Proteomic Analysis of Tardigrades: Towards a Better Understanding of Molecular Mechanisms by Anhydrobiotic Organisms

Elham Schokraie1, Agnes Hotz-Wagenblatt1, Uwe Warnken1, Brahim Mali2, Marcus Frohme2, Frank Förster3, Thomas Danekat3, Steffen Hengherr4, Ralph O. Schill4, Martina Schnölzer1*

1Functional Proteome Analysis, German Cancer Research Center, Heidelberg, Germany, 2Department of Molecular Biology and Functional Genomics, University of Applied Sciences Wildau, Wildau, Germany, 3Department of Bioinformatics, University of Würzburg, Würzburg, Germany, 4Department of Zoology, University of Stuttgart, Stuttgart, Germany

Abstract

Background: Tardigrades are small, multicellular invertebrates which are able to survive times of unfavourable environmental conditions using their well-known capability to undergo cryptobiosis at any stage of their life cycle. Milnesium tardigradum has become a powerful model system for the analysis of cryptobiosis. While some genetic information is already available for Milnesium tardigradum the proteome is still to be discovered.

Principal Findings: Here we present to the best of our knowledge the first comprehensive study of Milnesium tardigradum on the protein level. To establish a proteome reference map we developed optimized protocols for protein extraction from tardigrades in the active state and for separation of proteins by high resolution two-dimensional gel electrophoresis. Since only limited sequence information of M. tardigradum on the genome and gene expression level is available to date in public databases we initiated in parallel a tardigrade EST sequencing project to allow for protein identification by electrospray ionization tandem mass spectrometry. 271 out of 606 analyzed protein spots could be identified by searching against the publicly available NCBI nr database as well as our newly established tardigrade protein database corresponding to 144 unique proteins. Another 150 spots could be identified in the tardigrade clustered EST database corresponding to 36 unique contigs and ESTs. Proteins with annotated function were further categorized in more detail by their molecular function, biological process and cellular component. For the proteins of unknown function more information could be obtained by performing a protein domain annotation analysis. Our results include proteins like protein member of different heat shock protein families and LEA group 3, which might play important roles in surviving extreme conditions.

Conclusions: The proteome reference map of Milnesium tardigradum provides the basis for further studies in order to identify and characterize the biochemical mechanisms of tolerance to extreme desiccation. The optimized proteomics workflow will enable application of sensitive quantification techniques to detect differences in protein expression, which are characteristic of the active and anhydrobiotic states of tardigrades.

Introduction

Many organisms are exposed to unfavourable, stressful environmental conditions, either permanently or for just certain periods of their lives. To survive these extreme conditions, they possess different mechanisms. One of amazing adaptation is anhydrobiosis (from the Greek for “life without water”), which has puzzled scientists for more than 300 years. For the first time the Dutch microscopist Anton van Leeuwenhoek (1702) gave a formal description of this phenomenon. He reported the revival of “animalcules” from rehydrated moss samples. In extreme states of dehydration, anhydrobiotic invertebrates undergo a metabolic dormancy, in which metabolism decreases to a non-measurable level and life comes to a reversible standstill until activity is resumed under more favourable conditions [1]. One of the best known anhydrobiotic organisms are tardigrades. Tardigrades remain in their active form when they are surrounded by at least a film of water. By losing most of their free and bound water (>95%) anhydrobiosis occurs [2]. Tardigrades begin to contract their bodies and change their body structure into a so-called tun state (Figure 1). In the dry state these organisms are highly resistant to environmental challenge and they may remain dormant for a long period of time. Based on their amazing capability to undergo anhydrobiosis, tardigrades colonise a diversity of extreme habitats [3], and they are able to tolerate harsh environmental conditions in any developmental state [4]. Possessing the ability to enter anhydrobiosis at any stage of life cycle, tardigrades can extend their lifespan significantly [4,5]. Additionally, in the anhydrobiotic state, tardigrades are
extraordinary tolerant to physical extremes including high and subzero temperatures [6,7,8], high pressure [6,9], and extreme levels of ionizing radiation [10,11]. Interestingly, tardigrades are even able to survive space vacuum (imposing extreme desiccation) and some specimens have even recovered after combined exposure to space vacuum and solar radiation [12].

Anhydrobiosis seems to be the result of dynamic processes and appears to be mediated by protective systems that prevent lethal damage and repair systems. However, the molecular mechanisms of these processes are still poorly understood. Up to now investigations of mechanisms of desiccation tolerance have focused mainly on sugar metabolisms, stress proteins and a family of hydrophilic proteins called LEA (late embryogenesis abundant). The presence of non-reducing trehalose and its expression during anhydrobiosis has been reported for different anhydrobiotic species [13,14], which indicates the important role of trehalose in anhydrobiosis. However, the existence of anhydrobiotic animals that exhibit excellent desiccation tolerance without having disaccharides in their system [15,16] shows that sugars alone do not sufficiently explain these phenomena.

*Milnesium tardigradum* Doyère (1840) is a very well known species of carnivorous tardigrade. Different aspects of the life history of this species have been already described [17]. While some genetic studies of *M. tardigradum exist* [18] almost nothing is known about the proteome. Partial sequences of three heat shock protein (hsp70 family) genes and the housekeeping gene beta-actin have been described [18] and the relation of hsp70 expression to desiccation tolerance could be shown by real time PCR [18] and by de novo protein synthesis [6]. Since no trehalose could be detected in *M. tardigradum* [19], investigating proteins and posttranslational modifications is of particular importance to clarify surviving mechanisms during desiccation.

To gain insight into the unique adaptation capabilities of tardigrades on the protein level we aimed to establish a comprehensive proteome reference map of active *M. tardigradum* employing optimized protocols for protein extraction, generation of high-resolution 2D gels and high-throughput protein identification by electrospray ionization tandem mass spectrometry (ESI-MS/MS). The proteome reference map of *M. tardigradum* provides the basis for further studies in order to understand important physiological processes such as anhydrobiosis and stress resistance. The optimized proteomics workflow will enable application of sensitive quantification techniques to detect differences in protein expression, which are characteristic of active and anhydrobiotic states. Thus, our proteomic approach together with in-depth bioinformatic analysis will certainly provide valuable information to solve the over 300 years existing puzzle of anhydrobiosis.

**Results**

**Preparation of Protein Extracts from Active Tardigrades**

To establish and optimize a reliable and robust protocol for the extraction of proteins from tardigrades in the active state we applied different workup protocols and evaluated them by one-dimensional (1D) gel electrophoresis. Figure 2 shows the separation of protein extracts from whole tardigrades without any precipitation step (lane 2), after trichloroacetic acid/acetone precipitation (lane 3), after chloroform/methanol precipitation (lane 4) and after using a commercially available clean-up kit (lane 5). When using trichloroacetic acid/acetone precipitation we lost many proteins especially in the low molecular weight range. Chloroform/methanol precipitation and application of clean-up kit delivered satisfying results but also using the whole protein lysate directly without any further purification resulted in high yields across the entire molecular weight range. This workup protocol was therefore used throughout our proteome study. To evaluate the quality of our protocol especially with respect to proteolysis we performed Western blot analysis to detect any protein degradation. Since no proteins have been identified so far, we have chosen two polyclonal antibodies directed against the highly conserved proteins actin and alpha-tubulin. As shown in Figure 3A and 3B both proteins could be detected at their expected molecular weight at approx. 40 and 50 kDa, respectively, which is in agreement with the protein bands of the control lysate of HeLa cells. Importantly, no protein degradation could be observed during our sample preparation.

**Two Dimensional Gel Electrophoresis (2-DE)**

The establishment of an optimized workup protocol was a prerequisite for high quality 2D gels from tardigrades in the active state. The proteomics workflow is depicted in Figure 4. One important step in the workflow is the collection and preparation of the samples. To avoid contamination with food-organisms,
tardigrades were washed several times and starved over 3 days. Direct homogenization and sonication of deep-frozen tardigrades in ice cold lysis buffer without any previous precipitation step yielded protein extracts which were separated by high resolution 2D gel electrophoresis. For maximal resolution of protein spots and high loading capacity (330 μg proteins) we used pI 3–11 NL strips (24 cm) for the first dimension. Thus, high resolution separation could be achieved in the acidic as well as in the basic pH range as shown in the image of the silver stained preparative gel of whole protein extract (Figure 5).

Approximately 1000 protein spots were automatically detected on the 2D gel image using the Proteomweaver image software. A total of 606 protein spots were picked from the silver stained gel. These spots were digested with trypsin and after extraction of the tryptic peptides from the gel plugs peptide mixtures were analyzed by nanoLC-ESI-MS/MS.

Protein Identification

Identification of proteins depends on the representation of the sequence or a close homologue in the database. Since almost no genome or EST sequences of M. tardigradum are available to date in public databases we initiated the tardigrade EST sequencing project as outlined in figure 4 (Mali et al, submitted data). A cDNA library was prepared from tardigrades in different states (active, inactive, transition states). The cDNAs were sequenced as ESTs and clustered. Thereby, we obtained a nucleotide database containing 818 contigs and 2500 singlets. cDNA sequencing and generation of ESTs are still ongoing, thus the sequence coverage of M. tardigradum in the database is incomplete.

For protein identification we used the following databases: the database of M. tardigradum containing the clustered ESTs as outlined above, the tardigrade protein database, which was translated from the clustered EST database and thus represents a subdatabase containing only annotated proteins with known function and the publicly available NCBInr database. The selected 606 spots from the 2D gel correspond to some highly expressed proteins, but mostly to spots in the medium and low expression range. A total of 271 spots could be identified from the tardigrade protein and the NCBInr databases. Figure 6 demonstrates how identified proteins are distributed among these two databases. 56 unique proteins were successfully identified by searching the NCBInr database. It concerns proteins which are either highly conserved among different species e.g. actin or protein entries from M. tardigradum which are already available in the NCBInr database e.g. elongation factor 1-alpha. Further 73 unique proteins could be identified by searching the tardigrade protein database and another 15 unique proteins were present in both databases. Identical proteins that were identified from several spots were included only once in the statistics to avoid bias. Thus, the combination of the two databases was sufficient for the identification of 144 unique proteins. The corresponding protein spots are indicated by green circles in the 2D reference map shown in Figure 5. Table 1 shows an overview of identified proteins with annotation in different functional groups.
The individual ion score is included in brackets at the end of every peptide sequence. Following ion scores indicate a significant hit \( p < 0.05 \): 53 for NCBInr searches, 14 for searches in the tardigrade protein database and 27 by searching the EST clustered database. Identical proteins identified in different spots are listed only once in Table 2. In these cases the spot with the highest protein score (in bold) is ranked at the top whereas the other spots are listed below. All further information such as accession numbers, peptide sequences and sequence coverage refer to the top-ranked spot.

The 15 proteins which were identified in both databases are indicated with asterisk (e.g. spot A30*) and both accession numbers are listed. In these cases the listed peptide sequences belong to the hit with the highest score. Protein spots below the bold one are marked with \(^{\wedge}\), when only found in the NCBInr database or marked with \(^{\wedge}\), when only found in the tardigrade protein database.

Furthermore we were able to identify additional 150 protein spots by searching MS/MS data in the clustered EST database of *M. tardigradum*. These 150 proteins correspond to 36 unique contigs and ESTs. The protein information is listed in Table 3 and the protein spots are indicated by blue circles in the 2D reference map (Figure 5). Unfortunately, it was not possible to annotate them when performing a BLAST search. For these proteins of unknown function more information could be obtained by applying protein domain annotation methods. We ran all proteins through the DomainSweep pipeline which identifies the domain architecture within a protein sequence and therefore aids in finding correct functional assignments for uncharacterized protein sequences. It employs different database search methods to scan a number of
protein/domain family databases. 2 out of the 36 unique proteins gave a significant hit, whereas 28 proteins were listed as putative proteins in the respective database and the overlap between the two databases. In summary, we identified 421 (69.5%) out of 606 protein spots which were picked from the preparative 2D gel. 271 spots yielded 144 unique proteins with distinct functions whereas 150 spots were identified as proteins with yet unknown functions.

Functional Assignment of Proteins

The 144 unique proteins with annotation were further analysed using the Blast2GO program, which provides analysis of sequences and annotation of each protein with GO number to categorize the proteins in molecular function, biological process and cellular component. By analysing the proteins on the GO level 2 in the category molecular function we received a total of 9 subgroups as shown in Figure 7, upper middle chart. The majority of the identified proteins exhibit either binding (45%) or catalytic activity (35%). A more detailed analysis (GO level 3) revealed that 39% of the proteins with catalytic activity are involved in hydrolase activity (Figure 7, upper right chart) and 38% of binding proteins bind to other proteins (Figure 7, upper left chart).

Identified proteins are involved in diverse biological processes. A total of 16 subgroups of biological processes are represented (Figure 7, lower middle chart). 23% are involved in cellular processes and 18% in metabolic processes. Within the cellular processes a majority of 20% of tardigrade proteins are involved in cellular component organization and biogenesis. Within the metabolic processes 28% of proteins are involved in cellular metabolic processes, 26% in primary metabolic processes and 21% in macromolecule metabolic processes (Figure 7, lower right chart). A detailed GO description of all identified and annotated tardigrade proteins is included in Table S1.

Identified Proteins and Protein Families

In our proteomic study several heat shock proteins have been identified, namely hsp-1 (spot F27), hsp-3 (spot F2), hsp60 (spot F3), hsp70 (spot B146, B173, C131, C133), hsp82 (spot F13), hsp86 (spot F24, F25), hsp90 alpha (spot E64), hsp90 beta (spot F24) and hsp108 (spot F12). Hsp70 is already described in M. tardigradum as a molecular chaperone which could play a role in desiccation tolerance [18]. Hsp60 could be identified in spot F57 when searching the corresponding MS/MS data against the NCBI nr database. No hit was obtained in the tardigrade EST or protein database which is surprising, because hsp60 is an abundant protein.

Several protein spots have been identified as cytoskeletal proteins, including actin as most abundant protein spot (E40) on the 2D gel and tubulin. Actin and tubulin are highly conserved proteins and were used to control proteolytic degradation during our workup procedure by Western blotting. Four different actin proteins are found by MS/MS analysis, which play important roles in muscle contraction, cell motility, cytoskeletal structure and cell division. Tubulin is a key component of the cytoskeletal microtubules. Both alpha- and beta-tubulin could be identified on the 2D gel in spot D107, D110 and F6. Further proteins involved in motor activity and muscle contraction were found, namely tropomyosin (e.g. spot F35), myosin (e.g. spot F81), annexin A6 (e.g. spot D90) and myophilin (e.g. spot A128), which is a smooth-muscle protein and was described in the tapeworm Echinococcus granulosus [20].

In addition, several proteins have been identified which are known to have important roles in embryonic or larval development. Mitochondrial malate dehydrogenase precursor (e.g. spot B109), vitellogenin 1 and 2 (e.g. spot D62 and B88), GDP-mannose dehydratase (spot C87), protein disulphide isomerase 2 (e.g. spot F3), hsp-3 (spot F21), hsp-1 (spot F27), tropomyosin (spot F35) and troponin C (spot F87) belong to this group of proteins. Vitellogenin, a major lipoprotein in many oviparous animals, is known as the precursor of major yolk protein vitellin [21]. Vitellogenin is a phospholipido-glycoprotein which functions as a nutritional source for the development of embryos [22]. During developing oocytes vitellogenin and vitellin are modified through cleavage and by different posttranslational modifications (PTMs) like glycosylation, lipidation and phosphorylation. Interestingly we could identify vitellogenin in several spots on the 2D gel showing vertical [p] shifts most probably caused by PTMs.

 Peroxiredoxins identified first in yeast [23] are conserved, abundant, thioredoxin peroxidase enzymes containing one or two conserved cysteine residues that protect lipids, enzymes, and DNA against reactive oxygen species. Different isoforms of peroxiredoxins could be identified on the 2D gel: peroxiredoxin-4 (spot C132), peroxiredoxin-5 (spot B183) and peroxiredoxin-6 (spot D159). An important aspect of desiccation tolerance is protection against free radicals [24, 25]. Notably, the expression of 1-cysteine (1-Cys) peroxiredoxin family of antioxidant is reported in Arabidopsis thaliana and is shown to be related to dormancy [26]. Our results show the presence of important antioxidant systems, including superoxide dismutase (SOD) and peroxidases. Additionally different forms of glutathione S-transferases (spot A122, B153, B166, B169, D166, and D159) could be identified. Glutathione transferases (GSTs) constitute a superfamily of detoxifying enzymes involved in phase II metabolism. Detoxification occurs by either glutathione conjugation, peroxidase activity or passive binding [27]. Furthermore GSTs have cellular physiology roles such as regulators of cellular pathways of stress response and housekeeping roles in the binding and transport of specific ligands [28]. The consequence of this diversity in role is the expression of multiple forms of GST in an organism. It has been shown that the expression of the different isoenzymes is highly tissue-specific [29], and this heterogeneity of GSTs may be further complicated by posttranslational modifications such as glycosylation [30].

Some protein spots were identified as calreticulin (e.g. spot F14) which is a Ca++-binding protein and molecular chaperone. Calreticulin is also involved in the folding of synthesized proteins and glycoproteins [31].
| Cytoskeleton elements and modulators | Enzymes | Proteases and protease inhibitors |
|-----------------------------------|---------|----------------------------------|
| Alpha-III tubulin | Glucan endo-1,3-beta-glucosidase | Cathepsin K |
| Beta-tubulin class-V | Prostatic acid phosphatase | Cathepsin Z |
| Beta-tubulin class-I | Adenylate kinase isoenzyme 1 | Cathepsin L1 |
| Actin | Peptidyl-prolyl cis-trans isomerase | Neprilysin-2 |
| Actin-5C | Glutamate dehydrogenase | Peptidase M17 precursor |
| Actin | Lysozyme | Actinidain |
| Alpha actin | Mitochondrial malate dehydrogenase | Plasminogen |
| Actin, muscle-type (A2) | Arginine kinase | Aspartic protease inhibitor 8 |
| muscle actin | Aconitase, mitochondrial | AFG3-like protein 2 |
| Similar to alpha actinin CG4376-PB | Transaldolase | 26S proteasome non-ATPase regulatory subunit 8 |
| Myophilin | Aldolase A protein | Rab GDP dissociation inhibitor beta |
| Tropomyosin-1, isoforms 9A/A/B | Protein disulfide isomerase-3 | Gamma-glutamyltranspeptidase |
| Tropomyosin | Matrix metalloproteinase-17 | Response to stress or heat |
| Myosin regulatory light polypeptide 9 | Mitochondrial long-chain enoyl-CoA hydratase/3-hydroxysteroyl-CoA | NADP-dependent isocitrate dehydrogenase |
| Myosin, essential light chain | Dehydrogenase alpha-subunit | Heat shock 70 kDa protein II |
| **Heat shock proteins** | Peroxidase | similar to heat shock cognate 70 protein isoform 2 |
| Heat Shock Protein family member (hsp-3) | Methylmalonate-semialdehyde dehydrogenase | Short-chain dehydrogenase/reductase SDR YhdF |
| Heat Shock Protein family member (hsp-1) | Thioredoxin reductase 1 | Aspartic protease inhibitor 8 |
| Hsp 60 | Succinyl-CoA ligase [GDP-forming] subunit beta, Mitochondrial E | UspA |
| Hsp 70 | GTP-specific succinyl-CoA synthetase beta subunit | Rubber elongation factor protein (REF) (Allergen Hev b 1) |
| **Heat shock cognate 70** | Glyceraldehyde-3-phosphate dehydrogenase | Other Proteins |
| Heat shock cognate 70 protein isoform 2 | DEAD-box family (SNF2-like) helicase | Heat shock protein 90-beta |
| Heat shock 70 kDa protein II (HSP70 II) | Cysteine conjugate beta-lyase | Heat shock protein 83 |
| Hsp 90-beta | 26S proteasome non-ATPase regulatory subunit 13 | Heat shock protein 60 |
| Hsp90-alpha | Heat shock cognate 70 | Protein lethal(2)essential for life (member of Hsp20 family) |
| Hsp90 | Glycolysis | Phosphoglycerate kinase |
| Hsp 82 | Glyceraldehyde-3-phosphate dehydrogenase | Angiopoietin-related protein 1 |
| Hsp 83 | Triosephosphate isomerase | |
| Hsp108 | Enolase | |
| **Embryonic/larval development** | Spaghetti CG13570-PA | |
| Vitellogenin-1 | H(+)-transporting ATP synthase | Prohibitin |
| Vitellogenin-2 | ATP synthase subunit d, mitochondrial | Proteasome subunit alpha type-4 |
| Protein disulfide-isomerase 2 | ATP synthase beta subunit | 40S ribosomal protein S12 |
| Heat Shock Protein family member (hsp-3) | Mitochondrial ATP synthase alpha subunit precursor | Periostin |
| Heat Shock Protein family member (hsp-1) | Annexin A6 | Acetylcholine receptor subunit alpha-L1 |
| Troponin C | Antioxidant proteins | Nucleosome remodelling factor – 38kD CG4634-PA |
| Putative LEA III protein isoform 2 | Thiol-specific antioxidant protein | Coiled-coil domain-containing protein 25 |
| GDP-Mannose Dehydratase | Superoxide dismutase [Cu-Zn] | Calreticulin |
| Tropomyosin | Peroxiredoxin-5, mitochondria | Lipoprotein-related protein |
| **Dormancy related protein** | Peroxiredoxin-4 | 14-3-3 protein beta/alpha-2 (Protein 14-3-3B2) |
| Putative LEA III protein isoform 2 | Glutathione S-transferase | 60S ribosomal protein L26-1 |
| Peroxiredoxin-6 | Histone H4 | Histone H2B.2 |

Identified proteins with annotation are listed in 8 different groups with majority in protein enzymes. We also identified many heat shock proteins and proteins, which are involved in embryonic development, response to stress/heat and dormancy.

doi:10.1371/journal.pone.0009502.t001
| Spot no. | Protein name | NCBI nr Accession no. (|) | Tardigrade specific Accession no. (contig/EST) (|) | Total protein score | No. of unique/ significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|---------|--------------|---------------------------|-----------------------------------------------|-------------------|-------------------------------|---------------------------------------|------------------|
| A30*    | elongation factor 1-alpha [Milnesium tardigradum] | gi|4530101 | (EZ048811) | 544 | 5 | K. YAWVLDK. L (23) | 51% |
| A27*, A28*, A29*, A32*, A3*, A41* | | | | | | R. LIQDQVYK. I (52) | |
| A33*, A35*, A40*, A41* | | | | | | K. IGGGTVPGVQR. V (56) | |
| A42*, A49*, A50*, A59*, A81*, A84^, A85^, A88* | | | | | | R. EHALAYTLGVYK. Q (65) | |
| A107^, B78^, B80^, B81* | | | | | | K. YYVTIDAPGHKR. D (67) | |
| C28*, C36*, D120^ | | | | | | R. NGYPVLDCHTAVIACK. F (18) | |
| A79 | DB:Swissprot Frame:3 orf3 Homolog:Angiopoietin-related protein 1 | EZ048825 | 35 | 1 | K. VYQDSLVPDPRN. C (35) | 5% |
| A84 | DB:Swissprot Frame:1 orf8 Homolog:Cathepsin K | GH986829 | 58 | 1 | K. LSEEFPVR. D (13) | 16% |
| A85 | DB:Swissprot Frame:1 orf7 Homolog:Actinidain | EZ048769 | 33 | 1 | K. NSGPNMANK. G (33) | 18% |
| A90 | DB:Trembl Frame:3 orf1 Homolog:GF11309 | EZ048774 | 192 | 3 | K. GAVSCIDSFVNR. C (68) | 20% |
| spaghetti CG13570-PA [Drosophila melanogaster] | gi|17864228 | 55 | 1 | K. ILMGNFDSTFAFDLLR. T (55) | 2% |
| Spot no. | Protein name | NCBInr Accession no. | Tardigrade specific Accession no. (contig/EST) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|----------------------|-----------------------------------------------|---------------------|-----------------------------------|------------------------------------------|-----------------|
| A104     | DB:Swissprot Frame2orf3 Homolog:Prostatic acid phosphatase | GH986832              | 162                                           | 4                   | R.YSSYLGPK.F(53)                  | 48%                                      |
| A89      |              |                      |                                               |                     | R.YSSYLGPK.F(44)                  | 37%                                      |
| A111*    | prohibit [Aedes aegypti] | gi|157131967 | (E)Z048795                                   | 121                 | 2                                 | R.FNASQLITQR.Q(54)                        | 7%              |
| A121     | mitochondrial ATP synthase alpha subunit precursor [Strongylocentrotus purpuratus] | gi|47551121 |                                                | 97                  | 1                                 | R.VLSIGDIAR.Y(45)                         | 3%              |
| A122.    | DB:Swissprot Frame1orf3 Homolog:Glutathione S-transferase 1 | EZ048812              | 439                                           | 8                   | K.LSQVIER.I(38)                   | 45%                                      |
| B170, B175 | transferase 1 | Evalue:e-39 | Bitscore:164                                 |                     |                                   |                                          |                 |
| B153     | DB:Swissprot Frame3orf2 Homolog: Glutathione S- | EZ048805              | 260                                           | 7                   | K.YLIGNDVY.Y(19)                  | 55%                                      |
| B154     | transferase | Evalue:e-44 | Bitscore:177                                 |                     |                                   |                                          |                 |
| Spot no. | Protein name | NCBI nr Accession no. (contig/EST) (v) | Tardigrade specific Accession no. (contig/EST) (v) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|---------|--------------|----------------------------------------|---------------------------------|--------------------|-----------------|----------------------------------|------------------|
| B166    | DB:Swissprot Frame2 orf1 Homolog: Glutathione S- transferase | E048770 | 213 | 6 | K.TEEEQQQCDMVEALSDFR.Q + Oxidation (M) (25) |
| B158    | transferase | evaluated: 3e-4 Bitscore: 176 | EZ048770 | 213 | 6 | K.QILGSQD.RK.Y (25) |
| B169    | DB:Swissprot Frame1 orf3 Homolog: Probable glutathione S-transferase 9 | GH986911 | 96 | 2 | R.LFAAAQQR.K (48) |
| D166    | DB:Swissprot Frame3 orf3 Homolog: Glutathione S-transferase | GH986673 | 48 | 2 | R.DMLVAMQ.R.W (14) |
| D159    | DB:Swissprot Frame1 orf3 Homolog: Probable glutathione S-transferase 6 | GH986870 | 38 | 2 | R.MFEDQK.S (56) |
| A128    | DB:Swissprot Frame3 orf2 Homolog: Myophilin | E048783 | 273 | 6 | R.NFSDEQR.K (35) |
| A147    | DB:Swissprot Frame3 orf2 Homolog: Adenylate kinase isoenzyme 1 | E048787 | 19 | 1 | R.KFLDGFPP.R.E (19) |
| A148    | DB:Swissprot Frame1 orf1 Homolog: Peptidyl-proly1 | E048822 | 140 | 3 | K.TSKPVVDQGQ.L-(34) | 25% |
Table 2. Cont.

| Spot no. | Protein name | NCBInr Accession no. (1) | Tardigrade specific Accession no. (contig/EST) (2) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|--------------------------|-------------------------------------------------|-------------------|----------------------------------|----------------------------------------|------------------|
| B19      | DB:Swissprot Frame:2 orf:1 Homolog: Elongation factor 2 | GH986944 22 | K.TSKPVMVIAOCQQL. - (59) | K.HVVYQQQYTGELDVK. K (49) |
| B41      | PREDICTED: similar to CG8036-PA, isoform B | gi|66503776 125 | K.LDSDLRGHTPR. L (48) | 3% |
| B82      | glutamate dehydrogenase, short peptide [Drosophila melanogaster] | gi|458803 116 | K.IAEAAANGHTPAADK. I (50) | 9% |
| B88      | DB:Swissprot Frame:2 orf:1 Homolog: Vitellogenin-2 | EZ048823 200 | K.VSMINL.R. L (41) | 14% |
| B9, B6, B95, B96, C117, C122, C124 | Low coverage | | | |
| B92      | NADP-dependent isocitrate dehydrogenase [Homo sapiens] | gi|3641398 69 0 | K.DIFQEIYD.K.Q (42) | 4% |
| B95      | DB:Swissprot Frame:1 orf:1-3 | GH986689 17 1 | K.KDPEAHNETAK.W + Oxidation (M) (17) | 6% |
| C3orf33 homolog | | | | |
| B102     | DB:Swissprot Frame:3 orf:1 Homolog:Lysosomal | EZ048780 90 2 | K.FLEPLTVPR. A (52) | 46% |
| B101     | acid phosphatase | Ez049752 68 1 | K.FILYMTISALLAFAK. A (28) | | |
| B146     | hypothetical protein TRIADDRAFT | gi|198909922 342 4 | R.VEIJNTRG. N. (38) | 8% |
| B141, B144 | [Trichoplax adhaerens] | | | |

Table 2. Proteome Map of Tardigrades
| Spot no. | Protein name | NCBI/Accession no. (* | Tardigrade specific Accession no. (contig/EST) (†) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|-------|-------------|----------------------|------------------------------------------|------------------|-----------------------------------|-----------------------------------------|------------------|
| B173  | heat shock cognate 70 [Aedes aegypti] | gi|9446896 | 235 | 2 | K. IQVEYKGETK. N(38) | 8% |
| C131  | Heat shock 70 kDa protein II (HSP70 II) | gi|123622 | 154 | 2 | K. ITEAYLGGK. E (34) | 8% |
| C133  | PREDICTED: similar to heat shock cognate 70 protein isoform 2 [Acyrthosiphon pisum] | gi|193603576 | 153 | 2 | K. VEIQDNQGKR. T (60) | 8% |
| B148  | DB:Swissprot Frame:3 orf:2 Homolog:Malate dehydrogenase, cytoplasmic | GH986821 | 179 | 3 | R. AGMMAIQLK. N(52) | 24% |
| B152  | H(τ)-transporting ATP synthase (Rattus norvegicus) | gi|57029 | 92 | 1 | K. LELAQYR. E (31) | 8% |
| B164* | ATPase subunit (Beta vulgaris subsp. Vulgaris) | gi|11263 | (EZ048779) | 64 | 0 | K. LELAQYR. E (34) | 4% |
| B167* | DB:Swissprot Frame:1 orf:1 Homolog:Arginine kinase | (gi|124264768) | EZ048827 | 254 | 6 | R. FLQQAVVR. F (41) | 33% |
| B133\, B157\, B139* | kinase | (gi|124264768) | EZ048827 | 254 | 6 | R. LIDDHFLPK. E (39) | |
| B162\, C19\, C98\, C104\ | | | | | | K. LNFPNDPEGK. Y (60) | |
| C107\, C137\, C132* | | | | | | R. KMTPEIIOQK. L + Oxidation (M) (29) | |
| D98\, D159 | | | | | | R. SLQGFPNNPLNEQQYK. E (30) | |
| Spot no. | Protein name | NCBInr Accession no. (contig/EST) | Tardigrade specific Accession no. (contig/EST) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|-----------------------------------|-----------------------------------------------|--------------------|-------------------------------------|-----------------------------------------|------------------|
| B183     | D8:Swissprot Frame:2 of:5 Homolog: Peroxiredoxin | EZ048816 92 1 | R.HLSYVK.K(10) | 24% | K.DLFPIINDYHVGDIEK.G(55) |
|          | S, mitochondrial | | Evalue:3e-40 Bitscore:150 | | K.VHILLADPR.G(9) |
| C28      | pre-mRNA binding K protein, hnRNP K [Xenopus] | g|1299029 58 1 | R.ILSISADITIGEILK.K(58) | 4% | K.LNIEPDGOWECIAD.1(73) |
| C36      | laevi, Peptide, 396 aa | | | | |
| C42      | Heterogeneous nuclear ribonucleoprotein K [Mus musculus] | g|13384620 67 0 | R.ITAVALSP.I(43) | 7% | |
| C47      | PREDICTED: similar to acanthase, mitochondrial | g|156537745 58 0 | K.NTIVTSYNR.N(25) | 2% | |
|          | [Xenonia vitripennis] | | | | K.ILYSHLDPQK.Q(33) |
| C52      | peptidase M17 precursor [Onchorchis sinensis] | g|218429525 55 1 | K.GITYDGADVK.K(55) | 2% | |
| C60      | D8:Swissprot Frame:2 of:1 Homolog: Gamma-glutamyltranspeptidase 1 | GH986789 53 2 | K.DMSSPDGYHR.P + Oxidation (M) (31) | 11% | K.LKEFLTSQPVAQRST.R(22) |
|          | glutamyltranspeptidase 1 | Evalue:6e-49 Bitscore:194 | | | |
| C87      | GDP-Mannosate Dehydratase family member (gmd-2) | g|17507723 61 1 | K.FYQASTELGYK.V(61) | 3% | |
|          | [Caenorhabditis elegans] | | | | |
| C95      | Short-chain dehydrogenase /reductase SDR thdF | g|52079424 110 1 | K.GAIYAPF.T.S(51) | 7% | |
|          | [Bacillus licheniformis ATCC 14580] | | | | K.TAITTGDGSI.GR.A(59) |
| C98      | D8:Swissprot Frame:3 of:5 Homolog: Uncharacterized oxidoreductase yhdF | EZ048820 59 2 | R.GYRPEEVTLK.T(15) | 30% | |
|          | Evalve:3e-28 Bitscore:125 | | | | K.DGVLSVECPGLQSNR.L(44) |
|          | Homolog:Protein lethal?essential for life Evalue:2e-11 Bitscore:70.9 | | | | |
|          | D8:Swissprot Frame:1 of:1 | GH986571 35 1 | K.LLEELASTAK.V(24) | 18% | |
|          | Homolog:Probable transaldolase Evalue:6e-34 Bitscore:144 | | | | |
| Spot no. | Protein name | NCBInr Accession no. (’l) | Tardigrae specific Accession no. (contig/EST) (’l) | Total protein score | No. of unique/ significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|---------------------------|---------------------------------|-------------------|----------------------------------|---------------------------------|------------------|
| C110     | aldolase A protein [Homo sapiens] | gi|28595 | | 71 | 1 | K. GILAADESTGSIAK.R (71) | 12% |
| C111     | DB:Swissprot Frame:2 orf:1 | GH986712 | | 344 | 7 | K. ADVKEQDGQGLGINS.R.L(63) | 51% |
| C115     | DB:Swissprot Frame:1 orf:6 Homolog: Plasminogen | EZ048798 | | 63 | 2 | K. GODDEFR.I(34) | 19% |
| D159     |  | | | 121 | 2 | R. AYSGISADMLCGAAPGK.D(29) | 21% |
| C126     | DB:Swissprot Frame:2 orf:4 Homolog: Proteasome subunit alpha type-4 Evalue:4e-81 | GH986859 | | 47 | 1 | R. TITIFSPERK.L(47) | 4% |
| C128     | expressed hypothetical protein [Trichoplax adhaerens] | gi|196010133 | | 105 | 1 | K. VGASEATLINMLK.V(105) | 4% |
| F25H2.10 | [Caenorhabditis elegans] | gi|17506815 | | 97 | 1 | K. TSFPQAQIPTK.I(97) | 3% |
| DEAD-box family (SNF2-like helicase, putative [Theileria annulata] | gi|84996109 | | 54 | 1 | K. MELISNI.K(54) | 0% |
| ResB family protein [Hydrogenobaculum sp. Y04AAS1] | gi|195953863 | | 54 | 1 | K. MELISNI.K(54) | 1% |
| C132     | DB:Swissprot Frame:1 orf:1 Homolog: Peroxiredoxin-4 Evalue:4e-86 | EZ048818 | | 393 | 8 | R. GLFIDK.K(28) | 31% |
Table 2. Cont.

| Spot no. | Protein name | Accession no. | NCBInr Accession no. (contig/EST) | Tardigrade specific Accession no. (contig/EST) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|---------------|-----------------------------------|-----------------------------------------------|--------------------|-----------------------------------|------------------------------------------|------------------|
| C139     | DB:Swissprot:1 orf:1 Homolog:ATP synthase | EZ048797 | 81 | 2 | K. DYGVYLEDAGHTLR.G (81) |
| D166     | subunit d, mitochondrial Evaluation:27 Bitscore:121 | EZ048797 | 81 | 2 | K. VIAPPESPAK.I (33) |
| C143     | Glyceraldehyde-3-phosphate dehydrogenase | gi|7274154 | 107 | 1 | R. VPPVQGDVQFVR.L (107) |
| C145^    | [Achlya bisexualis] | EZ048781 | 113 | 3 | K. DVGAEWVILGHSER.K (37) |
| B163^    | Homolog:Triosephosphate isomerase B Evaluation:4e-65 Bitscore:247 | EZ048797 | 81 | 2 | K. AVLAVIISDWK.V (67) |
| C145     | DB:Swissprot:2 orf:3 Homolog:Peroxiredoxin | EZ048781 | 113 | 3 | K. AIAAVSDWKR.I (47) |
| C155     | DB:Swissprot:1 orf:1 Homolog:Superoxide dismutase [Cu-Zn] Evaluation:2e-48 Bitscore:192 | EZ048797 | 81 | 2 | K. LACGIVGVGTG.K (69) |
| D13      | DB:Swissprot:1 orf:1 Homolog:40S ribosomal protein S12 Evaluation:2e-34 Bitscore:144 | EZ048797 | 81 | 2 | K. LDAWLPQLR.K (22) |
| D53      | PREDICTED: similar to alpha actinin [Tribolium castaneum] | gi|91080533 | 120 | 1 | R. VGHEQLTSES.R (47) |

Notes:
- ^ indicates a predicted protein.
- Protein names followed by a homolog's species name or genus.
- Accession numbers for homologs are provided in parentheses.
| Spot no. | Protein name | NCBInr Accession no. (|) | Tardigrade specific Accession no. (contig/EST) (^) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|---------|--------------|----------------------|---------------------------------------------|---------------------|-----------------------------------|----------------------------------------|-----------------|
| D56     | DB:Swissprot Frame:1 orf3 Homolog:Periostin | EZ048782 | K.QTEGETVFIPOAAGF.J(22) | 6% |
| D57     | protein disulfide isomerase-3 [Haemaphysalis longicornis] | gi|148717319 | K.HVGGYPTLI.I(48) | 3% |
| D61     | Enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) | gi|1169533 | R.QNPYEVETTOK.G(74) | 6% |
| D66     | metalloproteinase-17 | EZ048784 | K.FGNNIIGQIEK.Y(46) | 34% |
| B19, C13, C15, C16, C28 | | | | |
| C33, D29, D30, (D32-D39) | | | | |
| D42, D50, D58, E11, E12 | | | | |
| E15, E45, E54, E65, E66 | | | | |
| E73, F26 | | | | |
| D81     | DB:Swissprot Frame:2 orf1 Homolog:Actin-1 | GH986913 | K.EISALAPNTIK.-(33) | 5% |
| D84     | UspA [Bacillus coagulans 36D1] | gi|124521548 | R.IVADGSK.M(56) | 6% |
| E85     | | | | |
| D90     | DB:Swissprot Frame:2 orf1 Homolog:Annexin A6 | EZ048803 | K.DLFDDLKK.E(24) | 48% |
| D89, D96 | | | | |
| E85     | | | | |
| Spot no. | Protein name | Accession no. (contig/EST) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|---------|--------------|---------------------------|---------------------|-----------------------------------|----------------------------------------|------------------|
| R.HLLFAIITTR.R(47) | | | | | | |
| R.AFOPPFONDAK.A(38) | | | | | | |
| R.EIDDIVTSGYFR.N(43) | | | | | | |
| K.AIAGAGTSEEDLIEIMLTR.N + Oxidation (M) (6) | | | | | | |
| D91* | elongation factor 1 gamma [Bombyx mori] | gi|112983898 (E2048793) | 60 | 1 | K.VPAFESADGK.V(58) | 2% |
| D104*, D106* | | | | | | |
| D92 | mitochondrial long-chain enoyl-CoA hydratase/3- hydroxycyl-CoA dehydrogenase alpha-subunit [Rattus n] | gi|510108 | 57 | 0 | K. ALTSEFER.D(7) | 3% |
| DB:Swissprot Frame:3 orf:1 Homolog: Peroxidase | EZ048773 | 30 | 1 | R. TGFYTDQMAILK.K + Oxidation (M) (30) | 7% |
| 7e-19 Bit | Evalue:7e-19 Bitscore:94.0 | | | | | |
| D96 | cysteine conjugate beta-lyase [Takifugu rubripes] | gi|5002565 | 84 | 1 | K. ALVINTPLNLGK.V(84) | 3% |
| D96 | DB:Swissprot Frame:3 orf:5 Homolog:proteasome non-ATPase regulatory subunit 13 | GH986860 | 100 | 3 | K. LLEEV.EK.K(12) | 15% |
| DB:Swissprot Frame:3 orf:5 Homolog:neprilysin-2 | | | | | | |
| 1e-29 Bit | Evalue:1e-29 Bitscore:133 | | | | | |
| D103 | DB:Swissprot Frame:1 orf:1 Homolog:Neprilysin-2 | EZ048722 | 32 | 1 | K. IIAQYSSHFR.Y(32) | 6% |
| D107 | Tubulin alpha-3 chain (Alpha-III tubulin) | gi|3915094 | 252 | 2 | R. LSVYDGK.K(27) | 18% |
| | | | | | | |
| R.QLFHPEQLITGR.K(31) | | | | | | |
| R.LIGQIVITTSLR.F(29) | | | | | | |
| R.AVFDVEQTEQIER.T(64) | | | | | | |
| R.NLDIERPYTNLN.R.L(81) | | | | | | |
| R.FDGALVPDLTEQTLNPYPR.I(21) | | | | | | |
| D110 | Tubulin beta-3 chain (Beta-tubulin class-IV) | gi|135464 | 152 | 1 | R. FPQGLNADL.R.K(70) | 7% |
| | | | | | | |
| K.LAVNMVPPFR.L(36) | | | | | | |
| Spot no. | Protein name | NCBInr Accession no. (1) | Tardigrade specific Accession no. (contig/EST) (2) | Total protein score | No. of unique/ significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|---------|--------------|--------------------------|--------------------------------------------------|-------------------|-----------------------------------|------------------------------------------|------------------|
| D111    | GH19645 [Drosophila grimshawi] | gi|195053606 | 120 | 2 | R. AVLVDLEPGMTDSVR. S(46) |                                                                 |
| D118    | D8:Swissprot Frame:1 orf:4 Homolog Histone H4 | GH986770 | 18 | 1 | R. ISGLIYETR. G(18) | 12% |
| D140    | D8:Swissprot Frame:2 orf:3 Homolog Acetylcholine | EZ048771 | 381 | 7 | K. LGWTFK.D(51) | 23% |
| D128    | receptor subunit alpha-L1 | EZ048771 | 381 | 7 | K. LGWTFK.D(51) | 23% |
| D159    | D8:Swissprot Frame:2 orf:1 Homolog Peroxiredoxin- | GH986904 | 403 | 8 | K. LAPEPEK. R(38) | 53% |
| E4      | phosphoglycerate kinase[Verrucomicrobia bacterium V4] | gi|161075769 | 54 | 1 | K. AIGFMKEKEL. Y + Oxidation (M) (54) | 2% |
| E5      | Rubber elongation factor protein (REF) (Allergen) | gi|132270 | 104 | 1 | R. SLASSLPQGTK. I(33) | 18% |
| D99     | Hev b 1) | gi|14423933 | 87 | 0 | K. AEQYAVITWR. A(43) | 14% |

Table 2. Cont.
| Spot no. | Protein name | NCBI no. Accession no. (contig/EST) (\^) | Tardigrade specific Accession no. (contig/EST) (\^) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | DB:Swissprot Frame:2 orf:1 Homolog:Liver | Sequence coverage |
|---------|--------------|----------------------------------------|----------------------------------------|-------------------|---------------------------------|------------------------------------------|------------------------------------------|------------------|
| E43*   | actin [Heliothis virescens] | gi|14010639 | (EZ048826) | 667 | 7 | K. EITALAPSTMK.I (41) | R. AVFPSIVGRPR.H (73) K. INHHTFNYELR.V (73) K. QYEDEGSPIVHR.K (94) K. SYEIPQVITIGNER.F (79) R. VAPEDHVILTEAPLNPK.A (90) K. YPIEGIIITMDDEK.I (56) | 41% |
| E47*   | Actin, muscle-type (A2) | gi|3121741 | (EZ048826) | 519 | 6 | R. DLTLYLMK.I (25) | R. GYSFPVTAER.E (38) K. EITALAPSTMK.I (41) K. SYEIPQVITIGNER.F (79) K. DLYANTVLSSTTMYGIAD.R.M (45) K. DLYANTVLSSTTMYGIAD.R.M Oxidation (M) (36) K. LCYVALDFQMAATAAASSLEK.S (39) R. TGTVLDSGDGVHTVPIEGYALPHAILR.L (55) | 37% |
| D99, D106, D108, E71, E72, E84, E92, E94, E99, E100, F44, F58, F61, F95 | | | | | | | | |
| F64*, F95* | | | | | | | | |

Table 2. Cont.
| Spot no. | Protein name | NCBInr Accession no. (|) | Tardigrade specific Accession no. (contig/EST) (|) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|----------------|----------------|----------------|------------------|-----------------------------------|----------------------------------|------------------|
| E48*     | actin 5C [Lycosa singoriensis] | gi|161661023 (E2048826) | 644 | 6 | K.AGFAGDAPR.A(80) | 35% |
| E44*     |            |              |               |                |                  | R.DLTDYLMK.I + Oxidation (M) (15) |
|          |            |              |               |                |                  | R.GSYFVTAE.R.E(44) |
|          |            |              |               |                |                  | R.AVFSIVGRPR.H(98) |
|          |            |              |               |                |                  | K.INHTFYNELR.V(66) |
|          |            |              |               |                |                  | K.QEYDEGSPSVHR.K(100) |
|          |            |              |               |                |                  | K.SYELPDQVITIGNER.F(70) |
|          |            |              |               |                |                  | R.VAPEEHPVILTEAFLNK.R.A(83) |
|          |            |              |               |                |                  | K.YPIENGITNWDMEK.I + Oxidation (M) (25) |
|          |            |              |               |                |                  | K.DLYANTVLSGTTMYPGIADR.M(47) |
|          |            |              |               |                |                  | R.KDLVANTVLSGTTMYPGIADR.M.Oxidation (M) (18) |
| E50*     | beta-actin [Rachycentron canadum] | gi|161376754 (E2048826) | 501 | 4 | K.AGFAGDAPR.A(59) | 33% |
| E6*, E42*, F27*, F28* |           |              |               |                |                  | R.DLTDYLMK.I (24) |
| E52*     | alpha-actin (aa 40-375) [Mus musculus] | gi|49364 (E2048826) | 106 | 0 | K.EITALAPSTMK.I(49) | |
| E85*, D103* |            |              |               |                |                  | R.GSYFVTAE.R.E(34) |
| E57      | muscle actin | gi|797290 | 290 | 1 | R.DLTDYLMK.I (24) | 25% |
| D104*, D97*, E69*, E60*, E57*, E56*, E55*, E53*, E51* |          |              |               |                |                  | R.GSYFVTAE.R.E(43) |
|          |            |              |               |                |                  | R.EITALAPSTMK.I (49) |
|          |            |              |               |                |                  | R.KELPDQVITIGNER.F(50) |
|          |            |              |               |                |                  | R.KELPDQVITIGNER.F(50) |
| E63      | DB:Swissprot Frame:1 orf:1 Homolog:AFG3-like protein 2 Evalue:9e-58 Bitscore:221 | gh986706 | 35 | 1 | K.CPELLEK.K(9) | 12% |
|          |            |              |               |                |                  | K.GLGYAQYLPR.R(27) |
| Spot no. | Protein name | NCBInr Accession no. (’?) | Tardigrade specific Accession no. (contig/EST) (’?) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. ion score) | Sequence coverage |
|----------|--------------|----------------------------|--------------------------------------------------|---------------------|-----------------------------------|------------------------------------------|-----------------|
| E64      | heat shock protein 90 alpha [Fundulus heteroclitus macrolepidotus] | gi|77999578 | | 62 | 1 | R. FTTSASGDENVSLK.D oxida (62) | 6% |
| E70*     | actin [Paraphidippus aurantius] | gi|167683068 (EZ048826) | | 376 | 2 | R. DLTDYLMK.I(15) | 42% |
| D102*, E51*, E56* | | | | | | | |
| E73      | predicted: similar to Nucleosome remodelling factor – 38kD CG4634-PA [Apis mellifera] | gi|66507623 | | 85 | 1 | K. NDNPDIVLEIGYK.V(85) | 1% |
| E74      | DB:Swissprot Frame:1 orf:4 Homolog: Protein | GH986548 | | 118 | 3 | K. SLAPEYAK.A(16) | 17% |
| E76, E88, E62 | dissulfide-isomerase Evaluate: e-44 Bitscore: 178 | | | | | | |
| E80      | DB:Swissprot Frame:1 orf:3-5 Homolog: Cathepsin Z | GH986945 | | 102 | 2 | K. VGFPGPISGR.E(47) | 11% |
| E83      | DB:Swissprot Frame:3 orf:1 Homolog:26S proteasome non-ATPase regulatory subunit 8 Evaluate:e-48 Bits | EZ048799 | | 154 | 4 | K. DLIPDGSLR.T(40) | 23% |
| E86      | DB:Swissprot Frame:3 orf:4 Homolog: domain-containing protein 25 Evaluate:e-46 Bitscore: 1 | EZ048088 | | 21 | 1 | K. LSSARNYLR.L(21) | 0% |
| E89      | ATP synthase beta subunit (Asteria miniata) | gi|46909233 | | 346 | 3 | K. AHGGSVFGVGER.T(33) | 24% |
| D140     | | | | | | | |
| Spot no. | Protein name | NCBInr Accession no. ('†') | Tardigrade specific Accession no. (contig/EST) ('‡') | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|---------|--------------|----------------------------|-----------------------------------------------|--------------------|---------------------------------|--------------------------------------|------------------|
| E46     |              |                            | 191 0                                          |                    |                                 | R. GIANLGYQVNPVLDSTSR.T (26)          |                  |
|         |              |                            |                                                |                    |                                 | R. EGNDDLYHEMLGGVLFKR.T (7)           |                  |
| E46     |              |                            |                                                |                    |                                 | R. IGPGGANVGVK.T (41)                | 13%              |
|         |              |                            |                                                |                    |                                 | R. IPQGPRELGR.T (34)                 |                  |
| E46     |              |                            |                                                |                    |                                 | R. VDGLAPYAK.T (40)                  |                  |
| E46     |              |                            |                                                |                    |                                 | R. TIAMDETEGLR.G + Oxidation (M) (44) |                  |
| E46     |              |                            |                                                |                    |                                 | R. FTQAGPNVSLGR.T (33)               |                  |
| E99     | DB:Swissprot Frame:2 orf:1 Homolog Rab GDP dissociation inhibitor beta | GH986887 | 152 3                                          |                    |                                 | R. GTPQVSEPVTK.V (55)                |                  |
|         |              |                            |                                                |                    |                                 | R. CICILDPHPMTK.D (6)                |                  |
|         |              |                            |                                                |                    |                                 | R. DALSTQILIPQVQNR.T (33)            |                  |
|         |              |                            |                                                |                    |                                 | R. MDYVLSVSYTHQVAK.G (4)             |                  |
| E92     | DB:Swissprot Frame:1 orf:1 | GH986892 | 75 2                                           |                    |                                 | R. TVTSILVR.E (22)                  | 11%              |
|         | Homolog: Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial |            |                                                |                    |                                 |                                      |                  |
| E93     | DB:Swissprot Frame:1 orf:1 | GH986518 | 50 1                                           |                    |                                 | R. TACTAEIGLDR.V (50)                | 5%               |
|         | Homolog: Thioredoxin reductase 1, cytoplasmic |            |                                                |                    |                                 |                                      |                  |
| E98     | DB:Swissprot Frame:3 orf:1 | EZ048794 | 527 9                                          |                    |                                 | R. IDSPPTLT.K (43)                   | 56%              |
| F3      | DB:Swissprot Frame:2 orf:1 Homolog:Protein |            |                                                |                    |                                 | R. ITPTEFLRGK.D (57)                |                  |
| F3      |              |                            |                                                |                    |                                 | K. NFDEVVMDK.S (56)                  |                  |
| F3      |              |                            |                                                |                    |                                 | R. LSLADELTVK.T (46)                |                  |
| F3      |              |                            |                                                |                    |                                 | K. GDNTPVEYEGGER.T (38)              |                  |
| F3      |              |                            |                                                |                    |                                 | K. MDATALEKEMK.T + Oxidation (M) (72)|                  |
| F3      |              |                            |                                                |                    |                                 | K. KGDNTPVEYEGGER.T (62)             |                  |
| F3      |              |                            |                                                |                    |                                 | K. LSFGYDEGLDFFK.D (79)              |                  |
| F3      |              |                            |                                                |                    |                                 | K. YKRGKLEDLPMTLTK.F (64)            |                  |
| F3      |              |                            |                                                |                    |                                 | K. LKPHIISQVPHEDMWAK.S (10)          |                  |
| F3      |              |                            |                                                |                    |                                 |                                      |                  |
| F6      | Tubulin beta-1 chain { Beta-tubulin class-I} | gi|57429 | 54 0                                           |                    |                                 | R. YLTVAAIFR.G (15)                  | 4%               |
| Spot no. | Protein name | NCBI ref Accession no. (\*) | Tardigrade specific Accession no. (contig/EST) (\*) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|---------|--------------|----------------------------|--------------------------------------------------|--------------------|-----------------------------------|------------------------------------------|------------------|
| F12     | hsp 108 [Gallus gallus] | gi|63509 | 84 | 0 | R. FPGQGSLADKLKR(39) | 4% |
| F13     | hsp 82 [Drosophila pseudoobscura] | gi|9069 | 77 | 1 | R. ELINSDSLALDKK(77) | 3% |
| F14*    | DB:Swissprot Frame:3 orf:2 Homolog:Calreticulin (gi|195107681) | GH986835 | 365 | 7 | K. FWTAGKFR(23) | 40% |
| B195\ B198^ | Evalute2e-69 Bitscore262 | | | | | |
| F38     | DB:Swissprot Frame:2 orf:1 Homolog:Calreticulin | GH986920 | 159 | 3 | R. EIIPMPAYK.G(19) | 47% |
| F21     | Heat Shock Protein family member (hsp-3) | gi|17568549 | 180 | 0 | R. LSPEEDIER.M(39) | 6% |
| F22, F23 | [Caenorhabditis elegans] | | | | | |
| F24     | heat shock protein 90-beta [Danio rerio] | gi|18858875 | 83 | 0 | K. HEFVEQLEFR.A(42) | 5% |
| F24     | heat shock protein 90 [Danio rerio] | gi|555574 | 62 | 1 | K. AIDLNNLTIAXS+Oxidation(M)(62) | 11% |
| F24     | DB:Swissprot Frame:2 orf:3 Homolog:Heat shock | EZ048788 | 214 | 3 | R. YMSITDPK.Q+Oxidation(M)(6) | 21% |
| F25     | protein 83 Evalute1e-84 Bitscore312 | | | | | |
| F25     | Heat shock protein HSP 90-alpha | gi|17865490 | 275 | 1 | R. ALLFPVR.M(20) | 8% |
| F27     | heat shock protein 70 [Liriomyza huidobrensis] | gi|99892741 | 225 | 3 | K. NQIHI6LVLGMSTAI.S(68) | 4% |

Table 2. Cont.
| Spot no. | Protein name | NCBInr Accession no. (a) | Tardigrade specific Accession no. (contig/EST) (b) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|--------------------------|-----------------------------------------------|-------------------|-----------------------------------|----------------------------------------|-----------------|
| Heat Shock Protein family member (hsp-1) | gi|17541098 | 219 | 2 | R. LSDKDDDR .M(52) | K. MDKQIIIHVDVLVGGSTR.I + Oxidation (M) (31) | 6% |
| [Caenorhabditis elegans]  | | | | | | | |
| F32    | DB:Trembl Frame:1 orf:3 | GH986605 | 19 | 1 | R. FEELCADFD. S (68) | K. VLASIDLTGK .T(19) | 9% |
| | Homolog:Lipoprotein-related protein Evaluate:/e-07 | | | | | | |
| | Bitscore:56.2 | | | | | | |
| | tropomyosin | gi|42559676 | 273 | 2 | R. EDSYERQIR .L (74) | K. IVELEEEER .V (55) | 17% |
| F35* | | | | | | | |
| F36*, F33* | homolog:: SwissProt Frame:1 orf:2 | GH986919 | 229 | 3 | K. ALKQREDSYQIR .L(18) | R. SSLQGSSQDEET .L (67) | 18% |
| | Homolog:Tropomyosin Evaluate:3e-11 Bitscore:66.6 | | | | | | |
| | 1, isoforms 9A/A/B Evaluate:3e-30 Bitscore:131 | | | | | | |
| | tropomyosin | homolog:: SwissProt Frame:2 orf:1 | GH98674 | 48 | 1 | R. LEDLQHEK .E (35) | K. LSEASQADESE .A (93) | 38% |
| | Homolog:Tropomyosin Evaluate:3e-11 Bitscore:66.6 | | | | | | |
| | 1, isoforms 9A/A/B Evaluate:3e-30 Bitscore:131 | | | | | | |
| | 1, isoforms 9A/A/B Evaluate:3e-30 Bitscore:131 | | | | | | |
| | tropomyosin | homolog:: SwissProt Frame:2 orf:1 | GH98674 | 48 | 1 | R. LEDLQHEK .E (35) | K. LSEASQADESE .A (93) | 38% |
| | Homolog:Tropomyosin Evaluate:3e-11 Bitscore:66.6 | | | | | | |
| | 1, isoforms 9A/A/B Evaluate:3e-30 Bitscore:131 | | | | | | |
| F36 | DB:Trembl Frame:2 orf1 Homolog:Putative | EZ048810 | 109 | 2 | R. DYPYNDISCT .R (51) | K. QGYFADEAQCGQVIR .R(58) | 28% |
| F33, F37, F64, F69, F71, | uncharacterized protein | | | | | | |
| | Evaluate:1e-18 Bitscore:97.1 | | | | | | |
| | tropomyosin | homolog:: SwissProt Frame:2 orf:1 | GH986609 | 53 | 2 | K. LDLQQAEFE .R (35) | K. LPIIIAANSLEDAAK .A (19) | 17% |
| | Homolog:: Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial E | | | | | | |
| F51 | DB:Trembl Frame:2 orf1 | EZ048819 | 58 | 1 | K. QGYFADEAQCGQVIR .R(58) | K. QGYFADEAQCGQVIR .R(58) | 16% |
| | Homolog:: AGAP009479- | | | | | | |
| E60, E63, E96, F31, F32, | | | | | | |
| Spot no. | Protein name | NCBInr Accession no. (contig/EST) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|----------------------------------|--------------------|-----------------------------------|-------------------------------------------|------------------|
| F52, F70, F73, F96 | Tardigrade specific Accession no. | EZ048790 | 62 | 1 | K.QAGFYADTEAQCVIR.R (62) | 18% |
| F40, F41, F66, F67, F68 | uncharacterized protein | Evaluate 2e-16 | 88.6 | | | |
| F55 | DB:Swissprot Frame:2 orf:1 Homolog:CD109 antigen | EZ048800 | 29 | 1 | R.TVVVYDYMQDR.K (29) | 20% |
| F56 | GTP-specific succiny-CoA synthetase beta subunit | gi|406564 | 59 | 1 | K.INFDONAER.R (59) | 5% |
| F57 | heat shock protein 60 [Salmon salar] | gi|6923167 | 63 | 1 | K.VGTSEVEVNEK.K (63) | 7% |
| F65* | 14-3-3 protein beta-alpha-2 (Protein 14-3-32) | gi|82089139 | 90 | 0 | R.NLLSVAYK.N (48) | 6% |
| F73 | DB:Swissprot Frame:2 orf:2 Homolog:60S ribosomal protein L26-1 Evaluate 6e-45 | GH98676 | 21 | 1 | K.LSKDSSSR.R (21) | 6% |
| F75 | DB:Swissprot Frame:2 orf:2 Homolog:Translationally-controlled tumor protein homolog Evaluate 1e-57 | GH98678 | 334 | 3 | K.LVDVFLFYTK.L (56) | 33% |
| F81 | DB:Swissprot Frame:2 orf:4 Homolog:Myosin | GH986792 | 331 | 6 | K.DTFAVLR.A (30) | 37% |
| F80, F89 | regulatory light polypeptide 9 Evaluate 4e-50 | GH986793 | 198 | | K.QGLDVYIAFAK.L (42) | |
| F81 | DB:Swissprot Frame:2 orf:1 Homolog:Cathepsin L1 Evaluate 2e-65 | GH98678 | 98 | 1 | K.LPDASENLVDCS.K (98) | 5% |
| F87 | DB:Swissprot Frame:3 orf:1 Homolog: Troponin C Evaluate 3e-60 | GH986791 | 24 | 1 | R.QGTLRR.T (24) | 4% |
| F88 | putative LEA III protein isoform 2 [Corylus avellana] | gi|14148981 | 69 | 1 | K.AGESQVDTANAAK.N (69) | 16% |
| Spot no. | Protein name | NCBInr Accession no. (\(\ast\)) | Tardigrade specific Accession no. (contig/EST) (\(\ast\)) | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage |
|----------|--------------|----------------------------------|----------------------------------|---------------------|-----------------------------------|--------------------------------------|------------------|
|          | glycosyl transferase, family 2 [Shewanella sediminis] | gi|157373461 | | 59 | 1 | R. HLLISLADK. Y (59) | 1% |
|          | HAW-EB3 | DB:Swissprot Frame:2 orf:5 Homolog: Myosin, | | | | | |
|          | essential light chain Evalue:8e-30 Bitscore:131 | EZ048813 | 404 | 10 | K. EVDEILR. L (20) | 58% |
|          |          |                                  | | | | R. HLLLISGER. L (59) | |
|          |          |                                  | | | | K. ESNUTIIAELR. H (68) | |
|          |          |                                  | | | | K. DVGTEDFMEAMR. V (73) | |
|          |          |                                  | | | | K. LTVIYWMYQQLSK. E (28) | |
|          |          |                                  | | | | R. VDKEKESNTIIAELR. H (26) | |
|          |          |                                  | | | | K. EDKVGTEDFMEAMR. V (23) | |
|          |          |                                  | | | | K. EVFGYDFPFDGDYIK. V (45) | |
|          |          |                                  | | | | K. KLTVIYWMYQQLSK. E (26) | |
|          |          |                                  | | | | K. EVFGYDFPFDGDYIK. V (37) | |
|          | F95 | SJCHGC06651 protein [Schistosoma japonicum] | gi|56759014 | | 103 | 0 | K. NTCCEFDGDILR. T (52) | 9% |
|          |          |                                  | | | | R. TVSIVAGPGVLDVYK. F (51) | |

Generated MS/MS data were searched against the NCBInr and tardigrade protein databases. Spot number, protein annotation, accession number, total protein score, number of matched peptides, peptide sequence and sequence coverage are listed. Identical proteins identified in different spots are listed only once and the spot with the highest protein score (in bold) is ranked at the top.

doi:10.1371/journal.pone.0009502.t002
Table 3. Identified proteins without annotation.

| Spot no. | Accession no. | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage | DomainSweep analysis |
|----------|---------------|---------------------|-----------------------------------|------------------------------------------|-------------------|---------------------|
| A11      | GH986700      | 52                  | 1                                 | -.VIAVSLPR.N(52)                         | 3%                | No hits             |
| A82, A88, B33, B41, B43, C50, D99, D105, E72, F87 |
| A11      | GH986755      | 32                  | 1                                 | -.LSISHNATLR.V(32)                       | 4%                | Putative            |
| A94      | GH986643      | 39                  | 1                                 | R.VDRSIPR.L(39)                          | 3%                | Putative            |
| A89, A95, A110, A123, A140, B49, B64, B83, B90, B98, B105, B155, B165, B173, B176, B185, B186, B187, B188, B189, B190, B191, B192, B193, B194, B195, C51, C128, C141, C153, D45, D46, D56, D57, D74, D123 |
| A100     | EZ048767      | 229                 | 4                                 | K.YDLIYK.G(15)                           | 20%               | Putative            |
| K.FLGFDTSAGK.T(61) | IPR017956 AT hook, DNA-binding, |
| K.IISFDVCNK.N(54) | conserved site |
| K.TDSGSCDVDTCRIPQTKJ.V(39) | IPR006689 ARF/SAR superfamily |
| K.AVDIEDIQPNNGQSIGYK.Y(60) | IPR005464 Psychosine receptor |
| A112     | GH986667      | 317                 | 5                                 | R.EQPTQPQGTCVR.G(61)                     | 22%               | Putative            |
| A114     | EZ048821      | 98                  | 2                                 | K.NADPLTLK.E(37)                         | 14%               | Putative            |
| K.IQSLCTPADLQ-PPQSTHDN.I(60) | IPR004835 Interferon-gamma receptor |
| K.IMEVCNPFNTENVNK.F + Oxidation (M) (44) | IPR000372 Leucine-rich repeat, cysteine-rich flanking region, N-terminal |
| K.IQSLCTPADLQ-PPQSTHDN.I(60) | IPR004825 Insulin/IGF/relexin |
| A112     | EZ048821      | 98                  | 2                                 | K.NADPLTLK.E(37)                         | 14%               | Putative            |
| A114     | EZ048817      | 49                  | 1                                 | R.IGETTDSFDYL.R.E(49)                    | 3%                | Putative            |
| A123     | EZ048785      | 221                 | 4                                 | K.FLDPTR.G(28)                           | 17%               | Putative            |
| R.AADDLTLK.L(57) | IPR000762 PTN/MK heparin-binding |
| R.YLMDQYDQ-MTR.S + Oxidation (M) (54) | protein |
| R.GTFTDAHILO-LQTALTTL.R.L(60) | |
| Spot no. | Accession no. | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage | DomainSweep analysis |
|---------|---------------|---------------------|-----------------------------------|----------------------------------------|------------------|---------------------|
| A136    | EZ048814      | 55                  | 1                                 | R. IMVLDLTDINS-APGMFDAAK.T + 2 - Oxidation (M) (23) | 5%               | Putative            |
| B48     | EZ048766      | 273                 | 5                                 | K. QVNAETFPQ.K.A(36)                   | 24%              | Putative            |
|         |               |                     |                                   | K. YSETVHYEGG.K.Q(39)                  |                  |                    |
| A157, A158, B49, B65 |         |                    |                                   | R. VDYVSYHTK.M(4)                     |                  |                    |
| B61     | GH986621      | 231                 | 6                                 | R. VLNNVGVLK.V(39)                    | 13%              | Putative            |
| B60, B62, B64, B65, B79, B84, |         |                    |                                   | R. VITVPKG1K.V(49)                    |                  |                    |
| B93, B112, B143 |         |                    |                                   | R. SLLGILPDITK.G(38)                  |                  | frame 4             |
| B84, B112, B138, B142, B143, |         |                    |                                   | R. VITVPKG1K.V(46)                    |                  |                    |
| B144, B161, B173 |         |                    |                                   | R. VITVPKG1KVESFK.S(26)               |                  | terminal            |
| B79, D67, D109 | GH986933     | 38                  | 1                                 | K. NGDVIQPR.Q(38)                     | 6%               | No hits             |
| B91     | GH986939      | 54                  | 1                                 | R. EALSAYTGG.R(62)                    | 9%               | No hits             |
| B102    | EZ048815      | 403                 | 6                                 | K. QVNAETFPQ.K.A(40)                  | 26%              | Putative            |
| A23, A24, A26, A112, A127, |         |                    |                                   | K. GGPWPAKDEK.F(17)                   |                  |                    |
| B99, B103, B105, B107, B108, |         |                    |                                   | K. IILFRPTLSAR.A(36)                  |                  |                    |

Table 3. Cont.
| Spot no. | Accession no. | Total protein score | No. of unique/ significant peptides | MS/MS peptide sequence (Indv. Ion score) | Sequence coverage | DomainSweep analysis |
|---------|---------------|---------------------|-----------------------------------|------------------------------------------|------------------|---------------------|
| B110, B111, B144 | | | | R.AQGLHEATTEGK.N(68) | | IPR002181 Fibrinogen, alpha/beta/gamma chain, C-terminal globular |
| | | | | R.LIPDELLGTTEFSQK.Q(92) | | |
| | | | | R.RLIPDELLGTTEFSQK.Q(36) | | IPR000463 Cytosolic fatty-acid binding |
| | | | | K.DYEFKEDGMQMTAK.F + Oxidation (M) (20) | | |
| | | | | K.EVEYSTNYMALDTVK.A(51) | | |
| | | | | R.EMLGWESTSEQ ENMELYK.A(22) | | |
| | | | | R.GDKPGLAAFGDNIIEYSETDSEGETGVLSQK.F(21) | | |
| B103 | EZ048768 | 40 | 1 | R.VTTTVSPR.I(40) | 3% | No hits |
| B150 | GH986581 | 108 | 3 | R.VFVEEQLK.A(33) | 14% | Putative |
| B151, B173 | | | | R.FNFLVFLGSTR.E(46) | | IPR000990 Innexin |
| B152 | EZ048775 | 42 | 1 | R.KLEFILXIF.-(42) | 5% | Putative |
| | | | | | | IPR003061 Colicin E1 (microcin) immunity |
| | | | | | | Protein |
| | | | | | | IPR000048 IQ calmodulin-binding region |
| B179 | GH986603 | 53 | 1 | R.AFEVPASECGK.S(53) | 5% | Putative |
| | | | | | | PR015880 Zinc finger, C2H2-like |
| B191 | EZ048789 | 26 | 1 | K.GSIGAPDVPK.N(26) | 4% | Putative |
| | | | | | | IPR001955 Pancreatic hormone |
| B186 | GH986708 | 468 | 6 | R.AFEVPASECGK.S(46) | 25% | Putative |
| A140 | | | | R.AFEVPASECGKSPK.R(82) | | IPR015880 Zinc finger, C2H2-like |
| | | | | | | IPR000436 Sushi/SCR/CCP |
| | | | | | | K.IVSKDVOGSSFPXR.K(90) | | |
| | | | | | | IPR008264 Beta-glucanase |
| | | | | | | R.SESGALNSEQTXAK.F(62) | | |
| | | | | | | IPR000008 C2 calcium-dependent |
| | | | | | | R.SESGALNSEQEKTHFR.D(137) | | |
| | | | | | | membrane targeting |
| | | | | | | R.VQVMKDVOGSSDLVEQ-PECLTPFLVSSR.S<oxidation (15) | | |
| C18 | EZ048777 | 46 | 1 | R.NLADQAMSMEGPNLFAK.A + 2 Oxidation (M) | 8% | Putative |
| | | | | | | IPR003569 Cytochrome c-type biogenesis |
| | | | | | | Protein CcbS |
| | | | | | | IPR002282 Platelet-activating factor |
| Spot no. | Accession no. | Total protein score | No. of unique/significant peptides | MS/MS peptide sequence (Indiv. Ion score) | Sequence coverage | DomainSweep analysis |
|---------|---------------|---------------------|------------------------------------|----------------------------------------|-------------------|---------------------|
| C78     | GH986847      | 32                  | 1                                  | K. SEVFPRIR.S(32)                      | 3%                | Putative            |
| B188, B173, C141 |            |                     |                                    |                                        |                   | NADH-ubiquinone     |
|         |               |                     |                                    |                                        |                   | oxidoreductase, chain 5 |
| C86     | GH986916      | 196                 | 4                                  | K. NPYLELTDPK.-(38)                    | 12%               | Putative            |
|          |               |                     |                                    | K. TPEESEAPQAIR.R(68)                  |                   | Sulfotransferase    |
|          |               |                     |                                    | K. TPEESEAPQAIR.R.K(58)                |                   | Glial cell line-derived |
|          |               |                     |                                    | K. VEKTPEESEAPQAIR.R(32)               |                   | neurotrophic factor receptor alpha 2 |
| C95     | GH986921      | 35                  | 1                                  | -. VIAVSLPR.N(30)                      | 2%                | No hits             |
| B18, B19, B47, B49, B138, C51, C62, C65, D107 |            |                     |                                    |                                        |                   |                     |
| C110    | GH986711      | 31                  | 1                                  | K. TALITGASTGIGR.A(31)                | 5%                | Significant         |
|          |               |                     |                                    |                                        |                   | Glucose/ribitol dehydrogenase |
|          |               |                     |                                    |                                        |                   | Short-chain dehydrogenase/reductase SDR |
|          |               |                     |                                    |                                        |                   | Putative            |
|          |               |                     |                                    |                                        |                   | 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase |
|          |               |                     |                                    |                                        |                   | 3-beta hydroxysteroid dehydrogenase/isomerase |
| C118    | EZ048824      | 45                  | 0                                  | K. DSVAIGFPK.D(24)                    | 7%                | Putative            |
|          |               |                     |                                    | K. ADEAGFTDAIK.A(21)                  |                   | Intimin bacterial adhesion mediator protein |
| C141    | EZ048801      | 395                 | 6                                  | R. NQYQSMER.H(34)                     | 22%               | Putative            |
| C117, C145 |           |                     |                                    |                                        |                   | Myogenic basic muscle- |
|          |               |                     |                                    |                                        |                   | specific protein    |
|          |               |                     |                                    |                                        |                   | Myogenic basic muscle- |
|          |               |                     |                                    |                                        |                   | Protein synthesis factor, GTP-binding |
| C141    | GH986597      | 27                  | 1                                  | -. TAVEAVVR.T(27)                     | 4%                | Putative            |
|          |               |                     |                                    |                                        |                   | Invasion protein B  |
| C156    | EZ048804      | 277                 | 5                                  | K. QFPFFISAK.H(43)                    | 27%               | Putative            |

Table 3. Cont.
Three different cathepsin proteins could be identified: cathepsin K (spot A84), cathepsin Z (spot E80) and cathepsin L1 (spot F81). Cathepsin L is a ubiquitous cysteine protease in eukaryotes and has been reported as an essential protein for development in *Xenopus laevis* [32], *Caenorhabditis elegans* [33] and *Artemia franciscana* [34].

Several protein spots are associated with ATP generation and consumption and may have important roles in the early development as described for *Artemia*, because many important metabolic processes require ATP [35,36]. ATP synthase (spot B152) regenerates ATP from ADP and Pi [37]. It consists of two parts: a hydrophobic membrane-bound part (CF0) and a soluble part (CF1) which consists of five different subunits, alpha, beta (spot E39), gamma, delta (spot C139) and epsilon. Arginine kinase (spot B167) is an ATP/guanidine phosphotransferase that provides ATP by catalyzing the conversion of ADP and phosphorylarginine to ATP and arginine [38]. The presence of arginine kinase has been shown in tissues with high energy demand [39].

Interestingly, we could identify the translationally controlled tumor protein (TCTP) (spot F75) on the 2D gel. TCTP is an important component of TOR (target of rapamycin) signalling pathway, which is the major regulator of cell growth in animals and fungi [40].
Evaluation of Heat Shock Proteins by Western Blot Analysis

To evaluate the highly conserved heat shock proteins 60 and 70, we performed Western blot analyses with antisera directed against these proteins. Hsp70 was found in several spots on the reference 2D proteome map, e.g. in spot B172, C31, C133 and F27. None of these spots fits well to the calculated molecular weight of approx. 70 kDa, most of them were considerably smaller. In contrast, the
imunoblot shows the strongest band at the expected position which is in agreement with the position of hsp70 in the control lysate of HeLa cells (Figure 8B). However, several additional bands can be observed at higher as well as at lower molecular weights. The lower bands might account for the identified spots on the 2D gel with lower molecular weight. The full-length protein might have escaped the spot picking procedure since only a limited number of detected spots were further processed.

Hsp60 was identified in spot F57 of the 2D map as described above. Since hsp60 was identified by only one peptide hit we confirmed this result by immunostaining using an antibody directed against a peptide in the C-terminal region of the entire protein. Only one band is visible on the Western blot at approx. 24 kDa whereas the protein band in the HeLa control lysate is located at its expected position (Figure 8A). The lower molecular weight is in accordance with the location of hsp60 (spot F57) on the 2D gel. Thus, in M. tardigradum hsp60 exists in a significantly shorter form. Whether the observed difference in the molecular weight indicates a different function and role of this protein in M. tardigradum needs to be investigated in future experiments. To test whether other tardigrade species show similar results we performed an immunoblot with protein lysates from 5 other species namely Paramacrobiotus richtersi blot with protein lysates from 5 other species namely Paramacrobiotus richtersi, Macrobiotus tonollii, Paramacrobiotus “richtersi group” 3, Macrobiotus tonollii, Paramacrobiotus “richtersi group” 2 and Paramacrobiotus “richtersi group” 1. Total protein lysate from HeLa cells was loaded as control (Figure 9A, lane 1). Actin served as loading control for all lysates (Figure 9B). Interestingly, some species also exhibit truncated forms of hsp60 on the Western blot whereas others show higher forms. The molecular weights are ranging from approx. 75 kDa for P. “richtersi group” 2 and P. “richtersi group” 1 lysates (Figure 9A, lane 4 and 6), 35 kDa for P. “richtersi group” 3 and P. richtersi lysates (Figure 9A, lane 3 and 5) down to 24 kDa for M. tardigradum and M. tonollii (Figure 9, lane 3 and 7).

Discussion

Establishing a Comprehensive Proteome Map of Milnesium tardigradum

The analysis of the proteome of M. tardigradum represents to our knowledge the first detailed study of tardigrades on the protein level. Our experimental strategy aimed to identify as many as possible proteins from tardigrades. Thus, we have not employed any subcellular fractionation steps to obtain specific subproteomes. We have tested various protocols for protein extraction from whole tardigrades. We could show that direct homogenisation of tardigrades in lysis buffer without any previous precipitation steps is most efficient and enables the generation of high quality 2D gels. Since nothing was known about the proteolytic activity in M. tardigradum special precautions were taken to avoid any protein degradation or proteolysis throughout the whole workup procedure. Integrity of proteins was carefully inspected by Western blot analysis of the two housekeeping proteins actin and tubulin where the sequence homology was assumed to be high enough to detect the proteins with commercially available antibodies. The development of a robust workup protocol laid the basis for the generation of a protein map from whole tardigrades in the active state. 56 unique proteins could be identified by searching high
quality MS/MS spectra against the publicly available NCBInr database. However, for many proteins we could not find any homologues in the NCBInr database and only by using our own newly generated tardigrade protein database it was possible to identify another 73 unique proteins. 15 proteins were present in both databases. In addition 36 unique proteins were found in the clustered tardigrade EST database which could not be annotated by BLAST search. This concerns new specific proteins of *M. tardigradum*.

**Performance of Database Searches**

When we started our study of the tardigrade proteome very little was known about tardigrades at the genome and gene expression level. To this day, only 12 proteins are recorded in the NCBInr database, which originate from *M. tardigradum*. For all of them only partial sequences ranging from as few as 43 amino acids for beta actin up to 705 amino acids for elongation factor-2 are available. Therefore, in parallel to our proteomic study a *M. tardigradum* EST sequencing project has been initiated. Subsequently, two tardigrade specific databases have been established: a clustered tardigrade EST database and a tardigrade protein database which was extracted from the clustered EST database and thus represents a subdatabase containing all tardigrade-specific proteins with annotated function. However, since cDNA sequencing is still ongoing sequence information remains incomplete. We assume that the tardigrade database currently covers approximately one tenth of the tardigrade specific genes comparing the unique clusters found in tardigrades to all known proteins of *Caenorhabditis elegans* or *Drosophila melanogaster* in Ensembl. This fact is greatly influencing our database searches. For most of the protein spots that were analysed by ESI-MS/MS high quality fragmentation spectra were obtained from MS/MS experiments. However, when we searched these MS/MS data against the tardigrade databases and the publicly available NCBInr database, only about 70% of the spots yielded in protein identification whereas the remaining spots gave no significant protein hit. In addition it was impossible to manually extract peptide sequences that were sufficient in length to perform BLAST searches with satisfactory results.

When we examined the protein hits obtained by the three databases in more detail we found that in the NCBInr database approximately one half of the proteins were identified by only one significant peptide hit (Figure 10). For about 25% of the proteins more than one significant peptide hit was obtained. For the remaining 25% only the protein score which is the sum of two or more individual peptides scores was above the significance threshold while none of the peptide scores alone reached this value. In contrast, proteins found in the tardigrade protein database were predominantly identified by more than one significant peptide hit whereas a smaller number was represented by only one peptide. In no cases a protein was identified by the sum of non-significant peptide matches. For proteins without annotation the number of proteins identified by only one peptide was only slightly higher than the number of proteins identified by two or more peptides.

These results are not surprising. Since the NCBInr database contains very few sequences originating from *M. tardigradum* e.g. elongation factor 1-alpha the identification relies predominantly on high homologies between tardigrade sequences and sequences from other more or less related species of other taxa. The chances for detecting more than one identical peptide is significantly higher when searching MS/MS data against the tardigrade EST and tardigrade protein databases since these databases contain only tardigrade specific sequences.

![Proteome Map of Tardigrades](image)

**Proteins Associated with Anhydrobiosis**

Among the numerous proteins which were identified in our proteomic study some proteins have already been reported to play an important role in anhydrobiotic organisms. Most importantly, spot F88 was identified as a protein belonging to the LEA (late embryogenesis abundant) family (group 3). This result was already known from Western blot analyses (Schill et al., 2005, poster presentation, ISEPEP, Denmark). At least six different groups of LEA proteins have been described so far. Group 1, 2 and 3 are the three major groups. Whereas group 1 is only found in plants and group 2 predominantly in plants, group 3 is reported in organisms other than plants. Although the precise role of LEA proteins has not yet been fully elucidated, different research groups have reported on their association with tolerance to water stress by desiccation [41,42]. LEA protein of group 3 could be already identified in nematodes *C. elegans*, *Strongylus vulgaris* and *Aphelenchus avenae*, and the prokaryotes *Deinococcus radiodurans*, *Bacillus subtilis* and *Haemophilus influenzae* [43,44,45].

**Proteins Exhibiting an Unusual Location on the 2D Map**

In general we identified some proteins which show a lower molecular weight than expected. As described above hsp60 is detected as a protein band at 24 kDa by Western blotting and the location of the corresponding protein on the 2D gel shows the same molecular weight. Comparison of different tardigrade species indicates the existence of short as well as long forms of hsp60.
Unique proteins, when analyzed on the 2D gel, often show multiple spots due to posttranslational modifications. Proteins of the vitellogenin family are widely distributed on the 2D gel and show pl as well as molecular weight shifts, which are due to modification through cleavage and to different PTMs like glycosylation and phosphorylation during development of oocytes. Ongoing experiments to detect PTMs using different fluorescence staining methods like ProQ-Emerald for the detection of glycoproteins and ProQ-Diamond for the detection of phosphoproteins indicate that these modifications indeed occur in tardigrades (data not shown).

Prediction of Functional Domains in Proteins with Yet Unknown Functions

36 proteins which could not be identified by BLAST searches were further examined looking for matching functional protein domains with DomainSweep. The function of the following two spots could be revealed with high confidence (Table 3): spot F63 seems to belong to the “tumor protein D52” interpro family (IPR007327). The hD52 gene was originally identified through its elevated expression level in human breast carcinoma, but cloning of D52 homologues from other species has indicated that D52 may play roles in calcium-mediated signal transduction and cell proliferation. Regarding the taxonomic neighbours of the tardigrades, one member in C. elegans and 10 members in Drosophila melanogaster are reported by Interpro for this family. Spot C95 seems to belong to the family “glucose/ribitol dehydrogenase” (IPR002347). 80 members both in C. elegans and in Drosophila melanogaster are reported for this family. 28 putative hits were found associated with other spots. These protein hits are putative candidates and therefore less reliable. A comprehensive protein database of M. tardigradum as the result of our ongoing cDNA sequencing will help us to evaluate these candidates.

Conclusion

In this study we present for the first time a comprehensive proteome map of M. tardigradum. A full description of proteins present in the active state provides a valuable basis for future studies. Most importantly, the protein reference map allows us to undertake quantitative proteomics analysis to detecting proteins with different expression levels in the active versus the anhydrobiotic state. In particular, our workflow is fully compatible with the application of 2D difference gel electrophoresis (2D DIGE), which is one technique allowing sensitive analysis of differences in the protein expression levels. This differential analysis on the protein level will help us to understand survival mechanisms in anhydrobiotic organisms and eventually to develop new methods for preservation of biological materials.

Materials and Methods

Tardigrade Culture and Sampling

Tardigrades of the species M. tardigradum Doyère 1840 were maintained in a laboratory culture. The culture was grown on agarose plates (3%) (peqGOLD Universal Agarose, peqLAB, Erlangen Germany) covered with Volvic™ water (Danone Waters, Wiesbaden, Germany) at 20°C. The juveniles were fed on green algae Chlorogonium elongatum, the adults with bdelloid rotifers Philodina citrina. The specimens for the experiments were all of middle-age, thus effects of age can be excluded. Tardigrades were starved over 3 days and washed several times with Volvic™ water to avoid contamination with food-organisms. Subsequently the animals were transferred to microliter tubes (200 individuals per tube) and surrounding water was reduced to approx. 1–2 µl.

An active state [1] according to Schill et al. [18] was investigated in this work. All samples were shock frozen in liquid nitrogen and stored at −80°C. 200 individuals are defined as one aliquot. Other tardigrade species (Paramacrobiotus richtersi, Paramacrobiotus “richtersi group” 3, Macrobiotus tonolithi, Paramacrobiotus “richtersi group” 2 and Paramacrobiotus “richtersi group” 1) used for immunodetection of hsp60 were prepared in the same way.

Sample Preparation for Gel Electrophoresis

To optimize the sample preparation different precipitation methods have been tested. Chloroform/methanol and TCA/acetone precipitations were performed as described by Wessel, Flueggé [46] and Gorg [47], respectively. We used also the commercially available precipitation kit (clean-up kit from GE Healthcare). Comparing the result of different precipitation protocols on a 1D gel we decided to homogenise the tardigrades directly in ice cold lysis buffer and avoid any precipitation steps. The animals (200 individuals) were homogenised directly in 60 µl lysis buffer (containing 8 M urea, 4% CHAPS, 30 mM Tris, pH 8.5) by ultrasonication (SONOPULIS, HD3100, Bandelin Electronic) with 45% amplitude intensity and 1–0.5 sec intervals. The lysis buffer contained a Protease Inhibitor Mix (GE Healthcare) to inhibit serine, cysteine and calpain proteases. After homogenisation the samples were stored at −80°C. For gel electrophoresis insoluble particles were removed by centrifugation for 2 min at 14,000 g and the supernatant was quantified using BCA mini-assay.

One Dimensional Gel Electrophoresis and Western Blotting

To compare the efficiency of different sample preparation methods we separated approx. 10 µg total protein extract on a 1D gel. The gel was stained with protein staining solution (PageBlue from Fermentas). For Western blotting a total protein extract of tardigrades (15–20 µg) was separated on a NuPAGE™ 4–12% Bis-Tris mini gel (Invitrogen) using MES running buffer. 200 V were applied until the bromophenol blue front had reached the bottom of the gel (approx. 40 min). Separated proteins were electro transferred onto PVDF membrane for 1.5 h at maximum 50 mA (0.8/cm²) in a semi-dry transfer unit (Hoeffer™ TE 77) using following transfer solution: 24 mM Tris, 192 mM glycine and 10% methanol. The PVDF membrane was incubated in a blocking buffer containing 5% non-fat milk, 0.1% Tween20 in PBS. As primary antibodies we used anti actin pan Ab-5 (dianova), anti hsp 60 Ab (D307) (Cell signaling), anti hsp70 Ab (BD Biosciences Pharmingen) and anti α-Tubulin Ab (Sigma).

For molecular weight determination of the target proteins on film we used ECL DualVue marker (GE-Healthcare). Immuno-reaction was detected using the ECL Western Blotting Detection kit from GE Healthcare. Images were acquired using an Image Scanner Model UTA-1100 (Amersham Biosciences).

Two Dimensional Gel Electrophoresis

For 2D gel preparation we added 60 µl 2x sample buffer (7 M urea, 2 M thiourea, 2% CHAPS, 2% DTT, 2% IPG-buffer 3–11 NL) to each aliquot and incubated by shaking for 30 min at 25°C. To avoid streaking on the gels we used 330 µl destreaking buffer (GE Healthcare) instead of rehydration buffer, to which we added 2% IPG-buffer (pI 3–11). Samples were incubated by shaking for 30 min at 25°C. We loaded 100 µg protein on analytical gels and 330 µg on preparative gel.

Strip loading. Loading of proteins was performed during strip rehydration with the recommended volume (450 µl for 24 cm strips) over night.
IEF conditions. First dimension isoelectric focusing (IEF) was performed, using 24 cm long IPG strips with non-linear gradients from pH 3–11 and an Etan IPGphor instrument and proceeded for 46.4 kVh with the following running protocol: 3 h at 300 V, 6 h at 500 V, 8-h gradient up to 1000 V, 3-h gradient up to 8000 V and 3 h at 8000 V. Strips were either immediately used for the second dimension or stored at −80°C.

Second dimension. Strips were equilibrated in 6 M urea, 2% SDS, 30% glycerol, 0.375 M Tris-HCl pH 8.8, 0.002% bromophenol blue and 10 mg/ml DTT for 15 min, followed by a second equilibration step with the same buffer containing 25 mg/ml iodoacetamide instead of DTT, also for 15 min.

Strips were loaded on 12% SDS-gels with an overlay of agarose solution (0.5 mg/100 ml electrophoresis buffer). The second dimension was performed using an Etan Daltwelve electrophoresis system (GE Healthcare). Separation was carried out at 1.5 watt/1.5 mm thick gel until the bromophenol blue reached the bottom of the gel (approx. 18 h).

Silver staining of proteins and image analysis. Proteins on analytical gels were visualized by destructive silver staining according to Blum [48]. Additionally, we performed a silver stain compatible with mass spectrometric analysis described by Sinha [49] for preparative gels. Images were acquired using an Image Scanner Model UTA-1100 (Amersham Biosciences).

Protein Identification

In-gel digestion. Protein spots were excised semi-manually with a spot picker (GePal, Genetix) following non-destructive silver staining and stored at −80°C after removing water. Gel pieces were reduced, alkylated and in-gel digested with trypsin. Briefly, after incubation with 150 µl water at 42°C for 8 min, water was removed (washing step) and gel pieces were shrunk by dehydration with 150 µl 40 mM NH4HCO3/ethanol 50:50 (v/v) at 42°C for 5 min in a thermo mixer (600 rpm). The solution was removed and the proteins were reduced with 50 µl 10 mM dithiothreitol in 40 mM NH4HCO3 for 1 h at 36°C. The solution was removed and gel pieces were incubated with 150 µl 40 mM NH4HCO3 for 5 min at 42°C. After removing the solution gel pieces were alkylated with 100 µl 55 mM iodoacetamide in 40 mM NH4HCO3 for 30 min at 25°C in the dark, followed by three alternating washing steps each with 150 µl of 40 mM NH4HCO3 and ethanol for 5 min at 37°C. Gel pieces were then dehydrated with 100 µl neat acetonitrile for 1 min at room temperature, dried for 15 min and subsequently rehydrated with porcine trypsin (sequencing grade, Promega, Mannheim, Germany) with the minimal volume sufficient to cover the gel pieces after rehydration (100 ng trypsin in 40 mM NH4HCO3). Samples were incubated over night at 37°C.

Extraction. After digestion over night the supernatant was collected in PCR-tubes while gel pieces were subjected to four further extraction steps. Gel pieces were sonicated for 5 min in acetonitrile/0.1% TFA 50:50 (v/v). After centrifugation the supernatant was collected and gel pieces were sonicated for 5 min in acetonitrile. After collecting the supernatant gel pieces were sonicated for 5 min in 0.1% TFA followed by an extraction step again with acetonitrile. The combined solutions were dried in a speed-vac at 37°C for 2 h. Peptides were redissolved in 6 µl 0.1% TFA by sonication for 5 min and applied for ES/MS/MS analysis.

ESI-MS/MS analysis and database search. NanoLC-ESI-MS/MS was performed on a Q10 Ultima mass spectrometer (Waters) coupled on-line to a nanoLC system (CapLC, Waters). For each measurement 5 µl of the digested sample was injected. Peptides were trapped on a Trapping guard C18-AQ, 10 mm×0.3 mm, particle size 5 µm (Dr. Maisch). The liquid chromatography separation was performed at a flow rate of 200 nL/min on a Reprosil C18-AQ column, 150 mm×75 µm, particle size 3 µm (Dr. Maisch GmbH). The following linear gradient was applied: 5% B for 5 min, from 5 to 15% B in 5 min, from 15 to 40% B in 25 min, from 40 to 60% B in 15 min and finally 60 to 95% B in 5 min. Solvent A contains 94.9% water, 5% acetonitrile, 0.1% formic acid, solvent B contains 95% acetonitrile, 4.9% water and 0.1% µl formic acid. The LC-ESI-MS/MS device was adjusted with a PicoTip Emitter (New Objective, Woburn, MA) fitted on a Z-spray nanoESI interface (Waters). Spectra were collected in the positive ion mode. The capillary voltage was set to 2400 V and the cone voltage was set to 80 V. Data acquisition was controlled by MassLynxTM 4.0 software (Waters). Low-energy collision-induced dissociation (CID) was performed using argon as a collision gas (pressure in the collision cell was set to 5×10−5 mbar), and the collision energy was in the range of 25–40 eV and optimized for all precursor ions dependent on their charge state and molecular weight. Mass Lynx raw data files were processed with Protein Lynx Global Server 2.2 software (Waters). Deisotoping was performed using the MaxEnt3 algorithm.

The obtained MS/MS spectra were searched against the publicly available NCBInr database using the MASCOT algorithm version 2.0 (Matrix Science, London, UK). The mass tolerance was set to 0.1 Da for fragment ions and 200 ppm for precursor ions. No fragment ions score cutoff was applied. The following search parameters were selected: variable modification due to methionine oxidation, fixed cysteine modification with the carbamidomethyl-side chain, one missed cleavage site in the case of incomplete trypsin hydrolysis. The following settings were applied: minimum protein score >53, minimum number of peptides ≥1. Furthermore, protein hits were taken as identified if a minimum of one peptide had an individual ion score exceeding the MASCOT identity threshold. Under the applied search parameters a sum MASCOT score of >53 refers to a match probability of p<0.05, where p is the probability that the observed match is a random event. Redundancy of proteins that appeared in the database under different names and accession numbers was eliminated. Additionally we searched against the M. tardigradum EST and protein database (see below) to identify sequences not present in the NCBInr databases. The following settings were applied: minimum protein score >14 for the EST and >27 for the clustered EST database (p<0.05). Other parameters were as described for the NCBInr searches.

Generation of the Tardigrade EST Database

cDNA libraries from mRNA from tardigrades in different states (active, inactive, transition states) were prepared and sequenced (Mali et al, submitted data). The obtained EST sequences were cleaned from vector sequences using Seqclean against UniVec-database from NCBI (version 12. September 2008, Kitts et al., unpublished). Repeats within the cleaned ESTs were masked using the online service RepeatMasker (version 3.2.6, RM-20080801, Smit et al., unpublished data) followed by a second Seqclean run to eliminate low quality and short sequences. The assembly was performed using cap3 [30] with clipping enabled and resulted in 3318 Unigenes (2500 singlets, 818 contigs). Identification of ribosomal sequences was done using a BlastN-search [51] against the Silva-DB (only eukaryotic sequences, Silva95, [52]) and an E-value cutoff of 1e-3 and resulted in 46 sequences which showed high similarity to ribosomal sequences. Unigenes coding for known proteins were identified using a BlastX search against Uniprot/Swissprot (version 14.1, September 2008), Uniprot/TrEMBL.
Protein Domain Analysis of Proteins without Annotation

specific Accession no.''.

indicated in the Tables 2, 3 and S1 in the column “Tardigrade substituted by an undefined amino acid (X). All new sequences mode of Virtual Ribosome was used. Afterwards stop codons were substituted by an undefined amino acid (X). All new sequences have been deposited in GenBank. The accession numbers are indicated in the Tables 2, 3 and S1 in the column “Tardigrade specific Accession no.”.

Classification of Proteins

For functional analysis of identified proteins we used Blast2GO software, which consists of three main steps: blast to find homologous sequences, mapping to collect GO-terms associated to blast hits and annotation to assign functional terms to query sequences from the pool of GO terms collected in the mapping step [53]. Function assignment is based on GO database. Sequence data of identified proteins were uploaded as a multiple FASTA file to the Blast2GO software. We performed the blast step against public database NCBI through blastp. Other parameters were kept at default values: e-value threshold of 1e-3 and a recovery of 20 hits per sequence. Furthermore, minimal alignment length (hash filter) was set to 25 to avoid hits with matching region smaller than 100 nucleotides. QBlast-NCBI was set as Blast mode. Furthermore, we have chosen an annotation configuration with e-value-Hit-filter of 1.0E-6, Annotation Cutoff of 55 and GO weight of 5. For visualizing the functional information (GO categories: Molecular Function and Biological process) we used the analysis tool of the Blast2GO software.

Protein Domain Analysis of Proteins without Annotation

Six frame translations of the Unigenes were run through the DomainSweep pipeline [56] and the significant and putative hits were collected. For each of the protein/domain databases used, different thresholds and rules were established [56]. Domain hits are listed as ‘significant’

i. if two or more hits belong to the same INTERPRO [57] family. The task compares all true positive hits of the different protein family databases grouping together those hits, which are members of the same INTERPRO family/domain.

ii. if the motif shows the same order as described in PRINTS [58] or BLOCKS [59]. Both databases characterize a protein family with a group of highly conserved motifs/segments in a well-defined order. The task compares the order of the identified true positive hits with the order described in the corresponding PRINTS or BLOCKS entry. Only hits in correct order are accepted.

All other hits above the trusted thresholds are listed as ‘putative’. By comparing the peptides which were identified by mass spectrometry with the six translations, the correct frame and the associated domain information was listed.

Supporting Information

Table S1 Blast2GO analysis of identified proteins. Spot number, protein annotation, accession number and GO information in all three categories molecular function, biological process and cellular component are listed.

Found at: doi:10.1371/journal.pone.0009502.s001 (0.16 MB XLS)

Acknowledgments

Many thanks to Kerstin Kammerer and Eva Zettler for optimal technical assistance and Tore Kempf for his helpful advice. This article is part of the Ph.D. thesis of Elham Schokraie, Darmstadt University of Technology.

Author Contributions

Conceived and designed the experiments: ES. Performed the experiments: ES. Analyzed the data: ES AHV. Contributed reagents/materials/analysis tools: AHV UW BM FF TD SH ROS. Wrote the paper: ES MS. Supervised the project: MS.

References

1. Keilin D (1959) The problem of anabiosis or latent life: history and current concept. Proc R Soc Lond B Biol Sci 130: 149–191.

2. Baumann H (1922) Die Anabiose der Tardigraden. Zoologischen Jahrbücher. pp 501–556.

3. Nelson DR (2002) Current status of the tardigrada: evolution and ecology. Integrative and Comparative Biology 42: 652–659.

4. Schill RO, Fritz GB (2008) Desiccation tolerance in embryonic stages of the tardigrade Milnesium tardigradum. Journal of Zoology (London) 276: 103–107.

5. Hengherr S, Brummer F, Schill RO (2008) Anhydrobiosis in tardigrades and its effects on longevity traits. Journal of Zoology (London): in press.

6. Ramlov H, Westh P (2001) Cryptobiosis in the eutardigrade Adelobiasis (Richardsia) conrovers. Tolerance to alcohols, temperature and de novo protein synthesis. Zoologischer Anzeiger 249: 517–523.

7. Ramlov H, Westh P (1992) Survival of the cryptobiotic eutardigrade: Adelobiasis Richardsia conovers during cooling to minus 196 C: effect of cooling rate, trehalose level, and short-term acclimation. Cryobiology 29: 125–130.

8. Hengherr S, Worland MR, Reuner A, Brummer F, Schill RO (2009) Freeze tolerance, supercooling points and ice formation: comparative studies on the subzero temperature survival of limno-terrestrial tardigrades. J Exp Biol 212: 802–807.

9. Seki K, Toyoshima M (1998) Preserving tardigrades under pressure. Nature 395: 853–854.

10. Horikawa DD, Sakashita T, Katagiri C, Watanabe M, Kikawada T, et al. (2006) Tardigrade Myosin chains. Proc R Soc Lond B Biol Sci 273: R729–R731.

11. Suzuki AC (2003) Life history of Milnesium tardigradum. International Journal of Radiation Biology 80: 1415–1422.

12. Jonsson KI, Rabbow E, Schill RO, Harms-Ringdahl M, Rethberg P (2008) Tardigrade survive exposure to space in low Earth orbit. Current Biology 18: R279–R371.

13. Clegg JS, Jackson SA (1992) Aerobic heat shock activates trehalose synthesis in embryos of Artemia franciscana. FEBS Lett 303: 45–47.

14. Caprioli M, Krabbe Katholm A, Melone G, Ramlov H, Ricci C, et al. (2004) Trehalose in desiccated rotifers: a comparison between a bdelloid and a monogonont species. Comp Biochem Physiol A Mol Integr Physiol 139: 527–532.

15. Lappin F, Tunnacliffe A (2003) Anhydrobiosis without trehalose in bdelloid rotifers. FEBS Lett 535: 387–390.

16. Tunnacliffe A, Lapinski J (2003) Resurrecting Van Leeuwenhoek’s rotifers: a reappraisal of the role of disaccharides in anhydrobiosis. Philos Trans R Soc Lond B Biol Sci 358: 1755–1771.

17. Suzuki AG (2003) Life history of Milnesium tardigradum. Doyere (tardigrada) under a rearing environment. Zoological Science (Tokyo) 20: 49–57.

18. Schill RO, Steinhuck GH, Kohler HR (2004) Stress gene (hp70) sequences and quantitative expression in Milnesium tardigradum (Tardigrada) during active and cryptobiotic stages. J Exp Biol 207: 1607–1613.

19. Hengherr S, Heyer AG, Kohler HR, Schill RO (2008) Trehalose and anhydrobiosis in tardigrades evidence for divergence in responses to dehydration. FEBS J 275: 201–208.

20. Martin RM, Chilton NB, Lightowlers MW, Gasser RR (1997) Echinococcus granulosus myophilin–relationship with protein homologues containing “calpo- nin-motifs”. Int J Parasitol 27: 1361–1367.

21. Kato Y, Tokishita S, Ohta T, Yamagata H (2004) A vitellogenin chain containing a superoxide dismutase-like domain is the major component of yolk proteins in cladoceran crustacean Daphnia magna. Gene 354: 157–165.

Supporting Information

Table S1 Blast2GO analysis of identified proteins. Spot number, protein annotation, accession number and GO information in all three categories molecular function, biological process and cellular component are listed.

Found at: doi:10.1371/journal.pone.0009502.s001 (0.16 MB XLS)

Acknowledgments

Many thanks to Kerstin Kammerer and Eva Zettler for optimal technical assistance and Tore Kempf for his helpful advice. This article is part of the Ph.D. thesis of Elham Schokraie, Darmstadt University of Technology.

Author Contributions

Conceived and designed the experiments: ES. Performed the experiments: ES. Analyzed the data: ES AHV. Contributed reagents/materials/analysis tools: AHV UW BM FF TD SH ROS. Wrote the paper: ES MS. Supervised the project: MS.
