Proteomic Analysis of *Schistosoma mansoni* Cercarial Secretions

Giselle M. Knudsen‡, Katalin F. Medzihradszky§¶, Kee-Chong Lim‡¶, Elizabeth Hansell‡, and James H. McKerrow‡¶**

Schistosomiasis is a global health problem caused by several species of schistosome blood flukes. The initial stage of infection is invasion of human skin by a multicellular larva, the cercaria. We identified proteins released by cercariae when they are experimentally induced to exhibit invasive behavior. Comparison of the proteome obtained from skin lipid-induced cercariae (the natural activator), a cleaner mechanical induction procedure, and an un-induced proteomic control allowed identification of protein groups contained in cercarial acetabular gland secretion versus other sources. These included a group of proteins involved in calcium binding, calcium regulation, and calcium-activated functions; two proteins (paramyosin and SPO-1) implicated in immune evasion; and protease isoforms implicated in degradation of host skin barriers. Several other protein families, traditionally found as cytosolic proteins, appeared concentrated in secretory cells. These included proteins with chaperone activity such as HSP70, -86, and -60. Comparison of the three experimental proteomes also allowed identification of protein contaminants from the environment that were identified because of the high sensitivity of the MS/MS system used. These included proteins from the intermediate host snail in which cercariae develop, the investigator, and the laboratory environment. Identification of proteins secreted by invasive larvae provides important new information for validation of models of skin invasion and immune evasion and aids in rational development of an anti-schistosome vaccine.  

Molecular & Cellular Proteomics 4:1862–1875, 2005.

Schistosomiasis is an infectious disease caused by trematode flatworms of the genus *Schistosoma*. It is endemic in 74 developing countries with more than 80% of infected people living in sub-Saharan Africa. It is estimated that 200 million people are infected worldwide, and 200,000 new infections were reported by the World Health Organization in 2003 (1). Although treatment with the drug praziquantel is effective and inexpensive (2), frequent schistosome reinfection occurs in endemic areas and may cause irreversible damage to the liver, kidneys, or urinary tract. Thus schistosomiasis has earned a Category II disease ranking next to malaria for importance as a target tropical disease by the World Health Organization Special Programme for Research and Training in Tropical Diseases (www.who.int/tdr/). There are three major species of *Schistosoma*. *Schistosoma mansoni*, a major species in Africa and South America, was used as a model in this study.

The life cycle of *S. mansoni* is complex (3) (Fig. 1). The initiation of infection of the human host by schistosome parasites involves penetration of skin by a multicellular larva (0.1 mm) called a cercaria(ae). Cercariae have forked tails that propel them through fresh water. Depending upon their specific vertebrate host, cercariae can respond to a variety of stimuli, including motion, light, and shadow, chemical gradients, and heat. Upon contact with human skin, cercariae are stimulated by the lipid on the surface of skin to begin penetration. Initial penetration involves mechanical entry into the superficial, cornified layer of skin, which presents little barrier in the aquatic environment. However, further entry requires degradation of intercellular bridges between epidermal cells, the dermal/epidermal basement membrane, and the extracellular matrix of the dermis. Ultimately the larvae, which have now shed their tails and are called schistosomula, enter small vessels in the superficial dermis where they complete their life cycle as described above.

Proteins secreted by cercariae play key roles in facilitating skin invasion and evading the immune response of the host. Microscopic and biochemical analyses have identified three potential sources of proteins released by cercariae. First, a carbohydrate-rich surface glycocalyx is released upon entry. This glycocalyx protects the organism from osmotic shock in fresh water but is a potent activator of complement and must be jettisoned prior to entry into the bloodstream (4, 5). Second, proteolytic secretions are produced by the set of two groups of acetabular “glands,” which are in fact clusters of cells with cytoplasmic processes extending into the anterior end of the organism (6). Acetabular cells release their contents beginning at the earliest stages of skin invasion and, at least in human skin, well into the superficial dermis. The escape glands, a third potential source of secretions, appear to re-
lease their contents late in invasion as the organisms enter dermal vessels.

The only consistently effective experimental vaccine against schistosome infection is that produced by irradiation of cercariae. Identifying the protein repertoire released during the initial stage of the skin invasion is key to both understanding the pathogenesis of infection and ultimately preventing it.

We carried out proteomic analysis of the secretions of cercariae stimulated to invade and transform into schistosomula by two independent techniques (4, 7, 8). In addition to identifying the major protein components released during this initial stage of infection, our analysis provides insights into how "background" protein species can be identified with mass spectrometry technology in a complex "ecosystem" of target organism, snail host, investigator, and laboratory environment.

EXPERIMENTAL PROCEDURES

Maintenance of the Schistosome Life Cycle—S. mansoni inhabit the blood stream of infected hosts. The life cycle also requires an aquatic or amphibious snail intermediate host. A Puerto Rican strain of S. mansoni was used for all experiments. This isolate was originally obtained from Dr. Fred Lewis of the Biomedical Research Institute, Bethesda, MD, but has been maintained in our laboratory for 10 years.

Biomphalaria glabrata snails are used as intermediate hosts and are maintained in a BSL2 laboratory in accordance with all approved biosafety protocols. Cercariae are obtained using a light induction method described previously (9). Cercarial secretions were prepared according to two methods reported previously (10) in response to human skin lipid and from cultures of mechanically transformed larvae. Proteins released without stimulation were also collected. Specifically these samples were prepared as follows.

Cercariae were "shed" from their host snails in glass culture dishes overnight (9). In the first method, which recapitulates the biological stimulus, cercariae were placed in a Fisher Petri dish coated with human skin lipid. The cercariae were monitored microscopically and allowed to secrete for 1.5–2 h into water warmed to 37 °C by immersion of the Petri dish in a water bath. Cercarial bodies and tails were then removed from the secretions by centrifugation at 3000 rpm, and the secreted protein mixture was lyophilized and stored at −20 °C. This sample is referred to as skin lipid-induced "cercarial secretion."

An alternative method for inducing secretion was through mechanical tail shearing to transform the cercariae into schistosomula (11, 12). This provides a cleaner protein sample known to contain acetabular gland secretions (11). Cercariae that were shed in aquarium water were chilled and allowed to settle in a 50-ml centrifuge tube, and the aquarium water was exchanged twice with cold distilled water by decantation/resuspension followed by gentle centrifugation at 3000 rpm to collect the cercariae. The cercariae were then sheared through a small bore syringe until >90% of the larval tails were released and transferred to a 10-cm Petri dish containing 50% serum-free schistosome culture medium 169 (13) at room temperature for 1.5–2 h. Finally the plate was swirled to remove the schistosomula and tail debris by pipette. The conditioned medium was collected and pooled, while the acetalbular secretions, visible in a dissecting microscope as vesicles, were retained on the plate surface. These vesicles could then be released from the plate surface by scraping using a cell scraper (Corning Costar, Acton, MA) and rinsing with a few milliliters of 0.1% SDS solution.

First Stage Preparation by SDS-PAGE—One-dimensional SDS-PAGE was performed on the cercarial secretion samples using NOVEX Tris-glycine 4–20% acrylamide gels from Invitrogen and SeeBlue Plus 2 standards from Invitrogen to calibrate the molecular weight range. The gels were stained with silver stain according to the method by Shevchenko (14) with modifications that reduce background staining (15).

In-gel Digestion and LC-MS/MS Analysis—The SDS-PAGE gel protein bands were excised, reduced with dithiothreitol, alkylated with...
S. mansoni Cercarial Secretions

| Protein name | Accession number | Molecular mass | Gel band index no. | Coverage | Unique peptides | Mascot protein score |
|--------------|------------------|----------------|-------------------|----------|-----------------|---------------------|
| Enolase (2-phosphoglycerate dehydratase)* (Sm) | Q27877 | 46.9 | 1, 3–11, 13–15 | 60 | 25 | 1447 |
| Fructose-bisphosphate aldolase* (Sm) | P53442 | 39.6 | 3–11, 13, 14 | 64 | 30 | 1379 |
| Hemoglobin (Bg) | CAH23231 | 53.5 | 4–11, 13–19 | 62 | 23 | 1250 |
| Actin 2* (Sm) | AAC46966 | 41.7 | 3–11, 13–15 | 52 | 22 | 908 |
| Ribulose-1,5-bisphosphate carboxylase, large (Ls) | P48706 | 52.9 | 5, 6, 9, 13, 14, 18, 20, 21 | 31 | 14 | 725 |
| Glutathione S-transferase,* 28 kDa (GST 28) (Sm) | P09792 | 23.8 | 3, 4, 5, 8, 9 | 51 | 13 | 564 |
| Triose-phosphate isomerase* (TIM) (Sm) | P48501 | 28.1 | 6–9 | 39 | 9 | 483 |
| Keratin, type II cytoskeletal 1 (Hs) | P04264 | 66.0 | 3–8, 10 | 15 | 10 | 460 |
| Keratin 10, type I, cytoskeletal (Hs) | KRHU0 | 59.5 | 6–8, 14, 15 | 31 | 17 | 508 |
| GST (Sm, 218 aa) | AAB21173 | 25.4 | 8 | 46 | 8 | 493 |
| 14-3-3 protein homolog 1a* (Sm) | Q26540 | 28.3 | 6, 7, 8 | 42 | 10 | 457 |
| Vaccine-dominant antigen Sm21.7* (Sm) | A45630 | 21.6 | 1–5 | 48 | 9 | 456 |
| Serine protease beta (Bg) | AAG40234 | 26.3 | 5, 8, 9 | 42 | 8 | 446 |
| Keratin 1; cytokeratin 1; hair protein (Hs) | NP_006112 | 66.0 | 14, 15, 20 | 17 | 9 | 425 |
| Thioredoxin* (Sm) | AAL79841 | 11.9 | 2, 3, 4 | 63 | 9 | 422 |
| GAPDH (major larval surface antigen)* (P-37) (Sm) | P20287 | 36.3 | 1–8, 10 | 40 | 11 | 397 |
| Tegumental protein Sm20.8* (Sm) | AAC79131 | 20.8 | 2, 3, 4 | 49 | 11 | 396 |
| Elastase 2a (Sm) | AAM43941 | 28.6 | 1, 3, 5, 8 | 41 | 8 | 366 |
| ATP:guanidino kinase SMC74* (Sm) | P16641 | 76.2 | 5–11, 13–17 | 19 | 11 | 361 |
| Pancreatic elastase precursor (elastase 1a) (Sm) | A28942 | 28.5 | 1, 3–6, 8 | 47 | 8 | 355 |
| Developmentally regulated albumin gland gene (Bg) | AAC83410 | 65.8 | 6, 7, 11 | 7 | 4 | 240 |
| Thioredoxin peroxidase 2* (Sm) | AAJ17299 | 21.0 | 7 | 13 | 5 | 231 |
| Elastase (elastase 1b) (Sm) | AAC46967 | 29.5 | 1–6, 9 | 23 | 5 | 296 |
| Fatty acid-binding protein Sm14* (Sm) | AAL15461 | 14.8 | 5 | 32 | 4 | 275 |
| Ubiquitin (Sm) | AAG49553 | 14.6 | 1, 2 | 41 | 4 | 254 |
| Phosphoenolpyruvate carboxykinase* (Sm) | AAD24794 | 70.4 | 3, 8–10, 14–16 | 14 | 10 | 252 |
| Epidermal cytokeratin 2 (Hs) | AAC83410 | 65.8 | 6, 7, 11 | 7 | 4 | 240 |
| Calpain (EC 3.4.22.17) large chain* (Sm) | A39343 | 86.9 | 6, 7 | 7 | 5 | 239 |
| Peptidyl-prolyl cis-trans isomerase* (PPIase) (Sm) | Q26565 | 17.6 | 3–7 | 13 | 4 | 231 |
| Phosphoglycerate kinase* (Sm) | P41759 | 44.5 | 2, 3, 5, 8, 9, 13 | 25 | 7 | 219 |
| Prolactin-inducd protein (Hs) | AAE51411 | 69.2 | 15, 20 | 6 | 3 | 109 |
| Homolog to phosphoglycerate mutase* (Sm) | TCR764 | 34.9 | 2–4, 8 | 17 | 5 | 210 |
| Weakly similar to lactate dehydrogenase* (Sm) | TC16735 | 35.9 | 3–5, 7–10 | 7 | 2 | 178 |
| Similar to malate dehydrogenase,* cytosolic (Sm) | TC17066 | 36.1 | 10 | 24 | 3 | 177 |
| Similar to histone H4 (Sm) | TCL4578 | 11.3 | 4, 5 | 51 | 4 | 183 |
| Probable dynein light chain (SM10) (Sm) | Q94748 | 10.4 | 1–3 | 60 | 5 | 172 |
| Albumin (Bt) | AAB86571 | 29.0 | 10, 11 | 28 | 4 | 158 |
| Similar to malate dehydrogenase,* mito (Sm) | AAU51411 | 69.2 | 15, 20 | 6 | 3 | 109 |
| Chlorophyll a/b-binding protein (Ee) | AAF26741 | 28.1 | 7 | 17 | 4 | 165 |
| Putative cys alkylpeptidase* (Sm) | AAA44142 | 56.4 | 15 | 6 | 3 | 165 |
| Ribulose-1,5-bisphosphate carboxylase, small (Ls) | Q40250 | 20.3 | 1, 2, 4–6 | 23 | 4 | 162 |
| Unknown (serpin) (Sm) | AAB86571 | 29.0 | 10, 11 | 28 | 4 | 158 |
| Ferritin-2 heavy chain (Sm) | P25319 | 20.1 | 7 | 17 | 3 | 152 |
| Homolog to calmodulin (Sm) | TCR612 | 16.8 | 6 | 19 | 2 | 108 |
| Similar to pyruvate kinase* (Sm) | TC7454 | 61.7 | 15 | 3 | 2 | 76 |
| Serine protease (Bg) | AAG40233 | 26.2 | 8 | 8 | 2 | 146 |
| SPO-1 protein* (anti-inflammatory protein 6) (Sm) | AAD26122 | 13.6 | 4–6 | 11 | 2 | 146 |
| S100 calcium-binding protein A7; psoriasis 1 (Hs) | NP_002954 | 11.4 | 4 | 12 | 2 | 121 |
| 70,000 molecular weight antigen/hsp70 homolog* (Sm) | CAA28976 | 68.0 | 7–9 | 5 | 3 | 114 |
TABLE I—continued

| Protein name                      | Accession number | Molecular mass | Gel band index no. | Coverage | Unique peptides | Mascot protein score |
|-----------------------------------|------------------|----------------|-------------------|----------|----------------|----------------------|
| Peptidyl-prolyl cis-trans isomerase B precursor<sup>a</sup> (Sm) | Q26551           | 23.2           | 7                 | 8        | 2              | 114                  |
| Similar to muscle glycogen phosphorylase<sup>a</sup> (Sm)     | TC13591          | 96.2           | 16                | 3        | 2              | 106                  |
| Calcium-binding protein (Sm)      | AAA29921         | 17.8           | 6, 7              | 7        | 1              | 95                   |
| Heat shock protein 86<sup>a</sup> (Sm) | A45529           | 50.9           | 13                | 3        | 1              | 93                   |
| Homolog to tubulin β-2 chain<sup>a</sup> (Sm)   | TC7336           | 47.7           | 6                 | 7        | 3              | 92                   |
| Similar to carbonyl reductase<sup>a</sup> (Sm) | AAC46898         | 30.6           | 8                 | 8        | 2              | 92                   |
| Keratin 10, type I, epidermal (Hs) | A31994           | 57.2           | 4, 5              | 4        | 1              | 91                   |
| Homolog to H2B histone (Sm)       | TC13606          | 13.5           | 6                 | 19       | 2              | 90                   |
| Similar to histone H3 (Sm)        | TC13658          | 15.3           | 5                 | 13       | 2              | 88                   |
| Cysteine protease inhibitor (Sm)  | AAQ16180         | 11.3           | 3, 5              | 27       | 2              | 88                   |
| ATP-diphosphohydrolase 1<sup>a</sup> (Sm) | AAP94734         | 61.3           | 15                | 3        | 1              | 81                   |
| Similar to ATP synthase β-chain mito<sup>a</sup> (Sm) | TC13604          | 57.5           | 13                | 3        | 1              | 77                   |
| Calponin homolog (Sm)             | AAB47536         | 38.3           | 5, 6              | 4        | 1              | 77                   |
| Elongation factor 1-α<sup>a</sup> (Sm) | CAA69721         | 50.9           | 7, 8              | 6        | 2              | 77                   |
| Actin-binding/filamin-like protein<sup>a</sup> (Sm) | AAR26703         | 106.0          | 6                 | 1        | 1              | 70                   |
| Ficolin (Hs)                     | BAA12120         | 34.3           | 7, 9              | 3        | 7              | 67                   |
| Cu,Zn-superoxide dismutase (Sm)   | AAC14467         | 15.7           | 6                 | 7        | 1              | 67                   |
| Lysozyme C (1,4-β-N-acetylglucosaminidase) (Hs) | NP_000230  | 16.5           | 5                 | 8        | 1              | 66                   |
| Lipocalin 1, tear prealbumin (Hs) | NP_002288        | 19.2           | 6                 | 6        | 1              | 64                   |
| Fimbrin<sup>a</sup> (Sm)          | AAA2988         | 274.2          | 16                | 2        | 1              | 63                   |
| Heat shock protein HSP60<sup>a</sup> (Sm) | AM69406         | 58.4           | 13                | 3        | 1              | 53                   |
| 6-Phosphofructokinase<sup>a</sup> (Sm) | Q27778           | 86.0           | 10                | 2        | 1              | 51                   |
| Similar to nucleoside-diphosphate kinase<sup>a</sup> (Sm) | TC11413          | 17.7           | 6                 | 14       | 2              | 55                   |

<sup>a</sup> These proteins were also identified in the tegumental subproteome reported by van Balkom et al. (37) and are presumed to originate from tegumental shedding. See “Discussion” for details.

iodoacetic acid, and then subjected to in-gel digestion<sup>1</sup> with side chain-protected porcine trypsin (Promega, Madison, WI) (16, 17). The resulting peptides were extracted and then analyzed by on-line liquid chromatography/mass spectrometry using an HPLC system consisting of a Famos autoinjector and an Eksigent nanoflow pump coupled to a quadrupole-orthogonal acceleration-time-of-flight hybrid tandem mass spectrometer, a QSTAR XL (Applied Biosystems, Foster City, CA). The reversed-phase chromatography was controlled with Eksigent software to develop a 5–50% acetonitrile gradient in 30 min using 0.1% formic acid as the ion pairing agent at a 350 nl/min flow rate. Data were acquired in information-dependent acquisition mode: 1-s mass measurements were followed by 3-s CID experiments for the multiply charged precursor ions were computer-selected and the collision conditions were adjusted to the charge state and the m/z values of the precursor ions. CID data were analyzed using Analyst QS service pack 6 software (Applied Biosystems) with the Mascot script 1.6b4 (Matrix Science, London, UK). Parameters used in the Mascot script were as follows: AutoCentroid for the TOF; 20-ppm merge distance, 10-ppm minimum width, 50% percentage height, and 100-ppm maximum width; PeakFinding for Spectrum: 0.5% default threshold, 400-gauss filter, and a gaussian filter limit of 10. Information-dependent acquisition survey scan centroid parameters were as follows: automatic charge state determination from survey scan, 50% percentage height, and 0.02-amu merge distance. MS/MS averaging of information-dependent acquisition dependents was as follows: reject spectra if less than eight peaks, 0.5 Da precursor mass tolerance for grouping, and 10 maximum and 1 minimum cycles between groups. MS/MS data centroid and threshold parameters were as follows: remove peaks with <0% of highest intensity, centroid all MS/MS data, no smoothing, 50% height percentage, and 0.05-amu merge distance.

Database searches were performed using Mascot Server version 2.0.01 (18, 19), and the MS-Tag and MS-Pattern modules of the internal Protein Prospector server version 4.11 (20) were applied to individual peptide sequences and CID data. Searches were performed first on the National Center for Biotechnology Information non-redundant (NCBiR) data bank (September 8, 2004), and the results were parsed into a working database. The Mascot Server search parameters were as follows. Only tryptic peptides were considered with one missed cleavage allowed. Variable modifications included carbamidomethylation or propionamidation (<i>i.e.</i> acrylamide addition) of Cys residues, protein N-acetylation, Met oxidation, and pyroglutamate formation from N-terminal Gln residues. Mass accuracy was within 100 ppm in MS and 0.2 Da for CID data. Peptide sequences matched to species other than schistosomes by Mascot were further BLAST searched against the expressed sequence tags available from The Institute for Genomic Research (TIGR) S. mansoni genome project in version 5.0 (www.tigr.org/tdb/tgi/smgi/). If the peptides were identical or matched to schistosome sequences, the Mascot scores were transferred. Gene Ontology (GO)<sup>2</sup> annotations were assigned based on sequence similarity searches against the GO annotated proteins in the Swiss-Prot and TrEMBL databases at European Bioinformatics Institute calculated using the GOtobl server (21).

Serine Protease Activity of the Cercarial Secretions—Activity was measured using a fluorescense end point assay detecting proteolytic
cleavage of specific substrates for cercarial elastase (also known as acetylable protease) versus background snail serine proteases: succinyl-AAPF-aminomethylcoumarin and carbobenzoxy-(Z)-FPR-aminomethylcoumarin, respectively, as described previously (22, 23). Total fluorescence was monitored for 6 h (excitation, 355 nm; emission, 460 nm) on a Molecular Devices Flex Station. Protein concentration was determined using the Bradford assay, and the activity values are reported in relative fluorescence units per second per microgram of protein.

RESULTS

The 72 proteins identified in the skin lipid-induced cercarial secretion sample (SDS-PAGE gel shown in Fig. 2) are listed in Table I. Microscopic examination of cercariae stimulated in this manner showed no significant morphologic damage, but tails were released, and acetalubular gland contents were secreted as documented by Alizarin Red staining (24) (Fig. 3). Proteins related to calcium binding and regulation were numerous in this sample. EF-hand motifs were found in calcium-regulated protein calponin (25). The presence of numerous Ca$^{2+}$-regulated proteins is consistent with reports of high (8–10 M) levels of Ca$^{2+}$ in cercarial acetalubular (secretion) glands (26) and the presence of calpain, a calcium-activated cysteine protease in acetalubular glands of the related schistosome S. japonicum (27).

The schistosome proteins in the lipid-induced sample also included many previously identified schistosome antigens such as the vaccine candidates GST (Sm28), triose-phosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase (GAPDH, Sm37), fatty acid-binding protein (Sm14), peptidy-ly prolyl cis-trans isomerase (cyclophilin, PN18), filamin (PR-52), and phosphoglycerate kinase (PL45) (28, 29). Twelve schistosome glycolytic proteins were found; those producing abundant peptides were enolase, fructose-bisphosphate aldolase, GAPDH, and phosphoenolpyruvate carboxykinase.

At the detection level of the QSTAR XL system used, it was possible to identify proteins from the laboratory environment in the lipid-induced sample. These included the snail digestive “trypstases” called serine proteases α and β, snail hemoglobin, and a developmentally regulated albumin gland gene characterized as part of the snail immune response to schistosome infection (30). Human investigator proteins included lysozyme C, a prolactin-induced protein, tear lipocalin 1, skin-related keratins, and psoriasin. Photosynthetic proteins such as ribulose-1,5-bisphosphate carboxylase/oxygenase, and chlorophyll aligned with sequences from Lactuca sativa, the lettuce used as a food source for the B. glabrata snails. These and other photosynthetic plant proteins as well as serine protease β, but not the developmentally regulated albumin gland gene protein, were found in a sample of aquarium water conditioned by uninfected snails (data not shown).

The lipid-induced cercarial secretions contained the three S. mansoni cercarial elastase (CE) gene isoforms that are known to be expressed (1a, 1b, and 2a) (22). The peptide sequences found are shown in an alignment of the CE isoforms in Fig. 4, showing 54, 42, and 57% sequence coverage, respectively. N-terminal prodomains for these enzymes are absent; the first peptides detected in the sequence are immediately adjacent to a known prodomain processing site at Leu$^{27}$ (CE1a numbering). As confirmation for these three isoforms, CID mass spectra for example isoform-unique peptides are shown in Figs. 5–7. CE peptides were found in bands 1–9 (Fig. 2, left panel). The lowest molecular weight bands contain peptides from the C terminus of the protein, whereas band 9 contained the full-length catalytic domain of the protease. The presence of CE2a in band 12, i.e. at molecular mass ~40 kDa, was also unambiguously established from a tryptic peptide, $^{119}$QTLSGFTIVMLAQVMNQLSGIR$^{142}$ (data not shown).

The tryptases of B. glabrata in this sample were sequenced with 13% sequence coverage for serine protease α but 45% sequence coverage for serine protease β, which had both termini included (23). These peptides were found in bands 7 and 8 with the exception of a single peptide in band 5. The amount of elastase relative to snail serine protease activity in the lipid-induced secretion sample was 1:40 (0.8:46.7 relative fluorescence units/s/μg using AAPF:FPR aminomethylcoumarin substrates).
To identify background proteins released without any stimulation to invade, cercariae incubated in water alone were also analyzed (Table II). As expected, there were reduced amounts of protein relative to lipid-stimulated cercariae (0.47 g/10 l by Bradford assay compared with 1.7 g/10 l). The schistosome proteins that yielded the most abundant peptides in this sample were abundant cytosolic proteins, actin, enolase, Sm20.8, thioredoxin, and triose-phosphate isomerase (31). Background laboratory environment proteins were present here as well: chlorophyll a/b-binding protein, ribulose-1,5-bisphosphate carboxylase/oxygenase, snail serine proteases and the developmentally regulated snail albumin gland gene, and bovine serum albumin.

To minimize contamination of cercaria secretion samples by investigator, snail, and laboratory environment components, cercariae were rinsed and then mechanically transformed to release their vesicular acetabular gland secretions. This “vesicular preparation” produced slightly reduced amounts of protein compared with the lipid-induced cercarial secretions (0.5 g/10 l by Bradford assay), but was enriched 40-fold in proportion to cercarial elastase activity over the snail serine protease background. The cercarial elastase peptides were found exclusively in bands 8–10 (Fig. 2, right panel), and unique peptides were again identified for the three CE isoforms 1a, 1b, and 2a. No prodomain peptides were detected, and the C-terminal peptide was detected in bands 8–10. However, in contrast to the lipid-induced sample, the N-terminal peptide starting at Ser27 appeared in band 9, whereas the Val84 peptide is the first N-terminal peptide in band 8. This suggests that the elastases are intact and active in this preparation, whereas they had been partially degraded in the lipid-induced sample. Of the 89 proteins identified in this sample (Table III), 84 were from S. mansoni.

Thirty-three proteins in the mechanically sheared cercarial preparation were not identified previously in the lipid-induced cercarial secretions sample. Among these were the motor proteins paramyosin (Sm97), myosin heavy and light chains (Sm62, Irv5), tropomyosins I and II, and myosin regulatory chain and a collection of refolding/chaperone proteins: the T-complex protein-1 components, major egg antigen P40 (a homolog to HSP20); and protein-disulfide isomerase. Furthermore peptides from the heat shock proteins were also present at significantly increased levels: HSP60 (from one to 19 peptides, cercarial secretions to vesicles preparations), HSP70 (from three to 10 peptides), and HSP86 (from one to three peptides). Also present were the 14-3-3 protein, calcium-binding protein (GenBank™ accession number A30792), and calreticulin. Increased numbers of peptides from cytoskeletal proteins - and -tubulin as well as a number of other metabolic proteins, arginase, citrate synthase, transketolase, and ATP-dependent proteins, including calcium ATPase 2, mitochondrial ADP/ATP translocator protein, and SNaK1 protein, were identified.

GO annotations for molecular function, biological process, and cellular component could be assigned for many of the schistosome-derived proteins using the program GOblet, which transfers GO annotations based on sequence similarity searches against the TrEMBL and Swiss-Prot databases (Table IV). The percentages of cercarial secretome proteins found in each GO category are compared with the percentages found for all the full-length S. mansoni proteins listed in the NCBI nr data bank. GO categories where there is an interesting enrichment of proteins in the cercarial secretions sample over all schistosome proteins are “binding function” (including the calcium-binding proteins) and “enzymatic activity” (including the hydrolytic enzymes). The percentage of chaperone activity was enriched more than twice over that generally observed in the NCBI nr data bank schistosomes. Among the biological processes, “physiological processes”
and "metabolism" seemed greatly enhanced in this sample. Finally the apparent cellular component of 60% of these proteins is intracellular as would be expected for the numerous glycolytic proteins identified.

DISCUSSION

Schistosomiasis is a global health problem affecting more than 250 million people. Infection is initiated when an infectious larval form, the cercaria, invades human skin and gains
FIG. 7. Low energy CID spectrum of tryptic peptide Ser224-Pro-Gln-Gly-Pro-Val-Leu-Gly-Val-Ser-His-Gly-Val-Thr-Leu-Ser-Asn-Arg242, a unique peptide for *S. mansoni* elastase 2a, NCBI accession number 21217531/H11549.AAM43941.1. The precursor ion was \( m/z \) 635.38(3+).

**TABLE II**

| Protein name                                      | Accession number | Molecular mass | Gel band index no. | Coverage | Unique peptides | Mascot protein score |
|---------------------------------------------------|------------------|----------------|--------------------|----------|-----------------|----------------------|
| Ribulose-1,5-bisphosphate carboxylase, large (Ls) | P48706           | 52.9           | 10                 | 27       | 12              | 688                  |
| Enolase (2-phosphoglycerate dehydratase)\(^a\) (Sm) | Q27877           | 46.9           | 9                  | 32       | 11              | 679                  |
| Developmentally regulated albumin gland gene (Bg) | AAB00448         |                |                    |          |                 |                      |
| Ribulose-1,5-bisphosphate carboxylase, small (Ls) | Q40250           | 20.3           | 6, 7               | 23       | 4               | 136                  |
| Myoglobin (Bg)                                    | AAC24318         | 16.1           | 6                  | 15       | 2               | 150                  |
| Glutathione S-transferase, 28 kDa (GST 28)\(^a\) (Sm) | P09792           | 23.8           | 8                  | 8        | 2               | 116                  |
| Hemoglobin (Bg)                                   | CAH23232         | 43.0           | 9                  | 7        | 3               | 144                  |
| Hemoglobin (Bg)                                   | CAH22321         | 53.5           | 9                  | 5        | 2               | 116                  |
| Homolog to elongation factor 1-\(\alpha\) (Sm)     | TC16831          | 50.9           | 9                  | 2        | 1               | 88                   |
| Serine protease \(\alpha\) (Bg)                   | AAC46966         | 41.7           | 9                  | 14       | 4               | 240                  |
| Hemoglobin (Bg)                                   | CAH23233         | 53.5           | 9                  | 5        | 2               | 116                  |
| Thioredoxin peroxidase 3\(^a\) (Sm)               | TC10839          | 24.9           | 8                  | 5        | 1               | 49                   |
| Triose-phosphate isomerase\(^a\) (Sm)             | TC16805          | 28.1           | 8                  | 6        | 1               | 46                   |

*As in Table I, these proteins were also identified in the tegumental subproteome reported by van Balkom *et al.* (37). In the absence of stimulation, these proteins are presumed to be released by incidental damage to cercarial or ambient shedding of tegument.

---

*S. mansoni* Cercarial Secretions
### TABLE III
Proteins released from acutabular glands

Eighty-five proteins were identified in the *S. mansoni* acutabular vesicles sample of which 82 were from schistosome and 33 proteins were not identified previously in the lipid-induced cercarial secretion sample of Table I. ER, endoplasmic reticulum; see Table I legend for other abbreviations.

| Protein name                             | Accession number | Molecular mass (kDa) | Gel band index | Coverage | Unique peptides | Mascot protein score |
|------------------------------------------|------------------|----------------------|----------------|----------|----------------|----------------------|
| Myosin heavy chain<sup>a</sup> (Sm)      | A59287           | 222.3                | 11–19          | 42       | 85             | 4230                 |
| Paramyosin<sup>a</sup> (Sm)              | P06198           | 100.4                | 11, 13–17      | 57       | 52             | 3091                 |
| Fructose-bisphosphate aldolase<sup>a</sup> (Sm) | P53442           | 39.6                 | 8, 10–14       | 66       | 23             | 1378                 |
| Actin 2 (Sm)                             | AAC6966          | 41.7                 | 9–13, 15–17    | 69       | 28             | 1355                 |
| Phosphoehnpyruvate carboxykinase (Sm)    | AAD24794         | 70.4                 | 15             | 43       | 22             | 1159                 |
| Enolase (2-phosphoglycerate dehydratase)<sup>a</sup> (Sm) | Q27877           | 46.9                 | 9–14           | 39       | 14             | 925                  |
| Heat shock protein HSP60<sup>a</sup> (Sm) | AAM69406         | 58.4                 | 14, 15, 18     | 33       | 19             | 814                  |
| GAPDH (major larval surface antigen) (P-37)<sup>a</sup> (Sm) | P20287           | 36.3                 | 8–13           | 49       | 18             | 789                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Tropomyosin 2<sup>a</sup> (TMI) (Sm)     | P42638           | 32.6                 | 11, 12         | 28       | 11             | 608                  |
| Tropomyosin 1 (TMI) (polypeptide 49) (Sm) | P42637           | 32.9                 | 11, 12         | 42       | 12             | 602                  |
| Fructose-bisphosphate aldolase<sup>a</sup> (Sm) | CA28976          | 68.0                 | 11, 14         | 17       | 10             | 596                  |
| Actin 2 (Sm)                             | AAC7383          | 47.7                 | 13             | 28       | 6              | 515                  |
| Fructose-bisphosphate aldolase<sup>a</sup> (Sm) | P41759           | 44.5                 | 12, 13         | 27       | 9              | 334                  |
| Myoglobin (Bg)                           | CAH23231         | 53.5                 | 16–18          | 19       | 8              | 462                  |
| Hemoglobin (Bg)                          | TC16844          | 36.3                 | 11, 12         | 22       | 6              | 453                  |
| GAPDH (P-37)                             | AAC7383          | 47.7                 | 13             | 28       | 6              | 515                  |
| Similar to GAPDH (major larval surface antigen) (P-37)<sup>a</sup> (Sm) | P41759           | 44.5                 | 12, 13         | 27       | 9              | 334                  |
| Phosphoglycerate kinase<sup>a</sup> (Sm) | P41759           | 44.5                 | 12, 13         | 27       | 9              | 334                  |
| Similar to Phosphoglycerate kinase<sup>a</sup> (Sm) | A45630          | 21.6                 | 3, 4, 6, 8, 9  | 41       | 8              | 412                  |
| Glutathione (Sm)                          | AAC7383          | 47.7                 | 13             | 28       | 6              | 515                  |
| Similar to Glutathione (Sm)              | AAC7383          | 47.7                 | 13             | 28       | 6              | 515                  |
| Triose-phosphate isomerase<sup>a</sup> (Sm) | P41759           | 44.5                 | 12, 13         | 27       | 9              | 334                  |
| Similar to Triose-phosphate isomerase<sup>a</sup> (Sm) | AAC7383        | 47.7                 | 13             | 28       | 6              | 515                  |
| Hemoglobin (Bg)                           | AAC7383          | 47.7                 | 13             | 28       | 6              | 515                  |
| Similar to Hemoglobin (Bg)               | AAC7383          | 47.7                 | 13             | 28       | 6              | 515                  |
| ATP:guanido kinase<sup>a</sup> (Sm)      | A45630           | 21.6                 | 3, 4, 6, 8, 9  | 41       | 8              | 412                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
| Calcium ATPase 2<sup>a</sup> (Sm)        | AAC72576         | 111.5                | 14, 17–19      | 12       | 11             | 616                  |
access to the bloodstream through superficial dermal vessels. Understanding the mode of larval invasion and, in parallel, the mechanism of immune evasion utilized by invading larvae is key to rational development of vaccines or drugs to prevent this disease.

We carried out a proteomic analysis of the secretory products of cercariae induced to invade by two well characterized experimental protocols. In the first, human skin lipid, the natural stimulant for cercarial invasion, induced invasive behavior in vitro and led to identification of a spectrum of proteins released from three potential cercarial sources: the large acetabular glands known to secrete proteases during invasion, the smaller head glands thought to play a role in late stages of vessel entry, and proteins released from turnover of the surface tegumental membrane complex.

Table I shows 72 proteins identified through LC-MS/MS analysis of lipid-induced secretions. A number of calcium-binding proteins and the cercarial elastase (also known as cercarial protease and acetabular gland protease), identified previously as acetabular gland secretions, were present. Absent from these samples were many of the tegument-bound antigens previously reported: Sm25, Sm23, Sm22, Sm15, Sm13, and Sm8 (29, 32). The only example of this type of protein found in lipid-induced secretions was the fatty acid-binding protein Sm14. Many of these proteins have transmembrane motifs, and therefore it seems likely that the contribution of the shed membrane-associated tegumental proteins to lipid-induced cercarial secretions is minimal in comparison with acetabular gland derivatives (33).

Other proteins that were abundant in this sample were soluble glycolytic proteins such as triose-phosphate isomerase, GAPDH, and phosphoglycerate kinase. There are three possibilities why these “cytosolic” proteins appear in secretions. First, they may merely represent proteins that have “leaked” from cercariae because of “artificial” damage during lipid stimulation. However, careful microscopic analysis of

| Protein name | Accession number | Molecular mass (kDa) | Coverage | Unique peptides | Mascot protein score |
|--------------|------------------|----------------------|----------|-----------------|---------------------|
| Phosphoglycerate mutasea (Sm) | TC7546 | 34.9 | 17, 18 | 8 | 2 | 141 |
| Actin-binding/filamin-like proteina (Sm) | AAC26703 | 106.0 | 16, 17 | 3 | 2 | 138 |
| 14-3-3 ε isoforma (Sm) | AAF21436 | 28.7 | 10, 11 | 14 | 3 | 138 |
| Ubiquitin (Bg) | AAG49553 | 14.6 | 2, 8 | 22 | 3 | 129 |
| Calreticulina (Sm) | AAA19024 | 43.0 | 13 | 5 | 2 | 127 |
| Similar to T-complex protein-1, β subunita (Sm) | TC13620 | 57.0 | 14 | 5 | 2 | 126 |
| Similar to citrate synthasea (Sm) | TC10655 | 47.9 | 13 | 4 | 2 | 119 |
| Protein-disulfide isomerase homologa (Sm) | CAA80520 | 54.2 | 14 | 3 | 2 | 119 |
| Similar to T-complex protein γa (Sm) | TC13671 | 51.7 | 15 | 2 | 1 | 119 |
| Similar to pyruvate kinasea (Sm) | TC7454 | 61.7 | 15 | 4 | 2 | 118 |
| Similar to HEL proteina (Sm) | TC7459 | 50.6 | 13 | 5 | 2 | 102 |
| Weakly similar to troponin T (Sm) | TC7449 | 22.2 | 3, 4 | 6 | 3 | 95 |
| Homolog to H2B histone (Sm) | TC13606 | 13.5 | 3, 6–9 | 16 | 2 | 90 |
| Similar to T-complex protein-1, γ subunita (Sm) | TC13671 | 51.7 | 11, 14, 15 | 2 | 1 | 88 |
| Similar to hypothetical Schistosoma japonicum protein (Sm) | TC17017 | 13.8 | 3, 4 | 10 | 1 | 87 |

...as in Table I, these proteins were also identified in the tegumental subproteome reported by van Balkom et al. (37) but might also represent acetabular cell cytoplasm (see “Discussion”).
**TABLE IV**

*Functional GO annotations of secreted proteins*

GO annotations for the 96 cercarial secretion (96 CS) *Schistosoma* sequences identified in this study compared with the ~650 non-fragmentary *schistosome* proteins of the NCBI nr data bank. GO annotations were assigned based on sequence similarity searches against the GO annotated proteins in the Swiss-Prot and TrEMBL databases calculated using the GOblet server at goblet.molgen.mpg.de. The three main GO categories are in boldface: Molecular function, Biological process, and Cellular component. Percentage values for GO categories that are apparently enriched in the 96 CS dataset are also highlighted in boldface.

| GO category | Name | 96 CS | Percentage 96 CS | 650 total | Percentage 650 total |
|-------------|------|-------|------------------|-----------|---------------------|
| GO:0003674  | Molecular function | 90    | 93.75            | 448       | 71.11               |
| GO:0016209  | Antioxidant activity | 2     | 2.08             | 26        | 4.13                |
| GO:0004601  | Peroxidase activity | 1     | 1.04             | 8         | 1.27                |
| GO:0005488  | Binding | 57    | 59.38            | 278       | 44.13               |
| GO:0030246  | Carbohydrate binding | 1     | 1.04             | 6         | 0.95                |
| GO:0008289  | Lipid binding | 2     | 2.08             | 15        | 2.38                |
| GO:006782   | Metal ion binding | 20    | 20.83            | 91        | 14.44               |
| GO:000166   | Nucleotide binding | 23    | 23.96            | 77        | 12.22               |
| GO:0005515  | Protein binding | 11    | 11.46            | 49        | 7.78                |
| GO:0003824  | Catalytic activity | 48    | 50.00            | 267       | 42.38               |
| GO:0016787  | Hydrolase activity | 19    | 19.79            | 90        | 14.29               |
| GO:0016853  | Isomerase activity | 4     | 4.17             | 20        | 3.17                |
| GO:0016301  | Kinase activity | 7     | 7.29             | 35        | 5.56                |
| GO:0016829  | Lyase activity | 5     | 5.21             | 9         | 1.43                |
| GO:0016491  | Oxidoreductase activity | 13    | 13.54            | 91        | 14.44               |
| GO:0016740  | Transferase activity | 10   | 10.42            | 73        | 11.59               |
| GO:0003754  | Chaperone activity | 15    | 15.63            | 39        | 6.19                |
| GO:003777   | Heat shock protein activity | 7    | 7.29             | 25        | 3.97                |
| GO:003234   | Enzyme regulator activity | 3    | 3.13             | 10        | 1.59                |
| GO:004857   | Enzyme inhibitor activity | 3    | 3.13             | 10        | 1.59                |
| GO:0019207  | Kinase regulator activity | 1    | 1.04             | 3         | 0.48                |
| GO:0005554  | Molecular function unknown | 1    | 1.04             | 4         | 0.63                |
| GO:0003774  | Motor activity | 7     | 7.29             | 29        | 4.60                |
| GO:003777   | Microtubule motor activity | 4    | 4.17             | 15        | 2.38                |
| GO:005198   | Structural molecule activity | 7    | 7.29             | 20        | 3.17                |
| GO:005200   | Structural constituent of cytoskeleton | 2    | 2.08             | 2         | 0.32                |
| GO:005212   | Structural constituent of eye lens | 1    | 1.04             | 3         | 0.48                |
| GO:003735   | Structural constituent of ribosome | 3    | 3.13             | 8         | 1.27                |
| GO:000215   | Transporter activity | 11    | 11.46            | 83        | 13.17               |
| GO:000386   | Carrier activity | 4     | 4.17             | 42        | 6.67                |
| GO:005489   | Electron transporter activity | 4    | 4.17             | 14        | 2.22                |
| GO:0015075  | Ion transporter activity | 4    | 4.17             | 34        | 5.40                |
| GO:0008150  | Biological process | 83    | 86.46            | 422       | 66.98               |
| GO:0007610  | Behavior | 2     | 2.08             | 9         | 1.43                |
| GO:0007611  | Learning and/or memory | 1    | 1.04             | 7         | 1.11                |
| GO:005795   | Regulation of behavior | 1    | 1.04             |           |                     |
| GO:0007622  | Rhythmic behavior | 1     | 1.04             | 1         | 0.16                |
| GO:0000004  | Biological_process unknown | 2    | 2.08             | 9         | 1.43                |
| GO:0009987  | Cellular process | 39    | 40.63            | 206       | 32.70               |
| GO:0007154  | Cell communication | 7     | 7.29             | 73        | 11.59               |
| GO:005875   | Cellular physiological process | 36    | 37.50            | 169       | 26.83               |
| GO:0007275  | Development | 13    | 13.54            | 76        | 12.06               |
| GO:000653   | Morphogenesis | 7     | 7.29             | 46        | 7.30                |
| GO:0007389  | Pattern specification | 1    | 1.04             | 5         | 0.79                |
| GO:000003   | Reproduction | 5     | 5.21             | 126       | 20.00               |
| GO:000782   | Physiological process | 80    | 83.33            | 330       | 52.38               |
| GO:000875   | Cellular physiological process | 36    | 37.50            | 134       | 21.27               |
| GO:000817   | Coagulation | 1     | 1.04             | 4         | 0.63                |
| GO:002592   | Homoeostasis | 3     | 3.13             | 11        | 1.71                |
| GO:0008152  | Metabolism | 60    | 62.50            | 217       | 34.44               |
| GO:000874   | Organismal physiological process | 5    | 5.21             | 30        | 4.76                |
| GO:005896   | Response to stimulus | 7    | 7.29             | 34        | 5.40                |
| GO:005789   | Regulation of biological process | 1    | 1.04             | 6         | 0.95                |
| GO:005795   | Regulation of behavior | 1    | 1.04             |           |                     |
| GO:0005575  | Cellular_component | 63    | 65.63            | 335       | 53.17               |
larvae during this collection procedure showed no significant morphologic damage, and the small group of proteins does not match the proteome of an extract of whole cercariae reported by Curwen et al. (31). The second possible source is the syncytium, or cytosolic component of the tegument, also reported to be shed as “vesicles” with tegumental membrane by developing larvae (5, 10). GST, phosphofructokinase (34), fructose-bisphosphate aldolase (35), and phosphoenolpyruvate carboxykinase (36) have all been localized to the Sm tegument. Very recently an analysis of the tegumental subproteome of schistosomes was reported (37). Of the proteins identified in this study, 95 were found in the tegument preparations as indicated in Tables I–III.

A third possible source of cytosolic proteins is the secretory acetabular glands. The acetabular glands are in fact a set of cells in the posterior portion of the cercaria head. Cell processes extend to the anterior of the larva to serve as “ducts” for passage of secretory material. Because these glands are in fact cells that release their cytoplasmic contents during secretion, at least some of the glycolytic enzymes identified may be from this cellular source. The acetabular cells are indeed a major volume of the cercaria (6).

The exact source of these cytosolic components including glycolytic and metabolic enzymes is a key issue for laboratories working on development of the subunit schistosome vaccine. Curwen et al. (31) recently reported a proteomic analysis of soluble sonicates from several stages of S. mansoni. The authors identified the 40 most abundant soluble proteins across the schistosome life cycle and reached the important conclusion that these primarily represent cytosolic enzymes, which appear to vary little with transition from stage to stage. They make the compelling argument that such proteins are less likely to be suitable vaccine components as it is unlikely for them to be “seen” by the host immune system in intact larvae or adult. Although our analysis of cercarial secretions raises the possibility that some of these glycolytic enzymes may be released from sources like acetabular cells, it is nevertheless imperative for groups working on specific vaccine candidates to confirm location within developing larvae and whether or not such proteins are accessible to an induced immune response. Taking into consideration the conclusions of Curwen et al. (31), it is instructive to note the striking differences between the GO category shown in Table IV for presumed secretion-related proteins versus all the non-fragmentary S. mansoni proteins listed in the NCBI data bank. There is an enrichment of binding function in the cercarial secretions, specifically metal ion nucleotide and protein binding categories. There are also enriched enzymatic activities including hydrolases (like the acetylalbund protease, isomerase, kinase, and lyase activities). The percentage of chaperone activity is enriched more than twice over that generally observed in the NCBI data bank of schistosome proteins. This may indicate that chaperone proteins are concentrated within secretion sources like the acetylalbund glands to facilitate or maintain folding of secreted proteins that are densely “packed” in acetylalbund gland vesicles (4).

In the absence of any stimulation, 15 proteins were identified. Two of these were snail digestive enzymes (see discussion below). The schistosome proteins identified represent a subset of those released by “shedding” of the tegument (e.g., GST), as discussed above, or merely of high abundance (enolase). Some of these proteins match those in the “soluble” sonicate analysis of Curwen et al. (31) and therefore have likely leaked from damaged organisms.

In the lipid-induced secretion analysis, an important observation was the identification of contaminating proteins from the laboratory environment due to the high sensitivity of the QSTAR XL system. This serves as a caveat for investigators to carefully characterize proteins to separate those that represent elements of the biological phenomenon being studied versus those that may invariably come from environmental contamination. The latter are more easily identified as the sensitivity of LC-MS/MS increases despite careful attention to sample preparation. In our analysis, one such group represented proteins from the snail host in which the cercariae develop. Although not schistosome proteins, these provide important biological information. First it is clear that the snail digestive enzymes (snail trypsinases) are released as cercariae emerge from the hepatopancreas adjacent to the snail digestive tract. In addition, a protein that is a component of the snail’s own defense response to developing parasites was identified (developmentally regulated albumin gland protein). The methods of analysis were so sensitive we could even detect proteins from the lettuce on which the snails feed.

The sensitivity of QSTAR XL also meant that trace amounts of protein from the investigator preparing the samples could

| GO category          | Name                        | 96 CS | Percentage | 650 total | Percentage |
|----------------------|-----------------------------|-------|------------|-----------|------------|
| GO:0005623           | Cell                        | 62    | 64.58      | 279       | 44.29      |
| GO:0042995           | Cell projection             | 1     | 1.04       | 4         | 0.63       |
| GO:0005622           | Intracellular               | 58    | 60.42      | 197       | 31.27      |
| GO:0016020           | Membrane                    | 13    | 13.54      | 92        | 14.60      |
| GO:0008372           | Cellular_component unknown  | 5     | 5.21       | 9         | 1.43       |
| GO:0005576           | Extracellular               | 2     | 2.08       | 13        | 2.06       |
| GO:0005578           | Extracellular matrix        | 1     | 1.04       | 7         | 1.11       |
| GO:0005615           | Extracellular space         | 1     | 1.04       | 4         | 0.63       |
be detected. This included keratin from sloughed epidermal cells during collection of skin lipid. Proteins were also identified from human tears, presumably from microscopic lachrymal gland droplets that exited the technician’s eye as samples were prepared and observed.

Finally proteins were identified that represent trace contamination from the laboratory environment. Principal among these was bovine serum albumin. We were able to identify this as a laboratory contaminant because, although our laboratory primarily uses BSA in preparation of tissue culture media, a second laboratory from which we obtained snails uses human serum albumin, which was the trace contaminant when those snails were used. This indicates how readily serum albumin can contaminate a laboratory environment and, with the increased sensitivity of LC-MS/MS, can contaminate protein samples as well.

To minimize environmental contamination and directly identify the protein contents of the acetabular gland secretions, a second secretion collection method was used. This involves mechanically shearing the tails off the cercariae, which has been shown to stimulate and induce invasion behavior, specifically release of acetalbul gland contents (11, 12). By this method, the number of schistosome-related proteins relative to snail proteins increased substantially. Using an internal standard of protease activity known to be released by cercariae (cercarial elastase) versus snails (snail trypetas), the change in ratio of these enzyme activities showed that the mechanical shearing method indeed gave a more direct analysis of the acetalbul gland proteins. Acetalbul secretions appeared on the plate as small vesicles, the form in which they are released from the acetalbul cells before they rupture in the host (4). An interesting observation was that the acetalbul gland protease isolated from the intact vesicles showed no autoproteolysis products but was present as an active, mature catalytic domain or proform prior to its release. N-terminal signal peptide sequences were found in all predicted sequences of the cercaral elastase species, calreticulin, SPO-1, endoplasm, protein-disulfide isomerase, GAPDH, and prohibitin using two different motif search methods, InterPro (38) and SignalP (39).

Proteomic analysis of isolated acetalbul gland vesicles also validated previous reports of specific proteins residing in these organelles by immunolocalization. Aside from the cercaral elastase, paramyosin and SPO-1 (Sm16) were also validated previous reports of specific proteins residing in these organelles by immunolocalization. Aside from the cercaral elastase, paramyosin and SPO-1 (Sm16) were also validated previous reports of specific proteins residing in these organelles by immunolocalization. Aside from the cercaral elastase, paramyosin and SPO-1 (Sm16) were also validated previous reports of specific proteins residing in these organelles by immunolocalization. Aside from the cercaral elastase, paramyosin and SPO-1 (Sm16) were also validated previous reports of specific proteins residing in these organelles by immunolocalization.

Proteomic analysis of cercaral secretions induced by two independent methods identifies a spectrum of proteins that can be mined for potential anti-schistosome vaccine components. Furthermore, identification of ‘functional’ proteins such as cercaral elastase, calcium-binding proteins, and paramyosin provides clues or validation of proposed mechanisms of host skin invasion and immune evasion. Finally sorting of environmental contaminants from schistosome proteins serves to alert investigators that the increasing sensitivity of LC-MS/MS may invariably result in identification of “environmental” contaminants. In our case, these came from the “biological ecosystem” of schistosome and snail, the investigator preparing the sample, and the laboratory environment.

Acknowledgment—We gratefully acknowledge Aenoch Lynn of the Mass Spectrometry Facility, University of California San Francisco, for Perl scripts that aided in sorting and parsing data from the Mascot analyses.

* This work was supported by the Sandler Family Supporting Foundation, by National Center for Research Resources, National Institutes of Health Grants RR001614, RR015804, and RR012961 to the University of California San Francisco (UCSF) Mass Spectrometry Facility (Director A. L. Burlingame), and by a Veterans Affairs merit award (to J. H. M.). Core support was from the UCSF Liver Center. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

REFERENCES
1. World Health Organization (2002) Strategic Direction for Research. TDR Strategic Direction: Schistosomiasis, World Health Organization, Geneva
2. Coi1, D., and Pica-Mattoccia, L. (2003) Praziquantel. Parasitol. Res. 90, Suppl. 1, S3–S9
3. He, Y. Y. (1993) Biology of Schistosoma japonicum. From cercaria penetrating into host skin to producing egg. Chin. Med. J. (Engl. Ed.) 106, 576–583
4. Fishelson, Z., Amir, P., Friend, D. S., Markovsky, M., Pettit, M., Newport, G., and McKerrow, J. H. (1992) Schistosoma mansoni: cell-specific expression and secretion of a serine protease during development of cercariae. Exp. Parasitol. 75, 87–98
5. Skelly, P. J., and Shoemaker, C. B. (2000) Induction cues for tegument formation during the transformation of Schistosoma mansoni cercariae. Int. J. Parasitol. 30, 625–631
6. Dorsey, C. H., Cousin, C. E., Lewis, F. A., and Stirewalt, M. A. (2002) Ultrastructure of the Schistosoma mansoni cercaria. Micron 33, 279–323
7. Fusco, A. C., Salafsky, B., Ellenberger, B., and Li, L. H. (1988) Schistosoma...
mansonii: correlations between mouse strain, skin eicosanoid production, and cercarial skin penetration. J. Parasitol. 74, 253–261
8. Fusco, A. C., Salafsky, B., Vanderkooi, G., and Shibuya, T. (1991) Schis-
tosoma mansoni: the role of calcium in the stimulation of cercarial proteinase release. J. Parasitol. 77, 649–657
9. Lim, K. C., Sun, E., Bahgat, M., Bucks, D., Guy, R., Hinz, R. S., Cullander, C., and McKerrow, J. H. (1999) Blockage of skin invasion by schistosome cercariae by serine protease inhibitors. Am. J. Trop. Med. Hyg. 60, 487–492
10. Stirewalt, M. A. (1974) Schistosoma mansoni: cercaria to schistosomule. Adv. Parasitol. 12, 115–182
11. Markovinsky, M., Fishelson, Z., and Arnon, R. (1988) Purification and char-
acterization of proteases secreted by transforming schistosomula of Schistosoma mansoni. Mol. Biochem. Parasitol. 30, 45–54
12. Basch, P. F. (1981) Cultivation of Schistosoma mansoni in vitro. I. Establish-
ishment of cultures from cercariae and development until pairing. J. Parasitol. 67, 179–185
13. Basch, P. F. (1991) Schistosomes. Development, Reproduction, and Host 
Relations, Oxford University Press, New York
14. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectro-
metric sequencing of proteins silver-stained polycrylamide gels. Anal. Chem. 68, 850–858
15. Mortz, E., Krogh, T. N., Vorum, H., and Gorg, A. (2001) Improved silver 
staining protocols for high sensitivity protein identification using matrix-
assisted laser desorption/ionization/time of flight analysis. Proteomics 1, 
1359–1363
16. Hellman, U., Wernstedt, C., Gonen, J., and Heldin, C. H. (1995) Improve-
ment of an ‘In-Gel’ digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. Anal. Biochem. 224, 451–455
17. Rosenfeld, J., Capdevieille, J., Guillomet, J. C., and Ferrara, P. (1992) In-gel digestion of proteins for internal sequence analysis after one- or two-
dimensional gel electrophoresis. Anal. Biochem. 203, 173–179
18. Choudhary, J. S., Blackstock, W. P., Creasy, D. M., and Cottrell, J. S. (2001) 
Matching peptide mass spectra to EST and genomic DNA databases. Trends Biotechnol. 19, S17–S22
19. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence data-
bases using mass spectrometry data. Electrophoresis 20, 3551–3567
20. Clauer, K. R., Baker, P. R., and Burlingame, A. L. (1999) Role of accurate mass measurement (<1 ppm) in protein identification strategies em-
ploying MS or MS/MS and database searching. Anal. Chem. 71, 2871–2882
21. Groth, D., Lehrrch, H., and Hennig, S. (2004) GOBlet: a platform for Gene 
Ontology annotation of anonymous sequence data. Nucleic Acids Res. 32, W313-W317
22. Salter, J. P., Choe, Y., Albrecht, H., Franklin, C., Lim, K. C., Craik, C. S., and 
McKerrow, J. H. (2002) Cercarial elastase is encoded by a functionally 
conserved gene family across multiple species of schistosomes. J. Biol. 
Chem. 277, 24618–24624
23. Salter, J. P., Lim, K. C., Hansell, E., Hsieh, I., and McKerrow, J. H. (2000) Schistosome invasion of human skin and degradation of dermal elastin 
are mediated by a single serine protease. J. Biol. Chem. 275, 38667–38673
24. Stirewalt, M. A. (ed) (1966) Skin Penetration Mechanisms of Helminths, 
Academic Press, New York, pp. 115–182
25. Gusev, N. B. (2001) Some properties of caldesmon and calponin and the participation of these proteins in regulation of smooth muscle contrac-
tion and cytoskeleton formation. Biochemistry (Mosc.) 66, 1112–1121
26. Dresden, M. H., and Edlin, E.M. (1975) Schistosoma mansoni: calcium

S. mansoni Cercarial Secretions

Molecular & Cellular Proteomics 4.12 1875