

# Cloning, expression and characterization of a metagenome derived thermoactive/thermostable pectinase

Rajvinder Singh · Samriti Dhawan ·

Kashmir Singh · Jagdeep Kaur

Received: 19 September 2011 / Accepted: 5 June 2012 / Published online: 19 June 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** The gene encoding a thermostable pectinase was isolated from a soil metagenome sample. The gene sequence corresponded to an open reading frame of 1,311 bp encoding a translation product of 47.9 kDa. It showed maximum (93 %) identity to a *Bacillus licheniformis* glycoside hydrolase. Deded amino acid analysis showed an absence of highly conserved cysteine residues in the N-terminal region at positions 24 and 42, and in the C-terminal region at positions 389, 394, 413 and 424. pQpecJKR01 (pQE30 expression vector containing the pectinase gene) was expressed in *Escherichia coli* strain M15 as a recombinant fusion protein containing an N-terminal 6× His tag. Biochemical properties of this pectinase were novel. The enzyme had temperature and pH optima of 70 °C and 7.0, respectively, but was active over a broad temperature and pH range. The enzyme was stable at 60 °C with a half-life of 5 h and the enzyme activity was inhibited by 0.1 % diethyl pyrocarbonate and 5 mM dicyclohexyl carbodiimide. The enzyme could be of great use in industrial processes due to its activity over a broad pH range and at high temperature.

**Keywords** Pectinase · GH family 28 · Metagenomics · Thermostability · Recombinant expression

**Electronic supplementary material** The online version of this article (doi:[10.1007/s11033-012-1685-x](https://doi.org/10.1007/s11033-012-1685-x)) contains supplementary material, which is available to authorized users.

R. Singh · K. Singh · J. Kaur (✉)  
Department of Biotechnology, Panjab University,  
Sector-14, Chandigarh 160014, India  
e-mail: jagsekhon@yahoo.com

S. Dhawan  
Department of Biotechnology, GGSD College,  
Chandigarh, India

## Introduction

Pectic polymers are one of the important polysaccharides present in plants and fungi. In plants, pectin is an important component of the middle lamella and primary cell wall. Pectin consists of a backbone of galacturonic acid residues linked by  $\alpha$ -1,4 glycosidic linkages. This backbone is to some extent methyl-esterified and interspersed with rhamnose, arabinan and galactan side chains [1]. Pectic polymers are degraded by the action of pectinolytic enzymes, which can be more or less divided into two categories; pectin esterases, which remove methyl groups and depolymerases that shorten the backbone. The depolymerases may be divided into hydrolases and lyases. A classification based on amino acid sequence similarity is now superseding this original classification scheme because it reflects structural features, reveals evolutionary relationships and provides a convenient tool to derive mechanistic information [2]. Based on this classification, around 300 families are described on the Carbohydrate Active Enzymes; <http://www.cazy.org> database [3], where all hydrolases including polygalacturonases (PGs) are classified in glycoside hydrolase (GH) family 28. Pectic enzymes play an important role in fruit ripening [4], fruit abscission, and plant diseases [5, 6]. They are produced by phytopathogenic fungi, bacteria, nematodes and higher plants. These enzymes are actively involved during penetration of pathogens and subsequent disease developmental stages. Some organisms like *Aspergillus niger* and *Bacillus* sp. secrete multiple forms of pectinases [7–10]. These forms generally have different pH optima and half-life [10]. The presence of various carbon sources and metal-ions in the growth media also leads to variation in the level of enzyme production [10]. Over the years, pectinases have been used in several conventional industrial processes, involved in the textile, tea, coffee, plant fiber processing, oil

extraction, industrial wastewater treatment, and paper and pulp industries [11, 12]. Temperature stability is regarded as one of the most important characteristics of a biocatalyst for use in industrial applications, hence novel enzymes are constantly being sought.

A metagenomic approach is often employed to isolate a large number of novel products like new antibiotics [13], antibiotic resistance genes [14], transporters [15] and novel biocatalysts [16–18]. This approach vastly increases the pool of biological diversity from which novel enzymes may be isolated. Therefore, an attempt has been made in the present investigation to clone a metagenome-derived gene from environmental soil samples collected from hot springs of northern India that encodes a highly thermoactive and thermostable pectinase. The gene was sequenced and compared with known sequences. The gene was expressed in *Escherichia coli* and the enzyme was characterized biochemically.

## Materials and methods

### Sample collection and DNA isolation

Soil samples were collected from hot springs of Manikaran (32.033N, 77.3500E), India, where the temperature varies from 70 to 100 °C. Metagenomic DNA was isolated [19] and purified [20]. Briefly, 1.3 ml of extraction buffer (100 mM Tris–Cl, pH 8.0; 100 mM EDTA, pH 8.0; 1.5 ml NaCl; 100 mM sodium phosphate buffer; CTAB 1 % w/v) was added to 0.5 g soil sample. After complete mixing 100 µg/ml of proteinase K was added and incubated for 45 min at 37 °C. SDS (2 % w/v) was added and incubated at 60 °C for 2 h, followed by a chloroform: isoamyl alcohol (24:1) extraction and precipitation by 0.6 volume isopropanol. Samples were centrifuged at 12,000 rpm for 15 min to pellet down the DNA. The pellet was air-dried and dissolved in TE (10 mM Tris–Cl and 1 mM EDTA, pH 8.0). For purification of DNA, a Q-Sepharose (Sigma, USA) anion exchanger was used to remove humic acid impurities following Sharma et al. [20].

### Cloning, sequencing and analysis of a pectinase gene

Primers were designed based on known sequences of pectinase genes from other species to amplify a metagenomic pectinase gene by PCR. (Forward primer 5'-GTG AGT CTG CAG AAA ATA AAA G-3' and reverse primer 5'-TTA GGC TTT GTG TGA GTC ATA G-3'). The PCR mix contained 30 ng metagenomic DNA, 1.5 mM MgCl<sub>2</sub>, 1X *Taq* buffer, 0.2 mM dNTPs, 0.2 µM of each primer and 1U *Taq* polymerase (Larova, Germany) in a final volume of 25 µl. The following PCR profile was used: 1 cycle 94 °C

for 3 min; 35 cycles of 94 °C for 30 s, 55 °C for 40 s, and 72 °C for 2 min; followed by a final extension at 72 °C for 7 min. Amplified products were analyzed on a 1.2 % agarose gel and then purified using the GenElute™ Gel Extraction Kit (Sigma, USA) according to the manufacturer's instructions. The PCR product was ligated into pGEM®-T vector (Promega, USA) as described by the manufacturer. The recombinant vector was then transformed into *E. coli* JM109 and recombinants were screened on LB ampicillin plates containing X-Gal (40 µg/ml) and IPTG (0.3 mM). Plasmid DNA was prepared according to Sambrook et al. [21].

Sequencing was carried out by Bangalore Genei, India, using an automated ABI 3100 Genetic Analyser.

BLAST ([www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)) and ClustalW [22] programs were used for similarity analysis and sequence alignment. SignalP-hidden markov models (HMM) [23, 24] was used to analyze signal sequences and the 3D structure was predicted using ESyPred3D [25].

### Expression of recombinant PecJKR01<sup>His</sup>

The primer pair, Forward 5'-AA GGATCC ATG AGT CTG CAG AAA ATA AAA G-3' Reverse 5'-AA GTC-GAC TTA GGC TTT GTG TGA GTC ATA G-3', was designed with *Bam*HI and *Sal*I restriction sites, respectively, for in frame ligation into pQE30 expression vector (Qiagen, Germany). PCR reaction conditions were the same as mentioned above. The PCR product was run on a 1.2 % agarose gel and desired amplification product was extracted using the GenElute™ Gel Extraction Kit (Sigma, USA). The gel-purified PCR product and pQE30 vector were digested separately using *Bam*HI and *Sal*I (NEB, USA). The digested fragments were purified from the gel and ligated using T4 DNA ligase (Bangalore Genei, India).

The recombinant plasmid, pQpecJKR01, was transformed into *E. coli* strain M15, containing the pREP4 plasmid and plated on LB ampicillin kanamycin plates. A colony of *E. coli* M15/pQpecJKR01 with insert was confirmed by colony PCR [26].

Effect of varying IPTG concentration on gene expression was studied using 0.05, 0.1, 0.2 and, 1.0 mM of IPTG. A colony was cultured overnight in 5 ml LB medium containing ampicillin (100 µg/ml) and kanamycin (30 µg/ml). The overnight culture was used as inoculum (10 % v/v) to inoculate 100 ml of LB, and grown till OD<sub>620nm</sub> reached 0.45. At this point IPTG of varying concentrations was used to induce the culture and grown for 3 h at 25 °C, after which cells were harvested by centrifugation at 8,000 rpm for 15 min at 4 °C. The cell pellet was resuspended in lysis buffer (10 mM Sodium Phosphate buffer, pH 7.0 and 0.2 % Triton-X 100) to release the induced enzyme which was analyzed by SDS-PAGE [21].

## Pectinase assay

Pectinase activity was assayed using the 3,5-dinitrosalicylic acid (DNS) method for estimating reducing sugar levels [27]. Polygalacturonic acid (PGA) (0.5 % w/v, HiMedia, India) was used as substrate in the reaction mixture with 50 mM phosphate buffer, pH 7.0 and an appropriate amount of enzyme (2.0  $\mu$ M). This reaction mixture was incubated at 70 °C for 20 min. A standard curve was prepared using varying concentrations of galacturonic acid. One unit activity was defined as the amount of enzyme forming 1  $\mu$ mol of galacturonic acid per minute at 70 °C. Specific activity of the enzyme was also calculated. Protein concentration was estimated using a protein estimation kit (Bangalore Genei, India) based on the BCA method.

## Biochemical characterization of the enzyme

Temperature and pH optima were determined by carrying out assays at different temperatures (10–90 °C) and buffers of varying pH (4–11), respectively. After determining temperature and pH optima, all other assays were performed at these conditions. For a thermostability assay the enzyme was incubated for 30 min at different temperatures and then kept for 15 min at 4 °C. The sample without any treatment served as control (100 %). The activity assay was then performed at optimum pH (7.0) and temperature (70 °C). Similarly for pH stability assay the enzyme was incubated for 1 h at 25 °C with buffers of different pH making the final strength of the reaction mix to 50 mM. Half life at 60 °C was determined by incubating the enzyme for 5 h at 60 °C. 30  $\mu$ l of enzyme was removed after every 1 h and kept on ice for 15 min. Similarly, half life at 70 °C was determined by incubating the enzyme for 1 h at 70 °C. 30  $\mu$ l of enzyme was removed after every 10 min and kept on ice for 15 min. Enzyme activity was then determined as described earlier. Effects of 0.1 % DEPC (diethyl pyrocarbonate) and varying concentrations of DCCD (dicyclohexyl carbodiimide) were studied to modify histidine residues and amino acids with carboxylate groups, respectively. Similarly the effect of 0.1, 1, 10, and 100 mM  $\beta$ -mercaptoethanol was also studied. The modifiers and enzyme were incubated for 30 min at 25 °C. The sample without any additive was taken as control (100 %).

## Substrate specificity

Pectin and polygalacturonase (each at 0.5 % w/v) were used to check the substrate specificity of the enzymes. To check pectin lyase activity 1 mM Ca<sup>2+</sup> was used in the enzyme reaction.

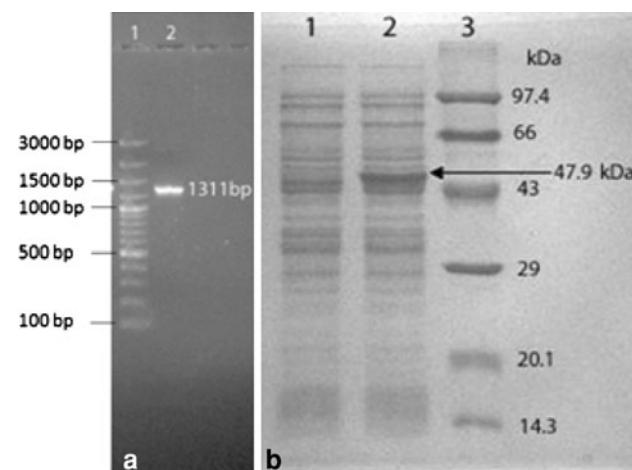
All the biochemical assays were performed in triplicate.

## Results and discussion

### Cloning and analysis of a pectinase

In the present study we cloned and characterized a pectinase gene and its product from metagenomic DNA isolated from hot springs soil. Metagenomic DNA was isolated and purified by Q-Sepharose. Purification reduced the humic acid content by 84 % making DNA suitable for use in PCR.

A PCR-based cloning strategy was used to clone a gene encoding a pectinase. After successful PCR an amplification product of 1,311 bp was obtained (Fig. 1a). The gene was sequenced and the gene sequence was submitted to GenBank (NCBI) as accession number FJ538208. Similarity analysis using BLASTn (NCBI) showed maximum 93 % identity to a gene encoding a pectinase from *Bacillus licheniformis* [Accession No.: CP000002] [28] and a putative polygalacturonase from *B. licheniformis* DSM13 [Accession No.: AE017333] [29]. BLAST results further showed 69 % [Accession No.: CP000813.1] [30], 66 % [Accession No.: AY836633] [31] and 65 % [Accession No.: CP000473] [32] sequence similarity to GH genes from *Bacillus pumilus* SAFR-032 and uncultured bacterium clone BD8804, respectively. The deduced amino acid sequence shared 93 % identity with a GH family protein (YP\_080606.1) from *B. licheniformis* ATCC14580 and a recently submitted hypothetical protein, HMPREF1012\_02989 (ZP\_08001950.1), from *Bacillus* sp. BT1B\_CT2 (isolated from an oral swab of a patient with Crohn's disease). A GH (YP\_001488197.1) from *B. pumilus* SAFR-032 showed significant similarity with a low E value of 4e<sup>-171</sup> with 96 % query coverage. Identity in this case was observed to be 66 %.



**Fig. 1** **a** PCR amplification of 1,311 bp product (lane 2), lane 1: 100 bp Plus ladder (Fermentas, Germany), **b** whole cell protein analysis of uninduced and IPTG induced samples. Protein samples were analyzed on a 12 % SDS-PAGE gel. Lane 1: uninduced sample, lane 2: 0.1 mM IPTG induced sample, lane 3: protein molecular weight marker

The gene encoded a protein of molecular weight 47.9 kDa, which was confirmed by SDS-PAGE (Fig. 1b). Molecular weight calculated by gel filtration was 33,550 kDa. The lower molecular weight observed by gel filtration could be due to a barrel-shaped structure of the protein as conformation (more appropriately hydrodynamic radius) plays an important role in elution profile. Similar variation has also been observed previously, where molecular weights according to SDS-PAGE and Gel filtration (Sephacryl S-100) were 60 and 40 kDa, respectively for a polygalacturonase from *Aspergillus kawachii* [33].

The PCR-based cloning from a soil metagenome sample resulted in isolation of a pectinase belonging to GH family 28. ClustalW alignment for deduced amino acid sequence of our newly isolated pectinase [PecJKR01<sup>His</sup>, Accession No.: FJ538208] with other pectinase sequences from GH 28 family, whose 3D structure and catalytic sites are known was carried out (Fig. 2). An endopolygalacturonase from *Calletotrichum lupini* [34], an *Aspergillus aculeatus* polygalacturonase [35] and an endopolygalacturonase II from *A. niger* [36] showed 25 % identity with PecJKR01<sup>His</sup>. An endopolygalacturonase [37] and endopolygalacturonase I from *A. niger* [38] showed 24 and 23 % identity, respectively. Sequence conservation was observed in catalytically important segments viz. 230\_NTD, 253\_DD, 287\_GHG and 320\_RIK. Conservation was invariably observed in Arg-320 and Lys-322, which are believed to be involved in substrate binding [39–41]. Conservation of Gly-293 and Tyr-355 which are conserved in all GHs was also observed. Tyr-355 is believed to be indispensable for enzyme activity [35, 41]. In addition other conserved sequences were observed (shown in black). Residues flanking the 253\_DD motif, glycine and cysteine are significant since all plant, bacterial and most fungal endopolygalacturonases have glycine N-terminal to this motif and all plant, insect and fungal PGs have cysteines immediately downstream [41]. From this we can conclude that our pectinase is most probably either of fungal or plant origin, however this is a very interesting result as the gene sequence showed maximum similarity with a bacterial gene from a *B. licheniformis*.

Sequence analyses showed some amino acids in novel positions e.g. presence of Leu-195 in place of a highly conserved proline and Val-234 instead of a strictly conserved phenylalanine. Analysis shows absence of highly conserved cysteine residues [39] in the N-terminal region at positions 24 and 42, and in the C-terminal region at positions 389, 394, 413 and 424.

#### Expression of pectinase

The ORF of the gene was inserted in the MCS of pQE30 vector and recombinant protein was purified. The effect of various IPTG concentrations showed that 0.1 mM IPTG

was optimum for the highest level of gene expression and enzyme production. Specific activity of the enzyme in lysate was determined to be 20 U/mg after induction with 0.1 mM IPTG for 3 h.

The enzyme was intracellularly expressed due to an absence of signal sequence in the protein, confirmed by analyzing the deduced polypeptide sequence using SignalP 3.0-HMM, based on neural networks and HMM.

#### Biochemical characterization

The recombinant enzyme demonstrated pH and temperature optima of 7.0 (Fig. 3a) and 70 °C (Fig. 3b), respectively, making it a highly thermoactive GH. The enzyme showed activity over a broad range of pH and temperature. The optimum temperature for the enzyme activity was observed to be 70 °C. However, the enzyme showed more than 90 % relative enzyme activity till 60 °C. Further the enzyme demonstrated more than 80 % relative enzyme activity till 50 °C. Even at extreme temperatures such as 10 and 90 °C, the enzyme showed 54 and 35 % relative enzyme activity, respectively (Fig. 3b). The enzyme was stable for 30 min till 60 °C (Fig. 3b). The enzyme retained 40.69 % enzyme activity after treatment at 70 °C for 30 min. The half-life of enzyme was determined to be 5 h at 60 °C and 23 min at 70 °C. The reduced thermostability at 70 °C could be attributed to the fact that with increasing temperature, the stability of any enzyme would decrease. Similarly enzyme activity was observed over a broad pH range of 5.0–9.0. At pH 5 and 9, 74 and 77 % relative enzyme activity was observed. The enzyme was nearly 90 % stable in the range of pH 4–11 for 1 h (Fig. 3a). To the best of our knowledge this is the first pectinolytic enzyme reported from a soil metagenome sample having optimal activity at 70 °C. Comparison with other enzymes mentioned in Table 2 shows that Polygalacturonate hydrolase [43] and PelB [40] had higher temperature optima than PecJKR01 but were active in acidic pH and in narrow pH range. The property, that made PecJKR01 special, was its high activity over a broad pH and temperature range. PelB showed optimum activity at 80 °C but demonstrated less than 10 % of the maximum value at 65 °C which was far less than the 94 % of PecJKR01 at the same temperature. As mentioned earlier PecJKR01 exhibited 55 % relative activity even at 10 °C. The EndoPG1 from the fungus *Stereum purpureum* showed 40 and 35 % relative activity at 40 and 90 °C, respectively [44]. The pH optimum of PelB was 6.4 but it showed less than 20 % relative activity at pH 5.4 and similarly at pH 7 the activity dropped rapidly. Contrary to this, the PG from *Thermoascus aurantiacus* CBMAI-756 showed less stability towards varying pH than PecJKR01 [46]. This PG also showed lower activity at pH other than optimum pH, showing 47

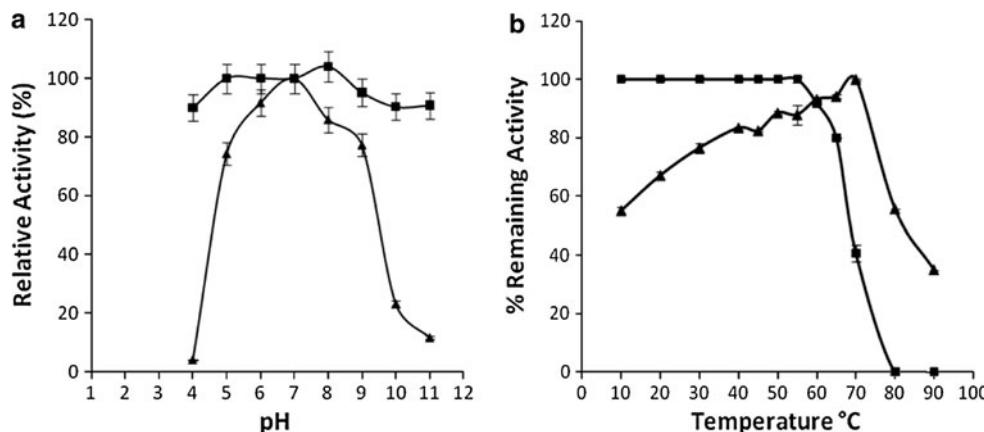
**Fig. 2** Multiple sequence alignment of PGs belonging to GH family 28 whose 3D structures are known.

Sequences:  
 endopolygalacturonase II from  
*A. niger* [PDB Id: 1CZF],  
 endopolygalacturonase G chain  
 from *C. lupini* [PDB Id: 2IQ7],  
 endopolygalacturonase I F chain  
 from *A. niger* [PDB Id: 1NHC].  
 polygalacturonase from *A. aculeatus* [PDB Id: 1IA5],  
 endopolygalacturonase A chain  
 from *Fusarium moniliforme*  
 [PDB Id: 1HG8], PecJKR01  
 from metagenome [GeneBank,  
 NCBI accession number  
 FJ538208]. Black regions show  
 complete identity while grey  
 regions partial similarity

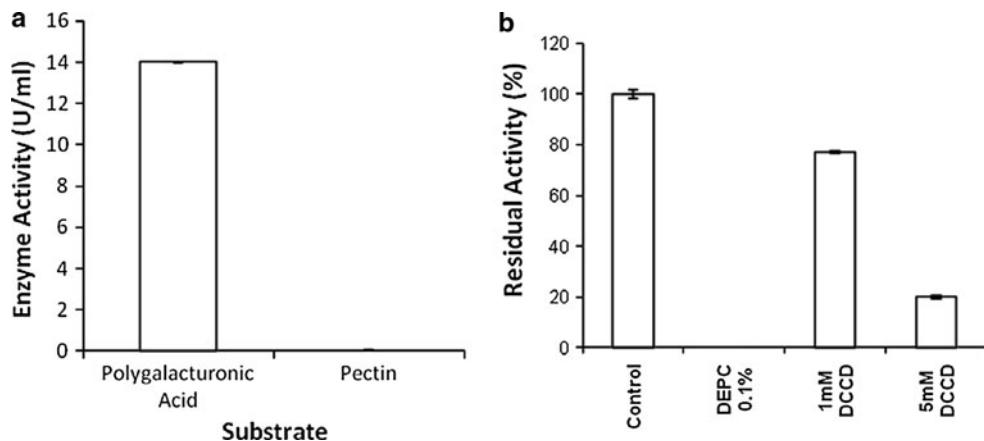
1CZF_B	MHSFASLLAYGLVAGATFASASPIEARDSCFTT--	AAAAGAKAKCST	47
2IQ7_G	--ASCTFTD--	AAAAIKGKACST	20
1NHC_F	--STCTFTS--	ASEASESISSSCSD	20
1IA5_A	--ATTCTFSGNSGASSASKSKTCST		24
1HG8_A	--DPCSVTE--	YGLATAVSSCKN	20
PecJKR01	--MSLQKIKEEILKTLKVPFPDRSFDTAFAFGADGGKKDATGAIQ		44
1CZF_B	ITLNNEIEVPAGTTLDLTG-----	LTSGTKVIFEG---	76
2IQ7_G	IILNGIVVPPACTTLDMTG-----	LKGSTTVTFQG---	49
1NHC_F	VVLSSIEVPAGETLTLDSL-----	AADGSTITFEFG---	49
1IA5_A	IVLSNVAVPSGTTLDLTK-----	LNDGTHVIFESG---	53
1HG8_A	IVLNGFQVPTGKQLDLSS-----	LQNDSVTPTKG---	49
PecJKR01	KAIDQAHKAGGGRVAVPEGVFLSGALRLKSNSVELHLAQQAVIKESQNPD		94
1CZF_B	-----TITFQEYWEGP--	LISMGSGEHITVTGASGHLINCDGARWNDK	118
2IQ7_G	-KITPGYKEWEGP--	LISFSGTNININGASGHSIDCOGSRWNDSK	91
1NHC_F	-TISEPGYKEWKGP--	LIRFGGKDLTVMADGAVIDGDGSRWNDSK	91
1IA5_A	-EITFEGYKEWSPGP--	LISVSGSDLTTTGASGHSINGDGSRWNDE	95
1HG8_A	-TITFATTADNDFN--	PIVISGSNITITGASGHVIDGNGQAYWDGK	92
PecJKR01	YLPVVLITREGVELYNYSPLIYAYEAENIAITGKGTLDQGQGDDEHWWPK		144
1CZF_B	-GTSG--KKTPKFFYAHGLDS-----	SSITGLNIKNT	147
2IQ7_G	-GSNGG-KTKPKFFYAHSLKS-----	SNIKGLNLINT	121
1NHC_F	-CTNGG-KTKPKFMYIHVED-----	STFKGINIKNT	121
1IA5_A	-GGNGG-KTKPKFFAAHSLTN-----	SVISGLKIVNS	125
1HG8_A	-GSNSNSNQKPDHFIVVQKTTG-----	NSKITNLNIQNW	125
PecJKR01	RCTNGQSSQEKDNRNALFEMAERGPVTERQFGKGHLRPNFICQPYRCKDI		194
1CZF_B	PIMAFSVQ-ANDITFTDVTINNADGDTQ-----	GGHTNTDAFDVGNSV	188
2IQ7_G	PVQAQSINSATTLGVDVIIIDSAGDSA-----	GGHTNTDAFDVGSS	163
1NHC_F	PVQAISVQ-ATNVHLNDFNTIDSAGDDN-----	GGHTNTDAFDISEST	162
1IA5_A	PVQVFSVAGSDYLTLDKDTITIDSAGDDN-----	GGHTNTDAFDIGST	167
1HG8_A	PVHCFDITGSSQLTISGLIILDRAGDKPNAKSGSLPAAHNTDGFDISSSD		175
PecJKR01	LIQGPVLNNSPMWQVHPVLCEBNTVDGIKVIG---HGPTNDGVNPESCK		240
1CZF_B	GVNI IKPWVHNQDDCLAVNSG-----	ENIWF TGGTC IGGHG-L	225
2IQ7_G	GVYI SGANVKNQDDCLAINSG-----	TNITFTGGTC SGGHG-L	200
1NHC_F	GVYI SGATVKNQDDCIAINSG-----	ESIISFTGGTC SGGHG-L	199
1IA5_A	YVTISGATVYNQDDCQAVANSG-----	ENIYFSGGYC SGGHG-L	204
1HG8_A	HVTLDNNHVYNQDDCQAVATSG-----	TNIVVSNMYC SGGHG-L	212
PecJKR01	NVVIKGCHFDSDGDCIAVKSGRNADARRINMPSENIVIEHNEKDGHGV		290
1CZF_B	**	*	*
2IQ7_G	SIG-SVGDRSNNVVKNVTIEHTVSNSENAVRIKTIISGATGSVSEITYSN		274
1NHC_F	SIG-SVGGRSNDTNTVKTFTISNSKIVNSDNGVRIKTVSGATGSVSGTVTSG		249
1IA5_A	SIG-SVGGRRDNTVKNVTISDSTSNSANGVRIKTIYKETGDVSEITYSN		248
1HG8_A	SIG-SVGGRSNDTNTVKTFTDSTIINSDNGVRIKTNIDTGSVSDVTYKD		253
PecJKR01	SIG-SVGGKSDNVVDGVQFLSQVNVNSQNGCRIKNSGATGTINNVITYQN		261
1CZF_B	TIGSEISGGVKNVIAEGNIMDS--PNLDRALRIKINSVRGGVLENIYPHK		338
2IQ7_G	IVMSGISDYGVIQDYEDGKPTGKPTNGVTIQDVKLESVTGSVDGATE		324
1NHC_F	ITLSNIAKYGVIEQDYENGSPGTGPTNGVPIITGLTLSKTTGSVASSGTN		299
1IA5_A	IQLSGITDYGIVIEQDYENGSPGTGPTSTGIPIITDVTVDGVTGTLEDDAQ		298
1HG8_A	ITLTSNIAKYGVIVQONY--GDTSSPTTGVPIITDVFVLDNVHGSVSSGTN		301
PecJKR01	IALTNISTYGVDVQDYLNNGGPTGKPTNGVKISNIKFIVTGTVAASSAQD		311
1CZF_B	NTVKSLSKREVIAIDMEYEFGDAGDFKPVVRNIDVEQLKSMGGQYG---I		384
2IQ7_G	IYLLCGSGSCSDWTWDDVKVTCGKKSTACKNFP--SVASC-----		362
1NHC_F	YVILLCASGACSNWKWSGVSVTGCCKSTKCSNIPSGSGAAC-----		339
1IA5_A	YVILCGDGSCSDWTWSGVDLSCGKTSDKCENVP--SGASC-----		336
1HG8_A	ILIISCGSGSCSDWTWTDVSVCGGTSSKCTNVP--SGASC-----		339
PecJKR01	WFILCGDGSCSGFTFSGNAITCGGKTCSSC-NYPTNTCP-----		349
1CZF_B	RVLAYDHS PVTGLKVTDSEIDCVDPIMELKHVKDPVFSNLINGKRYDSH		434
2IQ7_G	--		
1NHC_F	--		
1IA5_A	--		
1HG8_A	--		
PecJKR01	KA 436		

\*-Conserved catalytic site

**Fig. 3** **a** Effect of varying pH (filled triangle) on enzyme activity using 0.5 % PGA as substrate, in 50 mM Acetate buffer (pH 4 and 5), 50 mM Phosphate buffer (pH 6, 7 and 8), 50 mM Tris–Cl (pH 9), and 50 mM Glycine–NaOH buffer (pH 10 and 11). pH stability analysis (filled square), **b** effect of temperature on activity (filled triangle) and thermostability of enzyme (filled square)



**Fig. 4** **a** Effect of different substrates on enzyme activity. Enzyme activity checked using PGA and 9 % methyl esterified pectin. While activity of 14 U/ml was observed in case of PGA as substrate, no activity was observed in case of pectin as substrate, **b** effect of chemical modifiers on enzyme activity



and 10 % relative activity at 4.0 and 6.5, respectively in contrast to 75 and 77 % relative activity of PecJKR01 at 5.0 and 9.0, respectively. The EndoPG1 showed optimum activity in the pH range of 3.5–4.5. This enzyme demonstrated relative activity of 40 % at pH of 3.5 and 5.5. Similarly, APGase also showed very less activity at pH other than its optimum (10.0) [45]. The relative activities in this case were 50 and 60 % at 9.0 and 11.5, respectively. The PGase M from *Kluyveromyces marxianus* demonstrated highest activity in acidic pH at 5.0 [48]. The PGase showed 10 % relative activity at pH 3.0 and no activity at pH 8.0. The half-life of endo PG from *Penicillium capsulatum* was 3.8 min which was far less than 5 h of PecJKR01 at 60 °C [47].

In contrast to pectate lyases which have an absolute requirement for  $\text{Ca}^{2+}$  [49], PecJKR01 did not require  $\text{Ca}^{2+}$  for its activity. Therefore we inferred that PecJKR01 was a polygalacturonase. This was confirmed when enzyme activity was assayed on 9 % methyl esterified pectin substrate, when the enzyme showed no activity. Some enzymes had been found to show activity on both PGA and pectin as a substrate [49, 50] but PecJKR01 showed only polygalacturonase activity (Fig. 4a).

Generally the catalytic site for a particular family of enzymes was conserved. Therefore we used this rationale to predict the catalytic site in our cloned enzyme. Protein sequences were retrieved from RCSB Protein Data Bank ([www.rcsb.org](http://www.rcsb.org)). The conserved catalytic site was predicted using multiple sequence alignment and was found to be Asp-232, Asp-253, Asp-254, and His-287. These were further confirmed by using chemical modifiers for aspartate and histidine residues. Enzyme activity was completely inhibited by 0.1 % DEPC (a modifier of histidine residues), and reduced to 77 and 19 % by 1 and 5 mM DCCD, (a modifier of carboxylic acid residues) (Fig. 4b).

Most of the enzymes demonstrating a high similarity with this enzyme at the amino acid sequence level showed a high content of cysteine residues (8 residues). Two conserved cysteine residues in the N-terminus region and four conserved cysteine residues in the C-terminus region were absent in PecJKR01. Only one conserved cysteine residue in the 253\_DDC conserved motif was present in PecJKR01. In PecJKR01 a total of five cysteine residues were present between which disulfide bonds could be formed. It had been observed that when conserved cysteines were not present, they were generally replaced by

**Table 1** Effect of varying concentrations of  $\beta$ -ME on enzyme activity

Samples	Remaining activity (%)
Control	100 $\pm$ 7.32
0.1 mM	96 $\pm$ 0.45
1 mM	94 $\pm$ 0.25
10 mM	71 $\pm$ 1.30
100 mM	36 $\pm$ 2.4

hydrophobic residues [51]. In the case of PecJKR01 also cysteines were replaced with hydrophobic amino acids viz. phenylalanine, alanine and methionine. However, DISULFINDE [52–54] predicted no disulfide bonds in this protein while two probable disulphide bonds were predicted with DiANNA [55]. Moreover when different concentrations of  $\beta$ -mercaptoethanol were added to the reaction mix little inhibition in enzyme activity was observed with low concentration of  $\beta$ -mercaptoethanol (Table 1) apart from at 100 mM where notable inhibition was recorded. Conversely, Kaur et al. [56] observed nearly 51 and 42 % relative activity after treatment with 1 and 5 mM  $\beta$ -mercaptoethanol, respectively. Therefore this result justifies the prediction by the DiANNA bioinformatics tool. It was reported that highly conserved disulphide bonds play a crucial role in acquiring the appropriate folding state or for maintaining the active site conformation while other non-conserved disulphide bonds had a protein specific role [51]. In this case it might be speculated that these disulphide bonds are not important for maintaining the active site but might be important to maintain overall protein structure. On the other hand cysteine residues provided thermostability not only because of disulfide linkage but also due to higher hydrophobic effect [57]. The

latter fact could also be responsible for the high thermostability of PecJKR01.

The 3D structure was predicted using 3JUR (PDB Id) as a template which shared 48.5 % identity with PecJKR01 [58]. Parallel  $\beta$ -helix structural domains characteristic for GH family 28 members were observed (supplementary material). The secondary structure revealed 36 % extended  $\beta$  sheets, 59.4 % loops and the remainder being helical. Structural analysis showed that all the catalytically important residues viz. Asp-232, Asp-253, Asp-254, His-287, Arg-320, and Lys-322 were present on the surface and in close proximity to each other. Amino acid analysis showed the presence of 45 glycine residues, which account for the thermophilic nature of the enzyme because of the high conformational entropy of glycine [59]. Amino acid sequence also showed the presence of 21 proline residues which are frequently present in loop regions of thermostable enzymes [60]. Furthermore the high percentage of Glu, Lys, Ile, Val and Leu might account for the higher thermostability of the enzyme. The enzyme showed behavior which is contradictory to convention defined earlier [61]. The basic amino acid ratio Arg/Arg + Lys of this enzyme (0.38) is unexpectedly lower than that of thermophiles and even lower than that of mesophiles (0.40). This suggests that PecJKR01 is an exception and shows contradictory behavior from previously accepted conventions for thermostability.

In conclusion, PecJKR01 enzyme demonstrates unusual molecular characteristics that resulted in novel biochemical properties such as high enzyme activity over a broad pH and temperature range. It demonstrated high thermostability at 60 °C. Therefore the importance of this enzyme in industrial processes and for studies to elucidate structure/function relationships and to design inhibitors is warranted.

**Table 2** Comparison of PecJKR01 with thermostable PGs

Source	Enzyme name	Protein size (kDa)	pH optimum	Temperature optimum (°C)	Reference
<i>Thermotoga maritima</i>	PelB	50	6.4	80	[40]
<i>Thermoanaerobacterium thermosulfurogenes</i>	Polygalacturonate hydrolase	320	5.5	75	[43]
Metagenome	PecJKR01	47.9	7.0	70	This study
<i>S. purpureum</i>	Endo PG1	40	3.5–4.5	70	[44]
<i>Bacillus</i> sp.	APGase	60	10	65	[45]
<i>T. aurantiacus</i> CBMAI-756	PG	30	5.5	65	[46]
<i>P. capsulatum</i>	Endo PG	51.6	4.7	62	[47]
<i>K. marxianus</i>	PGase M	44	5	60	[48]

Information about other enzymes was obtained from BRENDA—The comprehensive enzyme information system [42]

**Acknowledgments** Fellowship granted to RS by University Grants Commission (UGC), India is acknowledged.

## References

- Yadav S, Yadav PK, Yadav D, Yadav KDS (2009) Pectin lyase: a review. *Process Biochem* 44:1–10
- Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 280:309–316
- Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The carbohydrate-active enzymes database (CAZy), an expert resource for glycogenomics. *Nucleic Acids Res* 37:D233–D238
- Brummell DA, Harpster MH (2001) Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Mol Biol* 47:311–340
- Abbott DW, Boraston AB (2008) Structural biology of pectin degradation by *Enterobacteriaceae*. *Microbiol Mol Biol Rev* 72: 301–316
- Herron SR, Benen JAE, Scavetta RD, Visser J, Jurnak F (2000) Structure and function of pectic enzymes, virulence factors of plant pathogens. *Proc Natl Acad Sci USA* 97:8762–8769
- Jyothi TC, Singh SA, Appu Rao AG (2005) The contribution of ionic interactions to the conformational stability and function of polygalacturonase from *A. niger*. *Int J Biol Macromol* 36:310–317
- Asif MH, Nath P (2005) Expression of multiple forms of polygalacturonase gene during ripening in banana fruit. *Plant Physiol Biochem* 43:177–184
- Levin L, Forchiassin F (1998) Culture conditions for the production of pectinolytic enzymes by the white-rot fungus *Trametes trogii* on a laboratory scale. *Acta Biotechnol* 18:157–166
- Dosanjh NS, Hoondal GS (1996) Production of constitutive, thermostable, hyperactive exo-pectinase from *Bacillus* GK-8. *Biotechnol Lett* 18:1435–1438
- Hadj-Taieb N, Ayadi M, Khelif M, Mrad K, Hassairi I, Gargouri A (2006) Fermentor production of pectinases on gruel, a local by-product and their use in olive oil extraction. *Enzyme Microb Technol* 39:1072–1076
- Kashyap DR, Vohra PK, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector, a review. *Bioresour Technol* 77:215–227
- Courtois S, Cappellano CM, Ball M, Francou FX, Normand P, Helyne G, Martinez A, Kolvek SJ, Hopke J, Osburne MS, August PR, Nalin R, Guerineau M, Jeannin P, Simonet P, Pernodet JL (2003) Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl Environ Microbiol* 69:49–55
- Diaz-Torres ML, McNab R, Spratt DA, Villedieu A, Hunt N, Wilson M, Mullany P (2003) Novel tetracycline resistance determinant from the oral metagenome. *Antimicrob Agents Chemother* 47:1430–1432
- Majernik A, Gottschalk G, Daniel R (2001) Screening of environmental DNA libraries for the presence of genes conferring  $\text{Na}^+(\text{Li}^+)/\text{H}^+$  antiporter activity on *Escherichia coli*, characterization of the recovered genes and the corresponding gene products. *J Bacteriol* 183:6645–6653
- Schipper C, Hornung C, Bijtenhoorn P, Quitschau M, Grond S, Streit WR (2009) Metagenome-derived clones encoding two novel lactonase family proteins involved in biofilm inhibition in *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 75:224–233
- Li G, Wang K, Liu YH (2008) Molecular cloning and characterization of a novel pyrethroid-hydrolyzing esterase originating from the metagenome. *Microb Cell Factories* 7:38
- Lee MH, Lee CH, Oh TK, Song JK, Yoon JH (2006) Isolation and characterization of a novel lipase from a metagenomic library of tidal flat sediments, evidence for a new family of bacterial lipases. *Appl Environ Microbiol* 72:7406–7409
- Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* 62:316–322
- Sharma PK, Capalash N, Kaur J (2007) An improved method for single step purification of metagenomic DNA. *Mol Biotechnol* 36:61–63
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York
- Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ (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* 22:4673–4680
- Bendtsen Dyrlov J, Henrik N, von Gunnar H, Soren B (2004) Improved prediction of signal peptides—SignalP 3.0. *J Mol Biol* 340:783–795
- Nielsen H, Krogh A (1998) Prediction of signal peptides and signal anchors by a hidden markov model. In: Proceedings of the sixth international conference on intelligent systems for molecular biology (ISMB6). AAAI Press, Menlo Park, pp 122–130
- Lambert C, Leonard N, De Bolle X, Depiereux E (2002) ESyPred3D: prediction of proteins 3D structures. *Bioinformatics* 18:1250–1256
- Singh K (2007) Cloning and characterization of gene(s) involved in catechin biosynthesis in tea (*Camellia sinensis*). PhD Thesis, Guru Nanak Dev University, Amritsar
- Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Chem* 31:426–428
- Rey MW, Ramaiya P, Nelson BA, Brody-Karpin SD, Zaretsky EJ, Tang M, de Leon AL, Xiang H, Gusti V, Clausen IG, Olsen PB, Rasmussen MD, Andersen JT, Jorgensen PL, Larsen TS, Sorokin A, Bolotin A, Lapidus A, Galleron N, Ehrlich SD, Berka RM (2004) Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species. *Genome Biol* 5(Suppl 10):R77
- Veith B, Herzberg C, Steckel S, Feesche J, Maurer KH, Ehrenreich P, Baeumer S, Henne A, Liesegang H, Merkl R, Ehrenreich A, Gottschalk G (2004) The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential. *J Mol Microbiol Biotechnol* 7:204–211
- Gioia J, Yerrapragada S, Qin X, Jiang H, Igboeli OC, Muzny D, Dugan-Rocha S, Ding Y, Hawes A, Liu W, Perez L, Kovar C, Dinh H, Lee S, Nazareth L, Blyth P, Holder M, Buhay C, Tirumalai MR, Liu Y, Dasgupta I, Bokhetache L, Fujita M, Karouia F, Eswara PM, Siebert J, Uzman A, Buzumbo P, Verma A, Zwiya H, McWilliams BD, Olowu A, Clinkenbeard KD, Newcombe D, Golebiewski L, Petrosino JF, Nicholson WL, Fox GE, Venkateswaran K, Highlander SK, Weinstock GM (2007) Paradoxical DNA repair and peroxide resistance gene conservation in *Bacillus pumilus* SAFR-032. *PLoS One* 2:E928
- Solbak AI, Richardson TH, McCann RT, Kline KA, Bartnek F, Tomlinson G, Tan X, Parra-Gessert L, Frey GJ, Podar M, Luginbuhl P, Gray KA, Mathur EJ, Robertson DE, Burk MJ, Hazlewood GP, Short JM, Kerovuo J (2005) Discovery of pectin-degrading enzymes and directed evolution of a novel pectate lyase for processing cotton fabric. *J Biol Chem* 280:9431–9438
- Copeland A, Lucas S, Lapidus A, Barry K, Detter JC, Glavina del Rio T, Hammon N, Israeni S, Dalin E, Tice H, Pitluck S, Thompson LS, Brettin T, Bruce D, Han C, Tapia R, Gilna P, Schmutz J, Larimer F, Land M, Hauser L, Kyrides N, Mikhailova N, Janssen PH, Kuske CR, Richardson P (2006) Complete sequence of *Solibacter usitatus* Ellin6076, accession

- CP000473. US DOE Joint Genome Institute, Walnut Creek, p 1698
33. Esquivel JCC, Voget CE (2004) Purification and partial characterization of an acidic polygalacturonase from *Aspergillus kawachii*. *J Biotechnol* 110:21–28
34. Bonivento D, Pontiggia D, Di Matteo A, Fernandez-Recio J, Salvi G, Tsernoglou D, Cervone F, De Lorenzo G, Federici L (2008) Crystal structure of the endopolygalacturonase from the phytopathogenic fungus *Colletotrichum lupini* and its interaction with polygalacturonase-inhibiting proteins. *Proteins* 70:294–299
35. Cho SW, Lee S, Shin W (2001) The X-ray structure of *Aspergillus aculeatus* polygalacturonase and a modeled structure of the polygalacturonase-octagalacturonate complex. *J Mol Biol* 311: 863–878
36. van Santen Y, Jacques AEB, Schroter KH, Kalk KH, Armand S, Visser J, Dijkstra BW (1999) 1.68-Å crystal structure of endopolygalacturonase II from *Aspergillus niger* and identification of active site residues by site-directed mutagenesis. *J Biol Chem* 274:30474–30480
37. Federici L, Caprari C, Mattei B, Savino C, Di Matteo A, De Lorenzo G, Cervone F, Tsernoglou D (2001) Structural requirements of endopolygalacturonase for the interaction with PGIP (polygalacturonase-inhibiting protein). *Proc Natl Acad Sci USA* 98:13425
38. van Pouderoyen G, Snijder HJ, Benen JA, Dijkstra BW (2003) Structural insights into the processivity of endopolygalacturonase I from *Aspergillus niger*. *FEBS Lett* 554:462–466
39. Martens-Uzunova ES, Zandlevan JS, Benen JAE, Awad H, Kools HJ, Beldman G, Voragen AGJ, Van den Berg JA, Schaap PJ (2006) A new group of exo-acting family 28 glycoside hydrolases of *Aspergillus niger* that are involved in pectin degradation. *Biochem J* 400:43–52
40. Kluskens LD, Gert-Jan van Alebeek WM, Walther J, Voragen AGJ, de Vos WM, van der Oost J (2005) Characterization and mode of action of an exopolygalacturonase from the hyperthermophilic bacterium *Thermotoga maritima*. *FEBS J* 272:5 464–5473
41. Markovic O, Stefan J (2001) Pectin degrading glycoside hydrolases of family 28, sequence—structural features, specificities and evolution. *Protein Eng* 14:615–631
42. Chang A, Scheer M, Grote A, Schomburg I, Schomburg D (2009) BRENDa, AMENDA and FRENDa the enzyme information system, new content and tools in 2009. *Nucleic Acids Res* 37: D588–D592. <http://www.brenda-enzymes.info/index>
43. Bernhard S, Zeikus JG (1983) Characterization of pectinolytic enzymes of *Clostridium thermosulfurogenes*. *FEMS Microbiol Lett* 17:295–298
44. Miyairi K, Okuno T, Sawai K (1985) Purification and properties of endopolygalacturonase I from *Stereum purpureum*, a factor inducing silver-leaf symptoms on apple trees. *Agric Biol Chem* 49(Suppl 4):1111–1118
45. Horikoshi K (1972) Production of alkaline enzymes by alkophilic microorganisms. *Agric Biol Chem* 36:285–293
46. Martins ES, Silva D, da Silva R, Leite RS, Gomes E (2007) Purification and characterization of polygalacturonase produced by thermophilic *Thermoascus aurantiacus* CBMAI-756 in submerged fermentation. *Antonie Van Leeuwenhoek* 91(Suppl 3):291–299
47. Gillespie AM, Cook K, Coughlan MP (1990) Characterization of an endopolygalacturonase produced by solid-state cultures of the aerobic fungus *Penicillium capsulatum*. *J Biotechnol* 17:279–292
48. Hirose N, Kishida M, Kawasaki H, Sakai T (1999) Purification and characterization of an endo-polygalacturonase from a mutant of *Saccharomyces cerevisiae*. *Biosci Biotechnol Biochem* 63: 1100–1103
49. Soriano M, Diaz P, Pastor FIJ (2005) Pectinolytic systems of two aerobic sporogenous bacterial strains with high activity on pectin. *Curr Microbiol* 50:114–118
50. Le Van T, Tuyen H, Helmke E, Le Tran B, Schweder T (2001) Cloning of two pectate lyase genes from the marine Antarctic bacterium *Pseudoalteromonas haloplanktis* strain ANT/505 and characterization of the enzymes. *Extremophiles* 5:35–44
51. Thangudu RR, Malini M, Srinivasan N, Cadet F, Sowdhamini R, Offmann B (2008) Analysis on conservation of disulphide bonds and their structural features in homologous protein domain families. *BMC Struct Biol* 8:55
52. Vullo A, Frasconi P (2004) Disulfide connectivity prediction using recursive neural networks and evolutionary information. *Bioinformatics* 20:653–659
53. Frasconi P, Passerini A, Vullo A (2002) A two-stage SVM architecture for predicting the disulfide bonding state of cysteines. In: Proceedings of IEEE workshop on neural networks for signal processing pp 25–34
54. Ceroni A, Frasconi P, Passerini A, Vullo A (2003) Predicting the disulfide bonding state of cysteines with combinations of kernel machines. *J VLSI Signal Process* 35:287–295
55. Ferre F, Clote P (2005) DiANNA: a web server for disulfide connectivity prediction. *Nucleic Acids Res* 33(web server issue):W230–W232
56. Kaur G, Kumar S, Satyanarayana T (2004) Production, characterization and application of a thermostable polygalacturonase of a thermophilic mould *Sporotrichum thermophile* Apinis. *Bioreactor Technol* 94:239–243
57. You C, Huang Q, Xue H, Xu Y, Lu H (2010) Potential hydrophobic interaction between two cysteines in interior hydrophobic region improves thermostability of a family 11 xylanase from *Neocallimastix patriciarum*. *Biotechnol Bioeng* 105:861–870
58. Pijning T, van Pouderoyen G, Kluskens L, van der Oost J, Dijkstra BW (2009) The crystal structure of a hyperthermoactive exopolygalacturonase from *Thermotoga maritima* reveals a unique tetramer. *FEBS Lett* 583:3665–3670
59. Vieille C, Burdette SD, Zeikus GJ (1996) Thermozyymes. *Biotechnol Annu Rev* 2:1–83
60. Watanabe K, Chishiro K, Kitamura K, Suzuki Y (1991) Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermoglucoamylase* KP1006. *J Biol Chem* 266:24287–24294
61. Nawani N, Kaur J (2007) Studies on lipolytic isoenzymes from a thermophilic *Bacillus* sp., production, purification and biochemical characterization. *Enzyme Microb Technol* 40:881–887