IJPS INTERNATIONAL JOURNAL OF PLANT SCIENCES Volume 9 | Issue 1 | January, 2014 | 148-153

In silico analysis of pectinase enzymes from ascomycetes fungus

■ ANUPSINGH THAKUR, ATUL S. HANDE AND MAHESH J. PATIL

SUMMARY

Pectinases are a big group of enzymes that break down pectic polysaccharides of plant tissues into simpler molecules like galacturonic acids. Pectinases are one of the most widely distributed enzymes in bacteria, fungi and plants. Some fungal species are used in industries as source of pectinases enzymes. It has long been used to increase yields and clarity of fruit juices. In the present investigation, total eleven protein sequences of pectinase enzyme from different fungal species of Ascomycetes were obtained from NCBI protein database and were considered for physico-chemical characterization including pI, EC, AI, GRAVY and instability index, motif discovery, motif family analysis and phylogenetic analysis. Three different motifs were discovered by MEME program where minimum motif width was 6 and maximum motif width was 50. All three discovered motifs were aligned using MAST tool which revealed the similarity between all of three submitted motif's sequence. The motif matches have shown a position p-value less than 0.0001. Each of the following 11 sequences has an E-value less than 100. There is only one major sequence clusters were constructed by phylogenetic analysis.

Key Words : In silico, Pectinase, Enzyme, Ascomycetes, Fungus

How to cite this article : Thakur, Anupsingh, Hande, Atul S. and Patil, Mahesh J. (2014). *In silico* analysis of pectinase enzymes from ascomycetes fungus. *Internat. J. Plant Sci.*, 9 (1): 148-153.

Article chronicle : Received : 07.10.2013; Revised : 26.10.2013; Accepted : 05.11.2013

The increasing energy demands have focused worldwide attention on the utilization of renewable resources, particularly agricultural and forest residues, the major components of which are cellulose, starch, lignin, xylan and pectin. These materials have attracted considerable attention as an alternative feedstock and energy source, since they are abundantly available. Several microbes are capable of using these substances as carbon and energy sources by producing a vast array of enzymes in different environmental niches

🛶 MEMBERS OF THE RESEARCH FORUM 🛶

Author to be contacted :

ANUPSINGH THAKUR, Dr. B.S. Konkan Krishi Vidyapeeth, Dapoli, RATNAGIRI (M.S.) INDIA Email: anupthakur34@yahoo.com

Address of the Co-authors: ATUL S. HANDE, Dr. B.S. Konkan Krishi Vidyapeeth, Dapoli, RATNAGIRI (M.S.) INDIA Email: ahande935@gmail.com

MAHESH J. PATEL, College of Pharmacy (MSBT), Sangulwadi, MUMBAI (M.S.) INDIA Email: mpatil.4859@gmail.com (Antranikian, 1992; Kaur *et al.*, 2004). Pectinase production occupies about 10% of the overall manufacturing of enzyme preparations. Pectinolytic enzymes are widely used in the food industry for juice and wine production (Semenova *et al.*, 2006).

This review describes the pectinolytic enzymes and their substrates, the microbial pectinase production and characterization, and the industrial application of these enzymes. *In silico* approaches provide a viable solution to major drawbacks of experimental analysis of protein in regards to physicochemical and structural properties. Time requirement, high cost and the methods without amenable to high throughput techniques are the major drawbacks of experimental analysis of protein whereas 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.

MATERIAL AND METHODS

Eleven amino acid sequences of pectinase from different organisms of different fungal species of Ascomycetes were

ANUPSINGH THAKUR, ATUL S. HANDE AND MAHESH J. PATIL

Table A : Protein sequences considered for the study							
Organism	Enzymes	Accession number					
Aspergillus kawachii	pectinase 2	gi 358373850 dbj GAA90446.1					
	pectinase 3	gi 358372427 dbj GAA89030.1					
	pectinase A	gi 358368561 dbj GAA85178.1					
Aspergillus aculeatus	pectin methylesterase	gi 3220207 gb AAC23565.1					
Aspergillus oryzae	polygalacturonase	gi 333943955 dbj BAA03244.2					
Penicillium expansum	pectinase	gi 7388004 sp O59925.1 PGLR_PENEN					
Penicillium digitatum	pectinase	gi 7388011 sp Q9Y718.1 PGLR_PENDI					
Penicillium janthinellum	pectinase	gi 7388002 sp O42824.1 PGLR_PENJA					
Emericella nidulans	Endopolygalacturonase B	gi 294958172 sp Q5B508.2 PGLRB_EMENI					
	Polygalacturonase D	gi 74657032 sp Q5AYH4.1					
	Pectinase AN8327	gi 74656855 sp Q5ATQ3.1 PGLR_EMENI					

obtained from NCBI protein database and only 347 characters of each sequence were considered for motif discovery, motif family analysis and phylogenetic analysis (Bairoch and Apweiler, 2000) as shown in Table 1.

For physico-chemical characterization, theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction co-efficient (Gill and Von Hippel, 1989), instability index (Guruprasad *et al.*, 1990), aliphatic index (Ikai, 1980) and grand average hydropathy (GRAVY) (Kyte and Doolottle, 1982) were computed using the Expasy's ProtParam server (Gasteiger, 2005) (*http://us.expasy.org/tools/protparam.html*).

Motifs were identified in sequences using MEME program (Bailey and Elkan, 1995). All three discovered motifs were aligned using MAST program (Bailey and Gribskov, 1998) to judge the similarity between three motifs discovered. The multiple sequence alignment was performed using MUSCLE programme and CLUSTAL-W (Thompson *et al.*, 1994) program before we construct phylogenetic tree. The phylogenetic analysis was performed by UPGMA method using MEGA4 programme (Kumar *et al.*, 2008).

RESULTS AND DISCUSSION

The accession number of retrieved sequences along with the organism is listed in Table 1. The physio-chemical characterization were computed by Expasy's ProtParam tool as shown in Table 2. Isoelectric point (pI) value was calculated because that value of protein is stable and compact. pI value is the pH at which protein has no net charge. The computed pI

Table 1 : Physio-chemical parameters of pectinases enzymes											
Organism	Emzymes	Accession number		M.wt	pI	-R	+R	EC	Π	AI	GRAVY
Aspergillus kawachii	pectinase 2	>gi 358373850 dbj GAA90446.1	362	37603.8	5.04	38	27	50920	31.58	78.12	-0.117
	pectinase 3	>gi 358372427 dbj GAA89030.1	384	404260	4.03	58	18	47495	31.12	77.4	-0.147
	pectinase A	gi 358368561 dbj GAA85178.1	370	38718.9	3.99	59	20	42900	34.95	73.24	-0.312
		gi 1223912 gb AAB42153.1	331	35681	4.23	38	16	64415	21.92	74.35	-0.191
Aspergillus aculeatus	pectin methylesterase	gi 3220207 gb AAC23565.1	378	38528.3	4.83	32	20	41410	25.44	77.57	-0.06
Aspergillus oryzae	polygalacturonase	gi 333943955 dbj BAA03244.2	363	37651.3	4.51	40	23	42900	32.31	75.7	-0.194
Penicillium expansum	pectinase	gi 7388004 sp O59925.1	378	38028	6.9	22	22	36940	26.72	81.48	0.038
Penicillium digitatum	pectinase	gi 7388011 sp Q9Y718.1	367	37651.2	8.04	30	32	36940	26.98	78.64	-0.134
Penicillium janthinellum	pectinase	gi 7388002 sp O42824.1	371	38326.9	4.27	46	23	45420	30.34	76.95	-0.174
Emericella nidulans	<i>Emericella</i> Endopolygalacturonase gi 294958172 sp Q5B5 nidulans B		364	37857.2	7.52	33	34	39920	23.67	72.88	-0.226
	Polygalacturonase D	gi 74657032 sp Q5AYH4.1	514	52731.6	4.51	54	19	34350	40.82	71.13	-0.132
	Pectinase AN8327	gi 74656855 sp Q5ATQ3.1	380	39359.7	3.89	56	16	49890	29.55	74.39	-0.182

Internat. J. Plant Sci., 9 (1) Jan., 2014 : 148-153 Hind Agricultural Research and Training Institute

value of Pectinase from *Penicillium expansum*, *Penicillium digitatum* and *Emericella nidulans* (Endopolygalacturonase B) were greater than 7 (pI<7) indicates that these proteins were considered as basic. The computed isoelctric point (pI) will be useful for developing buffer system for purification by isoelectric focusing method. EC value and II value are shown in Table 1. Higher EC revealed that presence of higher concentration of light absorbing amino acids *i.e.* Cys, Trp and Tyr. Protein with low Instability index (smaller than 40) are

stable (Guruprasad et al., 1990).

AI value of pectinases protein are shown in Table 1. The very high aliphatic index of all pectinase protein sequences indicates that these antioxidant proteins may be stable for a wide temperature range.

Motif discovery result revealed that three motifs were discovered (Fig. 1). Three conserved motifs were found as shown in Fig. 8. The motif matches shown have a position p-value less than 0.0001(Fig. 8). The multiple sequence

gi 3220207 gb AAC23565.1	212	1.23e-57	IGTSTYVTIS	GATVYNQDDCVAVNSGENIYFSGGYCSGGHGLSIGSVGGRSDNTVKNVTF	VDSTIINSDN
gi 358368561 dbj GAA85178.1	200	1.09e-56	ISESTYITIT	GATVKNQDDCVAINSGENIYFSGGTCSGGHGLSIGSVGGRDDNTVKNVTF	IDSTVSDSEN
gi 7388002 sp O42824.1 PGLR_PENJA	203	7.38e-53	ISESTGVTIR	NAVVKNQDDCIAINSGQNIYFTGGTCSGGHGLSIGSVGGRDDNTVKNVTI	TDSTVTDSAN
gi 294958172 sp Q5B508.2 PGLRB_EMENI	196	1.24e-52	IGSSTYITID	GATVYNQDDCIAINSGEHITFTNGYCSGGHGLSIGSVGGRSDNTVKSVTI	SNSKVVDSQN
gi 358372427 dbj GAA89030.1	215	5.04e-52	IGESTYITIT	GAEIYNQDDCVAINSGENIYFSASVCSGGHGLSIGSVGGRDDNTVKNVTF	YDVNVLKSQQ
gi 7388004 sp O59925.1 PGLR_PENEN	211	1.95e-51	IGDSDSITIT	GATVYNQDDCLAINSGTNIVFSGGYCSGGHGLSIGSVGGRSNNVVETVHI	SSTQVVNSQN
gi 333943955 dbj BAA03244.2	195	1.03e-50	VGSSTYINID	GATVYNQDDCLAINSGSHITFTNGYCDGGHGLSIGSVGGRSDNTVEDVTI	SNSKVVNSQN
gi 74656855 sp Q5ATQ3.1 PGLR_EMEN	208	1.59e-49	IGESSNVVIT	GAKVYNQDDCVAVNSGTSITFSGGTCSGGHGLSIGSVGGRDDNTVDTVTF	KDSTVSNSVN
gi 358373850 dbj GAA90446.1	194	2.38e-48	VGNSVGVNII	KPWVHNQDDCLAINSGENIWFTGGTCIGGHGLSIGSVGDRSNNVVKNVTI	EHSTVSNSEN

Fig. 1 : Site of Block one

The beight of the moht "block" is proportional to -log(p-value), truncated at the beight for a moht with a p-value of 1e-10. Click on any row to highlight sequence in all mohts. Mouse over the center of the moht blocks to see more information.



Fig. 2 : Block one show the motif location in each pectinase sequences

ites 🖾

Click on any row to highlight sequence in all motifs.

Name	Start	p-value	Sites ?
gi 73880C2 sp O42824.1 PGLR_PENJA	53	1.04e-45	AASASASKSS CSTIVLSNIEV AGKTLDLTDLKDGTKVIFEGTTTFGYKEW SGPLIKISGS
gi 333943955 dbj BAA03244.2	<mark>4</mark> 6	1.09e-44	AADAKSGKTS CSTITLSNIEV AGETLDLTGLNDGTTVIFSGETTFGYKEW EGPLISVSGT
gi 358372427 dbj GAA89030.1	63	1.61e-43	ASKASKSKTS CSTIYLSDVAV SGTTLDLTDLNDGT VIFQGETTFGYEEW TGPLVSVSGT
gi 3220207 gb AAC23565.1	61	4.90e-43	ASSASKSKTS CSTIVLSNVAV SGTTLDLTKLNDGTHVIFSGETTFGYKEW SGPLISVSGS
gi 294958172 sp Q53508.2 PGLRB_EMENI	47	1.63e-42	ASAAKSGASK CSTVTLKSIQV AGETLDLTGLKSGATVIFEGETTFGYKEW KGPLISMSGD
gi 7388011 sp Q9Y718.1 PGLR_PENDI	49	3.12e-42	AEAAKAGKGS CSTIVLDNIKVPAGETLDLTKLKSGTQVVFKGETSFGYKEW TGPLVSFSGS
gi 358373850 dbj GAA90446.1	45	3.36e-40	AAAAKAGKAK CSTITLDSIKV AGTTLDLTGLTSGTKVIFEGTTTFD EEW AGPLISMSGK
gi 73880C4 sp 059925.1 PGLR_PENEN	57	7.11e-39	AAAAIAGKAG CSSITLNNVVVPAGTTLDLTGLASGTKVIFEGTTTFGYKQW AGPLISISGT
gi 358368561 dbj GAA85178.1	50	4.94e-37	AATASESKTS CSDIVLKDITV AGETLNLKDLNDGTTVTFEGTTTWEYEEW DGPLLRISGK
gi 74656855 sp Q5ATQ3.1 PGLR_EMENI	57	1.89e-36	AASASRSQTD CATITLSDITVPSGTTLDLSDLEDDTTVIFEGTTSWEYEEW DGPLLQIKGN
gi 74657C32 sp Q5AYH4.1 PGLRD_EMENI	188	8.76e-27	YAGISSAVAS CSNIMLSDVYAPPSSTIDLQDLQTGAAVIFAGKTTFGDTSD SDFDPIVISG

Fig. 3 : Site of block two

Internat. J. Plant Sci., 9 (1) Jan., 2014 : 148-153 Hind Agricultural Research and Training Institute

ANUPSINGH THAKUR, ATUL S. HANDE AND MAHESH J. PATIL

Block Diagrams

The height of the motif "block" is proportional to -log(p-value), truncated at the height for a motif with a p-value of 1e-10. Clock on any row to highlight sequence in all motifs. Morse over the center of the motif blocks to see more information.



Time 4.7 secs

Fig. 4 : Block two show the motif location in each pectinase sequences

Sites ?

Click on any row to highlight sequence in all motifs.

Name	Start	<i>p</i> -value		Si	tes ?	
gi 333943955 cbj BAA03244.2	88	1.50e-36	ETTFGYKEWE	GPLISVSGTNIKVQQASG	AKIDGDGSRWWDGKGGNGG	KTKPKFFYAH
gi 7388002 sp 042824.1 PGLR_PENJA	95	2.46e-36	TTTFGYKEWS	GPLIKISGSDITVEAAD	AVINADGSRWWDGEGTNGG	KTKPKFFYAH
gi 3220207 gb AAC23565.1	103	3.14e-36	ETTFGYKEWS	GPLISVSGSDLTITGASG	HSINGDGSRWWDGEGGNGG	KTKPKFFAAH
gi 358373850 cbj GAA90446.1	87	4.51e-36	TTTFDYEEWA	G LISMSGKDITVTGAS	LINCDGSRWWDGKGTSGK	KKPKFFYAHG
gi 358372427 cbj GAA89030.1	105	1.03e-35	ETTFGYEEWT	GPLVSVSGTDITVEGES	AVLNGDGSRWWDGEGGNGG	KTKFKFFYAH
gi 294958172 sp Q5B508.2 PGLRB_EMENI	89	2.60e-35	ETTFGYKEWK	GPLISMSGDKITVKQAS	AKINCDGARWWDTKGSNGG	KTKFKFFSAH
gi 7388004 sp 059925.1 PGLR_PENEN	99	2.36e-34	TTTFGYKQWA	GPLISISGTNIQVSGAS	HLIDGQGSRWWDGEGSNSK	TNIKPKFFFA
gi 74656855 sp Q5ATQ3.1 PGLR_EMENI	99	1.56e-33	TTSWEYEEWD	GPLLQIKGNGITIKGADO	AKLNPDGSRWWDGEGSNGG	VTKFKFFYAH
gi 358368561 cbj GAA85178.1	92	1.17e-31	TTTWEYEEWD	GPLLRISGKDITVTQSSI	AVLNGNGAKWWDGEGTNGG	KTKFKFFYAH
gi 7388011 sp Q9Y718.1 FGLR_PENDI	91	6.82e-31	ETSFGYKEWT	GPLVSFSGSNI VSGAAG	VINGGG SWWDGKGTNGG	KKKFKFFYAH
gi 74657032 sp Q5AYH4.1 PGLRD_EMENI	231	2.39e-24	TTFGDTSDSD	FDPIVISGTNLTITGTEL	HVIDGNGQAYWDGQGSNGG	SDKFDHFIVL

Fig. 5 : Site of Block three

Block Diagrams

The height of the motif "block" is proportional to log(p value), truncated at the height for a motif with a p value of 1e 10. Cick on any row to highlight sequence in all motifs. Mouse over the center of the motif blocks to see more information.





Internat. J. Plant Sci., 9 (1) Jan., 2014 : 148-153 Hind Agricultural Research and Training Institute

In silico ANALYSIS OF PECTINASE ENZYMES FROM ASCOMYCETES FUNGUS

Non-overlapping sites with a p-value better than 0.0001. The height of the motif "block" is proportional to -log(p-value), truncated at the height for a motif with a p-value of 1e-10. Click on any row to highlight sequence in all motifs. The motif blocks have tool tips with more information.



Fig. 7 : Combined block diagram show the motif location of each block

Motif W			Similarity			
	Width	Best possible match	1	2	3	
1	50	GAWV:NQCDCLAINSGENIYFSGG/CSGG/GLSIGSVGGRODNTVMNVTF		0.25	0.22	
2	41	CSTIMLSNIEV AGETLOLIOLKOGT VIFEGETIFG KEW	0.25		0.20	
3	37	G-LIRMSGRNITVEQASG-VINGDGSRMWDGEGINGG	0.22	0.20		

Best possible match diagram showing similarity between Fig. 8 : submitted motif



Fig. 9 : Phylogenetic analysis of all retrieved sequences of pectinase from different organisms of different fungal species of ascomycetes



Fig. 10: Conserved motifs of pectinases

alignment result showed some conserved regions in all aligned sequences.

Phylogeneic analysis for pectinases proteins from different Ascomycetes fungi was done by MEGA 4 tool. Results were shown in Fig. 9. There is single cluster to all proteins revealed that the functional motif are conserved in evolution.

Conclusion:

Motifs identification and similarity in a group of related sequences of pectinases showed the evolutionary relationships of function features among different organisms of fungal species of ascomycetes. This suggests that these motifs have an important function in the evolution of pectinases in different fungal species of ascomycetes.

REFERENCES

- Antranikian G. (1992). Microbial degradation of starch, In: Winkelmann G, Ed. Microbial degradation of natural products. VCH, Weinheim pp. 27-51.
- Bailey, T.L. and Elkan, C. (1995). Unsupervised learning of multiple motifs in biopolymers using expectation maximization. Mach. Learn., 21 (51): 80-33.
- Bailey, T.L. and Gribskov, M. (1998). Combining evidence using p-values: application to sequence homology searches. Bioinformatics, 14(1):48-54.
- Bairoch, A. and Apweiler, R.(2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res., 28 (1): 45-48.
- Gasteiger, E. (2005). Protein Identifi cation and Analysis Tools on the ExPASy Server. In: John M. Walker ed, The proteomics protocols handbook, Humana Press 571-607.

Internat. J. Plant Sci., 9 (1) Jan., 2014 : 148-153 Hind Agricultural Research and Training Institute

- Gill, S.C. and Von Hippel, P.H. (1989). Extinction co-efficient. *Anal. Biochem.*, **182**: 319- 328.
- Guruprasad, K., Reddy, B.V.P. and Pandit, M.W. (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. Engg.*, 4: 155-164.
- Ikai, A.J. (1980). Thermo stability and aliphatic index of globular proteins. *J. Biochem.*, **88** (6) : 1895-1898.
- Kaur, G., Kumar, S. and Satyanarayana T. (2004). Production, characterization and application of a thermostable polygalacturonase of a thermophilic mould Sporotrichum thermophile Apinis. *Bioresour. Technol.*, **94** (3): 239-243.
- Kumar, S., Dudley, J., Nei, M., and Tamura, K. (2008). MEGA: a biologist-centric software for evolutionary analysis of

DNA and protein sequences. *Briefings Bioinformatics*, **9** (4) : 299-306.

- Kyte, J. and Doolottle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol., 157 (1): 105-132
- Semenova, M.V., Sinitsyna, O.A., Morozova, V.V., Fedorova, E.A., Gusakov, A.V., Okunev, O.N., Sokolova, L.M., Koshelev, A.V., Bubnova, Iu.P. and Sinitsyn, A.P. (2006). Use of a preparation from fungal pectin lyase in the food industry. *Appl. Biochem. Microbiol.*, 42 (6) : 598-602.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**(22) : 4673-4680.

9th ***** of Excellence *****