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*

Ranjit K. Deka1,2, Chad A. Brautigam1, Xiaofeng F. Yang1, Jon S. Blevins1, Mischa Machius6, Diana R. Tomchick6, and Michael V. Norgard12

From the Departments of 1Microbiology and 2Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75390 and the 6Department of Microbiology and Immunology, Indiana University School of Medicine, Indianapolis, Indiana 46202

Treponema pallidum, the bacterial agent of syphilis, cannot be cultivated in vitro. This constraint has severely impeded the study of the membrane biology of this complex human pathogen. A structure-to-function approach thus was adopted as a means of discerning the likely function of Tp0319, a 35-kDa cytoplasmic membrane-associated lipoprotein of T. pallidum formerly designated as TmpC. A 1.7-Å crystal structure showed that recombinant Tp0319 (rTp0319) consists of two αβ domains, linked by three crossovers, with a deep cleft between them akin to ATP-binding cassette (ABC) receptors. In the cleft, a molecule of inosine was bound. Isothermal titration calorimetry demonstrated that rTp0319 specifically binds purine nucleosides (dissociation constant (Kd) ~10^-7 M). This prediction for purine nucleosides by rTp0319 is consistent with its likely role as a receptor component of a cytoplasmic membrane-associated transporter system. Reverse transcription-PCR analysis of RNA isolated from rabbit tissue-extracted T. pallidum additionally showed that tp0319 is transcriptionally linked to four other downstream open reading frames, thereby supporting the existence of an ABC-like operon (tp0319–0323). We herein re-name tp0319 as purine nucleoside receptor A (pnrA), with its operonic partners tp0320–0323 designated as pnrB-E, respectively. Our study not only infers that PnrA transports purine nucleosides essential for the survival of T. pallidum within its obligate human host, but to our knowledge, this is the first description of an ABC-type nucleoside transport system in any bacterium. PnrA has been grouped with a functionally uncharacterized protein family (HBG016869), thereby implying that other members of the family may have similar nucleoside-binding function(s).

Treponema pallidum, the bacterial agent of syphilis, cannot be cultivated continuously in vitro (1). Historically, this constraint has severely hampered progress in understanding many features of this enigmatic pathogen. More specifically, key aspects of the membrane biology of T. pallidum, particularly as it pertains to the interaction of the pathogen with its human host, remain poorly defined (2–4). The initial discovery of membrane lipoproteins in T. pallidum (5, 6) spawned new avenues of investigation into treponemal membrane biology, inasmuch as in other bacteria, lipoproteins have importance as virulence factors, modular components of ATP-binding cassette (ABC)3 transporters, receptors, protective immune targets, and proinflammatory agonists that elicit innate immune responses (7–14). In fact, genome analysis predicts that T. pallidum devotes as much as 3% of its genetic coding capacity to lipoproteins (15). However, to date, with one exception (9, 16), comparative sequence analyses have not been fruitful for predicting the functions of treponemal lipoproteins, and the inability to cultivate (and thus genetically manipulate) the organisms has precluded using classical gene inactivation approaches (17) for investigating the functions of treponemal membrane lipoproteins.

As an alternative strategy for understanding the distinctive membrane biology of T. pallidum (2–4), we have been pursuing a structural biology approach that seeks to garner new insights into putative functions for a number of membrane lipoproteins of T. pallidum. This initiative previously has led to the characterization of an l-methionine-binding protein in T. pallidum (18). The current study focuses on Tp0319 (15), formerly designated as TmpC membrane protein C (TmpC), and shown to be a 35-kDa lipoprotein (19). We have obtained structural and biochemical evidence that Tp0319 likely is a periplasmic purine nucleoside receptor in T. pallidum. Hence, Tp0319 (TmpC) has been renamed herein as purine nucleoside receptor A (PnrA). Furthermore, PnrA appears to be a member of the ABC superfamily and, as such, represents the first description of a periplasmic purine nucleoside-binding protein of its type in bacteria. The overall importance of PnrA as an ABC-type receptor tethered to the cytoplasmic membrane (via its lipid anchors) and implications of PnrA for treponemal survival within the human host are discussed. These findings add to our knowledge of key physiological functions for the treponemal membrane lipoproteins and their relationship to T. pallidum membrane biology.

EXPERIMENTAL PROCEDURES

Construction of His6-tagged rTp0319—To create a non-acylated derivative of recombinant Tp0319 (rTp0319) in Escherichia coli, a fragment encoding amino acids 2–333 (residue 1 is the post-translationally

* This work was supported by Grant AI-56305 from the NIAID/National Institutes of Health and by Grant I-0940 from the Welch Foundation. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the United States Department of Energy, Office of Energy Research, under Contract W-31-109-ENG-38. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be indicated this fact.

¶1 These authors contributed equally to this work.

1 These abbreviations used are: ABC, ATP-binding cassette; GBP, glucose/galactose-binding protein; ITC, isothermal titration calorimetry; LF, ligase-free; Pnr, purine nucleoside receptor; RBP, ribose-binding protein; r.m.s., root mean square.
modified N-terminal cysteine) was amplified by PCR using genomic DNA isolated from the virulent Nichols strain of T. pallidum. The PCR primers were 5′-ctcGAATTCCTAAGAGCGACAGGCGG-3′ and 5′-ctcCTCGAGTTATTTGTCATCGGTGCAATGCT-3′. In the preceding sequences, the region of each primer complementary to the tp0319 sequence is underlined. The forward primer contained both a ctt overhang and an EcoRI site (bolded); the reverse primer contained a ctt overhang and a Xhol site (boldface). PCR was performed using TaKaRa Ex Taq DNA polymerase at an annealing temperature of 55 °C. The amplification product was directionally cloned into the pProEX HTa vector (Invitrogen), which adds an N-terminal His6 tag followed by a tobacco etch virus protease cleavage site. Ligation was transformed into E. coli XL1-Blue-competent cells (Stratagene); plasmids from colonies that tested positive by restriction digests were sequenced for verification.

Expression and Purification of rTp0319—E. coli XL1-Blue cells harboring the cloned tp0319 gene fusion were grown at 37 °C in LB medium containing 100 μg/ml of ampicillin until the cell density reached an A600 of 0.6. The temperature of the culture was then shifted to 30 °C, and overexpression of rTp0319 was achieved by induction for 5 h with 600 μM isopropyl β-D-thiogalactopyranoside. Cells were harvested by centrifugation and lysed at room temperature for 30 min using 50 ml of B-PER II (Pierce) per liter of culture. After centrifugation at 25,000 × g, the supernatant was immobilized on a Ni2+ affinity column and washed with 10 column volumes of 20 mM Tris-HCl, 20 mM NaCl, 20 mM imidazole (pH 8.5) (buffer A), followed by 10 column volumes of 20 mM Tris-HCl, 1 mM NaCl (pH 8.5) and 5 column volumes of buffer A. The rTp0319 protein was eluted from the column using 20 mM Tris-HCl, 200 mM NaCl, 200 mM imidazole (pH 8.5). The sample buffer was exchanged to 20 mM Heps (pH 7.5), 20 mM NaCl, 2 mM octyl β-glucoside (buffer B) by using a MW 10,000 exclusion Amicon concentration cake. For renaturation, protein bound to beads was suspended in 100 ml of denaturation buffer containing 8 M urea, 100 mM Tris-HCl (pH 8.0), and 15 mM β-mercaptoethanol. The beads were collected by centrifugation at 3,000 × g for 5 min and then were washed with 100 ml of the above buffer to remove unbound ligands. This step was repeated with 100 ml each of 1:1 and 1:3 diluted denaturation buffer. For renaturation, protein bound to beads was suspended in 100 ml of buffer containing 20 mM Tris-HCl, 20 mM NaCl, 20 mM imidazole (pH 8.5) and incubated for 1 h with gentle agitation. The slurry was then poured into a column, and the protein was eluted with the above buffer containing 200 mM imidazole. Protein buffer was exchanged with 20 mM Hepes (pH 7.5), 100 mM NaCl, and 2 mM octyl β-glucoside (buffer C) by using a M, 10,000 size exclusion concentration device (Amicon). Refolded, concentrated protein was further purified by gel filtration chromatography (as described above) and designated as ligand-free (LF)-rTp0319.

Isothermal Titration Calorimetry—Binding of ligands to LF-rTp0319 was analyzed by measuring heat changes during the titration of ligand into the protein solution (in buffer C) using a VP-ITC titration microcalorimeter (Microcal, Inc.). The concentration of LF-rTp0319 and ligands were obtained spectrophotometrically using a calculated extinction coefficient for LF-rTp0319, extinction coefficients for the individual ligands were provided by the supplier (Sigma). The titrations were carried out at 20 °C with the 1.4-ml stirred reaction cell containing ~20 μM protein in all experiments. A series of 30–32 injections (6–8 μl each) of ligand at concentrations ranging from 200 to 500 μM were applied to the cell, and the heat resulting from binding was monitored until the protein was saturated with ligand. Association constants and thermodynamic parameters were fitted to these data using the program Microcal-ORIGIN. The data reported are the average of at least three measurements.

Protein Crystallization—Crystals of rTp0319 or LF-rTp0319 with various nucleosides, or the selenomethionyl variant of rTp0319, were obtained by the hanging-drop vapor diffusion method (21). Sparse matrix crystallization schemes (Hampton Research) were used to screen preliminary conditions. In some conditions, long, rod-shaped crystals with dimensions up to 0.1 mm × 0.1 mm × 1 mm appeared within 2–4 days. The most suitable crystals were routinely obtained in drops containing 2.5 μl of protein (starting concentration of 15 mg/ml in buffer B) and 2.5 μl of 0.2 mM ammonium acetate, 0.1 mM sodium acetate (pH 4.6), and 30% (w/v) polyethylene glycol 4000 (20 °C). Crystals were cryoprotected in 10% (v/v) ethylene glycol-enriched crystallization reagent containing 32% (w/v) polyethylene glycol 4000, flash-cooled in liquid propane, and were stored in liquid nitrogen. Crystals exhibited the symmetry of space group P2221, with unit cell dimensions of a = 56 Å, b = 69 Å, c = 70 Å and contained one molecule in the asymmetric unit.

Structure Determination and Refinement—Diffraction data were collected at the beamlines 19-ID and 19-BM (Structural Biology Center Collaborative Access Team) at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) and processed with HKL2000 software (22). To facilitate phase calculations, diffraction data were collected from a crystal of the selenomethionyl variant of rTp0319. These data were collected at two wavelengths that corresponded to the peak and inflection points of an x-ray fluorescence scan of a selenomethionyl rTp0319 crystal conducted near the K-edge of x-ray absorption of selenium. Using diffraction data to 4.0 Å, 12 of 14 possible selenium sites were located using the program CNS (21). Utilizing data to 2.01 Å, phase refinement and density modification were performed via the programs MLPHARE and dm (Collaborative Computational Project 4). A preliminary model for 289 residues was automatically generated using the program ARP/wARP (23), and 241 of these were docked to the
sequence. A native data set was collected to a 1.71 Å resolution and employed for further model building via programs O (24) and XplorView (25); refinement using CNS resulted in an Rwork of 18.1% and an Rfree of 21.1% for 316 total residues. Data collection and refinement statistics are shown in Table 1. The initial electron density maps for the LF-Tp0319 loaded with exogenous adenosine (LF-A) or LF-Tp0319 loaded with exogenous guanosine (LF-G) were calculated after rigid-body refinement of the LF-Tp0319 model using the LF-A or LF-G data sets. Refinement proceeded as outlined above. Final statistics for these structures are also presented in Table 1.

RNA Isolation and RT-PCR—Approximately $1 \times 10^{11}$ T. pallidum bacteria were freshly harvested from rabbit testicular tissue (26, 27). RNA was extracted from the spirochetes using an RNasy kit (Qiagen). To remove residual contaminating DNA, the RNA was then subjected to further treatment with a TURBO DNA-free kit (Ambion). The quantity of RNA obtained was determined spectrophotometrically.

RT-PCR was used to evaluate the potential transcriptional linkage of tp0319 to four downstream open reading frames (tp0320–0323). Oligonucleotide primer pairs for RT-PCR were designed (Table 2) to enable synthesis of PCR products in the range of 400 to 800 bp that would span one or more of the intergenic regions of the open reading frames (Fig. 7). The StrataScript two-tube RT-PCR system (Stratagene) was used first to synthesize transcriptional cDNA via random priming. cDNA was generated by incubating ~400 ng of T. pallidum RNA in a 50-μl reaction for 90 min at 44 °C. One μl of the cDNA was then used as template for the specific oligonucleotide primer pairs (Table 2). Each RT-PCR reaction was accompanied by a positive control reaction that utilized the same set of primers and 20 ng of T. pallidum genomic DNA.

Template DNA was extracted from $5 \times 10^{10}$ spirochetes using a Qiagen DNA extraction kit. An additional negative control reaction was also performed in which 10 ng of T. pallidum RNA was used as template in the absence of RT. All resulting PCR products were analyzed by 1.5% agarose gel electrophoresis and staining with etidium bromide.

### Table 1

| Data collection | Crystal | rTp0319-AP | SeMet pk<sup>a</sup> | SeMet inf<sup>a</sup> | LF-G | LF-A |
|----------------|---------|------------|---------------------|---------------------|------|------|
| Energy (eV)    | 12,656.7| 12,659.9   | 12,657.8            | 12,659.9            | 12,659.9 | 12,659.9 |
| Resolution range (Å) | 49.0–1.71 (1.74–1.71) | 43.4–2.02 (2.07–2.02) | 43.6–2.02 (2.06–2.02) | 36.9–1.70 (1.76–1.70) | 36.9–1.70 (1.76–1.70) |
| Completeness (%) | 99.5 (93.8) | 99.2 (96.2) | 97.0 (66.3) | 96.3 (93.8) | 96.3 (93.8) |
| Rmerge (%)<sup>b</sup> | 103.0 (62.4) | 92.2 (27.0) | 8.8 (40.2) | 5.9 (35.2) | 6.8 (22.6) |
| Rfree (%) | 18.6 (2.0) | 19.6 (5.4) | 16.8 (2.9) | 18.2 (3.0) | 19.2 (3.0) |
| Wilson B-value (Å<sup>2</sup>) | 20.8 | 19.0 | 23.7 | 15.8 | 17.6 |

<sup>a</sup> Bijvoet-pairs were kept separate for data processing.

<sup>b</sup> Rmerge = $\sum_n\sum_k|I_{hk} - I_{hk}'|/\sum_n\sum_k I_{hk}$, 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.

<sup>c</sup> Using a modified Ramachandran plot (28).

### Table 2

| Primer name | 5’ to 3’ sequence |
|-------------|-------------------|
| Tp319–5’    | CGATCGTCGTCTCAATGGTCAGGAC |
| Tp320–3’    | CCGCAGAAAGAAGACGATCAC |
| Tp321–3’    | CGGACTACGGATGCAACAGC |
| Tp321–5’    | GTACCCTTCTTGAGGTAATGCAG |
| Tp322–3’    | CCTGAAGTTTCCACCGAGCTGCG |
| Tp322–5’    | GCACCGAGCATTTTCCGGATGT |
| Tp323–3’    | CCTCGGTTCCCTCGTGC |

### Results and Discussion

Expression and Purification of rTp0319—To obtain sufficient quantities of rTp0319 for crystallization, a recombinant (r) DNA-derived, soluble form of rTp0319 was generated lacking the N-terminal Cys acylation site. rTp0319 also contained an N-terminal His<sub>6</sub> fusion peptide linked to the mature protein via 23 spacers residues. This soluble derivative was expressed at high levels in E. coli, and the protein was purified to greater than 95% homogeneity by nickel-affinity and size-exclusion chromatography; typical yields were in the range of 10–15 mg per liter of cell culture. Size-exclusion chromatography also revealed that rTp0319 exists as a monomer in solution (data not shown).

Structural Overview of rTp0319—rTp0319 is comprised of 361 amino acid residues, 29 of which are vector derived. Crystals of rTp0319 were obtained readily and diffracted to a maximum resolution of 1.7 Å. After building a model into the electron density maps derived from x-ray diffraction data from rTp0319 crystals, it was apparent that a significant portion of the N-terminal part of rTp0319 was not present. To determine whether proteolytic processing and/or N-terminal disorder were responsible for this phenomenon, crystals of rTp0319 were dissolved and subjected to N-terminal sequencing. This analysis indicated that cleavage had occurred between residues 13 and 14 of the mature protein. The crystalline rTp0319 protein was therefore devoid of 41 N-terminal amino acid residues (12 of these from the mature rTp0319).
addition, residues 14–17 are not visible in electron density maps, presumably due to disorder. Residues 18–333 are visible in electron density maps and are included in the rTp0319 model.

The crystal structure of rTp0319, which was refined using diffraction data to 1.7 Å, is shown in Fig. 1. The crystal structure obtained for residues 18–333 revealed a polypeptide with two domains, designated herein as “N” or “C”, which denote the domain that harbors either the N terminus or the C terminus, respectively (Fig. 1). The N and C domains are structurally similar; for 90 equivalent Cα positions, the root mean square (r.m.s.) deviation is 1.6 Å. The 130-residue N domain (residues 18–122 and 262–286) is slightly smaller than the 178-residue C domain (residues 124–260 and 293–333). At the centers of the two domains are six-stranded β-sheets. In the N domain, the strands are parallel to one another, whereas one of the strands (strand 12) of the C domain sheet is antiparallel to the other strands. The C-terminal ends of the strands of the β-sheets of the N and C domains are oriented toward one another. In each domain, four α-helices flank the central β-sheet. Outside of these common secondary structural elements, the C domain has an additional span of amino acids that comprise an α-helix (helix J) and a short 3₁₀ helix (helix 3₁₀–K) near the C terminus. The model was analyzed using a Ramachandran plot; all residues but six were in “favored” regions of the plot (28). The main-chain atoms of all but one of these six (A60) were in regions of well-defined electron density.

The N and C domains are connected by three regions in rTp0319 that likely comprise a hinge (Fig. 1; see below). These connector regions are termed CR1 (residue 123), CR2 (residue 261), and CR3 (residues 287–292). All of the connector regions are devoid of any regular secondary structure; CR3 is the largest, and it forms a loop that protrudes from the body of the protein. The N and C domains contact each other, burying about 1300 Å² at their interface.

**Ligand Binding by rTp0319**—During the course of building the molecular model from electron density maps, non-protein density was observed at the interface between the N and C domains of rTp0319. The high quality of the electron density facilitated the assignment of the ligand as a purine nucleoside (Figs. 2 and 3). rTp0319 apparently appropriated this molecule from the cytoplasm of the host expression organism, *E. coli*, inasmuch as exogenous nucleoside was not added to any solution during protein purification. The procurement of this molecule from *E. coli* and its retention by rTp0319 through chromatographic purification steps provided the initial indication that rTp0319 likely bound purine nucleosides at high affinity. The hydrogen bonding pattern around the nucleoside and the resolution of the structure did not allow the precise identification of the base moiety. However, based on

---

**FIGURE 1.** The crystal structure of rTp0319. The α-helices (green), β-strands (purple), and loop regions (blue) of rTp0319 are shown in a ribbon representation. Atoms from the bound nucleoside are depicted as spheres: carbon, gray; oxygen, red; nitrogen, blue. Secondary structural elements are labeled (helices A through K; β-strands 1 through 12), as are the boundaries of the N and C domains and the three connector regions (CR).

**FIGURE 2.** Structures and numbering conventions for nucleosides used in this study.
the shape of the density, we surmised that the nitrogenous base was either adenine or hypoxanthine (the base moiety of the nucleoside inosine); the nucleoside was modeled as inosine, for reasons outlined below. A surface accessibility calculation revealed that the nucleoside was almost completely surrounded by protein (data not shown). The ribose moiety is not solvent-accessible, and the hypoxanthine has less than 1 Å² of accessibility to the solvent. The ribose moiety of the nucleoside adopts the C3′-endo conformation, with the base assuming the anti conformation (Fig. 3). Numerous polar and non-polar contacts are present between the nucleoside and both domains of rTp0319 (Fig. 3).

All oxygens of the ribose moiety of the nucleoside form polar interactions with the protein, except O4′. The side chains of Asp-108, Asp-238, and Lys-260 form hydrogen bonds with the exocyclic oxygen atoms of the ribose moiety. Additionally, the main chain nitrogen atom of Gly-212 makes a hydrogen bond with O5′ of the nucleoside. The base moiety of the inosine is sandwiched between the hydrophobic side chains of residues Phe-186 and Phe-36; the latter residue also contacts the ribose moiety. The hypoxanthine has fewer polar contacts to the protein than the ribose moiety; only the side chains of Asp-27 and Asn-37 have direct hydrogen bonds to the base. Water-mediated contacts to the base are made by the side chains of Asp-33, Asp-108, and Ser-187. To distinguish this structure from those detailed below, it shall hereafter be referred to as “rTp0319-AP” structure, the “AP” signifying that this is the structure of the protein as purified from the cytoplasm of E. coli.

**Nucleoside Binding Preferences of rTp0319—Crystallographic and calorimetric methods were used to explore the nucleoside binding preferences of rTp0319. To facilitate the characterization, LF-rTp0319 apoprotein was prepared. Unfortunately, attempts at crystallizing LF-rTp0319 thus far have been unsuccessful.** LF-rTp0319 was therefore titrated with either G or A and then crystallized as in earlier experiments. Such crystals displayed diffraction characteristics similar to the rTp0319-AP protein crystals (Table 1), and the structures (termed LF-G and LF-A, respectively) were determined using difference Fourier techniques.

The only appreciable difference between the main chain backbones of the rTp0319-AP and LF-G structures occurs in residues 83–87, where there is a shift in the position of this peptide. When the rTp0319-AP and LF-G structures are superposed using all Cα atoms, the r.m.s. deviation of these five shifted Cα atoms is 0.59 Å, compared with 0.19 Å for all Cα atoms. Because both of the LF structures harbored this change, it appears to be a consequence of either the treatment of the protein to render it ligand-free or of the nucleoside binding process. The nucleoside binding characteristics of the LF-G structure are very similar to those of the rTp0319-AP structure, with the exception that Asp-27 makes a bidentate interaction with the base moiety of the bound guanosine (Fig. 4A).

The LF-A structure exhibits additional features that are divergent from the rTp0319-AP and LF-G structures. Specifically, the hydrogen-bond network between the protein and adenosine is significantly different from that of the rTp0319-AP protein near position 6 of the respective nitrogenous bases (Figs. 2 and 4B). This change appears to be due to a shift in the position of the loop between β-strand 1 and helix A (residues 27–35, hereafter termed Lβ1–αA). In particular, the positions of Ser-28 and Gly-29 in LF-A deviate when compared with the homologous positions in the rTp0319-AP and LF-G structures. When superposed, the 316 Cα atoms of LF-A and the rTp0319-AP structure have an r.m.s. deviation of 0.6 Å. However, in this same superposition, the Cα atoms of Ser-28 are 1.8 Å apart, and those of Gly-29 are 3.3 Å apart (Fig. 4B). We designate the disposition of the loop in the LF-G and rTp0319-AP structures as "H" (hydrogen bond) and in the LF-A structure as "V" (van der Waals contact) according to the type of contacts that residue Ser-28 makes with bound nucleoside. When in the H position, the exocyclic atom at position 6 of the nucleoside no longer has a direct contact with Asn-37 nor does it have a water-mediated contact with Ser-187 (Fig. 3). Rather, the N-6 of adenine hydrogen bonds to the main chain carbonyl oxygen of Ser-28 and has a water-mediated contact with Asn-37 (Fig. 4B). In the V position, the side chain of Ser-28 has a van der Waals interaction with bound nucleoside (Fig. 3).

The LF-A and LF-G structures give insight into the nucleoside bound in the rTp0319-AP structure. The LF structures are identical except for the disposition of Lβ1–αA and the identity of the bound nucleoside. The atom at position 6 of the bound nucleoside is oxygen in LF-G but nitrogen in LF-A. The position of Lβ1–αA is most compatible with a nitrogen atom at position 6, because of the close proximity of this atom and the main chain carbonyl oxygen of Ser-28 (Fig. 4B). In the LF-G structure, Lβ1–αA assumes the V position, apparently as a response to the oxygen atom at position 6 of guanosine. The rTp0319-AP structure also features an Lβ1–αA in the V orientation, suggesting that the atom at position 6 of the bound nucleoside is oxygen. Given these observations and the shape of the electron density for the bound nucleoside, we conclude that inosine is the bound nucleoside in the rTp0319-AP structure.

The binding of various nucleosides to LF-rTp0319 was further characterized using isothermal titration calorimetry (ITC). As shown in Table 3 and Fig. 5, purines bind to LF-rTp0319 with higher affinity than pyrimidines. The isotherms obtained from ITC were fitted to a one-site binding model, which was judged to give the best fit to these data. The resulting fits gave values for the stoichiometry of binding between 0.7 and 1.0 (for different batches of LF-rTp0319), suggesting that one nucleoside binds per monomer of LF-rTp0319. Among the purines tested, there is a range of binding affinities. Guanosine appears to be preferred, because LF-rTp0319 binds to this nucleoside with a Kd of about 71 nM, which is about three times lower than any other purine. The superior binding of guanosine compared with adenosine or inosine is easily rationalized by the LF-G structure; an additional hydrogen bond between the nucleoside and the protein is evident in the LF-G structure compared with the LF-A and rTp0319-AP structures (Figs. 3 and 4). This interaction, which is between a carboxylate oxygen of Asp-27 and N-2 of guanosine, likely provides the added energy that accounts for the greater binding affinity of this nucleoside for LF-rTp0319. The binding affinity of LF-rTp0319 for the purine nucleoside xanthosine is significantly lower than for the other purines. This difference may signify the impor-
tance of the water-mediated interaction between a carboxylate oxygen of Asp-108 and N-3 of the bound purines. Of the tested purines, only xanthosine is protonated at N-3 (Fig. 2). We were unable to determine the structure of the LF-rTp0319/xanthosine complex, but we hypothesize that this alteration in the hydrogen bonding pattern around the nucleoside may be sufficient to eliminate the water-mediated interaction, thus accounting for the relatively weak affinity of LF-rTp0319 for xanthosine.

Two of the three pyrimidines tested bind to LF-rTp0319, but their affinities are substantially lower than those displayed by the purines (Table 3). The affinity of LF-rTp0319 for uridine is about 180-fold lower than that for guanosine. Cytidine binds to LF-rTp0319 about 2-fold more tightly than uridine. No binding of thymidine to LF-rTp0319 could be demonstrated.

Theoretical models of the various nucleosides bound to rTp0319 (data not shown) offer possible explanations for these observations. The models were constructed such that the ribose moiety was identical to that observed in our crystal structures. These models show that Asp-27, a key residue in the binding of the purines (Figs. 3 and 4), is not in a position that is compatible with hydrogen bonding to a pyrimidine. The water-mediated interaction between Asp-33 and the nucleoside observed in the rTp0319-AP and LF structures also may be absent when a pyrimidine is bound. Modeling a thymidine into the rTp0319-binding site shows that its C5 methyl group clashes with Gly-212. The clash cannot be relieved by the base adopting the syn conformation, because this would put the O-2 of thymidine in conflict with Gly-212. These large steric clashes explicate the inability of thymidine to bind to LF-rTp0319. These unfavorable parameters for pyrimidines binding to LF-rTp0319, combined with the purine nucleoside binding data above, prompt us to contend that in *T. pallidum*, Tp0319 likely is a purine nucleoside receptor.

The effect of removing hydroxyl groups from the ribose moieties of nucleosides binding to LF-rTp0319 was also examined (Table 3). The 2’-deoxy versions of guanosine and adenosine bind to LF-rTp0319 with affinities that are reduced about 5-fold in comparison with those of the nucleosides with unmodified ribose. It should be noted, however, that these affinities are at least 4-fold higher than those of any pyrimidine tested. The lack of detectable binding for 5’-deoxynucleosides and 3’-deoxynucleosides (Table 3) demonstrates that the 5’- and 3’-hydroxyl groups are crucial for the binding of nucleosides to LF-rTp0319. However, the loss of hydrogen bonding potential for these latter two modified nucleosides cannot fully explain the complete loss of binding to LF-rTp0319. It is more likely that these hydroxyls also are necessary for establishing the proper orientation or conformation of the nucleoside within the binding pocket. The 3’-hydroxyl group establishes hydrogen bonds to residues in both the N and C domains of rTp0319. Therefore, assuming the N and C domains are subject to rigid-body motions with respect to one another, the hydrogen bonds formed to this group may facilitate the closure of LF-rTp0319 around its ligand. LF-rTp0319 did not bind to either the sugar ribose or to the purine base adenine (Table 3). In addition, it should be noted that the $\Phi$ and $\psi$ angles of residue Asp-238 place it in a “disfavored but allowed” region of a modified Ramachandran plot (28). This location suggests that the favorable energy of hydrogen bonding to the 3’- and 5’-hydroxyl groups of a nucleoside compensates for the strain of assuming these $\Phi$ and $\psi$ angles. Without this hydrogen-bonding energy, Asp-238 may be unable to adopt the observed conformation, thus altering the nucleoside-binding site of rTp0319 in a way that discourages nucleoside binding.

The behavior of rTp0319 during purification and crystallization reveals other tantalizing clues concerning its interaction with nucleosides. Guanosine and adenosine apparently do not co-purify with rTp0319 through two chromatographic steps in which no exogenous ligand is supplied. This fact suggests that the off rate of this nucleoside must be very slow under the purification conditions employed (see “Experimental Procedures”). Guanosine and adenosine may not co-purify with rTp0319, Two possible explanations for these observations are that the off rates of guanosine and adenosine are rapid relative to the time frame for purification or that there are significantly lower concentrations of guanosine and adenosine (relative to inosine) in the cytoplasm of *E. coli*. Spectrophotometric and calorimetric experiments showed that rTp0319-AP is 60–80% occupied with nucleoside (data not shown). Because there is no

### Table 3

| Nucleoside purines     | $K_d$ (nM) |
|------------------------|------------|
| **Adenosine derivatives** |            |
| Adenosine              | 270 ± 40   |
| 2’-Deoxyadenosine      | 1,400 ± 90 |
| 5’-Deoxyadenosine      | NM         |
| **Guanosine derivatives** |          |
| Guanosine              | 71 ± 5     |
| 2’-Deoxyguanosine      | 370 ± 40   |
| 3’-Deoxyguanosine      | NM         |
| **Other purines**      |            |
| Inosine                | 210 ± 20   |
| Xanthosine             | 5,300 ± 80 |
| **Pyrimidines**        |            |
| Cytidine               | 5,900 ± 100|
| Thymidine              | NM         |
| Uridine                | 13,000 ± 700|
| **Purine bases**       |            |
| Adenine                | NM         |
| **Sugars**             |            |
| $\beta$-[1-(--)-Ribose]| NM         |

*NM, not measurable.*
evidence of partial occupancy of the nucleoside in the rTp0319-AP structure, ligand-bound rTp0319 must have preferentially crystallized in the presence of some ligand-free form. This observation implies that there are significant structural differences between liganded and unliganded rTp0319 that allow the former, but not the latter, to crystallize under the conditions tested. A large conformational change upon ligand binding could account for these findings.

Comparisons with Other Ligand-binding Proteins—The structure of rTp0319 was compared with those in the Brookhaven Protein Data Bank by submitting the coordinates to the DALI server (29). This search revealed that rTp0319 is structurally related to certain sugar-binding proteins, such as the *E. coli* ribose-binding protein (RBP) (30) or the glucose/galactose-binding protein (GBP) of *Salmonella typhimurium* (31). For example, the r.m.s. deviation between 200 homologous Cα coordinates of rTp0319 and RBP is 1.9 Å (Fig. 6). RBP and GBP are periplasmic components of well characterized ABC transport systems, wherein energy from the hydrolysis of ATP is used to actively transport small-molecule metabolites from the extracellular environment into the cytoplasm of the bacterial cell (32). All ABC-type receptors, as well as rTp0319, share the same structural organization: two globular domains connected by short linker regions (Figs. 1 and 6).

The similarities between the two periplasmic binding proteins noted and rTp0319 extend beyond their gross structural homology. Crystal structures of RBP and other periplasmic binding proteins demonstrated that these proteins undergo a large conformational change upon ligand binding (33–35). Specifically, the connector regions of the two α/β domains of the receptor serve as a hinge, allowing substantial rigid-body motions of the two domains with respect to one another during ligand binding. Such motion has been characterized as “Venus fly trap”-like (36). It is likely that rTp0319 undergoes a similar structural transition for the following reasons. The nucleosides in our structures are almost completely surrounded by the protein. The steric constraints to ligand ingress into and egress from the binding pocket of rTp0319 could be relieved by rigid-body movements of the N and C domains relative to one another, using the CR regions as a hinge. Circumstantial evidence for this hypothesis is provided by the crystallization behavior of rTp0319 (see above). Validation of the domain motions of rTp0319, however, awaits further structural characterization of LF-rTp0319.

The ligand binding characteristics of rTp0319 are also similar to those of RBP and GBP. Like rTp0319, bacterial periplasmic receptors...
frequently retain their respective ligands upon purification, a consequence of their high affinity binding and slow off rates (18, 32, 37). The ligand-binding cleft is located between the two globular domains in the periplasmic receptors and rTp0319. The ligand bound to rTp0319 occupies space comparable with those bound by RBP and GBP. Moreover, like these sugar-binding proteins, rTp0319 utilizes acidic amino acid side chains to contact some of the exocyclic oxygen atoms of the bound sugar moiety, and aromatic side chains are found in van der Waals contact with the respective ligands. It is noteworthy that there is precedence for multiple ligand binding and structural plasticity in response to ligand identity among periplasmic binding proteins. For example, the lysine-, arginine-, and ornithine-binding protein reorganizes binding site residues to accommodate the hydrogen bonding patterns of different bound ligands (38). The striking structural, conformational, and ligand-binding similarities between rTp0319 and periplasmic binding proteins thus allowed us to hypothesize that Tp0319 is the binding protein of an ABC-type nucleoside uptake system in T. pallidum. Tp0319 Is Encoded within an ABC-type Operon—The components of ABC transport systems typically include the receptor, an ATPase (for energy utilization), and an integral membrane permease for shuttling the ligand into the cytoplasm (32, 39, 40). In some cases, one or two additional gene products are accessories to the transport process (32, 41). The genes encoding all of these ABC proteins often are co-transcribed (32, 42). Genome sequencing has shown that rTp0319 (encoding a hypothetical protein), tp0320 (encoding a small hypothetical protein), tp0321 (encoding a hypothetical ATP-binding protein), and tp0322 (encoding a hypothetical permease) are linked in T. pallidum (15). These linkage data, in conjunction with the structural evidence presented above, prompted us to garner further genetic evidence to support that Tp0319 is a member of an ABC transport system. To this end, RT-PCR was performed on RNA isolated from rabbit tissue-extracted T. pallidum. Oligonucleotide primers were designed to span either three of the intergenic regions, or the entire sequence of tp0320 between tp0319 and tp0321 (Table 2). As shown in Fig. 7, the five treponemal genes were demonstrated by RT-PCR to be transcriptionally linked, similar to what has been observed for other ABC transport systems (32, 42).

To date, no ABC-type transport system for nucleosides has been described for any bacterial system. Based on earlier analyses, it was proposed that one or more of tp0319-tp0323 may encode members of an ABC transport system (43) and that Tp0319 may be a ligand-binding protein, but prior to our studies, there had been no indication that Tp0319 might represent an ABC transporter for nucleoside(s). In E. coli, nucleosides first traverse the outer membrane via Tsx, a 12-stranded β-barrel nucleoside-specific channel (44). Once in the periplasmic space, nucleosides then are transported across the cytoplasmic membrane by either of the high affinity transporters NupC or NupG, both of which are energized by the proton motive force (45). Orthologs of Tsx, NupC, or NupG are not purported to be present in T. pallidum (43, 46) nor is any other type of bacterial nucleoside transporters of the concentrative nucleoside transporter, major facilitator superfamily, or equilibrative nucleoside transporter types. Therefore, the ABC transport system comprised of Tp0319 may be the sole source of nucleosides for this bacterial pathogen.

Relationship of Tp0319 to Other Protein Families—Hogenprot, a protein data base derived from bacterial genome sequencing, places Tp0319 in a family designated as HBG016869. Heretofore, functions for proteins contained within this family have not been unambiguously assigned. Fig. 8 shows the results of an alignment of the 61 members of the HBG016869 family, revealing that all of the putative ligand-contacting residues are well conserved among these proteins. Thus, our structural, biochemical, and genetic analyses described herein, in conjunction with the fact that many of the proteins in this family likely are cytoplasmic membrane-associated, provide the first evidence to suggest that members of the HBG016869 family likely include nucleoside-binding or, more specifically, may be purine nucleoside-binding components of ABC transport systems. We thus contend that in addition to its newly described function in T. pallidum, Tp0319 likely also is representative of a newly recognized family of bacterial purine nucleoside receptors of the ABC-type.

The National Center for Biotechnology Information has placed Tp0319 into a cluster of orthologous groups (COG) of proteins denoted as COG1744. A similar analysis to that shown in Fig. 8 was performed on this group of 54 proteins. The nucleoside-binding residues that form the binding pocket of Tp0319 are not well conserved among the members of this COG. Many of these proteins, therefore, probably are not purine nucleoside receptors.

Summary and Implications—T. pallidum is an obligate pathogen of humans (47). Although long considered as one of the classical sexually transmitted disease agents, the inability to cultivate T. pallidum in vitro has rendered it among the most poorly understood of human infectious disease organisms. For example, it has taken decades for investigators to appreciate even the most basic aspects of treponemal membrane architecture (2–4, 48). In this regard, although T. pallidum possesses an outer membrane, the structure bears little resemblance to the outer membranes of Gram-negative bacteria; the treponemal outer membrane has a paucity of protein(s), lacks lipopolysaccharide, and seems to be devoid of channel-forming (e.g. porin) proteins (48–51) despite early reports to the contrary (52, 53).

Aside from their importance as potential carbon sources, nitrogen sources, and precursors of nucleic acid biosynthesis, T. pallidum lacks the capacity for de novo purine biosynthesis, consistent with the limited genomic coding capacity of the organism (15). As such, T. pallidum must rely on the uptake of preformed nucleosides, nucleotides, or nitrogenous bases to satisfy its replicative needs. The human host is able to synthesize purines (54, 55) and thereby provide them to T. pallidum as it replicates within human tissues and body fluids. That T. pallidum presumably lacks, however, porins or other substrate-specific outer membrane channels, prompts the question as to how vital nutrients, such as nucleosides, traverse the treponemal outer membrane. Simple diffusion of nutrients across the treponemal outer membrane, ostensi-
bly an inefficient process, has been a popular hypothesis consistent with the long mean generation time of 30–33 h postulated for T. pallidum (56). A more recent study by Hazlett et al. (57), however, has provided new evidence to suggest that Tp0453, a lipoprotein likely tethered to the inner leaflet of the treponemal outer membrane, has membrane-inserting amphipathic α-helices that transiently perturb the outer membrane, thereby culminating in enhanced outer membrane permeability. Inasmuch as T. pallidum does not possess orthologs of Txs (15, 44) or other porins or substrate-specific channels (4, 15, 58), it remains likely that the organism first avails itself of nucleosides by allowing them to traverse the outer membrane via some variation(s) of simple diffusion.

After traversing the treponemal outer membrane via simple diffusion, it is likely that Tp0319 serves as the receptor for purine nucleosides within the treponemal periplasm. There, nucleoside binding and transport across the cytoplasmic membrane via Tp0319 and its ABC partners presumably occurs. Unfortunately, the constraint of not being able to cultivate or genetically manipulate T. pallidum precludes the direct testing of this hypothesis. Once within the treponemal cytoplasm, however, the purine nucleosides likely are interconverted by salvage pathways; genomic analyses predict that T. pallidum encodes all of the enzymes required for nucleoside interconversion (15).

The finding that Tp0319 is part of a purine nucleoside ABC receptor/transporter system also is consistent with the increasing body of evidence that T. pallidum exploits lipoproteins as periplasmic receptors (18). In fact, all of the lipoproteins identified in T. pallidum to date are believed to be tethered via their three long-chain fatty acids either to the outer leaflet of the cytoplasmic membrane or to the inner leaflet of the outer membrane. In the former case, such lipoproteins are ideally suited as substrate-specific receptors, functioning analogously to receptors that are cytoplasmic membrane- or peptidoglycan-tethered in Gram-positive bacteria (to compensate for the lack of a periplasm) (59). In Gram-negative bacteria, periplasmic receptors tend to be free within the periplasm (32). That T. pallidum, despite its dual membrane system, has membrane architectural features that more closely resemble those of Gram-positive rather than Gram-negative bacteria further underscores the peculiar membrane biology of this enigmatic human pathogen (2–4, 48).

ABC transporters typically are named on the basis of the type of ligand bound by the receptor. Inasmuch as Tp0319 preferentially binds purine nucleosides, we propose re-naming Tp0319 as a purine nucleoside receptor A (PnrA). Furthermore, given transcriptional linkage of tp0319 to tp0320, tp0321, tp0322, and tp0323, we suggest denoting this operon of tp0319, tp0320, tp0321, tp0322, and tp0323 as pnrABCDE, respectively. The pnrABCDE operon is the first ABC-type transporter of purine nucleosides described for any bacterium. Finally, discerning for the first time the structure/function of an HBG016869 protein family member (Tp0319) may shed new light on the heretofore unknown function(s) of other members of the HBG016869 protein family.

Acknowledgments—We thank Lori Neil for technical assistance and critical review of the manuscript, Sandra Hill for technical assistance, and the staff of the Protein Chemistry Technology Center (University of Texas Southwestern) for mass spectrometry and protein sequencing.

REFERENCES
1. Norris, S. J. (1993) Microbiol. Rev. 57, 750–779
2. Cox, D. L., Chang, P., McDowell, A., and Radolf, J. D. (1992) Infect. Immun. 60, 1076–1083
3. Radolf, J. D. (1994) Trends Microbiol. 2, 307–311
4. Radolf, J. D. (1995) Mol. Microbiol. 16, 1067–1073
5. Chamberlain, N. R., Brandt, M. E., Erwin, A. L., Radolf, J. D., and Norgard, M. V. (1989) Infect. Immun. 57, 2872–2877
6. Schols, L. M., Moet, R., Dekker, J., and Van Embden, J. D. A. (1989) Microbial. Pathogen. 7, 175–188
7. Hayashi, S., and Wu, H. C. (1990) J. Bioenerg. Biomembr. 22, 451–471
8. Sutcliffe, I. C., and Russell, R. B. (1995) J. Bacteriol. 177, 1123–1128
9. Becker, P. S., Akins, D. R., Radolf, J. D., and Norgard, M. V. (1994) Infect. Immun. 62, 1381–1391
