MCP Papers in Press. Published on November 23, 2006 as Manuscript M600372-MCP200

Molecular & Cellular Proteomics

1 Proteomic analysis and protein expression during sporozoite excystation of
2 *Cryptosporidium parvum* (coccidia, Apicomplexa)

3

4 William J. Snelling[1], Qishan Lin[2], John E. Moore[3], B. Cherie Millar[3], Fabio Tosini[4],
5 Edoardo Pozio[4], James S.G. Dooley[1], and Colm J. Lowery[1*]

6

7 [1]Centre for Molecular Biosciences, School of Biomedical Sciences,
8 University of Ulster, Cromore Road, Coleraine, County Londonderry, N. Ireland,
9 BT52 1SA.
10 [2]UAlbany Proteomics Facility, Centre for Functional Genomics, University of
11 Albany, One Discovery Drive, Rensselaer, NY 12144.
12 [3]Northern Ireland Public Health Laboratory, Department of Bacteriology,
13 Belfast City Hospital, Belfast, N. Ireland, BT9 7AD.
14 [4]Department of Infectious, Parasitic, and Immunomediated Diseases, Istituto
15 Superiore di Sanità, Rome, Italy.

16

17 [*]: To whom correspondence may be addressed: Tel: +44 (0)28 70323528; Fax: +44
18 (0)28 7032 4375; Email cj.lowery@ulster.ac.uk.

19

20 Running title: Analysis of *Cryptosporidium parvum* proteome

21

22 Key words: *Cryptosporidium parvum*, excystation, proteomics, sporozoites, oocysts,
23 vaccine development

Downloaded from www.mcponline.org by on November 24, 2006

Copyright 2006 by The American Society for Biochemistry and Molecular Biology, Inc.

Molecular & Cellular Proteomics

1 **Abbreviations:**

2 AC: apical complex

3 AdhE: acetaldehyde reductase plus alcohol dehydrogenase

4 AP: anterior pole

5 ARM: armadillo repeat domain

6 CPV: capsid protein

7 DAPI-PI: 4',6'-diamidino-2-phenylindole-propidium iodide

8 DG: dense granules

9 FA: formic acid

10 GIPLs: glycosylinositol phospholipids

11 GPI: glycosylphosphatidylinositol

12 HDAC: histone deacetylase interacting domain

13 HSP: heat shock protein

14 IAA: iodoacetamide

15 IDA: integrated data appliance

16 LCCL: Limulus factor C, cochlear protein Coch-5b2 and late gestation lung protein

17 MBF1: multiprotein bridging factor type 1

18 M: micronemes

19 MT: mitochondrion

20 N: nucleus

21 OW: oocyst wall

22 PP: posterior pole

23 R: rhoptry

24 RB: refractile body

25 S: scissure

Downloaded from www.mcponline.org by on November 24, 2006

1 SPO: sporozoite

2 TBP: tributylphosphine

3 TCTP: translationally controlled tumor protein

Molecular & Cellular Proteomics
ASBMB
MCP

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics
ASBMB
MCP

1 **SUMMARY**

2 Cryptosporidiosis, caused by coccidian parasites of the genus *Cryptosporidium*, is a
3 major cause of human gastrointestinal infections and poses a significant health risk,
4 especially to immunocompromised patients. Despite intensive efforts for more than
5 20 years, there is currently no effective drug treatment against these protozoa. This
6 study examined the zoonotic species *Cryptosporidium parvum* at two important stages
7 of its life cycle; the non-excysted (transmissive) and excysted (infective) forms. To
8 increase our understanding of the molecular basis of sporozoite excystation, LC-
9 MS/MS coupling with a stable isotope N-terminal labelling strategy using iTRAQ^{TM}
10 reagents was employed on soluble fractions of both non-excysted and excysted
11 sporozoites, i.e. sporozoites both inside and outside oocysts were examined.
12 Sporozoites are the infective stage which penetrate small intestinal enterocytes. Also,
13 to increase our knowledge of the *C. parvum* proteome, shotgun sequencing was
14 performed on insoluble fractions from both non-excysted and excysted sporozoites.
15 In total 303 *C. parvum* proteins were identified, 56 of which, hitherto described as
16 being only hypothetical proteins, are expressed in both excysted and non-excysted
17 sporozoites. Importantly, we demonstrated that the expression of 26 proteins increase
18 significantly during excystation. These excystation induced proteins included
19 ribosomal proteins, metabolic enzymes, and heat shock proteins. Interestingly, three
20 Apicomplexa-specific proteins and five *Cryptosporidium*-specific proteins augment in
21 excysted invasive sporozoites. These eight proteins represent promising targets for
22 developing vaccines or chemotherapies that could block parasite entry into host cells.

Downloaded from www.mcponline.org by on November 24, 2006

1 **INTRODUCTION**

2 The protozoan parasite *Cryptosporidium parvum*, which causes acute
3 gastroenteritis in both humans and animals, is considered an important pathogen all
4 over the world (1). *Cryptosporidium* is one of several genera in the phylum
5 Apicomplexa, whose members share a common apical secretory apparatus mediating
6 locomotion and tissue or cellular invasion (2). There are currently 16 recognised
7 species in this genus (3), all of which have 4 naked sporozoites, contained within a
8 thick walled oocyst (Fig 1A); the resistant and transmissive form of these parasites
9 (4). Species of the genus *Cryptosporidium* have a complex life cycle. Upon
10 ingestion, excystation of viable oocysts, 4-6 µm diameter for intestinal species and
11 gastric species are slightly larger, is triggered (5,6). Following excystation,
12 sporozoites (Fig 1B; the infective stage) attach to the intestinal epithelium, are
13 enveloped by the apical membrane and reside in an intracellular, extracytoplasmic
14 parasitophorous vacuole (5,7). In this peculiar niche, the parasite undergoes an
15 endogenous asexual and sexual reproductive development (Fig. 1C), culminating in
16 the production of an encysted stage discharged in the faeces of their host (6,8).
17 Coccidian oocyst stages are highly resistant to environmental stress and treatments,
18 e.g. chemical disinfection, which is attributed to a durable oocyst wall; a complex
19 protective barrier consisting of a double layer of a protein-lipid-carbohydrate matrix,
20 allowing the parasite to stably persist outside hosts (2,6). *Cryptosporidium parvum*
21 causes self-limited watery diarrhoea in immunocompetent subjects, but is life-
22 threatening in immunocompromised patients (1,9,10). Cryptosporidiosis is also
23 responsible for significant neonatal morbidity in farmed livestock and causes weight
24 loss and growth retardation, leading to large economic losses (11). *Cryptosporidium*
25 *parvum* is transmitted by fecal contamination of food or water or by accidental

Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006

1 ingestion of infectious oocysts during water-related recreational activities (12). In
2 1993, the largest *Cryptosporidium* spp. outbreak was registered in Milwaukee,
3 Wisconsin, USA where 403,000 people were infected through contaminated drinking
4 water (13), with associated economic costs of $31.7 million in medical costs and
5 $64.6 million in productivity losses (14). Over the last two decades increasing
6 numbers of cryptosporidiosis outbreaks have been recorded in developed countries
7 and the importance of the zoonotic species *C. parvum* is being recognized by both
8 government agencies and the global scientific community (15,16). Unlike many
9 organisms belonging to the phylum Apicomplexa, such as *Plasmodium* spp. and
10 *Toxoplasma gondii*, there is no clinically proven effective drug treatment against
11 *Cyrptosporidium* spp. (17). Proteomic profiling is a useful approach for obtaining a
12 global overview of the proteins present in a system under differing conditions and can
13 aid in understanding the molecular determinants involved with pathogenesis and
14 vaccine development (18). Compared with other Apicomplexans, e.g. *Toxoplasma*
15 *gondii,* both the limited supply of purified parasite material and the lack of
16 transfection systems have restricted analyses of proteins in parasites of the genus
17 *Cryptosporidium* (15). The genomic analysis of *C. parvum* and the closely related *C.*
18 *hominis* (19) have revealed extremely streamlined metabolic pathways and a reliance
19 on hosts for nutrients (20). *Cryptosporidium parvum* has an animal-type *O*-linked
20 glycosylation pathway and >30 predicted surface proteins with mucin-like segments
21 (21).
22
23 The mechanisms involved in *Cryptosporidium* spp. excystation are incompletely
24 understood, although roles for both host and parasite-derived components are
25 recognized, little is known about their precise involvements (5). In vitro excystation

Molecular & Cellular Proteomics MCP ASBMB

Downloaded from www.mcponline.org by on November 24, 2006

ASBMB Molecular & Cellular Proteomics MCP

protocols for *C. parvum* mimic host-derived signals, including temperature (37°C), pH fluctuations, and bile salts (5). We adopted an accepted *in vitro* excystation method (22) and analyzed *C. parvum* proteins at two key stages, both the transmissive (oocysts) and infective stages (sporozoites), of its life cycle (5,6). Thus, as well as evaluating the proteome of *C. parvum*, a major aim of our study was to identify proteins which demonstrated significant increases of expression during oocyst excystation, a vital step of the pathogenesis of these coccidian (6). This was achieved by employing LC/MS and iTRAQ™ reagent isotope labelling, providing a powerful tool for the identification and quantification (23,24) of soluble *C. parvum* proteins. Using LC/MS with iTRAQ the protein expression in non-excysted and excysted soluble *C. parvum* oocyst fractions was compared. Then using shotgun peptide sequencing (25) the insoluble protein content of non-excysted and excysted oocysts was also analyzed.

Downloaded from www.mcponline.org by on November 24, 2006

1 **EXPERIMENTAL PROCEDURES**

2 **Oocyst isolation and sporulation.** To recover high numbers of *C. parvum* oocysts,
3 with minimal faecal and bacterial contaminants (26), 4 x 10^8 *C. parvum* oocysts
4 (strain code ISSC162) were recovered from experimentally infected calves using
5 faeces purification with sucrose and Percoll density gradients (27). Purified oocysts
6 were stored at 4°C for no longer than three weeks. Using 4',6'-diamidino-2-
7 phenylindole-propidium iodide assays (DAPI-PI) assays (28), oocyst viability was
8 determined to be at least 92%. To obtain free sporozoites, half of the oocysts (2 x
9 10^8) were excysted as previously described (22). Briefly, oocysts were resuspended
10 in 10 mM HCl (Sigma, St. Louis, MO) and incubated at 37°C for 10 min. The
11 suspension was centrifuged at 3,000 x *g* for 5 min. The pellet was then resuspended
12 in 2 mM sodium taurocholate in PBS (Sigma), incubated at 15°C for 10 min, followed
13 by 37°C for 8 min and a high level of excystation i.e., at least 90%, was confirmed
14 (29).
15
16 ***C. parvum* lysis, protein precipitation and fraction preparation.** All excysted and
17 non-excysted oocycts were solubilzed (30 min at 4°C) in lysis buffer [50 mM Tris
18 (BioRad, Hercules, CA) 5 mM EDTA (Sigma), 5 mM iodoacetamide (IAA) (Sigma),
19 0.1 mM N-α-p-tosyl-L-lysyl chloromethyl ketone (Sigma), 1 mM
20 phenylmethylsulfonyl fluoride (Sigma), 1% (wt/vol) octyl-glucoside (Roche,
21 Hertfordshire, UK)], and stored at -80°C (30). Upon thawing, samples were sonicated
22 with twenty 10 s pulses (50 W) on ice, with 1 min intervals and were then
23 ultracentrifuged at 4°C (20,000 x g, 30 min) to separate the soluble and insoluble
24 fractions (30). Insoluble fractions were then rinsed 3 times with ice-cold PBS and
25 stored at -80°C. Proteins in soluble lysate fractions were then precipitated by mixing

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics

oocyst lysates with eight volumes of ice-cold acetone, 1 volume of TCA and 0.7% (wt/vol) 2-Mercaptoethanol at -20°C for 1 h, and then centrifuged at 18,000 x *g* for 15 min at 4°C (31,32). Pellets of precipitated soluble proteins were then washed with 1 ml of ice-cold acetone, centrifuged at 18,000 x *g* for 15 min at 4°C, covered in a layer of ice-cold acetone, and stored at -80°C.

Trypsin digestion, shotgun peptide sequencing and iTRAQ isobaric labelling. The pellets (soluble and insoluble fractions) were re-suspended in 400 µl sample preparation buffer [100 mM Tris-HCl, 8 M urea, 0.4% SDS, 5 mM tributylphosphine (TBP), pH 8.3] and placed on ice (33). Sample mixtures were sonicated with five 20 s pulses (50 W), with 1 min intervals on ice, followed by IAA alkylation at room temperature for 1 h (34). Samples were then ultracentrifuged at 100,000 x g for 20 min at room temperature. Protein concentrations were measured using a MicroBCA protein assay kit (Pierce, Rockford, IL) according to the manufacturer's instruction (35), before reactions were quenched by the addition of DTT (33). The protein mixtures were diluted 8-fold in 50 mM Tris-HCl. Modified trypsin (Sigma, MO) was added to a final substrate-to enzyme ratio of 30:1 and the trypsin digests were incubated overnight at 37°C. The peptides from each digest solution were acidified to pH 3.0 with formic acid (FA) and loaded onto a Discovery DSC-18 Cartridge (Sigma, MO). Then peptides were desalted (5 ml of 0.1% FA) and eluted with 5 ml of a solution composed of 50% ACN with 0.1% FA. Peptides derived from insoluble fractions were analyzed qualitatively using shotgun peptide sequencing as previously described (36). For peptides derived from soluble fractions, equal amounts (100 µg) of sample were labelled with two iTRAQ reagents (Appled Biosystems, CA) following the manufacture's instruction; TRAQ-114 (non-

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics ASBMB MCP

1 excysted) and iTRAQ-115 (excysted). Briefly, after desalting on a C18 cartridge the
2 peptide mixture was lyophilized and re-suspended in 30 µl of 0.5 M
3 triethylammonium bicarbonate (N(Et)3HCO3), pH 8.5. The appropriate iTRAQ
4 reagent (dissolved in 70 µl ethanol) was added, allowed to react for 1 h at room
5 temperature, and then quenched with 300 µl of ddH2O.
6
7 **Off-line strong cation exchange chromatography.** iTRAQ labelled peptides were
8 then concentrated, mixed, and acidified to a total volume of 2.0 ml. They were then
9 injected into an Agilent 1100 HPLC system with a Zorbax 300-SCX column (4.6 ID
10 ×250 mm) (Agilent, Waldbronn, Germany). Solvent A was 5 mM KHPO4 and 25%
11 ACN (pH 3.0) and solvent B was 350 mM KCL in solvent A. Peptides were eluted
12 from the column with a 40 min mobile phase gradient B of solvent B. A total of 30
13 fractions were collected and samples were dried by a speed-vac prior to LC-MS/MS
14 analysis.
15
16 **On-line nano-LC ESI QqTOF MS analysis.** A nanobore LC system (Dionex,
17 Sunnyvale, CA) which was interfaced to a QSTAR XL QqTOF mass spectrometer
18 with a NanoSpray ion source (Applied Biosystems, Foster City, CA) was used for
19 mass spectrometry. The Magic C18 100 A pore 75 µm ID ×150 mm (Picofrit
20 Woburn, MA) was packed in-house. Solvent A was 3% ACN, 0.1% FA and 0.01%
21 TFA and solvent B was 98% ACN, 0.1% FA and 0.01% TFA. Peptide mixtures
22 (reconstituted in 200 µl of 5% FA) were injected and eluted from the column with a
23 110 min mobile phase solvent B gradient (5-5 % B in 5 min, 5-18% B in 10 min, 18-
24 30% B in 65 min, 30 –60% B in 10 min, 60-90 % B in 10 min, and 90-90% in 5 min)
25 at a flow rate of 250 nl/min. The mass spectrometer was operated in an information

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics
ASBMB

dependent acquisition mode whereby following the interrogation of MS data (m/z 350-1500) using a 1 s survey scan, ions were selected for MS/MS analysis based on their intensity (>15 cpm) and charge state (+2, +3, and +4). A total of 3 product ion scans (2,3,3 s each) were set from each survey scan. Rolling collision energies were chosen automatically based on the m/z and charge-state of the selected precursor ions. The integrated data appliance (IDA) Extensions II script was set to one repetition before dynamic exclusion.

Protein identification and quantification. Identification and quantitation was performed using Pro ID (37) and ProQUANT software 1.1 (38) (Applied Biosystems) using the non-redundant database (NRDB) *C. parvum* sub-database (9811 entries) with an MS and MS/MS mass tolerance of 0.15 Da (39). Variable modification on methionine (oxidation, +16 Da) and fixed modification on cysteine (carbamidomethylation, +57 Da) residues were considered during the searching (39). Protein identification with confidence scores of >95% were considered significant and the false positive rate was determined from decoy database (39) searching (data not shown). The identifications of all detected *C. parvum* proteins were then verified using BlastP at http://www.ncbi.nlm.nih.gov/BLAST/, or where appropriate at CryptoDB 3.3 (http://cryptodb.org/cryptodb/).

Using ProQUANT software, quantification was based upon the signature peak areas (m/z: 114, 115) and corrected according to the manufacturer's instructions to account for isotopic overlap. Briefly, non-excysted ooysts (control) were labelled with isobaric tag 114 and relative quantification ratios of the identified proteins were calculated, averaged, and corrected for systematic error in labelling from the iTRAQ

Downloaded from www.mcponline.org by on November 24, 2006

peptides. For proteins in which every peptide's (labelled with tag 115) iTRAQ ratio was 0, the average iTRAQ ratio was set to 0; an indication a protein which may be highly down-regulated relative to the control. Similarly for those proteins where every (tag 115) ratio was 9999, the average was set to 9999 (indicating a protein which may be highly up-regulated). Statistically significant changes were weighted by the error factor (a measure of the variation between the different iTRAQ ratios for the reagent pair), and *p* value (95% confidence interval, which defined a range in which the true average iTRAQ ratio was 95% likely to fall into). For instance, for a ratio of 1.0 (indicating no change in expression levels), with an error factor of 2 (and the corresponding 95% confidence interval of 0.5-2.0), the protein expression ratio and error were reported as 1.0 ($p<0.05$ with an error factor of 2). Ratios in which the total peak area was <40 counts were omitted. If there were fewer than 2 peptides contributing towards an iTRAQ ratio average, the error factor and *p* values were not valid, and were not calculated. Proteins which augmented significantly during excystation had a $p<0.05$ and an error factor < 2.

Bioinformatic analysis of 26 proteins which significantly augmented during excystation. Significantly over-expressed proteins were examined using BLASTP both on the NCBI database and on ApiDB (http://www.apidb.org/blast/) to find similarity with other Apicomplexan proteins or with proteins from unrelated species. Only proteins identified by the BLASTP search with an expected value lower than 10^{-10} were considered homologs. Searches were also performed for molecular traits and domains of the putative proteins using SMART (http://smart.embl-heidelberg.de/). The predicted coding sequence for the hypothetical protein cgd6_4460 (N46) was extended (results are currently unfinished in CryptoDB), by downloading a larger

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics

1 portion of the AAEE01000002 contig. Gene sequences were predicted using Genscan

2 (http://genes.mit.edu/GENSCAN.html).

Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006

1 **Results**

2 **Protein expression in the soluble fractions of excysted and non-excysted ooycsts.**

3 Using LC/MS with iCAT, 142 proteins were detected in both soluble fractions of
4 excysted and non-excysted ooycsts (supplemental table 1). Ribosomal proteins
5 constituted a significant proportion (35: 24.7%) of proteins which were detected in
6 both of the soluble fractions. Twenty-five proteins (17.6%) were detected which had
7 hitherto been described as being only hypothetical proteins, e.g. N77, (Fig 3; Table 1).
8 Six heat shock proteins and 17 secreted proteins were also expressed in both excysted
9 and non-excysted ooycsts. Many of the detected proteins are involved in
10 infection/pathogenesis, energy pathways, e.g. glycolsis, cellular division and
11 replication and DNA modification. The protein expression of 26 proteins was
12 significantly greater (p<0.05 and error factor <2) in excysted oocysts (Table 1).

13

14 **Bioinformatics analysis of the 26 proteins which augmented during excystation.**

15 There were 18 ubiquitous proteins, which are highly conserved in many species from
16 different phyla. Seven proteins which had hitherto been described as only being
17 hypothetical proteins augmented significantly during excystation. Four are associated
18 with ribosomes and four are heat shock proteins. A number of house keeping, e.g.
19 lactate dehydrogenase (N102), or structural genes, e.g. histone H4 (N118), showed
20 significant increases of expression. The protein Chro.50226 (N35) was a ubiquitous
21 ribosomal protein L3. Eight proteins are restricted to the *Cryptosporidium* genus or to
22 Apicomplexa, and consequently could play a specific role in the invasion process
23 (Table 1). Five of these proteins appear to be restricted to the *Cryptosporidium* genus
24 (Table 1). The hypothetical protein with signal peptide EAK88888 (N77) is highly
25 conserved both in *C. parvum* and *C. hominis* (Table 1). EAK88888 demonstrates a

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics

low similarity with proteins of the genus *Plasmodium,* however, the E value (1 x 10^{-8}) is higher than the fixed cut off point. Therefore, EAK88888 is probably more closely related to similar proteins present in the genus *Plasmodium* than it appears from *C. parvum* and *C. hominis*. However, these proteins contain repetitive motifs that make the identification of a consensus sequence difficult. Three proteins [hypothetical protein cgd6_4460 (N46), hypothetical protein cgd7_4280 (N51), and Cpa135 protein (a135; N64)] demonstrate similarity with proteins from other Apicomplexa genera (*Plasmodium* and/or *Toxoplasma*), but not with proteins from other phyla (Table 1).

Identification of proteins from the insoluble fractions of oocoysts. Using shotgun sequencing 122 proteins were detected from the non-excysted insoluble ooycst fraction (supplemental table 2). Thirty-four (27.9%) of these detected proteins were ribosomal and 26 of the proteins (21.3%) had hitherto been described as being only hypothetical proteins. Five heat shock proteins were also detected in non-excysted ooycsts. As expected a number of surface antigens and oocyst wall associated proteins were also detected in this fraction. As per the soluble oocyst fractions, many of the detected proteins in the insoluble non-excysted oocyst fraction are involved in glycolsis, cellular division and replication and DNA modification. Also, 14 proteins (with unique accession numbers) were only detected from the excysted insoluble oocyst fractions (supplemental table 3).

Surface, cytoskeletal and extra-cellular proteins. During our investigation, in both soluble and insoluble fractions, we detected a number of *C. parvum* cytoskeletal and surface proteins. Detected cytoskeletal proteins included actin (N36) and alpha-tubulin (N277; Table 2). A number of proteins which are associated with extra-

Molecular & Cellular Proteomics
Downloaded from www.mcponline.org by on November 24, 2006

1 cellular protein secretion were detected in both soluble and insoluble fractions. These
2 proteins included a predicted secreted protein signal peptide (N65), aminopeptidase N
3 (N109), and the secreted GP900 (N88; Table 2) (40). Moreover, among the
4 sporozoite secreted proteins which were identified, all the *C. parvum* predicted LCCL
5 containing proteins namely CCp2 (N69) and CCp3 (N70; Table 2), as well as the
6 LCCL containing protein Cpa135 (N64; Table 1), were detected. Several oocyst wall
7 and surface proteins were also detected, including glycoprotein Cp17 (N223), an
8 oocyst wall protein precursor (N224), COWP1 (N225), and 15 kDa glycoprotein
9 (N286; Table 2).
10
11 **Metabolic enzymes.** A number of proteins involved in metabolic reactions were
12 detected in both soluble and insoluble fractions. These enzymes included enolase (2-
13 phosphoglycerate dehydratase) (N98), glycerol-3-phosphate dehydrogenase (N106),
14 and fructose-1,6-bisphosphate aldolase (N236; Table 2). A number of adenosine
15 triphosphate (ATP) related enzymes were also detected, including secreted
16 nucleoside-diphosphate kinase (N73; Table 2). A Pdr17p-like protein (N110), which
17 may regulate lipid synthesis, and a pleckstrin homology (PH domain) containing
18 protein (N124) were also detected (Table 2).
19
20 **Nucleic acids and protein synthesis and modification and replication.** In both
21 soluble and insoluble oocyst fractions a number of enzymes which are involved in
22 nucleic acid synthesis, repair and modification were detected. Examples of these
23 included a small GTP binding protein rab1a (N91) and guanine nucleotide-binding
24 protein (N115; Table 2). A number of enzymes were detected that would have key
25 functions during the *Cryptosporidium* sp. replication, protein synthesis and

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics

1 modification. Detected proteins with functions during replication included a GTP-
2 binding nuclear protein ran/tc4 (N136), ribonucleotide reductase (N131), transcription
3 regulator (N133), and a SEC14 domain containing protein (N264; Table 2). As well
4 as, a number of ribosomal proteins, translation elongation factor 2 (EF-2; N137), and
5 ubiquitin-activating enzyme E1 (N123) were detected, which have key roles during
6 protein synthesis and in proteasome/ubiquitin systems, respectively (41) (Table 2). A
7 number of enzymes associated with the protein structural modification were also
8 detected, e.g. protein disulfide isomerase (N121), and Pre3p/proteasome regulatory
9 subunit beta type 6, NTN hydrolase fold (N122; Table 2).

Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics
ASBMB
MCP

1 **DISCUSSION**

2 The first and most extensive proteome maps for all the life cycle stages of an
3 Apicomplexan parasite were published in 2002 for *Plasmodium falciparum* (42).
4 Proteome maps are available for the invasive stages of development of the other
5 Apicomplexan parasites including the sporozoite of *Eimeria tenella*, as well as the
6 tachyzoite of *T. gondii,* and *Neospora caninum* (42). Our study is the first major
7 proteomic investigation of any *Cryptosporidium* species, and in particular of *C.*
8 *parvum*, a major global human and animal pathogen. We examined the proteome of
9 *C. parvum* oocysts at two important stages of its life cycle; the transmissive non-
10 excysted oocyst and infective sporozoite stages. The identified proteins represent
11 roughly 8% of the predicted proteome (19,20). Together with the sequencing of the
12 *C. parvum* (20) and *C. hominis* (19) genomes, our study identified several potential
13 targets both for immunotherapy and chemoteraphy. Because of the high sequence
14 similarity between the *C. hominis* and *C. parvum* genomes (19), the major pathogenic
15 *Cryptosporidium* species for humans (10,11), it is conceivable that the development of
16 an effective vaccine or a drug for one of these species would work on both of them.
17 In the *C. parvum* life cycle, there are two motile stages, the sporozoite and the
18 merozoite, which invade the host-cell membrane (Fig. 1). Sporozoites initiate the
19 infection process, slipping out of the oocyst and invading the epithelial cells of the
20 small intestine (2). It is conceivable that sporozoites as well as merozoites are more
21 susceptible to drugs and the host immune system because they are exposed in the
22 intestinal lumen, whereas the other parasite stages are intracellular.
23
24 One approach to the development of anticryptosporidial agents, has been to identify
25 sporozoite and merozoite surface antigens involved in recognition, attachment, and

Downloaded from www.mcponline.org by on November 24, 2006

1 invasion of the host epithelial cells in order to block these interactions (423). We
2 detected a number of surface and extra-cellular proteins. Glycosylinositol
3 phospholipids (GIPLs) and glycosylphosphatidylinositol (GPI) protein anchors are
4 abundant in the surface membranes of a wide variety of parasitic protozoa and are
5 increasingly recognized as important modulators of the immune function during
6 infection (12). The immunodominant *C. parvum* 17-kDa surface antigen (N223;
7 Table 2) is GPI-anchored and parasite surface GIPLs are recognized by serum
8 antibodies from infected patients (12). Furthermore, this work also confirms the
9 expression of antigenic proteins, namely Cpa135 (N64) and GP900 (N88) (Tables 1
10 and 2, respectively), which other than induce specific antibodies (40,44) can also elicit
11 a cellular Th1 response (45).
12
13 Importantly, during oocyst excystation, we identified 26 proteins which
14 demonstrated significant expression increases. Thus, future vaccine development
15 strategies could potentially select for antigens blocking the recognition and the
16 attachment to host cells. Increased ribosomal protein expression may be due to the
17 requirement of a rapid activation protein synthesis after a variable period of
18 quiescence at the oocyst stage. The regulation of the ribosome and of the associated
19 proteins has been observed in other species of the phylum Apicomplexa, e.g. *E.*
20 *tenella* (46-48). Heat shock proteins (HSPs) are probably important for stress
21 tolerance in the host environment and for the folding of newly synthesised proteins.
22 Previous studies have found that the expression of HSPs occurred during the
23 development of Apicomplexan parasites, and that HSPs are interesting targets for the
24 host immune system (49). The level of HSP70 expression in *T. gondii* has been found
25 to be higher in mouse-virulent than in mouse-avirulent strains (50). Recently, it has

Molecular & Cellular Proteomics
Downloaded from www.mcponline.org by on November 24, 2006

been reported that *E. tenella* HSP90 is essential for the invasion of the host cell and that its expression increases in sporozoites which secrete this protein (51). Interestingly, we identified seven proteins which are unique for the genus *Crytopsoridium* and/or the phylum Apicomplexa (Table 1). Five of these proteins have a signal peptide at their N-terminus indicating that are secreted proteins. These proteins probably play a specialized role in the invasion machinery of the parasite. Apart from its signal peptide, the cgd7_4280 protein (N51; Table 1) is an Apicomplexan protein which does not demonstrate identifiable traits. There are paralogs of this gene both in *C. parvum* and in *C. hominis* [hypothetical protein cgd7_1270 (N49; table 2) and hypothetical protein Chro.70152], suggesting that these proteins belong to a more extensive protein family. The cgd6_4460 (N46) protein belongs to an Apicomplexan family sharing a peculiar architecture, because they contain both a histone deacetylase (HDAC) interacting domain and an armadillo repeat (ARM) domain. Histone deacetylases are involved in higher order chromatin assembly (52) and ARM repeat proteins are involved in various processes, including intracellular signalling and cytoskeletal regulation (53).

Cpa135 has been already characterised, it is secreted by the apical complex prior to host cell invasion and Cpa135-related proteins are a distinct family among the apicomplexan (LCCL) proteins (52). Accordingly with the present data, the expression of Cpa135 increases four-fold within 30 min after excystation (52). On the whole, three proteins containing LCCL domains (including Cpa135), with peculiar domain architecture, e.g. CpCC2 (N69) and CpCC3 (N70) were detected. These proteins are composed of various adhesive motifs (ricin B, LCCL and discoidin), suggesting these proteins are involved in some adhesion process. They also have

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics ASBMB MCP

1 closely related homologs in *P. falciparum,* namely PfCCp1, PfCCp2 and PfCCp3 that
2 are orthologs to Cpa135 (also named CpCCp1), CpCCp2 and CpCC3 respectively
3 (52,54). Interestingly, our results demonstrated that these proteins are expressed
4 simultaneously in both oocysts and free sporozoites, whereas all the *P. falciparum*
5 orthologs have gametocyte stage specific expression of these genes (54). This fact
6 suggests that these proteins require a coordinate expression to exert their role.
7
8 This study is the first major proteomic investigation of its kind on any
9 *Cryptosporidium* species. We ascertained the expression of many formerly putative
10 proteins and the presence of several secreted or surface proteins in sporozoites which
11 are restricted to the Apicomplexa or to *Cryptosporidium* genus. These findings could
12 favor further studies for the identification and role of specific molecules involved in
13 the motility, host recognition and invasion. Interestingly, three Apicomplexa-specific
14 proteins and five *Cryptosporidium* sp.-specific proteins augment in excysted invasive
15 sporozoites. These eight proteins represent promising targets for developing vaccines
16 or chemotherapies that could block parasite entry into host cells. Therefore our study
17 represents a substantial step forward towards increasing our understanding of
18 *Cryptosporidium* sp. biology, the molecular mechanisms of parasite entry, as well as
19 the development of therapy to combat cryptosporidiosis.
20
21
22
23

Downloaded from www.mcponline.org by on November 24, 2006

ASBMB Molecular & Cellular Proteomics MCP

Molecular & Cellular Proteomics
ASBMB
MCP

1 **ACKNOWLEDMENTS**

2 W.J.S. was supported by The Northern Ireland Centre of Excellence in Functional
3 Genomics, with funding from the European Union (EU) Program for Peace and
4 Reconciliation, under the Technology Support for the Knowledge-Based Economy.
5 E.P. and F.T. were partially supported by the grant of the Istituto Superiore di Sanità,
6 5th National AIDS Project, contract 50F/H. We are grateful to Daniele Tonanzi for
7 their excellent oocyst purification technical support. We thank Lorenza Putignani for
8 generating transmission electron microscopy images.

Downloaded from www.mcponline.org by on November 24, 2006

REFERENCES

1. Doganci T, Araz E, Ensari A, Tanyuksel M, Doganci L (2002) Detection of *Cryptosporidium parvum* infection in childhood using various techniques. *Med Sci Monit,* **8**, 223-226.

2. Fayer R, Speer CA, Dubey JP. General biology of *Cryptosporidium*: *In:* R. Fayer (ed.), *Cryptosporidium* and cryptosporidiosis. Boca Raton: CRC Press; 1997. p. 1-42.

3. Xiao L, Fayer R, Ryan U, Upton SJ (2004) *Cryptosporidium* taxonomy: Recent advances and implications for public health. *Clin Microbiol Rev,* **17**, 72-97.

4. Blackman MJ, Bannister LH (2001) Apical organelles of Apicomplexa: biology and isolation by subcellular fractionation. *Mol Biochem Parasitol*, **117**, 11-25.

5. Smith HV, Nichols RAB, Grimason AM (2005) *Cryptosporidium* excystation and invasion: getting to the guts of the matter. *Trends Parasitol,* **21**, 133-142.

6. Sunnotel O, Lowery CJ, Moore JE, Dooley JS, Xiao L, Millar BC, Rooney PJ, Snelling WJ (2006) Under the microscope: *Cryptosporidium*. *Lett Appl Microbiol*, **42**, 7-14.

7. McCole DF, Eckmann L, Laurent F, Kagnoff MF (2000) Intestinal epithelial cell apoptosis following *Cryptosporidium parvum* infection. *Infect Immun*, **68**, 1710–1713.

8. Fayer R, Morgan U, Upton SJ (2000) Epidemiology of *Cryptosporidium*: transmission, detection, and identification. *Int J Parasitol*, **30,** 1305–1322.

9. Caccio SM (2005) Molecular epidemiology of human cryptosporidiosis. *Parassitologia*, **47**, 185-192.

10. Chen XM, O'Hara SP, Huang BQ, Splinter PL, Nelson JB, LaRusso NF (2005) Localized glucose and water influx facilitates *Cryptosporidium parvum* cellular

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics

1 invasion by means of modulation of host-cell membrane protrusion. *Proc Natl Acad*
2 *Sci USA*, **102**, 6338-6343.
3 11. McDonald V (2000) Host cell-mediated responses to infection with
4 *Cryptosporidium*. *Parasite Immunol*, **22**, 597-604.
5 12. Priest JW, Mehlert A, Moss DM, Arrowood MJ, Ferguson MA (2006)
6 Characterization of the glycosylphosphatidylinositol anchor of the immunodominant
7 *Cryptosporidium parvum* 17-kDa antigen. *Mol Biochem Parasitol*, **149**, 108-112.
8 13. MacKenzie WR, Hoxie NJ, Proctor ME, Gradus MS, Blair KA, Peterson DE,
9 Kazmierczak JJ, Addiss DG, Fox KR, Rose JB (1994) A massive outbreak in
10 Milwaukee of cryptosporidium infection transmitted through the public water supply.
11 *N Engl J Med*, **331**, 161-167.
12 14. Corso PS, Kramer MH, Blair KA, Addiss DG, Davis JP, Haddix AC (2003) Cost
13 of illness in the 1993 waterborne *Cryptosporidium* outbreak, Milwaukee, Wisconsin.
14 *Emerg Infect Dis*, **9**, 426-431.
15 15. Zhu G, LaGier MJ, Hirose S, Keithly JS (2000) *Cryptosporidium parvum*:
16 functional complementation of a parasite transcriptional coactivator CpMBF1 in
17 yeast. *Exp Parasitol*, **96**, 195-201.
18 16. Craun GF, Calderon RL, Craun MF (2005) Outbreaks associated with recreational
19 water in the United States. *Int J Environ Health Res*, **15**, 243-262.
20 17. Putignani L (2005) The unusual architecture and predicted function of the
21 mitochondrion organelle in *Cryptosporidium parvum* and *hominis* species: the strong
22 paradigm of the structure-function relationship. *Parassitologia*, **47**, 217-225.
23 18. Adkins JN, Mottaz HM, Norbeck AD, Gustin JK, Rue J, Clauss TR, Purvine SO,
24 Rodland K, Heffron F, Smith RD (2006) Analysis of the *Salmonella typhimurium*

Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006

1 proteome through environmental response towards infectious conditions. *Mol Cell* 2 *Proteomics*, **5**, 1450-1461.

3 19. Xu P, Widmer G, Wang Y, Ozaki LS, Alves JM, Serrano MG, Puiu D, Manque P, 4 Akiyoshi D, Mackey AJ, Pearson WR, Dear PH, Bankier AT, Peterson DL, 5 Abrahamsen MS, Kapur V, Tzipori S, Buck GA (2004) The genome of 6 *Cryptosporidium hominis. Nature*, **431**, 1107-1112.

7 20. Abrahamsen MS, Templeton TJ, Enomoto S, Abrahante JE, Zhu G, Lancto CA, 8 Deng M, Liu C, Widmer G, Tzipori S, Buck GA, Xu P, Bankier AT, Dear PH, 9 Konfortov BA, Spriggs HF, Iyer L, Anantharaman V, Aravind L, Kapur V (2004) 10 Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science*, 11 **304**, 441-445.

12 21. Templeton TJ, Iyer LM, Anantharaman V, Enomoto S, Abrahante JE, 13 Subramanian GM, Hoffman SL, Abrahamsen MS, Aravind L (2004) Comparative 14 analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Res,* **14**, 1686- 15 1695.

16 22. Gut J, Nelson RG (1999) *Cryptosporidium parvum*: synchronized excystation in 17 vitro and evaluation of sporozoite infectivity with a new lectin-based assay. *J* 18 *Eukaryot Microbiol*, **46**, 56S-57S.

19 23. Wu XD, Kircher RA, McVerry PH, Malinzak DA (2000) Characterization of post- 20 translationally modified recombinant protein using liquid chromatography/mass 21 spectrometry. *Dev Biol (Basel)*, **103**, 61-67.

22 24. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, 23 Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet- 24 Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in

Downloaded from www.mcponline.org by on November 24, 2006

1 *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell*
2 *Proteomics*, **3**, 1154-1169.
3 25. Kislinger T, Gramolini AO, Pan Y, Rahman K, MacLennan DH, Emili A (2005)
4 Proteome dynamics during C2C12 myoblast differentiation. *Mol Cell Proteomics*, **4**,
5 887-901.
6 26. Truong Q, Ferrari BC (2006) Quantitative and qualitative comparisons of
7 *Cryptosporidium* faecal purification procedures for the isolation of oocysts suitable
8 for proteomic analysis. *Int J Parasitol*, **36**, 811-819.
9 27. Arrowood JM, Sterling CR (1987) Isolation of *Cryptosporidium* oocysts and
10 sporozoites using discontinuous sucrose and isopycnic Percoll gradients. *J Parasitol*,
11 **73**, 314-319.
12 28. Bukhari Z, Marshall MM, Korich DG, Fricker CR, Smith HV, Rosen J, Clancy JL
13 (2000) Comparison of *Cryptosporidium parvum* viability and infectivity assays
14 following ozone treatment of oocysts. *Appl Environ Microbiol*, **66**, 2972-2980.
15 29. Mele R, Gomez Morales MA, Tosini F, Pozio E (2004) *Cryptosporidium parvum*
16 at different developmental stages modulates host cell apoptosis in vitro. *Infect Immun*,
17 **72**, 6061-6067.
18 30. Riggs MW, Cama VA, Leary HL Jr, Sterling CR (1994) Bovine antibody against
19 *Cryptosporidium parvum* elicits a circumsporozoite precipitate-like reaction and has
20 immunotherapeutic effect against persistent cryptosporidiosis in SCID mice. *Infect*
21 *Immun*, **62**, 1927-1939.
22 31. Mandal P, Chakraborty P, Sau S, Mandal NC (2006) Purification and
23 characterization of a deoxyriboendonuclease from *Mycobacterium smegmatis*. *J*
24 *Biochem Mol Biol*, **39**, 140-144.

Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006

1 32. Tantipaiboonwong P, Sinchaikul S, Sriyam S, Phutrakul S, Chen ST (2005)
2 Different techniques for urinary protein analysis of normal and lung cancer patients.
3 *Proteomics*, **5**, 1140-1149.
4 33. Ranish JA, Yi EC, Leslie DM, Purvine SO, Goodlett DR, Eng J, Aebersold R
5 (2003) The study of macromolecular complexes by quantitative proteomics. *Nat*
6 *Genet*, **33**, 349-355.
7 34. Gu Z, Kaul M, Yan B, Kridel SJ, Cui J, Strongin A, Smith JW, Liddington RC,
8 Lipton SA (2002) S-nitrosylation of matrix metalloproteinases: signaling pathway to
9 neuronal cell death. *Science*, **297**, 1186-1190.
10 35. Prozialeck WC, Fay MJ, Lamar PC, Pearson CA, Sigar I, Ramsey KH (2002)
11 *Chlamydia trachomatis* disrupts N-cadherin-dependent cell-cell junctions and
12 sequester β-catenin in human cervical epithelial cells. *Infect Immun,* **70,** 2605-2613.
13 36. Wolters DA, Washburn MP, Yates Jr 3rd (2001) An automated multidimensional
14 protein identification technology for shotgun proteomics. *Anal Chem*, **73**, 5683-5690.
15 37. Smith JA, Blanchette RA, Burnes TA, Jacobs JJ, Higgins L, Witthuhn BA, David
16 AJ, Gillman JH. (2006) Proteomic comparison of needles from blister rust-resistant
17 and susceptible *Pinus strobus* seedlings reveals upregulation of putative disease
18 resistance proteins. *Mol Plant Microbe Interact*, **19**, 150-160.
19 38. Shadforth IP, Dunkley TPJ, Lilley KS and Bessant C (2005) i-Tracker: For
20 quantitative proteomics using iTRAQ™. *BMC Genomics*, **6**, 145.
21 39. Tang WH, Halpern BR, Shilov IV, Seymour SL, Keating SP, Loboda A, Patel
22 AA, Schaeffer DA, Nuwaysir LM (2005) Discovering known and unanticipated
23 protein modifications using MS/MS database searching. *Anal Chem*, **77**, 3931-3946.

Downloaded from www.mcponline.org by on November 24, 2006

40. Barnes DA, Bonnin A, Huang JX, Gousset L, Wu J, Gut J, Doyle P, Dubremetz JF, Ward H, Petersen C (1998) A novel multi-domain mucin-like glycoprotein of *Cryptosporidium parvum* mediates invasion. *Mol Biochem Parasitol*, **96**, 93-110.

41. Gonzalez J, Bai G, Frevert U, Corey EJ, Eichinger D (1999) Proteasome-dependent cyst formation and stage-specific ubiquitin mRNA accumulation in *Entamoeba invadens*. *Eur J Biochem*, 264, 897-904.

42. Belli SI, Walker RA, Flowers SA (2005) Global protein expression analysis in apicomplexan parasites: current status. *Proteomics*, **5**, 918-924.

43. O'Connor RM, Kim K, Khan F, Ward HD (2003) Expression of *Cpgp40/15* in *Toxoplasma gondii*: a surrogate system for the study of *Cryptosporidium* glycoprotein antigens. *Infect Immun*, **71,** 6027-6034.

44. Tosini F, Caccio S, Tamburrini A, La Rosa G, Pozio E (1999) Identification and characterisation of three antigenic proteins from *Cryptosporidium parvum* sporozoites using a DNA library expressing poly-histidine tagged peptides. *Int J Parasitol*, **29**, 1925-1933.

45. Gomez Morales MA, Mele R, Ludovisi A, Bruschi F, Tosini F, Rigano R, Pozio E (2004) *Cryptosporidium parvum*-specific CD4 Th1 cells from sensitized donors responding to both fractionated and recombinant antigenic proteins. Infect Immun, **72**, 1306-1310.

46. Waters AP, Syin C, McCutchan TF (1989) Developmental regulation of stage-specific ribosome populations in *Plasmodium*. *Nature*, **342**, 438-440.

47. Li J, McConkey GA, Rogers MJ, Waters AP, McCutchan TR (1994) *Plasmodium*: the developmentally regulated ribosome. *Exp Parasitol*, **78**, 437-441.

Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics

1 48. Schaap D, Arts G, van Poppel NF, Vermeulen AN (2005) De novo ribosome
2 biosynthesis is transcriptionally regulated in *Eimeria tenella*, dependent on its life
3 cycle stage. *Mol Biochem Parasitol*, **139**, 239-248.
4 49. Sharma YD (1992) Structure and possible function of heat-shock proteins in
5 *Falciparum* malaria. *Comp Biochem Physiol B*, **102**, 437-444.
6 50. Lyons RE, Johnson AM (1995) Heat shock proteins of *Toxoplasma gondii*.
7 *Parasite Immunol*, **17**, 353-359.
8 51. Peroval M, Pery P, Labbe M (2006) The heat shock protein 90 of *Eimeria tenella*
9 is essential for invasion of host cell and schizont growth. *Int J Parasitol*, **36**, 1205-
10 1215.
11 52. Tosini F, Agnoli A, Mele R, Gomez Morales MA, Pozio E. (2004) A new
12 modular protein of *Cryptosporidium parvum*, with ricin B and LCCL domains,
13 expressed in the sporozoite invasive stage. *Mol Biochem Parasitol*, **134**, 137-147.
14 53. Nakayama J, Xiao G, Noma K, Malikzay A, Bjerling P, Ekwall K, Kobayashi R,
15 Grewal SI (2003) Alp13, an MRG family protein, is a component of fission yeast Clr6
16 histone deacetylase required for genomic integrity. EMBO J, **22**, 2776-2787.
17 54. Pradel G, Wagner C, Mejia C, Templeton TJ (2006) *Plasmodium falciparum*: Co-
18 dependent expression and co-localization of the PfCCp multi-adhesion domain
19 proteins. *Exp Parasitol*, **112**, 263-268.

Downloaded from www.mcponline.org by on November 24, 2006

1 **Table 1**: Summary of the bioinformatic analysis of the 26 proteins which showed significant expression increases during excystation.

Ribosomal				
N	**Accession number**	**Protein name**	**Homologs**	**Domains**
5	gi|66358116	40S ribosomal protein S7	LC R	none detected
17	gi|32398896	60S ribosomal protein-like, probable	LC R	none detected
31	gi|46229456	60S ribosomal protein L35A, transcript identified by EST	LC R	none detected
35	gi|67594773	hypothetical protein Chro.50226 (ribosomal protein L3), note: it is not present in the *C. parvum* annoted proteins	LC R	none detected
Undetermined function				
N	**Accession number**	**Protein name**	**Homologs**	**Domains**
44	gi|66357812	hypothetical protein cgd5_1370	*C. hominis*	not detected
46	gi|66475922	hypothetical protein cgd6_4460	Apicomplexa	HDAC interact domain and ARM domain
Heat shock				
N	**Accession number**	**Protein name**	**Homologs**	**Domains**
58	gi|46229711	HSP70, transcripts identified by EST	LC S	none detected
59	gi|8515208	HSP70	LC S	none detected
61	gi|3273568	HSP90	LC S	none detected
62	gi|66359492	HSP90	U S	none detected
Secreted				
N	**Accession number**	**Protein name**	**Homologs**	**Domains**
42	gi|66359420	hypothetical protein cgd3_3370	*C. hominis*	signal peptide
51	gi|66363232	hypothetical protein cgd7_4280	Apicomplexa	signal peptide

Downloaded from www.mcponline.org by on November 24, 2006

N	Accession number	Protein name	Homologs	Domains
52	gi|32398968	hypothetical garp protein, possible	*C. hominis*	signal peptide
64	gi|20513131	Cpa135 protein	Apicomplexa	signal peptide, ricin B, and LCCL
68	gi|46229371	protein with signal peptide and 2 *Cryptosporidium*-specific paralogs, putative secreted protein EAK90189	*C. hominis*	signal peptide
77	gi|46227968	hypothetical protein with signal peptide EAK88888	*C. hominis*	signal peptide
Metabolic enzymes				
N	**Accession number**	**Protein name**	**Homologs**	**Domains**
102	gi|10444017	lactate dehydrogenase	LC H	none detected
105	gi|46229140	glyceraldehyde 3-phosphate dehydrogenase	LC H	none detected
107	gi|66359800	acetaldehyde reductase plus alcohol dehydrogenase (AdhE) of possible bacterial origin	LC H	none detected
108	gi|46229859	phosphoglycerate kinase 1	LC H	none detected
Nucleic acids and protein synthesis and modification and replication				
N	**Accession number**	**Protein name**	**Homologs**	**Domains**
84	gi|51951320	thioredoxin peroxidase-like protein	LC H	none detected
118	gi|46228775	histone H4	LC H	none detected
119	gi|32398654	protein disulphide isomerase, probable	LC H	none detected
120	gi|32398904	Mov34/MPN/PAD-1 family proteasome regulatory subunit, probable	LC H	none detected
127	gi|46229603	EF-1-gamma (glutathione S-transferase family)	LC R	none detected
135	gi|32398975	EF-1-alpha	LC R	none detected

1 LC: largely conserved among different phyla, R: proteins involved in protein synthesis and/or associated with ribosomes, H: house keeping
2 genes or structural genes, S: heat shock proteins, LCCL *Limulus* factor C, Coch 5B2, LG11, HDAC: histone deacetylase, and ARM: armadillo
3 repeat.

Downloaded from www.mcponline.org by on November 24, 2006

1 **Table 2:** Data from of a subset of 28 proteins, from the total 303 *Cryptosporidium parvum* proteins, which were identified during this

2 investigation. Two proteins (N115 and N136) were not present in the GenBank database, thus their accession numbers refer to *C. parvum*

3 predicted proteins (prefix cgd) in CryptoDB.

Cytoskeletal				
N	**Unused ProtScore**	**Total ProtScore**	**Accession number**	**Protein name**
36	5.96	5.96	gi|20799249	actin
277	2	2.01	gi|2522417	alpha-tubulin
Undetermined function				
N	**Unused ProtScore**	**Total ProtScore**	**Accession number**	**Protein name**
49	2.01	2.06	gi|66362692	hypothetical protein cgd7_1270
Secreted				
N	**Unused ProtScore**	**Total ProtScore**	**Accession number**	**Protein name**
65	5.6	5.6	gi|46227507	predicted secreted protein, signal peptide
69	7.25	7.26	gi|46229704	CpCCp2; extracellular protein with a signal peptide, ricin, discoidin, NEC, LCCL, 2 levanase lexan-specific cysteine-rich repeat (apicA repeat)
70	2.78	2.82	gi|46227948	CpCCp3, multidomain extracellular protein with a signal peptide, and the following architecture:LH2, LCCLCCL, pentraxin, and 2xLCCL
73	2	2.01	gi|46227296	signal peptide containing protein having a NDK (nucleoside-diphosphate kinase) domain, transcript identified by EST
88	10.16	10.17	gi|4063042	GP900; mucin-like glycoprotein
109	4.02	4.02	gi|46228639	zinc in/aminopeptidase N like metalloprotease

Downloaded from www.mcponline.org by on November 24, 2006

Surface and oocyst wall				
N	**Unused ProtScore**	**Total ProtScore**	**Accession number**	**Protein name**
223	8.01	8.02	gi|30268708	Cp17 antigen precursor
224	2.48	2.49	gi|32398662	oocyst wall protein precursor, possible
225	6.19	6.98	gi|46229286	CpCOWP1, oocyst wall protein with type I and type II cysteine-rich repeats
286	6.12	6.12	gi|8671728	15 kDa glycoprotein gp15
Metabolic enzymes				
N	**Unused ProtScore**	**Total ProtScore**	**Accession Number**	**Protein name**
98	6	6	gi|66357920	enolase (2-phosphoglycerate dehydratase)
106	2	2.01	gi|66356424	glycerol-3-phosphate dehydrogenase
110	2.01	2.02	gi|66475304	Pdr17p-like protein
124	24.2	24.2	gi|66360002	pleckstrin homology (PH domain) containing protein
236	11.7	11.7	gi|46227620	fructose-1,6-bisphosphate aldolase [EC:4.1.2.13] [*Cryptosporidium parvum*]
Nucleic acids and protein synthesis and modification and replication				
N	**Unused ProtScoree**	**Total ProtScore**	**Accession Number**	**Protein name**
91	2	2.01	gi|32398899	small GTP binding protein Rab1a, probable
115	2	2	cgd2_1870	guanine nucleotide-binding protein, putative
121	2	2.02	gi|46229649	protein disulfide isomerase, signal peptide plus ER retention motif
122	2.01	2.01	gi|66358282	Pre3p/proteasome regulatory subunit beta type 6, NTN hydrolase fold
123	2.03	2.05	gi|46226970	ubiquitin-activating enzyme E1 (UBA)
131	1.52	1.53	gi|46228920	ribonucleotide reductase small subunit, duplicated adjacent gene

Downloaded from www.mcponline.org by on November 24, 2006

133	2.02	2.03	gi|66362450	transcription regulator
136	2	2	cgd7_220	GTP-binding nuclear protein ran/tc4
137	4.02	4.03	gi|46228834	Eft2p GTpase; translation elongation factor 2 (EF-2)
264	1.7	1.7	gi|46228972	SEC14 domain containing protein

Downloaded from www.mcponline.org by on November 24, 2006

1 **Figure legends.**

2 **Fig. 1 (A through C):** The structure of *Cryptosporidium parvum* and its life-cycle
3 stages.

4

5 **(A)** A transmission electron microscopy image of a non-excysted *C. parvum* oocyst,
6 which contains 4 sporozoites. Three of the sporozoites (out of four) are visible, one in
7 a longitudinal plane, and the other two in transverse planes; AP: anterior pole, OW:
8 oocyst wall, N: nucleus, PP: posterior pole, S: scissure, and SPO: sporozoite.

9

10 **(B)** Transverse section of a free excysted sporozoite showing the apical complex and
11 the organelles; AC: apical complex, DG: dense granules, M: micronemes, MT:
12 mitochondrion, N: nucleus, R: rhoptry, and RB: refractile body.

13

14 **(C)** The stages of *Cryptosporidium* spp. life-cycle (5). During this investigation
15 oocysts were analysed at both non-excysted (x) and (excysted) (y) stages of this cycle.
16 Upon ingestion by the host, sporozoites are released and adhere directly to the
17 intestinal epithelial cells of the host. Cell invasion by sporozoite is followed by
18 intracellular development to trophozoite. Trophozoite undergo schizogony to form
19 schizonts. Asexual replication occurs by re-infection of merozoites, released by type I
20 schizont. Development of type II from type I schizont is the initial step of the asexual
21 reproductive cycle. Type II merozoites are released and re-infect neighbouring cells
22 were they develop into microgametocytes (male) or macrogametocytes (female). The
23 macrogametocyte is fertilised by released microgametes and matures into a zygote,
24 which undergoes further development into an oocyst. Two types of oocysts are

Downloaded from www.mcponline.org by on November 24, 2006

Molecular & Cellular Proteomics

1 released: (I) thick-walled oocysts, which are excreted in the feaces, or, (II) thin walled

2 oocysts for endogenous re-infection (auto-infection).

3

4 **Fig. 2:** Scheme of the overall workflow used for the proteomic analysis of

5 *Cryptosporidium parvum*.

6

7 **Fig. 3:** Example of MS/MS spectrum of peptide FPTLTGFNR (hypothetical protein

8 with signal peptide EAK88888; N77) from a protein digestion mixture prepared by

9 labelling two separate digests with tag 114 and 115, respectively and combining the

10 reaction mixtures in a 1:1 ratio. (A) Isotopic distribution of a doubly charged

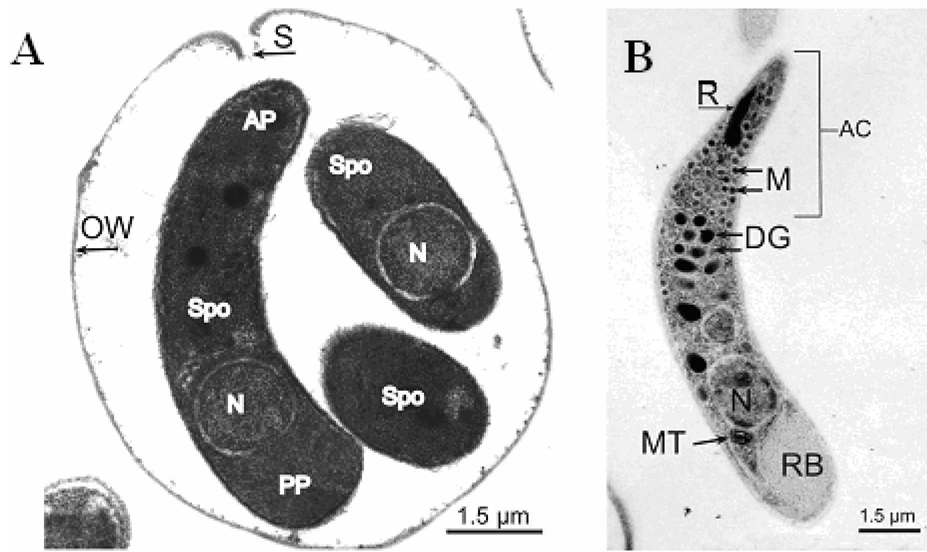
11 precursor [M+2H]2+, m/z 598.8164, (B) low mass region showing the signature ions

12 used for quantification. The peptide was labelled by isobaric tags at the N-terminus.

Molecular & Cellular Proteomics
ASBMB
MCP
Downloaded from www.mcponline.org by on November 24, 2006

1 **Figs. 1A and 1B:**

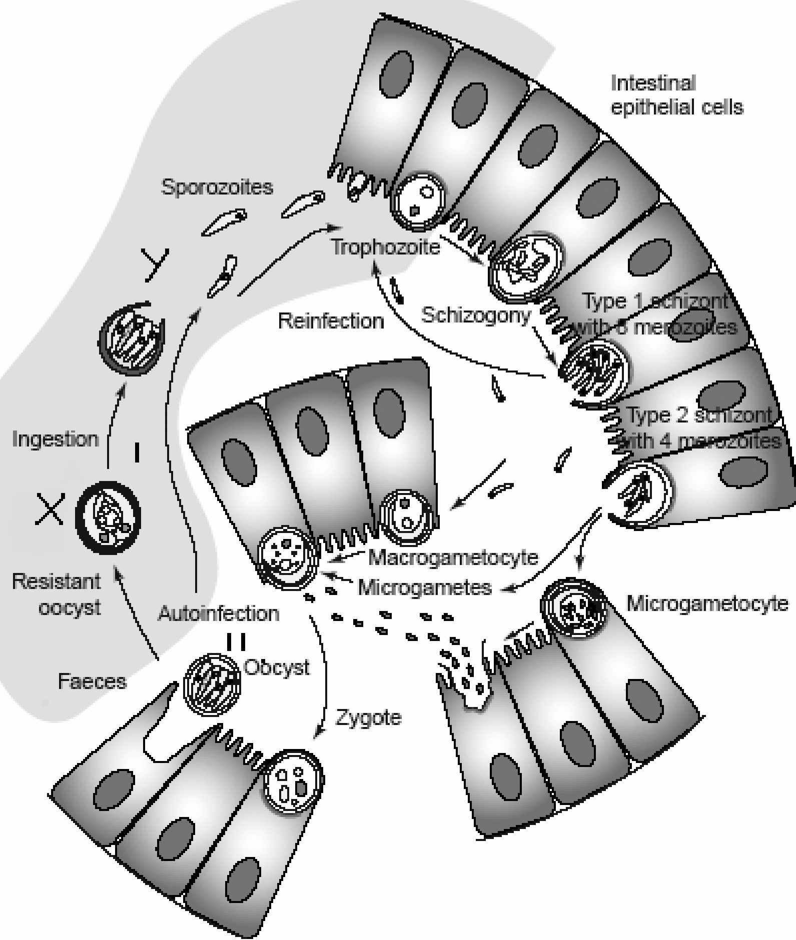
2



3

Downloaded from www.mcponline.org by on November 24, 2006

1 **Fig. 1C:**



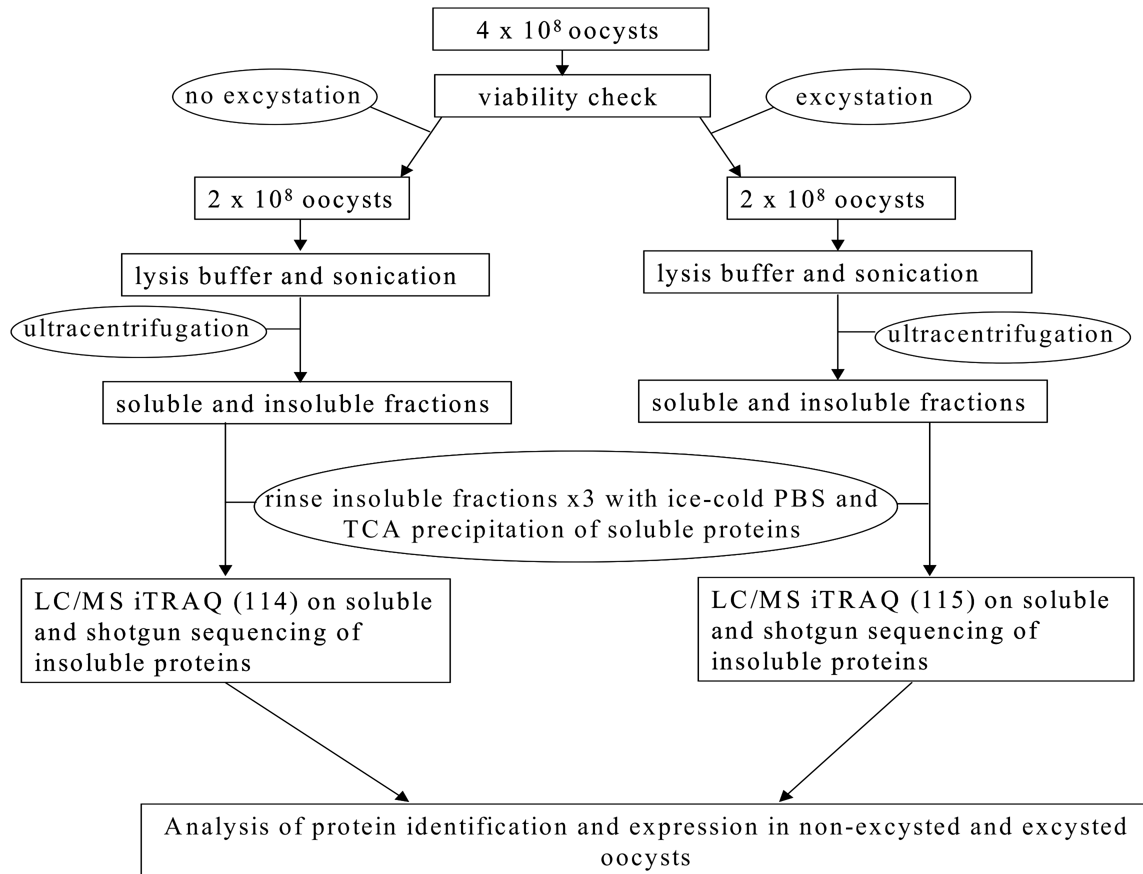
2

Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006

1 **Fig. 2:**

2

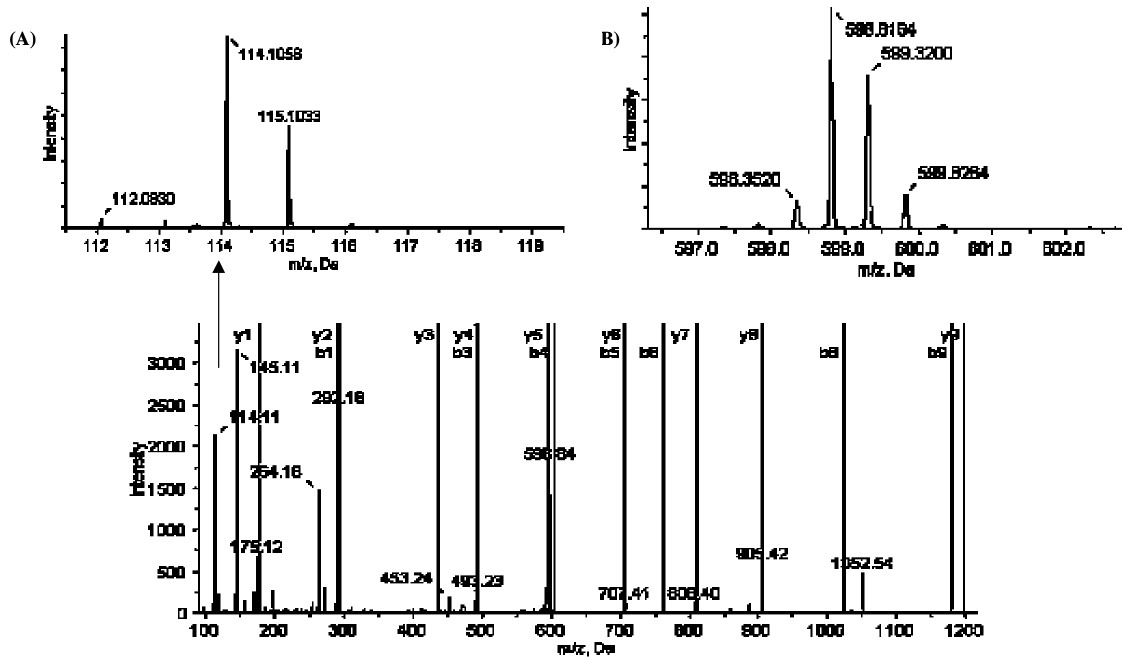


Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006

1 **Fig. 3:**

2



3

4

Molecular & Cellular Proteomics

Downloaded from www.mcponline.org by on November 24, 2006