Final Version

Trichomonas vaginalis surface proteins: a view from the genome

Hirt R. P.[1]*, Noel C. J.[1], Sicheritz-Ponten, T.[2], Tachezy J.[3] and Fiori P.-L.[4]

[1]Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK.

[2]Center for Biological Sequence Analysis, BioCentrum-DTU, Technical University of Denmark, DK-2800 Lyngby, DK.

[3]Department of Parasitology, Charles University, Vinicna 7, 128 44 Prague 2, Czech Republic

[4]Department of Biomedical Sciences, Division of Experimental and Clinical Microbiology, University of Sassari, Italy, 07100 Sassari, Italy.

*Corresponding author:

Dr Robert P. Hirt

Institute for Cell and Molecular Biosciences

Cathrine Coockson Building

Framlington Place

Newcastle University

Newcastle upon Tyne, NE2 4HH, UK

Tel: +44-191-222-6963

Fax: +44-1910246-7424

E-mail: R.P.Hirt@ncl.ac.uk

Abstract

Surface proteins of mucosal microbial pathogens play multiple and essential roles in initiating and sustaining the colonization of the heavily defended mucosa. The protist *Trichomonas vaginalis* is one of the most common human sexually transmitted pathogens that colonize the urogenital mucosa. However, little is known about its surface proteins. The recently completed draft genome sequence of *T. vaginalis* provides an invaluable resource to guide molecular and cellular characterization of surface proteins and to investigate their role in pathogenicity. Here we review the existing data on *T. vaginalis* surface proteins and summarize some of the main findings from the recent *in silico* characterization of its candidate surface proteins.

Surface proteins of mucosal pathogens

In order to initiate the invasion of host mucosa, and later thrive within them, mucosal microbial pathogens must be able to bind specifically to the host tissue and bypass the initial innate defence systems[1,2]. Pathogens must also have the ability to subvert or oppose attacks from the proteins and cells of the innate and later adaptive immune responses[3,4]. *Trichomonas vaginalis* is a sexually transmitted, non-self-limiting pathogen in females, infecting the urogenital tract of both sexes. It has been recognized as an important pathogen due to its high prevalence and its link with other disease, in particular inducing higher susceptibility to HIV (see for example, Ref[5,6]). Infection is also correlated with higher prevalence of invasive cervical cancer (see for example, Ref [7]) and more recently an association with prostate cancer was also suggested[8]. In addition to the morbidity caused to carriers, *Trichomonas* infections during pregnancy also to lead to increased risk for preterm and low weight babies[7,9], which in poor sanitary and nutritional contexts in particular, is likely to translate into higher rates of babies early death. Furthermore, only ~50% of *T. vaginalis* infections lead to symptomatic inflammatory responses and together with the low sensitivity of the

commonly used diagnostic method makes it an elusive pathogen (see for example, Refs[6,10,11]). This causes underestimation of infection rates and also complicates its clinical follow up[6]. For reviews on various biological and medical aspects of *T. vaginalis* see Refs.[5,7,12].

T. vaginalis successful colonization of the host mucosa is thought to depend on multiple mechanisms including: (i) binding to, and degradation of components from, the mucus and extracellular matrix (ECM) proteins (reviewed in[12]), (ii) binding to host cells including vaginal epithelial cells (VEC) and immune cells such as neutrophils, leading to cytotoxic effects[13,14], (iii) phagocytosis of vaginal bacteria and host cells (VEC, erythrocytes and immune cells)([15]and reference herein), (iv) endocytosis of host proteins and (v) degradation of IgG and IgA antibodies and complement proteins (reviewed in Ref.[16]). In addition, *T. vaginalis* can internalize viable viruses such as HIV via endocytosis[17] and bacteria such as *Mycoplasma*, with *Mycoplasma* being able to survive and multiply intracellularly[18]. *T. vaginalis* may represent a Trojan horse, or vector, for these pathogenic agents, as previously suggested for the human papillomavirus (reviewed in Ref.[7]). *T. vaginalis* surface proteins are thought to be essential for these different activities and have been shown in other organisms to be important virulence factors central to the host-pathogen interface (see for example, Refs[19,20]). Pathogen surface proteins are also involved in host immune recognition and stimulation (innate and adaptive responses) and *T. vaginalis* surface molecules were recently shown to induce dendritic cell (DC) maturation regulating the development of the immune response[21].

The principle published data on *T. vaginalis* surface proteins, accumulated during the past three decades, suggest that several gene families encoding classic enzymes known to localize and function in hydrogenosomes, a mitochondrial homologue mediating anaerobic metabolism[22], can also be expressed on the cell surface and are adhesins - so called moonlighting proteins[23] having two or more unrelated functions[24]. These data are controversial in that, for example, these 'adhesins' include the well established hydrogenosomal marker malate dehydrogenase decarboxylating enzyme

(malic enzyme, ME)[25]. More recently the genome sequencing project identified a surprisingly large genome for *T. vaginalis* (strain G3, ~160 Mbp) encoding the staggering amount of ~60,000 hypothetical proteins[26]. Over 300 candidate surface proteins were identified with at least one inferred transmembrane (TM) domain and sharing one or more features with surface proteins from other pathogens known to contribute to mucosal colonization or other pathogenic processes. Here we briefly reconsider the published *T. vaginalis* adhesins data and other described surface proteins and review the more recent genome data.

T. vaginalis surface proteins: adhesins, cystein proteinases and P270 proteins

A collection of *T. vaginalis* proteins has been described in the literature with different pieces of, often indirect, data supporting the notion that they are expressed on the cell surface and play important roles in regulating *T. vaginalis*-host interactions. Several hydrogenosomal enzymes (including ME, α and β subunits of succinyl CoA synthetase [SCS] and pyruvate:ferredoxin oxidoreductase [PFO]), have been claimed to have dual cellular localizations with an alternative localization on the cell surface where they could play an adhesin function (adhesins designated AP65, AP33, AP51 and AP120, respectively – Table 1, Table S1)[24,27]. However these data are controversial because other papers have demonstrated a uniquely hydrogenosomal localization for these enzymes by means of immunocytochemistry and/or cell fractionation[28-30] – for further pros and cons on these data see **Box 1**. More recently a transcription initiation factor-like protein was also claimed to be expressed on the cell surface[31]. All these proposed surface proteins lack detectable sequence features known to target and anchor eukaryotic proteins to the cell surface and have homologues with well-established functions in intracellular compartments in other systems. Assuming that these proteins can be expressed on the cell surface, it will be important to establish the cellular trafficking and membrane anchorage of these candidate surface proteins by performing

detailed molecular cell studies. Assuming some of these also play adhesin roles, it will be very interesting to investigate in more detail the function of these proteins by identifying their host binding partners.

In contrast to these debated data, a family of immunoreactive surface proteins, called P270, has been shown to be expressed on the surface of *T. vaginalis* cells and to possess the structural organization of classic eukaryotic transmembrane proteins (reviewed in Ref.[12]). That is, they have a TM domain that can be identified using well-established software such as TMHMM2.0. The function of these proteins is currently unknown but their structure and predicted topology (presence of repeats facing the outside of the cell) is consistent with some form of binding activities on the cell surface.

As an alternative to TM domains several important surface proteins of significant parasitic protists (including *Plasmodium*, *Trypanosoma*, *Leishmania* and *Entamoeba*) are known to possess glycosylphosphatidylinositol (GPI)-anchors[32,33]. However, none of the genes encoding enzymes known to synthesise GPI-anchors and to mediate their anchoring to proteins was identified in the genome of *T. vaginalis*[26]. *T. vaginalis* is the first eukaryote that does not seem to generate GPI-anchored surface proteins. This further highlights the importance of identifying proteins with TM domains when hunting for *T. vaginalis* candidate surface proteins.

T. vaginalis genome encodes numerous candidate surface proteins

The *T. vaginalis* genome data provide a unique opportunity to investigate the features of the proteome *in toto* and by doing so identify candidate surface proteins with structural features consistent with cell surface localisation and functions[26]. Indeed, over 5100 *T. vaginalis* proteins were inferred to possess one or more TM domain(s) using TMHMM2.0. Of these a total of 3347 proteins have a single hypothetical TM domain that could anchor these to a membrane. In eukaryotic cells such proteins face two cellular compartments and when linked to membrane of the

secretory/endocytic pathways, including the plasma membrane, have a cytosolic and an 'extracellular domain' with the extracellular domain facing the outside of the cell or the lumen of an organelle (topologically equivalent environments). If expressed on the cell surface, the extracellular domain can interact with external components mediating binding functions (e.g. adhesins binding to host molecules) or perform exo-enzymatic activities (e.g. degrading host proteins or polysaccharides). Some surface proteins have multiple functions and combine binding to one or more partners and enzymatic activities such as the GP63 surface proteases from *Leishmania* (reviewed in Ref.[34]). Initial analyses focused on those proteins families that encode proteins with one TM domain and possess sequence features, including one or more well-established protein domains, shared with surface proteins from eukaryotic or prokaryotic pathogens. In addition, proteins with protease domains and at least one TM domain were also investigated. This lead to the identification of 10 different protein families comprising in excess 300 proteins (Table 2) that can be considered as representing strong candidate *T. vaginalis* surface proteins[26].

BspA-like proteins (TpLRR-containing proteins)

The largest gene family encoding potential surface proteins shares a specific type of leucine-rich repeat (LRR), the TpLRR[35] (Table 2, Figure 1). These were named BspA-like proteins due to high sequence similarity between the TpLRR of the BspA surface protein from *Tannarella forsythensis* and the first hypothetical protein identified from *T. vaginalis*[36] (Figure 1). TpLRR containing proteins in different mucosal bacteria were shown to mediate binding to host epithelial cells and/or ECM proteins and were also implicated in bacteria co-aggregation[37]. The taxonomic distribution of TpLRR-containing proteins is broad, including all three domains of life (Eukaryota, Archaea and Bacteria), but very patchy with currently only two eukaryotic genuses (*Trichomonas* and *Entamoeba*), two Archaea and various Bacteria (~16 bacterial species and strains), all known to

live on mucosal surfaces or body cavities. This strongly suggests that these proteins are specific to microorganisms that thrive in such environments and that the genes encoding these were subjected to lateral gene transfers between distantly related organisms occupying the same niche[36]. The only other eukaryotes currently known to encode BspA-like proteins are *Entamoeba histolytica*[38] and *E. dispar*[39], with *E. histolytica* being another important parasitic protist attacking human mucosa[33]. Like *T. vaginalis*, the *E. histolytica* genome has a large gene family with over 90 entries[40] – in contrast to prokayotes which have only one or few genes. One of the *E. histolytica* protein was demonstrated to be expressed on the cell surface, consistent with binding function(s) mediated at the cell surface[40] as demonstrated for two bacterial BspA-like proteins[37]. *E. histolytica* BpsA-like proteins do not possess identifiable TM domains and seem to be anchored to the plasma membrane with a unusual type of surface lipid anchor that could resembles the one added to the CaaX motif of membrane proteins facing the cytosol of other eukaryotes[40]. None of the identified *T. vaginalis* BspA-like proteins possess such a CaaX motif nor do the hydrogenosomal 'adhesins' discussed above.

Among the 656 BspA-like hypothetical proteins identified in *T. vaginalis*, 178 possess an identifiable TM domain with an inferred topology exposing the TpLRR domain extracellularly, consistent with a binding function on the cell surface. One member of the *T. vaginalis* BspA-like protein with a TM was shown to be expressed on the cell surface (TvBspA-625[36]) (C. J. Noel *et al.* unpublished) (Fig. 1c). There is a marked diversity in the length and other sequence features between members of this large protein family, including variation in the number of TpLRR (from three to over 30 repeats). The proteins with TM often possess other repeats, typically proline-rich repeat (PRR) (Figure 1)[36]. The large number of BspA-like genes invites us to suggest that these proteins represent an array of diverse surface proteins with multiple binding properties. This diversity could also contribute to host immune evasion if differentially expressed during infection - as is well

established for several gene families in other pathogens[19]. Interestingly, TpLRRs of prokaryotic proteins were also suggested to represent pattern recognition molecule for the innate immune response, influencing the evolution of microbial infection[41]. Taken together these different pieces of information strongly suggest that BspA-like proteins are expressed on the cell surface and play important roles in *T. vaginalis* interactions with its host.

Surface proteases

Proteases from parasites are thought to play numerous and important role in host-pathogen interactions[16]. *T. vaginalis* encodes an impressive repertoire of candidate proteases with over 400 entries from multiple Clans and families[26], greatly exceeding the degradome of currently known microbial eukaryotic parasites such as *Encephalitozoon cuniculi* (41), *E. histolytica* (51), *Plasmodium falciparum* (92) or *Leishmania major* (153)[26,42]. Among these, 122 *T. vaginalis* entries possess one or more TM domains, so called transmembrane proteases (TP)[43]. Such proteases are better known in the human system where they fulfil multiple functions including degrading ECM proteins, cell-cell and cell-ECM adhesion, and are thought to be important in neoplastic, inflammatory and infection sites[43]. *T. vaginalis* is known to invade the human mucosal tissue, where it induces inflammation and degrades host proteins including those from the mucus and ECM. *T. vaginalis* TP are likely to play important roles in these processes[16]. A selection of these TP is discussed here (Table 2 and Figure 2).

The second largest gene family of candidate surface proteins encodes GP63-like proteins (77 paralogues in total) of which 53 possess potential TM domains (Table 2, Figure 2). The GP63 proteins (or leishmanolysins, or major surface protein - MSP) were characterized as the most abundant surface proteins (GPI-anchored) in *Leishmania* sp. promastigotes (life cycle stage in the gut of the insect host) and these are also expressed in the amastigote stage (life cycle stage in the

mammal host) playing important roles in the parasites survival in both hosts[34]. GP63 are metalloproteinase belonging to the metzincin class (EC 3.4.24.36) characterized by the motif HExxHxxGxxH (x represents any amino acid residues) forming an extended zinc-binding motif and the catalytic site. The majority of the 53 *T. vaginalis* GP63-like sequences (40 entries) with a TM domain possess a similar but shorter motif (missing the final three residues xxH) that contains the minimal motifs HExxH for Zincins[44] (Figure 2). GP63 in *Leishmania* are involved in binding to host cells, degradation of various host proteins, including proteins from the immune system (complement) and ECM proteins, and survival in phagolysosomes[34]. The genome of *Leishmania major* encodes seven GP63 paralogues, whereas those of *Trypanosoma cruzii* and *T. brucei* encode 13 and >100 respectively[45]. Phylogenetic analysis of the *T. vaginalis* GP63-like proteins with homologues from several eukaryotes recovered the *T. vaginalis* sequences as monophyletic, indicating that the *T. vaginalis* gene family was generated by relatively recent gene duplications, further suggesting that these are playing an important role in *T. vaginalis* biology. These GP63-like paralogues probably derived from 'adaptive gene amplifications' (e.g. to accommodate a broad substrate diversity and provide immunovariants) with the gene family size not correlating with the organism complexity as measured by cellular diversity[46]. Taken together, the sequence features and the diversity of *T. vaginalis* GP63-like proteins and the functional data from other parasitic protists suggest that these proteins are likely to play important roles in *T. vaginalis* pathogenicity by degrading and binding to various host components.

Additional notable candidate TP families are 28 subtilisin-like serine proteases, nine other serine proteases and five calpain-like cysteine proteases (Figure 2). Together with the GP63-like proteins they represent an impressive array of candidate TP that could degrade a broad range of host proteins or provide alternative immunovariant proteins if differentially expressed during infection. The calpain-like cysteine proteases possess 22-23 identifiable TM domains. This structural

organisation suggest that these TP could function as surface proteases and transporters importing protein fragments or small peptides, if not single amino acids, providing for *T. vaginalis* a source of amino acids, which is thought to rely heavily on some amino acid for energy generation and redox balancing[26,47].

In addition to these *in silico* inferred TP, several publications have characterized secreted and surface cysteine proteases[16,48], none of these seems to possess an identifiable TM domain or to be identifiable among the TP discussed above. Nonetheless, based on existing data, these are appealing candidate virulence factors potentially important in binding host tissue and in cytotoxicity[48].

Candidate surface proteins sharing domains with other microbial pathogens

Other *T. vaginalis* hypothetical surface protein families share domains with the surface proteins of other mucosal pathogens and these include domains from: *Chlamydia* polymorphic membrane proteins[49], *Giardia* variant surface proteins (VSP)[50], *Entamoeba* immunodominant variable surface antigen M17[51] and candidate lectins (Table 2 and Figure 3). In all cases, several family members were identified suggesting that: (i) the alternative variant could be expressed during different stages of the infection (upon contact to mucus, VEC surface, or ECM proteins); (ii) they play different roles in binding diverse target molecules from the host or other microbes from the mucosal surfaces; or (iii) they are differentially expressed during infection contributing to escaping the host adaptive immune responses.

Among the two immunodominant variable surface antigen M17 encoded by the *E. histolytica* genome, one was recently shown to be located in phagosomes[51]. It was suggested that this surface protein could represent an amoebic receptor for host cells inducing their phagocytosis by *E. histolytica*[51]. The *T. vaginalis* proteins have a similar structural organisation to the *E. histolytica* one and since the former is also a phagocytic cell[15] this protein family could be involved in this process. The 11 genes encoding potential surface lectins (Figure 3) could be involved in important adhesin

function via binding to sugar moieties of host glycoconjugates from the mucus, glycocalix from the VEC or from or host immune cells, and from the ECM[1,2]. Lectins are considered as important virulence factors in other pathogens, e.g. in *E. histolytica*[33]. Notably the structural organization of the *Entamoeba* lectins is clearly distinct from *Trichomonas* proteins sharing no sequence similarity. *T. vaginalis* lectins could also be involved in binding and endocytosing HIV particles[17]. HIV are known to be internalized by human mucosal DC lectins and can later be released from these cells once they migrated to lymph nodes where the virus can then infect CD4+ target cells[52]. Similarly, following transfer from one person to another *T. vaginalis* (from a dually infected *T. vaginalis*-HIV individual) having internalized viruses could release HIV particles upon contact to host tissue in the newly infected mucosa[53], and thus facilitate the spread of the virus.

Several of these *T. vaginalis* candidate surface proteins also possess repeats, such as PRR, a feature shared with several of the *T. vaginalis* BspA-like and GP63-like protein families and numerous other surface proteins from prokaryotes and eukaryotes involved in protein-protein binding activities[35,54]. These structural features are consistent with these different proteins playing a role on the *T. vaginalis* cell surface and performing binding functions.

Concluding remarks

In contrast to early suggestions[55] that *T. vaginalis* needs to diversify the function of a small set of proteins (due to an allegedly small genome) through moonlighting[23] to adapt to its environment, we now know that *T. vaginalis* genome is large and encodes a massive proteome with a considerable and diverse repertoire of candidate surface proteins[26]. Future work will establish the relative contribution of the *in silico* predictions versus the existing enzymatic 'adhesins' data and other surface molecules[56] in regulating host-pathogen interactions (see **Box 2**). The identified candidate transmembrane proteins share domains and/or sequence motifs with surface proteins

known to be important for other pathogens, which together are likely to mediate several of the important functions underlying *T. vaginalis* pathogenicity. As such, they represent promising targets to study *T. vaginalis* pathogenesis, to develop protective mucosal vaccines[4], and for the development of new antimicrobial strategies based on adhesins currently thought off to circumvent emerging antibiotic resistance[2]. As some of the candidate surface proteins are also candidate proteases, specific drugs targeting their proteolytic activities could also be thought off to control infections to complement existing drugs[16,57]. Finally, due to the repetitive nature of several members of the identified gene families encoding candidate surface proteins, some of these could also represent valuable markers for genotyping *T. vaginalis* clinical strains.

Acknowledgments

We greatly acknowledge the sequencing effort of the *T. vaginalis* draft genome by Jane Carlton and colleagues at TIGR. We thank Dr Martin Embley and two anonymous reviewers for constructive comments on an earlier version of the manuscript.

References

1 Nataro, J.P. et al., eds (2005) *Colonization of mucosal surfaces*, ASM Press

2 Preissner, K.T. and Chhatwal, G.S. (2004) Extracellular matrix and host cell surfaces: potential sites for pathogen interaction. In *Cellular microbiology* (Cossart, P. et al., eds.), pp. 87-104, ASM press

3 Cole, A.M. (2006) Innate host defense of human vaginal and cervical mucosae. *Curr Top Microbiol Immunol* 306, 199-230

4 Neutra, M.R. and Kozlowski, P.A. (2006) Mucosal vaccines: the promise and the challenge. *Nat Rev Immunol* 6, 148-158

5 Schwebke, J.R. and Burgess, D. (2004) Trichomoniasis. *Clin Microbiol Rev* 17, 794-803

6 Soper, D. (2004) Trichomoniasis: under control or undercontrolled? *Am J Obstet Gynecol* 190, 281-290

7 Nanda, N. et al. (2006) Trichomoniasis and its treatment. *Expert Rev Anti Infect Ther* 4, 125-135

8 Sutcliffe, S. et al. (2006) Plasma antibodies against Trichomonas vaginalis and subsequent risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 15, 939-945

9 Cotch, M.F. et al. (1997) Trichomonas vaginalis associated with low birth weight and preterm delivery. The Vaginal Infections and Prematurity Study Group. *Sex Transm Dis* 24, 353-360

10 Huppert, J.S. et al. (2007) Rapid antigen testing compares favorably with transcription-mediated amplification assay for the detection of Trichomonas vaginalis in young women. *Clin Infect Dis* 45, 194-198

11 Sena, A.C. et al. (2007) Trichomonas vaginalis infection in male sexual partners: implications for diagnosis, treatment, and prevention. *Clin Infect Dis* 44, 13-22

12 Lehker, M.W. and Alderete, J.F. (2000) Biology of trichomonosis. *Curr. Opin. Infect. Dis.* 13, 37-45

13 Fiori, P.L. et al. (1997) Contact-dependent disruption of the host cell membrane skeleton induced by Trichomonas vaginalis. *Infect Immun* 65, 5142-5148

14 da Costa, R.F. et al. (2005) Trichomonas vaginalis perturbs the junctional complex in epithelial cells. *Cell Res* 15, 704-716

15 Pereira-Neves, A. and Benchimol, M. (2006) Phagocytosis by Trichomonas vaginalis - New Insights. *Biol Cell*

16 Klemba, M. and Goldberg, D.E. (2002) Biological roles of proteases in parasitic protozoa. *Annu Rev Biochem* 71, 275-305

17 Rendon-Maldonado, J. et al. (2003) Trichomonas vaginalis: in vitro attachment and internalization of HIV-1 and HIV-1-infected lymphocytes. *J Eukaryot Microbiol* 50, 43-48

18 Dessi, D. et al. (2005) Long-term survival and intracellular replication of Mycoplasma hominis in Trichomonas vaginalis cells: potential role of the protozoon in transmitting bacterial infection. *Infect Immun* 73, 1180-1186

19 Casadevall, A. and Pirofski, L. (2001) Host-pathogen interactions: the attributes of virulence. *J Infect Dis* 184, 337-344

20 Smith, H. (1977) Microbial surfaces in relation to pathogenicity. *Bacteriol Rev* 41, 475-500

21 Scott, K. et al. (2005) Qualitatively distinct patterns of cytokines are released by human dendritic cells in response to different pathogens. *Immunology* 116, 245-254

22 Embley, T.M. (2006) Multiple secondary origins of the anaerobic lifestyle in eukaryotes. *Philos Trans R Soc Lond B Biol Sci* 361, 1055-1067

23 Jeffery, C.J. (2004) Molecular mechanisms for multitasking: recent crystal structures of moonlighting proteins. *Curr Opin Struct Biol* 14, 663-668

24 Alderete, J.F. et al. (2001) Enzymes on microbial pathogens and *Trichomonas vaginalis*: molecular mimicry and functional diversity. *Cell Microbiol* 3, 359-370

25 Hrdy, I. et al. (2004) Trichomonas hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. *Nature* 432, 618-622

26 Carlton, J.M. et al. (2007) Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. *Science* 315, 207-212

27 Moreno-Brito, V. et al. (2005) A Trichomonas vaginalis 120 kDa protein with identity to hydrogenosome pyruvate:ferredoxin oxidoreductase is a surface adhesin induced by iron. *Cell Microbiol* 7, 245-258

28 Drmota, T. et al. (1996) Iron-ascorbate cleavable malic enzyme from hydrogenosomes of Trichomonas vaginalis: purification and characterization. *Mol Biochem Parasitol* 83, 221-234

29 Brugerolle, G. et al. (2000) Immunolocalization of two hydrogenosomal enzymes of Trichomonas vaginalis. *Parasitol Res* 86, 30-35

30 Hrdy, I. and Muller, M. (1995) Primary structure and eubacterial relationships of the pyruvate:ferredoxin oxidoreductase of the amitochondriate eukaryote Trichomonas vaginalis. *J Mol Evol* 41, 388-396

31 Mundodi, V. et al. (2006) A novel surface protein of Trichomonas vaginalis is regulated independently by low iron and contact with vaginal epithelial cells. *BMC Microbiol* 6, 6

32 Ropert, C. and Gazzinelli, R.T. (2000) Signaling of immune system cells by glycosylphosphatidylinositol (GPI) anchor and related structures derived from parasitic protozoa. *Curr Opin Microbiol* 3, 395-403

33 Petri, W.A., Jr. et al. (2002) The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite Entamoeba histolytica. *Annu Rev Microbiol* 56, 39-64

34 Yao, C. et al. (2003) The major surface protease (MSP or GP63) of *Leishmania* sp. biosynthesis, regulation of expression, and function. *Mol Biochem Parasitol* 132, 1-16

35 Kobe, B. and Kajava, A.V. (2001) The leucine-rich repeat as a protein recognition motif. *Curr. Op. Str. Biol.* 11, 725-732

36 Hirt, R.P. et al. (2002) A novel potential surface protein in *Trichomonas vaginalis* contains a leucine-rich repeat shared by micro-organisms from all three domains of life. *Mol Biochem Parasitol* 125, 195-199

37 Ikegami, A. et al. (2004) Multiple functions of the leucine-rich repeat protein LrrA of *Treponema denticola*. *Infect Immun* 72, 4619-4627

38 Loftus, B. et al. (2005) The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433, 865-868

39 Shire, A.M. and Ackers, J.P. (2007) SINE elements of Entamoeba dispar. *Mol Biochem Parasitol* 152, 47-52

40 Davis, P.H. et al. (2006) Identification of a family of BspA like surface proteins of *Entamoeba histolytica* with novel leucine rich repeats. *Mol Biochem Parasitol* 145, 111-116

41 Hajishengallis, G. et al. (2002) Dependence of bacterial protein adhesins on toll-like receptors for proinflammatory cytokine induction. *Clin Diagn Lab Immunol* 9, 403-411

42 Rawlings, N.D. et al. (2006) MEROPS: the peptidase database. *Nucleic Acids Res* 34, D270-272

43 Bauvois, B. (2004) Transmembrane proteases in cell growth and invasion: new contributors to angiogenesis? *Oncogene* 23, 317-329

44 Gomis-Rüth, F.X. (2003) Structural aspects of the metzincin clan of metalloendopeptidases. *Mol Biotechnol* 24, 157-202

45 Ivens, A.C. et al. (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436-442

46 Vogel, C. and Chothia, C. (2006) Protein family expansions and biological complexity. *PLoS Comput Biol* 2, e48

47 Coombs, G.H. et al. (2004) The amitochondriate eukaryote Trichomonas vaginalis contains a divergent thioredoxin-linked peroxiredoxin antioxidant system. *J Biol Chem* 279, 5249-5256

48 Solano-Gonzalez, E. et al. (2006) Location of the cell-binding domain of CP65, a 65kDa cysteine proteinase involved in Trichomonas vaginalis cytotoxicity. *Int J Biochem Cell Biol* 38, 2114-2127

49 Gomes, J.P. et al. (2006) Polymorphisms in the nine polymorphic membrane proteins of Chlamydia trachomatis across all serovars: evidence for serovar Da recombination and correlation with tissue tropism. *J Bacteriol* 188, 275-286

50 Adam, R.D. (2001) Biology of Giardia lamblia. *Clin Microbiol Rev* 14, 447-475.

51 Marion, S. and Guillen, N. (2006) Genomic and proteomic approaches highlight phagocytosis of living and apoptotic human cells by the parasite Entamoeba histolytica. *Int J Parasitol* 36, 131-139

52 Marsh, M. (2005) Cell biology of virus infection. In *Cellular microbiology* (Cossart, P. et al., eds.), pp. 517-542, ASM press
53 Lal, K. et al. (2006) Dramatic reorganisation of Trichomonas endomembranes during amoebal transformation: a possible role for G-proteins. *Mol Biochem Parasitol* 148, 99-102
54 Williamson, M.P. (1994) The structure and function of proline-rich regions in proteins. *Biochem. J.* 297, 249-260
55 Garcia, A.F. et al. (2003) Iron and contact with host cells induce expression of adhesins on surface of Trichomonas vaginalis. *Mol Microbiol* 47, 1207-1224
56 Bastida-Corcuera, F.D. et al. (2005) Trichomonas vaginalis lipophosphoglycan mutants have reduced adherence and cytotoxicity to human ectocervical cells. *Eukaryot Cell* 4, 1951-1958
57 Dunn, L.A. et al. (2007) The activity of protease inhibitors against Giardia duodenalis and metronidazole-resistant Trichomonas vaginalis. *Int J Antimicrob Agents* 29, 98-102
58 Kucknoor, A.S. et al. (2005) Heterologous expression in Tritrichomonas foetus of functional Trichomonas vaginalis AP65 adhesin. *BMC Mol Biol* 6, 5
59 Mundodi, V. et al. (2004) Silencing the ap65 gene reduces adherence to vaginal epithelial cells by Trichomonas vaginalis. *Mol Microbiol* 53, 1099-1108
60 Addis, M.F. et al. (2000) Host and tissue specificity of Trichomonas vaginalis is not mediated by its known adhesion proteins. *Infect Immun* 68, 4358-4360
61 Dolezal, P. et al. (2005) Giardia mitosomes and trichomonad hydrogenosomes share a common mode of protein targeting. *Proc Natl Acad Sci U S A* 102, 10924-10929
62 Tao, X. et al. (2003) Crystal structures of substrate complexes of malic enzyme and insights into the catalytic mechanism. *Structure* 11, 1141-1150
63 Krogh, A. et al. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567-580.
64 Rouault, T.A. and Tong, W.H. (2005) Iron-sulphur cluster biogenesis and mitochondrial iron homeostasis. *Nat Rev Mol Cell Biol* 6, 345-351
65 Mueller, J.C. et al. (2004) Mechanisms for multiple intracellular localization of human mitochondrial proteins. *Mitochondrion* 3, 315-325
66 Alderete, J.F. et al. (1995) Cloning and molecular characterization of two genes encoding adhesion proteins involved in *Trichomonas vaginalis* cytoadherence. *Mol Microbiol* 17, 69-83
67 Rigden, D.J. et al. (2004) The PA14 domain, a conserved all-beta domain in bacterial toxins, enzymes, adhesins and signaling molecule. *Trends Biochem Sci* 29, 335-339

Boxes: Box 1 and Box 2

Box 1. Can *T. vaginalis* hydrogenosomal enzymes also function as adhesins?

Putative adhesins were first identified using an *in vitro* binding assay (see references cited in Ref.[24]) based on exposure of *T. vaginalis* total protein extracts (not intact cells) to human cell cultures. Candidate *T. vaginalis* adhesins bound to the human cells were then revealed using Western blot analyses of total protein human cell extracts using specific antibodies raised against *T. vaginalis* total protein extracts. Consistent with an 'adhesin' roles for these different proteins (AP65, AP33, AP51 and AP120) several publications have suggested cell surface localization by immunocytochemistry and antibodies specific to the different proteins affect *T. vaginalis* cell binding to host cell cultures (see for example, Refs[27,55]). More recently, down regulation of the expression of ME(AP65) or its heterologous expression in *Tritrichomonas foetus* (a bovine trichomonad[5]) resulted in binding phenotypes consistent with adhesin functions for this hydrogenosomal enzyme[58,59]. However several features are at odd with this adhesin model, including:

- ME(AP65) and other 'adhesins' bind non-specifically to different cells surfaces of different histological derivation (from human cells to bacteria), including protease treated cells, using the same assay originally used to identify these proteins[60]. This demonstrated that these proteins can bind to membranes non-specifically even after being protease treated. This also suggests that the assay used to identify *T. vaginalis* adhesins is not adequate
- No detailed confocal microscopy analyses is available to demonstrate convincingly the dual localization for these enzymes - hydrogenosome and cell surface
- All these enzymatic 'adhesins' possess N-terminal presequences targeting the proteins to the hydrogenosomes and these are cleaved off by specific matrix metalloproteases during the protein maturation within organelles[61](Table 1). No alternative targeting signals have been identified to so far on any of these proteins

- These proteins have no identifiable structural features suggesting classic membrane anchorage. Recent structural data for ME[62] make it difficult to structurally contextualize the two suggested TM domains of ME(AP65) identified using an unspecified bioinformatic tool[55] - and which cannot be confirmed using reliable software such as TMHMM2.0[63]
- Mechanisms such as alternative splicing, transcriptional or translational initiation, can generate alternative targeting of mitochondrial proteins to various cell compartments in other systems (see for example, Refs[64,65]). Currently, however, there is neither published evidence nor sequence features from the recent genome sequence data covering these hydrogenosomal enzymes suggesting that such mechanisms might operate in *T. vaginalis*
- To our knowledge, no human receptors have yet been identified and characterized for these enzymatic 'adhesins', although, there was a recent claim for a 130 kDa host protein at the surface of HeLa cells binding PFO(AP120)[27]

Based on these considerations, it is currently unknown how these enzymes could be expressed and linked to the cell surface under physiological conditions, and how they could mediate specific binding to host tissue *in vivo*. One possible explanation for cell surface localization of these hydrogenosomal enzymes is that during stress (e.g. nutrient shortage or exposure to high levels of iron) hydrogenosomes could be degraded via authophagy and some hydrogenosomal proteins reach the cell surface via trafficking through endocytic compartments linked with phagosomes. Indeed, hydrogenosomal markers ME and βSCS were observed in vacuoles containing hydrogenosomal material in the process of degradation[20]. Interestingly, N-terminal amino acid sequencing of ME(AP65), which binds to HeLa cells in a ligand assay[66], revealed an absence of the targeting pre-sequence, consistent with hydrogenosomal processing of the protein to its mature form. Subsequent binding of the proteins from degraded hydrogenosomes to the membrane could be explained by their overall hydrophobic nature and propensity to bind membranes non-specifically[60].

End of Box 1

Box 2: Some outstanding questions for *T. vaginalis* surface proteins

- What proportion of the *in silico* identified candidate surface gene families are translated and expressed on the cell surface?
- What are the differences in surface proteins expression patterns during the infection of the female versus male urogenital tracts?
- Which surface molecules contribute to the initiation of the innate and adaptive immune response?
- Is there any correlation between strain virulence and the expression level of the different gene families encoding surface proteins?
- Are there any correlations between variations or primary sequences (e.g. number of repeats) and protein family sizes of surface proteins and strain virulence?

End of box 2

Figure legends

Figure 1. *Trichomonas vaginalis* BspA-like proteins. (a) The structural organisation of BspA-like proteins with a TM domain (178 entries for a total of 656 entries) is shown with the TpLRR domain (green rectangle – 14 repeats are shown here). Other repeats are also often present such as proline-rich repeats (PRR) or serine and threonine-rich repeats (pink rectangle). Some BspA-like proteins have also detectable signal peptides. A majority of hypothetical BspA-like proteins have no TM. Some of these have a signal peptide - could be either secreted or anchored to the membrane in an unknown fashion or be derived from pseudogenes. (b) BlastP alignment (E-value = $2e^{-78}$, identity =

52%) between one *T. vaginalis* BspA-like protein (top sequence, TvBspA625, accession: AAM51159) and BspA from *Tannerella forsythensis* (accession: AAC82625 – note that the species former name was *Bacteroides forsythus*) showing the high level of sequence identity between their shared TpLRR. The first and last aligned residues for each protein are indicated. The TpLRR consensus sequence, made of 23 residues (with some known variation in length), is also shown (bottom right) – see Refs[35,36]. The conserved residues (identical or similar) defined in the consensus sequence are highlighted in the same way in the alignment. The + and – signs indicate similar and distinct amino acid residues, respectively. (c) Indirect immunocytochemistry on fixed and permeabilized *T. vaginalis* cells using confocal microscopy demonstrating the expression and cellular localization of one BspA-like protein (TvBspA-625[36]) using a specific mouse anti-peptide antisera (green) compared to the localisation of malic enzyme (a classic hydrogenosomal marker - using a specific rabbit antisera – red) and the nucleus (DAPI staining – blue) (C.J. Noel *et al*. unpublished).

Figure 2. **Structural organisation of candidate transmembrane protease.** In all cases the TMHMM2.0 analyses recovered the protease domains as being extracellular, they would face the outside of the cell if the protein were to be expressed on the cell surface. The accession numbers of the identified domains by InterProScan analyses are shown for each family. (a) Among the 77 *T. vaginalis* GP63-like entries the majority have an identifiable TM domain (52) and functional Zincin domain (HExxHxxG). Several proteins also possess identifiable signal peptides (SP). Repeats linking the protease domain to the C-terminal TM domain (pink box) are also present in several entries. (b) There are 28 subtilisin-like serine proteases with an identifiable TM domain. Some entries have also EGF-like domain, consistent with these entries representing surface proteins interacting with host proteins. (c) Similarly, some of the entries of the nine serine proteases with TM domain possess an

EGF-like domain. (d) The five calpain-like proteases with 22-23 TM domains, could make up a transporter. The drawings are not to scale between the different families and length variation exists between members of a given family.

Figure 3. Structural organisation of candidate surface proteins sharing domains or motifs with other mucosal pathogens. Similar to the BspA-like proteins and the TP, all shown proteins have their identified domains inferred to be facing the outside of the cell if they were to be expressed on the cell surface. (a) A total of 17 *T. vaginalis* proteins with TM domain also possess at least two domains defined from *Chlamydia* polymorphic membrane proteins (named Pmp proteins – e.g.[49]). These Pmp domains are characterized by the conserved GGAI motif, which are present in the *T. vaginalis* proteins (up to six GGAI copies per *T. vaginalis* protein). The GGAI motifs are thought to play an important role in the *Chlamydia* Pmps and were shown to be part of the Pmp segment that is under functional constrains such as immune pressure[49]. (b) Eight *T. vaginalis* proteins possess an identifiable TM domain and show significant sequence similarity to *E. histolytica* immunodominant variable surface antigen M17. They possess the domain named PA14 domain, which has been shown or thought to be sugar-binding domain and found in bacterial and fungal adhesins or other surface proteins or toxins[67]. (c) 11 proteins are made up a legume-like lectin domain and one TM domain. If expressed on the cell surface these proteins could mediate binding to sugar moieties of various host glyco-conjugates. (d) Five proteins were also shown to possess two VSP-like (variant surface proteins) domains made of cysteine-rich repeats that were defined from *Giardia* VSPs contributing to the dipomonad evading the vertebrate host adaptive immune responses by differential expression[50]. Several members of these protein families (in a, b and d) also possess repeats such as PRR (pink rectangles) (as for BspA-like and some TP proteins, Fig. 1-2) consistent with these having a binding function.

Tables

Table 1. *T. vaginalis* hydrogenosomal enzymes suggested to play adhesins roles

Table 2. *T. vaginalis* candidate surface proteins identified *in silico* discussed in this review.

Table S1. *T. vaginalis* hydrogenosomal enzymes suggested to be adhesins.

Figure 1

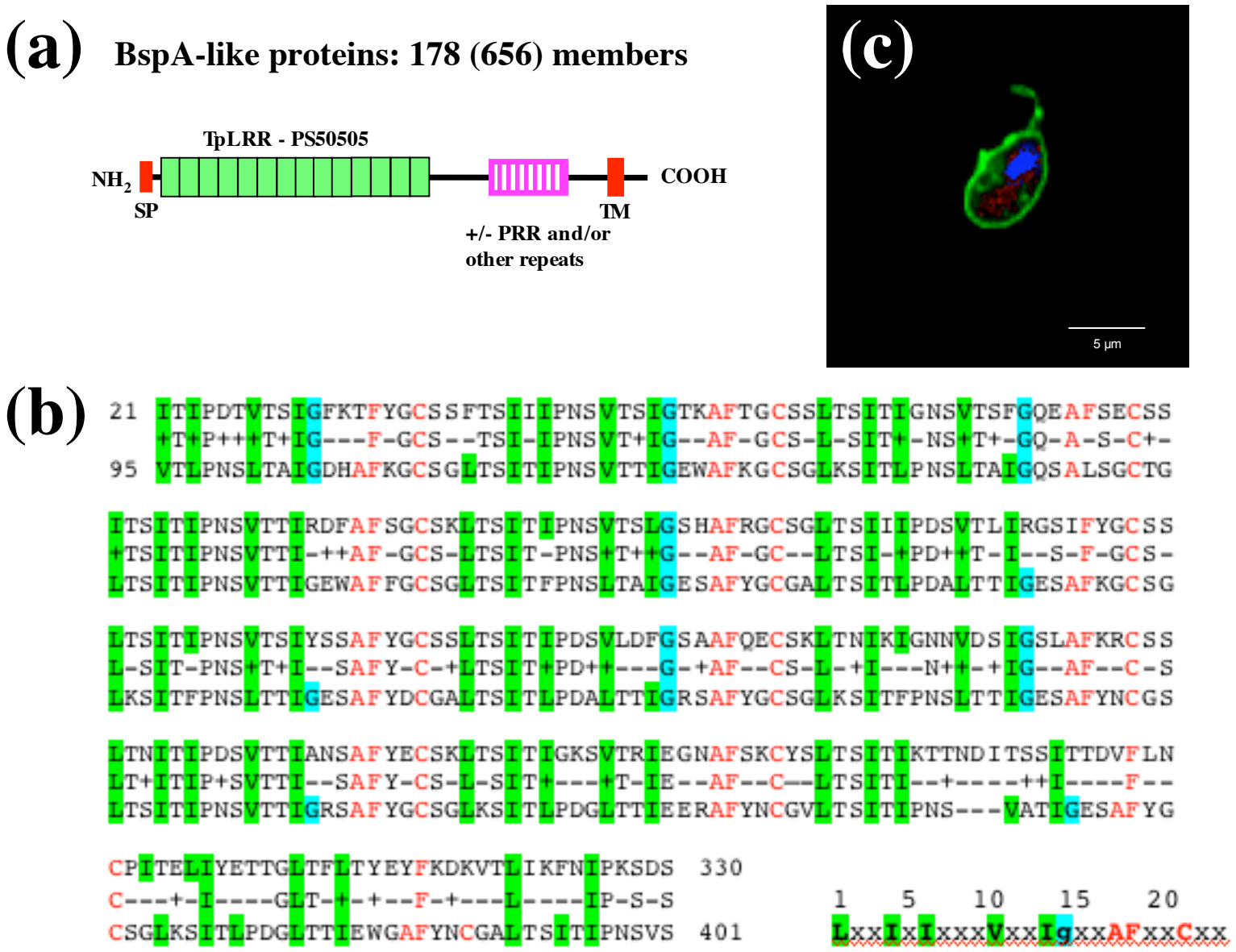
(a)
BspA-like proteins: 178 (656) members
TpLRR - PS50505
NH2
COOH
SP
TM
+/- PRR and/or
other repeats
(c)
5 µm
(b)
21 ITIPDTVTSIGFKTFYGCSSFTSIIIPNSVTSIGTKAFTGCSSLTSITIGNSVTSFGQEAFSECSS
+T+P+++T+IG---F-GCS--TSI-IPNSVT+IG--AF-GCS-L-SIT+-NS+T+-GQ-A-S-C+-
95 VTLPNSLTAIGDHAFKGCSGLTSITIPNSVTTIGEWAFKGCSGLKSITLPNSLTAIGQSALSGCTG
ITSITIPNSVTTIRDFAFSGCSKLTSITIPNSVTSLGSHAFRGCSGLTSIIIPDSVTLIRGSIFYGCSS
+TSITIPNSVTTI-++AF-GCS-LTSIT-PNS+T++G--AF-GC--LTSI-+PD++T-I--S-F-GCS-
LTSITIPNSVTTIGEWAFFGCSGLTSITFPNSLTAIGESAFYGCGALTSITLPDALTTIGESAFKGCSG
LTSITIPNSVTSIYSSAFYGCSSLTSITIPDSVLDFGSAAFQECSKLTNIKIGNNVDSIGSLAFKRCSS
L-SIT-PNS+T+I--SAFY-C-+LTSIT+PD++---G-+AF--CS-L-+I---N++-+IG--AF--C-S
LKSITFPNSLTTIGESAFYDCGALTSITLPDALTTIGRSAFYGCSGLKSITFPNSLTTIGESAFYNCGS
LTNITIPDSVTTIANSAFYECSKLTSITIGKSVTRIEGNAFSKCYSLTSITIKTTNDITSSITTDVFLN
LT+ITIP+SVTTI--SAFY-CS-L-SIT+---+T-IE--AF--C--LTSITI--+----++I----F--
LTSITIPNSVTTIGRSAFYGCSGLKSITLPDGLTTIEERAFYNCGVLTSITIPNS---VATIGESAFYG
CPITELIYETTGLTFLTYEYFKDKVTLIKFNIPKSDS 330
C---+-I----GLT-+-+--F-+---L----IP-S-S
CSGLKSITLPDGLTTIEWGAFYNCGALTSITIPNSVS 401
1 5 10 15 20
LxxIxIxxxVxxIgxxAFxxCxx

Figure 2

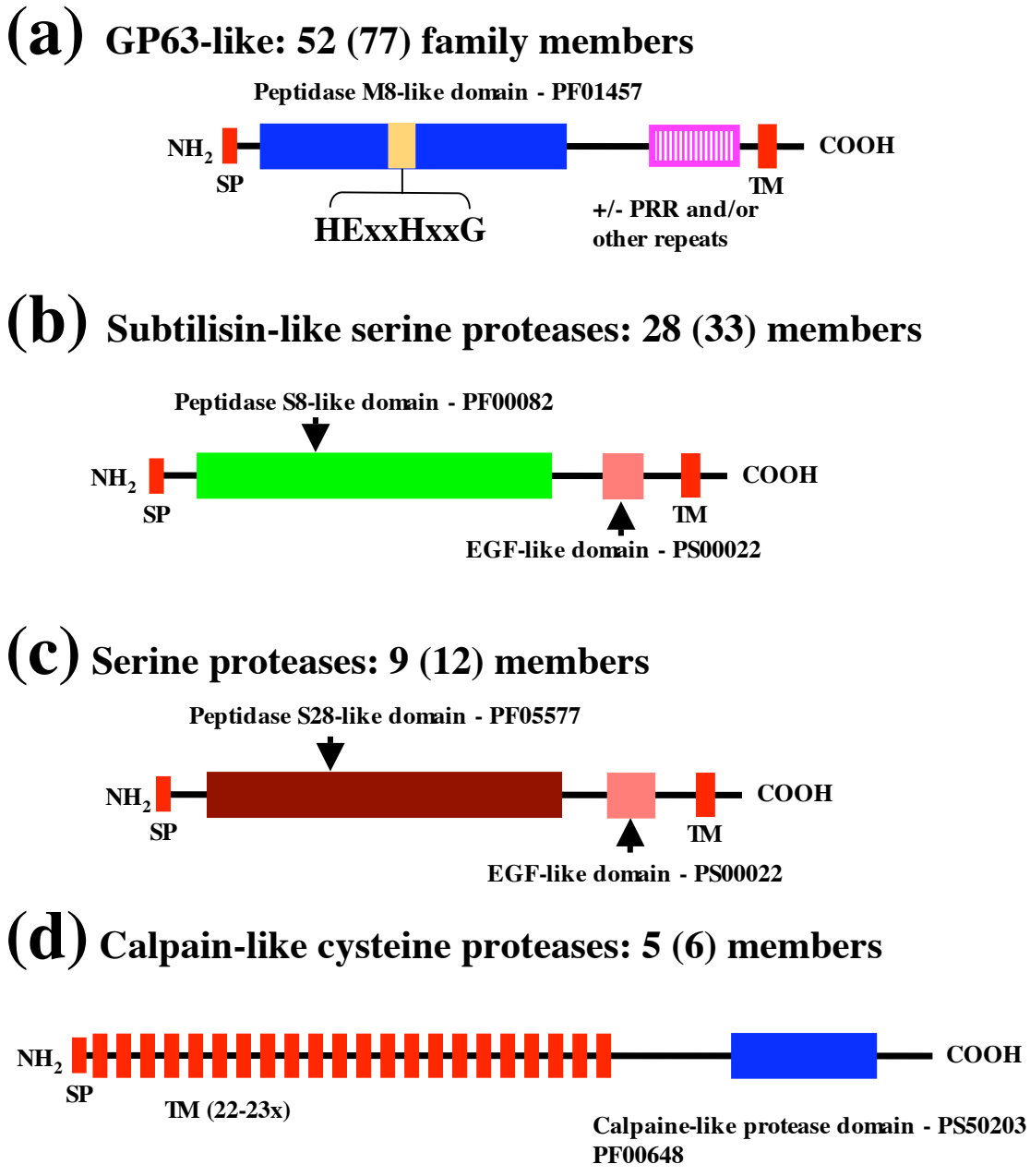
(a) GP63-like: 52 (77) family members
Peptidase M8-like domain - PF01457
NH2
SP
HExxHxxG
+/- PRR and/or
other repeats
TM
COOH
(b) Subtilisin-like serine proteases: 28 (33) members
Peptidase S8-like domain - PF00082
NH2
SP
EGF-like domain - PS00022
TM
COOH
(c) Serine proteases: 9 (12) members
Peptidase S28-like domain - PF05577
NH2
SP
EGF-like domain - PS00022
TM
COOH
(d) Calpain-like cysteine proteases: 5 (6) members
NH2
SP
TM (22-23x)
Calpaine-like protease domain - PS50203
PF00648
COOH

Figure 3

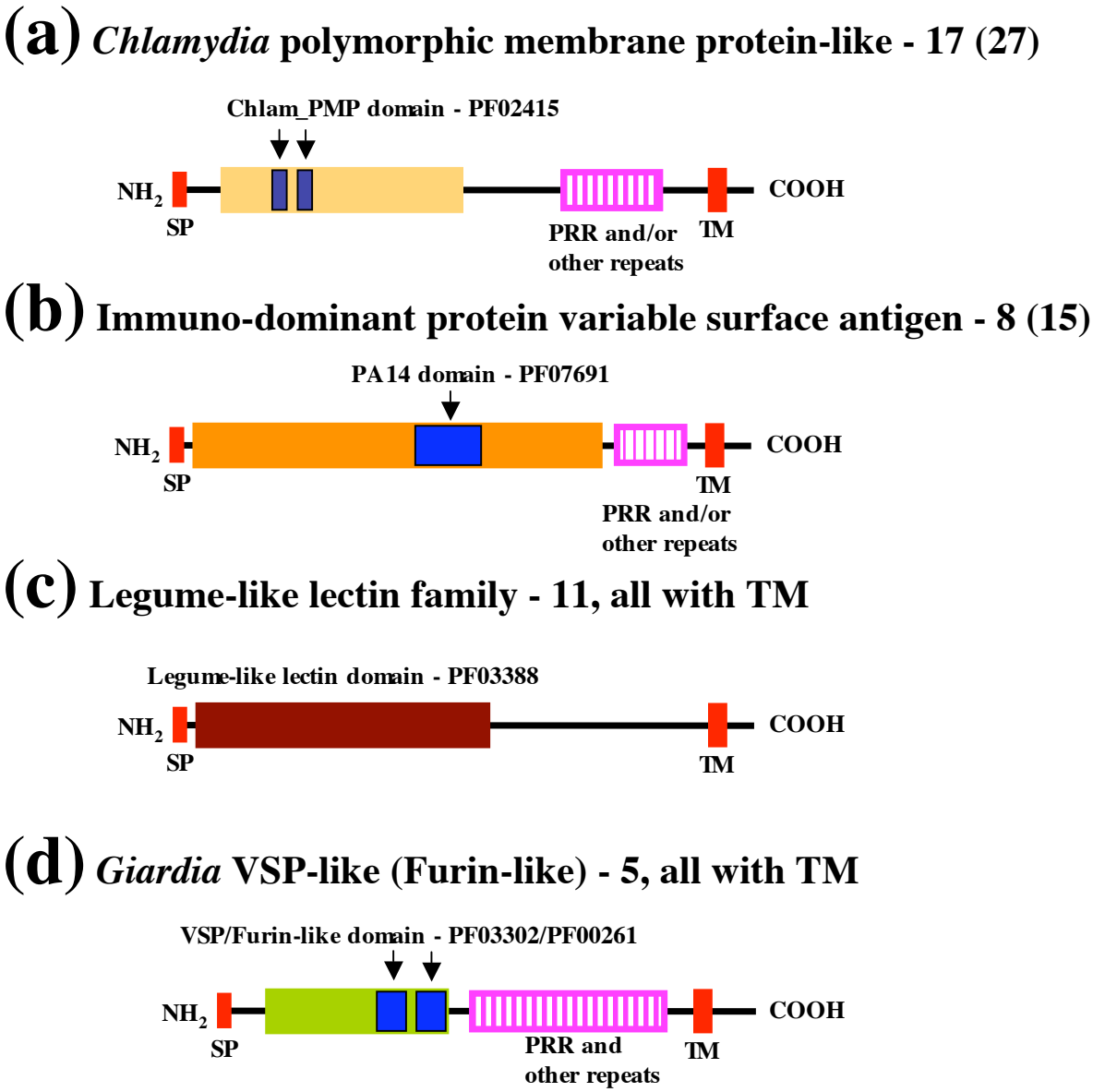
(a) Chlamydia polymorphic membrane protein-like - 17 (27)
Chlam_PMP domain - PF02415
NH2
SP
PRR and/or
other repeats
TM
COOH
(b) Immuno-dominant protein variable surface antigen - 8 (15)
PA14 domain - PF07691
NH2
SP
TM
PRR and/or
other repeats
COOH
(c) Legume-like lectin family - 11, all with TM
Legume-like lectin domain - PF03388
NH2
SP
TM
COOH
(d) Giardia VSP-like (Furin-like) - 5, all with TM
VSP/Furin-like domain - PF03302/PF00261
NH2
SP
PRR and
other repeats
TM
COOH

Table 1. Prediction of N-terminal extensions targeting metabolic enzymes to hydrogenosomes

Locus tag[a]	N-terminal sequence[b]	Enzyme name[c] (alternative adhesin name)
TVAG_267870	**MLTSSVSVPVRN**	Malic enzyme A* (AP65-2)
TVAG_238830	**MLTSSVNFPARE**	Malic enzyme B*
TVAG_416100	**MLTSVSYPVRN**	Malic enzyme C*
TVAG_412220	**MLTSVSLPVRN**	Malic enzyme D*
TVAG_340290	MLTSSVSLPARE	Malic enzyme H* (AP65-1)
TVAG_183790	MLASSVAAPVRN	Malic enzyme* (AP65-3)
TVAG_318670	**MLAGDFSRN**	Succinyl-CoA synthetase alpha-chain* (AP33-1)
TVAG_165340	**MLSSSFERN**	Succinyl-CoA synthetase alpha-chain* (AP33-2)
TVAG_047890	**MLSSSFERN**	Succinyl-CoA synthetase alpha-chain* (AP33-3)
TVAG_259190	**MLSSSFARN**	Succinyl-CoA synthetase beta-chain* (AP51-1)
TVAG_144730	MLSNGSFARN	Succinyl-CoA synthetase beta-chain* (AP51-2)
TVAG_183500	MLSASSNFARN	Succinyl-CoA synthetase beta-chain (AP51-3)
TVAG_198110	**MLRSF**	Pyruvate:ferredoxin oxidoreductase A* (AP120)
TVAG_230580	MLRNF	Pyruvate:ferredoxin oxidoreductase B1
TVAG_242960	MLRNF	Pyruvate:ferredoxin oxidoreductase B2

[a]Locus tag from the genome data (deposited at the NCBI). For references see Table S1
[b]Cleavage of extension confirmed by N-terminal amino acid sequencing are in bold
[c]Name of gene products with experimentally verified hydrogenosomal localization have a*

Table 2. *Trichomonas vaginalis* candidate surface protein families discussed in this review

Protein family name[a]	Domain/motif[b]	Origin	Family size	With TM[c]
Bacterial-like adhesins				
BspA-like (TpLRR)	PS50505	Likely prokaryotic	656	178
Proteases				
GP63-like (Clan MA, Family M8) (leishmanolysine-like)	PF01457	Eukaryotic	77	52
Subtilisin-like serine protease (Clan SB, Family S8)	PF00082	Eukaryotic	33	28
Serine protease (Clan SC, Family S28)	PF05577	Eukaryotic	12	9
Calpain-like cysteine protease	PS50203	Eukaryotic	6	5
Other (adhesins or else)				
Chlamydia polymorphic membrane protein	PF02415	Ambiguous	27	17
Immuno-dominant variable surface antigen (*Entamoeba*)	PF07691	Ambiguous	15	8
Legume-like lectin family	PF03388	Eukaryotic	11	11

Giardia VSP-like membrane protein	PF03302	Eukaryotic	5	5
P270 surface immunogen	na	*Trichomonas* specific	5	5
Total entries			847	318

a TpLRR, *Treponema pallidium* leucine-rich repeat
b PSXXXXX, PROSITE entries, PFXXXXX, Pfam entries
c Transmembrane domain (TM)

Table S1. *T. vaginalis* hydrogenosomal enzymes suggested to be adhesins

Names given for identical gene in different publications and corresponding locus tag are in row

Malate dehydrogenase (decarboxylating) = Malic enzyme (putative adhesin AP-65)

[1]Hrd_ & Muller 1995	[2]Dyall et al. 2004	[3,4]Alderete et al. 1995;1996	Locus tag (NCBI)
MEA	MAE sub.A	AP65-2	TVAG_267870
MEB	MAE sub.B		TVAG_238830
MEC			TVAG_416100
MED	MAE sub.G		TVAG_412220
	MAE sub.H	AP65-1	TVAG_340290
		AP65-3	TVAG_183790

Succinyl CoA synthase (SCS) α subunit (putative adhesin AP-33)

[5]Lahti et al.1994a	[6]Enbring & Alderete1998	Locus tag (NCBI)
SCS-1 α sub.	AP33-1	TVAG_318670
SCS-2 α sub.	AP33-2	TVAG_165340
SCS-3 α sub.	AP33-3	TVAG_047890

Succinyl CoA syntase (SCS) β subunit (putative adhesin AP-51)

[7]Lahti et al.1994b	[8]Alderete et al.1998	Locus tag (NCBI)
SCS β sub.	AP51-1	TVAG_259190
	AP51-2	TVAG_144730
	AP51-3	TVAG_183500

Pyruvate:ferredoxin oxidoreductase (PFOR) (putative adhesine AP-120)

[9]Hrd_ & Muller, 1995	[10]Upcroft et al. 2006	[11]Moreno-Brito et al., 2005	Locus tag (NCBI)
PFOA	PFOA	AP120	TVAG_198110
PFOB	PFOB1		TVAG_230580
	PFOB2		TVAG_242960

Predictions of N-terminal presequence targeting the proteins to hydrogenosomes

Targeting presequence*	Hydrogenosomal protein**	Locus tag (NCBI)
MLTSSVSVPVRN[1]	Malic enzyme A (AP65-2)[1]	TVAG_267870
MLTSSVNFPARE[1]	Malic enzyme B[1]	TVAG_238830
MLTSVSYPVRN[1]	Malic enzyme C[1]	TVAG_416100
MLTSVSLPVRN[1]	Malic enzyme D[1,2]	TVAG_412220
MLTSSVSLPARE	Malic enzyme H (AP65-1)[2]	TVAG_340290
MLASSVAAPVRN	Malic enzyme (AP65-3)[13]	TVAG_183790
MLAGDFSRN[5]	Succinyl-CoA synthetase alpha-chain (AP33-1)[5]	TVAG_318670
MLSSSFERN[5]	Succinyl-CoA synthetase alpha-chain (AP33-2)[5]	TVAG_165340
MLSSSFERN[5]	Succinyl-CoA synthetase alpha-chain (AP33-3)[5]	TVAG_047890
MLSSSFARN[7]	Succinyl-CoA synthetase beta-chain (AP51-1)[7]	TVAG_259190
MLSNGSFARN	Succinyl-CoA synthetase beta-chain (AP51-2)[7]	TVAG_144730
MLSASSNFARN	Succinyl-CoA synthetase beta-chain (AP51-3)	TVAG_183500
MLRSF[9]	Pyruvate:ferredoxin oxidoreductase A (AP120)[9]	TVAG_198110
MLRNF	Pyruvate:ferredoxin oxidoreductase B1[13]	TVAG_230580
MLRNF	Pyruvate:ferredoxin oxidoreductase B2[13]	TVAG_242960

*Processing of the targeting sequences in bold has been experimentaly confirmed (see reference).
** Presence of the protein in hydrogenosomes has been demonstrated in given reference.

References

[1]Hrdy I, Müller M. Primary structure of the hydrogenosomal malic enzyme of *Trichomonas vaginalis* and its relationship to homologous enzymes. *J. Eukaryot. Microbiol.* 1995; **42:** 593-603.

[2]Dyall SD, Yan W, Delgadillo-Correa MG *et al.* Non-mitochondrial complex I proteins in a hydrogenosomal oxidoreductase complex. *Nature* 2004; **431:** 1103-7.

[3]Alderete JF, O'Brien JL, Arroyo R *et al.* Cloning and molecular characterization of two genes encoding adhesion proteins involved in *Trichomonas vaginalis* cytoadherence. *Mol. Microbiol.* 1995; **17:** 69-83.

[4]O'Brien JL, Lauriano CM, Alderete JF. Molecular characterization of a third malic enzyme-like AP65 adhesin gene of Trichomonas vaginalis. *Microb. Pathog.* 1996; **20:** 335-49.

[5]Lahti CJ, Bradley PJ, Johnson PJ. Molecular characterization of the alpha-subunit of *Trichomonas vaginalis* hydrogenosomal succinyl CoA synthetase. *Mol. Biochem. Parasitol.* 1994; **66:** 309-18.

[6] Engbring JA, Alderete JF. Three genes encode distinct AP33 proteins involved in *Trichomonas vaginalis* cytoadherence. *Mol. Microbiol.* 1998; **28:** 305-13.

[7]Lahti CJ, d'Oliveira CE, Johnson PJ. Beta-succinyl-coenzyme A synthetase from *Trichomonas vaginalis* is a soluble hydrogenosomal protein with an amino-terminal sequence that resembles mitochondrial presequences. *J. Bacteriol.* 1992; **174:** 6822-30.

[8]Alderete JF, Engbring J, Lauriano CM, O'Brien JL. Only two of the *Trichomonas vaginalis* triplet AP51 adhesins are regulated by iron. *Microb. Pathog.* 1998; **24:** 1-16.

[9]Hrdy I, Müller M. Primary structure and eubacterial relationships of the pyruvate:ferredoxin oxidoreductase of the amitochondriate eukaryote *Trichomonas vaginalis*. *J. Mol. Evol.* 1995; **41:** 388-96.

[10]Upcroft JA, Delgadillo-Correa MG, Dunne RL, Sturm AW, Johnson PJ, Upcroft P. Genotyping *Trichomonas vaginalis*. *Int. J. Parasitol.* 2006; **36:** 821-8.

[11]Moreno-Brito V, Yanez-Gomez C, Meza-Cervantez P *et al.* A *Trichomonas vaginalis* 120 kDa protein with identity to hydrogenosome pyruvate:ferredoxin oxidoreductase is a surface adhesin induced by iron. *Cell Microbiol.* 2005; **7:** 245-58.

[13] Tachezy J, unpublished data.