Published in final edited form as:
J Insect Physiol. 2009 March ; 55(3): 264–272. doi:10.1016/j.jinsphys.2008.12.008.

Proteomic analysis of the mosquito *Aedes aegypti* midgut brush border membrane vesicles

Alexandra Popova-Butler[a] and **Donald H. Dean**[a,b,*]

[a] Department of Biochemistry, The Ohio State University, 484 West 12th Avenue, Columbus, OH 43210, USA

[b] Department of Entomology, The Ohio State University, 484 West 12th Avenue, Columbus, OH 43210, USA

Abstract

We analyzed brush border membrane vesicle proteins from isolated midguts of the mosquito *Aedes aegypti*, by two proteomic methods: two-dimensional gel electrophoresis (isoelectric focusing and SDS-PAGE) and a shotgun two-dimensional liquid chromatographic (LS/LS) approach based on multidimensional protein identification technology (MudPIT). We were interested in the most abundant proteins of the apical brush border midgut membrane. About 400 spots were detected on 2D gels and 39 spots were cored and identified by mass spectrometry. 86 proteins were identified by MudPIT. Three proteins, arginine kinase, putative allergen and actin are shown to be the most predominant proteins in the sample. The total number of 36 proteins detected by both methods represents the most abundant proteins in the BBMV.

Keywords

Midgut proteome; Mosquito

1. Introduction

The mosquito larval midgut is the largest organ of the organism and is responsible for maintaining ion transport, amino acid, lipid and sugar absorption. The midgut consists of a single layer of columnar epithelial cells resting on continuous basal lamina or basement membrane (Billingsley and Lehane, 1996). The laminar surface of the midgut is enhanced by extensive microvillae, the brush border membrane, in which digestive enzymes, ion channels and various extra cellular matrices are located. The structure of the microvillar projections is maintained by intracellular actin filaments. Preparations of brush border membrane vesicles (BBMV) have allowed analysis and studies of insect membrane proteins and enzymes, however, most of that research has been performed on lepidopteran insects. For example a recent proteomic analysis of *Manduca sexta* BBMV proteins has been conducted (McNall and Adang, 2003). Proteomic analysis of insect BBMV is a valuable prerequisite for determination of potential receptors for insecticidal proteins such as the *Bacillus thuringiensis* Cry proteins. To date only a few studies of toxin interaction with mosquito midgut BBMV have been done (Dronina et al., 2006; Fernández et al., 2006), but no extensive proteomic analysis has been reported. Here we report a partial proteome analysis of *Aedes aegypti* midgut BBMV. This is

*Corresponding author at: Department of Biochemistry, The Ohio State University, 484 West 12th Avenue, Columbus, OH 43210, USA. Tel.: +1 614 292 8829; fax: +1 614 292 6773. E-mail address: dean.10@osu.edu (D.H. Dean).

the first step in the understanding of complete protein composition of mosquito epithelial membrane and will be helpful in an understanding of toxin–midgut interactions.

To obtain and analyze BBMV protein spectra from *A. aegypti* we used two complementary proteomics workflows: (1) two-dimensional (2D) gel separation followed by liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis and protein identification and (2) shotgun approach which is a gel-free method based on multidimensional liquid chromatography separation of complex peptide coupled to mass spectrometry (MudPIT analysis (Wolters et al., 2001)).

2. Materials and methods

2.1. Mosquito rearing

Mosquitoes *A. aegypti*, were maintained at 28 °C with 80% relative humidity under a photoperiod 14:10 light/dark hours. Adult insects were maintained with a diet consisting of cow blood and 20% sucrose solution. Larvae were maintained on 2:1 ratio of ground dog food (Purina Dog Chow) and dried torula yeast (Ohio State University Stores).

2.2. Midgut dissection

Fourth instar larvae of *A. aegypti* were chilled on ice for at least 20 min. Forth instar larvae were used for dissections because of their size. Larvae were dissected under a microscope and midguts were collected in English–Readdy buffer (English and Readdy, 1989) (50 mM sucrose, 2 mM Tris–HCl, 1 mM PMSF, pH 7.4) with PIC Complete Protease Inhibitor (Roche) in a microfuge tube. The head was removed with a scapula and the midgut removed by grasping the body at the thorax, and at the base of the gills and pulling. The peritrophic membrane and gut contents were removed from the midgut. Then, midguts were centrifuged at 8960 × *g* for 5 min at 4 °C in a microfuge, buffer was discharged and midguts stored at −80 °C.

2.3. Brush border membrane vesicles purification

For BBMV purification, about 0.3–0.4 g of frozen midguts was suspended in 2 ml English–Readdy buffer with 1-X PIC-EDTA free Complete Protease Inhibitor (Roche) and homogenized by 30–40 strokes of a motorized Potter-Elvehjem pestle at setting 15 on ice. $CaCl_2$ was added to the homogenate to a final concentration 0.01 M and kept on ice for 15 min. This homogenate was centrifuged at 2240 × *g* at 4 °C for 10 min in JA-17 rotor. The supernatant was collected and centrifuged again at 35850 × *g* for 10 min at 4 °C. The pellet was re-suspended in 1 ml English–Readdy buffer and used for further analysis. BBMV concentration (mg protein/ml) was determined by the Bradford assay. BBMV proteins were subjected to 2D-Clean-up Kit (GE Health Sciences, Piscataway, NJ) as described by the manufacturer. The BBMV protein sample was re-suspended in buffer containing 2 M thiourea, 5 M urea, 2% CHAPS, 2% SB3-10, 65 mM DTT and 0.2% Bio-lyte ampholytes (pH 3–10, BioRad). Reconstituted proteins were centrifuged with a microcentrifuge at 15,140 × *g* for 5 min to remove any insoluble material.

2.4. Two-dimensional gel electrophoresis

A total 150 μg of midgut BBMV protein was applied on 11 cm IPG strips pH 3–10 (BioRad, Hercules, CA) for overnight rehydra-tion. The IPG strips were subjected to isoelectric focusing using a Protean IEF Cell (BioRad). Focusing was performed as follows: 400 V for 20 min, 8000 V for 2.5 h and then up to 20,000 V/h. Current did not exceed 50 μA per strip.

After isoelectric focusing, IPG strips were equilibrated for 15 min in Equilibration Buffer I (EB I) (6 M urea, 2% SDS, 0.375 M Tris–HCl (pH 8.8), 20% glycerol and 2% (w/v) DTT) followed by 15 min 6 M urea in EB II (same as EB I but containing 2.5% iodoacetamide instead

of DTT). For the second dimension IPG strips were placed across precast Ready Gel (BioRad), over-laid with agarose. Electrophoresis was run under constant voltage $V = 200$ for about 1 h. Gels were fixed overnight in a solution of 50% ethanol and 10% acetic acid, stained with Bio-safe Coomassie stain (BioRad) for at least an hour and washed in water. Stained protein spots were cut out by Proteome Works Spot Cutter (BioRad) at the Plant-Microbe Genomic facility at The Ohio State University (OSU). Cores from the gel were sent for protein identification to the Mass Spectrometry and Proteome Facility at Campus Chemical Instrument Center (OSU).

2.5. Protein identification

Thirty-nine of the most intensively stained spots were cored. These were digested with sequencing grade trypsin (Promega, Madison, WI) by the Montage In-Gel Digestion Kit (Millipore, Bedford, MA) protocol. Gel cores were washed with 50% methanol, 5% acetic acid, then dehydrated with acetonitrile and reconstituted with dithiothertol (DTT) to reduce cysteines. Iodoacetamide was added to alkylate cysteine residues. In gel protein digestion was carried out with trypsin overnight at room temperature. The peptides were extracted from the gel with 50% acetonitrile and 5 % formic acid. Capillary-liquid chromatography-nanospray tandem mass spectrometry (Nano-LC/MS/MS) was performed on a Thermo Finnigan LTQ mass spectrometer (New Objective, Inc. Woburn, MA). Protein digests from each spot were applied on 5 cm, 73 μm ID ProteoPrep C-18 column (Sigma–Aldrich). Peptides were eluted directly into an LTQ system with a gradient of 2–80% acetonitrile with a flow rate of 300 nl/min.

2.6. Multidimensional liquid chromatography

A total of 140 μg of BBMV protein, prepared from dissected midguts by the same protocol as for 2D gels, was separated into two parts; one part was digested with trypsin and another with chymotrypsin. Peptides from both digestive processes were combined and subjected to two-dimensional liquid chromatography coupled with tandem mass spectrometry. To perform 2D Capillary LC, a strong cation exchange (SCX) column 10 cm 300 μm ID Poros 10S (LC Packings Sunnyvale, CA) was utilized as the first dimension in series with a reverse phase column, 5 cm, 75 μm ID ProteoPep II C18 column (New Objective, Inc. Woburn, MA) packed directly in the nanospray tip as the second dimension. Five microlitres of each sample was injected. Peptides initially not retained on the SCX column were eluted to a C-18 trapping column (LC-Packings A Dionex Co, Sunnyvale, CA), and washed with 50 mM acetic acid to desalt the peptides. The injector port was switched to inject and the peptides were eluted from the trapping column onto the C-18 column into the LTQ system for separation. Elution was performed with solvent A (50 mM aqueous acetic acid) and solvent B (acetonitrile gradient of 2–80%) using a gradient of 2–80% B over 30 min, with a flow rate of 300 nl/min. The total run time was 58 min. Ammonium acetate injections (salt plugs) were used to elute peptides stepwise from the SCX and then onto the C-18 as described above. Twenty-microlitre injections of 10, 25, 50, 100, 200, 500, 1000 mM ammonium acetate were utilized.

2.7. Data search and protein assignment

Mass spectra were transformed into data files which were used for searching *A. aegypti* database that were used for search and assignment with Mascot MS/MS Ion search algorithm. Search parameters were as follows: peptide tolerance – 2 Da, fixed modification – carbamidomethyl, variable modification – oxidation, MS/MS tolerance 0.5 Da, allowed miscleavages – 2.

The *A. aegypti* genome has been sequenced. However the *A. aegypti* database deposited in NCBI is not defined as a separate taxonomy. Therefore, the *A. aegypti* protein database available at ftp://ftp.ensembl.org was downloaded and used as the database for mass spectral searches.

The Mascot score is estimated by the probability that at 0.95 significance level, a match between a theoretical peptide from the database and the experimental spectrum is a random event. In Mascot, the score is calculated as the negative logarithm of such probability, hence the higher score indicated the lower the probability of a random match. Protein assignment was done by at least two peptide matches (Carr et al., 2004). Some peptides matched to more than one protein. In such cases, we manually verified mass spectra for presence of unique peptides for each homologous assignment. An example of such verification is given in Fig. 1. In this figure we demonstrate the alignment of two homologous aldehyde dehydrogenases, gi|108873899 and gi|108869413 and peptides which have been identified in each of those two proteins. Those proteins have 70 identity in amino acid sequences. However peptides **K.VCGETVPSDGPHLTYTR.K** (score 66 exp 3.7e–005), **K.VGNPFSQGIQHGPQIDDIQFK.K** (score 53 exp 0.00064) and **R.EMGKEGIEEYLETK.T** (score 61, expectancy 0.00013) are specific for protein gi +108873899 (Table 1) while peptides **K.IHGSTVPSDGPVMTYIR.K** (score 77, exp 3.6e–006), **K.VGDAFAEGTQQGPQVDEEQLNK.I** (score 81 exp 1.3e–006) and **R.ELGYDGIELYTETK.T** (score 85 exp 5.8e–007) are specific for aldehyde dehydrogenase gi|108869413. Similar verification has been done for each set of homologous proteins.

3. Results

3.1. Two-dimensional gel separation and identification of A. aegypti BBMV proteins

About 400 spots were detected on a 11 cm 2D gel, after 150 μg BBMV protein was applied for two-dimensional separation (Fig. 2). Approximately 10% of spots, 39 in total, were cored and identified by mass spectrometry. Gel cores were subjected to in-gel trypsin digestion and the resulting peptide mixture from each core was analyzed by LC–MS/MS spectrometry.

Protein identifications are summarized in Table 1. Total number of unidentified proteins is 69, many of which appear in multiple spots. Protein masses are shown in Table 1 with consideration of cysteine residues, alkylated with iodoacetimide, which adds 59 Da to each cysteine residue. In the case of homologous proteins such as two aldehyde dehydrogenases (spots 8 and 9), or two actins (spot 10), we manually make the protein assignment to assure that each of homologous proteins contains unique peptides as shown above (Fig. 1).

When the estimated molecular weight of the protein matches its localization on the gel we presume what this protein is in the intact form (32 proteins shown in bold font in Table 1). Otherwise, proteins are detected by their fragments.

3.2. Protein identification of A. aegypti BBMV proteins by the shot gun approach and comparison to 2D gel separation method

The shotgun method, or MudPIT (Multidimensional Protein Identification Technology) method (Wu and Yates, 2003), is a complementary approach for protein identification, wherein a BBMV sample was subjected to trypsin and chymotrypsin digestion without preliminary sample separation. By examining the mass spectra of shotgun method with Mascot search algorithm, 86 *A. aegypti* proteins were matched in the *Aedes* database. Results are shown in Table 2. We considered only mass spectra of two or more peptides. Also, all proteins identified as mosquito trypsin-like protease were excluded as potential contaminants, since trypsin was used to digest the sample and Mascot search would force matching of added trypsin to the *Aedes* protein dataset.

4. Discussion

4.1. Two-dimensional gel separation and identification of A. aegypti BBMV proteins

Of the approximate 400 spots detected by 2D gel electrophoresis we selected 39 for further analysis. These were selected since we were interested in most abundant proteins and therefore we focused on the largest spots for protein identification after visual inspection. Of the 39 spots selected, 69 proteins were identified.

Proteins were identified with Mascot search engine using the *A. aegypti* protein database in FASTA format. This is a more specific search compared to a general NCBI search. Using the *A. aegypti* database we excluded redundancies and contaminations in the search results. Probabilistic Mascot score for the proteins selected was 34, which means that there is 95% probability that protein assignment is not random. We did not consider protein identified by just one peptide as reliable and excluded them from the further consideration.

A number of proteins recovered from spots on the 2D gel were identified as fragments of the complete protein. Such misplacement of protein on the gel may be due to the following reasons:

1. Proteins were fragmented during purification and separation, and fragment of various length appear in different spots because of proteolysis or acidic hydrolysis during sample preparation.
2. Various post-translational modifications (PTM) may shift protein in both, horizontal and vertical directions. For example, phosphorylation and other PTMs which do not change protein mass significantly but change total charge and therefore affect the protein's isoelectric point (p*I*). In those cases proteins may appear as horizontal "chain" of spots. Glycosylation can change mass of protein as well as p*I*. Mass of attached sugar chain can be significant and place the protein at higher MW position.
3. Hydrophobic protein due to poor solubilization may have a tendency to form large complexes.

There was no match found for peptides from spot 1. Only a few spots on the gel – spots 2, 3, 7, 9, 10, 11, 13, 16, 38 – represent just one protein. Among spots 7, 9, 10, 13, 16, and 38 corresponds to the molecular mass of whole, intact protein (Table 1, Fig. 2). Proteins in spot 2 and 3 seem to be positioned above their predicted mass and spot 11, was identified as a fragmented form. Spot 2 was detected as a hypothetical protein of unknown function with MW about 61 kDa. This protein was found also in spots 4 and 6, where it seems to be either in intact form (spot 4) or fragmented form (spot 6). We can presume what this hypothetical protein is prone to various modifications and therefore can be found possibly found in several, intact, fragmented and agglomerated forms. A few other proteins were found in intact and fragmented form as well: protein disulfide isomerase (spot 7, 30 32), aldehyde dehydrogenase (spot 8, 22, 24) and others (see Table 1). In each case protein assignment has been verified manually. Some major spots contain only truncated proteins. Thus in spots 24–29, 33, 34 and 39 (all located at molecular weight less then 23 kDa), contain only fragments of proteins. All these spots were reproducible on 3 replicas (data not shown), suggesting what such truncation is attributed to our sample preparation procedures.

At spot 3 only putative allergen was found. This protein was also detected in various spots – 38, 29, 12, 31, 34, 27, 35, 28. Apparently, only spot 38 corresponds to the observed molecular weight of this protein, while all other spots presumably contain the protein in agglomerated or modified form. This protein belongs to large family of lipocalin. The most permanent characteristics of lipocalins are their ability to bind small hydrophobic molecules (fatty acids, hormones), binding to specific receptors and formation of complexes with soluble macromolecules (Flower, 1996). It is likely that the ability to form complexes and binding to

hydrophobic molecules makes this relatively small protein appear on the gel in multiple spots. We used rehydration buffer of 5 M urea 2 M thoiurea, 2%Chaps and 2% SB3-10 as it was recommended for membrane proteins (Molloy, 2000; Santoni et al., 2000), however we were not able to solubilize very hydrophobic proteins efficiently. Enhancing the solubilization power of rehydration buffer by adding more urea (up to 7 M) and adding small amount of nonionic detergent (TX100 for example) in combination with zwitterionic surfactants can diminish this problem.

Possible examples of post-translational modification, when modified proteins form horizontal "chains", can be found in spots 16, 18, 19 and 20. All of them contain V ATPase ε-subunit, 25 kDa protein. All those spots are located on the same molecular weight level, but have significantly different p*I*.

Proteins in mixtures, such as our BBMV samples, have extremely heterogeneous physical property. Therefore not all proteins are completely solubilized while other proteins from the mixture become truncated. In addition, we observe multiple examples of protein co-migration (Table 1), when one spot on gel contains few proteins with unknown copy number, which is common of 2D gel separation (Gygi et al., 2000;Peng and Gygi, 2001). All these factors complicate proteomic analysis using two-dimensional electrophoresis. However, when large and intense spots on the gel contain just one protein this protein can be regarded as abundant. Therefore, arginine kinase (spot 13), fatty acid binding protein (spot 38), actin (spot 10), aldehyde dehydrogenase (spot 9) and protein disulfide isomerase (spot 7) are predominant proteins in *Aedes* BBMV.

4.2. Protein identification of A. aegypti BBMV proteins by the shot gun approach and comparison to 2D gel separation method

All homologous assignments were manually verified for presence of unique peptides. Trypsin digestion generated total 83 matches, while chymotrypsin generated only 37 with 33 overlapping with trypsin-digested proteins ID. Only 4 protein IDs were unique for chymotrypsin digestion, marked with '**' in Tables 1 and 2. Chymotrypsin digestion generates smaller number of spectra, because its target site for cleavage (tyrosine, tryptophan and phenylalanine) produces peptides that are less amenable to electro spray ionization, compared to trypsin-generated peptides (Kinter and Sherman, 2000).

Successful protein identification from MS/MS data depends on several factors:

1. Accuracy of protein sequences databases. The sequenced *A. aegypti* genome was very beneficiary in data searching and protein assignment.
2. Amino acid composition of protein/peptides defines the peptides detection. Ability of a peptide to carry positive charge is critical for ESI MS/MS analysis. Only charged peptides and peptide fragments are detected by mass spectrometry. For that reason, trypsin is the most popular protease: lysine or arginine at the C-terminal of the peptide is always amenable to protonation. However, the distribution of arginine or lysine in an amino acid sequence is the critical factor: when peptides are too short or too long then fragmentation of peptides into sufficient amount of y-and b-ions in the collision cell becomes problematic. For that reason, we used two proteases, trypsin and chymotrypsin to increase the probability for protein identification.
3. Occurrence and detection of a sufficient number of peptides (with appropriate amounts) to generate high probability match. Abundant proteins are more likely to generate larger number of peptides copies and therefore increase chances of successful identification. Protein identification scores are the sum of individual peptide matching scores. Therefore, the larger the number of peptides identified from a particular

protein, the higher the score that this protein will have; while proteins present in smaller amount and generating a smaller number of peptides have less chances to have high score or to be identified at all. Also the score and number of identified peptides cannot be used for quantification; it can serve, however, as “circumstantial evidence” of proteins abundance (at least protein presence) (Ishihama et al., 2005).

The amount of sample is an important factor in MudPIT analysis as well. Koller et al. (2002) identified more than 2500 proteins from 3 mg of rice plant material; Wolters et al. (2001) detected about 1500 proteins from a total amount of about 1.3 mg of fractionated protein; Ostrowski et al. (2002) detected only 66 proteins by multidimensional LC–MS/MS from just 30 μg of human cilia. Of course, the number of detected proteins depends on sample nature, exact protocol and other factors, but evidently, larger amount of starting material gives better results. We had a relatively small sample, 140 μg of a proteins mixture, which were divided into 2 parts for trypsin and chymotrypsin digestion, and we identified 86 proteins by MudPIT.

Tables 1 and 2 contain the results of our MudPIT experiment. Table 1 shows 36 proteins, found in both shotgun and 2D analysis. Arginine kinase, three different actins (108872511, 108879764 and 108878285) and putative allergen, detected with highest score in shotgun method, also found as single proteins in the largest spots 10, 13 and 38 on 2D gels, can be regarded as the most predominant proteins in the sample. Vacuolar ATP synthase subunits alpha and beta having very high score in MudPIT detection, were found on 2D gel at spots 32 and 34. We suppose, that many other proteins from Table 1 which were detected by MudPIT with high scores and significant protein coverage but found on 2D gel in truncated form for similar reasons: either their intact spot was not picked for analysis or those protein do not appear on 2D gel as predominate spots due to severe fragmentation during 2D protocol.

4.3. Comparison of 2D gel and MudPIT methods

In some cases, protein detection by the 2D gel approach was more successful than by the MudPIT method. For instance, disulfide isomerase, a single protein (spot 7) on the 2D gel identified by 18 peptides with 36% coverage and a high score, was detected in the shot gun approach only by 2 peptides with a relatively low score of 40. This spot does not appear to be predominant in the mixture; it is rather small compared to other spots. But it was the only protein in gel cores and therefore was detected by mass spectrometry with higher score than if it was mixed with other proteins. Glutamine synthetase was detected from large gel spots 12 and 15 with scores 832, 15 peptides and 265, 5 peptides, respectively, while in the shotgun method it has a relatively low score of 60 and 2 peptides. There are co-migrating proteins – 4 in spot 12 and 3 in spot 15. In this case it is hard to tell which protein in each spot is predominant. However even less abundantproteins can bedetectedeasier froma mixtureof justa few proteins cored from a single gel spot than from whole sample containing hundreds of proteins. This illustrates how 2D gel separation can be more advantageous in detecting low abundance proteins than the shotgun method.

Table 2 contains 50 proteins which had been found in the shotgun method only. Metallopeptidase with aminopeptidase activity, 112 kDa, and three sterol carrier proteins, which were detected by the MudPIT approach with high score and more then 30% coverage of amino acid sequences were not found on 2D gel. As it was mentioned before, high and low molecular weight proteins may be difficult to resolve by 2D electrophoresis. In general, it is concern in case of extreme size: over 150 kDa and lower then 10 kDa. However, in our experiment with rather complicated sample on 11 cm gels, the resolution of proteins around 100 kDa and above becomes challenging. Other large proteins (metalloprotease 102.5 kDa; aconitase 88.1 and 99.2 kDa; putative heat shock protein, 107 kDa; estradiol 17 beta-dehydrogenase 78 kDa; heat shock protein 81.8 kDa; kine-sin-like protein KIF3A 77.8 kDa) and small proteins (sterol carrier protein, 11.1, 11.2 and 11.3 kDa; Heat shock protein 10.7

kDa; diazepam binding inhibitor, 9.8 kDa; Cytochrome C, 11.8 kDa; and, putative thioredoxin, 12.1 kDa) were not found on 2D gels but were identified by the MudPIT method. Carbonic anhydrase (26) and cd-36 antigen are membrane proteins, which are also known to be difficult for 2D electrophoresis separations, and possibly for that reason were not present on the gel but were identified by MudPIT.

Some homologous proteins were detected on either 2D or shotgun data set, but not on both. Alpha amylase (gi|108873259) was found by MudPIT and alpha amylase (gi|108873258) by 2D gel. Our analysis (shown in Fig. 3) indicates that these are different proteins. From four different aldehyde dehydrogenases, two were detected on 2D gel, while other two were not. Again, our analysis indicates that these are different proteins (data not shown).

Many other proteins from Table 2 may be present on 2D but in lower abundance spots and were not picked for coring. To excise and analyze every single spot on 2D gel is very costly, and this is one of the reason why the shotgun approach is considered preferable for proteome analysis of protein mixtures.

A total of 400 proteins were detected via 2D gel. Thirty-three of the most abundant of these were selected for analysis. Out of 33 proteins detected on 2D gel but not by shotgun method, only 4 proteins were found in more than one spot and almost all of them co-migrated with other proteins which were identified by the MudPIT approach (Table 1 and 2). Because of the co-migration problem in 2D, it is not known which protein from the spot is predominant and responsible for the expression level. However, when certain proteins from such spots were found by both methods, we can presume that those proteins are more amenable to detection, due to higher expression level, and proteins identified by 2D only are likely to be less abundant than proteins, detected by both methods.

There are two proteins: hypothetical protein #108883797, and NADH ubiquinone oxyreductase which are particularly interesting. Hypothetical protein #108883797 appears to be a single protein in relatively large spot 2. Proteins identified by high scores, e.g., 13, and proteins appearing alone in spots indicate rather abundant protein in the sample. However, this protein was not detected by shotgun approach, because its peptides co-eluted with peptide spots of more predominant proteins. Even though this is not typical situation when protein appears predominant on 2D gel and not identified by MudPIT, it demonstrates that these two approaches are complementary to each other. Enzyme NADH ubiquinone oxireducatse was found to be just one protein in spot 11, one of the predominant spots on the gel as well. However, low score and peptide number and protein coverage indicate that this protein is unlikely to be predominant even in this particular spot.

5. Conclusions

Analyzing a total of 270 μg of BBMV (140 μg by MudPIT and 130 μg by 2D gels) by two complementary proteomic approaches, we identified 119 proteins: 86 total by MudPIT method and 69 by 2D gel electrophoresis. The combination of these two methods gives the advantage of overcoming the shortcomings of each of the individual methods. 2D gel electrophoresis, the most popular method for proteomic analysis, provides a visual display but has well known limitations (Santoni et al., 2000): (1) narrow dynamic range of the proteins, (2) difficulty with solubilization of hydrophobic proteins, and (3) precipitation of soluble proteins at their isoelectric point (p*I*) and it is problematic to detect high and low molecular weight proteins as well as proteins with extreme p*I* on 2D gels. These problems all applied to our samples. A satisfactory complimentary method is two-dimensional liquid chromatography (LC/LC) followed by tandem mass spectrometry.

Three proteins, arginine kinase, putative allergen and actin are shown to be the most predominant proteins in the sample. Total number of 36 proteins detected by both methods represents the most abundant proteins in the *A. aegypti* BBMV.

Acknowledgments

The authors thank Dr. Oscar Alzate for comments on the manuscript. This project was supported in part by NIH grant R01 AI 29092 to D.H.D.

References

Billingsley, PF.; Lehane, MJ. Structure and ultrastructure of the insect midgut. In: Lehane, MJ.; Billinggsley, PF., editors. The Insect Midgut. Chapman & Hall; London: 1996. p. 3-30.

Carr S, Aebersold R, Baldwin M, Burlingame A, Clauser K, Nesvizhskii A. The need for guidelines in publication of peptide and protein identification data. Mol Cell Proteom 2004;3:531–533.

Dronina MA, Revina LP, Kostina LI, Ganushkina LA, Zalunin IA, Chestukhina GG. Toxin-binding proteins from midgut epithelium membranes of *Anopheles stephensi* larvae. Biochemistry (Mosc) 2006;71:133–139. [PubMed: 16489916]

English LH, Readdy TL. Delta endotoxin inhibits a phosphatase in midgut epithelial membranes of *Heliothis virescens*. Insect Biochem 1989;19:145–152.

Fernández LE, Aimanova KG, Gill SS, Bravo A, Soberón M. A GPI-anchored alkaline phosphatase is a functional midgut receptor of Cry11Aa toxin in *Aedes aegypti* larvae. Biochem J 2006;394:77–84. [PubMed: 16255715]

Flower DR. The lipocalin protein family: structure and function. Biochem J 1996;318(Pt 1):1–14. [PubMed: 8761444]

Gygi S, Corthals GYZ, Rochon Y, Aebersold R. Evaluation of two-dimensional gel electrophoresis based proteome analysis technology. Proc Natl Acad Sci USA 2000;97:9390–9395. [PubMed: 10920198]

Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteom 2005;4:1265–1272.

Kinter, M.; Sherman, N. Protein Sequencing and Identification Using Tandem Mass Spectrometry. Wiley-Interscience Inc; 2000.

Koller A, Washburn MP, Lange BM, Andon NL, Deciu C, Haynes PA, Hays L, Schieltz D, Ulaszek R, Wei J, et al. Proteomic survey of metabolic pathways in rice. Proc Natl Acad Sci USA 2002;99:11969–11974. [PubMed: 12163647]

McNall RJ, Adang MJ. Identification of novel *Bacillus thuringiensis* Cry1Ac binding proteins in *Manduca sexta* midgut through proteomic analysis. Insect Biochem Mol Biol 2003;33:999–1010. [PubMed: 14505693]

Molloy M. Two-dimensional electrophoresis of membrane proteins using immobilized pH gradient. Anal Biochem 2000;280:1–10. [PubMed: 10805514]

Ostrowski LE, Blackburn K, Radde KM, Moyer MB, Schlatzer DM, Moseley A, Boucher RC. Proteomic analysis of human cilia. Mol Cell Proteom 2002;1:451–465.

Peng J, Gygi S. Proteomics: the move to mixtures. J Mass Spectrom 2001;36:1083–1091. [PubMed: 11747101]

Santoni V, Malloy M, Rabbiloud T. Membrane proteins and proteomics: un amour impossible? Electrophoresis 2000;21:1054–1070. [PubMed: 10786880]

Wolters D, Washburn M, Yates JRI. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 2001;73:5683–5690. [PubMed: 11774908]

Wu CC, Yates JR. The application of mass spectrometry to membrane proteomics. Nat Biotechnol 2003;21:262–267. [PubMed: 12610573]

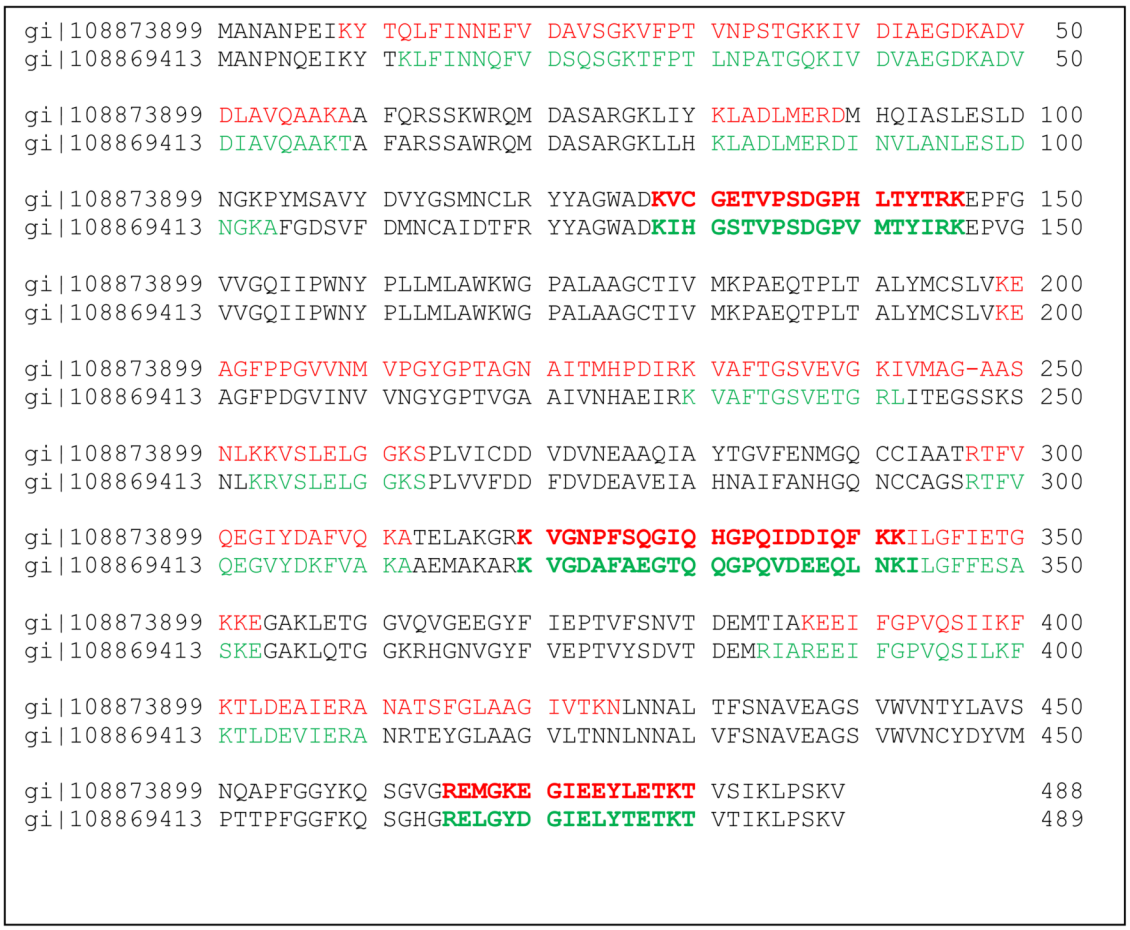

Fig. 1.

Pairwise alignment of two aldehyde dehydrogenases identified in spot 8 and 9 on 2D gel. Top line is the amino acid sequences of the protein gi|108873899, identified peptides shown in red; bottom line corresponds to amino acid sequence of gi|108869413, identified peptides are shown in green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

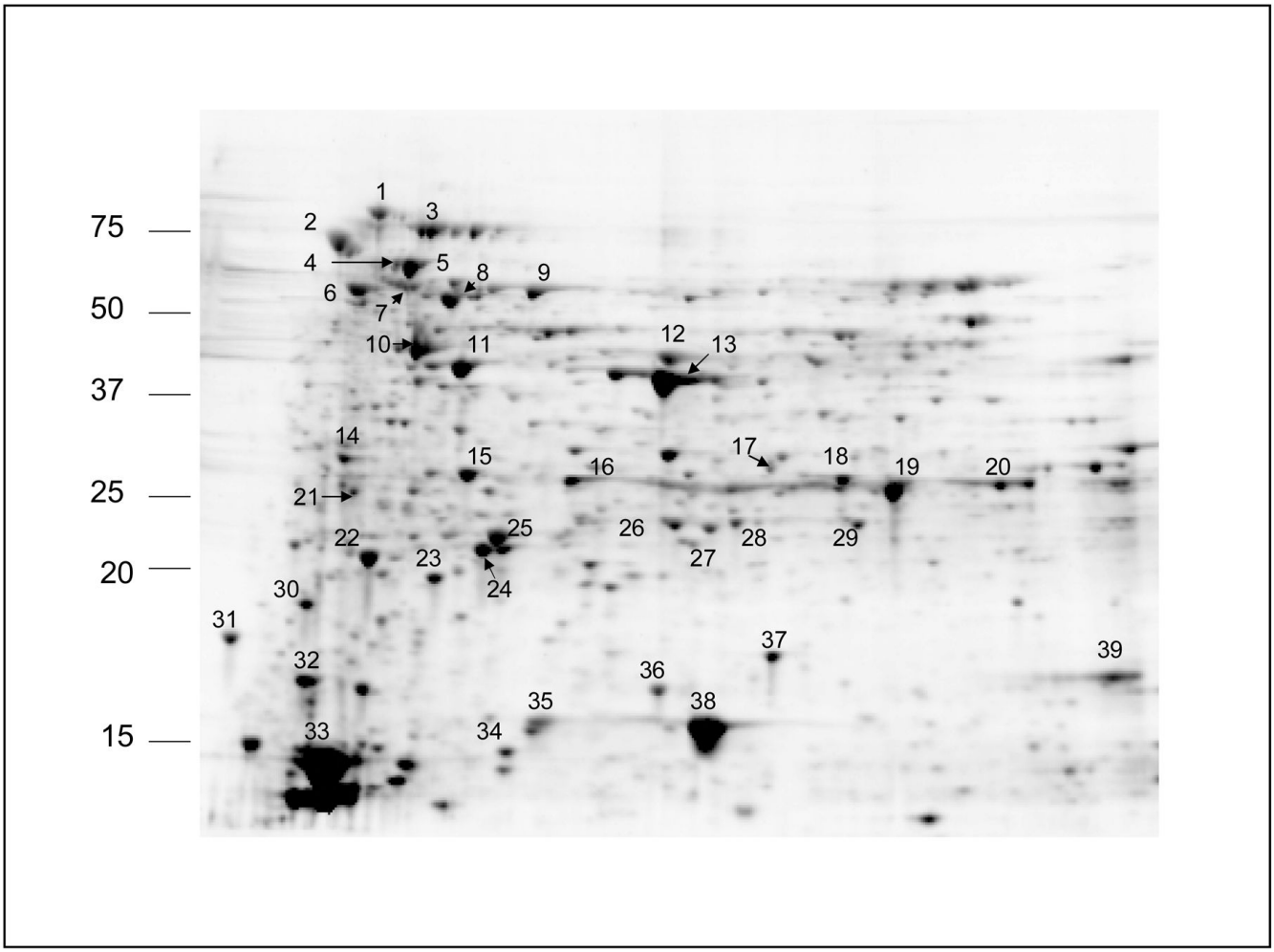

Fig. 2.

A. aegypti brush border membrane vesicle proteins from dissected midguts. The horizontal dimension was isoelectric focusing and the vertical dimension was SDS-PAGE. Molecular weights of standard are indicated on the left. The numbered protein spots were cored and identified.

```
gi|108873258 MKKCAVVCLL GLLALAGAKS AVKQDGHDHD MPELDWWEGG VFYQIYPRSF  50
gi|108873259 -----MISLR VPLLLLVATA VLLATAQDSD QKE--WWETT VFYQIYPRSF  50

gi|108873258 KDTNNDGVGD IAGIMEKLDH LVDLGVTGVW FSPLFKSPMK DFGYDISDFK 100
gi|108873259 FDSNDDGTGD IKGITAKLQH LKDTGFEATW LSPIFQSPQE DFGYDVSDFV 100

gi|108873258 DVDPTFGTLE DLKALIKKAK ELGIKVILDF VPNHTSDEHE WFKKALAD-D 150
gi|108873259 SVDPLFGSNS DLEELFAEAE KLGIKIILDF VPNHSSNEHE WFVKSENRVD 150

gi|108873258 PDYIDYYVWK DG--NAEGG- --PPNNWQSV FHTDAWTKPA GKSKYYLHQF 200
gi|108873259 P-YTDYYMWH NGKPNPQGGR PLPPNNWQSV FYGSAWEWSE KRQQYYLHQF 200

gi|108873258 DKGQPDLNYE NPKVKAEMEE MLHFWFELGV DGFRIDAINH AYEDAGFLDE 250
gi|108873259 AVGQPDLNYR NEAVIKEFDE ILRYWMKKGA SGFRIDAINH MFEVEDLRDE 250

gi|108873258 PIIDENKGLF YENMEHIYTM NQNESYELIY DWRVVFDEWS EKSNQ-TKLM 300
gi|108873259 PINDPSDPNS YGYTHHIYTK DLPDTYEVIA RWRKVIDDYV KESDSDTIIM 300

gi|108873258 MTEAYANMEQ TMRWY--GDG KRNGSHFPFN FAMINRIESS SNAADFKEVI 350
gi|108873259 MTEAYANLTM TMKFYESDDG TQPRAHFPFN FAMIEDLNDG SKASNFKYII 350

gi|108873258 DEWLDNMPEG GNANWVLGNH DRPRIASRFG RDRAASFAIM EMTLPGIAVV 400
gi|108873259 DRWLDNMPRG KITNWVLGNH DKPRMASRYG RDRIDGMALI LMTLPGVAVV 400

gi|108873258 YYGEEIGMED YRDISFEDTQ DPQAANTNKE IYQLYTRDPV RTPFQWDNTT 450
gi|108873259 YNGEEIGMED YRDMSYEDSK DPQGCNLGPD NYKWASRDPQ RTPFQWDDSF 450

gi|108873258 YAGFTGSAAE KTWLPVHPNY KELNLAAQKE DPKSLFTLYK NLIQLRKDHT 500
gi|108873259 NAGF--SKAA KTWLPMHPLY RQTNLLKQTE ADYSTYHFYV DCMKLRKERI 500

gi|108873258 FKYGSFESKA LVNNVFGFTR KLDDHKS--- ---YAVVVNM NSMEAQLNLK 550
gi|108873259 LTHGEFRSRA FNDDVFAFVR FLRENEDREL DPYYVTLVNF HGETYTVDVT 550
```

Fig. 3.

Pairwise alignment of two different alpha-amylases identified by 2D Protein gi#108873258 and MudPit gi#108873259 experiments. Identified peptides are shown in red color for alpha-amylase 108873258 and in green for gi#108873259. Those proteins have 48% identity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 1

Aedes aegypti BBMV proteins identified by 2D gel separation and shotgun approach

Spot number	Protein identified	Accession number	Predicted mass, Da	Identification data from 2D gel			Identification data from MudPIT			Proteins detected on 2D gel only
				Mascot score	No. of peptides	% of protein coverage	Mascot score	No. of peptides	% of protein coverage	
2	hypothetical protein	108883797	60680	934	16	35				*
3	allergen, putative	108878248	14908	346	5	58	1901	16	81	
4	**hypothetical protein**	**108883797**	**60680**	**487**	**8**	**14**				*
	alkaline phosphatase	**108881200**	**61117**	**99**	**2**	**3**				*
5	**chaperonin-60kD, ch60**	**108872102**	**61155**	**1337**	**18**	**40**	187	6	14	
	cyclohex-1-ene-1-carboxyl-CoA hydratase, putative	108880435	31909	147	3	15	346	8	35	
6	hypothetical protein	108883797	60680	225	3	6				*
	peptidyl-prolyl cis-trans isomerase f, ppif	108871914	22639	111	2	14	498	3	27	
7	**protein disulfide isomerase**	**108884061**	**56260**	**1108**	**18**	**36**	40	2	6	
8	**aldehyde dehydrogenase**	**108869413**	**53305**	**1084**	**18**	**39**	339	6	17	
	glutamate carboxypeptidase	**108883076**	**53689**	**549**	**11**	**27**				*
	chaperonin-60kD, ch60	108872102	61155	110	2	5	187	6	14	
9	**aldehyde dehydrogenase**	**108873899**	**53037**	**1478**	**20**	**47**	252	6	17	
10	**actin**	**108872511**	**42194**	**761**	**11**	**34**	1722	18	61	
	actin	**108879764**	**42058**	**590**	**9**	**29**	1657	13	51	
11	NADH-ubiquinone oxidoreductase flavoprotein 1	108876370	54040	64	2	3				*
12	**aliphatic nitrilase, putative**	**108873526**	**43923**	**849**	**16**	**53**				*

Spot number	Protein identified	Accession number	Predicted mass, Da	Identification data from 2D gel			Identification data from MudPIT			Proteins detected on 2D gel only
				Mascot score	No. of peptides	% of protein coverage	Mascot score	No. of peptides	% of protein coverage	
	glutamine synthetase 1, 2	**108882715**	**45100**	**832**	**15**	**37**	60	2	5	
	allergen, putative	108878248	14908	310	5	51	1901	16	81	
	actin	**108878285**	**41958**	**87**	**2**	**7**	1202	14	46	
13	**arginine or creatine kinase**	**108874753**	**40191**	**1598**	**24**	**61**	2364	27	67	
14	actin	108872511	42194	691	8	33	1722	18	61	
	actin	108879764	42058	626	9	30	1657	13	51	
	acyl-coa dehydrogenase	108876518	46836	260	4	11	480	9	42	
	chaperonin-60kD, ch60	108872102	61155	181	3	5	187	6	14	
	myo inositol monophosphatase	**108884401**	**30332**	**116**	**2**	**5**				*
15	**beta lactamase domain**	**108874176**	**33271**	**643**	**12**	**41**	136	4	17	
	arginine or creatine kinase	108874753	40191	425	6	21	2364	27	67	
	glutamine synthetase 1, 2	108882715	45100	265	5	10	60	2	5	
16	**vacuolar ATP synthase subunit e**	**108871609**	**25728**	**987**	**18**	**58**	95	4	20	
17	arginine or creatine kinase	108874753	40191	1488	18	55	2364	27	67	
	conserved hypothetical protein	**108884639**	**31569**	**407**	**9**	**34**				*
	conserved hypothetical protein	**108872537**	**31227**	**121**	**3**	**12**	105	3	3	
	phosphatidylethanolamine-binding protein	108876530	23313	96	2	21				*
18	conserved hypothetical protein	108872537	31227	654	8	38	105	3	3	
	vacuolar ATP synthase subunit e	**108871609**	**25728**	**254**	**4**	**19**	95	4	20	
	aldehyde dehydrogenase	108873899	53037	174	3	8	252	6	17	
	cyclohex-1-ene-1-carboxyl-CoA hydratase, putative	108880435	31909	152	3	19	346	8	35	
	glutathione s-transferase	**108871931**	**27034**	**142**	**3**	**17**	478	10	58	

Spot number	Protein identified	Accession number	Predicted mass, Da	Identification data from 2D gel			Identification data from MudPIT			Proteins detected on 2D gel only
				Mascot score	No. of peptides	% of protein coverage	Mascot score	No. of peptides	% of protein coverage	
	catalase	108870108	57148	138	2	5	117	3	7	
	conserved hypothetical	108883378	33060	114	3	9				*
	protein nadp-specific isocitrate dehydrogenase	108883996	46264	89	2	5	541	9	27	
	hypothetical protein	108871193	37184	56	2	5				*
19	cyclohex-1-ene-1-carboxyl-CoA hydratase, putative	108880435	31909	1071	18	53	346	8	35	
	3-hydroxyisobutyrate dehydrogenase	108869599	34126	628	10	48	257	4	20	
	adenylate kinase 1, putative	**108870854**	**26438**	**381**	**7**	**31**	95	2	9	
	catalase	108870108	57148	191	3	6	117	3	7	
	vacuolar ATP synthase subunit e	**108871609**	**25728**	**149**	**3**	**16**	95	4	20	
	conserved hypothetical protein	108872537	31227	141	3	16	105	3	3	
	electron transport oxidoreductase	108869776	34440	97	2	7	259	5	29	
20	3-hydroxyisobutyrate dehydrogenase	108869599	34126	901	13	55	257	4	20	
	cyclohex-1-ene-1-carboxyl-CoA hydratase, putative	108880435	31909	529	10	40	346	8	35	
	electron transfer flavoprotein beta-subunit	108879274	22906	428	8	42				*
	3-2trans-enoyl-CoA isomerase, putative	108879601	31541	239	5	15				*
	short-chain dehydrogenase	108877992	29579	140	3	10				*
	vacuolar ATP synthase subunit e	**108871609**	**25728**	**116**	**3**	**9**	95	4	20	
	conserved hypothetical protein	108880463	17323	95	2	20	189	4	39	
	hypothetical protein	**108872539**	**26361**	**52**	**2**	**6**				*
21	alpha-amylase	108873258	68905	407	9	13				*
	cathepsin l	108881694	38479	262	3	13				*
	14-3-3 protein sigma, gamma, zeta, beta/alpha	**108877244**	**28324**	**239**	**4**	**17**				*
	glutathione s-transferase	108871931	27034	212	3	21	478	10	58	
	actin	108882963	42149	170	3	10				*
	acyl-coa dehydrogenase	108876518	46836	158	2	5	480	9	42	

Spot number	Protein identified	Accession number	Predicted mass, Da	Identification data from 2D gel			Identification data from MudPIT			Proteins detected on 2D gel only
				Mascot score	No. of peptides	% of protein coverage	Mascot score	No. of peptides	% of protein coverage	
	chaperonin-60kD, ch60	108872102	61155	143	2	4	187	6	14	
	proteasome subunit alpha type	**108880550**	**27839**	**86**	**2**	**9**				*
22	**glutathione s-transferase**	**108871931**	**23**	**839**	**13**	**64**	478	10	58	
	aldehyde dehydrogenase	108869413	53305	105	2	6	339	6	17	
23	**phosphatidylethanolamine-binding protein**	**108876530**	**23313**	**590**	**7**	**47**				*
	acyl-coa dehydrogenase	108876518	46836	105	2	8	480	9	42	
24	peroxiredoxin 6, prx-6	108882310	25108	974	17	68	189	9	60	
	beta lactamase domain	108874176	33271	283	6	21	136	4	17	
	juvenile hormone-inducible protein, putative	108883478|	50240	222	5	10	131	6	16	
	glutathione-s-transferase theta, gst	108884710	24931	207	5	24	92	2	13	
	aldehyde dehydrogenase	108869413	53305	107	2	4	339	6	17	
	DNA-directed RNA polymerase II	108872649	24544	47	2	10				*
25	acyl-coa dehydrogenase	108876518	46836	1182	17	53	480	9	42	
	arginine or creatine kinase	108874753	40191	87	2	6	2364	27	67	
	eukaryotic translation elongation factor	108879886	95301	59	2	2	63	5	8	
26	oxidoreductase	108876391	26898	626	9	39	198	3	18	
	cyclohex-1-ene-1-carboxyl-CoA hydratase, putative	108880435	31909	188	3	16	346	8	35	
	oxidoreductase	108883208	29224	112	2	10				*
	triosephosphate isomerase	108882001	26705	108	2	9	70	2	17	
27	triosephosphate isomerase	108882001	26705	835	13	64	70	2	17	
	allergen, putative	108878248	14908	212	4	34	1901	16	81	
	3-hydroxyisobutyrate dehydrogenase	108869599	34126	190	3	11	257	4	20	

Spot number	Protein identified	Accession number	Predicted mass, Da	Identification data from 2D gel			Identification data from MudPIT			Proteins detected on 2D gel only
				Mascot score	No. of peptides	% of protein coverage	Mascot score	No. of peptides	% of protein coverage	
28	allergen, putative	108878248	14908	442	6	51	1901	16	81	
	oxidoreductase	108876391	26898	187	2	9	198	3	18	
29	arginine or creatine kinase	108874753	40191	1205	16	47	2364	27	67	
	allergen, putative	108878248	14908	913	12	78	1901	16	81	
	citrate synthase	108881547	51852	134	3	6	686	7	25	
	conserved hypothetical protein **	108872537	31227	54	2	8	105	3	3	
30	**translationally controlled tumor protein**	**108880570**	**19704**	**856**	**10**	**71**	125	3	19	
	protein disulfide isomerase	108884061	56260	273	7	15	40	2	6	
31	**calmodulin**	**108871289**	**16800**	**866**	**13**	**65**				*
	allergen, putative	108878248	14908	112	3	31	1901	16	81	
32	translationally controlled tumor protein	108880570	19704	473	5	52	125	3	19	
	prefoldin, subunit, putative	**108882254**	**15963**	**239**	**5**	**39**				*
	conserved hypothetical protein	**108879062**	**15572**	**169**	**4**	**54**				*
	arp2/3 complex 16 kd subunit (P16-arc)	108874879	16946	165	3	27				*
	acyl-coa dehydrogenase	108876518	46836	144	3	10	480	9	42	
	ATP synthase alpha subunit vacuolar	108875173	68528	143	3	7	1379	19	46	
	arginine or creatine kinase	108874753	40191	123	4	15	2364	27	67	
	protein disulfide isomerase	108884061	56260	113	3	8	40	2	6	
	hypothetical protein	108873119	61622	88	2	5				*
	ribosomal RNA small subunit methyltransferase b (sun)	108872435	49137	55	2	4				*
33	acyl-CoA oxidase	108875415	75711	179	3	5				*
	conserved hypothetical protein	108879061	21865	108	2	17	395	5	44	

Spot number	Protein identified	Accession number	Predicted mass, Da	Identification data from 2D gel			Identification data from MudPIT			Proteins detected on 2D gel only
				Mascot score	No. of peptides	% of protein coverage	Mascot score	No. of peptides	% of protein coverage	
34	allergen, putative	108878248	14908	700	9	59	1901	16	81	
	ATP synthase subunit beta vacuolar	108878452	55466	210	4	9	1326	16	55	
	vacuolar ATP synthase subunit f	**108882111**	**14229**	**74**	**1**	**9**				*
	protease m1 zinc metalloprotease	108875835	211798	59	2	1	412	12	10	
35	**allergen, putative**	**108878248**	**14908**	**977**	**14**	**82**	1901	16	81	
	peroxiredoxin 6, prx-6	108882310	25108	314	7	39	189	9	60	
36	**superoxide dismutase**	**108884477**	**15616**	**739**	**9**	**73**				*
	calmodulin	108871289	16800	198	5	36				*
	thioredoxin reductase	108881629	57184	116	2	4				*
37	**conserved hypothetical protein**	**108880463**	**17323**	**566**	**9**	**68**	189	4	39	
38	**allergen, putative**	**108878248**	**14908**	**1277**	**18**	**82**	1901	16	81	
39	peptidyl-prolyl cis-trans isomerase f, ppif	108871914	22639	534	8	48	498	3	27	
	acetyl-coa acetyltransferase, mitochondrial	108877532	43577	110	2	4	149	4	15	
	esterase, putative	108873192	32558	87	2	5				*

Table 2

Proteins identified by MudPIT method only ** proteins detected with chymotrypsin digestion

Protein identified	Accession number	Predicted mass, Da	Mascot score	No. of peptides	Percent of protein coverage
protease m1 zinc metalloprotease	108870804	112249	1148	21	32
sterol carrier protein-2, putative	108870885	11387	864	8	80
sterol carrier protein-2, putative	108870882	11107	856	8	73
sterol carrier protein-2, putative	108870883	11260	633	5	50
acetyl-coa acyltransferase	108873073	41941	417	5	24
acetyl-coa acetyltransferase 2, putative	108877211	41313	353	6	27
carbonic anhydrase	108879424	31630	232	3	15
alpha-amylase	108873259	70411	226	6	15
cd36 antigen	108874482	56096	226	2	6
aconitase, mitochondrial	108880755	88134	221	6	13
malate dehydrogenase	108875864	45044	191	7	26
transketolase	108879967	68461	188	3	5
nucleoside-diphosphate kinase NBR-A, putative	108871239	18501	172	6	51
methylmalonate-semialdehyde dehydrogenase	108883539	57220	166	7	20
nadp-specific isocitrate dehydrogenase	108870975	37184	149	5	18
aspartate aminotransferase	108882223	47688	149	6	21
serine protease inhibitor 4, serpin-4	108876285	43500	146	2	9
protease m1 zinc metalloprotease	108870802	102538	134	7	10
glutathione-s-transferase theta, gst	108883606	25026	133	2	11
serine-type endopeptidase, putative	108877559	27076	130	2	13
heat shock protein, putative	108883661	10737	118	2	29
zinc carboxypeptidase	108875384	48322	106	3	9
diazepam binding inhibitor, putative	108874714	9878	106	2	37
sulfotransferase (sult)	108877850	38472	105	2	13
elongation factor 1-beta2	108883773	24708	104	2	14
serine-type enodpeptidase, putative	108871734	26996	102	2	10
conserved hypothetical protein	108873250	68905	101	4	10