Journal of
Proteomics & Bioinformatics

Vaseeharan and Valli J Proteomics Bioinform 2011, 4:1
DOI: 10.4172/jpb.1000166

Research Article Open Access

In silico Homology Modeling of Prophenoloxidase activating factor Serine Proteinase Gene from the Haemocytes of *Fenneropenaeus indicus*

Baskaralingam Vaseeharan* and Sivakama Valli . J

Department of Animal Health and Management, Alagappa University, Karaikudi-630 003, Tamil Nadu, India

Abstract

In invertebrates, immune system consist of encapsulation, phagocytosis, and nodule formation while humoral responses includes clotting, synthesis of antimicrobial peptides, and activation of the Prophenoloxidase (proPO) system. Serine proteinases (SPs) constitute one of the largest families of proteolytic enzymes involved in the activation of Prophenoloxidase. The major feature of serine proteinase is interlinked by three pairs of conserved disulfide bridges. Although the exact function of the clip domain presently remains unclear, there are certain speculations about its function. The present study reports the three-dimensional structures of novel immune related gene serine proteinase predicted by *in silico* homology modelling studies. Physico-chemical characterization interprets properties such as pI, EC, AI, GRAVY and instability index and provides valuable data about this clip domain serine proteinase. Prediction of motifs, patterns, disulfide bridges and secondary structure were performed for functional characterization of the serine proteinases. Three dimensional structures for these proteins were not available as yet at PDB. Therefore, a homology model for this serine proteinase protein was developed. The modelling of the three dimensional structure of the proteins showed that models generated by Modeller9V8 were more acceptable in comparison to that by Swiss Model. The models were validated using protein structure checking tools PROCHECK and WHAT IF. The structures will provide a good foundation for functional analysis of experimentally derived crystal structures. The better results of the in silico modelling study are presented, and may help lead to the discovery of new synthetic immune related peptides or derivatives of serine proteinases that could be useful to understand the mechanism of serine proteinase involvement in the Prophenoloxidase activating system of crustaceans. The crystal structure prediction of the immune related proteins serine proteinase of shrimps will help to explore the other life sciences Pharmacokinetics and toxicology, Drug designing and chemo informatics etc.

Keywords: Serine Proteinase; Homology modeling; *Fenneropenaeus*; Prophenoloxidase; Ramachandran plot

Introduction

Serine proteinases belong to a class of proteolytic enzymes, which are characterized by the presence of a unique serine residue within the active site. The members of this class are further classified into distinct families according to their structure and mechanism of action. The two largest families, which have been well studied, are the (chymo) trypsin and subtilisin. Serine proteinases (SPs) feature a wide array of important physiological functions, including digestion, blood coagulation, fibrinolysis, fertilisation, embryonic development and immunity [1,2]. X-ray crystal structural study indicates that the active center of bovine chymo trypsin is composed of His57, Asp102, and Ser195, which are responsible for the acyl transfer mechanism during catalysis. Known as a catalytic triad, these residues actually form two diads, Ser–His and His–Asp, which operate in concert [3]. Substrate binding clefts near the active site largely determine the substrate specificity of SPs. To execute their extracellular functions, these enzymes are generally synthesised as zymogens and stored in vesicles or granules from which they are released and converted to the active enzyme by a proteolytic cleavage at a particular peptide bond. Through specific protein-protein interactions, several SP zymogens can form a cascade pathway in which one protease activates the zymogen of another to mediate a rapid, local reaction. The blood clotting cascade in human plasma, the prophenoloxidase activating system (proPO system) in arthropods and the clotting cascade in *Limulus* haemolymph are classical examples of such SP system [4-6]. Serine Proteinase is an important enzyme in the needed for the induction of prophenoloxidase, tissue damage and microbial infection, respectively and signal transduction cascade [7]. Serine Proteinase catalyzes the conversion of inactive zymogen Prophenoloxidase into active prophenoloxidase and is catabolically repressed by higher concentrations of serine proteinase [8,9]. However, a 3D structure for serine proteinase is yet not available. Hence, we constructed the model structure for *Fenneropeaneaus indicus* serine proteinase using known structural templates and describe its structural features to understand molecular function.

Crustacean serine proteinase most likely contributes in various protection responses including haemolymph coagulation, melanotic encapsulation, induction of antimicrobial peptide synthesis, and activation of cytokines. However, there is no information about the precise role of shrimp haemocyte SP and their presence has been only assumed as for example SPs are required for proteolytic activation of the proPO-system [10]. Prophenoloxidase induces the innate immune system of the crustaceans. That proPO activation was revealed by the factor Serine Proteinase. Structure prediction by homology modelling (HM) can aid understanding the three dimensional (3D) structure of a given protein. This in turn will help elucidate the mechanisms behind protein function, since function is determined by 3D structure [11,12]. Based on the structure the function of the protein will be predicted. (Structure based *in silico* modelling).

***Corresponding author:** Dr. B.Vaseeharan, Associate Professor & Head Department of Animal Health and Management, Alagappa University, Karaikudi-630 003, Tamil Nadu, India, Tel: + 91 4565 225205, 225206, 225207; Fax: + 91 4565 225202; E-mail: vaseeharanb@gmail.com

Received October 05, 2010; **Accepted** December 28, 2011; **Published** January 05, 2011

Citation: Vaseeharan B, Valli SJ (2011) In silico Homology Modeling of Prophenoloxidase activating factor Serine Proteinase Gene from the Haemocytes of *Fenneropenaeus indicus*. J Proteomics Bioinform 4: 053-057. doi:10.4172/jpb.1000166

Copyright: © 2011 Vaseeharan B, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Homology Modelling has proved very useful in the prediction of the 3D structure and function of proteins. Comparative structure prediction produces an all-atom model of a sequence based on its alignment to one or more related protein structures. Comparative model building includes either sequential or simultaneous modeling of the core of the protein, loops, and side chains. The methodology of homology modeling relies on the observation that the structural conformation of a protein is more highly conserved than its amino acid sequence. Homology modelling can be divided into four steps: template identification, alignment, model building and refinement, and validation with various computational tools. In bioinformatics, computational tools were used to identify the biomolecules like proteins, physiochemical property and the structure prediction of Serine Proteinase. Major demerits of the experimental area is high cost and time taking process and are not amenable to high throughput techniques. *In silico* approaches provide a viable solution to these problems. The amino acid sequence provides most of the information required for determining and characterizing the molecule's function, physical and chemical properties. Computationally based characterization of the features of the proteins found or predicted in completely sequenced proteomes is an important task in the search for knowledge of protein function. In this paper the *in silico* analysis and homology modelling studies on prophenoloxidase activating factor serine proteinase protein from the haemocytes of *Fenneropenaeus indicus* was reported. Three dimensional structures for these proteins were yet not available. Hence to describe its structural features and to understand molecular function, the model structures for these proteins were also constructed.

Materials and Methods

Prophenoloxidase activating factor serine proteinase gene sequence

Indian white shrimp *Fenneropenaeus indicus* ranged from 15.7 to 23.2 g, averaging 18.75 ± 3.60 g (mean ± SD) were obtained from the coastal area of Nagapattinam to Chennai, Tamil Nadu, India. *F.indicus* were stocked and maintained in FRP tanks with flow-through sea water (35 % salinity) at 28°C and fed twice daily with a formulated shrimp diet. The hemolymph was collected from the ventral sinus of an individual shrimp using anti coagulant solution (0.45 M NaCl, 0.1M glucose, 30 mM sodium citrate, 26 mM citric acid, 10 mM EDTA, pH 7.5, Osmolality 780 mOsm kg^{-1}) and immediately centrifuged at 500 g at 4°C for 20 min to separate the hemocytes from the plasma. The resulting haemocyte pellet was used for total RNA isolation. Molecular approaches were used to clone the Propheloxidase activating factor serine proteinase gene in the hemocytes of Indian white shrimp *Feneropenaeus indicus.* The full length sequence of serine proteinase gene was determined by RT-PCR, cloning and sequencing of overlapping PCR and rapid amplification of cDNA ends (RACE) method. The sequences were submitted to NCBI GenBank (Accession No: HM368165). The same sequence was used for the further studies *in silico* homology modeling analysis studies.

Theoretical calculations of physico-chemical characterization

For physico-chemical characterization, theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient [13], instability index Guruprasad et al. [14], aliphatic index [15] and grand average hydropathy (GRAVY) [16] were computed using the Expasy's ProtParam server [17].

Functional characterization

The SOSUI server performed the identification of transmembrane regions in the clip domain serine proteinase. The transmembrane region identified for these two proteins. The unknown function of the protein was predicted using the Prosite [18] by Domain prediction. Table 3 represents the output of Prosite that was recorded in terms of the length of amino residues of protein with specific profiles and patterns.

3D structure generation

Modeller 9V8: Modeller was used for homology or comparative modeling of protein three-dimensional structures. Alignment of a sequence to be modeled is provided with known related structures and modeller automatically calculates a model containing all non-hydrogen atoms. Modeller [19] implements comparative protein structure modeling by satisfaction of spatial restraints, and can perform many additional tasks, including de novo modeling of loops in protein structures, optimization of various models of protein structure with respect to a flexibly defined objective function, multiple alignment of protein sequences and/or structures, clustering, searching of sequence databases, comparison of protein structures, etc. The novel sequence of Serine Proteinase sequenced and it was searched against selection of the related homologues of query sequence in PDB. The homology modelling requires sequences of known 3D structure and the target having above 35% of similarity. In order to confirm the selection pair wise alignment of template and target was performed.

Template identification and sequence alignment

Template identification is an important step. It lays the foundation by identifying appropriate homologues of known protein structure, called template, which are sufficiently similar to the target sequence to be modelled. Template sequence were selected by a simple search submits the target sequence to programs BLASTP [20] search along with default parameters was performed against the Brook Heaven Protein Data Bank (PDB). Based on the high identity, lowest e-value and low gaps the high resolution having sequence was selected as a template. To ensure the high accuracy of the structure, the target and the template sequence can be aligned

Model building and refinement

Although the theory behind building a protein homology model is complicated, using available programs is relatively easy. Several modelling programs are available, using different methods to construct the 3D structures. In segment matching methods, the target is divided into short segments, and alignment is done over segments rather than over the entire protein. Satisfying spatial restraints is the most common method. It uses either distances or optimization techniques to satisfy the spatial restraints. The method is implemented using the popular program, Modeller and [21] which includes the CHARMM energy terms that ensure valid stereochemistry is combined with spatial restraints [22]. The academic version of MODELLER9V8 [19] was used for 3D structure generation based on the information obtained from sequence alignment.

Validation

The best validation combines common sense, biological knowledge and results from analytical tools. Most refinement involves adjusting the alignment. We used PROCHECK [23] to calculate the main-chain torsion angles, i.e. the Ramachandran plot [24], for our predicted structures. Three models were predicted using different templates among those the one that shows the good resolution factor and R-factor was used as a template and evaluated by Procheck (http://www.biotech.ebi.ac.uk/) performing full geometric analysis with a resolution of 1.5 Å. The validation for structure models obtained from the three software tools was performed by using PROCHECK [25]. The models were further checked with WHAT IF [26]. Structural analysis was performed

and figures representations were generated with Swiss PDB Viewer [27]. Ramachandran plot statistics was used to evaluate the best model. The root mean square deviation (RMSd) values were calculated using the modeller by fitting the carbon backbone of the predicted. Finally the all-atom models were subjected to a short run of energy minimization by using AMBER [28] to relieve unfavorable steric interactions and to optimize the stereochemistry. Based on the template prophenoloxidase the crystal structure of serine proteinase was predicted. Fold recognition energy minimization was done by using several tools available in SPDBV. Some of them are energy minimization, loop building, Side chain fixing [29] setting phi/psi angles, fixing the clashing amino acids.

Result and Discussion

Parameters computed using Expasy's ProtParam tool was represented in Table 1. The calculated isoelectric point (pI) will be useful because at pI, solubility is least and mobility in an electro focusing system is zero. Isoelectric point (pI) is the pH at which the surface of protein is covered with charge but net charge of protein is zero. At pI, proteins are stable and compact. The computed pI value of serine proteinase (HM368165) was less than 7 (pI<7) indicates that these proteins were considered as acidic. The computed isoelectric point (pI) will be useful for developing buffer system for purification by isoelectric focusing method. Although Expasy's ProtParam computes the extinction coefficient for 276, 278, 279, 280 and 282 nm wavelengths, 280 nm is favored because proteins absorb light strongly there while other substances commonly in protein solutions do not. Extinction coefficient of serine proteinase homologue at 280 nm is ranging from 45840 from 46715 M−1 cm−1 with respect to the concentration of Cys, Trp and Tyr. The high extinction coefficient of Serine Proteinase indicates presence of high concentration of Cys, Trp and Tyr. The computed extinction coefficients help in the quantitative study of protein–protein and protein–ligand interactions in solution. The instability index provides an estimate of the stability of protein in a test tube. There are certain dipeptides, the occurrence of which is significantly different in the unstable proteins compared with those in the stable ones. This method assigns a weight value of instability. Using these weight values it is possible to compute an instability index (II). A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable [14]. The instability index value for the immune related proteins was found to be ranging from 21.90 to 47.14. The result classified SP as stable protein (Table 1).The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chains (A, V, I and L) is regarded as a positive factor for the increase of thermal stability of globular proteins. Aliphatic index for the serine proteinase and serine proteinase homologue protein sequences ranged from 74.40 to 83.18. The very high aliphatic index of all serine protein sequences indicates that these proteins may be stable for a wide temperature range. The lower thermal stability of serine proteins (HM368165) was indicative of a more flexible structure when compared to other protein. The Grand Average hydropathy (GRAVY) value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence. GRAVY indices of serine proteinases are ranging from - 0.172. This low range of value indicates the possibility of better interaction with water. As disulphide bridges play an important role in determining the thermo stability of these proteins. CYS_REC was used to determine the Cysteine residues and disulphide bonds. Disulfide bonds prediction of *F.indicus* serine proteinase depicts the 18 cysteine residues were occurred among 400 residues of the protein [30].

Serine proteinase Protein	Accession Number	Sequence Length	M.wt	pI	-R	+R	EC	II	AI	GRAVY
SP	HM368165	380	40780.0	5.07	43	31	32430	38.09	83.18	-0.172

Table 1: Parameters computed using Expasy's ProtParam tool. Theoretical isoelectric point (pI), Molecular weight, Total number of positive and negative residues, Extinction coefficient, Instability index, aliphatic index and Grand average hydropathy of Serine Protease was depicted in table 1.

No.	N terminal	transmembrane region	C terminal	type	length
1	7	LWLGTVLLFVYVEGGVGPYCVSL	29	PRIMARY	23

Table 2: Transmembrane regions identified by SOSUI server. This amino acid sequence is of a MEMBRANE PROTEIN which have 1 transmembrane helix.

Name of the Protein	**Accession No**	**Motif Found**	**Profile**	**Position in the protein**	**Description**
Serine protease	HM368165	VTAAHC	TRYPSIN_HIS	**157 - 162**	Transmembrane protein

Table 3: Functional characterization of proteins of SP at Prosite.

Functional characterization of serine proteinase

Functional analysis of these proteins includes prediction of transmembrane region, disulfide bond and identification of important motifs. SOSUI distinguishes between membrane and soluble proteins from amino acid sequences, and predicts the transmembrane helices for the former. The Trans membrane regions and their length were tabulated in Table 2. The server SOSUI [31] classifies serine proteinase as membrane protein and other these proteins as soluble proteins. SOSUI server has identified one transmembrane region in these proteins. The transmembrane regions are rich in hydrophobic amino acids. Table 3 showed the output of Prosite that was recorded in terms of the length of amino residues of protein with specific profiles and patterns. Secondary structure analysis revealed alpha helixes were dominated among secondary structure elements followed by random coils, extended strand and beta turns for all sequences (Table 4).

3D structure generation

According to the identity three templates were chosen .Based on the Ramachandron Plot value and, over all quality factor Among that best one model was selected as suitable template . The stereo chemical quality of the predicted models and accuracy of the protein model was evaluated after the refinement process using Ramachandran Plot Map calculations computed with the PROCHECK program. The assessment of the predicted models generated by modeler was shown in Figure1. The main chain parameters plotted are Ramachandran Plot quality, peptide bond planarity, Bad no bonded interactions, main chain hydrogen bond energy, C-alpha chirality and over-all R factor. In the Ramachandran Plot analysis, the residues were classified according to its regions in the quadrangle. The red regions in the graph indicate the most allowed regions whereas the yellow regions represent allowed regions. Glycine is represented by triangles and other residues are represented by squares. The result revealed that the modeled structure for serine proteinase was 96.8% of residues in allowed region. The distribution of the main chain bond lengths and bond angles were found to be within the limits for these proteins. Such figures assigned by Ramachandran plot represent a good quality of the predicted models. The modeled structures of serine proteinase protein was also validated by other structure verification servers WHAT IF. Standard bond angles of the model were determined by using WHAT IF. Results were shown in Table 5.The analysis revealed RMS Z-scores were almost equal to 1

suggesting high model quality. The predicted structures conformed well to the stereochemistry indicating reasonably good quality.

Protein structure validation

To validate the homology modeled serine proteinase structure, a Ramachandran plot was drawn and the structure was analyzed by PROCHECK, a well known protein structure checking program. It was found that the phi/psi angles of 90% of residues fell in the most favored regions, 6.8% residues fell in the additional allowed regions, and 3.2% fell in generously allowed regions; none of the residues fell in the disallowed conformations (Figure 2). The results of PROCHECK and WHAT IF analysis was shown in the Figure 3. These observations indicate that an increase in the number of bad dihedral angles of the modeled structure had occurred. This may be due to MD simulation causing an unfavorable dihedral angle, allowing the protein to overcome high energy barriers.

Secondary structure	SOPMA	GOR
Alpha helix	19.27%	16.86%
310 helix	0.00%	0.00%
Pi helix	0.00%	0.00%
Beta bridge	0.00%	0.00%
Extended strand	18.35%	22.89%
Beta turn	3.21%	0.00%
Bend region	0.00%	0.00%
Random coil	59.17%	60.26%
Ambiguous states	0.00%	0.00%
Other states	0.00%	0.00%
Sequence length	218	380

Table 4: Secondary structure of serine proteinase by **SOPMA & GOR.**

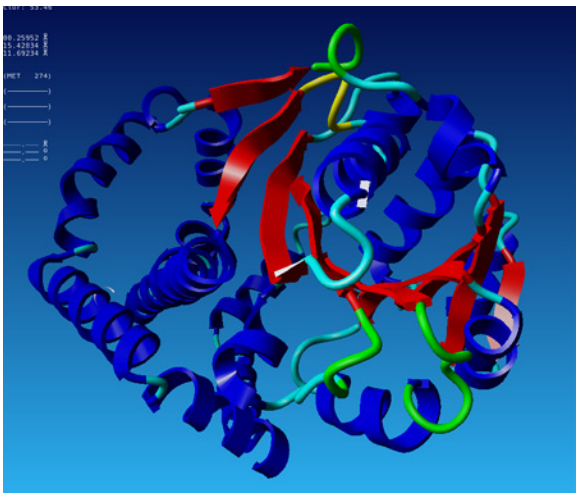

Figure 1: Structure of Serine Proteinase gene from *Fenneropenaeus indicus* modelled by modeller9v8 and the ribbon structure was visualized by YASARA. The protein serine proteinase was modelled using homology modelling based on the template obtained from PDB. The closest homologue with the highest sequence identity is 36%.

Sl. no	Plot statitics	values
1	Residues in most favoured regions [A,B,L]	96.8%
2	Residues in additional allowed regions [a,b,l,p]	3.2%
3	Number of non-glycine and non-proline residues	100.0%
4	Number of end-residues (excl. Gly and Pro)	2
5	Number of glycine residues (shown as triangles)	38
6	Number of proline residues	10

Table 5: RamachandranPlot calculation and comparative analysis of the models from Modeller computed with the **PROCHECK** program.

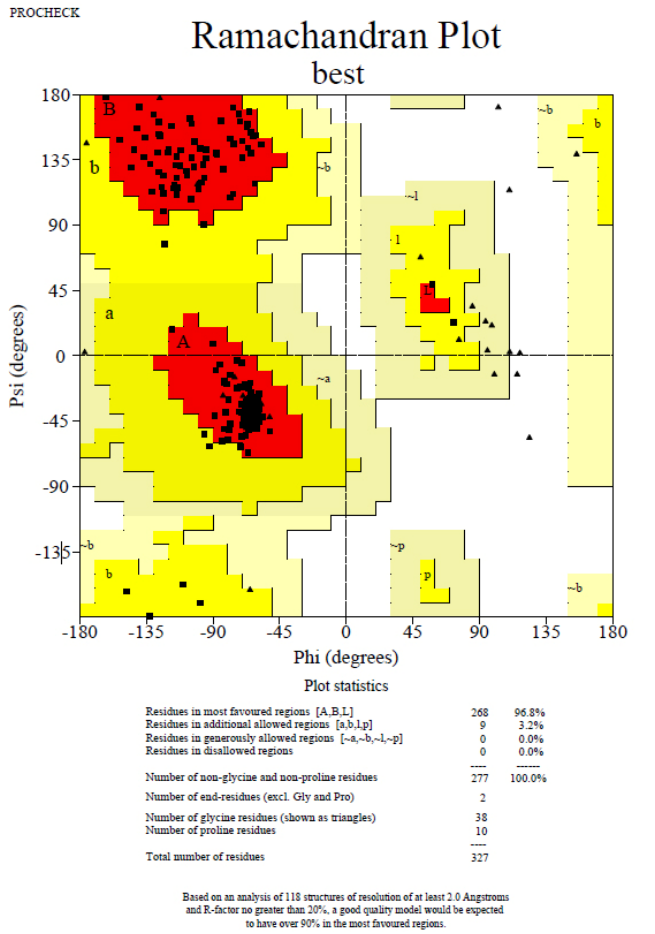

Figure 2: Ramachandran plot validation for the immune related protein Serine Protease from *Fenneropenaeus indicus.*

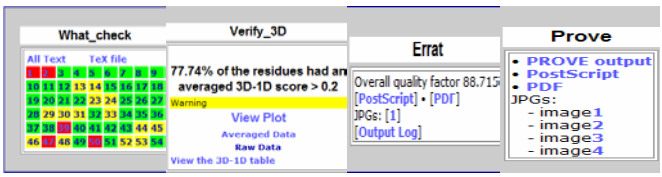

Figure 3: RMS *Z*-score for bond angles of modelled protein structure using **WHAT IF.** The validation values of the Ramachandran plot is given below in the table Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Conclusion

In this study prophenoloxidase activating factor serine proteinase was selected. Physicochemical characterization were performed by computing theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY). Functional analysis of these proteins was performed by SOSUI server. For these proteins disulphide linkages, motifs and profiles were predicted. Secondary structure analysis revealed that random coils dominated among secondary structure elements followed by alpha helix, extended strand and beta turns for all sequences. The modelling of the three dimensional structure of the protein was performed by homology programs, Swiss model and Modeller. The

models were validated using protein structure checking tools Figure 2, Ramachandran's map of serine proteinase. PROCHECK and WHAT IF. These structures will provide a good foundation for functional analysis of experimentally derived crystal structures.The refined model was analyzed by different protein analysis programs including PROCHECK for the evaluation of the Ramachandran plot quality, and WHATIF for the calculation of packing quality. The Verify 3D plot show the 77.74% of residue had an average 3D-1D score. This structure for the corresponding coordinates in PDB format was found to be satisfactory based on the above results. The predicted 3-D model of the serine proteinase protein of Fenneropeaneaus indicus will be very useful in wet laboratory while studying the real structure of the protein. The total energy values of the predicted 3-D model were calculated as 96.8% of Ramachandran plot value in 30 and 40 steepest descent and conjugate gradient, respectively. The protein model has more beta-sheets with only three short a-helices. This transmembrane protein involved in the pattern recognition mechanism and signal transduction pathways.

Acknowledgement

This work was supported by Department of Biotechnology (DBT), New Delhi, India, under the Project grants code: BT/PR11907/AAQ/03/459/2009.

References

1. Rawlings RD, Barrett AJ (1993) Evolutionary families of peptidases. Biochem J 290: 205-18.
2. Krem M, Di Cera E (2002) Evolution of enzyme cascades from embryonic development to blood coagulation. Trends Biochem Sci 27: 67-74.
3. Perona JJ, Craik CS (1995) Structural basis of substrate specificity in the Serine Protease. Protein Sci 4:337-360.
4. Muta T, Iwanaga S (1996) Clotting and immune defense in Limulidae. Prog Mol Subcell Biol 15:154-189.
5. Bennett B, Ratnoff OD (1972) The normal coagulation mechanism. Med Clin North Am 56: 95-104.
6. Soderhall K, Cerenius L (1998) Role of the Prophenoloxidase-activating system in invertebrate immunity. Curr Opin Immunol 10:23-28.
7. O'Brien D, McVey J (1993) Blood coagulation, inflammation, and defense. In: Sim E (ed), The Natural Immune Systems Humoral Factors, pp 257-280.
8. Gillespie JP, Kanost MR, Trenczek T (1997) Biological mediators of insect immunity. Annu Rev Entomol 42: 611-643.
9. Cerenius L, Soderhall K (2004) The prophenoloxidase-activating system in invertebrates. Immunol Rev 198: 116-126.
10. Gollas-Galva´n T, Herna´ ndez-Lo´pez J, Vargas-Albores F(1997) Effect of calcium on the Prophenoloxidase system activation of the brown shrimp (*Penaeus californiensis*, *Holmes*). Comp Biochem Physiol A 117: 419-25.
11. Wade D, Boman A, Wahlin B, Drain CM, Andreu D, et al. (1990) All-D amino acid-containing channel-forming antibiotic peptides. Proc Natl Acad Sci USA 88: 4761-4765.
12. Cociancich S, Ghazi A, Hetru C, Hoffmann JA, Letelliers L (1993) Insect defensin, an inducible antibacterial peptid forms voltage-dependent channels in Micrococcus luteus. J Biol Chem 268: 19239-19245.
13. Gill SC, Von Hippel PH (1989) Extinction coefficient. Anal Biochem 182: 319-328.
14. Guruprasad K, Reddy BVP, Pandit MW (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Prot Eng 4: 155-164.
15. Ikai AJ (1980) Thermo stability and aliphatic index of globular proteins. J Biochem 88: 1895-1898.
16. Kyte J, Doolottle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105- 132.
17. Gasteiger E (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: John M Walker ed, The Proteomics Protocols Handbook, Humana Press 571-607.
18. Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJA, et al. (2002) The PROSITE database, its status in 2002. Nucleic Acids Res 30: 235-238.
19. Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol 234: 779-815.
20. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.
21. Marti-Renom MA, Stuart A, Fiser A, Sánchez R, Melo F, et al. (2000) Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 29: 291-325.
22. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, et al. (1993) CHARMM: a program for macromolecular energy minimization and dynamics calculations. J Comput Chem 4: 187-217.
23. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereo chemical quality of protein structures. J Appl Crystallography 26: 283-291.
24. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. J Mol Biol 7: 95-99.
25. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8: 477-486.
26. Vriend G (1990) WHAT IF: A molecular modeling and drug design program. J Mol Graph 8: 52-56.
27. Guex N, Manuel CP (1997) Data Modeling, Analysis and Classification SWISSMODEL and the Swiss-Pdb Viewer: An environment for comparative protein.
28. Weiner SJ, Kollman PA, Nguyen DT, Case DA (1986) An all atom force field for simulations of proteins and nucleic acids. J Comput Chem 7: 230-252.
29. Baker D, Sali A (2001) Protein Structure Prediction and Structural Genomics. Science 294: 93-96.
30. Ferre F, Clote P (2005) Disulfide connectivity prediction using secondary structure information and diresidue frequencies. Bioinformatics 21: 2336-2346.
31. Hirokawa T, Boon-Chieng S, Mitaku SB (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatic 14: 378-379.
32. Christian JAS, Cerutti L, Hulo N, Gattiker A, Falquet L (2002) PROSITE: A documented database using patterns and profiles as motif descriptors. Brief Bioinform 3: 265-274.
33. Eisenberg, D, Lüthy R, and Bowie JU (1997) VERIFY3D:Assessment of protein models with three-dimensional profiles. Methods Enzymol 277: 396-404.
34. Fariselli P, Casadio R (2001). Prediction of disulfide connectivity in proteins. Bioinformatics 17: 957-964.
35. Geourjon C, Deléage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 11: 681-684.
36. Krieger E, Nabuurs SB, Vriend G (2003) Homology modeling. Struct bioinform 44: 507-521.
37. Laskowski RA, Watson JD, Thornton JM (2005) ProFunc: A server for predicting protein function from 3D structure. Nucleic Acids Res 89-93.
38. Suresh A, Verma C (2006) Modelling study of dimerzation in mammalian defensins. BMC Bioinformatics 5: S17.