0.2 M KSCN, 2 mM β-OG, 20% PEG 3350, and 5% (v/v) ethylene glycol. Stabilized crystals were serially transferred into similar solutions that contained increasing concentrations of ethylene glycol; the final [ethylene glycol] was 25%. After about 5 min. in this final solution, the crystals were flash-cooled in liquid nitrogen.

The crystals, which had the symmetry of space group $P2_1$, were diffracted as noted above for Tp0956. The X-ray diffraction data were indexed, integrated, and scaled using the same program and procedures as used for Tp0956.

Phase determination, structure refinement & analysis

Phases for Tp0956 were obtained from a single-wavelength anomalous diffraction experiment using selenomethionine-substituted protein (Table 1). SHELXD [57] was used to locate the selenium atoms. Phases were refined with MLPHARE [58] and were further improved by density modification with DM [59]. An initial model containing 82% of all Tp0956 residues was generated by alternating cycles of Resolve [60] and ARP/wARP (Langer, 2006). Additional residues were manually modeled in O [61]. Refinement (Table 1) was performed with the data collected on a native crystal using PHENIX [62].

The Tp0957 structure was determined using molecular replacement. The search model was that of Tp0957 that had been obtained from a co-crystal structure of the Tp0956:Tp0957 complex (C.A.B., R.K.D., and M.V.N, unpublished results). Two molecules of Tp0957 were located in the asymmetric unit of the Tp0957 crystals using Phaser [63]. The model was subjected to Cartesian simulated-annealing, positional, and individual B-factor refinement in PHENIX. At this point, missing parts of the molecules could be built using Coot [64]. The model refinement was finalized using PHENIX; riding hydrogens were added to the model in the final stages of refinement. Electrostatic calculations were performed with APBS [65], using a protein dielectric constant of 1, and monovalent ion concentrations of 0.15 M. All structure figures were made using the PyMOL version 1.2 (Schrödinger, LLC). For both proteins, MolProbity was used to verify the structures[66].

Bioinformatics

The amino-acid sequence of Tp0956 was subjected to PSI-BLAST [42;67]. Five iterations were performed. We discarded “hits” with *E*-values above 1e-32, because such proteins were deemed to be non-orthologous (other TPR-containing proteins). The sequences of 218 P proteins were selected for comparison. The selection criteria were:

1. All P proteins with Tp0956-like genetic neighbors were included ($n = 36$).
2. All P proteins with known crystal structures were included ($n = 8$).
3. The remaining proteins ($n = 175$) were selected from the TRAP database [43]. One protein was included from each organism present. In cases where more than one P protein was present in an organism, one was chosen at random.

PROMALS-3D [68] was used to align the sequences; the structures of the proteins from criterion 2 above were used to provide structural constraints for the algorithm (the structure of Tp0957 was not included in this group). The unedited alignment was converted to PHYLIP format, and the phylogenetic tree was calculated using PhyML [69]; the LG model of amino-acid substitution [70] was employed. The tree was viewed and output using Archaeopteryx (C.M. Zmasek, www.phylosoft.org).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Zhiming Ouyang for technical assistance with the RT-PCR analyses, the scientists in the UT Southwestern Protein Chemistry Core for protein-sequence and mass analyses, and Dr. Lisa Kinch for helpful comments on the bioinformatics. This research was supported by an NIH grant (AI056305) and a Welch

Foundation grant (I-0940) to M.V.N. This work was also supported in part by the Intramural Research Program of the National Institute of Biomedical Imaging and Bioengineering (P.S.). 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.

Abbreviations

AUC	analytical ultracentrifugation
β-OG	*n*-octyl β-D-glucoside
COG	cluster of orthologous group
CDD	conserved domain database
HA	hydrophobic antechamber
SBP	substrate-binding protein
SEC	size-exclusion chromatography
SV	sedimentation velocity
TRAP-T	tripartite ATP-independent periplasmic transporter
TPR	tetratricopeptide repeat
TPAT	TPR-protein associated TRAP transporter

References

1. Fraser CM, Norris SJ, Weinstock GM, White O, Sutton GG, Dodson R, Gwinn M, Hickey EK, Clayton R, Ketchum KA, Sodergren E, Hardham JM, McLeod MP, Salzberg S, Peterson J, Kalak H, Richardson D, Howell JK, Chidambaram M, Utterback T, McDonald L, Artiach P, Bowman C, Cotton MD, Fujii C, Garland S, Hatch B, Horst K, Roberts K, Sandusky M, Weidman J, Smith HO, Venter JC. Complete genome sequence of Treponema pallidum, the syphilis spirochete. Science. 1998; 281:375–388. [PubMed: 9665876]
2. Norris SJ. Polypeptides of Treponema pallidum: progress toward understanding their structural, functional, and immunologic roles. Microbiological Reviews. 1993; 57:750–779. [PubMed: 8246847]
3. Davidson AL, Maloney PC. ABC transporters: how small machines do a big job. Trends in Microbiology. 2007; 15:448–455. [PubMed: 17920277]
4. Abramson J, Wright EM. Structure and function of Na^+-symporters with inverted repeats. Current Opinion in Structural Biology. 2009; 19:425–432. [PubMed: 19631523]
5. Kelly DJ, Thomas GH. The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. FEMS Microbiology Reviews. 2001; 25:405–424. [PubMed: 11524131]
6. Forward JA, Behrendt MC, Wyborn NR, Cross R, Kelly DJ. TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of Rhodobacter capsulatus and by homologs in diverse gram-negative bacteria. Journal of Bacteriology. 1997; 179:5482–5493. [PubMed: 9287004]
7. Shaw JG, Hamblin MJ, Kelly DJ. Purification, characterization and nucleotide sequence of the periplasmic C4-dicarboxylate-binding protein (DctP) from Rhodobacter capsulatus. Molecular Microbiology. 1991; 5:3055–3062. [PubMed: 1809844]
8. Mulligan C, Geertsma ER, Severi E, Kelly DJ, Poolman B, Thomas GH. The substrate-binding protein imposes directionality on an electrochemical sodium gradient-driven TRAP transporter. Proceedings of the National Academy of Sciences, USA. 2009; 106:1778–1783.
9. Fischer M, Zhang QY, Hubbard RE, Thomas GH. Caught in a TRAP: substrate-binding proteins in secondary transport. Trends in Microbiology. 2010; 18:471–478. [PubMed: 20656493]

10. Winnen B, Hvorup RN, Saier MH Jr. The tripartite tricarboxylate transporter (TTT) family. Research in Microbiology. 2003; 154:457–465. [PubMed: 14499931]
11. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Research. 2011; 39:D225–D229. [PubMed: 21109532]
12. Saïd-Salim B, Mostowy S, Kristof AS, Behr MA. Mutations in Mycobacterium tuberculosis Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in Mycobacterium bovis. Molecular Microbiology. 2006; 62:1251–1263. [PubMed: 17064366]
13. Deka RK, Brautigam CA, Tomson FL, Lumpkins SB, Tomchick DR, Machius M, Norgard MV. Crystal Structure of the Tp34 (TP0971) lipoprotein of Treponema pallidum: implications of its metal-bound state and affinity for human lactoferrin. Journal of Biological Chemistry. 2007; 282:5944–5958. [PubMed: 17192261]
14. Deka RK, Brautigam CA, Yang XF, Blevins JS, Machius M, Tomchick DR, Norgard MV. 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. Journal of Biological Chemistry. 2006; 281:8072–8081. [PubMed: 16418175]
15. Deka RK, Machius M, Norgard MV, Tomchick DR. Crystal structure of the 47-kDa lipoprotein of Treponema pallidum reveals a novel penicillin-binding protein. Journal of Biological Chemistry. 2002; 277:41857–41864. [PubMed: 12196546]
16. Machius M, Brautigam CA, Tomchick DR, Ward P, Otwinowski Z, Blevins JS, Deka RK, Norgard MV. Structural and biochemical basis for polyamine binding to the Tp0655 lipoprotein of Treponema pallidum: putative role for Tp0655 (TpPotD) as a polyamine receptor. Journal of Molecular Biology. 2007; 373:681–694. [PubMed: 17868688]
17. D'Andrea LD, Regan L. TPR proteins: the versatile helix. Trends in Biochemical Sciences. 2003; 28:655–662. [PubMed: 14659697]
18. Sampathkumar P, Roach C, Michels PAM, Hol WGJ. Structural Insights into the Recognition of Peroxisomal Targeting Signal 1 by Trypanosoma brucei Peroxin 5. Journal of Molecular Biology. 2008; 381:867–880. [PubMed: 18598704]
19. Severi E, Randle G, Kivlin P, Whitfield K, Young R, Moxon R, Kelly DJ, Hood D, Thomas GH. Sialic acid transport in Haemophilus influenzae is essential for lipopolysaccharide sialylation and serum resistance and is dependent on a novel tripartite ATP-independent periplasmic transporter. Molecular Microbiology. 2005; 58
20. Setubal JC, Reis M, Matsunaga J, Haake DA. Lipoprotein computational prediction in spirochaetal genomes. Microbiology. 2006; 152:113–121. [PubMed: 16385121]
21. McPherson A, Koszelak S, Axelrod H, Day J, Williams R, Robinson L, McGrath M, Cascio D. An experiment regarding crystallization of soluble proteins in the presence of beta-octyl glucoside. Journal of Biological Chemistry. 1986; 261:1969–1975. [PubMed: 3944122]
22. Löwe J, Stock D, Jap B, Zwickl P, Baumeister W, Huber R. Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. Science. 1995; 268:533–539. [PubMed: 7725097]
23. Lima CD, Wang JC, Mondragón A. Three-dimensional structure of the 67K N-terminal fragment of E. coli DNA topoisomerase I. Nature. 1994; 367:138–146. [PubMed: 8114910]
24. Tajkhorshid E, Nollert P, Jensen MO, Miercke LJW, O'Connell J, Stroud RM, Schulten K. Control of the selectivity of the aquaporin water channel family by global orientational tuning. Science. 2002; 296:525–530. [PubMed: 11964478]
25. Sui H, Han B-G, Lee JK, Walian P, Jap BK. Structural basis of water-specific transport through the AQP1 water channel. Nature. 2001; 414:872–878. [PubMed: 11780053]
26. Fu D, Libson A, Miercke LJW, Weitzman C, Nollert P, Krucinski J, Stroud RM. Structure of a glycerol-conducting channel and the basis for its selectivity. Science. 2000; 290:481–486. [PubMed: 11039922]

27. Zeth K, Diederichs K, Welte W, Engelhardt H. Crystal structure of Omp 32, the anion-selective porin form Comamonas acidovorans, in complex with a periplasmic peptide at 2.1 A resolution. Structure. 2000; 8:981–992. [PubMed: 10986465]
28. Holm L, Rosenstrōm P. Dali server: conservation mapping in 3D. Nucleic Acids Research. 2010; 38:W545–W549. [PubMed: 20457744]
29. Lim H, Kim K, Han D, Oh J, Kim Y. Crystal structure of TTC0263, a thermophilic TPR protein from Thermus thermophilus HB27. Molecules and Cells. 2007; 24:27–36. [PubMed: 17846496]
30. Smith DF. Tetratricopeptide repeat cochaperones in steroid receptor complexes. Cell Stress & Chaperones. 2004; 9:109–121. [PubMed: 15497498]
31. Vodermaier HC, Gieffers C, Maurer-Stroh S, Eisenhaber F, Peters J-M. TPR Subunits of the Anaphase-Promoting Complex Mediate Binding to the Activator Protein CDH1. Current Biology. 2003; 13:1459–1468. [PubMed: 12956947]
32. Biegert A, Mayer C, Remmert M, Soding J, Lupas AN. The MPI Bioinformatics Toolkit for protein sequence analysis. Nucleic Acids Research. 2006; 34:W335–W339. [PubMed: 16845021]
33. Krissinel E, Henrick K. Inference of Macromolecular Assemblies from Crystalline State. Journal of Molecular Biology. 2007; 372:774–797. [PubMed: 17681537]
34. 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. Journal of Biological Chemistry. 2008; 283:32812–32820. [PubMed: 18723845]
35. Gonin S, Arnoux P, Pierru B, Lavergne J, Alonso B, Sabaty M, Pignol D. 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 Structural Biology. 2007; 7:11. [PubMed: 17362499]
36. Tam R, Saier MH Jr. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. Microbiological Reviews. 1993; 57:320–346. [PubMed: 8336670]
37. Müller A, Severi E, Mulligan C, Watts AG, Kelly DJ, Wilson KS, Wilkinson AJ, Thomas GH. Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from Haemophilus influenzae. Journal of Biological Chemistry. 2006; 281:22212–22222. [PubMed: 16702222]
38. Johnston JW, Coussens NP, Allen S, Houtman JCD, Turner KH, Zaleski A, Ramaswamy S, Gibson BW, Apicella MA. Characterization of the N-acetyl-5-neuraminic acid-binding site of the extracytoplasmic solute receptor (SiaP) of nontypeable Haemophilus influenzae strain 2019. Journal of Biological Chemistry. 2008; 283:855–865. [PubMed: 17947229]
39. Holm L, Park J. DaliLite workbench for protein structure comparison. Bioinformatics. 2000; 16:566–567. [PubMed: 10980157]
40. Akiyama N, Takeda K, Miki K. Crystal Structure of a Periplasmic Substrate-Binding Protein in Complex with Calcium Lactate. Journal of Molecular Biology. 2009; 392:559–565. [PubMed: 19631222]
41. Kuhlmann SI, van Scheltinga ACT, Bienert R, Kunte H-J, Ziegler C. 1.55 A structure of the ectoine binding protein TeaA of the osmoregulated TRAP-transporter TeaABC from Halomonas elongata. Biochemistry. 2008; 47:9457–9485.
42. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990; 215:403. [PubMed: 2231712]
43. Mulligan C, Kelly DJ, Thomas GH. Tripartite ATP-independent periplasmic transporters: application of a relational database for genome-wide analysis of transporter gene frequency and organization. Journal of Molecular Microbiology and Biotechnology. 2007; 12:218–226. [PubMed: 17587870]
44. Stolp H, Starr MP. Bacteriolysis. Annual Review of Microbiology. 1965; 19:79–104.
45. Baer ML, Ravel J, Piñeiro SA, Guether-Borg D, Williams HN. Reclassification of salt-water Bdellovibrio sp. as Bacteriovorax marinus sp. nov. and Bacteriovorax litoralis sp. nov. International. Journal of Systematic and Evolutionary Microbiology. 2004; 54:1011–1016.

46. Oldham ML, Khare D, Quiocho FA, Davidson AL, Chen J. Crystal structure of a catalytic intermediate of the maltose transporter. Nature. 2007; 450:515–520. [PubMed: 18033289]
47. Oldham ML, Chen J. Crystal structure of the maltose transporter in a pretranslocation intermediate state. Science. 2011; 332:1202–1205. [PubMed: 21566157]
48. Schweikhard ES, Kuhlmann SI, Kunte H.-J. r. Grammann K, Ziegler CM. Structure and Function of the Universal Stress Protein TeaD and Its Role in Regulating the Ectoine Transporter TeaABC ofHalomonas elongataDSM 2581T. Biochemistry. 2010; 49:2194–2204. [PubMed: 20113006]
49. Malinverni JC, Silhavy TJ. An ABC transport system that maintains lipid asymmetry in the Gram-negative outer membrane. Proceedings of the National Academy of Sciences. 2009; 106:8009–8014.
50. Schuck P. Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophysical Journal. 2000; 78:1606–1619. [PubMed: 10692345]
51. Brautigam CA. Using Lamm-equation modeling of sedimentation velocity data to determine the kinetic and thermodynamic properties of macromolecular interactions. Methods. 2011; 54:4–15. [PubMed: 21187153]
52. Dam J, Velikovsky CA, Mariuzza RA, Urbanke C, Schuck P. Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modeling and sedimentation coefficient distributions c(s). Biophysical Journal. 2005; 89:619–634. [PubMed: 15863475]
53. de la Torre JG, Huertas ML, Carrasco B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structures. Biophysical Journal. 2000; 78:719–730. [PubMed: 10653785]
54. Brautigam CA, Deka RK, Schuck P, Tomchick D, Norgard MV. Structural and thermodynamic characterization of the interaction between two periplasmic Treponema pallidum lipoproteins that are components of a unique ATP-independent periplasmic transporter. Accompanying paper. 2011
55. Skare JT, Ahmer BMM, Seachord CL, Darveau RP, Postle K. Energy transduction between membranes: TonB, a cytoplasmic membrane protein, can be chemically cross-linked in vivo to the outer membrane receptor FepA. Journal of Biological Chemistry. 1993; 268:16302–16308. [PubMed: 8344918]
56. 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 Crystallographica Section D Biological Crystallography. 2006; 62:859–866.
57. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. Acta Crystallographica Section D Biological Crystallography. 2002; D58:1772–1779.
58. Otwinowski Z. Maximum likelihood refinement of heavy atom parameters. Proceedings of the CCP4 Daresbury Study Weekend. 1991:80–86.
59. Cowtan K, Main P. Miscellaneous algorithms for density modification. Acta Crystallographica Section D Biological Crystallography. 1998; D54:487–493.
60. Terwilliger TC. Automated main-chain model building by template matching and iterative fragment extension. Acta Crystallographica. Section D: Biological Crystallography. 2003; 59:38–44.
61. Jones TA, Zou J-Y, Cowan SW. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallographica. Section A, Crystal Physics, Diffraction, Theoretical and General Crystallography. 1991; 47:110–119.
62. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure determination. Acta Crystallographica. 2010; D66:213–221.
63. Read RJ. Pushing the boundaries of molecular replacement with maximum likelihood. Acta Crystallographica. Section D: Biological Crystallography. 2001; 57:1373–1382.
64. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. Acta Crystallographica. Section D: Biological Crystallography. 2004; 60:2126–2132.

65. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:10037–10041. [PubMed: 11517324]

66. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Research. 2007; 35:W375–W383. [PubMed: 17452350]

67. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research. 1997; 25:3389–3402. [PubMed: 9254694]

68. Pei J, Kim B-H, Grishin NV. PROMALS3D: a tool for mulotiple sequence and structure alignment. Nucleic Acids Research. 2008; 36:2295–2300. [PubMed: 18287115]

69. 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. Systematic Biology. 2010; 59:307–321. [PubMed: 20525638]

70. Le SQ, Gascuel O. An Improved General Amino Acid Replacement Matrix. Molecular Biology and Evolution. 2008; 25:1307–1320. [PubMed: 18367465]

71. Petřek M, Otyepka M, Banás P, Kosinová P, Koča J, Damborský J. CAVER: a new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinformatics. 2006; 7:316. [PubMed: 16792811]

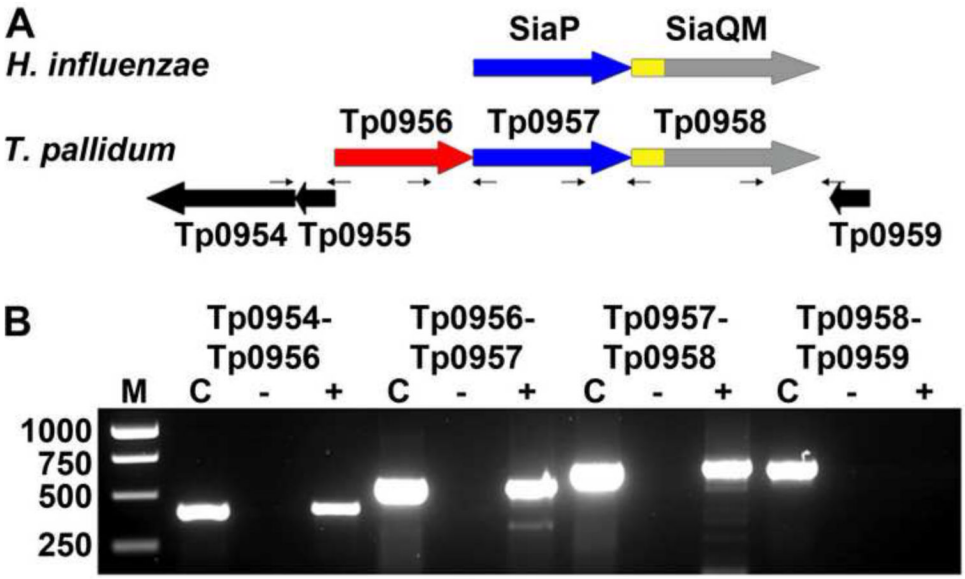

Figure 1. Transcriptional linkage of the putative treponemal TRAP operon

(A) The genomic organizations of TRAP systems from *H. influenzae* (upper) and *T. pallidum* (lower). The genes for the P components are colored blue, those for the Q components are yellow, and those for the M components are gray. These latter two components are fused in the examples shown, but not in all TRAP-Ts. (B) Evidence of transcriptional linkage of Tp0956 and Tp0957. RT PCR was performed on *T. pallidum* RNA using primer pairs (small arrows) listed in Table S1, specific for either the intergenic regions of listed gene pairs or the entire *tp0955* gene (pseudo gene). The lanes are: Lane M, DNA molecular weight markers; Lane C, PCR reaction with indicated primer pair served as positive control using *T. pallidum* genomic DNA as template in place of cDNA; Lane '-', PCR reaction with indicated primer pair using RNA as template (lacking RT), which served as a negative control for DNA contamination; Lane '+', RT PCR products with indicated primer pairs.

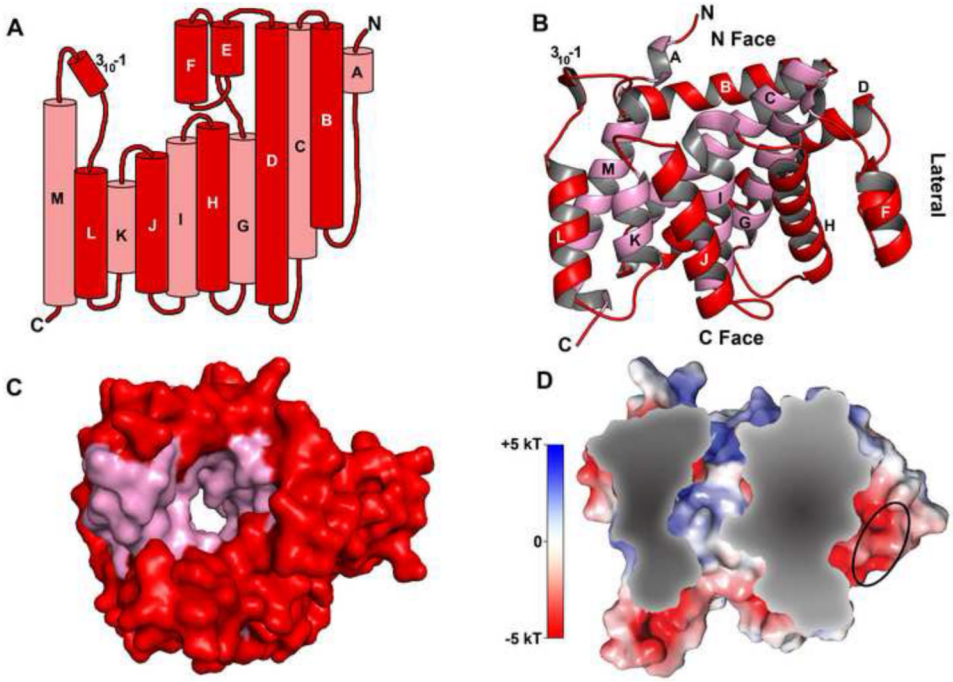

Figure 2. The crystal structure of Tp0956

(A) A topological diagram of Tp0956. The cylinders are α-helices, and are labeled with their respective alphabetical designations. The protein forms a ring with a pore running through it, but in this diagram, the ring as been opened and "flattened", with the helices most proximal to the pore shown behind the others and color-coded pink. All other helices are colored red. This color-coding holds for parts B and C as well. (B) A "side view" of the protein, with the C, lateral, and N faces denoted. The model has been rotated approximately 90° around the horizontal axis compared to that shown in (A). (C) A surface representation of the protein, viewed from the C face. (D) A cross-section through the protein, which has been colored according to its electrostatic potential. The scale for the coloration is also shown. An area of negative potential on the lateral surface of the protein is circled.

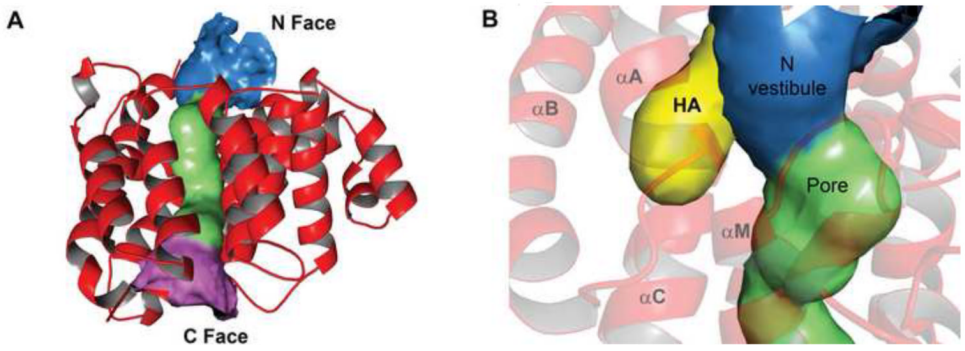

Figure 3. Pore characteristics of Tp0956

(A) A ribbons representation of the protein with a surface representation of the pore. The N, middle, and C portions of the pore are colored blue, green, and pink, respectively. (B) The hydrophobic antechamber (HA). A different view of the pore than that in (A) is shown, and the HA is depicted as a yellow surface. Secondary structure is shown semi-transparently for clarity, and the α-helices that contribute residues to the HA are labeled. The HA surface was calculated with CAVER 2.0 [71].

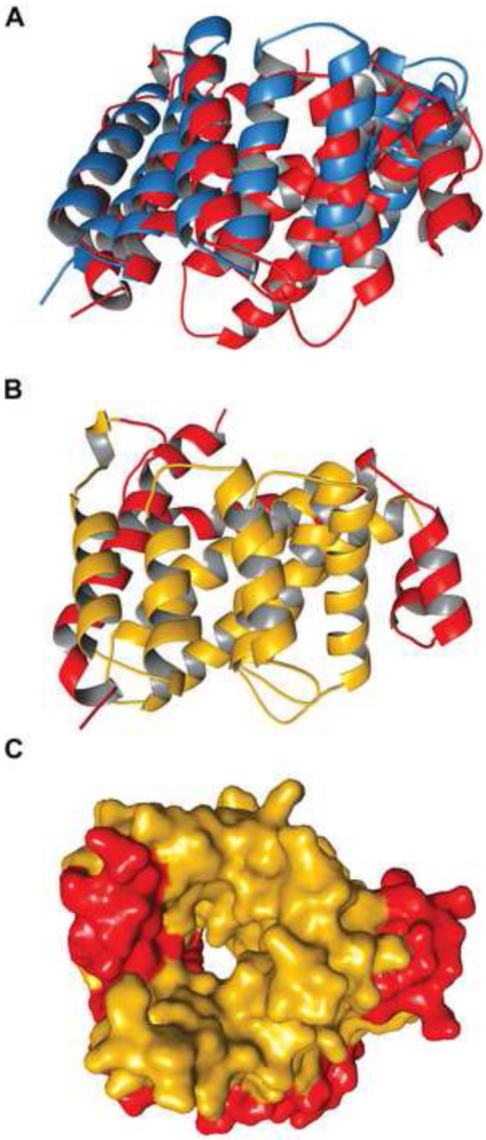

Figure 4. The cTPR motifs of Tp0956

(A) A superposition of Tp0956 (red) and DrR162B (blue). (B) The positions of the cTPR motifs. The motifs are colored gold, while the remainder of the protein is red. (C) The contributions of the cTPR motifs to the pore of Tp0956. A surface representation is shown, with the protein colored as in (B).

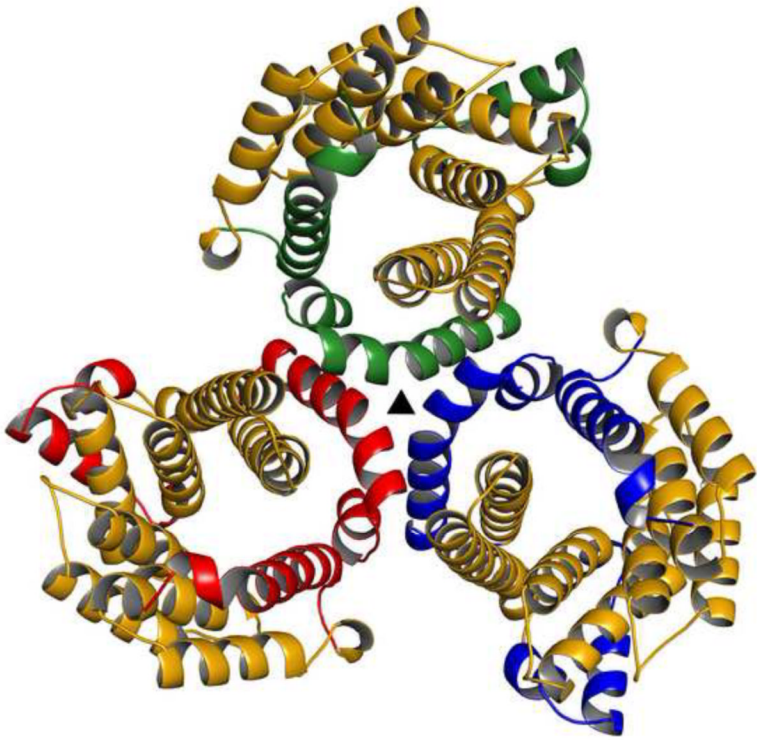

Figure 5. The trimer of Tp0956

The cTPRs are colored as in Fig. 4B, but the remainders are colored differently to distinguish the monomers. The triangular symbol represents the crystallographic three-fold axis that relates the monomers.

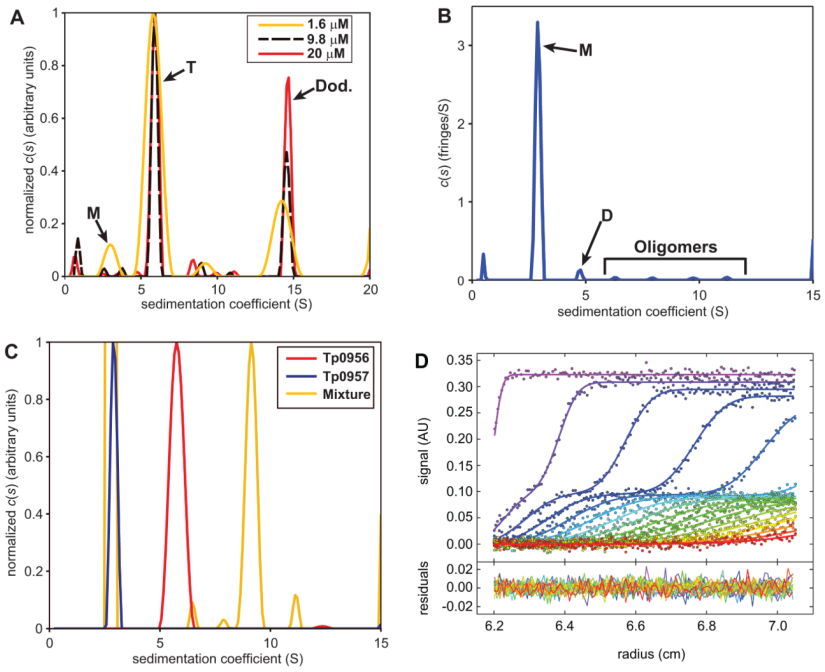

Figure 6. The solution behavior of Tp0956 and Tp0957

(A) The solution oligomerization state of Tp0956. Three *c*(*s*) distributions are shown; they have been normalized such that the maximum value in each distribution is 1. Our assignments of the oligomeric states are shown: monomer (M), trimer (T), and dodecamer (Dod.). The inset shows three concentrations used for this experiment. (B) The oligomerization status of Tp0957. The peaks for monomer (M) and dimer (D) are marked. Probable oligomers are also marked. (C) Hydrodynamic evidence for the interaction of Tp0956 and Tp0957. Again, three normalized distributions are shown. Here, the distributions were normalized such that the peak of interest would have a maximum value of one. (D) The Lamm-equation fit to the SV data for wild-type Tp0956 and Tp0957 mixtures. The concentration of Tp0956 was held constant at 0.7 μM, while that of Tp0957 was 0.7, 1.4, 2.1, 3.5, 7, 10.5, and 14 μM. Only the 3.5 μM experiment is shown. The x-axis of the graph is the distance from the center of rotation in cm, and the y-axis in the upper part is absorbance. The y-axis in the lower part is the residual between the data and the fits, also in absorbance units. The filled-in circles are individual data points, and the lines are the fits to these data; both are color-coded such that early scans are shown in violet, and late scans in red. The fits are global, i.e. a single set of parameters (see Materials & Methods) was used to fit to all seven data sets. For the sake of clarity, only every third data point and every third scan used in the analysis are shown. Additional concentrations are shown in Fig. S6.

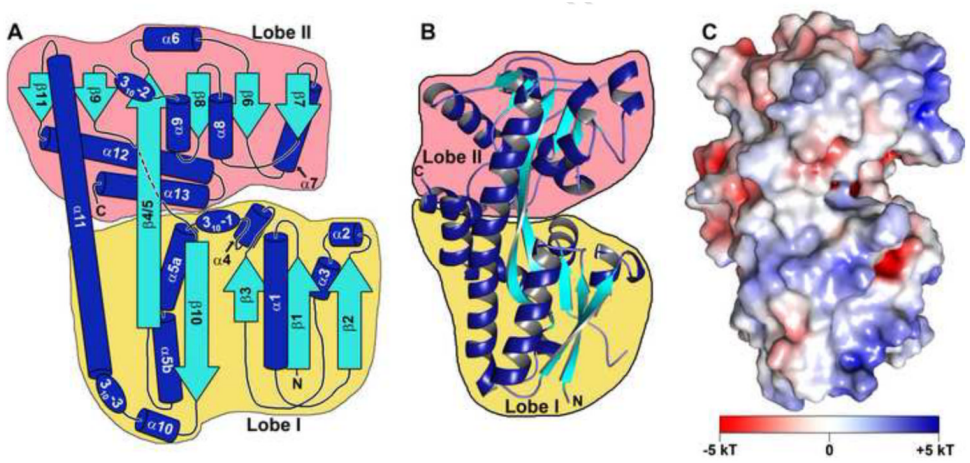

Figure 7. The topology, structure, and surface features of Tp0957

(A) A topology diagram of Tp0957. β-strands are shown as cyan arrows, while α-helices are depicted as blue cylinders and 3_{10} helices are shown as ovals. The β-strands and α-helices are shown to scale, but the connecting regions and 3_{10} helices are not. All elements are labeled with the names that will be used throughout the text. Lobe I is outlined in yellow, Lobe II in pink. The amino-and carboxyl-termini are marked with “N” and “C”, respectively. (B) A ribbons-style drawing of the structure of Tp0957. The secondary structure and lobes are color-coded as in part (A). (C) The surface electrostatic features of Tp0957. The electrostatic potential is color-coded on the surface of Tp0957. A key to the coloring scale is shown. The orientation of Tp0957 is identical to that in part (B).

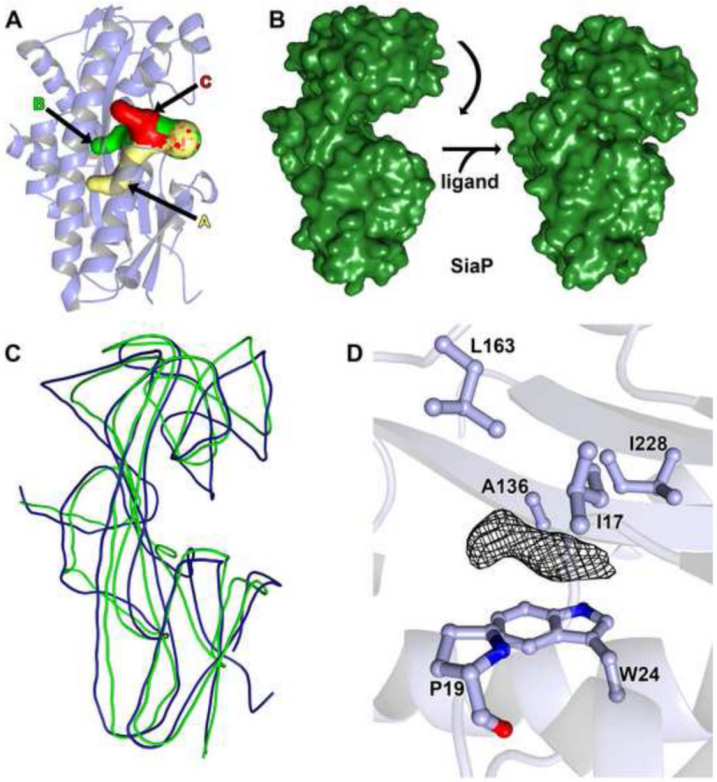

Figure 8. Features of the Tp0957 cleft

(A) Cavities proximal to the Tp0957 cleft. Cavity A is shown in cream, Cavity B in green, and Cavity C in red, and the three cavities are also labeled. All three fuse near to the surface of the protein; the mottled appearance of the surface here demonstrates that all three of the calculated paths contribute to this opening. The secondary structure of Tp0957 is shown semi-transparently for clarity. The orientation is slightly different from that drawn in Fig. 7. The surfaces were calculated using Caver 2.0 [71]. (B) Observed closure of SiaP [37]. The left side shows the unliganded form of SiaP, in the analogous orientation to that shown for Tp0957 in Fig. 7B. Upon ligand binding, the upper lobe moves down (right side), enveloping the ligand. (C) Superposition of Tp0957 and SiaP (open form). Both structures are shown as smoothed traces through their respective C_α atoms; Tp0957 is colored blue, and SiaP is colored green. (D) Unassigned electron density in Tp0957's cleft. A $2mF_o$-DF_c map contoured at the 1-σ level is shown for the density. Nearby side chains are shown and labeled. The secondary structure of Tp0957 is shown semi-transparently for clarity.

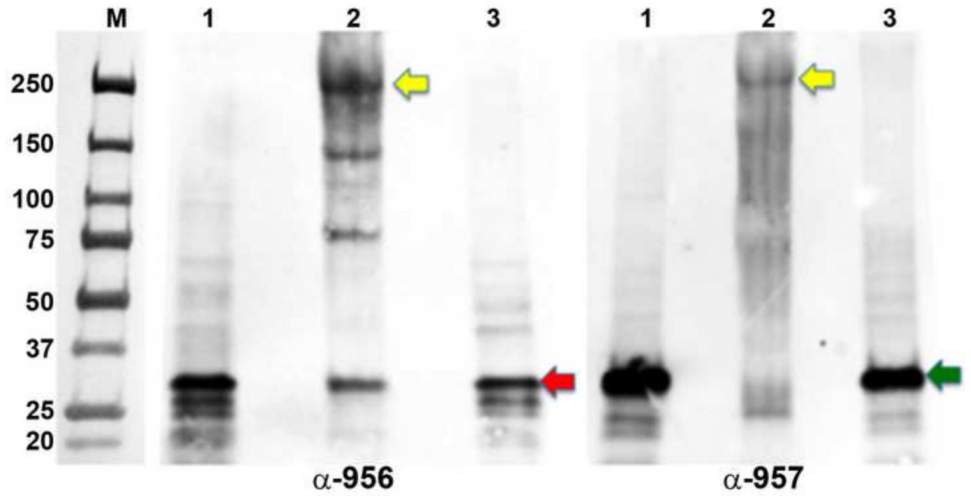

Figure 9. ***In vivo*** **formaldehyde cross-linkage of Tp0956 and Tp0957**

Samples were resolved in two identical gels and then electrotransferred to two separate membranes. The membranes were individually probed with either an anti-Tp0956 (α-Tp0956; left panel) or an anti-Tp0957 (α-Tp0957; right panel) antibody and visualized by an enhanced chemiluminescent detection using an anti-goat IgG-HRP. Lane 1 shows *T. pallidum* not exposed to the formaldehyde; lane 2, formaldehyde cross-linked treponemes, and lane 3, boiled treponemes to break the cross-links prior to SDS PAGE. The yellow arrows represent the putative positions of the cross-linked hetero-hexameric complex, while the red and green arrows denote the positions of monomeric Tp0956 and Tp0957, respectively. The positions of pre-stained molecular-weight standards (M; values are in kDa) are indicated in the image of the blot.

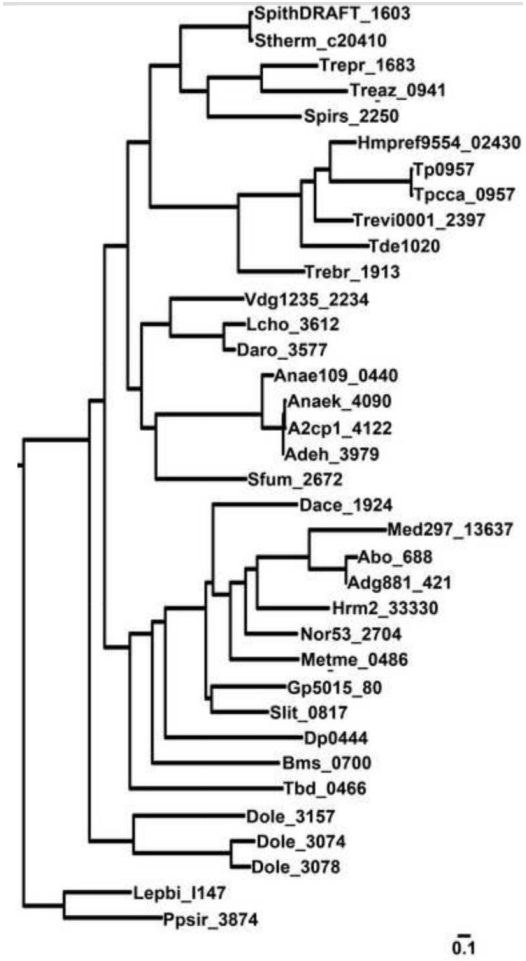

Figure 10. Phylogenetic comparison of 218 P-protein sequences

Only a portion of the entire phylogenetic tree is shown. The locus tag protein names are provided. The scale, given in substitutions per site, is shown at the bottom.

Table 1
Data collection, phasing and refinement statistics

Crystal	Tp0956, Se[a] *peak*	Tp0956, native	Tp0957, native
Space group	C2	I23	$P2_1$
Energy (eV)	12,687.7	12,682.7	12,675.0
Resolution range (Å)	34.1 - 2.30 (2.34 - 2.30)	40.2 - 2.30 (2.34 - 2.30)	37.8-1.40 (1.42-1.40)
Unique reflections	86,064 (3,818)[b]	153,999 (769)	125,807(6,229)
Multiplicity	3.0 (2.3)	19.2 (14.3)	4.1 (3.6)
Data completeness (%)	91.8 (81.9)	99.9 (99.7)	99.8 (99.7)
R_{merge} (%)[b]	6.3 (46.4)	7.7 (89.9)	3.7 (56.6)
I/σ(I)	25.1 (2.10)	42.7 (2.89)	33.7 (2.03)
Wilson B-value (Å^2)	44.6	52.4	18.0
Anomalous scatterers	selenium, 24 out of 24 possible sites	None	None
Figure of merit (90.5 - 2.30 Å)	0.13	N/A	N/A
Figure of merit post-DM (90.5 - 2.3 Å)	0.69	N/A	N/A
Crystal		Tp0956, native	Tp0957, native
Resolution range (Å)		29.9 - 2.30 (2.48 - 2.30)	37.8-1.40 (1.41-1.40)
No. of reflections R_{work}/R_{free}		15,250/758 (2,868/146)	125,752/6,322 (3,800/198)
Data completeness (%)		99.5 (100.0)	99.7 (97.0)
Atoms (non-H protein/solvent/ions)		2,153/61/20	5,634/431/0
R_{work} (%)		19.7 (29.9)	17.8 (30.0)
R_{free} (%)		24.2 (41.1)	20.1 (36.2)
R.m.s.d. bond length (Å)		0.008	0.009
R.m.s.d. bond angle (°)		1.0	1.2
Mean B-value (Å^2) (protein/solvent/ions)		49.2/48.9/108.0	21.9/29.1/38.5
Ramachandran plot (%) (favored/additional/disallowed)[d]		98.9/1.1/0.0	99.0/1.0/0.0
Maximum likelihood coordinate error		0.31	0.36
Missing residues		A: 1-30	A: 7-8, 65-66, 326-328 B: 7-8, 325-328

[a] Bijvoet-pairs were kept separate for data processing

[b] Data for the outermost shell are given in parentheses.

[c] $R_{merge} = 100\ \Sigma_h\Sigma_i|I_{h,i} - \langle I_h \rangle| / \Sigma_h\Sigma_i I_{h,i}$, where the outer sum (h) is over the unique reflections and the inner sum (i) is over the set of independent observations of each unique reflection.

[d] As defined by the validation suite MolProbity[71].

J Mol Biol. Author manuscript; available in PMC 2013 March 9.

Table 2

Putative TPAT T-components and their genomic neighbors

Organism Name	Class	Locus Tag of Tp0956-like Protein (T)	*E*-value[a]	Locus Tag of Nearby DctP-like Protein	Presence of DctM and/or DctQ
Treponema pallidum	Spirochaetes	*tp0956*	3e-99	*tp0957*	Yes
Treponema paraluiscuniculi	Spirochaetes	*tpcca_0956*	3e-99	*tpcca_0957*	Yes
Treponema denticola	Spirochaetes	*tde1021*	7e-92	*tde1020*	Yes
Treponema phagedenis	Spirochaetes	*hmpref9554_2429*	1e-87	*hmpref9554_2430*	Yes
Treponema vincentii	Spirochaetes	*trevi0001_2396*	1e-87	*trevi0001_2397*	Yes
Treponema primitia ZAS-2	Spirochaetes	*trepr_1682*	9e-85	*trepr_1683*	Yes
Verrucomicrobiae bacterium DG1235	Verrucomicrobiae	*vdg1235_4783*	2e-84	*vdg1235_2234*	No
Treponema azotonutricum ZAS-9	Spirochaetes	*treaz_0940*	3e-83	*treaz_0941*	Yes
Treponema brennaborense DSM 12168	Spirochaetes	*trebr_1914*	2e-80	*trebr_1913*	Yes
Alcanivorax borkumensis SK2	γ-proteobacteria	*abo_0687*	2e-75	*abo_0688*	Yes
Syntrophobacter fumaroxidans MPOB	δ-proteobacteria	*sfum_2673*	3e-75	*sfum_2672*	Yes
Spirochaeta thermophila DSM 6578	Spirochaetes	*spithDRAFT_1602*	3e-74	*spithDRAFT_1603*	Yes
Anaeromyxobacter dehalogenans 2CP-C	δ-proteobacteria	*adeh_3980*	3e-74	*adeh_3979*	Yes
Alcanivorax sp. DG881	γ-proteobacteria	*adg881_3113*	4e-72	*adg881_421*	Yes
Anaeromyxobacter sp. Fw109-5	δ-proteobacteria	*anae109_0439*	9e-70	*anae109_0440*	Yes
Anaeromyxobacter sp. K	δ-proteobacteria	*anaeK_4091*	4e-69	*anaeK_4090*	Yes
Anaeromyxobacter dehalogenans 2CP-1	δ-proteobacteria	*a2cp1_4123*	6e-69	*a2cp1_4122*	Yes
Desulfotalea psychrophila LSv54	δ-proteobacteria	*dp0443*	3e-68	*dp0444*	Yes
Spirochaeta smaragdinae DSM 11293	Spirochaetes	*spirs_2251*	1e-67	*spirs_2250*	Yes
Sideroxydans lithotrophicus ES-1	β-proteobacteria	*slit_0818*	5e-66	*slit_0817*	Yes
Dechloromonas aromatica RCB	β-proteobacteria	*daro_3578*	7e-65	*daro_3577*	Yes
Desulfuromonas acetoxidans DSM 684	δ-proteobacteria	*dace_1925* *dace_2628*	9e-63 4e-33	*dace_1924* *N/A*	Yes No
Desulfobacterium autotrophicum HRM2	γ-proteobacteria	*hrm2_33310*	2e-62	*hrm2_33300*	Yes
gamma proteobacterium HTCC5015	δ-proteobacteria	*gp5015_96*	2e-61	*gp5015_80*	Yes

Organism Name	Class	Locus Tag of Tp0956-like Protein (T)	*E*-value[a]	Locus Tag of Nearby DctP-like Protein	Presence of DctM and/or DctQ
Desulfatibacillum alkenivorans AK-01	δ-proteobacteria	*dalk_2688* *dalk_2187*	4e-59 2e-52	N/A N/A	No Yes
Reinekea sp. MED297	γ-proteobacteria	*med297_13632*	3e-57	*med297_13637*	Yes
Desulfococcus oleovorans Hxd3	δ-proteobacteria	*dole_3154* *dole_3077*	2e-56 8e-54	*dole_3157* *dole_3078*, *dole_3074*	Yes Yes
gamma proteobacterium NOR5-3	γ-proteobacteria	*nor53_2500*	2e-56	*nor53_2704*	Yes
Methylomonas methanica MC09	γ-proteobacteria	*metme_0487*	2e-55	*metme_0486*	Yes
Thiobacillus denitrificans ATCC 25259	β-proteobacteria	*tbd_0465*	2e-55	*tbd_0466*	Yes
Leptothrix cholodnii SP-6	β-proteobacteria	*lcho_3613*	3e-53	*lcho_3612*	Yes
Spirochaeta thermophila DSM 6192	Spirochaetes	*stherm_c20400*	6e-50	*stherm_c20410*	Yes
Bacteriovorax marinus SJ	δ-proteobacteria	*bms_0699*	2e-44	*bms_0700*	Yes
Plesiocystis pacifica SIR-1	δ-proteobacteria	*ppsir1_38751*	1e-38	*ppsir1_38741*	Yes
Leptospira biflexa serovar Patoc strain ‘Patoc 1 (Paris)’	Spirochaetes	*lepbi_I1477*	4e-32	*lepbi_I1476*	Yes

[a] As defined by PSI-BLAST[42; 67]