

Structural prediction and analysis of VIH-related peptides from selected crustacean species

Ganji Purna Chandra Nagaraju^{1*}, Nunna Siva Kumari^{2S}, Ganji Lakshmi Vara Prasad^{3S}, Balney Rajitha^{1S}, Madan Meenu^{4S}, Manam Sreenivasa Rao^{5S}, Bannoth Reddya Naik^{3S}

¹Department of Biology, ⁴Department of Molecular Biology and Microbiology, University of Central Florida, Orlando, FL-32816, USA; ²Department of Bioinformatics and ³Department of Zoology, Osmania University, Hyderabad, Andhra Pradesh, 500 007, India; ⁵Department of Chemistry, Chundi Ranganayakulu Post Graduate College, Chilakaluripet, Andhra Pradesh, 522 616, India; ^SEqual Contribution; Ganji Purna Chandra Nagaraju - EMail: gpcnraju@rediffmail.com; Phone: +1 - 321 - 305- 8721; Fax: +1 - 407 - 823 - 0715;

* Corresponding author

Received March 24, 2009; Revised April 21, 2009; Accepted May 27, 2009; Published August 17, 2009

Abstract:

The tentative elucidation of the 3D-structure of vitellogenesis inhibiting hormone (VIH) peptides is conversely underprivileged by difficulties in gaining enough peptide or protein, diffracting crystals, and numerous extra technical aspects. As a result, no structural information is available for VIH peptide sequences registered in the Genbank. In this situation, it is not surprising that predictive methods have achieved great interest. Here, in this study the molt-inhibiting hormone (MIH) of the kuruma prawn (*Marsupenaeus japonicus*) is used, to predict the structure of four VIH-related peptides in the crustacean species. The high similarity of the 3D-structures and the calculated physicochemical characteristics of these peptides suggest a common fold for the entire family.

Keywords: Growth inhibiting hormone(GIH); Vitellogenin protein(VTG); Vitellogenesis inhibiting hormone (VIH); molt-inhibiting hormone (MIH); 3D-structure; homology model; shrimp or prawn.

Background:

Reproduction in female crustaceans is distinguished by production and accumulation of yolk proteins in the eggs. These yolk proteins are essential for giving nutrition to the growing embryos. Vitellogenesis refers to the progression of molecular biosynthesis of yolk proteins, their way of transport and storage in the ovary. Vitellogenesis is synchronized by the complex endocrine system in the female crustaceans. Several experiments have established that removal of eyestalk in crustaceans speeds up the timing of gonadal maturation and stimulates growth [1, 2]. It was understood that the eyestalk of neuroendocrine system, the X-organ sinus gland complex secreted inhibiting hormone (called VIH) embattled against the gonad. The inhibitory neuropeptide (VIH) control was confirmed by evidence that (unilateral and bilateral) eyestalk ablation arouses ovarian maturation [1]. The VIH inhibited the synthesis of 97 kDa yolk proteins but not general protein synthesis, implying that the hormone exclusively controlled the timing of vitellogenesis in crustaceans [2]. Consequently, the function of VIH is to hold down yolk production in female crustaceans until rich organic reserves are present in ovary or hepatopancreas to feed the growing egg. In spite of the isolation and depiction of VIH, the clear-cut mechanism of its action on the vitellogenesis is still mysterious.

The X-organ-sinus gland (XO-SG) complex of eyestalk in crustaceans is considered as the key rich source of the CHH family peptide production. VIH is a member of CHH family peptides. Molecular analysis of VIHs isolated from various crustacean species (lobsters, crabs, shrimps, and prawns) shows that they are peptides of 78-83 residues. VIH peptides have considerable degree of sequence similarity with MIH, including the conservation of 6 cysteine residues at the same relative positions. So VIH is considered as a CHH-II family group. RP-HPLC analysis revealed that this neuropeptide is dimorphic in *Homarus americanus*; both isoforms have the same sequence, mass and isoelectric point but varies in hydrophobicity [3]. Ollivaux et al., [4] conducted experiments to analyze VIH proteins. They extracted VIHs from the European lobster *Homarus gammarus* and *H. americanus* by RP-HPLC, and analyzed the peptide by Fourier transform ion cyclotron resonance (FT-ICR) MS. The N-terminal of CHHs and VIHs were determined by ELISA. In both species the VIH had the same retention time and molecular masses, and hence

suggested that the VIH peptide sequence was highly conserved between the two species. The functional significance of such structural isoforms is far from clear. Immunocytochemistry and in situ hybridization studies indicate that there are no remarkable differences in the number of GIH neuroendocrine cells of XO of female and male lobsters, suggesting that this hormone has an endocrine role in male reproduction [5]. Some of the CHH family peptides have inhibitory effects on protein synthesis (including VTG and shrimp ovarian peritrophin, SOP) and VTG mRNA synthesis *in vitro*-incubated ovarian fragments [6-8]. Thus, the synthesis of VTG and SOP seems to be mainly under control by CHH family peptides. However, evidence to support this hypothesis is limited and further studies are necessary.

Homology-forming can offer a knowledgeable speculation about the structure of a protein (the target) from its sequence. Achieving high resolution structures of proteins remains a tricky and low yield procedure, but homology forming based on existing structures has gained with significant success. The NMR structure of MIH provides insight into how homologous members of CHH family (VIHs) are modeled. Homology forming results include precious information about how well each region of the model fits its templates, and thus how much confidence is justified. Homology modeling methods also have numerous rewards; they are quick, reproducible, and clearly accessible. These methods have been shown to be satisfactorily correct [9] for optimizing experiments. The idea of this study was to examine amino acid sequence, composition, properties of VIH variations. Gene organizations of the VIH sequences were also examined and discussed. By using modeler software, we design the structure of 4 VIH peptides are predicted.

Methodology:

Homology modeling of VIH:

The VIH like sequences for selected crustaceans used in this work were published by Ollivaux et al.[4], Yang and Rao[10] and Treerattrakool et al. [11]. The reference used for model creation was the solution structure of the molt-inhibiting hormone from the kuruma prawn, *Marsupenaeus japonicus* [*Marja MIH*; Protein Data Bank Code: 1J0T; [12]. The VIH like sequences *Homarus gammarus* (*Homga_VIH*; accession number: ABA42181; VIH, *Macrobrachium*

rosenbergii (*Macro_VIH1*, *Macro_VIH2* accession numbers: AF432346, AF432347), and *Penaeus monodon* (*Penmo_VIH*; accession number: ABG33898) were accessed through GenBank at NCBI and homologous sequences were found by BLAST program (<http://www.ncbi.nlm.nih.gov/ilsprod.lib.neu.edu/BLAST/> basic local alignment search tool) [13] against PDB.

Homology models of VIHs were constructed using the MODELLER software, a program for relative protein structure modeling. The obtained results were expressed as probability density functions (pdfs) for the restrained features. The Pdfs were used to describe intra and inter residue inter atomic distances ($C^{\alpha}-C^{\alpha}$), as well as bond angles (N-O), dihedral angles and other more spatially extensive regions of the protein structure. The 3-D model of a protein is attained by optimizing the resulting model so that it violates the input restraints as least as possible. Modeller uses a combination of bond angles and dihedral angle pdf's to optimize the protein structure models it builds. Modeller's internal OPTIMIZE routine can be used for PDF based structure optimization. The optimization procedure is a variable target function method that applies a conjugate gradients algorithm to positions of all non-hydrogen atoms [14]. The coordinates for the structurally conserved regions (SCRs) of VIHs were allocated from the template using Homstrad sequence alignment (<http://www-cryst.bioc.cam.ac.uk/homstrad/>) based on the structure-based alignments.

From the best alignment of reference structures to target sequences, 3-D models containing all non-hydrogen atoms were obtained using MODELLER software [14]. Minimizations of the models were obtained automatically using the method implemented in the program. The spatial restraints and CHARMM energy terms enforcing proper stereochemistry [15] were combined into one objective function. The model corresponding to the lowest value of the objective function was used for further analysis. The cycle of realignment, modeling and structural validation was repeated until no further improvement on the structure was observed.

Validation of VIH like peptides:

The final structure was analyzed by Ramachandran's map using PROCHECK (<http://biotech.ebi.ac.uk/>, a program to check the stereochemical quality of protein structures) [16]. Reliability of the VIH-related structures was subjected to VERIFY3D [17], available from NIH MBI Laboratory Servers (http://nihserver.mbi.ucla.edu/Verify_3D/). The assessment with VERIFY3D derives a "3D-1D" profile based on the local environment of each residue, described by the statistical preferences for: the area of the residue that is hidden, the fraction of side-chain area that is covered by polar atoms (oxygen and nitrogen), and the local secondary structure. It is also substantiated the reliability of the three dimensional structure. A structure-based sequence alignment of the template structure was obtained using the program JOY [12]. The alignment is shown in **Figure 1**.

Analysis of amino acid residues in VIH like peptides:

Analysis of amino acid content, molecular weight Isoelectric point and phosphorylation sites of peptides was performed using software provided by SDSC Biologyworkbench3.2 (website <http://workbench.sdsc.edu/>). Predictions of HTH Motifs in Protein/peptide chains were determined according to Dodd and Egan [18].

Discussion:

Homology modeling of VIH:

The results of BLAST search with the PDB revealed that three reference proteins, including molt inhibiting hormone (*Marja_MIH*) from kuruma prawn [12], Crystal structure of *Toxoplasma gondii* micronemal protein (2JH1), and Crystal structure of *Toxoplasma gondii* micronemal protein 1 bound to 3' sialyl-N-acetylactoseamine

(2JHD), had a considerable sequence identity and a similarity of 51%, 48%, 51% and 70% respectively. Structurally conserved regions (SCRs) for the model and the template were determined by superimposition of the two structures and the multiple sequence alignment from CLUSTALW (**Figure 1**). Coordinates from the reference protein (*Marja_MIH*) to the SCRs, Structural variable regions (SVRs), N-termini and C-termini were assigned to the target sequence based on the satisfaction of spatial restraints. An initial model was thus generated using the above procedure. In this study, VIH signal peptide *Homga_VIH*; residues 1-33, *Macro_VIH1*; residues 1-40, *Macro_VIH2*; residues 1-33 and *Penmo_VIH* residues 1-17 region were removed from the model because no homologous region occurred in template (*Marja_MIH*). All models contain 79-80 residues corresponding to the A chain of the template MIH (Figure 1). HPLC analysis revealed that this VIH like peptides dimorphic in several crustaceans [19-20] and the significance of such structural dimorphism is far from clear. Nevertheless, they have shown that the difference was at the level of the chirality of one amino acid residue. The present VIH data is also in agreement with the presence of different CHH and MIH isoforms of peptide demonstrated in other shrimps such as *P. serratus*, *P. Japonicus* and *Macrobrachium rosenbergii* [19-20].

Validation of VIH like peptides:

Validation of the model was carried out using Ramachandran plot calculations computed with PROCHECK program. PROCHECK compares the orientation of residues in the model to those permitted by Ramachandran plots, chi-chi2 plots, main and side chain torsions, and residue properties. The PROCHECK G factor ranks values above -0.5 as positive candidates for homology models (data not shown). The models generated by MODELLER (9.1 version) [14], based on the HOMSTRAD alignment, had values exceeding -0.5. The G values indicated here predict that the secondary structures of the models are in the high probability regime for chi-chi² values and the models are realistic. The Φ and Ψ distributions of the Ramachandran plots of non-glycine, non-proline residues are observed. Altogether, 100% of the residues were in favored and allowed regions (data not shown). All residues appeared to be reasonable from Figure 2. The overall PROCHECK G-factor [16] and VERIFY-3D environment profile was good (**Figure 2**).

RMS comparisons of the backbone fold:

Root mean square (RMS) deviation of each model was calculated from the PDB structure *Marja_MIH*, based on the carbons (C_{α}), using application SPDBV [22] (Swiss PDB viewer). As shown in **Figure 2**, the average RMS deviations are generally within 1 Å. This further indicates that the homology model is reliable. Thus, the model was built with considerable template structure and most of the parameters were correlated well throughout the study.

Secondary structure prediction:

The amino acid sequences of template (*Marja_MIH*), VIH like peptides (models generated from the MODELLER) are generated using the JOY program (protein sequence structure representation and analysis [17]). The JOY program uses 3D coordinates from PDB files to calculate secondary structural and local environmental features and displays the same in a color-coded format in the sequence alignment. The secondary structures of template, VIH like peptides consists highly conserved 6 cysteine residues (Cys8 and Cys45, Cys25 and Cys41 and Cys28 and Cys54), which showed that the models are realistic as seen in **Figure 1**, and **Table 1**. Finally, all of the VIH peptides and *Marja_MIH* were predicted to contain five α -helix regions (**Figure 1**).

Identification of phosphorylation site:

Prosites patterns illustrate the specificity pattern of a few well known kinases. Nevertheless, Prosites patterns are the most extensively used approach for predicting phosphorylation sites in novel peptides or

proteins. Phosphorylation sites of proteins or peptides play a very important role in signal transduction, being responsible for protein-membrane and protein-protein interactions [23]. Phosphorylation usually results in functional changes of the target protein by changing the enzyme activity. Amino acid sequence analysis predicts putative Protein kinase C (PKC) phosphorylation site (amino acid residues from 07-10), Casein kinase II (CK-2) phosphorylation site (amino acid residues from 02-06), and two N-myristoylation sites (amino acid residues from 10-16 and 39-45) appeared in MIH and VIH mature peptides (Table 2 in supplementary material).

Amino acid composition and Isoelectric point of peptides:

The basic amino acid residues (H, K, and R) in MIH mature peptides were lower than VIH-like mature peptides (Table 3). Interestingly, all VIH like peptides contain histidine residue except MIH. The acidic amino acid residues (D and E) content in VIH peptides is higher than the MIH mature peptides except *Penmo_VIH*. Separation of proteins may be influenced by isoelectric point (pI), electric charge and molecular weight of protein or peptide. Theoretical isoelectric points (pI) values for MIH-like mature peptide and some of VIH-like peptides were different than VIH peptides (*Homga_VIH*, *Macro_VIH* and *Macro_VIH2*); because these mature peptide have different amino acid sequences (relative numbers of anionic and cationic groups).

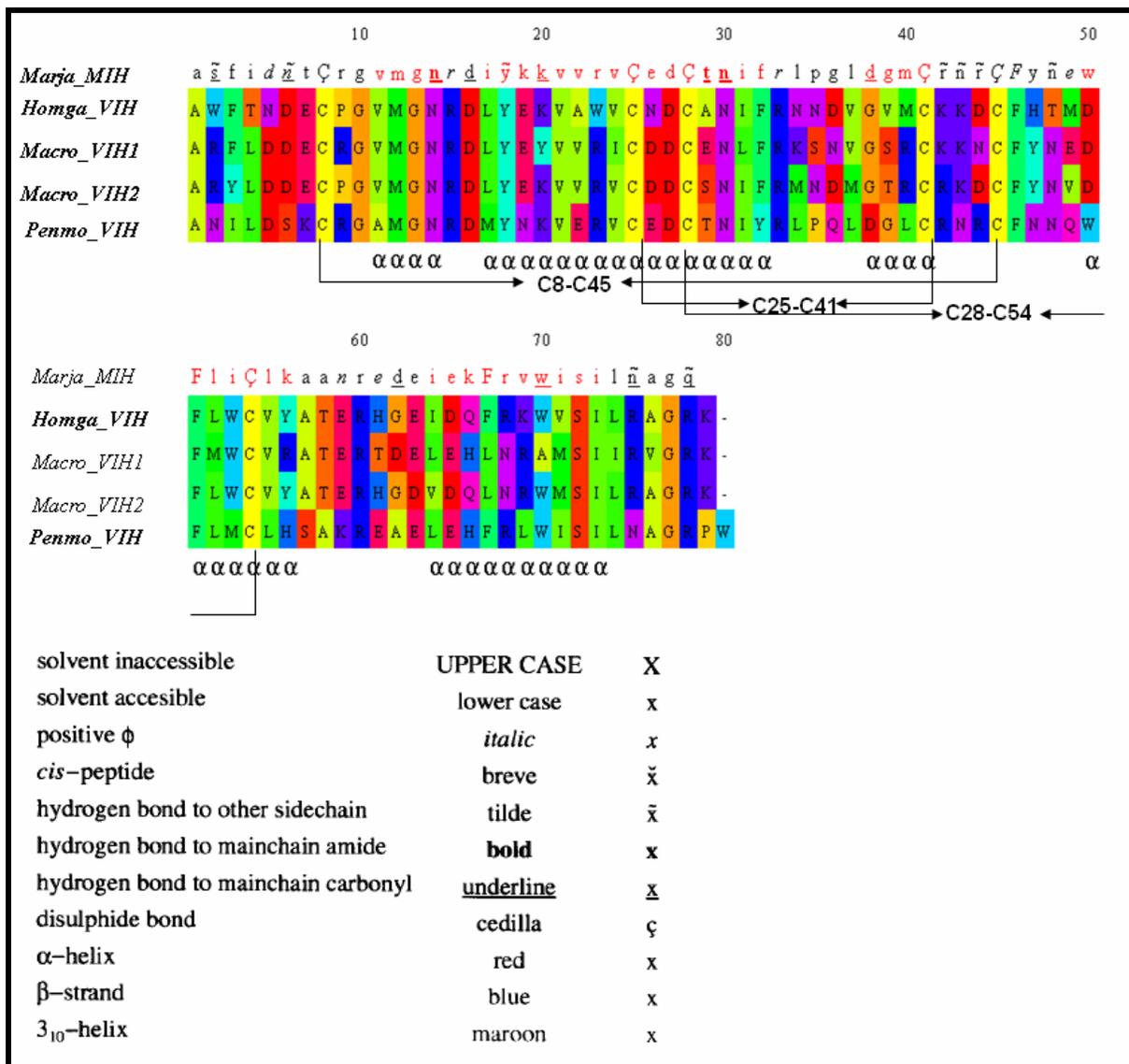


Figure 1: The conformation of VIH peptides and MIH (*Marja_MIH*) were stabilized by a number of conserved hydrophobic interactions in the cluster of ten hydrophobic amino acid residues (Y18, V21, I/V24, I/L31, F/Y32, F46, F51, V/L55, I73 and I/L74) and the three disulfide bonds. In addition four acidic amino acid residues were strongly conserved in both MIH and VIH-like mature peptides at position 16, 27, 63, and 65. Similarly, four basic amino acid residues were strongly conserved in both MIH and VIH-like mature peptides at position 15, 33, 42, and 60. These ten hydrophobic and cysteine residues are conserved in all of the peptides as shown in this figure, suggesting that the VIH like peptides and MIH peptides harbor similar folding.

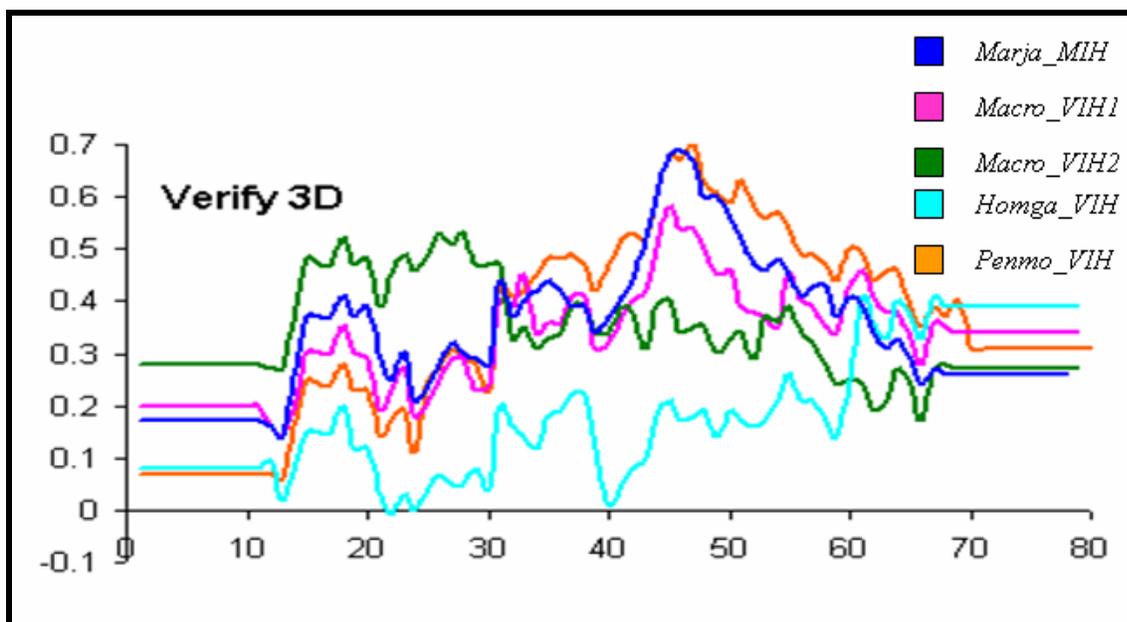


Figure 2: Maximum Verify3D scores for each VIH peptides and template and residues with positive compatibility score are reasonably folded.

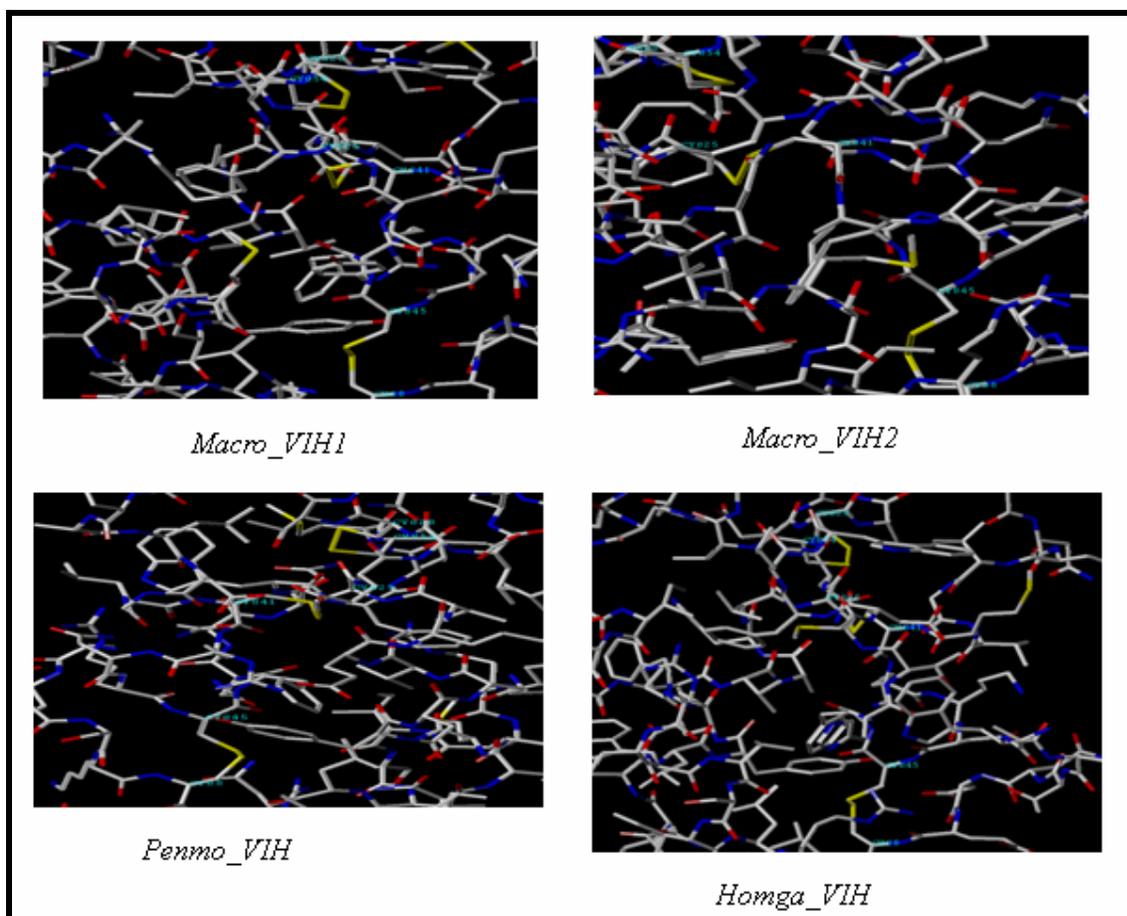


Figure 3: Molecules appear in stick representation in a black window. The default atoms colors are the following. ‘C’ white ‘O’ red ‘N’ blue ‘S’ yellow ‘other’ gray Disulfide Bridges for the residues C8, C45, C25, C41, C28 and C54.

Molecular weight and extinction coefficient of mature peptides:

The molecular weight of MIH mature peptide is 9.1 kDa and all VIH-like peptides have between 9.3 - 9.5 kDa (**Table 3 in supplementary material**). The extinction coefficients of all peptides were determined based on number of tryptophan (W), tyrosine (Y) and cysteine (C) residues. The extinction coefficient values and absorbance at 280nm for MIH mature peptide and some of VIH-like peptides were higher in value than the VIH peptide (*Macro_VIH2*) (see **Table 3 in supplementary material**). The MIH and all VIH contain Proline (P) and Glutamine (Q) residues except in *Macro_VIH2*. The extinction coefficient of a protein is pH dependent, because the conformation of a protein changes due to its surrounding pH and so does absorption and scattering of light through a solution containing that protein. In the current study, MIH have lower number of tryptophan and tyrosine residues. This is the explanation for MIH and VIH (*Macro_VIH2*) showing lower value of extinction coefficient and absorbance at 280 nm.

Conclusion:

Among these eyestalk neuro-peptides, some peptides have similar protein structure and share conserved special characteristics in amino acid sequence and the peptides do form a unique family. Katayama *et al.*, [12] have reported that the r-MIH was stabilized by nine hydrophobic residues in kuruma prawn. In the present investigation, we observed ten hydrophobic and six cysteine residues are conserved in all of the peptides as shown in Figure1, suggesting that the VIH-like peptides and MIH peptides harbor similar folding. Several reports understood that all members of the VIH/MIH gene family derived from a common ancestral gene in the course of gene mutation and duplication. The 3-D structure of kuruma MIH shows above 40% homology, with the VIH like peptides. Nevertheless, consideration of the roles of various members of this VIH family peptides, together with sequence information recently reported, strongly suggests that these VIH peptides may be multifunctional. This suggests a common fold for the entire family. On the other hand, site-directed mutagenesis, antibody recognition information and photoaffinity labeling are helpful experiments to authenticate and improve the proposed models. To know whether the MIH or VIH peptides are phosphorylated at any specific site might be useful to one would have to be investigating *in vitro* using phospho-specific antibodies. In conclusion, we have presented detailed structural analyses of the VIH like neuropeptides. The structural identification of these VIH-like peptides provides the basis for future functional studies of these novel signaling molecules. The comparative study of VIH-like neuropeptides will offer new

insights into the functional consequences of peptide diversity and peptide evolution.

Acknowledgements:

I thank NSF; IBN-0611447 (USA) for the financial support during my investigation.

References:

- [1] GPC Nagaraju, *Aquaculture* **272**: 39 (2007)
- [2] LS Quackenbush, *Amer. Zool.* **41**: 458 (2001)
- [3] D Soyez *et al.*, *Neuropeptides* **20**: 25 (1991) [PMID: 1791922]
- [4] C Ollivaux *et al.*, *Febs. J.* **273**: 2151 (2006) [PMID: 16649992]
- [5] G Rotllant *et al.*, *Cell. Tissue. Res.*, **271**: 507 (1993)
- [6] JC Avarre *et al.*, *Gen. Comp. Endocrinol.* **124**: 257 (2001) [PMID: 11742508]
- [7] M Khayat *et al.*, *Gen. Comp. Endocrinol.* **110**: 307 (1998) [PMID: 9593651]
- [8] N Tsutsui *et al.*, *Gen. Comp. Endocrinol.* **144**: 232 (2005) [PMID: 16055123]
- [9] H Nielsen *et al.*, *Protein Eng.* **12**: 3 (1999) [PMID: 10065704]
- [10] WJ Yang, RR Rao, *Biochem. Biophys. Res. Commun.* **289**: 407 (2001)
- [11] S Treerattrakool *et al.*, *Febs. J.* **275**: 970 (2008) [PMID: 18221491]
- [12] H Katayama *et al.*, *J. Biol. Chem.* **278**: 9620 (2003) [PMID: 12519766]
- [13] SF Altschul *et al.*, *Nucleic Acids Res.* **50**: 3389 (1997) [PMID: 9254694]
- [14] A Sali, TL Blundell, *J. Mol. Biol.* **234**: 779 (1993) [PMID: 8254673]
- [15] BR Brooks *et al.*, *J. Comp. Chem.* **4**: 187 (1983)
- [16] RA Laskowski *et al.*, *J. Appl. Cryst.* **26**: 283 (1993)
- [17] R Lüthy *et al.*, *Nature* **356**: 83 (1992) [PMID: 1538787]
- [18] IB Dodd and JB Egan, *Nucleic Acids Res.* **18**: 5019 (1990) [PMID: 2402433]
- [19] WJ Yang *et al.*, *Peptides* **18**: 479 (1997) [PMID: 9210164]
- [20] SH Chen *et al.*, *Mar. Biotechnol.*, **6**: 83 (2004) [PMID: 15933902]
- [21] A Van Wormhoudt *et al.*, *Comp. Biochem. Physiol.* **79**: 353 (1984)
- [22] N Guex, MC Peitsch, *Electrophoresis* **18**: 2714 (1997) [PMID: 9504803]
- [23] H Taniguchi, *Biophys. Chem.* **82**: 129 (1999)

Edited by P. Kanguane

Citation: Nagaraju *et al.*, *Bioinformation* 4(1): 6-11 (2009)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

Supplementary material

Table 1: Characteristics of the disulfide bonds for the models of VIH like peptides

Disulfide bonds	Marja MIH	Homga VIH	Macro VIH1	Macro VIH2	Penmo VIH
C08 and C45	2.26Å	2.022 Å	2.026Å	2.017Å	2.869Å
C25 and C41	2.242Å	2.010Å	2.031Å	2.031Å	2.028Å
C28 and C54	1.986Å	2.031Å	2.027Å	2.030Å	2.019Å

Table 2: Identification of phosphorylation site on template and target sequences

Phosphorylation site name and Access#	Marja_MIH MP	Homga_VIH MP	Macro_VIH1 MP	Macro_VIH2 MP	Penmo_VIH MP
PKC_PHOSPHO	07 to 10	58 to 61	58 to 61	58 to 61	57 to 60
PS00005	AAR	AAR	AAR	AAR	AAR
CK2_PHOSPHO	02 to 06	04 to 08	-----	-----	-----
PS00006	AAR	AAR	40 to 48	-----	-----
TYR_PHOSPHO	-----	-----	AAR	10 to 16	10 to 16
PS00007	10 to 16	10 to 16	10 to 16		39 to 45
MYRISTYL	39 to 45			AAR	AAR
	AAR	AAR	AAR	76 to 79	
PS00008	-----	76 to 79	76 to 79	AAR	
AMIDATION		AAR	AAR		
PS00009					

MP - mature peptides; AAR - amino acid residues; PKC_PHOSPHO_SITE - protein kinase C phosphorylation site; CK2_PHOSPHO_SITE - casein kinase II phosphorylation site; TYR_PHOSPHO_SITE - Tyrosine kinase phosphorylation site; MYRISTYL_SITE - N-myristoylation site; AMIDATION: amidation site

Table 3: Amino acid composition, molecular weight, extension coefficient, isoelectric point of peptides

Name of amino acid	Marja_MIH	Homga_VIH	Macro_VIH1	Macro_VIH2	Penmo_VIH
Alanine A	5	5	3	3	5
Cysteine C	6	6	6	6	6
Aspartic acid D	5	7	10	7	4
Glutamic acid E	5	4	3	7	5
Phenylalanine F	5	5	3	4	3
Glycine G	5	5	5	4	4
Histidine H	0	2	1	1	2
Isoleucine I	7	3	2	3	4
Lysine K	4	5	3	4	3
Leucine L	5	3	5	5	9
Methionine M	2	3	4	3	3
Asparagines N	7	6	5	6	8
Praline P	1	1	1	0	2
Glutamine Q	1	1	1	0	2
Arginine R	8	6	10	11	9
Serine S	2	1	2	3	3
Threonine T	2	3	2	2	1
Valine V	5	7	7	6	2
Trpptophan W	2	4	2	1	3
Tyrosine Y	2	2	4	3	2
No. of Amino acid residues	78	79	79	79	80
Molecular weight (Da)	9141.08	9326.11	9407.21	9462.28	9506.48
Extcoef at 280nm (1/M 1/cm)	14300	25680	16860	9890	19990
Absorbance at 280 nm (Abs0.1%=1g/l)	1.564	2.753	1.792	1.045	2.102
Isoelectric Point	1.524	2.714	1.753	1.007	2.064
	8.012	6.920	6.772	7.631	8.304