Proteome analysis of *Cryptosporidium parvum* and *C. hominis* using two-dimentional electrophoresis, image analysis and tandem mass spectrometry

AMAM Zonaed Siddiki

Department of Pathology and Parasitology, Chittagong Veterinary and Animal Sciences University, Khulshi-4202, Chittagong, Bangladesh

Received: 28 July 2011
Accepted: 21 Jan 2012

Abstract
Until recently, *Cryptosporidium* was thought to be a single species genus. Molecular studies now show that there are at least 10 valid species of this parasite. Among them, two morphologically identical species, *C. hominis* and *C. parvum* are the most pathogenic identified to date and share 97% of identical genomes. Post-genomic analyses is therefore necessary to explore further the genetic variations among them. During this study, a comparative proteomic approach was applied to analyze the differential expression of sporozoite proteins in both *C. hominis* and *C. parvum*. Using 2-DE gels with different pH ranges (3-10 and 4-7) and automated three dimensional (3D) image analyses, a total of 20 protein spots were shown to be differentially expressed between the two species. Mass spectrometry analyses of these spots identified one hypothetical protein, however, identification of the remaining spots was unsuccessful. Further characterization of this hypothetical protein along with all differentially expressed proteins could provide crucial information in understanding the differential biology of *Cryptosporidium spp*.
Keywords: Comparative proteomics; *Cryptosporidium*; bioinformatics; 2 DE image analysis; mass spectrometry

INTRODUCTION

Comparative proteomics is based upon the comparison of protein profiles with the aim of identifying specific changes between the respective proteomes. There have been many reports of comparative proteomic analyses involving different approaches including comparing the position and intensity of spots in a 2-DE gel (Tonge *et al*., 2001; Westbrook *et al*., 2001), comparing mass spectral properties of spots or protein fractions (Davidson *et al*., 2001), or quantification of isotopically coded tryptic peptides from proteome samples (Wang and Regnier, 2001). Comparative proteomics has also been undertaken to study parasites such as *Neospora* and *Toxoplasma sp*. (Lee *et al*., 2005; Shin *et al*., 2005).

Up until the end of 2002, *C. hominis* was regarded as a genotype of *C. parvum* (Human genotype; genotype H; genotype 1). Morgan *et al*. (2002) first proposed the new species, *C. hominis* and reviewed respective molecular, structural and biological evidences. Molecular and phylogenetic studies of a number of unlinked loci have shown the difference between *C. hominis* and *C. parvum* and provided evidence for the existence of two different species. Phylogenetic analysis of *C. hominis* and *C. parvum* at the 18S rRNA locus has identified a similarity of 99.7% (Morgan *et al*., 1999 a), while at the ITS locus, it was only 82.23% (Morgan *et al*., 1999 b). Similar studies at the HSP70 locus indicated a genetic similarity of 98.50% between the two species (Sulaiman *et al*., 2000) while it was 97% at the dihydrofolate reductase-thymidylate synthase (DHFR) gene. A preliminary survey (based on 2% of whole genome) also indicated 95% of genome similarity between *C. hominis*

*Correspondence to: **AMAM Zonaed Siddiki, Ph.D.**
Tel: +88 031 659093; Fax: +88 031 659620
E-mail: zsiddiki@gmail.com

and *C. parvum* (Widmer *et al*., 2000). In another study by Tanriverdi and Widmer (2006), the differential evolution of repetitive sequences (in *C. hominis* and *C. parvum*) were analysed where significant differences were observed in frequency distributions of alleles (in 6 out of 14 loci analysed) in both species. Until now only few studies have been reported that analysed partial proteome of *Crytosporidium* spp. (Siddiki and Wastling, 2009; Sanderson *et al*., 2008, Snelling *et al*., 2007) while no report of comparative proteome analysis is available.

Use of comparative proteomics can identify the global protein changes during any invasion and thereby characterize the proteins responsible for host sensitivity and selectivity by *C. hominis* and *C. parvum*. This can be facilitated by identifying the proteins expressed in invasive sporozoite stage of *C. hominis* and *C. parvum*. Therefore, the present study was undertaken to identify the differentially expressed proteins (DEP) in these two species using the 2-DE with subsequent image analysis by ImageMaster 2D Platinum® 6.0 software in combination with tandem mass spectrometry and bioinformatics.

MATERIALS AND METHODS

Apparatus and chemicals: All equipments and reagents (if not otherwise stated) for IEF and SDS electrophoresis were purchased from Amersham Biosciences Ltd. (Bucks, UK). The iodoacetamide and Coomassie Brilliant blue G250 were purchased from Sigma Aldrich Ltd. (Poole, Dorset, UK). The first dimension SDS-PAGE equipment and 30% polyacrylamide solutions were purchased from BioRad Ltd. (Hertfordshire, UK).

Source and excystation of *Cryptosporidium* oocysts: Oocysts of *C. parvum* and *C. hominis* were propagated at Tuft University School of Veterinary Medicine (TUSVM), USA and supplied as a kind gift from Dr. Giovanni Widmer and Prof. Saul Tzipori. The *C. hominis* isolates were cultivated in gnotobiotic piglets and amplified in calves while the *C. parvum* isolates were passaged in cattle. Once collected, oocysts were concentrated by sucrose density centrifugation, washed and resuspended in Phosphate-buffered saline (PBS; pH7.2). The parasite suspension was stored at 4°C in the presence of 1000 U/ml penicillin and 1000 µg/ml streptomycin. The oocysts of both *C. parvum* and *C. hominis* were excysted using the similar method described by Campbell *et al*. (1992). Briefly, 10 µl of Sodium deoxycholate solution (1% sodium deoxycholate in Hanks minimal essential medium) and 10 µl of sodium hydrogen carbonate (2.2% sodium hydrogen carbonate in HBSS) were added to 100 µl of purified oocyst suspension. The oocysts and the excystation medium were mixed thoroughly prior to incubation at 37°C and thereafter every 15 min. After 60 min of incubation, the excystation materials were checked for excystation efficiency. It was observed that approximately 75% excystation was achieved after 1 h incubation at 37°C. The excysted sporozoites were then pelleted down by centrifugation at 10,000 ×g for 5 min. and the supernatant is discarded. This pellet containing sporozoites, empty oocysts and other unexcysted oocysts is then washed by resuspending in 500 µl of PBS and centrifuged at 10,000 ×g for 3 min. After removing the supernatant, the pellet is then collected and washed 3 times with PBS. The final pellet is then stored at -20°C for future use.

Sample preparation for 2-DE: Several replicates of 2-DE gels (Table 2) were prepared and the best gels were used for comparison of spot volume and intensity. The reproducibility of the gel spots were assessed by visual observation after image analysis while the sample preparation was identical for all of them. Similar amounts of parasite materials (~90 µg, 5×10^7 sporozoites) were used for each 2-DE gel while protein quantitation assay were performed according to section 2.3.5. For 2-DE, frozen sporozoite pellets were dissolved in 100 µl lysis buffer (8 M urea, 4% (w/v) CHAPS and 40 mM Tris hydrochloride (pH 6.8), disrupted at least 5 times by a freeze-thaw-vortexing cycle in liquid nitrogen, and the sporozoite lysate was then incubated at room temperature for 1 h for improved solubilization. The lysate was then added with 22 mM DTT (added fresh) and 0.5% (v/v) immobilised pH gradient (IPG) buffer to a final sample volume of 135 µl (for 7 cm gel). Insoluble material was removed by centrifugation at 13,000 ×g for 5 min at 4°C before loading into immobilized pH gradient (IPG) strip.

First dimension IEF: The separation of protein was performed in the first dimension using the IPGPhor® isoelectric focusing system and immobilised pH gradient (IPG) strips of different pH ranges. Once the sample was added with fresh DTT and relevant IPG buffer, appropriate volume (135 µl for 7 cm strip, 450 µl for 24 cm strip) was evenly pipetted in each slot of the

reswelling tray that will hold a strip. The IPG strips were then laid gently in each slot with the gel facing down and the pointed end of the strip near the number labels of the tray. Any bubbles that were trapped in between the strip and solution were removed. To minimize evaporation and urea crystallization, each Strip was then overlayed with 3-4 ml of DryStrip Cover Fluid® (Amersham). The lid of the reswelling tray was then placed and the gels were allowed to rehydrate overnight (10-20 h) at room temperature. During the following morning, the Ettan IPGphor Cup Loading Manifold® were placed on to the IPGPhor® unit and was levelled by using the round sprit level on the center of the manifold tray. The DryStrip Cover Fluid® (108 ml) were added evenly in the 12 manifold channels. The strips were removed from the reswelling tray and washed with distilled deionized water with subsequent removal of any fluid by using a filter paper. This was done by resting the strip vertically while only the plastic backing will touch the filter paper. The strips were then placed under the cover fluid face up in the tray with the anodic (+) end of the gel resting on the etched mark in the bottom of the manifold channel. Finally, they were adjusted in place in the centre of the manifold channel manually. Now the IEF electrode paper wicks were prepared (two for each strip) and 150 µl of distilled water was added to each wick. They were placed at both the anodal and cathodal end of the rehydrated IPG strips overlapped at the ends of the gel.

Table 1. Isoelectric focusing steps used to focus different IPG strips.

IPG	Step	Voltage	Duration (h)	Gradient type
3-11	Rehydration	-	12	-
(nonlinear)	1	300	0.30*	Step-n-hold
	2	1000	0.30	Gradient
	3	5000	1.30	Gradient
	4	5000	0.30#	Step-n-hold
4-7	Rehydration	-	12	-
(linear)	1	300	0.30*	Step-n-hold
	2	1000	0.30	Gradient
	3	5000	1.30	Gradient
	4	5000	0.30#	Step-n-hold
6-11	Rehydration	-	12	-
(linear)	1	300	0.30*	Step-n-hold
	2	1000	0.30	Gradient
	3	5000	1.20	Gradient
	4	5000	0.30#	Step-n-hold

*step that was extended upto 4 h if voltage is low. #These steps were extended as necessary until the final cumulative volt-hour reached a minimum of 6,000Vh (for 7 cm strips).

They absorb excess water, salts and proteins with pIs that lie outside the pH range of the IPG strip. Afterwards, the electrode assembly were placed in on the top of the wicks where electrode must contact the wicks and the cams are used to leave it in a closed position. The proteins were then focused to their isoelectric points according to the programmed conditions in Table 1.

Following isoelectric focusing, the proteins were reduced and bound to sodium dodecyl sulfate (SDS) by equilibrating each strip for 15 min in 10 ml of SDS equilibration buffer (50 mM Tris-HCl, 6 M urea, 30% v/v glycerol, 2% w/v SDS) containing 100 mg dithiothreitol (added fresh before use). A second equilibration step in SDS equilibration buffer containing 250 mg alpha-iodoacetamide (added fresh before use) instead of dithiothreitol was performed in order to prevent protein re-oxidation and to minimise reactions of cysteine residues.

Gel electrophoresis: Polyacrylamide gels (12%) were made using mini gel apparatus (BioRad). The gel consisted of 30% acrylamide in 1.5 M Tris-HCl (pH 8.8), 10% (w/v) SDS, 10% (w/v) ammonium persulphate (APS) and 10 µl TEMED. The SDS electrophoresis buffer was prepared by dissolving 25 mM Tris-base, 192 mM glycine and 0.1% (w/v) SDS in double distilled deionised water. Separation was performed by electrophoresis at 120 V for 2 h. Gels were then ready to stain after the romophenol blue front had completely migrated out of the SDS gel.

Gel staining and image analysis: The gels were stained with silver staining or colloidal coomassie staining technique. The silver staining was performed using the Protein Silver Staining kit (Amersham). Images of gels were acquired at 300 dots per inch using Vistascan® (version 3.5) software (UMAX Data Systems) on a Umax flatbed scanner (OD maximum 3.4) with integrated transparency adapter. The tiff format images (.tiff) were generated, then saved and converted to melanie format (.mel) before analysing with ImageMaster 2D Platinum® software (GE Healthcare). The authenticity and outline of each spot were validated by visual inspection and edited wherever necessary. The images were then compared for expression differences and significant spots were selected for in gel tryptic digestion with subsequent MS and bioinformatics analysis.

In gel tryptic digestion: The selected protein spots

Table 2. Summary of different 2-DE analyses used for comparative proteomic investigation of *C. hominis and C. parvum*. The numbers of resolved protein spots were recorded from Coomassie- and Silver-stained gels (using 7 cm IPG strips) after automatic spot detection using digital images processed by ImageMaster 2D Platinum® 6.0 software.

IPG strip pH range	No. of Sporozoites	Species	Stain	No. of spots resolved		Sample size
				Average	Range	
3-10 (nonlinear)	$5x10^7$	*C. parvum*	Coomassie	93	41-129	4
3-10 (nonlinear)	$5x10^7$	*C. hominis*	Coomassie	191	143-234	4
4-7 (linear)	$5x10^7$	*C. parvum*	Coomassie	104	90-127	3
4-7 (linear)	$5x10^7$	*C. hominis*	Coomassie	166	125-240	3

were excised from the SDS-PAGE gel and processed for MS analysis. Gel spots from coomassie stained gels were taken in separate tubes and to them 10 µl of solution containing equal amounts of 50 mM Ambic (Ammonium bicarbonate) and 50% acetonitrile was added. After incubating at 37°C for 10 min, the destain was removed with gel loading tip and discarded. The procedure was repeated several times until the gel band was fully destained. After the final wash, the liquid was removed with gel loading tip. The gels were then reduced by adding 50 µl DTT and incubating at 37°C for 30 min. After removing the DTT with gel loading tip, 50 µl iodoacetamide (IAA) was added with the gel pieces to perform alkylation. They were then incubated at 37°C for 60 min in the dark. The IAA was then removed and discarded and the gel slices were soaked with 10 µl of 100% acetonitrile and incubated at 37°C for 15 min. The solvents were then removed with a gel loading tip and further incubated at 37°C with open lids to allow the remaining solvent to evaporate. In gel digestion of proteins was then performed by incubation of the gel slice in 10 µl trypsin solution (sequencing grade trypsin diluted in 25 mM ammonium bicarbonate with final concentration of 100 µg/ml) overnight at 37°C. The reaction was stopped by adding 2 µl of 2.6M formic acid. The resulting peptides were desalted and concentrated using ZipTip® pipette tips. The resulting concentrated desalted sample was then ready for MS analysis.

Mass spectrometry and bioinformatics: LC-MS/MS data were acquired using a LTQ ion-trap mass spectrometer (Thermo-Electron, Hemel Hempstead, UK) coupled on-line to a Thermo-Finnigan surveyor HPLC system equipped with a BioBasic C_{18} reversed-phase column (100×0.18 mm 5 µm). Ionised peptides were analysed in the mass spectrometer (0-10^6 m/z, global and Ms^x) using the "triple play" mode, consisting initially of a survey (MS) spectrum from which the three most abundant ions were determined (threshold = 200-500 TIC). Collision energy was set at 35% for 30 ms. The charge state of each ion was then assigned from the C_{13} isotope envelope "zoom scan" and finally subjected to a third MS/MS scan. The LTQ was tuned using a 500 fmol/µl solution of glufibrinopeptide (m/z 785.8, $[M+2H]^{2+}$) and calibrated according to manufacturer's instructions. The resulting MS/MS spectra were merged into an mgf file submitted to MASCOT, to search CryptoDB database. Database search parameters included: fixed modification: carbamidomethyl of cysteine (C) residues; variable modification: oxidation of methionine (M) ; a peptide tolerance of up to ± 2 Da; MS/MS tolerance ± 0.8 Da; +1, +2, +3 peptide charge state; and a single missed trypsin cleavage.

RESULTS

Comparative proteomic analysis of *C. parvum* and *C. hominis* on nonlinear pH range 3-10 gels: To ensure reproducibility, identical sample preparation steps were followed for both species and multiple 2-DE analyses were performed under the same IEF conditions (Table 1). Three different approaches were used to ensure that a similar quantity of protein (~90 µg) was loaded onto each 2-DE gel. First, a similar number of oocyst materials were used for each run. Second, protein quantitation assays were performed (before loading the sample onto the IPG strip) for each protein sample derived from same number of sporozoite materials (5×10^7). Finally, the distributions of more intense protein spots over acidic and basic areas of the gels were visually checked to assess the volume of corresponding protein spots. For the final selection of differentially expressed spots, changes in protein expression were only noted if they recurred in at least 2/3 gels for each species.

In order to standardize the intensities of

Coomassie staining among spots present in several 2-DE maps, analysis was carried out by taking into account the standardized relative intensity volume of spots (or % vol, i.e. the volume of each spot over the volume of all spots in the gel). The comparative analyses of three different areas revealed unique expression of 13 spots in *C. hominis* and one spot in *C. parvum* (Table 3, Fig. 1). There were also subtle differences in the level of expression for some spots in comparing gels which might be attributed to small discrepancies in the amount of protein loaded onto each gel. Some of the differentially expressed spots were then excised manually and after digestion with trypsin, were analysed by LC-MS/MS.

Comparative proteomic analysis of *C. parvum* and *C. hominis* on linear pH 4-7 gels: To identify the differentially expressed proteins in narrow pH range, IPG strip with pH 4-7 was used to focus the sporozoite protein samples in small 7 cm 2-DE gels. The colloidal coomassie stained gels were then used for comparative image analyses and the automated spot detection was achieved by the ImageMaster 2D Platinum® 6.0 software. Protein expression was compared based on the gel images of *C. hominis* and *C. parvum* (Fig. 2) and selective areas of the gel representing similar protein spots were further analysed by creating a 3D view (data not shown) using the same image analysis tool. Based on the image analyses from three different regions of a 7 cm gel, a total of five spots were found to be uniquely expressed in *C. hominis*, while one spot was found only in *C. parvum* (Table 4, Fig. 2). There were also significant differences in the level of expression for some spots in both gels (data not shown). All the 6 spots were excised and further analysed by tandem MS and MALDI-ToF MS. Unfortunately, none of these samples were successfully identified by MS possibly due to the relatively low concentration of protein materials present in the 2-DE gels.

Identification of differentially expressed *Cryptosporidium* proteins by LC- MS/MS and MS BLAST homology searching: A total of 10 spots were selected from four different gels which showed significant levels of differential expression between the two species. For protein identification, after excision of the selected spots and tryptic digestion, all 10 spots were subjected to LC-MS/MS analysis (Table 5). Unfortunately among the 10 spots analysed, only one spot was successfully identified with a significant MASCOT score. This was a hypothetical protein of *C. hominis* (gi. 54659570) with two significant peptide hits and a MASCOT score of 98. Although LC-MS/MS analysis was unsuccessful in identifying all the digested protein spots, it did detect some peptides from each spot, the score of which was in between 17 and 25 (Table 5). The contig sequence containing the MS/MS identified peptides was used for a homology based MS BLAST search to characterize the putative proteins matching the peptides (Table 6). Hits included a number of uncharacterised and hypothetical proteins of *Cryptosporidium*.

Table 3. Summary of protein spots differentially expressed in *C. parvum* and *C. hominis* with pH 3-10 NL IPG strip. The numbers of resolved protein spots were recorded from coomassie stained 2-DE gels (Fig. 1).

Sl. No.	Gel ID	Spot ID	Expressed in *C. hominis*	Expressed in *C. parvum*
1	Ch1-Cp1	E	Yes	No
2	Ch1-Cp1	F	Yes	No
3	Ch1-Cp1	H	Yes	No
4	Ch2-Cp2	A	Yes	No
5	Ch2-Cp2	D	Yes	No
6	Ch2-Cp2	E	Yes	No
7	Ch2-Cp2	F	Yes	No
8	Ch2-Cp2	G	Yes	No
9	Ch3-Cp3	B	Yes	No
10	Ch3-Cp3	E	Yes	No
11	Ch3-Cp3	F	Yes	No
12	Ch3-Cp3	G	Yes	No
13	Ch3-Cp3	I	No	Yes
14	Ch3-Cp3	A	Yes	No

DISCUSSION

Determination of accurate protein concentration is a

Table 4. Summary of different protein spots differentially expressed in *C. parvum* and *C. hominis* with pH 4-7 L IPG strip. The numbers of resolved protein spots were recorded from Coomassie stained 2-DE gels (Fig. 2).

Sl. No.	Gel ID	Spot ID	Expressed in *C. parvum*	Expressed in *C. hominis*
1	Ch4-Cp4	E	No	Yes
2	Ch5-Cp5	F	Yes	No
3	Ch5-Cp5	I	Yes	No
4	Ch5-Cp5	K	Yes	No
5	Ch6-Cp6	A	Yes	No
6	Ch6-Cp6	B	Yes	No

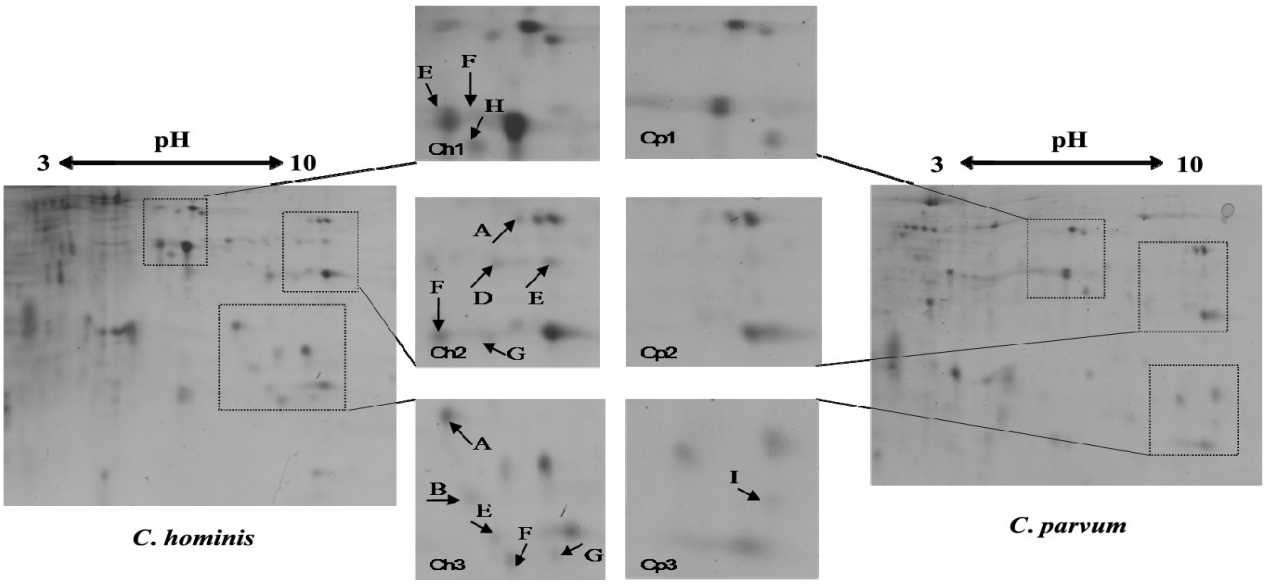

Figure 1. Colloidal Coomassie stained 2-DE gels of sporozoite proteins from *C. hominis* and *C. parvum* using broad pH range isoelectric focusing (pH 3-10 nonlinear). Different focal areas were matched and analysed by ImageMaster 2D Platinum® 6.0 software. Soluble proteins (~90 μg) extracted from $5x10^7$ sporozoites of each species were loaded on to small 7 cm gel. Only spots identified as differentially expressed are shown by arrows.

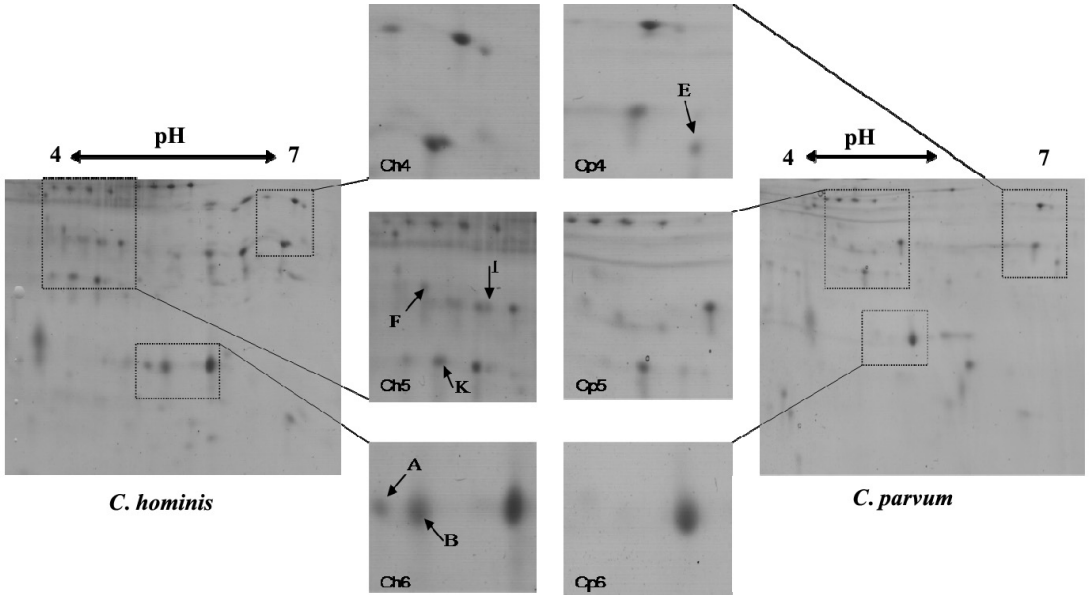

Figure 2. Colloidal Coomassie stained 2-DE gels of sporozoite proteins from *C. hominis* and *C. parvum* using narrow pH range isoelectric focusing (pH 4-7 linear). Different focal areas were matched and analysed by ImageMaster 2D Platinum® 6.0 software. Soluble proteins (~90 μg) extracted from $5x10^7$ sporozoites of each species were loaded on to small 7 cm gel. Only spots identified as differentially expressed are shown by arrows.

crucial step in sample preparation for comparative proteomic analyses. A number of protein assays are available for this purpose. However, they have variable compatibility which needs to be considered for any specific sample material to be assessed. As the protein sample might contain some sample preparation reagents, the selection of a suitable quantitation should be based on compatibility of the assay to determine the protein concentration in the presence of those reagents.

Analysis of the *Cryptosporidium* genomes showed that a highly streamlined glycolysis-based metabolism is evident in both species. Their apparent similarity in

Table 5. Mascot search result after LC-MS/MS data analysis of 10 selected spots differentially expressed in *C. hominis* and *C. parvum*. Each peptide identified by MS is listed for each spots analysed (column 2). The sequence contig of the whole genome of *Cryptosporidium* in which the peptide was detected is indicated in column 3. Column 6 shows the calculated molecular weight of the peptides while column 7 gives the MASCOT probability score for respective peptides.

Gel ID	Spot ID	CryptoDB sequence location	Peptide and possible modification	Mol. Wt.	Mass	Mascot score
Ch1	E	ChTU502_I_AAEL01000084-2-7259-7474	SSSFWSSIPR	1152.56	8325	19
Ch1	H	cgd_1540	LIGEVDLNILETTKDTFDK	2163.13	75577	23
Cp1	I	ChTU502_EAL36438	NGIGNISK	801.43	57013	25
Ch2	F	cgd4_3650	RNYNIVR	932.82	28217	18
Ch2	E	cgd7_5470	ASSSVDFTVIK	1152.60	45872	18
Ch3	A	ChTU502_EAL38141	FSDMNVFK + Oxidation (M)	1002.45	161770	98*
			SSPLLTSILVK	1156.71		
Cp4	E	ChTU502_I_AAEL01000084-2-7259-7474	SSSFWSSIPR	1152.56	8325	20
Ch6	B	cgd7_5470	ASSSVDFTVIK	1152.60	45872	20
Ch6	A	cgd7_5470	ASSSVDFTVIK	1152.60	45872	21
Ch5	L	cgd5_2500	IFRQMETGR + Oxidation (M)	1152.57	186070	17

*Only significant Mascot score with two peptide hits identifying the spot as hypothetical protein of *C. hominis* (gi.54659570).

unusual features such as absence of fully functional plastid and mitochondria are also supported by the comparative genomics (Abrahamsen *et al*., 2004; Xu *et al*., 2004). During this study, many protein spots in comparing 2-DE gels were found as identical (in terms of position and spot intensity) which supports the hypothesis that both *C. hominis* and *C. parvum* have a large number of common metabolic enzymes or proteins. Still, a number of them are unique (for any one species) which could be responsible for different functional aspects (such as signal transduction or selective response to host-mediated immunity) which needs to be determined. The complete characterization of these proteins is crucial to elucidate the actual biological variation between *C. hominis* and *C. parvum*.

In the Apicomplexa, hypothetical proteins constitute a large part of the putative proteome as predicted through genome sequencing projects. It includes 69.6%, 64.9% and 60.9% of *C. hominis*, *C. parvum* and *P. falciparum* genome respectively. During this study, one out of ten spots analysed spot was identified as hypothetical protein of *C. hominis* by LC-MS/MS analysis. Interestingly, the hypothetical protein was uniquely identified from *C. hominis* sporozoite sample while BLAST homology analyses revealed no homologous entries for this protein in *C. parvum* genome. Thus, it was possible in this study to visualise and identify specific protein spots which could be useful to be a species-specific marker for *C. hominis*. Moreover, additional study on this potential hypothetical protein could help explore its significance in host specificity and pathogenicity of *Cryptosporidium spp*.

Bioinformatic analyses using sequence similarity based MS BLAST approach also matched several other differentially expressed hypothetical proteins in *C. hominis* and *C. parvum*. While most of the hypothetical proteins are merely the result of conceptual translation of the genome (which is again predicted), the ultimate expression and their functional role could provide important information about the biology of *Cryptosporidium*.

One of the significant aims of proteomics is to understand the cellular function at the protein level under any given condition, known as functional proteomics (Adam *et al*., 2002). Functional proteomics usually focuses on those proteins that are associated with a specific condition (e.g., a diseased state of an organ, life cycle stage of an organism or any pathophysiological state). Comparative proteomic investigations during this study have identified a total of 20 differentially expressed protein (DEP) spots in both *C. hominis* and *C. parvum*. Complete characterization of these DEPs by additional approaches (e.g. raising specific antibody and localization studies, metabolomics or transfection analyses) may shed light on their functional and metabolic roles in different species of *Cryptosporidium* and on the molecular basis of host specificity.

While comparative proteomics has the potential to identify DEPs, it has to overcome several important

Table 6. Result of MS BLAST sequence similarity search using sequence containing MS-identified peptides from 10 differentially expressed spots. Contig sequence with the identified peptides (Table 5) were BLAST (protein *versus* protein) searched for homology. Column 6 gives the BLAST score probability of the top sequence matches (column 3) for each peptide identified by MS.

Gel ID	Spot ID	Accession no.	Protein ID/BLAST homology	Species	Score/ probability
Ch1	E		No significant hit		
Ch1	H	gi.66362060	a. Dynein intermediate chain	*C. parvum Iowa II*	3235/e=0
		gi.46227504	b. Dynein intermediate chain	*C. parvum*	
Cp1	I	gi.66362628	a. Hypothetical protein cgd7_920	*C. parvum Iowa II*	2074/e=0
		gi.46229860	b. Protein with 2 possible TPR domains, possible n-terminal acetyltransferase	*C. parvum*	
Ch2	F	gi.66357526	a. Hypothetical protein cgd4_3650	*C. parvum Iowa II*	1188/6e-129
		gi.46226791	b. Hypothetical protein with signal peptide, possible *Cryptosporidium* specific paralog	*C. parvum*	
Ch2	E	gi.66363448	a. Hypothetical protein cgd7_5470	*C. parvum Iowa II*	1735/e=0
		gi.46229678	b. Uncharacterized protein	*C. parvum*	
Ch3	A	gi.67624155	a. Hypothetical protein Chro.70043	*C. hominis TU502*	7257/e=0
		gi.54659570	b. Hypothetical protein Chro.70043	*C. hominis*	
Cp4	E		No significant hit		
Ch6	B	gi.66363448	a. Hypothetical protein cgd7_5470	*C. parvum Iowa II*	1735/e=0
		gi.46229678	b. Uncharacterized protein	*C. parvum*	
Ch6	A	gi.66363448	a. Hypothetical protein cgd7_5470	*C. parvum Iowa II*	1735/e=0
		gi.46229678	b. Uncharacterized protein	*C. parvum*	
Ch5	L	gi.66358018	a. Ubiquitin C-terminal hydrolase of the cysteine proteinase fold	*C. parvum Iowa II*	4093/e=0
		gi.46227023	b. Ubiquitin C-terminal hydrolase of the cysteine proteinase fold	*C. parvum*	

issues. One of the drawbacks is the purification of sample material which can give rise to complexity in data output (for example, giving many false positive hits from contaminating bacteria, mycoplasma or fungus). The purification of sporozoite materials of *Cryptosporidium sp.* for proteome analysis was a crucial part of this study where it was found difficult to achieve highly pure sample materials. In addition, all the limitations of global proteomic analyses of *Cryptosporidium* are also applicable for comparative proteomic analyses. They include limited dynamic range and limited ability to resolve both very small and very large proteins and proteins that occur in low copy number in whole cell lysates (e.g., signalling proteins and transcription factors) and hydrophobic or membrane proteins (Santoni *et al*., 2000 a,b).

The position of a protein on corresponding region of gels is sufficient for its identification under carefully standardized conditions (Govorun and Archakov, 2002). Comparison of maps of 2-DE gels obtained in experiments with previously characterized reference maps (for which protein spots have already been identified through database searching) may allow the determination of proteins expressed under different states. However, for samples without reference 2-DE maps, this could be achieved by comparing experimental gels to characterize the differentially expressed proteins. As there is no complete sporozoite 2-DE map available for either *C. parvum* or *C. hominis*, an attempt has been made to compare two proteomes based on the image analysis tools. The developments of new improved gel matching algorithms and relevant image analysis software have significantly aided this type of comparative proteomic analysis.

Two different categories of gel matching algorithms are available for gel matching in any comparative proteomic study. The first group match protein spots from different gels according to the characteristic of each spot image (eg. in programs like PDQuest, Phoretix 2D, Melanie) while the other group uses algorithms based on direct comparison of images by distribution of intensity (eg. in programs like Z3, MIR). During this study, recently developed ImageMaster 2D Platinum® 6.0 software was used which is based on the Melanie analysis package (Amersham). The software has superior spot detection ability and efficient matching algorithm that facilitates the extraction of statistically valid differences between two or more 2-D gels. The production of 3D pictures of each representative area of the gel also helps show differences in

protein expression. In addition to the software tools, the genome database is essential for successful bioinformatic analysis. Development of suitable bioinformatic programs as well as a completely annotated database is therefore an important prerequisite to compare the proteome of *C. hominis and C. parvum* as well as with other organisms.

Approximately 90 µg of proteins (from 5×10^7 sporozoites) were used for all preparative gels during this study. All the gels were run in small strips (7 cm) and use of large gels was not possible due to limitations of sample availability. However, use of small gels in this study clearly indicates that use of high sample concentration in large gel formats (eg. 24 cm) could provide more MS-derived protein sequence information than small gels. This might be the reason as to why most of the spots analysed by MS during this study were not successfully identified by MS analyses.

Basic proteins have long been difficult to successfully separate via 2-DE, and therefore are under-represented in global reference map (Cordwell *et al*., 2001). Although advancement of the IPG technology for the separation of basic proteins is reported earlier (Gorg, 1999; Gorg *et al*., 1999), the use of pH 6-11 IPG strip was not successful during this study. Different approaches involving loading of sample into the first dimension IEF step have been tried for both species of *Cryptosporidium* and subsequent SDS-PAGE and coomassie staining revealed that the proteins did not migrate efficiently (from paper bridge or sample cup) into the IPG strip. Two sample loading options have been tested namely, the paper bridge loading and cup loading method where for IPG strip with pH 6-11, the cup loading is recommended by the manufacturers (Amersham). Several factors could be attributed to the difficulty of pH 6-11 strip in this study, including higher concentration of Tris (40 mM) in sample solution, sample contaminants such as high salt from PBS and deoxycholate etc. While paper bridge loading method allows large sample volume (up to 500 µl), sample preparation became more critical for high sample load where the concentrations of the contaminants are also increased. Further study is necessary to evaluate a suitable sample preparation protocol to be used with pH 6-11 IPG strips for *Cryptosporidium* sporozoites. In addition, sample loading should also be assessed for better representation of highly basic proteins.

Reproducibility of 2-DE gels is another important issue concerning the successful comparative proteomic investigation. Preparation of sporozoites protein of *Cryptosporidium* has been simplified as much as possible to achieve reproducible 2-DE map. The source of parasite, excystation medium and subsequent washing steps (to remove deoxycholate) have been kept constant. Similarly, electrophoresis running conditions for first and second dimension separation were also standardized. The use of non-gel based separation and identification technology could be a useful alternative for this problem where pre-fractionation of samples can be achieved by liquid chromatography or MudPIT technology. The modern tools like DIGE and ICAT could be used as alternative approaches for such investigations. In addition, recently developed DIGE technology could be used to compare different experimental gels quantitatively.

In conclusion, the present study illustrates considerable differences in the sporozoite proteome between *C. hominis* and *C. parvum* while showing DEPs in comparing gels. Elucidating their possible functional significance by comparing genomic, transcriptomic, proteomic and metabolomic data is crucial to comprehensively clarify the basic molecular mechanisms in each species of *Cryptosporidium*. Ultimately, this will lead to future stage specific comparative proteomic investigations of different species of *Cryptosporidium* towards identification of potential therapeutic targets.

Acknowledgements

The author is grateful to Prof. Jonathan Wastling and Dr. Morag Nelson for useful discussion and guidance on this study. Thanks also for Drs. Andy Pitt and Richard Burchmore of Sir Henry Welcome Functional Genome Facility, University of Glasgow, Scotland, UK for technical support with the mass spectrometry and data analysis software. The work was supported by a Commonwealth Scholarship under Commonwealth Scholarship Commission, UK.

References

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

Adam GC, Sorensen EJ, Cravatt BF (2002). Chemical strategies for functional proteomics. *Mol Cell Proteomics*. 1: 781-791.

Campbell AT, Robertson LJ, Smith HV (1992). Viability of *Cryptosporidium parvum* oocysts: correlation of *in vitro* excystation with inclusion or exclusion of fluorogenic vital dyes. *Appl Environ Microbiol*. 58: 3488-3493.

Cordwell SJ, Nouwens AS, Walsh BJ (2001). Comparative proteomics of bacterial pathogens. *Proteomics* 1: 461-472.

Davidson B, Risberg B, Berner A, Nesland JM, Trope CG, Kristensen, GB, Bryne M, Goscinski M, van de PG, Florenes VA (2001). Expression of cell cycle proteins in ovarian carcinoma cells in serous effusions-biological and prognostic implications. *Gynecol Oncol*. 83: 249-256.

Gorg A (1999). IPG-Dalt of very alkaline proteins. *Methods Mol Biol*. 112: 197-209.

Gorg A, Obermaier C, Boguth G, Weiss W (1999). Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide pH gradients up to pH 12, longer separation distances and simplified procedures. *Electrophoresis* 20: 712-717.

Govorun VM, Archakov AI (2002). Proteomic technologies in modern biomedical science. *Biochem Mosc*. 67: 1109-1123.

Lee EG, Kim JH, Shin YS, Shin GW, Kim YR, Palaksha KJ, Kim DY, Yamane I, Kim YH, Kim GS, Suh MD, Jung TS (2005). Application of proteomics for comparison of proteome of *Neospora caninum* and *Toxoplasma gondii* tachyzoites. *J Chromatogr B Analyt Technol Biomed Life Sci*. 815: 305-314.

Morgan UM, Fall A, Ward LA, Hijjawi N, Sulaiman I, Fayer R, Thompson RC, Olson M, Lal A, Xiao L (2002). *Cryptosporidium hominis* n. sp. Apicomplexa: Cryptosporidiidae from Homo sapiens. *J Eukaryot Microbiol*. 49: 433-440.

Morgan UM, Xiao L, Fayer R, Graczyk TK, Lal AA, Deplazes P, Thompson RC (1999a). Phylogenetic analysis of *Cryptosporidium* isolates from captive reptiles using 18S rDNA sequence data and random amplified polymorphic DNA analysis. *J Parasitol*. 85: 525-530.

Morgan UM, Buddle JR, Armson A, Elliot A, Thompson RC (1999b). Molecular and biological characterisation of *Cryptosporidium* in pigs. *Aust Vet J*. 77: 44-47.

Sanderson SJ, Xia D, Prieto H, Yates J, Heiges M, Kissinger JC, Bromley E, Lal K, Sinden RE, Tomley F, Wastling JM (2008). Determining the protein repertoire of Cryptosporidium parvum sporozoites. *Proteomics* 8: 1398-1414.

Santoni V, Kieffer S, Desclaux D, Masson F, Rabilloud T (2000a). Membrane proteomics: use of additive main effects with multiplicative interaction model to classify plasma membrane proteins according to their solubility and electrophoretic properties. *Electrophoresis* 21: 3329-3344.

Santoni V, Molloy M, Rabilloud T (2000b). Membrane proteins and proteomics: un amour impossible? *Electrophoresis* 21: 1054-1070.

Shin YS, Shin GW, Kim YR, Lee EY, Yang HH, Palaksha KJ, Youn HJ, Kim JH, Kim DY, Marsh AE, Lakritz J, Jung TS (2005). Comparison of proteome and antigenic proteome between two *Neospora caninum* isolates. *Vet Parasitol*. 134: 41-52.

Siddiki AM, Wastling JM (2009). Charting the proteome of *Cryptosporidium parvum* sporozoites using sequence similarity based BLAST searching. *J Vet Sci*. 10: 203-210.

Snelling WJ, Lin Q, Moore JE, Millar BC, Tosini F, Pozio E, Dooley JS, Lowery CJ (2007) Proteomics analysis and protein expression during sporozoite excystation of *Cryptosporidium parvum* (Coccidia, Apicomplexa). *Mol Cell Proteomics*. 6: 346-355.

Sulaiman IM, Morgan UM, Thompson RC, Lal AA, Xiao L (2000). Phylogenetic relationships of *Cryptosporidium* parasites based on the 70-kilodalton heat shock protein HSP70 gene. *Appl Environ Microbiol*. 66: 2385-2391.

Tanriverdi S, Widmer G (2006). Differential evolution of repetitive sequences in *Cryptosporidium parvum* and *Cryptosporidium hominis*. *Infect Genet Evol*. 6: 113-122.

Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M (2001). Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* 1: 377-396.

Wang S, Regnier FE (2001). Proteomics based on selecting and quantifying cysteine containing peptides by covalent chromatography. *J Chromatogr*. A 924: 345-357.

Westbrook JA, Yan JX, Wait R, Welson SY, Dunn MJ (2001). Zooming-in on the proteome: very narrow-range immobilised pH gradients reveal more protein species and isoforms. *Electrophoresis* 22: 2865-2871.

Widmer G, Akiyoshi D, Buckholt MA, Feng X, Rich SM, Deary KM, Bowman CA, Xu P, Wang Y, Wang X, Buck GA, Tzipori S (2000). Animal propagation and genomic survey of a genotype 1 isolate of *Cryptosporidium parvum*. *Mol Biochem Parasitol*. 108: 187-197.

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