INTERNATIONAL JOURNAL OF ADVANCED BIOLOGICAL RESEARCH

www.scienceandnature.org

BIOINFORMATICS OF MICROBIAL PROTEINASE THROUGH DIFFERENT TOOLS

[a]Rajamurugan, J., [b]Brindha, S., [c]Maragathavalli, S., [d]Annadurai, B. & [d]Gangwar,S K.
[a]Department of Plant science, Bharathidasan Government College for Women, Puducherry – 605 003 India.,
[b]Department of Bioinformatics, New Prince Arts and Science College, Chennai,
[c]Department of Biochemistry, King Nandhivarman College, Thellar,
[d]Department of Crop and Horticultural Sciences, Biotechnology Team, College of Dryland Agriculture and Natural Resources, Mekelle University, Mekelle, Ethiopia.

ABSTRACT
Currently Bioinformatics includes various analytical tools to determine via computer vs experimental things like gene location within a chromosome, finding similar genes or proteins from other species and determining three dimentional structure, physiochemical properties and functions of different proteins. This analysis can enable or greatly accelerate drug target identification, drug lead validation and optimization, pharmacognostic studies and many other biotechnological application. The methodology of bioinformatics is based on the informational analysis of nucleotide and protein sequences. Molecular modeling represents a powerful experimental approach that provides understanding of numerous physical and chemical aspects of protein molecules. Such modeling methods resulted in appearance of adequate software such as Prot Param, SOSUI, Compute pI/Mw, Prot scale and Signal P. These software's are easy in use. These proteomics expasy tools show the different characters of proteinase. This result shows that large number of microbes under study is a good producer of extracellular proteinase and its characters which can be beneficiary for industries

KEYWORDS: Bioinformatics, proteinase, microbes, proteomics tools, prot param, SOSUI, Compute pI/Mw, Prot scale and Signal P

INTRODUCTION

Enzymes are delicate protein molecules necessary for life. Protease are the single class of enzymes which occupy a pivotal position due to their wide applications in detergent, pharmaceutical, photography, leather, food and agricultural industries. An important biotechnological application of protease is in bioremediation processes (Gupta *et al*. 2002). Among the various proteases, bacterial proteases are the most significant, compared with animal and fungal proteases. These proteases have wide applications in pharmaceutical, leather, laundry, food and waste processing industries (Pastor *et al*., 2001). In plants, proteases execute an important role in the overall process of protein turnover during seed development, germination and senescence has investigated by Panicker *et al*. (2009). Expression and production of microbial proteinase and principles of methodology have been published in previous paper.

For purposes of this review, we define bioinformatics as the backbone computational tools and databases that support genomic and related research, which broadly encompasses the study of DNA structure/function, gene expression and protein production/structure/function. Computer methods are widely employed in modern biochemistry and molecular biology for various purposes. During the last decade bioinformatics become popular and intensively developing areas. The methodology of bioinformatics is based on the informational analysis of nucleotide and protein sequences. We believe that the bioinformatic approaches promote better understanding of the fundamentals of biocatalysis. In the present paper we have used proteomics expasy tools of bioinformatics for different characteristics of these useful enzyme proteinase from microorganisms.

METERIALS AND METHODS

Determination of Protein sequence of Proteinase

To know the protein sequence of proteinase first the web page namely NCBI home page (http://www.ncbi.nlm.nih/gov) has to be opened. In the search column, the proiten has to be selected for proteinase. Then "Go" button has to be clicked.

The protein sequence of proteinase in different species with their Id will appear. In this Micro organism name has to be selected with Id No is AAB72063. It has to be clicked. The common features of selected protein sequence will appear. In this fragment 1-60, 61-120, 121-180, 181-240, 241-300, 300-360, 361-420, 421-460, 461-466 is made retrieved the whole sequence is converted in to Fasta format.

SYSTEMS AND METHODS

Prot Param

In Prot param program the computer shows various physiochemical properties of protein. To know physiochemical proteies of proteinase first we go to protemic expasy tools by the help of http://www.expasy.ch/tools. The home page of proteomic expasy tools will appear. In this primary structure analysis tool, another characters are also appear like compute pI/MW. Alycan mass, peptide cutter, protparam etc., In this, protparam has to be selected and clicked. The protein submission page will then be visible. Now the protein sequence of proteinase has to be submitted in Fasta format with the help of submit button. Physiochemical properties

of proteinase will appear (Gill. S.C. and Hippel, P.H., 1989).

Method for SOSUI

SOSUI tool is one of proteomic expasy tools. This tool predicts transmembrance region present in the protein sequence. This tool also calculates length of the transmembrane region of protein sequence.

To know transmembrane region of proteinase first we go to proteomic expasy tools with the help http://www.expasy.ch/tool. The home page proteomic expasy tool will appear. In this topology prediction tool also will appear such as Tmpred, Topred, SOSUI, etc. In this SOSUI has to be selected and clicked, the protein submission page of SOSUI will be visible. Now the protein sequence pasted in the submission place and submitted by using submit button. The result page for proteins will appear. (Mitaku *et al* 2002)

Compute pI/Mw

Compute pI/Mw is one of the proteomics expasy tools, specifically it is a primary structure analysis tools.

To know isoelectric point and molecular weight of proteinase first we go to proteomics expasy tools through http: www.expasy.ch/tools. The homepage of proteomics expasy tools will appear. In this primary structure analysis tools also appear like compute pi/Mw, glycan mass, peptide mass, peptide cutter, prot param, prof scale etc., from these compute pi/mw tool has to be selected and clicked. The protein sequence submission page of compute pi/mw will be visible. Now the protein sequence of proteinase has to be pasted in faster format & submit button. The result page of compute pi/mw will appear.(Bjellquist, *et al*. 1994)

Prot scale

Prot scale is one of the proteomics expasy tools, specifically. It is primary structure of protein analysis tool. To know hydrophobic value for individual amino acid present in protein sequence. First we go proteomics expasy tools through http:// www.expasy.ch/tools. The home page of proteomics expasy tools will appears, like compute pI/MW, glycan mass, peptide cutter, protparam, prot scale etc., from there the prot scale tools has to be selected and clicked the protein submission page of protscale will be visible after that the protein sequence of proteinase has to be submitted by using submit button. The result page of prot scale of protinase will appear (Gasteiger *et al*., 2005).

Signal P

Signal P tool is one of the expasy proteomic tools. Specifically it is a post translation modification tool.

To known the signal peptide present in proteinase, first we go to proteomics expasy tools specifically we go to post translation modification tools through the web address of http : www.expasy.ch/tools. These are several post translation modification tools like chlorop, lipo P, net Ogly, Net N gly, Gpi-som, DGPI, signal P, Multistolator, Net phos,net pico RNA, Sulfinalor, sumoplot. In these signal P has to be selected and clicked. The protein submission page the protein segment of proteinase has to be passed in the faster format in protein sequence submission place after that submitted the protein segment by the help of submit button. The result page of signal P will appear (Henrik Nielsen *et al*. 1997).

RESULTS

Fig.1 shows the theoretical isoelectric point and molecular weight of the enzyme proteinase from this program the molecular weight of the proteinase is conformed as 43387.78 and the isoelectric point of the protenase is 7.82.

FIGURE 1. Compute pI / Mw

Theoretical pI/Mw (average) for the user-entered sequence: proteinase

```
        10         20         30         40         50         60
AIKGMIQSLD DPYSTYMDQE QAKSFDETIS ASFEGIGAQV EEKDGEILIV SPIKGSPAEK
        70         80         90        100        110        120
AGIKPRDQII KVNGKSVKGM NVNEAVALIR GKKGTKVKLE LNRAGVGNID LSIKRDTIPV
       130        140        150        160        170        180
ETVYSEMKDN NIGEIQITSF SETTAKELTD AIDSLEKKGA KGYILDLRGN PGGLMEQAIT
       190        200        210        220        230        240
MSNLFIDKGK NIMQVEYKNG SKEVMKAEKE RKVTKPTVVL VNDGTASAAE IMAAALHESS
       250        260        270        280        290        300
NVPLIGETTF GKGTVQTAKE YDDGSTVKLT VAKWLTADGE WIHKKGIKPQ VKAELPDYAK
       310        320        330        340        350        360
LPYLDADKTY KSGDTGTNVK VAQKMLKALG YKVKVNSMYD QDFVSVVKQF QKKEKLNETG
       370        380        390
ILTGDTTTKL MIELQKKLSD NDTQMEKAIE TLKKEM
```

Theoretical pI/Mw: 7.82 / 43387.78

Fig. 2 shows the Prot param it is the user provided sequence of proteinase. It shows a number of amnio acid in proteinase, molecular weight, theoretical pI, amino acid composition, total number of negatively charged residues (Asptculor)., total number of positively charge residues, atomic composition formula, total number atom present in the protenase. Extinct air co-efficient, estimated of life instability indere (II), and grand average hydropathicity (gravy) of proteinase. By this, the number amino acid in proteinase found to be 466.

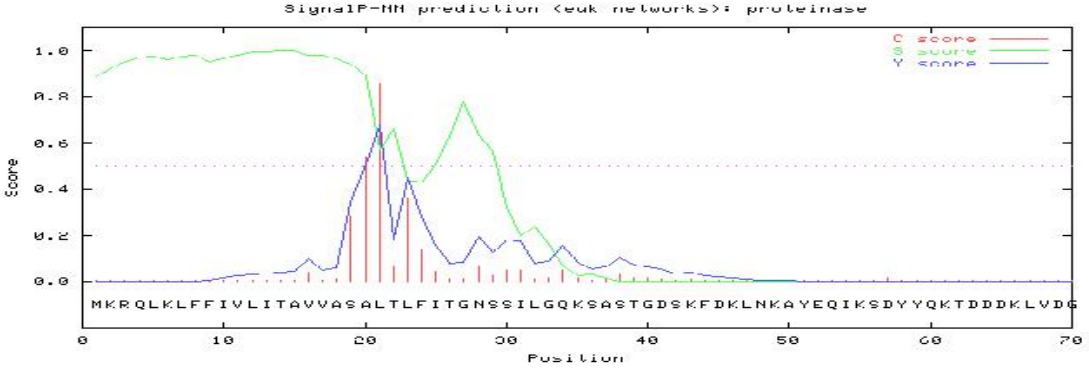
FIGURE 2. ProtParam result
User-provided sequence: proteinase

```
      10                  20          30          40              50          60
MKRQLKLFFI VLITAVVASA  LTLFITGNSS ILGQKSASTG DSKFDKLNKA YEQIKSDYYQ

      70                  80          90          100         110         120
KTDDDKLVDG AIKGMIQSLD DPYSTYMDQE QAKSFDETIS ASFEGIGAQV EEKDGEILIV

      130         140         150             160         170             180
SPIKGSPAEK AGIKPRDQII KVNGKSVKGM NVNEAVALIR GKKGTKVKLE LNRAGVGNID

      190         200         210         220         230         240
LSIKRDTIPV ETVYSEMKDN NIGEIQITSF SETTAKELTD AIDSLEKKGA KGYILDLRGN

      250         260         270         280         290             300
PGGLMEQAIT MSNLFIDKGK NIMQVEYKNG SKEVMKAEKE RKVTKPTVVL VNDGTASAAE

      310         320         330         340         350         360
IMAAALHESS NVPLIGETTF GKGTVQTAKE YDDGSTVKLT VAKWLTADGE WIHKKGIKPQ

      370         380         390         400             410         420
VKAELPDYAK LPYLDADKTY KSGDTGTNVK VAQKMLKALG YKVKVNSMYD QDFVSVVKQF

      430         440         450         460         466
QKKEKLNETG ILTGDTTTKL MIELQKKLSD NDTQMEKAIE TLKKEM
```

Number of amino acids: 466

Figure 3. shows the signal P using the neural networks by this it is understood the proteinase length 70, measure position value cut of is a 0.860 to 0.819 and minimum C is 0.32 to 0.43, the most likely cleavage site between position 20 and 21: ASA - L & T

FIGURE 3. Signal P- result




```
# data
>proteinase          length = 70
# Measure  Position  Value  Cutoff  signal peptide?
  max. C    21       0.860   0.32   YES
  max. Y    21       0.674   0.33   YES
  max. S    14       0.997   0.87   YES
  mean S     1-20    0.963   0.48   YES
      D      1-20    0.819   0.43   YES
# Most likely cleavage site between pos. 20 and 21: ASA-LT
```

Fig. 4 shows SOSUI result for proteinase this stable shows the transmembrane helix region of proteinase and the type of protein and length of the transmembrane region. By this it is known transmembrane region of the proteinase is LEFILITAVVASALTLITNGS. The N-terminal end of the transmembrane regions 7th position, C-terminal end of the transmembrane region 29th position, type of proteinase is primary type. The length of the transmembrane region in proteinase is 23. from the above results it is concluded that there are 23 amino acids are present in the transmembrane regions. And that too from the 7th position of N-terminal to 29th position of C-terminal of the

proteins. Hence, it is concluded the protease is the membrane protein.

SOSUI Result for proteinase
This amino acid sequence is of a MEMBRANE PROTEIN
which has 1 transmembrane helix.

No.	N terminal	Transmembrane region	C- terminal	Type	Length
1	7	LFFIVLITAVVASALTLFITGNS	29	PRIMARY	23

FIGURE 5. Shows the ball and stick model of proteinase

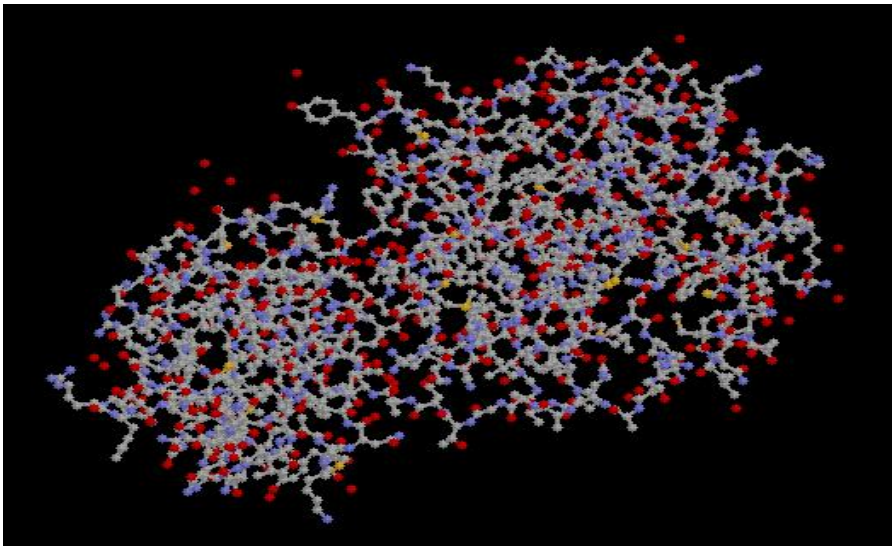


FIGURE 6. Shows the wire frame model of proteinase.

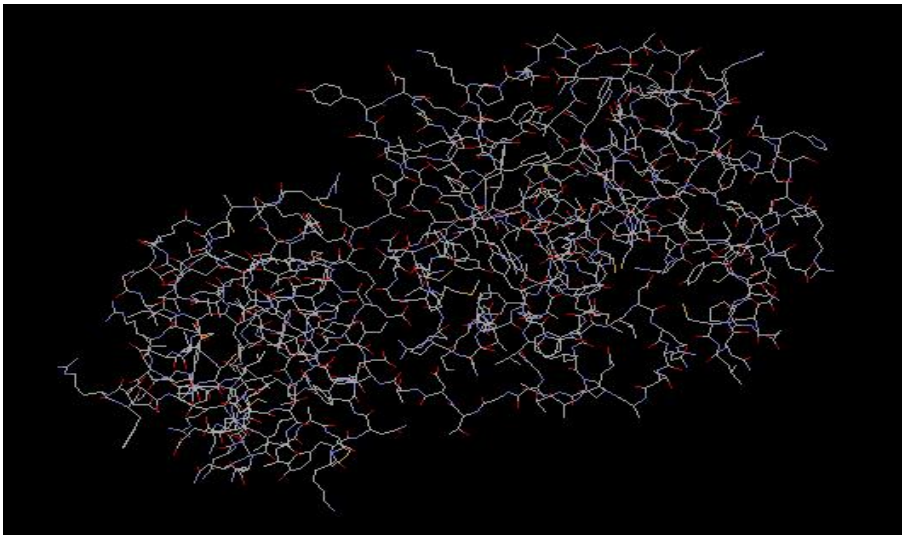
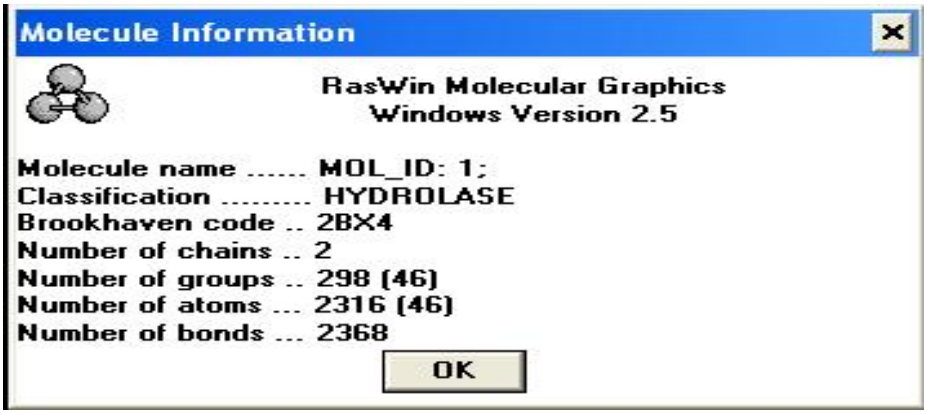


FIGURE 7. shows the RasWin molecular Graphics of proteinase. The Molecular name is: MOL_ID:1; Classification is : Hydrolase, Brookhaven code is : 2BX4, Number of Chains are : 2, Number of groups are : 298 [46], Number of atoms are: 2316 [46], Number of bonds are : 2368.

FIGURE 7. RasWin molecular Graphics of proteinase



DISCUSSION

Proteases are one of the most important industrial enzymes accounting for nearly 60% of the total world wide enzyme sales. microbial proteinases are receiving much attention, because of their increasing applications in many industries like food, chemical, leather and pharmaceutical industries. The new field of enzyme engineering has in a short period of time made striking contribution to industry, medicine, agriculture and in pollution control. In prot param program the computer shows various physiochemical properties of protein such as theoretical isoelectric point and moklecular wight of the enzyme proteinase is conformed as 43387.78 and the isoelectric point of the proteinase is 7.82. The number of aminoacids in proteinase found to be 466. The signal P using the neural networks is understood the proteinase length 70, measure position value cut of is a 0.860 to 0.819 and minimum C is 0.32 to 0.43, the most likely cleavage site between position 20 and 21:ASA-L&T. SOSUI results for proteinase shows the transmembrane helix region of proteinase and the type of protein and length of the transmembrane region. The ball and stick model of proteinase, wire frame model of proteinase, and RasWin molecular Graphics was established.

Thus at last it can be said that this Different Characters of Microbial Proteinase through Bioinformatics tools has wide application in various industrial and medical fields and it can be produced in large scale from different microorganism.

REFERENCES

Bjellqvist, B., Basse, B., Olsen, E. and Celis, J. E. *Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions.* Electrophoresis 1994, 15, 529-539.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005).pp.571-607.

Gill, S.C. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. 182:319-326(1989).

Gupta, R., Beg, Q.K., Lorenz, P. (2002) Bacterial alkaline proteases: molecular approaches and industrial applications. Appl. Microbiol. And Biotechnol. 59(1): 15-32

Henrik Nielsen, Jacob Engelbrecht, Søren Brunak and Gunnar von Heijne. Protein Engineering, 10:1-6, 1997.

Mitaku S., Hirokawa T., and Tsuji T., *Bioinformatics*, (2002) Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. 18 608-16

Paster, M.D., Lorda, G.S., Balatti, A. (2001). Protease observation using Bacillus subtitis3411 and amaranth seed meal medium at different aeration rates. Brazilian J, Microbiol, 32: 1-8.

Panicker, L. M., Usha, R., Roy, S., Mandal, C. (2009) Purification and characterization of a serine protease (CESP) from mature coconut endosperm, BMC Res Notes, 9,2-81.