# A proteomics approach for the analysis of hemolymph proteins involved in the immediate defense response of the soft tick, *Ornithodoros savignyi*, when challenged with *Candida albicans*

Elaine Stopforth, Albert W.H. Neitz and Anabella R.M. Gaspar*

Department of Biochemistry, University of Pretoria, Pretoria 0002, South Africa
* Corresponding author: anabella.gaspar@up.ac.za

## Abstract

A proteomics approach was employed to identify proteins secreted into the hemolymph of *O. savignyi* ticks two hours after immune-challenge with the yeast, *Candida albicans*. Profiling of the proteins present in hemolymph of unchallenged ticks versus ticks challenged with heat-killed yeast revealed 5 proteins to be differentially expressed. The modulated protein spots were subjected to tandem mass spectrometry (MS/MS) analysis, but could not be positively identified. These proteins can be assigned to the immune response as they were not induced after aseptic injury. In an attempt to identify hemolymph proteins that recognize and bind to yeast cells, hemolymph obtained from both unchallenged and challenged ticks was incubated with *C. albicans.* Elution of the bound proteins followed by SDS-PAGE analysis indicated that three proteins (97, 88 and 26 kDa) present in both unchallenged and challenged hemolymph samples bind to yeast cells. The constant presence of these three proteins in tick hemolymph leads us to believe that they may be involved in non-self recognition and participate in yeast clearance from tick plasma. The analyzed yeast-binding proteins could also not be positively identified, suggesting that all the tick immune proteins investigated in this study are novel.

Keywords: *Ornithodoros savignyi*, Hemolymph, Differential protein expression, Yeast, Yeast-binding proteins

## Introduction

The soft tick, *Ornithodoros savignyi* more commonly known as the sand tampan, is distributed throughout the North Western parts of Southern Africa as well as Egypt, Ethiopia, Kenya and Zimbabwe (Paton and Evans, 1929). The bites of these ticks cause irritation, blisters and bruising (Estrada-Pena and Jongejan, 1999). They cause large losses of livestock especially young calves and lambs in the areas where they occur and are of special relevance in Southern Africa (Mans and Neitz, 2003). Pathogenesis associated with these ticks is caused by secretion of a toxin during feeding and can lead to cardiac failure (Mans et al., 2004). The toxin also induces a serious allergic reaction in humans (Howell et al. 1975). No infectious diseases are transmitted by these ticks, although it has been reported that *O. savignyi* caused relapsing fever in humans (Paton and Evans, 1929).

Invertebrates only have an innate immune system which is the first line of defense against infections. It is activated by host proteins that recognize conserved surface determinants of pathogens, such as lipopolysaccharide (LPS), lipoteichoic acid (LTA) and peptidoglycan (PGN) from bacteria and β-1,3-glucans from fungi (Medzhitov and

Janeway, 2002). Both cellular and humoral immune responses are activated to eliminate pathogens (Imler and Hoffmann, 2000). In contrast to the well investigated immune response in insects, research into tick innate immunity is rather limited (Sonenshine and Hynes, 2008). In ticks, cellular responses include phagocytosis and encapsulation, while humoral responses lead to the secretion of antimicrobial peptides and proteins. Most of the research in this field has focused on the identification and characterization of antimicrobial peptides (AMPs) in various hard tick species and in one soft tick, *O. moubata.*

AMPs and other immune-related proteins are synthesized within the fat bodies or hemocytes and secreted into the plasma. Both mRNA-based and proteomic approaches should be combined to decipher the complexity of innate immunity (Levy et al., 2004), however, proteomic analysis allows the investigation of innate immunity at the protein level, the real mediators of physiological function. Moreover, genomic approaches cannot be used for fluids such as hemolymph. The proteins in the hemolymph of *Drosophila melanogaster* have been extensively researched. The second dimension (2D) -map of hemolymph proteins of *Drosophila* was constructed to serve as a reference database for researchers investigating the changes that occur at protein level in different developmental stages, physiological conditions or after infection (Vierstraete et al., 2003). The 2D-database was updated by de Morais Guedes et al. (2003). In one of the first studies of this kind, Vierstraete et al. (2004) employing 2D difference gel electrophoresis (DIGE) identified 10 differential proteins that appear in the fruit fly hemolymph very early after immune-challenge with LPS. Using 2D-gel electrophoresis (2DE), Levy et al. (2004) found that more than 70 out of a total of 160 hemolymph proteins from *Drosophila* were modulated by at least 5-fold after microbial infection. MS and data bank searches revealed the identity of proteins directly involved in innate immunity such as proteases, protease inhibitors and recognition molecules. Proteins with a potential function in the immune response as well as molecules resulting from the cleavage of proteins were also identified in this study.

Since the first paper was published by Madden et al. (2002) highlighting the complications associated with using proteomics approaches to characterize protein profiles of two related tick species, *Amblyomma americanum* and *A. maculatum*, several papers using the above approach have appeared in the tick field (Untalan et al., 2005; Vennestrom and Jensen, 2007; Oleaga et al., 2007; Francischetti et al., 2008a, 2008b; Rachinsky et al., 2007 & 2008; Wickramasekara et al., 2008). Most of these papers described the identification of either salivary gland or larval proteins from a few hard and soft tick species. Only two papers have reported using differential protein expression to identify proteins modulated in *Babasia bovis*-infected *Rhipicephalus (Boophilus) microplus* ovaries and midgut (Rachinsky et al., 2007, 2008).

Hemolymph molecules associated with tick immunity that have been described thus far include lysozyme, defensins, histidine-rich peptides, non-cationic peptides as well as an alpha-macroglobulin like glycoprotein (reviewed in Sonenshine and Hynes, 2008; Taylor, 2006). To date no information on innate immunity has been published for *O. savignyi*. To our knowledge no proteomics investigations have been performed to identify hemolymph proteins involved in tick innate immunity. This paper is the first report of a differential analysis performed in an attempt to identify

hemolymph proteins involved in the immediate defense response of *O.savignyi* to *C. albicans* challenge. In addition preliminary characterization of hemolymph proteins involved in the recognition and binding of the yeast cells is also described.

## Materials and methods

*Ticks*

*O. savignyi* ticks were collected from the Upington region (Northern Cape, South Africa), by sifting of sand, placed in plastic containers containing sterile sand and kept at room temperature. Adult female ticks were used in all experiments.

*Culturing of microorganisms*

*C. albicans* in pellet form was obtained from the American type culture collection (ATCC catalog no. 90028$^{TM}$, USA) and cultured according to the supplier's specifications. For long term storage, 20 ml of the yeast culture was aliquoted into Eppendorf tubes (500 µl of yeast each) and 500 µl of 25% glycerol. The tubes were stored at -70ºC. For the challenge of ticks with *C. albicans*, an aliquot of the yeast culture was thawed at room temperature and 500 µl of the culture added to 9.5 ml of potato dextrose broth (PDB) and grown overnight at 30$^{o}$C in a shaking incubator. The culture (1 ml) was then centrifuged at 8200 x g for 15 minutes, the supernatant removed and the pellet reconstituted in 1 ml of sterile saline (0.9% NaCl) to obtain an $OD_{600nm}$= 0.25. The suspension was inactivated by boiling at 95$^{o}$C for 5 minutes (heat killed).

*Challenge of ticks and hemolymph collection*

Ticks were immobilized onto Petri-dishes with their ventral sides up using double sided tape. Before inoculation, tick integuments were by wiped with ethanol. Ticks were injected into their idiosoma between their second and third coxa using a 10 µl Hamilton micro syringe (Hamilton Co., Reno, Nevada) and a 30 G1/2 sterile needle (Becton Dickinson and Co.). Hemolymph was collected from ticks by piercing their legs with a needle after they were immobilized as described above. The hemolymph was extracted with a PCR micropipette (1-5 µl, Drummond scientific Co., USA) and collected into Eppendorf tubes containing either lysis solution or tick bleeding buffer (TBB).

For 2D differential expression experiments, 25 ticks were injected with either 1 µl of sterile saline (aseptic injury control; S2) or heat-killed yeast cells (Y2) and the hemolymph collected after 2 hours into a lysis solution containing 8 M urea, 5% (w/v) CHAPS and a mixture of protease inhibitors (Complete, Roche Diagnostics) at 4 ºC. Hemolymph obtained from 25 non-injected ticks served as the unchallenged control (UN).The samples were centrifuged for 15 minutes at 8200 x g at 4 ºC and the protein content of the supernatants was determined using the Bradford method (1976).

For the isolation of yeast-binding proteins from hemolymph plasma, 50 ticks were injected with either 1 µl of sterile saline (S2), 5.4 µg β-1,3-glucan (B2) or heat killed

yeast (Y2) and the hemolymph collected from the ticks 2 hours post challenge. Hemolymph obtained from 50 non-injected ticks served as the unchallenged control (UN). For each of the four experimental conditions, the collected hemolymph (approximately 100 µl) was added to an equal volume of TBB (50 mM sodium hydrogen phosphate, 0.15 mM sodium chloride, 2 mM EDTA, 0.02% sodium azide, 2 mM phenylthiourea, protease inhibitor cocktail, pH 7.0) at 4 ºC. The hemolymph samples were centrifuged for 15 minutes at 8200 x g (4 ºC) to obtain plasma samples.

*Isolation of yeast- binding proteins from unchallenged and challenged hemolymph plasma samples*

The procedure followed was adapted from the method described by Zhu et al. (2005). For saturated cultures, *C. albicans* was grown overnight at 30ºC in a shaking incubator. Overnight cultures were washed once with 0.9% NaCl and twice with phosphate buffered saline (PBS) repeated centrifugation at 12000 x g for 2 minutes. The washed cells were re-suspended in a small volume of PBS (1/4 of original volume) and incubated with the unchallenged or challenged hemolymph samples. As a control, yeast cells were incubated with TBB alone. After incubation of samples with rotation for 10 minutes, the yeast cells were pelleted by centrifugation for 2 minutes at 12000 x g and washed twice with PBS. Bound proteins were eluted with 25 µl 0.15 M triethylamine, pH 11.5. The cells were pelleted and the eluted proteins were concentrated with acetone precipitation (1 part protein: 4 parts acetone) and re-suspended in 10 ul first-dimensional electrophoresis (1DE) reducing buffer (0.06 M Tris, pH 6.8, 2% SDS, 0.1% glycerol, 0.05% β-mercaptoethanol and 0.025% bromophenol blue). Samples were analyzed by 1DE according to the method described by Laemmli, (1970) using 12.5% polyacrylamide gels. Controls representing total hemolymph proteins (HL) and total yeast proteins (YC) were included. For the total hemolymph protein control, unchallenged hemolymph was used, while for the yeast protein control, heat killed yeast cells were sonified (Branson Sonic Power Company, USA), centrifuged at 8200 x g for 15 minutes and the collected supernatant used for analysis. Electrophoresed proteins were detected by staining gels either with MS-compatible silver stain (Yan et al., 2000) or colloidal Coomassie G-250 (Candiano et al., 2004) for MS/MS analysis.

*2DE: Isoelectric focusing (IEF)*

Hemolymph sample volumes were adjusted in order to analyze the same amount of protein (60 µg) for each set of experiments. Each sample was applied by including it in the rehydration solution (8M urea, 0.5% (w/v) CHAPS, 0.125% (v/v) IPG buffer). DTT was added to the rehydration solution prior to use to give a final concentration of 18 mM. Rehydration of 7 cm immobilized pH gradient strips (IPG-strips; non-linear pH gradient 3-10) and IEF were performed at 20 ºC using the Ettan IPGphor II system (Amersham Bioscience). After rehydration for 12 hours, IEF was performed at 100 V for 30 minutes, 500 V for 30 minutes, 1000 V for 30 minutes and 5 000 V for 1 hour 40 minutes. The current was limited to 50 µA/ IPG strip.

In a separate experiment, 60 µg of the 2D protein standards: conalbumin (MW= 76 kDa, pH 6.6), albumin (MW= 66 kDa, pH 5.5), actin (MW= 43 kDa, pH 5.1), glyceraldehydes-3-phospate dehydrogenase (MW= 36 kDa, pH 8.8), carbonic

anhydrase (MW= 31 kDa, pH 5.9), trypsin inhibitor (MW= 21.5 kDa, pH 4.5) and myoglobin (MW= 17.5 kDa, pH 7.0) were loaded onto an IPG strip as described above.

*2DE: SDS-PAGE and image analysis*

Focused IPG strips were subjected to reduction/alkylation prior to second dimensional electrophoresis. Strips were incubated for 15 minutes in an equilibration buffer containing 50 mM Tris-HCl (pH 8.8), 6 M Urea, 2% (w/v) SDS and 1% (w/v) DTT. The incubation was repeated for another 20 minutes with the same solution substituting the DTT with 2.5% (w/v) iodoacetamide. 2DE was carried out on 12.5% gels using a Hoefer mini vertical electrophoresis system (miniVE) at 60 V for 30 minutes and then 120 V until the bromophenol-blue front reached the bottom of the gel. Thereafter, the gels were placed into a fixing solution containing 30% ethanol, 10% acetic acid, before staining. The spots were detected with Flamingo fluorescent stain (BioRad) according to the manufacturer's instructions for quantitative analysis.

Each experimental condition was repeated four times. The 2D gels were scanned with a Pharos FX Plus Molecular imager (BIORAD). The four biological repeats were used to produce a master image for each of the three experimental conditions. Spots that were present in 2 or more of the gels were included in the respective master images. The master images obtained for the challenged hemolymph (aseptic injury and yeast-injected) were compared with the master image obtained for unchallenged hemolymph. Generation of the master images as well as statistical analysis of the modulated spots was performed using PDQuest$^{TM}$ software. Differentially expressed spots were analyzed using the Student's t- test ($p<0.05$).

*In-gel tryptic digestion of proteins*

Selected protein spots or bands were excised from gels for MS/MS analysis at the University of York (UK). Briefly, the gel pieces were washed twice with 50% (v:v) aqueous acetonitrile containing 25 mM ammonium bicarbonate, then once with acetonitrile and dried in a vacuum concentrator for 20 min. Sequencing-grade, modified porcine trypsin (Promega) was dissolved in the 50 mM acetic acid supplied by the manufacturer, then diluted 5-fold by adding 25 mM ammonium bicarbonate to give a final trypsin concentration of 0.01 µg/µl. The gel pieces were rehydrated with 10 µl of trypsin solution and covered with 25 mM ammonium bicarbonate solution for 30 minutes. These digests were incubated overnight at 37$^{o}$C.

*Matrix-assisted laser desorption ionization- time-of-flight mass spectrometry (MALDI-TOF/TOF) and database searches*

For the MS analysis 1 µl aliquot of each fraction was applied to the ground steel MALDI target plate. An equal volume of freshly prepared 5 mg/ml 4-hydroxy-α-cyano-cinnamic acid ([Sigma, in 50% aqueous (v:v), acetonitrile containing 0.1% trifluoroacetic acid (v:v)] was then added to the fractions. A positive-ion MALDI-MS spectra was obtained (Bruker Ultraflex III in reflectron mode, equipped with a Nd:YAG smart beam laser) and the MS spectra were acquired over a m/z range of 800-4000 (peptide mass fingerprints). The final mass spectra were externally calibrated against an adjacent spot containing 6 peptides (des-Arg$^{1}$-bradykinin,

904.681; angiotensin I, 1296.685; Glu$^{1}$-fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929.). Monoisotopic masses were obtained using a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) and a S/N threshold of 2.

The ten strongest peaks of interest for each spot were then selected for MS/MS fragmentation (signal to noise ratio greater than 10). The fragmentation was performed without the introduction of a collision gas. A default calibration was used for the MS/MS spectra, which were first baseline-subtracted and smoothed (Savitsky-Golay, width 0.15 m/z, cycles 4); monoisotopic peak detection used a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) with a minimum S/N of 3. Bruker FlexAnalysis software was used to perform the spectral processing and peak list generation for both the MS and MS/MS spectra.

MS/MS ion searches of the NCBInr database were performed using a locally-running copy of the Mascot program (Matrix Science Ltd., version 2.1), through the Bruker BioTools interface (version 3.1). Search criteria included: carbamidomethyl (C) as a fixed modification for all alkylated samples, enzyme, trypsin; variable modifications, oxidation (M); peptide tolerance, 200 ppm; MS/MS tolerance, 0.8 Da; instrument, MALDI-TOF-TOF.

For the *de novo* sequence interpretations from the tandem mass spectra PredictSequence routine (BioMultiview 1.4 software, MDS Sciex) was used. The complete and partial peptide sequences obtained were then searched with MS BLAST using PAM30 matrix and the nrdb95 database (NCBI). The search was performed with the WU-BLAST2 program (Gish W (1996-1999) http://blast.wustl.edu) on an EMBL server http://dove.embl-heidelberg.de/Blast2/.

Additionally *de novo* sequences were compared with virtual translations of the sequences deposited in the current tick EST databases at the Gen Index databases (http://compbio.dfci.harvard.edu/tgi/), namely, *A. variegatum*, *R. microplus*, *Ixodus scapularis* and *Rhiphicephalus appendiculatis* using the tBLASTn program.

## Results and discussion

*Comparative 2D-gel analysis of hemolymph proteins*

Previous 1DE of hemolymph plasma proteins (results not shown) indicated that the most abundant proteins have molecular weights above 76 kDa, while fewer proteins were observed in the molecular weight range of 14-66 kDa. The fact that large (~ 200 kDa) proteins, hydrophobic proteins and basic (pI > 8) proteins can all precipitate during IEF and are poorly resolved by 2DE (Yarmush and Jayaraman, 2002) may explain why protein spots > 76 kDa were absent from 2D gels. In this study the 2D comparative analysis was thus restricted to proteins in the molecular weight range of 17.5-76 kDa.

To determine which proteins are differentially expressed in the hemolymph, the profile obtained for unchallenged hemolymph (UN) was compared to the profiles obtained for ticks injected with saline (S2) or heat killed *C. albicans* (Y2),

respectively. Differentially expressed spots ($p<0.05$) are indicated on the 2DE master images (Fig. 1) as determined by PDQuest. Two proteins were found to be down-regulated in the saline control, while four protein spots were up-regulated and one protein spot down-regulated in the hemolymph of ticks challended with yeast (Table. 1). The modulated proteins are all acidic proteins (pI 4.8–6.4) and their molecular weights range from 22–64 kDa. The proteins modulated by yeast challenge can be assigned to the immune response as they were not induced after aseptic injury.

It was previously observed that inoculation of *O. savignyi* ticks with either heat killed Gram-positive or Gram-negative bacteria increased the Gram-positive antibacterial activity in the hemolymph up to 2 hours after which the activity decreased (results not shown). Based on the above observation hemolymph was collected 2 hours post challenge as hemolymph proteome changes would identify proteins involved in the tick's immediate response to challenge before the induction of biosynthesis of defense peptides and proteins. The five modulated proteins analyzed in this study are most likely involved in the very first line of the tick's defense against yeast. The four up-regulated proteins are probably released into the hemolymph plasma from storage sites in either hemocytes or from the fat body. The decrease in the level of the one protein suggests that due to its involvement in the immediate response it is rapidly depleted. In a differential proteomic analysis of *Drosophila* hemolymph, Levy et al. (2004) found that at least 70 proteins were up- or down-regulated after a bacterial or fungal challenge. These studies were performed on hemolymph collected 6 hours and 72 hours, post bacterial and fungal challenge, respectively. Maximum alterations in protein expression occurred at these selected time points. However, similar to the results obtained in our study, Vierstraete et al. (2004) identified only 10 differentially expressed proteins in *Drosophila* hemolymph 25 minutes after challenge with LPS, suggesting that these proteins are involved in the very early defense response of the fruit fly.

A total of five spots (no 3-7, Fig. 1b) were subjected to MS/MS analysis. Peptide mass fingerprint searches were not performed, instead the MS/MS ion spectra data obtained for each protein were used to search for homologous proteins using MASCOT software. The results (Table. 2) show that MS/MS ion searches failed to positively identify any of the five modulated proteins. For proteins not available in databases, it is necessary to determine partial or complete amino acid sequences (Shevchenko et al., 2001; Samyn et al., 2006). *De novo* sequencing based on MS/MS data provided amino acid sequence data for the proteins (Table. 3). BLAST searching of the NCBInr database with these sequences also failed to identify these proteins. For instance, in the case of spot 3, each of the four derived amino acid sequences for this protein matched a different protein in the database. No significant alignments were observed between the deduced peptide sequences of the soft tick, *O. savignyi*, and the sequences in the tick EST databases. The fact that the proteins in this study do not match the sequences deposited in databases for salivary gland proteins of *O. parkeri* and *O. coriaceus* (Francischetti et al., 2008a, 2008b), suggests that these proteins are most likely not expressed in salivary glands. The limited number of tick sequences available in databases is a serious drawback to the identification of tick proteins using proteomic approaches. Rachinsky et al. (2007) investigated the proteins that were differentially expressed in ovarian tissues collected from *Babesia bovis*-infected and uninfected *R. microplus* and could only identify 19 of their 43 differentially expressed proteins. Similarly, Oleaga et al. (2007)

could only identify 27 spots out a total of 97 analyzed in the salivary glands of *O. moubata* and *O. erraticus*.

In this study 7 cm strips were employed because hemolymph samples were extremely limited in quantity. A higher protein load using longer strips would have resulted in more spots and could have improved protein identification. Another point worth mentioning is the presence of high- abundance proteins of high molecular weight present in hemolymph samples which prevented optimal focusing and limited loading capacity of low-abundance proteins. In an attempt to remove high-abundance proteins initial hemolymph sample preparation included filtration of samples through a size-exclusion filter (cut-off,100 kDa). 1DE analysis of the filtrate revealed that the filter excluded all proteins of molecular weight above 30 kDa (results not shown). This additional sample preparation step was thus omitted as it resulted in substantial protein loss and was unsuitable for enriching the proteins of interest.

*Yeast binding proteins from hemolymph*

Due to the limitations associated with 2DE, alternative approaches are required to detect additional hemolymph proteins involved in the immune response. *C. albicans* cells were used as affinity beads to adsorb proteins from the hemolymph plasma. The strategy used in our study was based on the method described by Zhu et al. (2005) who were able to identify a functional homolog of vertebrate complement 3 in the hemolymph of the horseshoe crab, *Carcinoscorpius rotundicauda*, using Gram-positive *Staphylococcus aureus*. Using a similar strategy Koizumi et al. (1997) isolated a protein involved in the clearance of *E. coli* from the larval hemolymph of the silkworm *Bombyx mori*.

In these experiments, yeast cells were incubated with unchallenged hemolymph (UN) and with hemolymph collected from ticks two hours after injection with saline (S2) or β-1,3-glucan (B2) or heat killed yeast (Y2). In each case the bound proteins were eluted from the yeast cells with triethylamine, pH 11.5, and analyzed with 1DE (Fig. 2). The results showed that several proteins present in unchallenged hemolymph as well as hemolymph collected from both injured and yeast-challenged ticks bound to the yeast cells. These proteins were not present when yeast cells were incubated with buffer alone, confirming that the eluted proteins were of hemolymph origin. The 97 kDa and 88 kDa proteins correspond to high-abundance proteins present in the hemolymph plasma control and may be binding to yeast non-specifically. However, the 26 kDa protein showed a clear differential pattern and appeared mainly in the hemolymph fraction collected after β-1,3-glucan injection. For these experiments, hemolymph was collected directly into TBB followed by centrifugation to remove the hemocytes. It can thus be assumed that the plasma samples prepared in this manner represent mainly plasma proteins as lysis of hemocytes was prevented. One can speculate that these proteins may play an important role in the recognition and removal of yeast cells in tick plasma. They are possibly pattern-recognition proteins that specifically bind to β-1,3-glucans, conserved surface determinants, of fungi. Similar experiments using Gram-negative *Esherichia coli* as affinity beads resulted in the analysis of two different high molecular weight bacteria-binding proteins from tick hemolymph (results not shown). However, these Gram-negative binding proteins were only detected in hemolymph after either aseptic or bacterial challenge. These

preliminary findings suggest that tick hemolymph plasma contains different proteins involved in pathogen recognition, some being permanently present while others are released following either aseptic or septic activation of the immune response.

For their identification, the experiment was repeated and the three proteins (97, 88 and 26 kDa) were excised from a colloidal Coomassie stained 1DE gel and subjected to MS/MS analysis. MASCOT searches using the MS/MS ion spectra (Table. 4) as well as BLAST searches using derived *de novo* sequences (Table. 5), failed to positively identify any of the *C. albicans* binding proteins.

## Conclusion

This study enabled the analysis of hemolymph proteins involved in the immediate immune response of the tick when challenged with *C. albicans*. The proteins analyzed by differential expression are released from their storage sites and are most likely involved in the very first line of the tick's defense against yeast. These proteins can be assigned to the immune response as they were not induced after aseptic injury. Using the yeast cells as affinity beads allowed the analysis of additional hemolymph proteins involved in the yeast-response, not detected with differential expression analysis. For these investigations a combination of strategies is therefore necessary in order to obtain the maximum amount of information on the proteins involved in a process as complex as innate immunity.

None of the proteins analyzed in this study could be identified. The limited number of tick protein sequences available in databases is a drawback to the identification of tick hemolymph proteins using a proteomics approach. However, the number of sequence information is rapidly growing and the MS/MS data as well as *de novo* sequences obtained for these proteins may be archived and used to search databases in the future. Until then, *de novo* peptide sequences may be used for the design of degenerate primers for cloning of the genes of the corresponding proteins by a PCR-based approach. RNA interference studies can be performed to investigate the functional role of these proteins in tick innate immunity. Such investigations are in progress.

In conclusion, results obtained from these and similar studies will lead to a better understanding of tick innate immunity. Moreover, information obtained from structural and functional characterization of tick proteins involved in pathogen defence may be applied in future for the design of novel anti-infective agents.

## Acknowledgements

We acknowledge the National Research Foundation of South Africa and the University of Pretoria for financial support. We thank the University of York, Department of Biology for MS/MS analysis of proteins as well as C. Gains from the University of York for her assistance in the interpretation of the MS/MS data.

## References


Bradford, MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248-254.

Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri G.M, Carnemolla B, Orecchia P, Zardi L and Righetti PG (2004). Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis. Electrophoresis 25: 1327-1333 doi:10.1002/elps.2002305844

De Morais Guedes S, Vitorino R, Torner K, Domingues MRM, Correia AJF, Amado F and Domingues P (2003). *Drosophila melanogaster* larval hemolymph protein mapping. Biochem Biophys Res Commun 312: 545-554 doi:10.1016/j.bbrc.2003.10.156

Estrada-Pena A and Jongejan F (1999). Ticks feeding on humans: a review of records on human-biting *Ixodidea* with special reference to pathogen transmission. Exp Appl Acarol 23: 685-715 doi:10.1023/A:1006241108739

Francischetti IBM, Meng Z, Mans BJ, Guderra N, Hall M, Veenstra TD, Pham VM, Kotsyfakis M and Ribeiro JMC (2008a) An insight into the salivary transcriptome and proteome of the soft tick and vector of epizootic bovine abortion, *Ornithodoros coriaceus*. J Proteomics 71: 493-512 doi:10.1016/j.jprot.2008.07.006

Francischetti IBM, Mans BJ, Meng Z, Guderra N, Veenstra TD, Pham VM and Ribeiro JMC (2008b). An insight into the sialome of the soft tick, *Ornithodoros parkeri*. Insect Biochem Mol Biol 38: 1-21 doi:10.1016/j.ibmb.2007.09.009

Howell CJ, Neitz AWH and Potgieter DJJ (1975). Some toxic and chemical propterties of the oral secretion of the sand tampan, *Ornithodoros savignyi* Audouin (1825). *Onderstepoort J Vet Res* 43: 99-102 PMID:1196584

Imler JL and Hoffmann JA (2000). Signaling mechanisms in the antimicrobial host defence of *Drosophila.* Curr Opin Microbiol 3: 16-22 doi:1061016/S1369-5274(99)00045-4

Koizumi N, Morozumi A, Imamura M, Tanaka E, Iwahana H and Sato R (1997). Lipopolysaccharide-binding proteins and their involvement in the bacterial clearance from the hemolymph of the silkworm *Bombyx mori*. Eur J Biochem 248: 217-224 doi:10.111/j.1432-1033.1997.ta-1-00217.x

Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685 PMID:5432063

Levy F, Bulet P and Ehret-Sabatier L (2004). Proteomic analysis of the systemic immune response of Drosophila. Mol Cell Proteomics 3: 156-166 doi:10.1074/MCP.M300114-MCP200

Madden RD Sauer JF and Dillwith JW (2002). A proteomics approach to characterizing tick salivary secretions. Exp Appl Acarol 28: 77-87 doi 10.1023/B:APPA.0000018241.81636.91

Mans BJ and Neitz AWH (2003) Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. Insect Biochem Mol Biol 34: 1–17 doi:10.1016/j.ibmb.2003.09.002

Mans BJ, Venter JD, Coons LB, Louw AI and Neitz AWH (2004). A reassessment of argasid tick salivary gland ultrastructure from an immuno-cytochemical perspective. Exp Appl Acarol 33: 119-129 doi:10.1023/B:APPA.0000030012.47964.b3

Medzhitov R and Janeway Jnr, CA (2002). Decoding the patterns of self and nonself by the innate immune system. Science 296: 298-300 doi:10.1126/science.1068883

Oleaga A, Escudero-Poblacion A, Camafeita E, and Perez-Sanchez R, (2007). A proteomic approach to the identification of salivary proteins from the argasid ticks *Ornithodoros moubata* and *Ornithodoros erraticus,* Insect Biochem Mol Biol 37: 1149-1159 doi:10.1016/j.ibmb.2007.07.003

Paton WS and Evans AM (1929). Insects, ticks, mites and venomous animals of medical and veterinary importance, Part 1. HR Grubb LTD. Croydon Great Britain, pp 628-630 doi:10.1016/j.ibmb.2007.07.003

Rachinsky A, Guerro FD and Scoles GA (2007). Differential protein expression in ovaries of uninfected and *Babesia*-infected southern cattle ticks, *Rhipicephalus (Boophilus) microplus*. Insect Biochem Mol Biol 37: 1291-1308 doi:10.1016/j.ibmb.2007.08.007

Ranchinsky A, Guerro FD and Scoles GA (2008). Proteomic profiling of *Rhipicephalus (Boophilus) microplus* midgut response to infection with *Babesia bovis*. Vet Parasitol 152: 294-313 doi:10.1016/j.vetpar.2007.12.027

Samyn B, Sergeant K, Memmi S, Debyser G, Devreese B and Van Beeumen J (2006). MALDI-TOF/TOF *de novo* sequence analysis of 2-D PAGE-separated proteins from *Halirhodospira halophila*, a bacterium with unsequenced genome. Electrophoresis 27: 2702-2711 doi:10.1002/elps.200500959

Shevchenko A, Sunyaev S, loboda A, Shevchenko A, Bork P, Ens W and Standing KG (2001). Charting the proteomes of organisms with unsequenced genomes by MALDI-Quadrople Time-of-Flight Mass Spectrometry and BLAST homology searching. Anal Chem 73: 1917-1926

Sonenshine DE and Hynes WL (2008). Molecular characterization and related aspects of the innate immune response in ticks. Front Biosci 13: 7046-7063 PMID:18508715

Taylor D (2006). Innate immunity in ticks: A review, J Acarol Soc Jpn 15 (2): 109-127 ISSN:0918-1067

Untalan PM, Guerrero FD, Haines LR and Pearson TW (2005). Proteome analysis of abundantly expressed proteins from unfed larvae of the cattle tick, *Boophilus microplus*. Insect Biochem Mol Biol 35: 141-151 doi:10.1016/j.ibmb.2004.10.009

Vierstraete E, Cerstiaens A, Baggerman G, Van den Bergh G, De Loof A and Schoofs L (2003). Proteomics in *Drosophila melanogaster*: first 2D database of larval hemolymph proteins. Biochem Biophys Res Commun 304 (4): 831-838 doi:10.1016/S0006-291x(03)00683-1

Vierstraete E, Verleyen P, Sas F, van den Bergh G, De Loof A, Arckens L and Schoofs L (2004). The instantly released *Drosophila* immune proteome is infection-specific. Biochem Biophys Res Commun 317: 1052-1060 doi:10.1016/j.bbrc.2004.03.150

Vennestrøm J and Jensen PM (2007). *Ixodes ricinus*: The potential of two-dimensional gel electrophoresis as a tool for studying host-vector-pathogen interactions. Exp Parasitol 115: 53-58 doi:10.1016/j.exppara.2006.05.008

Wickramasekara S, Bunikis J, Wysocki V and Barbour AG (2008). Identification of residual blood proteins in ticks by mass spectrometry proteomics. Emerg Infect Dis 14: 1273-1275 doi:10.3201/eid1408.080227

Yan JX, Wait R, Berkelman T, Harry RA, Westbrook JA, Wheeler CH and Dunn MJ (2000). A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. Electrophoresis 21: 3666-3672 doi:10.1002/1522-2683(200011)21:17<3666>3.0.C.O;2-6

Yarmush M and Jayaraman A (2002). Advances in proteomic technologies. Annu Rev Biomed Eng 4: 349-373 doi:10.1146/annurev.bioeng.4.020702.153443

Zhu Y, Thangamani S, Ho B and Ding JL (2005). The ancient origin of the complement system. EMBO J 24: 382-394

## List of figures

**Fig. 1** Comparative 2D- gel analysis of hemolymph proteins. Comparison of the master image obtained for hemolypmh proteins from unchallenged ticks with (a.) the image for proteins from saline-injected ticks and (b.) the image for proteins from yeast injected ticks. Each experimental condition was repeated four times. The four biological repeats of each experimental condition were used to produce a master image and this was compared to the master image from the four biological repeats obtained from unchallenged ticks, respectively. Spots that are present in 2 or more of the gels are included in the master image. The numbered dark spots are differentially expressed ($p<0.05$) as determined by PDQuest software.

**Fig. 2** 1D- gel analysis of hemolymph proteins bound to *C. albicans*. Yeast cells were incubated with only buffer (C), unchallenged hemolymph (UN) or hemolymph collected 2 hours after injecting ticks with either saline (S2), β-1,3-glucan (B2) or heat-killed yeast (Y2). Eluted proteins were precipitated with acetone, re-suspended and analyzed on a 12.5% gel. In each case 10 ul of protein was loaded and visualized by silver staining. Controls representing total hemolymph plasma proteins (HL) and total yeast proteins (YC) were included. The molecular masses of the standard proteins are on the left.

## List of tables

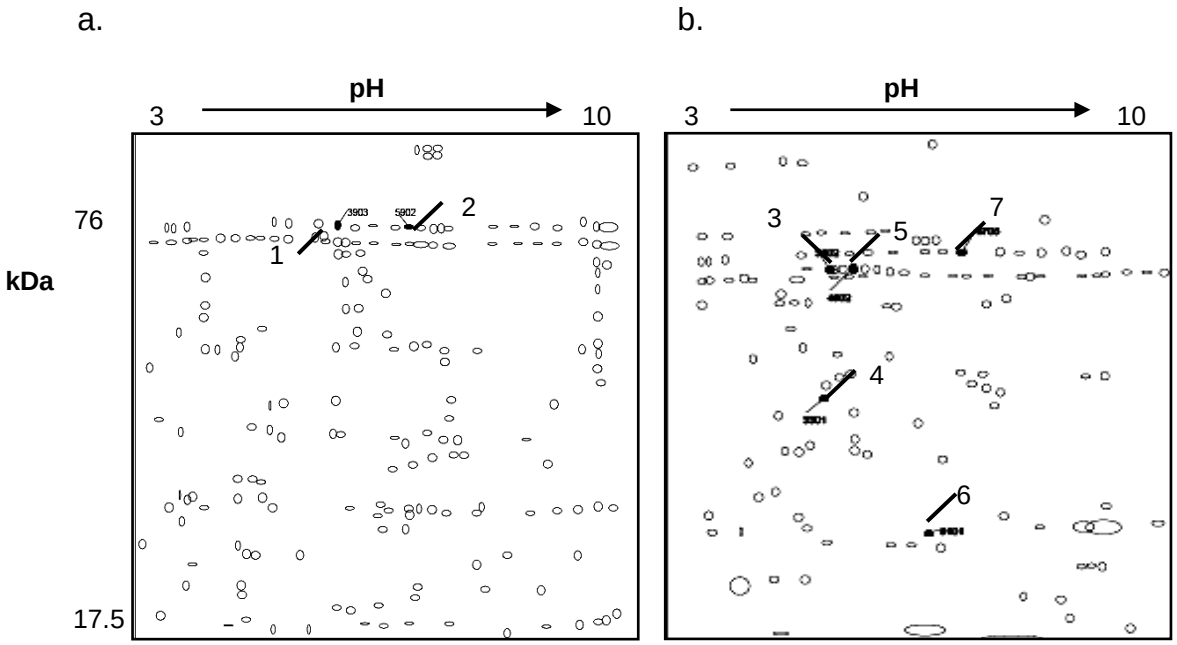


**Fig. 1**

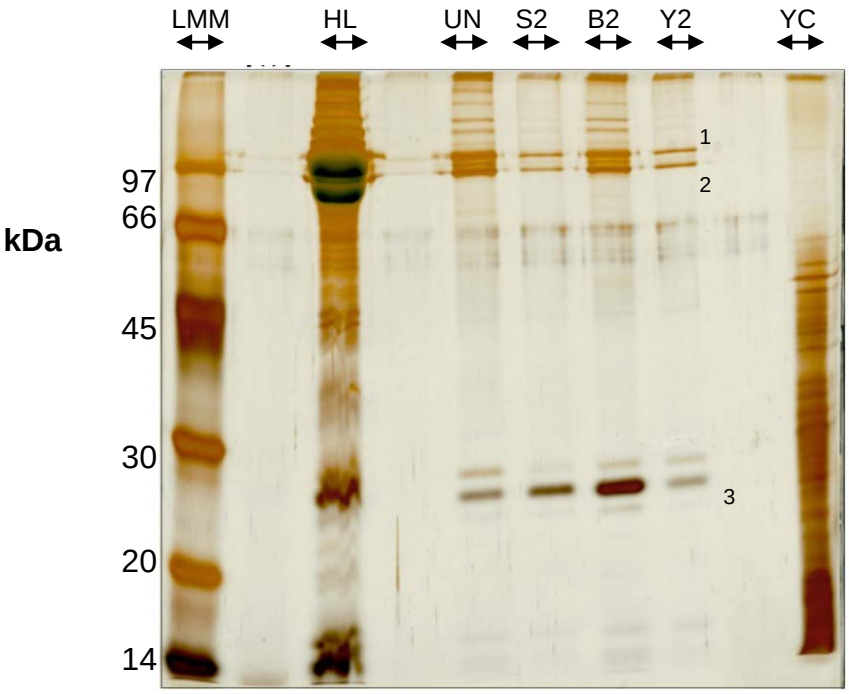


**Fig. 2**

**Table. 1**

| Unchallenged vs. Challenged | Up-regulated spots | Down-regulated spots | pI[a] | MM (kDa)[a] |
|---|---|---|---|---|
| Saline | | 1 | 5.4 | 54 |
| | | 2 | 6.2 | 54 |
| Yeast | 3 | | 4.9 | 53 |
| | 4 | | 4.8 | 35 |
| | | 5 | 5.2 | 53 |
| | 6 | | 6.0 | 22 |
| | 7 | | 6.4 | 64 |

a. Experimental values

**Table. 2**

| Spot | Hit | Protein[a] | Accession number | Peptide sequence[b] | Observed mass (M+H+)/ Da | Theoretical mass (M)/ Da | Mass error/ ppm | Mowse Score | Expect Score |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | Conserved hypothetical protein (*Pseudomonas aeruginosa*) | gi\|187939647 | R.NHAIADAVLLSGDEDVR.V | 1794.8259 | 1793.8187 | -40.20 | 25 | 22 |
| | 2 | Phytanol-CoA diowygenase (*Solibacter usitatus* Ellin6076) | gi\|116622764 | M.SNFIPSAEDLHR.V | 1385.7280 | 1384.7208 | 34.1 | 25 | 32 |
| | 3 | Methylase, putative (*Chlorobium tepidum TLS*) | gi\|21673792 | -.MQIHAGRYR.G + Oxidation (M) | 1147.5603 | 1146.5530 | -16.25 | 18 | 1.8e+02 |
| | 4 | General secretion pathway protein K (*Desulfuromonas acetocidans DSM 684*) | gi\|95931262 | M.SVFFAKKPLGNER.G | 1492.7182 | 1491.7109 | -72.98 | 18 | 1.5e+02 |
| 4 | 1 | Cantactin-3 precursor (Brain-derived immunoglobulin superfamily protein 1, BIG-1) Plasmacytoma-associated neuronal glycoprotein) (*Monodelphis domestica*) | gi\|126336353 | R.EPSNNIFPVGSEEK.K | 1546.7206 | 1545.7133 | -11.50 | 24 | 32 |
| | 2 | Hypothetical protein FBALC1_16342 (*Flavobacteriales bacteruim ALC-1*) | gi\|163786772 | K.HLEPETSFR.S | 1115.6164 | 1114.6091 | 61.3 | 25 | 35 |
| | 3 | Sensor histidine kinase, putative (*Oceanibulbus indolifex HEL-45*) | gi\|163746438 | R.RDEATRPFSR.L | 1234.6850 | 1233.6777 | 45.7 | 18 | 1.5e+02 |
| | 4 | Hypothetical protein PFL_4711 (*Pseudomonas fluorescens Pf-5*) | gi\|70732040 | K.ALMLNRMLITGSR.E | 1475.7624 | 1474.7551 | -38.02 | 18 | 1.4e+02 |
| 5 | 1 | To-component sensor histidine kinase (*Bacillus cereus B4264*) | gi\|168134092 | R.HLEDRILVK.R | 1122.6266 | 1121.6193 | -32.41 | 38 | 1.4 |
| | 2 | Unnamed protein product *(Tetraodon nigroviridis*) | gi\|47226843 | R.HLEILGYDR.V | 1115.6024 | 1114.5952 | 16.2 | 28 | 16 |
| | 3 | Polyphosphate kinase (*unidentified Eubacterium SCB49*) | gi\|149371389 | K.TNYALIEIPR.S | 1189.6152 | 1188.6079 | -35.58 | 28 | 19 |
| | 4 | Hypothetical protein, conserved in *Plasmodium* species (*Plasmodium knowlesi strain H*) | gi\|193807613 | K.MYNYLMKLLDR.E + 2 Oxidation (M) | 1491.7352 | 1490.7279 | 1.20 | 28 | 28 |
| 6 | 1 | Predicted protein (*Aspergillus terreus NIH2624*) | gi\|115388339 | R.LHELYPGMR.Q | 1115.6256 | 1114.6183 | 52.9 | 32 | 7.1 |
| | 2 | Thiamine monophosphate kinase (*Idiomarina loihiensis L2TR*) | gi\|56461240 | R.LHYPTPR.V | 883.4456 | 882.4383 | -37.26 | 29 | 10 |
| | 3 | Hypothetical protein BPSS0216 (*Burkholderia pseudomallei K96243*) | gi\|53721251 | R.LLVGAQPGR.A | 910.5931 | 909.5859 | 50.9 | 25 | 10 |
| | 4 | Katanin p60 subunit A-like 1 (*Gallus gallus*) | gi\|50730993 | R.KESPGLQPR.G | 1011.5299 | 1010.5226 | -27.93 | 25 | 35 |
| 7 | 1 | Conserved hypothetical protein (*Comamonas testosteroni KF-1*) | gi\|118053819 | R.DLLVQDVINMMER.H | 1575.8525 | 1574.8452 | 41.6 | 37 | 1.6 |
| | 2 | Iron permease FTR1 (E*xiguobacterium sibiricum 255-15*) | gi\|172057621 | R.QTGLAMNLAGR.S + Oxidation (M) | 1147.5370 | 1146.5298 | -45.16 | 36 | 2.6 |
| | 3 | Hypoxia associated factor (*Tribolium castaneum*) | gi\|91090962 | R.YTGPIMEFK.E | 1085.5952 | 1084.5879 | 56.8 | 30 | 9.4 |
| | 4 | Hypothetical protein LOC780085 *(Xenopus tropicalis*) | gi\|118404668 | K.STMERNLVLEER.S + Oxidation (M) | 1492.6934 | 1491.6861 | -32.85 | 30 | 8.8 |

a. Protein to which the peptide sequence was matched with MASCOT

b. Sequence information obtained for matched peptide from tandem mass spectrometry as determined by searching with MASCOT

**Table. 3**

| Spot no | Hit | Score[a] | Accession number | Peptide sequence[b] | Protein name[c] |
|---|---|---|---|---|---|
| 3 | 1 | 110 | gi\|89200423\|ref\|ZP_01179176.1\| | 451 DTFRPHIWNAV 461 | Membrane bound O-acyl transferase, MBOAT (*Bacillus cereus* subsp.) |
| | 2 | 108 | gi\|118789505\|ref\|XP_317464.3\| | 225 WDMFVQDVTM 234 | ENSANGP00000012356 (*Anopheles gambiae* str. PEST) |
| | 3 | 106 | gi\|108760113\|ref\|YP_633239.1\| | 48 FTWTASPPR 56 | Putative phosphomethylpyrimidine kinase (*Myxococcus xanthus* ) |
| | 4 | 90 | gi\|473161\|emb\|CAA81555.1\| | 178 TTWRSST 184 | Microneme antigen (*Sarcocystis muris*) |
| 4 | 1 | 99 | gi\|125806480\|ref\|XP_001346220.1\| | 572 YHNYNDNYYC 581 | Hypothetical protein (*Danio rerio)* |
| | 2 | 96 | gi\|125846263\|ref\|XP_001342573.1\| | 86 YYYYDNNYDY 95 | Hypothetical protein (*Danio rerio)* |
| | 3 | 88 | gi\|124802919\|ref\|XP_001347634.1\| | 318 SDAEHYND 325 | Hypothetical protein PF10_0350 (*Plasmodium falciparum* 3D7) |
| | 4 | 88 | gi\|110800410\|ref\|YP_695004.1\| | 105 SNWYYS 110 | Cell wall binding repeat domain protein (*Clostridium perfringens*) |
| 5 | 1 | 127 | gi\|115609728\|ref\|XP_001185319.1\| | 155 EGHLDVVECL 164 | Similar to ankyrin 2,3/unc44 (*Strongylocentrotus purpuratus*) |
| | 2 | 95 | gi\|57238755\|ref\|YP_179891.1\| | 214 AHDIEHAV 221 | Hypothetical protein Erum0250 (*Ehrlichia ruminantium* str) |
| | 3 | 93 | gi\|118752713\|ref\|ZP_01600528.1\| | 96 EGHMDVEAML 105 | Conserved hypothetical protein (*Metallosphaera sedula*) |
| 6 | 1 | 73 | gi\|47496945\|dbj\|BAD20015.1\| | 287 AGDEQCRRWAT 297 | Hypothetical protein [*Oryza sativa* (japonica cultivar-group)] |
| | 2 | 69 | gi\|88806673\|ref\|ZP_01122190.1\| | 1 MNYRRTNPTGI 11 | Hypothetical protein RB2501_01336 (*Robiginitalea biformata*) |
| | 3 | 68 | gi\|119356379\|ref\|YP_911023.1\| | 479 NRIGYKRLLGSGV 491 | ABC-type branched-chain amino acid transport systems periplasmic component-like (Chlorobium phaeobacteroides) |
| 7 | 1 | 112 | gi\|70950671\|ref\|XP_744639.1\| | 121 SFVTVIMPYLTVTM 134 | Hypothetical protein (*Mus musculus*) |
| | 2 | 75 | gi\|94369673\|ref\|XP_996645.1\| | 45 YTTLMDIMITTMT 57 | Hypothetical protein PC103780.00.0 (*Plasmodium chabaudi chabaudi)* |

a. Score (high scoring pair, HSP) as determined by BLAST, using sequences obtained from *de novo* sequencing. Matches higher than 64 are considered significant

b. Peptide sequence of the *de novo* data (only sequences with highest scores are given). The numbers give the residues to which the query sequence matched to the protein sequence in the database

c. Identity of the proteins to which the *de novo* sequences were matched

**Table. 4**

| Protein band | Hit | Protein[a] | Accession number | Peptide sequence[b] | Observed mass (M+H+)/ Da | Theoretical mass (M)/ Da | Mowse score | Expect Score |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | Peptidase M48, Ste24p [*Burkholderia* sp. 383] | gi\|78067922 | R.LPDLMPLYAK.A | 1176.6140 | 1175.6067 | 56 | 0.025 |
| | 2 | Hypothetical protein Bxe_A4392 [*Burkholderia xenovorans* LB400] | gi\|91781454 | R.LPDMLPLYAK.A | 1176.6140 | 1175.6067 | 54 | 0.035 |
| | 3 | Gamma-butyrobetaine,2-oxoglutarate dioxygenase, putative [uncultured bacterium 578] | gi\|40063312 | R.LVSELVAYR.A | 1049.6051 | 1048.5978 | 51 | 0.06 |
| | 4 | Cytochrome P450 monooxygenase [*Nocardia farcinica* IFM 10152] | gi\|54026491 | R.LVSELVMSR.S | 1049.6051 | 1048.5978 | 51 | 0.06 |
| 2 | 1 | Similar to Katanin p60 subunit A-like 1 [*Gallus gallus*] | gi\|50730993 | R.KESPGLQPR.G | 1011.4938 | 1010.4865 | 40 | 0.79 |
| | 2 | vacJ lipoprotein [*Methylococcus capsulatus* str. Bath] | gi\|53803895 | K.QEGPEAGAPR.T | 1011.4938 | 1010.4865 | 40 | 0.79 |
| | 3 | hypothetical protein Bpse1_03001421 [*Burkholderia pseudomallei* 1655] | gi\|99912487 | K.KADSLPKPR.Y | 1011.4938 | 1010.4865 | 40 | 0.79 |
| | 4 | NT02FT1147 [synthetic construct] | gi\|54114569 | R.IHEILTAVVK.A | 1122.6151 | 1121.6078 | 39 | 0.92 |
| 3 | 1 | Hypothetical protein LOC735239 | gi\|148233280 | R.TPTISKER.T | 931.5029 | 930.4956 | 45 | 0.36 |
| | 2 | Hypothetical protein Bd2369 | gi\|42523815 | K.TPVDSKER.N | 931.5029 | 930.4956 | 44 | 0.44 |
| | 3 | Hypothetical protein MGG_13328 | gi\|145609013 | R.TPLTSQER.A | 931.5029 | 930.4956 | 43 | 0.58 |
| | 4 | Predicted protein | gi\|156384031 | -.MAARLPTK.A | 903.5832 | 902.5759 | 41 | 0.39 |

a. Protein to which the peptide sequence was matched with MASCOT

b. Sequence information obtained for matched peptide from tandem mass spectrometry as determined by searching with MASCOT

**Table 5**

| Band | Hit | Score[a] | Accession number | Peptide sequence[b] | Protein name[c] |
|---|---|---|---|---|---|
| 1 | 1 | 107 (70 +37) | trembl\|AACY01484588\|AACY01484588_1 | 69 DYFVRMTNVKVAFL 82<br>120 FLWVY 124 | Unknown |
| | 2 | 72 | sptrembl\|Q9F0C0\|Q9F0C0 | 144 KPTFCWNSPFLHP 156 | Beta(1,3)galactosyltransferase (*Lactobacillus delbrueckii*) |
| | 3 | 70 | sptrembl\|Q62LI9\|Q62LI9 | 26 AGTDIPNNIAANMIP 40 | N-formylglutamate amidohydrolase (EC3.5.1.68) (*Burkholderia Mallei*) |
| | 4 | 69 | sptrembl\|Q7UJE7\|Q7UJE7 | 185 MEYLFWDER 193 | Hypothetical protein.//:sptrembl\|Q7UJE7\|Q7UJE7_RHOBA |
| | 5 | 67 | trembl\|X04942\|HSTCRA08_1 | 51 ELFWYVNYAR 60 | Human mRNA for T-cell receptor alpha-chain HAVP36 |
| | 6 | 67 | trembl\|Z47370\|HSXPF10A2_1 | 46 ELFWYVNYAR 55 | TCRAV9S1"; product: "T-cell receptor alpha-chain V-region (V-J-C) |
| 2 | 1 | 69 | sptrembl\|Q8S638\|Q8S638 | 42 WRWFGLHG 49 | Hypothetical protein OSJNBa0004P12.8 (*Oryza sativa* (japonica cultivar-group)) |
| | 2 | 69 | sptrembl\|O32454\|O32454 | 29 WPELFDEY 36 | ORF.//:sptrembl\|O32454\|O32454_9ACTO ORF (*Actinomadura hibisca*) |
| | 3 | 68 | trembl\|AACY01070275\|AACY01070275_2 | 4 YLVFKLLHVLF 14 | Unknown |
| | 4 | 65 | trembl\|AAEK01000004\|AAEK01000004_8 | 34 VHKDVEDFE 42 | Conserved hypothetical protein (*Bacillus cereus*) |
| | 5 | 65 | trembl\|AAEK01000004\|AAEK01000004_173 | 21 VHKDVEDFE 29 | Conserved hypothetical protein (*Bacillus cereus*) |
| | 6 | 65 | swiss\|Q9Z5C9\|NRDI_STAAM | 42 LVTGHIGFGEVEE 54 | NrdI protein.//:sptrembl\|Q6GBA0\|Q6GBA0 Hypothetical protein (*Staphylococcus aureus*) |
| | 7 | 64 | sptrembl\|Q9LP94\|Q9LP94 | 73 SEVIFDEYYR 82 | T32E20.26.//:sptrembl\|Q9LP94\|Q9LP94_ARATH T32E20.26. (*Arabidopsis thaliana*) |
| 3 | 1 | 71 | trembl\|AACY01071833\|AACY01071833_2 | 373 AGIEFNTIVGAHIKK 387 | Unknown |
| | 2 | 70 | sptrembl\|Q61M93\|Q61M93_CAEBR | 1026 YLAKLGTVEH 1035 | Hypothetical protein CBG08584 |
| | 3 | 69 | sptrembl\|Q5WNK3\|Q5WNK3_CAEBR | 232 IEYTAFPAV 240 | Hypothetical protein CBG07969 |
| | 4 | 67 | trembl\|AACY01759246\|AACY01759246_1 | 213 GRLPNTLVATLK 224 | Unknown |
| | 5 | 67 | sptrembl\|Q83YL6\|Q83YL6 | 137 TAEADITHR 145 | Putative transposase (*Mycobacterium fortuitum)* |
| | 6 | 66 | trembl\|AACY01718843\|AACY01718843_1 | 243 KELFNTLVGAHLK 255 | Unknown |

a. Score (high scoring pair, HSP) as determined by BLAST, using sequences obtained from *de novo* sequencing. Matches higher than 64 are considered significant

b. Peptide sequence of the *de novo* data (only sequences with highest scores are given). The numbers give the residues to which the query sequenced matched to the protein sequence in the database

c. Identity of the proteins to which the *de novo* sequences were matched