5-2009

Protozoan Vaccine Candidate Homologues in Tetrahymena thermophila

Justin Schumacher
University of Rhode Island

An Orthologue of the Protozoan Vaccine Candidate Phosphoprotein p0 in *Tetrahymena thermophila*

J. Schumacher, J. Corriveau , S. Canton, and L.A. Hufnagel

Department of Cell and Molecular Biology, University of Rhode Island

Abstract:

Tetrahymena thermophila (Ciliophora) is a member of the Alveolata, the major eukaryotic clade containing apicomplexan parasites responsible for serious human and animal diseases. We are using online genomic and gene expression resources to discover and characterize *T. thermophila* orthologues of apicomplexan genes coding for vaccine candidates. Phosphoprotein p0 belongs to a family of highly conserved ribosomal proteins that also are found at the cell surface. Antibodies against p0 block parasite invasion of the host, explaining why p0 is a target for vaccine development. Using BLAST functions at NCBI and TGD (Tetrahymena Genome Database), we identified the Tetrahymena orthologue of *Plasmodium falciparum* p0: a predicted protein coded by the gene TTHERM_ 00636970 *(*e-value: 2.7x10-35) with domains of strong homology with apicomplexan p0. We plan to determine if Ttp0 is also found at the cell surface. If so, Tetrahymena may be useful to discover how p0 goes to the cell surface rather than to ribosomes.

Background:

Phosphoprotein p0 is a member of the L10 family of ribosomal proteins that are found in all eukaryotes (Machler-Bauer, 2007). It is closely related to proteins P1 and P2, which are also ribosomal phosphoproteins. P0 forms complexes with P1 and P2, to serve as the GTPase stalk binding region of the 60s ribosomal subunit. Studies have shown that P0 is essential for protein synthesis and can function even in the absence of P1 and P2 (Rajeshwari, 2004). In addition to acting as a ribosomal protein, P0 has been shown to be able to localize on the cell surface of many organisms. Also, the pathway by which phosphoprotein p0 translocates to the cell surface remains currently unknown as no transmembrane domain or canonical signal sequence has been identified (Singh, 2002). When apicomplexan parasites are exposed to antibodies against phosphoprotein p0, their ability of cellular invasion is inhibited, indicating that phosphoprotein p0's role on the cell surface is that of invasion of host cells (Rajeshwari, 2004). For this reason, phosphoprotein p0 is now a vaccine candidate for many apicomplexan diseases and protozoan parasite diseases.

Tetrahymena thermophila is a member of the phylum Ciliophora and is currently of great interest due to its unique qualities that make it useful as a model for research directed towards understanding how eukaryotic cells function. Recently, the *T. thermophila* genome was sequenced (Eisen et al, 2006), and made available online (Stover et al, 2006). Since Tetrahymena is a member of the Alveolata, the major eukaryotic clade that also contains the apicomplexan parasites, we sought to determine whether the genes coding for apicomplexan vaccine candidate phosphoprotein p0 was represented in the genome of *Tetrahymena*. If present in the genome, the identified orthologue would then undergo sequence alignment, analyzed for motifs and included in a neighbor-joining tree.

Methods:

Indentifying Protein Sequences:

In order to identify homologues of protozoan vaccine candidate phosphoprotein p0 in *T.thermophila,* the amino acid sequences of the vaccine candidates had to be first identified. The sequences of these vaccine candidates were in apicomplexans as well as in non-apicomplexans. These sequences were found by searching the NCBI protein database (Entrez Protein). By using this database, the sequences of phosphoprotein p0 in *Babesia gibsoni* (Terkawi et al, 2007)*, Babesia microti* (Terkawi et al, 2007)*, Theileria annulata* (Pain et al, 2005)*, Neospora caninum* (Zhang et al, 2007)*, Toxoplasma gondii* (Sehgal et al, 2003)*, Plasmodium falciparum* (Goswami et al, 1996)*, Homo sapien(* Rich et al, 1987)*, Saccharomyces cerevisiae* (Albuquerque et al, 2008)*, Eufolliculina uhligi* (Markmann-Mulisch et al, 1999)*, Euplotes raikovi* (Pucciarelli et al, 2005)*, Euplotes minuta* (Pucciarelli et al, 2005)*, Paramecium tetraurelia* (Aury et al, 2006)*, Trypanosoma bruci* (El-Sayed et al, 2005)*, Trypanosoma cruzi* (Skeiky et al, 1992)*, Leishmania braziliensis* (Peacock et al, 2007) and *Tetrahymena thermophila* (Pucciarelli et al, 2005) were found.

Identifying Homologues in Tetrahymena:

The vaccine candidate phosphoprotein p0 sequences from the apicomplexans, *Plasmodium falciparum, Babesia gibsoni, Toxoplasma gondii* were entered into the BLASTP function against the TIGR Gene Predictions. This compared the vaccine candidate amino acid protein sequences to the predicted amino acid sequences of the genes in the Tetrahymena Genome Database (Stover et al, 2006). The genes with predicted amino acid sequences with the most homology to the inputted protein sequence were then recorded along with their corresponding E-value. The complete homologous *Tetrahymena* gene amino acid sequence was then found at the NCBI database (Entrez Protein).

Sequence Aligning and Protein Motif Identification:

The protein sequences were then aligned to find areas of high conservation. The alignment was created by entering the protein sequences into the ClustalW2 alignment program (Larkin et al, 2007). The areas of the protein sequences that were shown to have functional motifs were then checked against the alignment to see if these motifs were conserved. Three alignments were created. One with apicomplexans and *Tetrahymena*, a second with non-apicomplexans and *Tetrahymena*, and another with both apicomplexans and non-apicomplexans with *Tetrahymena.* This was done to see if the conservation of the different parts of the protein varied between apicomplexans and non-apicomplexans.

The predicted protein sequences from the Tetrahymena genome that showed homology to the vaccine candidates were then analyzed for functional motifs. Functional motifs were found by searching for known motif sequences in the phosphoprotein p0 amino acids sequences. The known motifs were found by using the motif finder at genome.jp (GenomeNet, 1991).

Neighbor-joining Tree formation

The ClustalW program at the Kyoto University Bioinformatics Center (GenomeNet, 1991) was utilized to create a neighbor joining tree for the phosphoprotein sequences. This neighbor joining tree consisted of organisms across the tree of life in order to determine where the *Tetrahymena* phosphoprotein p0 sequence branched off evolutionarily.

Results:

Protein BLAST function:

The apicomplexan phosphoprotein p0 sequences and non-apicomplexan phosphoprotein p0 sequences showed the most homology to the same predicted gene amino acid sequence in *Tetrahymena thermophila.* This gene is TTHERM_00636970. In a previous unrelated study this protein was identified as 60S Ribosomal Protein (Pucciarelli, 2005). This gene has a predicted amino acid sequence length of 324 amino acids and has been reported in a microarray analyses to be expressed 250x background continuously throughout all stages of the *Tetrahymena* life cycle (Miaw, 2009). Below are the e-values from the protein BLAST results:

Orgamism	E-value From Tetrahymena p0 Blast
Babesia Bovis	2.5x10^-32
Babesia Gibsoni	5.1x10^-32
Babesia Microti	5.4x10^-37
Cryptosporidium hominis	8.7x10^-37
Cryptosporidium parvum	1.1x10^-36
Leishmania braziliensis	1.3x10^-33
Leishmania infantum	2.2x10^-31
Leishmania major	5.9x10^-31
Neospora caninum	1.4x10^-36
Plasmodium falciparum	2.7x10^-35
Plasmodium vivax	9.7x10^-38
Theileria annulata	4x10^-32
Toxoplasma gondii	1.8x10^-36
Trypanosoma brucei	3.1x10^-32
Trypanosoma cruzi	1.1x10^-31

Table 1: E-values from the protein BLAST function at the Tetrahymena Genome Database (Stover, 2006)

The immunogen of commercially made antibodies (RPLP0, 1998) created against phosphoprotein p0 were then compared against the *Tetrahymena* predicted gene amino acid sequences by protein BLAST function. The resulting e-value was 4.0x10^-32. The immunogen

was shown to have the highest degree of homology towards the 85 to 271 amino acid section of the sequence.

Protein Sequence Alignment, Motifs And Neighbor-Joining Tree

The protein sequences for phosphoprotein p0 in apicomplexans and in *Tetrahymena* was aligned to search for areas of high conservation and several were noted. These regions went from the 7-68, 127-230, and 242-323 amino acids in the *Tetrahymena* phosphoprotein p0 sequence. The protein sequences used in the alignments were then inputted into a motif finder (GenomeNet, 1991) to see if the conserved regions consisted of any known motifs. The conserved region from the 7-68 amino acid section contained a motif for ribosomal L10 family proteins and the 242-323 amino acid section comprised a motif characteristic of 60S Ribosomal proteins. The conserved region from the 127-230 amino acids are still currently being analyzed for motifs. Similar alignments were performed with non-apicomplexans to see if these conserved regions were found in other organism across the tree of life. See below figures for the sequence alignments and highlighted motifs.

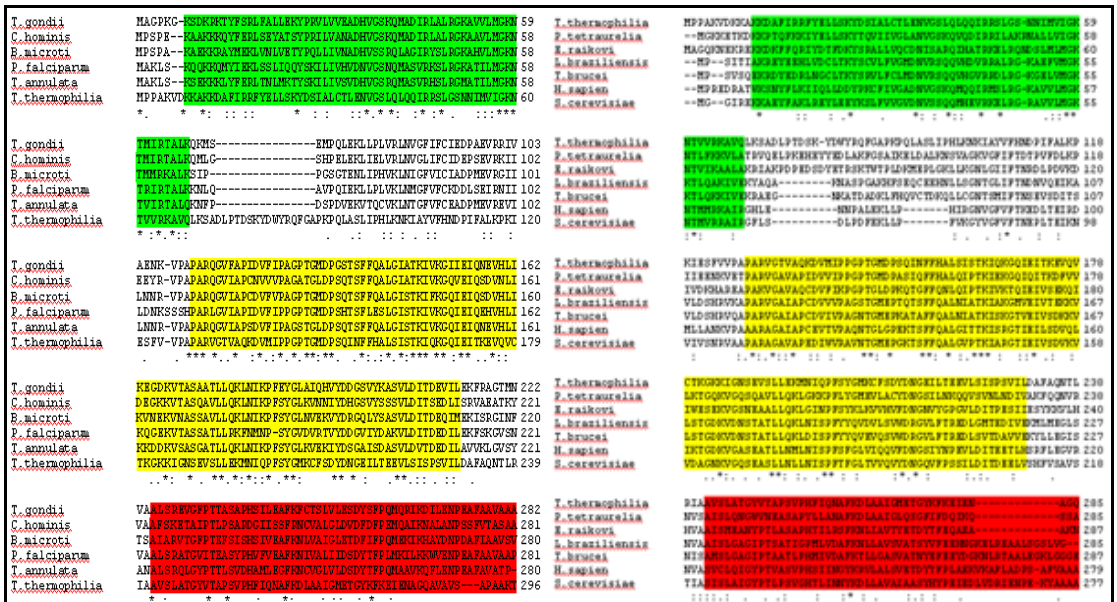

Figure 2: (a) Alignment of apicomplexan phosphoprotein p0 sequences with Tetrahymena thermophila on the left. (b) Alignment of non-apicomplexan phosphoprotein p0 sequences with Tetrahymena thermophila on the right. (c)Green highlighted section is a motif characteristic of ribosomal L10 family. The red highlighter is characteristic of 60S Ribosomal proteins. (d)Gene index numbers: T. gondii (GI:14579678), C. hominis (GI:54659401), B. microti (GI:124271139), P. falciparum (GI:4191737), T. annulata (GI:65305428), P. tetraurelia (GI:124416010), E. raikovi

(GI:62902643), L. braziliensis (GI:134063228),T. Brucei (GI:70833676), H. sapien (GI:16933546), S. cerevisiae (GI:133046), and T. thermophila (GI:62902647)

An unedited neighbor-joining tree was created with the ClustalW program located at the Kyoto University Bioinformatics Center (GenomeNet, 1991) in order to see how the Ttp0 branched off evolutionarily. The created tree showed that Ttp0 was most closely related to that of ciliophora. The ciliophora were then branched of the apicomplexans and eukaryotes. The kinetoplastids were then branched off from the very begining. See the below figure:

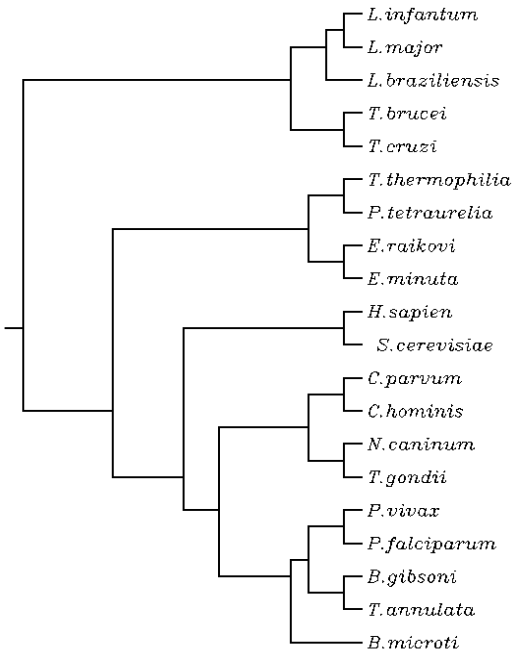

Figure 3: Neighbor-joining tree for the phosphoprotein p0 sequences

Discussion:

Phosphoprotein p0 has recently been considered as a vaccine candidate for many apicomplexan and other parasitic protozoan caused diseases. Because *Tetrahymena thermophila* is a member of the phylum Ciliophora, which branches off evolutionarily from the same point as apicomplexans, we searched for orthologues of phosphoprotein p0 in *Tetrahymena*. By using protein BLAST functions, the predicted amino acid sequence for the gene TTHERM_00636970 shown to have a high degree of homology. In an unrelated previous studies, TTHERM_00636970 was identified as a 60S Ribosomal protein (Pucciarelli et al, 2005) and is known to be expressed 250x background continuously throughout the life cycle of *Tetrahymena (*Miaw, 2009*).*

When the phosphoprotein p0 sequences were aligned, several areas of high conservation were identified. Two of these areas were then discovered to contain protein motifs. One motif being characteristic of L10 ribsomal family proteins and the second being characteristic of 60S ribosomal proteins. The third area of high conservation is still be analyzed for motifs though none have been found yet. Discovery of any motifs here may indicate that sections purpose in the functionality of the protein. That third area of high conservation is also the area of phosphoprotein p0 that contains the immunogen that commercially made antibodies against phosphoprotein p0 target. With such a high degree of homology in this region, it may be possible for antibodies targeting one organism to interact with phosphoprotein p0 in another. How well antibodies overlap may depend on how well related the phosphoprotein p0 sequences are related. How well different phosphoprotein p0 sequences are related can be determined based on the neighbor-joining tree above.

The neighbor-joining tree results came out as expected. The main branches on the tree were separated based on phylum. The main phylums included in the tree were apicomplexans, kinetoplastids and ciliophora. The kinetoplastids, containing *Leishmania spp.* and *Trypanosoma spp.* branched off early on. Tetrahymena and the other ciliosphora organisms showed the greatest similarity to the apicomplexans. This was as expected because ciliophora and apicomplexans branch off from the Alveolates.

Future Research:

In the future, we hope to perform immunocytochemistry with commercially made antibodies to discover if phosphoprotein p0 is present on the cell surface of Tetrahymena like it is on other organisms and also to determine how highly expressed it is there. To date, it is currently unknown how phosphoprotein p0 translocates to the cell surface. Because *Tetrahymena* is an excellent model in how eukaryotic cells function, we plan to use it to identify this translocation pathway if we find that phosphoprotein p0 is on the cell surface. By doing so, it may be possible to discover a method to inhibit this translocation pathway in Tetrahymena and as well in other organisms. If this pathway was inhibited, phosphoprotein would be unable to move to the cell surface and therefore would not be able to assist organisms in the process of cellular invasion.

References:

[1] Albuquerque, C., Smolka, M., Payne, S., Bafna, V., Eng, J., & Zhou , H. (2008). A multidimensional chromatography technology for in-depth phosphoproteome analysis. *Mol. Cell Proteomics, 7*(7), 1389-1396.

[2] Aury,J.-M., Jaillon,O., Duret,L., Noel,B., Jubin,C., Porcel,B.M., Segurens,B., Daubin,V., Anthouard,V., Aiach,N., Arnaiz,O., Billaut,A., Beisson,J., Blanc,I., Bouhouche,K., Camara,F., Duharcourt,S., Guigo,R., Gogendeau,D., Katinka,M., Keller,A.-M., Kissmehl,R., Klotz,C., Koll,F., Le Mouel,A., Lepere,C., Malinsky,S., Nowacki,M., Nowak,J.K., Plattner,H.,

Poulain,J., et al. (2006). Global trends of whole-genome duplications revealed by the genome sequence of the ciliate Paramecium tetraurelia. *Nature*.

[3] Chatterjee, S., Singh, S., Sohoni, R., & et al. (2000). Characterization of domains of the phosphoriboprotein P0 of Plasmodium falciparum. *Mol Biochem Parasitol,* (107), 143-154.

[4] Eisen, J. A., Coyne, R. S., Wu, M., Wu, D., Thiagarajan, M., & et al. (2006). Macronuclear Genome Sequence of the Ciliate Tetrahymena thermophila, a Model Eukaryote. *PLoS Biol, 4*(9), e286.

[5] El-Sayed,N.M.A., Myler,P.J., Blandin,G., Berriman,M., Crabtree,J.,Aggarwal,G., Caler,E., Renauld,H., Worthey,E.A., Hertz-Fowl, Ghedin,E., Peacock,C., Bartholomeu,D.C., Haas,B.J., Tran,A.-N., Wortman,J.R., Alsmark,U.C.M., Angiuoli,S., Anupama,A., Badger,J., Bringaud,F., Cadag,E., Carlton,J.M., Cerqueira,G.C., Creasy,T., Delcher,A.L., Djikeng,A., Embley,T.M., Hauser,C., Ivens,A.C., Kummerfeld,S.K., Pereira-Leal,J.B., Nilsson,D., Peterson,J., Salzberg,S.L., Shallom,J., Silva,J.C., Sundaram,J., et al. (2005). Comparative genomics of trypanosomatid parasitic protozoa. *Science, 309*(5733), 404-409.

[6] *Entrez Protein.* (n.d.). Retrieved April 23, 2009, from National Center of Biotechnology Information Entrez Protein Web site: http://www.ncbi.nlm.nih.gov/sites/entrez?db=Protein&itool=toolbar

[7] *GenomeNet.* (1991, September). Retrieved April 22, 2009, from Kyoto University Bioinformatics Center GenomeNet Web site: http://www.genome.jp/

[8} Goswami, A., Chatterjee, S., & Sharma, S. (1996). Cloning of a ribosomal phosphoprotein P0 gene homologue from Plasmodium falciparum. *Mol. Biochem. Parasitol, 82*(1), 117-120.

[9] Larkin, M., Blackshields, G., Brown, N., Chenna, P., McWilliam, H., Valentin, F., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*(21), 2947-2948.

[10] Marchler-Bauer A., J. Anderson, M. Derbyshire, C. Deweese-Scott, N. Gonzales, M. Gwadz, L. Hao, S. He, D. Hurwitz, J. Jackson, Z. Ke, D. Krylov, C. Lanczycki, C. Liebert, C. Liu, F. Lu, S. Lu, G. Marchler, M. Mullokandov, J. Song, N. Thanki, R. Yamashita, J. Yin, D. Zhang, & S. Bryant. (2007), *"CDD: a conserved domain database for interactive domain family analysis." Nucleic Acids Res.* 35(D) p.237-240.

[11] Markmann-Mulisch, U., Reiss, B., & Mulisch, M. (1999). Cell type-specific gene expression in the cell cycle of the dimorphic ciliate Eufolliculina uhligi. *Mol. Gen. Genet, 262*(2), 390-399.

[12] Miaw, W., Xiong, J., Bowen, J., Wang, W., Liu, Y., Braguinets, O., et al. (2009). Microarray analyses of gene expression during the Tetrahymena thermophila life cycle. *PLoS ONE, 4*(2), e4429.

[13] Pain,A., Renauld,H., Berriman,M., Murphy,L., Yeats,C.A., Weir,W., Kerhornou,A., Aslett,M., Bishop,R., Bouchier,C., Cochet,M., Coulson,R.M., Cronin,A., de Villiers,E.P., Fraser,A., Fosker,N., Gardner,M., Goble,A., Griffiths-Jones,S., Harris,D.E., Katzer,F., Larke,N., Lord,A., Maser,P., McKellar,S., Mooney,P., Morton,F., Nene,V., O'Neil,S., Price,C., Quail,M.A., Rabbinowitsch,E., et al. (2005). Genome of the host-cell transforming parasite Theileria annulata compared with T. parva. *Science, 309*(5731), 131-133.

[14] Peacock, C., Seeger, K., Harris, D., Murphy, L., Ruiz, J., Quail, M., et al. (2007). Comparative genomic analysis of three Leishmania species that cause diverse human disease. *Nat. Genet, 39*(7), 839-847.

[15] Pucciarelli, S., Marziale, F., Di Giuseppe, G., & Miceli, C. (2005). Ribosomal cold-adaptation: Characterization of the genes encoding the acidic ribosomal P0 and P2 proteins from the Antarctic ciliate Euplotes focardii. *Gene 360,* (2), 103-110.

[16] Rajeshwari, K., K. Patel, S. Nambeesan, M. Mehta, A. Sehgal, T. Chakraborty, & S. Sharma. (2004), "The P Domain of the P0 protein of *Plasmodium falciparum* Protects against Challenge with Malaria Parasites". Infection and Immunity Vol. 72 No. 9 p.5515-21.

[17] Rich, B., & Steitz, J. (1987). Human acidic ribosomal phosphoproteins P0, P1, and P2: analysis of cDNA clones, in vitro synthesis, and assembly. *Mol. Cell. Biol, 7*(11), 4065-4074.

[18] *RPLP0 antibody.* (1998). Retrieved April 23, 2009, from GenWay Biotech, Inc GenWay Web site: http://www.genwaybio.com/product_info.php?products_id=44160

[19] Sehgal, A., Kumar, N., Carruthers, V., & Sharma, S. (2003). Translocation of ribosomal protein P0 onto the Toxoplasma gondii tachyzoite surface. *Int. J. Parasitol, 33*(14), 1589-1594.

[20] Singh, S., Sehgal, A., Waghmare, S., Chakraborty, T., Goswami, A., & Sharma, S. (2002). Surface expression of the conserved ribosomal protein P0 on parasite and other cells. *Molecular & Biochemical Parasitology,* (119), 121-124.

[21] Skeiky,Y.A., Benson,D.R., Parsons,M., Elkon,K.B.Reed,S.G. (1992). Cloning and expression of Trypanosoma cruzi ribosomal protein P0 and epitope analysis of anti-P0 autoantibodies in Chagas' disease patients. *J. Exp. Med, 176*(1), 201-211.

[22] Stover, N., Krieger, C., Binkley, G., Dong, Q., Fisk, D., Nash, R., et al. (2006). Tetrahymena Genome Database (TGD): a new genomic resource for Tetrahymena thermophila research. *Oxford Journals, 34*, 500-503.

[23] Terkawi, M., Jia, H., Zhou, J., Lee, E., Igarashi, I., & Fujisaki, K. (2007). Babesia gibsoni ribosomal phosphoprotein P0 induces cross-protective immunity against B. microti infection in mice. *Vaccine, 25*(11), 2027-2035.

[24] Zhang, H., Lee, E., Liao, M., Compaore, M., Zhang, G., Kawase, O., et al. (2007). Identification of ribosomal phosphoprotein P0 of Neospora caninum as a potential common vaccine candidate for the control of both neosporosis and toxoplasmosis. *Mol. Biochem. Parasitol, 153*(2), 141-148.