ORIGINAL RESEARCH PAPER

Biotechnology

Volume - 7 | Issue - 1 | January - 2017 | ISSN - 2249-555X | IF : 3.919 | IC Value : 79.96

Identification and in-silico analysis of a small heat shock protein (HSP26) expressed in the *Bombyx mori* due to heat shock

KEYWORDS

Bombyx mori, small heat shock protein, thermal stress.

Punyavathi

Proteomics and Genomics Laboratory, Department of Studies in Sericulture Science, University of Mysore, Mysore 570006, Karnataka, India.

H. B. Manjunatha

Proteomics and Genomics Laboratory, Department of Studies in Sericulture Science, University of Mysore, Mysore 570006, Karnataka, India. corresponding author

ABSTRACT *Protein extracted from different instars male and female larvae of Bombyx mori exposed to heat shock at 40°C was subjected for identification and in-silico analysis. Heat shock induced protein bands with molecular mass of 25 kDa excised from the 1-DE gel of male and female larvae was analyzed by MALDI-TOF-MS. The resultant peptide mass fingerprints denoted it as small heat shock protein 26 kDa and found localized on the chromosome 27 of B. mori. The multiple sequence alignment and phylogenetic tree clearly indicates that BmHSP26 is species-specific. The Epiloc program while determine the location of BmHSP26 as cytoplasmic, the physico-chemical properties of BmHSP26 with 80.81 aliphatic index and disulphide bridge between Cys78 and Cys85 indicate its thermo-stability in association with two major HSPs - Hsp90 and Hsc70-4. Accordingly, we suppose that BmHSP26 along with other HSPs help the silkworm larvae acquire tolerance to overcome thermal stress.*

1. Introduction

Proteomics is relatively new but fast growing discipline with in life science research to understand the biological process and functions through large scale study of proteins encoded by the genome, because proteins are the ultimate operating molecules producing the physiological effect in an organism. For this reason, protein identification is a key and essential step in the field of proteomics. Without identifying the proteins that are critically involved in the system, it is impossible to probe into the biological explanation for the pattern and/or hypothesis underlying in the system of interest. However, protein identification may often be overlooked or taken for granted while the initial key steps in elucidating the biological process is depends on how accurately, quickly and completely full complement of proteins in a cell/organism be identified. Hence, protein identification using mass spectrometry (MS) is emerged as an indispensable computational tool.

Mass spectrometry and electrophoresis are powerful methods that can provide micro scale analysis of peptides and proteins. The matrix assisted laser desorption-ionization (MALDI) provides rapid, sensitive and accurate molecular masses of peptides and proteins that are resolved through gel electrophoresis, which is a widely used approach for separation of complex mixture of proteins. In addition, peptide MS is also a powerful tool to discover isoforms and identifying protein-coding genes and post-translational modifications that would facilitate annotation as one of the key goals of genome sequencing projects. Now, there are number of bioinformatics tools are available for the analysis of MALDI-TOF-MS data (Palagi et al., 2006). Hence, advances in proteomics tools and techniques would facilitate identification and analysis proteins expressed in the silkworm, *Bombyx mori* L.

B. mori has been exploited extensively for production of silk over 5000 years through domestication due to its great economic importance and has been considered as a model for genetics and molecular biology research. As a consequence, the silkworm genome sequence generated has become fragmented due to shallow coverage which is a limiting factor in identification and annotation of genes effectively (Yamamoto *et al.* 2008). This gap opened ample scope for integration of genomics and proteomics not only to identify the genes encoding for proteins but also assign the function to it. However, the progress made in proteomics unlike genomics has slow phase in *B. mori*, because, it is highly dynamic and complex with respect to the environmental changes that the larvae experience during rearing in the laboratory and field conditions. The proteome differentially expressed during different developmental stages offers systematic approach for comparative analysis not only in different strains/breeds but also different stages and sex of *B. mori.*

2. Materials and methods

2.1 Experimental material

The eggs of silkworm *B. mori* strains - *pere* and NB_4D_2 were incubated under optimum temperature of 25±1°C and 75±5% relative humidity until hatching and larvae were reared on mulberry leaves following standard procedure (Jolly, 1987).

2.2 Induction of heat shock

Different instars larvae of *pere* (male and female) and NB_4D_2 (undifferentiated in early stages and male and female in V instar) were exposed separately to heat shock (HS) temperature of 40°C with 75% ± 5% relative humidity in the water bath for 2 h. The heat induced larvae were transferred to room temperature for recovery and after 2 h they were macerated for extraction of total protein for analysis.

2.3 Protein isolation and Sodium dodecyl sulphate - poly acrylamide gel electrophoresis (SDS- PAGE)

The whole larval protein extraction and SDS-PAGE was performed as described by Weber and Osborn (1969) with necessary modifications. Briefly, 2 h after recovery period, the HS-induced and control whole larvae were homogenized separately in 50mM Tris-HCl (pH 6.8) containing DTT and PMSF and centrifuged at 6700×g (4000 rpm). The clear protein sample was subjected for SDS-PAGE (12%). The resultant gel was stained with Coomassie brilliant blue R-250 (Vasudha et al., 2006).

2.4 In-gel tryptic digestion and MALDI-TOF-MS analysis

In-gel trypsin (Sigma) digest was performed as described by Shevchenko et al., (1996) with some modifications. A protein band excised from the gel was placed in1.5 mL eppendorf tube and dried in vacuum centrifugation. The dried gel pieces were re-hydrated in enzyme solution (with 5ng/mL of trypsin in 25mM NH_4HCO_3). The resulted peptides were extracted twice with 200 mL 50% ACN containing 5% TFA (20min for each extraction was followed). A protein-free gel piece was treated as above and used as control to identify autolysis products derived from trypsin. Peptide masses were analyzed by MALDI-TOF-MS using Ultraflex TOF/TOF (Bruker Daltonics, Germany). Saturated matrix solution was prepared in 50% ACN with 0.1% TFA. After centrifugation, the supernatant was applied for spotting. A pulsed N2 laser-337nm of 50 Hz was used for taking an average of 100–150 laser shots. The resulted spectrum was analyzed using Bruker Daltonics Flex Analysis Software.

2.5 Protein identification

Swiss-Prot and NCBInr databases were used for protein identifica-

Volume - 7 | Issue - 1 | January - 2017 | ISSN - 2249-555X | IF : 3.919 | IC Value : 79.96

tion through the search engine Matrix Science – Mascot (http://www.matrixscience.com). The parameters used in all searches were - the maximum number of missed cleavages allowed was 1; taxonomy used was *Drosophila* followed by all entries (as *Bombyx* is not listed in); enzyme – trypsin; the fixed and variable modifications considered in the search were carbamidomethylation of cysteine and oxidation of methionine respectively (Aitken et al., 2005).

2.6 Multiple sequence alignment and functional analysis of HSPs

The resultant amino acid sequences of all the peptides derived from Mascot search were computed for multiple sequence alignment employing clustal omega (http://www.ebi.ac.uk) for comparative analysis among related HSPs in *Bombyx* and between *Bombyx* and *Drosophila* and also with other organisms.

Protein-Protein interaction analysis was performed employing STRING software (Version10) (http://string-db.org). The sub cellular location of the protein was detected using Epiloc bioinformatic tool (http://epiloc.cs.queensu.ca/).

2.7 *hsp* gene annotation

The spectral signature of mass spectrometry was used for a search of related cds in the silkworm database for its annotation using SilkDB (http://silkworm.swu.edu.cn/silkdb/doc) and KAIKObase (http://sgp.dna.affrc.go.jp/KAIKObase/).

3. Results and discussion

A wide variety of living organisms tend to express sHSPs in various tissues and organs and the number of sHSPs is variable between species. The advances in 1-DE, 2-DE, MS, etc., coupled with bioinformatics tools facilitated not only to identify the protein of interest also determine its functions, reactions, interactions and location of the gene products in tissues and/or cells of living organisms. Hence, we have made an attempt to identify proteins that over expressed in different instars male and female larvae due to HS and uncover sex-specific variation if any and their interaction with other proteins.

3.1 Expression of HSPs in the male and female larvae of *Bombyx mori*

Over expression of 26 kDa HSP was obvious in I, II, III, IV and V instars larvae of *pere* and NB_4D_2 silkworm strains heat shocked at 40°C through SDS-PAGE gel (Figure 1). Since, it forms the first report no information are available on this protein but such strategy was adapted for identification of few HSPs in *B. mori* (Vasudha et al 2006, Manjunatha et al 2010).

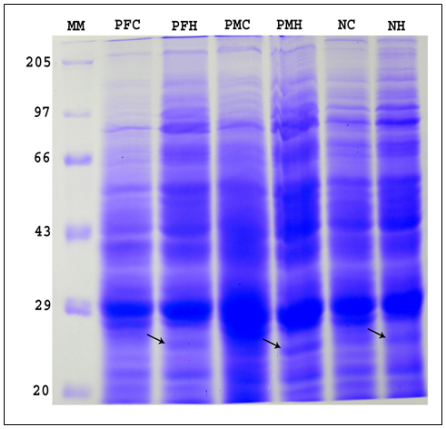

Figure 1: Over expression of 26 kDa small heat shock protein in the II instar male and female larvae of *pere* and undifferentiated stage of NB_4D_2.

(MM-Molecular marker, PFC- *pere* female control, PFH- *pere* female heat shock, PMC- *pere* male control, PFH- *pere* female heat shock, NC-NB_4D_2 control, NH-NB_4D_2- heat shock).

3.2 Identification of HSPs by MALDI-TOF-MS against *Drosophila* database

26 kDa protein bands of male and female larvae of *pere* and NB_4D_2 (undifferentiated) excised from 1-DE gel were subjected to MALDI-TOF-MS analysis to uncover sex-specific variations if any either at protein or peptides. Because, MS can measure other types of information that are invaluable for annotation of genes including post-translational modifications (Jensen, 2006) while some isoforms of biological significance differ by presence or absence of single amino acid (Tadokoro et al., 2005). Thus, resultant peptide mass fingerprints (PMF, Figure 2) were analyzed using a search engine MASCOT search and top hits of the result revealed it as 26 kDa HSP (Table 1) with a highest search score of 27, 238 amino acids and 39% protein sequence coverage to *Drosophila melanogaster* heat shock protein 67B1. Interestingly, the PMF of male and female larvae of *pere* and undifferentiated larvae of NB4D2 did not showed any significant variation because all the hit results were the same. Thus, in the present investigation we have identified over expression of a 26 kDa sHSP in male and female larvae of all the instars for the first time, while most of the proteome research were confined to tissue specific, which includes silk gland (Zhang et. al., 2006), colleterial gland (Jin et al., 2004), hemolymph (Li et. al., 2006), midgut (Kajiwara et. al., 2005) and fat body (Xu et. al., 2000).

Table 1: Peptide mass finger prints of the protein band excised from *Bombyx mori match* to *Drosophila melanogaster* heat shock protein 67B1 (HSP26).

Start	End	Observed	Mr (expt)	Mr (calc)	Delta	Missed cleavage sites	Peptides
20.00	47.00	2994.83	2993.82	2994.46	-0.64	0.00	R.SLAMDIDDSAGFGLYPLEATSQLPQLSR.G
48.00	74.00	2804.49	2803.48	2802.37	1.11	1.00	R.GVGAWECNDVGAHQGSVGGHRSIAIIR.T
69.00	82.00	1651.04	1650.03	1649.96	0.07	1.00	R.SIAIIRTIVWPEPR.L
75.00	82.00	997.13	996.12	996.54	-0.42	0.00	R.TIVWPEPR.L
160.00	168.00	970.14	969.13	968.47	0.66	0.00	K.EDGHGVISR.H
216.00	227.00	1427.69	1426.68	1426.78	-0.09	0.00	R.IVDIQQISQQQK.D
216.00	229.00	1671.97	1670.96	1669.90	1.06	1.00	R.IVDIQQISQQQKDK.D
230.00	236.00	868.98	867.97	868.43	-0.46	1.00	K.DAHRQSR.Q

Identity: Small Heat shock 26 kDa	Gene Name: HSP67B1 protein 67B1
Species: *Drosophila melanogaster*	Match to: HS6B_DROME
Score: 27	Nominal Mass (Mr): 26689
Calculated pI: 8.91	Protein sequence coverage: 39

Volume - 7 | Issue - 1 | January - 2017 | ISSN - 2249-555X | IF : 3.919 | IC Value : 79.96

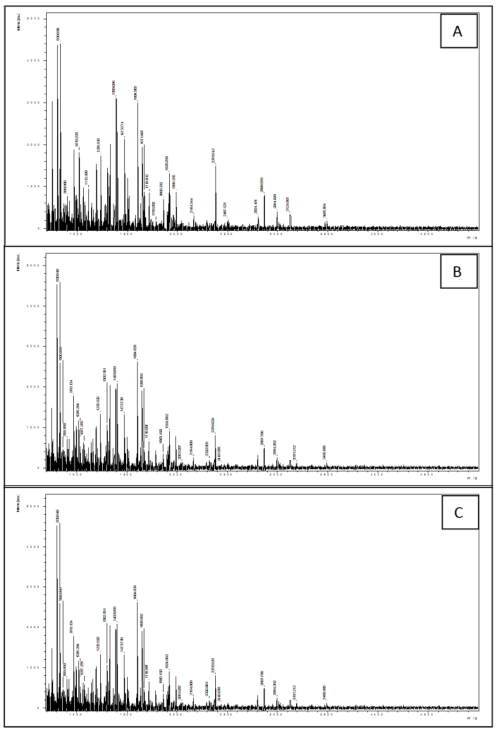

Figure 2: MALDI-TOF mass spectrum of the tryptic digest of heat shock-induced protein bands excised from 1-DE gel. A. *pere* female, B. *pere* male and C. NB_4D_2.

3.3 BmHSP26 multiple sequence and phylogenetic analysis

Matched amino acid sequence for BmHSP26 based on m/z values of mascot search was retrieved in the FASTA format and subjected for sequence similarity search at NCBI BLAST. Consequently, similarity search was performed with other organisms that includes *D. melanogaster, D. buzzatii, Saccharomyces cerevisiae* and *Apostichopus japonicas* and their amino acid sequence was retrieved for multiple sequence alignment. Multiple sequence analysis revealed the evolutionary relationship with 03 residues that completely conserved column (*), 21 residues very conservative substitutions (:) and 10 residues mostly conservative substitutions (.) (figure not shown). Accordingly, phylogenetic tree constructed for matched sequences revealed a separate branch which equally shared with other two clusters *Drosophila sp. and S. cerevisiae and A. japonicas* (Figure 3). This clearly indicates that BmHSP26 could be *Bombyx* specific protein, which offers further investigation to determine its biological significance.

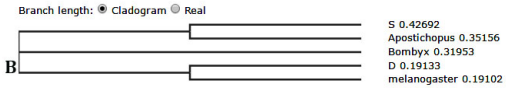

Figure 3: Phylogenetic tree of *Bombyx mori* HSP26 with the matched sequences of other organisms.(S-*Saccharomyces cerevisiae*, D- *Drosophila buzzatii*)

3.4 *hsp* gene annotation in silkworm

After the identification of HSP26 expressed in all the instar silkworm larvae of *pere* and NB_4D_2 against *Drosophila* genome database, the same spectral signature was used for search against *Bombyx* genome database in NCBInr. We have extracted amino acid sequence from resultant NCBI-BLAST search and subjected it for multiple sequence alignment. Based on these inferences and confirmation, the amino acid sequence of *Bombyx mori* was subjected for a search against the coding sequence of silkworm genome in silkdb (http://silkworm.swu.edu.cn/silkdb/). Consequently, silkmap results revealed localization of BmHSP26 on the chromosome 27 with the best hit for both the sequences but their loci are still undetermined (Figure 4).

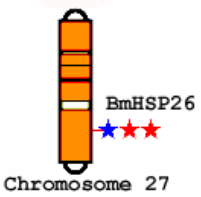

Figure 4: BmHSP26 annotated on the Chromosome 27 of Bombyx mori genome.

3.5 Structural and functional analyses of HSP26 using bioinformatics tools

BmHSP26 sequence was characterized and accordingly the secondary structure predicted posses 31.51% alpha helix, 20.17% extended strands, 11.34% β turn and 36.97% random coil (Figure not shown). Based on this features 3-D structure was determined for BmHSP26 using Phyre2 (Figure 5), which shall be used for molecular docking studies to identify binding sites.

Further, interaction of BmHSP26 with other proteins was predicted using STRING database (http://string-db.org/cgi/network.pl?taskId=bag0omI0NWrc) that explicit interaction of BmHSP26 with two major HSPs - HSP90 (716 aa) and Hsc70-4 (649 aa) besides Uncharacterized proteins (BGIBMGA008273-TA, 348 aa, BGIBMGA008272-TA, 469 aa and BGIBMGA001750-TA, 437 aa), EcRB1, HSFb (Heat shock factor-c-536 aa) Br-c (Broad-complex A-NZ4 isoform, Broad-complex B-NZ4 isoform, 288 aa), Muscular protein 20 (LOC733016, 184 aa) and Bmo.5617 (Uncharacterized protein, 210 aa) as *Bombyx* mori functional partners (Figure 6).

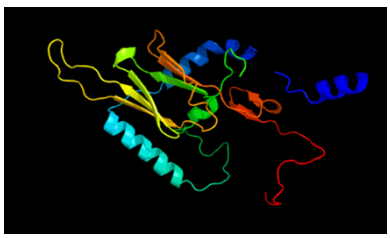

Figure 5: 3D Structure of BmHSP26

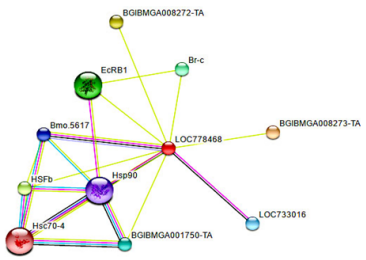

Figure 6: BmHSP26 interaction with other proteins

The physico-chemical properties of the 26 kDa protein sequence such as theoretical isoelectric point (*pI*), molecular weight, and total number of positive and negative residues, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY) were computed by the Expasy's ProtParam server. Accordingly, the *pI* of the protein was 7.80 and it reveals that the protein is basic in nature. The instability index, which provides an estimate of the stability of protein, was 43.62, which is more than 40 indicating that membrane protein was unstable protein. The grand average of hydropathicity (GRAVY) value of protein was 0.672 that indicate less interaction with water. The aliphatic index was 80.81, which indicates the protein may be stable for a wide range of temperature.

Volume - 7 | Issue - 1 | January - 2017 | ISSN - 2249-555X | IF : 3.919 | IC Value : 79.96

Predicting the sub cellular location of proteins is an active research area, as protein's location within the cell provides meaningful clues about its function, Epiloc program was employed to find out the location of BmHSP26. As a result, the output revealed the location of BmHSP26 as cytoplasmic with the score of 0.61 for the first time. Disulphide bridge between Cys78 and Cys85 was also predicted which is invariably used to determine the thermo-stability of protein and the functional domain of this protein was identified as Rhodanase using Scanprosite.

4. Conclusion

Taken together, we have clearly demonstrated that integration of proteomic assays - 1-DE coupled with MALDI-TOF-MS or MS/MS through bioinformatics tools and techniques – facilitated identification of a protein over expressed in different instars male and female larvae of *B. mori* as BmHSP26 first the time. Notably, the physico-chemical properties of BmHSP26 explicit aliphatic index of 80.81 that indicates stability of protein for a wide range of temperature and disulphide bridge between Cys78 and Cys85 which indicate thermo-stability of a protein. Because of this salient feature of BmHSP26, we predict that this protein might actively participate in enhancing the tolerance level in both the sex and strains of *B. mori* against harsh climatic conditions.

Acknowledgement

Ms. Punyavathi is thankful to Department of Science and Technology, New Delhi for award of DST-INSPIRE fellowship (IF-120339).

5. References

1. Aitken, A., 2005. Protein identification by peptide mass finger printing. In: The proteomics protocols hand book, John M. Walker (Ed.). Humana Press, Totowa, New Jersey. 319-324.
2. Jensen, O.N., 2006. Interpreting the protein language using proteomics. Nature Reviews Molecular Cell Biology. 7, 391–403.
3. Jin, Y.X., Chen, Y.Y., Xu, M.K., Jiang, Y.H., 2004. Studies on middle silkgland proteins of cocoon colout sex limited silkworm (*Bombyx mori*) using two-dimensional polyaccralamide gel electrophoresis. Journal of Biosciences. 29, 45-49.
4. Jolly, M.S., 1987. Appropriate sericulture techniques. 63-106.
5. Kregel, K.C., 2002. Heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. Journal of Applied Physiology. 92, 2177-2186.
6. Kajiwara, H., Ito, Y., Imamaki, A., Nakamura, M. et al., 2005. Protein profile of silkworm midgut of fifth-instar day-3 larvae. Journal of Electrophoresis. 49, 61–69.
7. Manjunatha, H.B., Rajesh, R.K., Aparna, H.S., 2010. Silkworm thermal biology: A review of heat shock response, heat shock proteins and heat acclimation in the domesticated silkworm, *Bombyx mori*. Journal of Insect Science. 10,189.
8. Palagi, P.M., Hernandez, P., Walther, D., 2006. Proteome informatics 1: bioinformatics tools for processing experimental data. Proteomics. 6, 5435–5444.
9. Shevchenko, A., Wilm, M., Vorm, O., Mamm, M., 1996. Mass spectrometric sequencing of proteins from silver stained polyaccralamide gels. Analytical Chemistry 68, 850-858.
10. Tadokoro, K., Yamazaki-Inoue, M., Tachibana, M., 2005. Frequent occurrence of protein isoforms with or without single amino acid residue by subtle alternative splicing: The case of gln in drpla affects subcellular localization of the products. Journal Human Genetics 50, 382–394.
11. Vasudha, B.C., Aparna, H.S., Manjunatha, H.B., 2006. Impact of heat shock on heat shock proteins expression, biological and commercial traits of *Bombyx mori*. Insect Science. 13, 243-250.
12. Weber, K., Osborn, M., 1969. The reliability of molecular weight determination by Dodecyl Sulphate Polyacrylamide Gel Electrophoresis. Journal Biological Chemistry. 244, 4406-4412.
13. Xu, Y., Xu, J., Hideki, K., 2000. Studies on changes of protein synthesis in fat body of silkworm (*Bombyx mori*). Acta Sericologica Sinica. 26, 239–243.
14. Yamamoto, K., Nohata, J., Okuda, et al., 2008. A BAC-based integrated linkage map of the silkworm *Bombyx mori*. Genome biology. (doi:10.1186/gb-2008-9-1-r21)
15. Zhang, P., Aso, Y., Yamamoto, K., Banno, Y. et al., 2006. Proteome analysis of silk gland proteins from the silkworm, *Bombyx mori*. Proteomics. 68, 2586–2599.