JOURNAL OF BACTERIOLOGY, Apr. 2009, p. 2585–2592 Vol. 191, No. 8
0021-9193/09/$08.00+0 doi:10.1128/JB.01680-08

Structural Insights into the Extracytoplasmic Thiamine-Binding Lipoprotein p37 of *Mycoplasma hyorhinis*▽‡

Katherine H. Sippel,[1] Arthur H. Robbins,[1] Robbie Reutzel,[1]† Susan K. Boehlein,[2] Kazunori Namiki,[3] Steve Goodison,[4] Mavis Agbandje-McKenna,[1] Charles J. Rosser,[3] and Robert McKenna[1]*

Department of Biochemistry and Molecular Biology,[1] Program in Plant Molecular and Cellular Biology and Horticultural Sciences,[2] and Department of Urology,[3] University of Florida, Gainesville, Florida 32610, and Department of Surgery Jacksonville, Shands Health Science Center, Jacksonville, Florida 32209[4]

Received 1 December 2008/Accepted 9 February 2009

The *Mycoplasma hyorhinis* protein p37 has been implicated in tumorigenic transformation for more than 20 years. Though there are many speculations as to its function, based solely on sequence homology, the issue has remained unresolved. Presented here is the 1.6-Å-resolution refined crystal structure of *M. hyorhinis* p37, renamed the extracytoplasmic thiamine-binding lipoprotein (Cypl). The structure shows thiamine pyrophosphate (TPP) and two calcium ions are bound to Cypl and give the first insights into possible functions of the Cypl-like family of proteins. Sequence alignments of Cypl-like proteins between several different species of mycoplasma show that the thiamine-binding site is likely conserved and structural alignments reveal the similarity of Cypl to various binding proteins. While the experimentally determined function of Cypl remains unknown, the structure shows that the protein is a TPP-binding protein, opening up many avenues for future mechanistic studies and making Cypl a possible target for combating mycoplasma infections and tumorigenic transformation.

Mycoplasmas (class *Mollicutes*) are the smallest organisms capable of unassisted replication. These tiny, pleomorphic, wall-free bacteria survive in tandem with eukaryotes, either attached to the cell membrane or intracellularly. Only 0.2 to 0.3 μm in diameter, they also have one of the smallest functional genomes, 500 to 800 kbp. *Mycoplasma hyorhinis* is generally considered to be a swine pathogen, causing lung lesions, chest and abdominal lining inflammation, and arthritis (30). However, in recent years, an increasing body of work suggests that chronic mycoplasma infections, including *M. hyorhinis*, may be a source of oncogenic transformation (5, 15, 16, 29, 36, 52, 56, 57).

M. hyorhinis p37 was originally identified as a 37-kDa protein that induced invasiveness in mouse sarcoma cells. This invasivity could be reversed if the cells were pretreated with the antibody to this protein (48, 49). Several years later, a study was conducted on patients immunized intralymphatically with their own surgically debulked tumors looking for antibodies that induced remission and their associated antigens. This study also identified p37 (13). The presence of this antigen as an oncogenic marker was confirmed by Ilantzis et al. (24).

p37 is a misnomer; the actual protein is 43.5 kDa, composed of 403 residues. The operon it resides in bears some sequence similarity to a periplasmic binding protein-dependent transport system found in gram-negative bacteria. It is believed to function analogously as an extracytoplasmic binding lipoprotein, given that mycoplasma have only one membrane (9, 18). Since p37 was identified, several studies have indicated a correlation between the bacterium *M. hyorhinis* and gastric carcinoma, colon carcinoma, esophageal cancer, and lung cancer (23, 41). p37 alone is sufficient to increase the invasivity of cancer cells (26, 48, 49). It has also been shown that p37 may also be associated with cancer malignancy and metastasis (20, 21, 28).

In an effort to give p37 a more descriptive, identifiable, and accurate name, the protein is now designated the extracytoplasmic thiamine-binding lipoprotein (Cypl). Presented here is a 1.6-Å-resolution refined crystal structure of Cypl. Utilizing this structure, as well as sequence alignments to other mycoplasmal Cypl-like proteins and structural alignments to other proteins, the putative role of Cypl is addressed. Evidence will be offered which supports the current hypothesis that Cypl is an extracytoplasmic binding lipoprotein (9, 18). This paper proposes that Cypl is, specifically, a thiamine pyrophosphate (TPP)-binding protein rather than an enzyme utilizing TPP as a cofactor. Furthermore, sequence data suggest that the other mycoplasmal Cypl-like proteins may also bind thiamine or similar molecules.

MATERIALS AND METHODS

Expression and purification. An N-terminally truncated form of Cypl (residues 24 to 403) was expressed and purified using methods previously described by Ketcham et al. (26). Plasmid pMH38-113 contained the entire coding sequence for Cypl. The first 23 residues of Cypl contain the signal sequence, and the lipoprotein moiety was removed to enhance purification. The PCR primers used to amplify the Cypl coding region minus the leader sequence were 5′-ATGTGTTCTAACACCGGTGTAGTTAAG-3′ and 5′-GAATTCTTATTTAATGGCTTTTTCATAAAC-3′. Additionally, all of the TGA codons (mycoplasmal codon for Trp) were changed to TGG to optimize its expression in *E. coli*. This product was ligated into the expression vector pET31f1m1 (gift of P. Laipis,

* Corresponding author. Mailing address: 100 Newell Dr. LG-171, Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Gainesville, FL 32610. Phone: (352) 392-5696. Fax: (352) 392-3422. E-mail: rmckenna@ufl.edu.

‡ Supplemental material for this article may be found at http://jb.asm.org/.

† Present address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454.

▽ Published ahead of print on 20 February 2009.

TABLE 1. Mycoplasma Cypl-like proteins used for sequence alignment

Organism	Entrez Protein description	Entrez Protein accession code	Host organism	Sequence length (aa)	Reference or source
M. hyorhinis	p37	CAA32357	Swine	403	9
M. genitalium	G37, phosphonate ABC transporter, substrate binding protein, putative	NP_072955	Human	368	19
M. hyopneumoniae 7448	p37-like ABC transporter, substrate binding lipoprotein	YP_287754	Swine	421	53
M. pneumoniae M129	High-affinity transport system protein p37	NP_110103	Human	380	7
M. penetrans HF-2	High-affinity transport system protein	NP_758050	Human	402	42
M. pulmonis	p37-like ABC transporter, substrate binding lipoprotein	CAC13398	Murine	412	2
M. gallisepticum	High-affinity transport protein A	AAF78031	Avian	366	DS[a]
M. agalactiae PG2	Alkylphosphonate ABC transporter, substrate binding protein	YP_001256411	Ovine, caprine	438	45
M. synoviae 53	Putative p37-like ABC transporter, substrate binding lipoprotein	YP_278219	Avian	404	53
M. arthritidis 158L3–1	p37-like ABC transporter, substrate binding lipoprotein	ACF07046	Murine	380	10
M. mycoides subsp. *mycoides* SC PG1	ABC transporter, substrate-binding protein	NP_975778	Bovine	477	55
M. capricolum subsp. *capricolum* ATCC 27343	ABC transporter, substrate-binding protein	YP_424691	Caprine	486	19

[a] DS, direct submission.

Department of Biochemistry, University of Florida, Gainesville, FL). *E. coli* strain BL21(DE3)/pLysS was subsequently transformed with the plasmid.

The protein was expressed by inoculating 1 liter of minimal medium (37°C) supplemented with tryptone (10 g/liter), ampicillin (100 μg/ml), chloramphenicol (30 μg/ml), and glucose (0.75%, wt/vol). The culture was induced at an optical density at 600 nm of 0.7 to 1.0 with a final concentration of 1 mM isopropyl-L-thio-β-D-galactopyranoside (IPTG). The cells were grown for an additional 2.5 h at 37°C and harvested.

The cells were lysed by vortexing the pellet in 1/10 the original volume of 20 mM phosphate buffer (pH 7.8), followed by three 15-s sonication cycles. The resulting crude cell lysate was centrifuged at 40,000 × *g* for 20 min at 4°C. The supernatant underwent ion-exchange chromatography using a 5-ml Bio-Rad Econo-Pac S cation-exchange column attached to the bottom of a 50-ml Bio-Rad anion-exchange Q column and equilibrated with 20 mM sodium phosphate buffer (pH 7.95) at a flow rate of 2.5 ml/min. Approximately 125 mg of total protein was loaded on the column. The flowthrough containing Cypl was adjusted to pH 6.1 with 2 M acetic acid, loaded on a 5-ml cation exchanger (Econo-Pac S cartridge; Bio-Rad), and equilibrated with 20 mM sodium acetate (pH 6.1) (buffer A). The column was washed with 5% buffer B (20 mM sodium acetate [pH 6.1], 1 M NaCl), and the Cypl protein was eluted with 15% buffer B. The eluted sample was then concentrated using a Centriprep 10 spin column (Millipore, Bedford MA). The yield of Cypl was estimated to be ~30 mg of purified Cypl from 1 liter of *E. coli*.

Crystallization and diffraction data collection. The crystallization conditions for Cypl were those previously reported by Reutzel et al. (38). Briefly, crystals were grown using the microbatch method under oil at room temperature (4). Crystallization drops were prepared by mixing 7 μl protein at a concentration of ~10 mg ml^{-1} in 50 mM Tris-HCl (pH 7.5) with 3 μl of the precipitant solution (100 mM citric acid at pH 4.0 containing 40% PEG 4000 and 100 mM NH_4Br) under paraffin oil. Useful crystals appeared within 7 days.

Diffraction data were collected at Cornell High Energy Synchrotron Source (CHESS), beamline F2 at a wavelength of 0.9169 Å. The highest-resolution data resulted from a crystal soaked in 10 mM NaBr for 10 s during an extensive search for phasing derivatives. Data were collected using the oscillation method in intervals of 1° steps on an ADSC Quantum 4 charge-coupled-device detector, with a crystal-to-detector distance of 120 mm. A total of 130 images were collected on a flash-cryocooled crystal at 100 K. Indexing, integration, and scaling were performed using HKL2000, resulting in a 94.6% complete data set to 1.6-Å resolution with an R_{sym} of 7.1% (35). Complete processing statistics and the equation for calculating R_{sym} are given in Table S1 in the supplemental material.

Structural refinement. The structure was solved exclusively by molecular replacement. The starting model for the phasing was the previously solved 1.9-Å-resolution structure of Cypl (PDB code 3E78) with solvent molecules and TPP removed (44). SHELXL from the SHELX97 program suite was used to refine the model, alternating with rounds of model building using the COOT molecular graphics package (11, 43). Protein restraints used in SHELXL were from Engh and Huber, and restraints for TPP and glycerol were generated in SHELXPRO (12, 43). No restraints were used for the interactions involving calcium or bromine ions. The temperature factor for the more tightly bound calcium ion was refined anisotropically. The final model contains Cypl, TPP, one glycerol molecule, two calcium ions, seven bromide ions, and 285 water molecules, with an $R_{all\ data}$ of 18.4% and with good stereochemistry, as shown by PROCHECK (27). The final refinement statistics are presented in Table S2 in the supplemental material. There is a phenylalanine at position 256 in the deposited sequence of Cypl; however, the density at this position is consistent with a serine (44). To confirm this assertion, the plasmid was sequenced (Genomics Core, Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL); additionally, a simulated annealing omit map was generated, and it clearly showed electron density at position 256 corresponding to serine (9; data not shown).

Sequence alignments and phylogenic tree analysis. Sequence alignment and phylogenetic analysis were performed on 12 mycoplasma Cypl-like proteins. Species, description, host organism, and Entrez Protein accession number are listed in Table 1 (http://www.ncbi.nlm.nih.gov/sites/entrez?db=protein).

Sequence alignments were performed using ClustalW in the SDSC Biology Workbench (50, 51). Pairwise and multiple sequence alignments were performed using the Gonnet Series weight matrix, with an open gap penalty of 11.0 and a gap extension penalty of 0.85. All other parameters were set at default. An unrooted tree was generated using PHYLIP, also within the Biology Workbench (14, 50).

Structural alignment. A structural alignment of Cypl was performed using the DALI server (22). The top four classes of proteins were chosen for secondary structural matching superposition in COOT (11). The name, PDB ID, and DALI statistics are listed in Table 2.

Protein coordinate accession number. Coordinates and structure factors have been submitted to the Protein Data Bank (code 3EKI).

RESULTS

Overall fold of Cypl. The Cypl structure, as described previously (365 ordered amino acids), is arranged into two approximately equal-sized compact domains (domains I and II), separated by a deep cleft at the interface of the domains. Domain I consists of a six-stranded β sheet flanked by seven helices, and domain II has five β strands and five helices (Fig. 1A). The domains are connected by two linker regions and the

TABLE 2. Proteins with structural similarity to Cypl

PDB ID (chain ID)	Description	Z score	Root mean square deviation (Å) (no. of Cα)	No. of residues	% Identity	Reference
1JQF(A)	Human transferrin N-lobe mutant H249Q	16.1	4.1 (250)	329	13	1
3B7D(H)	GLUR2 ligand binding core (HS1S2J) in complex with CNQX	8.9	4.3 (210)	258	10	30
2QRY(C)	Periplasmic thiamine binding protein	6.9	4.3 (190)	318	10	47
1Z7N(G)	ATP phosphoribosyl transferase (HisZG ATP-PRTase) from *Lactococcus lactis* with bound PRPP[a] substrate	7.7	3.6 (164)	203	9	3

[a] PRPP, 5-phosphoribosyl-1-pyrophosphate.

C-terminal helix. A hydrophilic loop extends partially above the cleft entrance (residues 55 to 63) (44). A map of the electrostatic surface of Cypl shows a typical protein electrostatic profile, with no significant exposed hydrophobic patches (Fig. 1B). Of note is the patch of basic residues along the inside of the cleft corresponding to the TPP-binding site characterized below.

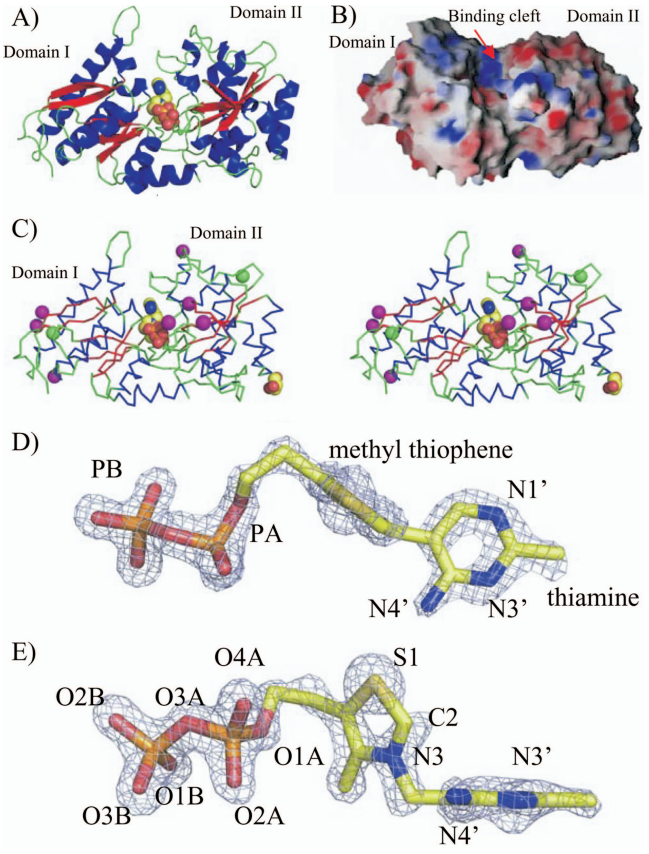

FIG. 1. Overall structure of Cypl. (A) Ribbon diagram depicting the domain (domains I and II) and secondary structural topology organization (α-helices, blue; β-strands, red; and coils, green). TPP is shown in spheres (carbon, yellow; nitrogen, blue; oxygen, red; and phosphorus, orange). (B) Surface charge distribution. Blue indicates regions of positive charge and red negative charge. (C) Stereo coil Cα tracing of Cypl, indicating the binding sites of TPP (coloring as described above), Ca^{2+} (green spheres), Br^{-1} (pink spheres), and glycerol (carbon, yellow; oxygen, red). (D and E) Fo-Fc electron density (blue mesh) omit map contoured at 1.5σ of TPP. View E is a 90° rotation of the orientation shown in D. Panels A, C, D, and E were generated using Pymol (8). Panel B was generated using Grasp (33).

TPP-binding site. The most surprising finding from the crystal structure determination of Cypl was the presence of TPP (Fig. 1A and C) (44). At 1.6-Å resolution, little doubt remains that it is truly TPP, as shown in the simulated annealing omit electron density map (Fig. 1D and E). The TPP is bound between the two domains and makes numerous interactions with both the main chain and the side chains of the adjacent amino acids. In brief, the methyl amino pyrimidine ring interacts with Asn200, Val201, Glu308, Trp314, and Leu379; the thiazole ring interacts with Leu379, Tyr343, Asp344, and Tyr215; and the pyrophosphate interacts with Lys258, Lys129, Tyr215, Ser255, Ser256, and Ser257 (Fig. 2 and Table 3).

Calcium-binding sites. Two calcium-binding sites are located toward the distal ends of domains I and II (Fig. 1C, green spheres). The calcium ion with lower temperature factors (site A) is located on domain I, coordinated by both carboxyl oxygen atoms of Asp121 and one oxygen atom of Asp43, the carbonyl oxygen atoms of Pro118 and Ser45, and two water molecules (Fig. 3A; Table 3). The observed well-ordered electron density and low refined temperature factor indicate that this calcium is very tightly bound to Cypl.

The other calcium-binding site (site B) is located in a more solvent-accessible region of domain II coordinated by the carbonyl oxygen atoms of Lys327 and Glu330, the OG1 atom of Thr326, and two water molecules (Fig. 3B and Table 3). The temperature factor of this calcium ion is approximately twofold higher than that of the site A calcium ion, consistent with its more exposed location. As no calcium was used during either purification or crystallization, this leads to the possibility that calcium ions, particularly site A, may play a functional role in Cypl activity.

Glycerol- and bromine-binding sites. As glycerol and bromine were additives soaked into crystal for X-ray diffraction data collection, they most likely do not contribute to functional analysis and therefore are not discussed further. However, a description of both the glycerol- and bromine-binding sites are provided in the supplemental material.

Comparative structural analysis. The Cypl structure was submitted to the DALI server to identify several varieties of proteins with similar structural homologies (22). All of the proteins listed were some form of binding protein. There were 504 structures with Z-scores of greater than two, and of those, none had greater than 16% identity to the sequence of Cypl. Of the structures obtained, there was no root mean square

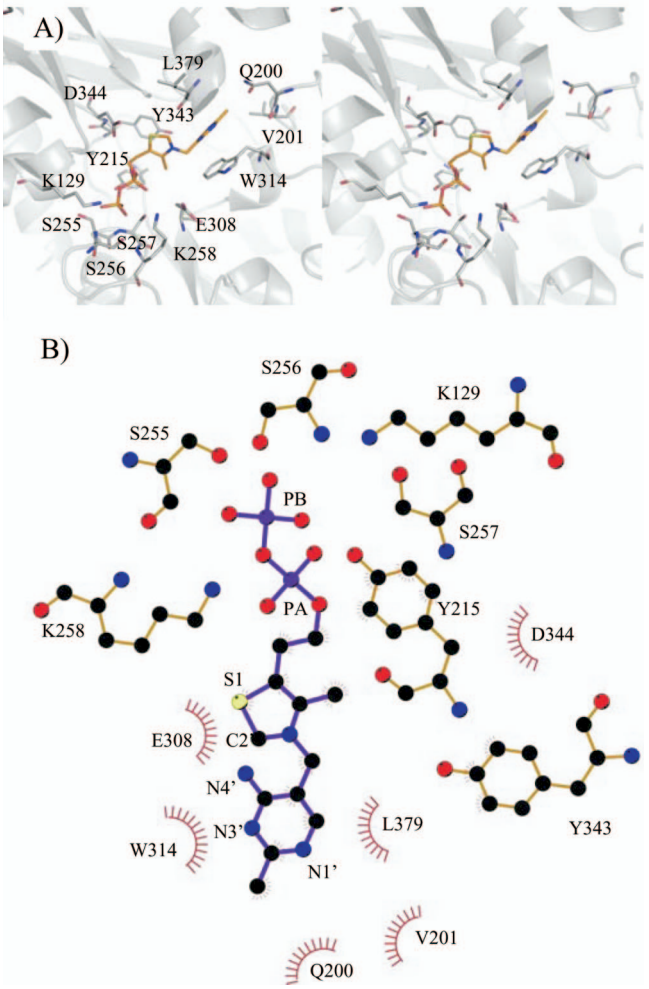

FIG. 2. TPP-binding site. (A) Stereo image showing the location of TPP within the Cypl cleft. (B) Protein-ligand interactions. Panel A was generated using Pymol (8). Panel B was generated using Ligplot (54).

deviation of less than 3.4 Å. Interestingly, the first 76 matches were to some variant of transferrin (sero-, ovo-, lacto-). The top match was to the H249Q mutant N-lobe of human serum transferrin (PDB code 1JFQ), with a root mean square deviation of 4.1 Å for 250 Cα positions (1). Several other proteins, including a type II ionotropic glutamate-binding domain (ADMA receptor), a periplasmic thiamine-binding protein (TbpA), and an ATP phosphoribosyl transferase (PRT), were selected for comparison and were structurally superpositioned onto Cypl using COOT (3, 11, 31, 47). A sample superposition of Cypl and TbpA is shown in Fig. 4. The PDB ID codes, descriptions, and statistics are given in Table 2.

DISCUSSION

Cypl and other Cypl-like proteins. The sequence of Cypl was compared to the sequences of Cypl-like proteins in 11 other species of mycoplasmas (Table 1). Five of the aligned sequences are shown in Fig. 5. There is no significant clustering of sequence identity between the various species, with most pairwise agreements ranging between 20% and 40% identity. However, there are a few exceptions. *M. mycoides*, *Mycoplasma capricolum*, and *Mycoplasma agalactiae*, all of which infect ruminants, show significant identity, up to 70% in the case of *M. mycoides* and *M. capricolum*, as well as 58% identity between *Mycoplasma genitalium* and *Mycoplasma pneumoniae*, two human pathogens (see Table S3 in the supplemental material). This is reflected in the phylogenetic tree (Fig. 6 and Table 1; see also Table S3 in the supplemental material). Each species deviates in a similarly diverse manner from the others. The tree is almost completely centrally rooted with little grouping. Given that two sequences only need 25% identity to be structurally homologous, it is likely that all of these proteins will share very similar structures overall (39).

It is important to note that several residues associated with TPP binding and calcium site A in Cypl are identical or conserved between species (Table 3 and Fig. 2, 3, and 5). All residues described use Cypl numbering. Within the TPP-binding site residues, Ser255, Tyr343, Tyr215, Trp314, Asp344, Glu308, and Lys258 were identical or showed significant con-

TABLE 3. Ligand interactions of Cypl

Ligand	Bond or contact[a]	Bording residue(s) (distance [Å])
TPP	O1B HB	Tyr215 (2.6), Ser257 (2.9 and 2.7)
	O2B HB	Lys129 (3.0), Ser256 (2.9)
	O3B HB	Ser255 (2.5), Lys258 (3.0), Lys129 (3.5)
	O1A HB	Tyr215 (3.5)
	O2A HB	Lys258 (2.7), HOH 1062 (2.85)
	O3A HB	HOH 1073 (3.5)
	O4A HB	HOH 1073 (2.9), HOH 1048 (2.8)
	S1 HB	Asp344 (3.5)
	N3 HB	Tyr343 (3.3)
	N4′ HB	HOH 1126 (2.9)
	N3′ HB	HOH 1270 (3.1)
	N1′ HB	HOH 1027 (2.7)
	HC	Asn200, Val201, Glu308, Ser257, Ser255, Lys258, Trp314, Leu379, Asp344, Tyr215, HOH 1062
Ca^{2+} site A	HB	Asp121 (2.3 and 2.7), Pro118 (2.3), Asp43 (2.4), Ser45 (2.4), HOH 1011 (2.5), HOH 1075 (2.6)
Ca^{2+} site B	HB	Thr326 (2.4), Lys327 (2.5), Glu330 (2.3), HOH 1153 (2.3), HOH 1154 (2.3)

[a] HB, hydrogen bond; HC, hydrophobic contact.

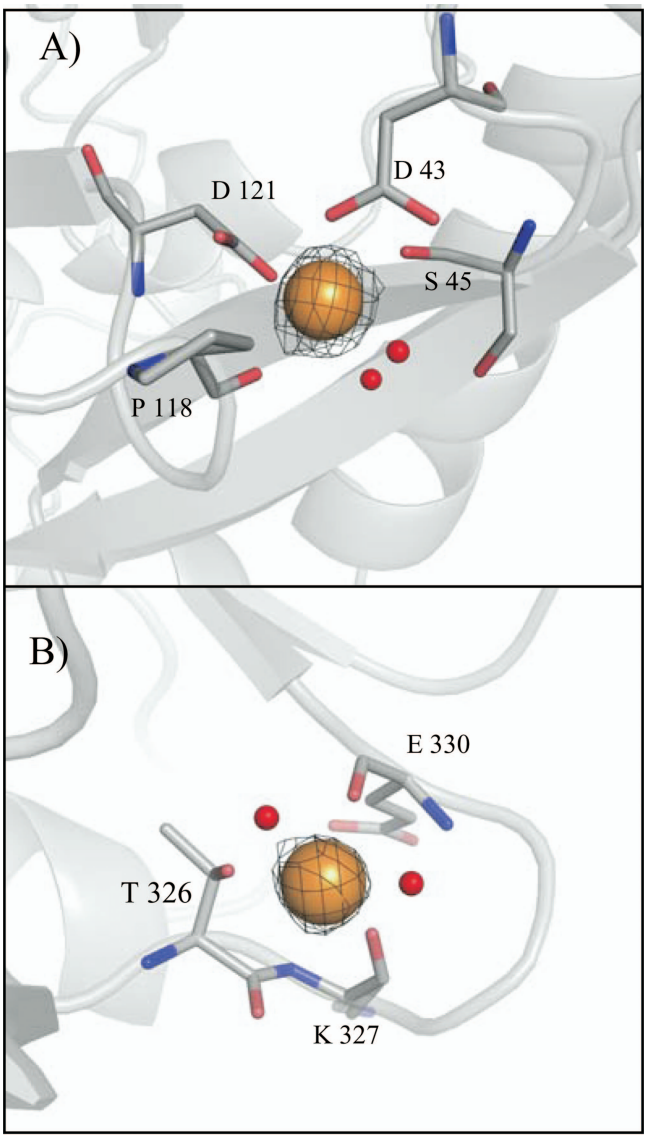

FIG. 3. Calcium-binding sites. Calcium site A (A) and site B (B) interactions. Waters shown as red spheres. Calcium shown in orange spheres. 2Fo-Fc electron density contoured to 1σ. The figure was generated using Pymol (8).

servation for all 12 species. Of interest are two of the residues necessary for pyrophosphate stabilization in the Cypl structure, Ser256 and -257. Ser256 (confirmed by sequencing) is a conserved Ala in eight sequences. Ser257 is only conserved in four sequences, the remaining eight being Gly. Given the role of Ser257 in stabilizing the second phosphate, it is likely that the other Cypl-like proteins bind thiamine monophosphate or another similar molecule. Although there is low sequence identity between the Cypl-like proteins, the overall degree of conservation between the thiamine-interacting residues is a strong indicator that they also bind thiamine in either the mono- or pyrophosphate variety (Fig. 5).

The calcium-binding sites in Cypl are less conserved than the thiamine-binding site. In calcium-binding site A, Asp121 is identical in 11 of 12 species and Asp43 is conserved in 9 of 12.

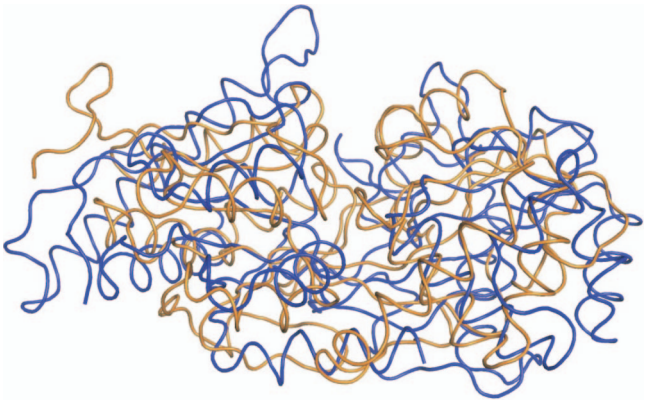

FIG. 4. Structural similarity of Cypl to *E. coli* TbpA. Coil C-alpha tracing of Cypl (blue), with TbpA (PDB code 2QRY chain C; color, orange). The figure was generated using Pymol (8).

However, the remaining two residues of site A (Pro118 and Ser45) and three residues of site B (Lys327, Glu330, and Thr326) show little sequence conservation (Fig. 3 and 5). This is not surprising considering that the calcium is bound to the carbonyl oxygens of these residues, with the exception of Thr326. Given this information and the proximity of these residues to flexible loop regions, a conserved calcium-binding site among the various Cypl-like proteins seems unlikely but cannot be ruled out.

Presence of TPP in Cypl. The conservation of thiamine-binding residues in the various species of mycoplasma reinforce the idea that thiamine could be present in all Cypl-like proteins. In an attempt to elucidate the function of TPP binding in Cypl, the crystal structures of several enzymes that require TPP for catalysis were examined. It was found that in all TPP-dependent enzymes, the TPP cofactors assume a similar V shape (25, 32). When the TPP conformation from yeast transketolase (PDB code 1TRK) is compared to the TPP conformation observed in Cypl, it is sterically misoriented for catalytic activity (Fig. 7A) (34). The TPP of Cypl more closely resembles the thiamine monophosphate observed in the crystal structure of a periplasmic thiamine-binding protein (TbpA) from *Escherichia coli* (PDB code 2QRY; Fig. 7B) (47). In gram-negative bacteria, the periplasmic binding protein is part of an ATP-binding cassette (ABC) transporter which also includes a membrane permease and an ATPase.

Why Cypl is probably not an enzyme. With these observations, it becomes apparent that TPP is probably not a cofactor for an active enzyme. The Cypl TPP is not in a catalytically active conformation. Also, Cypl lacks a conserved sequence necessary for loading the TPP into the protein as a cofactor. The orientation of the TPP as well as the surrounding protein environment renders it chemically inert. A more extensive discussion of these points is available in the supplemental material.

Why Cypl is probably a transport/binding protein. Genetic analysis performed by Dudler et al. in 1988 shows that Cypl is located on an *M. hyorhinis* 5.2-kb operon, including p29 and p69, which putatively codes for parts of an ABC transporter similar to those found in gram-negative bacteria (9, 18). The first open reading frame codes for Cypl and contains a 23-amino-acid export signal peptide which appears to be cleaved as the protein matures. The C-S-N sequence of the N terminus

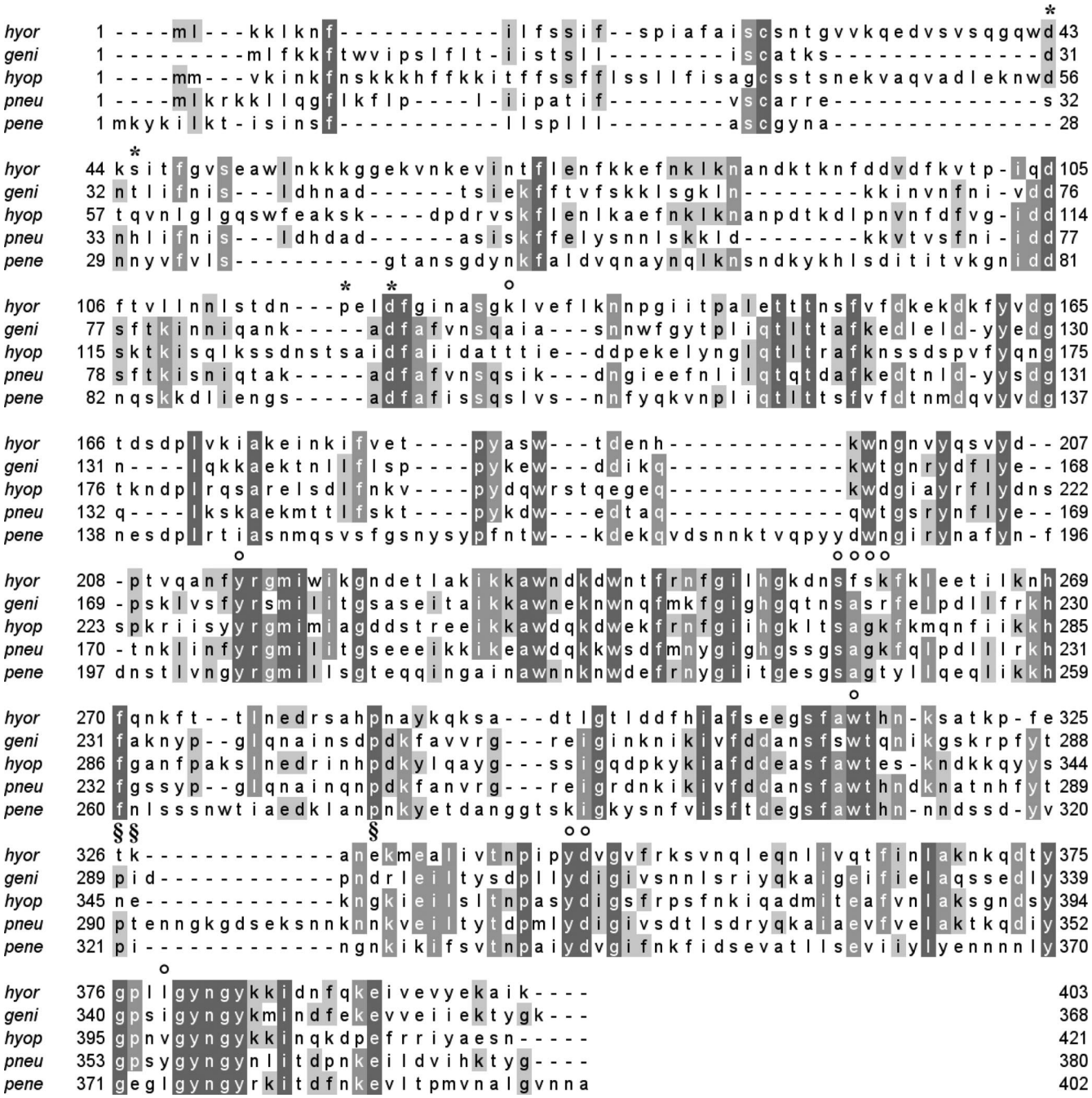

FIG. 5. Sequence alignment of Cypl-like proteins. hyor, *M. hyorhinis*; geni, *M. genitalium*; hyop, *M. hyopneumoniae*; pneu, *M. pneumoniae*; pene, *Mycoplasma penetrans*. Sequences are highlighted to indicate percentage identity and consensus. Symbols above columns denote residues involved in hydrogen bonding interactions: ○, TPP binding; ∗, Ca site A; and §, Ca site B. Dark, medium, and light grey indicate 100%, 80%, and 60% identity, respectively. The figure was generated using Jalview (6).

of the mature protein is consistent with a lipoprotein. The other two open reading frames code for p29 and p69, respectively. The overall operon is flanked by inverted repeats, indicating rho-independent transcription termination. p29 appears to have sequence similarity to bacterial proteins that interact with both substrate-binding proteins and integral membrane proteins. p69 bears a hydropathy plot similar to that of integral membrane proteins of ABC transport systems, though it does lack the EAA sequence typically seen in these proteins (9).

The overall structural topology of Cypl is fairly consistent with the type II periplasmic binding protein (Fig. 4) (17, 37). This is supported by the overwhelming majority of DALI matches being a binding protein. It has strong structural similarity to the N-lobe of transferrin but not the entire protein (44). However, there are some critical differences. Transferrin is a secreted protein and not membrane associated like Cypl (40). The other structurally similar proteins are not ideal functional models either, given their propensity for complex for-

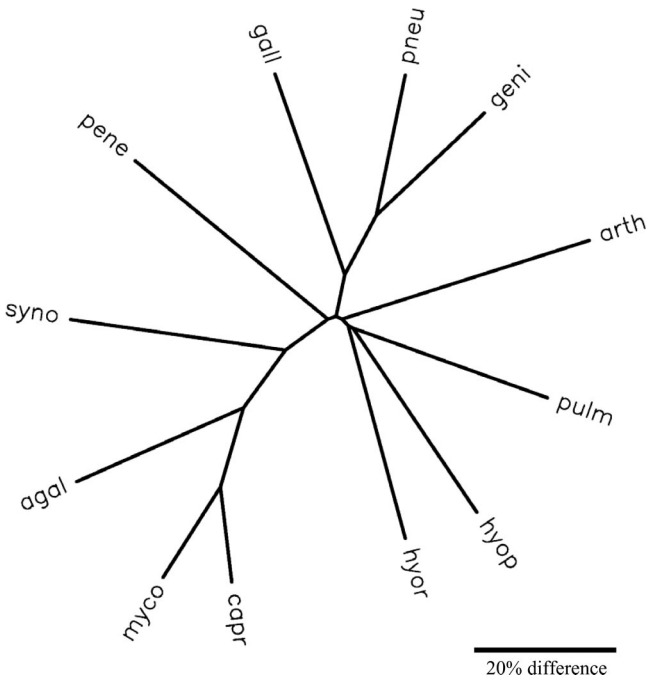

FIG. 6. An unrooted phylogenetic tree of mycoplasma Cypl-like proteins. hyor, *M. hyorhinis*; geni, *M. genitalium*; hyop, *M. hyopneumoniae*; pneu, *M. pneumoniae*; pene, *Mycoplasma penetrans*; pulm, *M. pulmonis*; gall, *M. gallisepticum*; agal, *M. agalactiae*; syno, *M. synoviae*; arth, *M. arthritidis*; myco, *M. mycoides*; capr, *M. capricolum*. The figure was generated in PHYLIP (14).

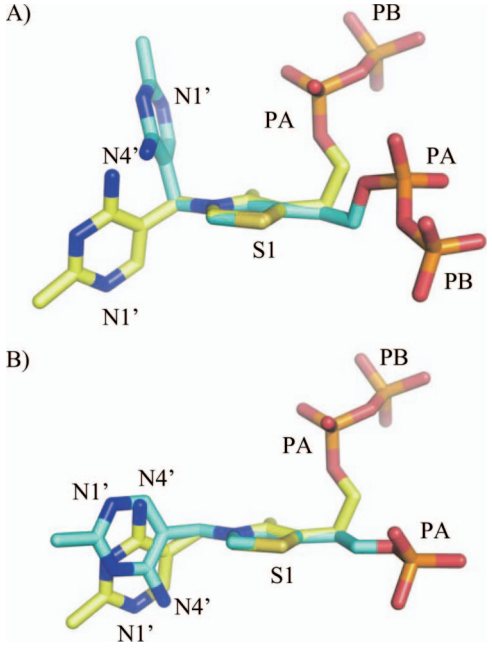

FIG. 7. Conformation of TPP. (A) TPP from yeast transketolase (PDB code 1TRK) superposed onto TPP of Cypl. (B) TPS from periplasmic thiamine-binding protein (PDB code 2QRY) superposed onto TPP of Cypl. The carbons of Cypl TPP are shown in yellow. The carbons of transketolase TPP and TPS are colored cyan. All other atoms are colored as follows: red, oxygen; blue, nitrogen; and orange, phosphorus. The figure was made using Pymol (8).

mation. The ADMA receptor and PRT are octamers (3, 31). Even the TbpA crystallizes as a dimer, which may or may not be physiologically relevant (47). Cypl has shown no indication of functioning as a dimer. In native gels (data not shown), the protein migrates predominantly as a single monomeric band and the electrostatic surface of the protein shows no hydrophobic patches along the opening of the binding cleft, indicative of periplasmic binding protein dimers (Fig. 1B) (3, 47). While it may behave as part of an ABC-binding cassette in mycoplasma, its role both in the virulence of mycoplasma and in tumorigenic transformation remains unclear. Cypl's role in mycoplasma infection has never been addressed. As for its effect on cancer cells, it has been shown that Cypl alone is sufficient to cause increased tumor invasivity, to the same degree as whole mycoplasma (26, 48, 49). Perhaps in vivo Cypl is involved in the transport of thiamine into the mycoplasma, which may sequester thiamine away from surrounding host cells, weakening them and creating an opportunity for infection. A precedent for this scenario exists and has been described for *Bacillus thiaminolyticus*, which is pathogenic because it degrades thiamine, thus depleting the supply for its host (47). Research has also shown that TPP or thiamine is a necessary component of lactate oxidation in *Mycoplasma gallisepticum* (46). It is not a great leap of logic to consider that a thiamine-binding protein would be an essential component for mycoplasma survival.

Additional studies are required to adequately identify a function for Cypl. Further biological and crystallographic studies are necessary to fully characterize this protein and its role in infection as well as the tumorigenic transformation of human cells. It is apparent that Cypl-like proteins may offer an important therapeutic target in future drug studies, either for the inhibition of Cypl-induced tumor progression or for ameliorating persistent pathogenic infections.

ACKNOWLEDGMENTS

We greatly appreciate Sheldon M. Schuster for introducing us to this fascinating protein.

This work was funded in part by American Cancer Society grant 58229 to C.J.R., the MacCHESS grant (U.S. NIH grant RR001646), U.S. DOE grant DE-FG02-97ER62443, and CHESS, which is supported by the U.S. NSF and NIH-NIGMS through NSF grant DMR-0225180. K.H.S. was also funded by the University of Florida, College of Medicine, Alumni Fellowship.

REFERENCES

1. **Baker, H. M., A. B. Mason, Q. Y. He, R. T. MacGillivray, and E. N. Baker.** 2001. Ligand variation in the transferrin family: the crystal structure of the H249Q mutant of the human transferrin N-lobe as a model for iron binding in insect transferrins. Biochemistry **40:**11670–11675.
2. **Chambaud, I., R. Heilig, S. Ferris, V. Barbe, D. Samson, F. Galisson, I. Moszer, K. Dybvig, H. Wroblewski, A. Viari, E. P. Rocha, and A. Blanchard.** 2001. The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis. Nucleic Acids Res. **29:**2145–2153.
3. **Champagne, K. S., M. Sissler, Y. Larrabee, S. Doublie, and C. S. Francklyn.** 2005. Activation of the hetero-octameric ATP phosphoribosyl transferase through subunit interface rearrangement by a tRNA synthetase paralog. J. Biol. Chem. **280:**34096–34104.
4. **Chayen, N. E., P. D. S. Stewart, D. L. Maeder, and D. M. Blow.** 1990. An automated system for micro-batch protein crystallization and screening. J. Appl. Crystallogr. **23:**297–302.
5. **Cimolai, N.** 2001. Do mycoplasmas cause human cancer? Can. J. Microbiol. **47:**691–697.
6. **Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton.** 2004. The Jalview Java alignment editor. Bioinformatics **20:**426–427.
7. **Dandekar, T., M. Huynen, J. T. Regula, B. Ueberle, C. U. Zimmermann, M. A. Andrade, T. Doerks, L. Sanchez-Pulido, B. Snel, M. Suyama, Y. P. Yuan, R. Herrmann, and P. Bork.** 2000. Re-annotating the Mycoplasma

pneumoniae genome sequence: adding value, function and reading frames. Nucleic Acids Res. **28:**3278–3288.
8. **DeLano, W. L.** 2002. The Pymol molecular graphics system. Delano Scientific, Palo Alto, CA.
9. **Dudler, R., C. Schmidhauser, R. W. Parish, R. E. Wettenhall, and T. Schmidt.** 1988. A mycoplasma high-affinity transport system and the in vitro invasiveness of mouse sarcoma cells. EMBO J. **7:**3963–3970.
10. **Dybvig, K., C. Zuhua, P. Lao, D. S. Jordan, C. T. French, A. H. Tu, and A. E. Loraine.** 2008. Genome of *Mycoplasma arthritidis*. Infect. Immun. **76:**4000–4008.
11. **Emsley, P., and K. Cowtan.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D **60:**2126–2132.
12. **Engh, R. A., and R. Huber.** 1991. Accurate bond and angle parameters for X-ray protein-structure refinement. Acta Crystallogr. A **47:**392–400.
13. **Fareed, G. C., E. Mendiaz, A. Sen, G. J. Juillard, T. H. Weisenburger, and T. Totanes.** 1988. Novel antigenic markers of human tumor regression. J. Biol. Response Mod. **7:**11–23.
14. **Felsenstein, J.** 1993. PHYLIP (phylogeny inference package), 3.5c ed. Department of Genetics, University of Washington, Seattle, WA.
15. **Feng, S. H., S. Tsai, J. Rodriguez, and S. C. Lo.** 1999. Mycoplasmal infections prevent apoptosis and induce malignant transformation of interleukin-3-dependent 32D hematopoietic cells. Mol. Cell. Biol. **19:**7995–8002.
16. **Fogh, J., and H. Fogh.** 1965. Chromosome changes in Pplo-infected Fl human amnion cells. Proc. Soc. Exp. Biol. Med. **119:**233–238.
17. **Fukami-Kobayashi, K., Y. Tateno, and K. Nishikawa.** 1999. Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. J. Mol. Biol. **286:**279–290.
18. **Gilson, E., G. Alloing, T. Schmidt, J. P. Claverys, R. Dudler, and M. Hofnung.** 1988. Evidence for high affinity binding-protein dependent transport systems in gram-positive bacteria and in Mycoplasma. EMBO J. **7:**3971–3974.
19. **Glass, J. I., N. Assad-Garcia, N. Alperovich, S. Yooseph, M. R. Lewis, M. Maruf, C. A. Hutchison III, H. O. Smith, and J. C. Venter.** 2006. Essential genes of a minimal bacterium. Proc. Natl. Acad. Sci. USA **103:**425–430.
20. **Gong, M., L. Meng, B. Jiang, J. Zhang, H. Yang, J. Wu, and C. Shou.** 2008. p37 from Mycoplasma hyorhinis promotes cancer cell invasiveness and metastasis through activation of MMP-2 and followed by phosphorylation of EGFR. Mol. Cancer Ther. **7:**530–537.
21. **Goodison, S., K. Nakamura, K. A. Iczkowski, S. Anai, S. K. Boehlein, and C. J. Rosser.** 2007. Exogenous mycoplasmal p37 protein alters gene expression, growth and morphology of prostate cancer cells. Cytogenet. Genome Res. **118:**204–213.
22. **Holm, L., and C. Sander.** 1996. Alignment of three-dimensional protein structures: network server for database searching. Methods Enzymol. **266:**653–662.
23. **Huang, S., J. Y. Li, J. Wu, L. Meng, and C. C. Shou.** 2001. Mycoplasma infections and different human carcinomas. World J. Gastroenterol. **7:**266–269.
24. **Ilantzis, C., D. M. Thomson, A. Michaelidou, S. Benchimol, and C. P. Stanners.** 1993. Identification of a human cancer related organ-specific neoantigen. Microbiol. Immunol. **37:**119–128.
25. **Jordan, F.** 2003. Current mechanistic understanding of thiamin diphosphate-dependent enzymatic reactions. Nat. Prod. Rep. **20:**184–201.
26. **Ketcham, C. M., S. Anai, R. Reutzel, S. Sheng, S. M. Schuster, R. B. Brenes, M. Agbandje-McKenna, R. McKenna, C. J. Rosser, and S. K. Boehlein.** 2005. p37 induces tumor invasiveness. Mol. Cancer Ther. **4:**1031–1038.
27. **Laskowski, R. A., D. S. Moss, and J. M. Thornton.** 1993. Main-chain bond lengths and bond angles in protein structures. J. Mol. Biol. **231:**1049–1067.
28. **Liu, W., T. Ren, B. Jiang, M. Gong, and C. Shou.** 2007. Mycoplasmal membrane protein p37 promotes malignant changes in mammalian cells. Can. J. Microbiol. **53:**270–276.
29. **Macpherson, I., and W. Russell.** 1966. Transformations in hamster cells mediated by mycoplasmas. Nature **210:**1343–1345.
30. **Maniloff, J., R. N. McElhaney, L. R. Finch, and J. B. Baseman (ed.).** 1992. Mycoplasmas: molecular biology and pathogenesis. American Society for Microbiology, Washington, DC.
31. **Menuz, K., R. M. Stroud, R. A. Nicoll, and F. A. Hays.** 2007. TARP auxiliary subunits switch AMPA receptor antagonists into partial agonists. Science **318:**815–817.
32. **Muller, Y. A., and G. E. Schulz.** 1993. Structure of the thiamine- and flavin-dependent enzyme pyruvate oxidase. Science **259:**965–967.
33. **Nicholls, A., K. A. Sharp, and B. Honig.** 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins **11:**281–296.
34. **Nikkola, M., Y. Lindqvist, and G. Schneider.** 1994. Refined structure of transketolase from Saccharomyces cerevisiae at 2.0 A resolution. J. Mol. Biol. **238:**387–404.
35. **Otwinowski, Z., and W. Minor.** 1997. Processing of X-ray diffraction data collected in oscillation mode, vol. 276. Yale University, New Haven, CT.
36. **Paton, G. R., J. P. Jacobs, and F. T. Perkins.** 1965. Chromosome changes in human diploid-cell cultures infected with Mycoplasma. Nature **207:**43–45.
37. **Quiocho, F. A., and P. S. Ledvina.** 1996. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. Mol. Microbiol. **20:**17–25.
38. **Reutzel, R., S. K. Boehlein, L. Govindasamy, R. B. Brenes, M. Agbandje-McKenna, S. M. Schuster, and R. McKenna.** 2002. Crystallization and preliminary X-ray analysis of the tumor metastasis factor p37. Acta Crystallogr. D **58:**2141–2144.
39. **Rost, B.** 1999. Twilight zone of protein sequence alignments. Protein Eng. **12:**85–94.
40. **Sargent, P. J., S. Farnaud, and R. W. Evans.** 2005. Structure/function overview of proteins involved in iron storage and transport. Curr. Med. Chem. **12:**2683–2693.
41. **Sasaki, H., H. Igaki, T. Ishizuka, Y. Kogoma, T. Sugimura, and M. Terada.** 1995. Presence of Streptococcus DNA sequence in surgical specimens of gastric cancer. Jpn. J. Cancer Res. **86:**791–794.
42. **Sasaki, Y., J. Ishikawa, A. Yamashita, K. Oshima, T. Kenri, K. Furuya, C. Yoshino, A. Horino, T. Shiba, T. Sasaki, and M. Hattori.** 2002. The complete genomic sequence of Mycoplasma penetrans, an intracellular bacterial pathogen in humans. Nucleic Acids Res. **30:**5293–5300.
43. **Sheldrick, G. M.** 2008. A short history of SHELX. Acta Crystallogr. A **64:**112–122.
44. **Sippel, K. H., A. H. Robbins, R. Reutzel, J. Domsic, S. K. Boehlein, L. Govindasamy, M. Agbandje-McKenna, C. J. Rosser, and R. McKenna.** 2008. Structure determination of the cancer-associated Mycoplasma hyorhinis protein Mh-p37. Acta Crystallogr. D **64:**1172–1178.
45. **Sirand-Pugnet, P., C. Lartigue, M. Marenda, D. Jacob, A. Barre, V. Barbe, C. Schenowitz, S. Mangenot, A. Couloux, B. Segurens, A. de Daruvar, A. Blanchard, and C. Citti.** 2007. Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. PLoS Genet. **3:**e75.
46. **Smith, S. L., P. J. Vandemark, and J. Fabricant.** 1963. Respiratory pathways in the *Mycoplasma*. I. Lactate oxidation by *Mycoplasma gallisepticum*. J. Bacteriol. **86:**893–897.
47. **Soriano, E. V., K. R. Rajashankar, J. W. Hanes, S. Bale, T. P. Begley, and S. E. Ealick.** 2008. Structural similarities between thiamin-binding protein and thiaminase-I suggest a common ancestor. Biochemistry **47:**1346–1357.
48. **Steinemann, C., M. Fenner, H. Binz, and R. W. Parish.** 1984. Invasive behavior of mouse sarcoma cells is inhibited by blocking a 37,000-dalton plasma membrane glycoprotein with Fab fragments. Proc. Natl. Acad. Sci. USA **81:**3747–3750.
49. **Steinemann, C., M. Fenner, R. W. Parish, and H. Binz.** 1984. Studies of the invasiveness of the chemically induced mouse sarcoma FS9. I. Monoclonal antibodies to a 37,000 dalton membrane glycoprotein inhibit invasion of fibroblasts in vitro. Int. J. Cancer **34:**407–414.
50. **Subramaniam, S.** 1998. The Biology Workbench—a seamless database and analysis environment for the biologist. Proteins **32:**1–2.
51. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
52. **Tsai, S., D. J. Wear, J. W. Shih, and S. C. Lo.** 1995. Mycoplasmas and oncogenesis: persistent infection and multistage malignant transformation. Proc. Natl. Acad. Sci. USA **92:**10197–10201.
53. **Vasconcelos, A. T., H. B. Ferreira, C. V. Bizarro, S. L. Bonatto, M. O. Carvalho, P. M. Pinto, D. F. Almeida, L. G. Almeida, R. Almeida, L. Alves-Filho, E. N. Assuncao, V. A. Azevedo, M. R. Bogo, M. M. Brigido, M. Brocchi, H. A. Burity, A. A. Camargo, S. S. Camargo, M. S. Carepo, D. M. Carraro, J. C. de Mattos Cascardo, L. A. Castro, G. Cavalcanti, G. Chemale, R. G. Collevatti, C. W. Cunha, B. Dallagiovanna, B. P. Dambros, O. A. Dellagostin, C. Falcao, F. Fantinatti-Garboggini, M. S. Felipe, L. Fiorentin, G. R. Franco, N. S. Freitas, D. Frias, T. B. Grangeiro, E. C. Grisard, C. T. Guimaraes, M. Hungria, S. N. Jardim, M. A. Krieger, J. P. Laurino, L. F. Lima, M. I. Lopes, E. L. Loreto, H. M. Madeira, G. P. Manfio, A. Q. Maranhao, C. T. Martinkovics, S. R. Medeiros, M. A. Moreira, M. Neiva, C. E. Ramalho-Neto, M. F. Nicolas, S. C. Oliveira, R. F. Paixao, F. O. Pedrosa, S. D. Pena, M. Pereira, L. Pereira-Ferrari, I. Piffer, L. S. Pinto, D. P. Potrich, A. C. Salim, F. R. Santos, R. Schmitt, M. P. Schneider, A. Schrank, I. S. Schrank, A. F. Schuck, H. N. Seuanez, D. W. Silva, R. Silva, S. C. Silva, C. M. Soares, K. R. Souza, R. C. Souza, C. C. Staats, M. B. Steffens, S. M. Teixeira, T. P. Urmenyi, M. H. Vainstein, L. W. Zuccherato, A. J. Simpson, and A. Zaha.** 2005. Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. J. Bacteriol. **187:**5568–5577.
54. **Wallace, A. C., R. A. Laskowski, and J. M. Thornton.** 1995. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. **8:**127–134.
55. **Westberg, J., A. Persson, A. Holmberg, A. Goesmann, J. Lundeberg, K. E. Johansson, B. Pettersson, and M. Uhlen.** 2004. The genome sequence of Mycoplasma mycoides subsp. mycoides SC type strain PG1T, the causative agent of contagious bovine pleuropneumonia (CBPP). Genome Res. **14:**221–227.
56. **Zhang, B., J. W. Shih, D. J. Wear, S. Tsai, and S. C. Lo.** 1997. High-level expression of H-ras and c-myc oncogenes in mycoplasma-mediated malignant cell transformation. Proc. Soc. Exp. Biol. Med. **214:**359–366.
57. **Zhang, B., S. Tsai, J. W. Shih, D. J. Wear, and S. C. Lo.** 1998. Absence of mycoplasmal gene in malignant mammalian cells transformed by chronic persistent infection of mycoplasmas. Proc. Soc. Exp. Biol. Med. **218:**83–89.