Molecular studies on protein- and carbohydrate-converting enzymes from thermophilic bacteria

Leon D. Kluskens

Promotor	**prof. dr. W.M. de Vos** Hoogleraar in de Microbiologie Wageningen Universiteit
Co-promotor	**dr. J. van der Oost** Universitair docent bij de Leerstoelgroep Microbiologie Wageningen Universiteit
Leden van de promotie commissie	**prof. dr. dr. h.c. G. Antranikian** Technical University Hamburg-Harburg, Germany
	prof. dr. J.A.M. Leunissen Wageningen Universiteit
	prof. dr. R.J. Siezen Katholieke Universiteit Nijmegen
	prof. dr. A.G.J. Voragen Wageningen Universiteit
	dr. ir. G. Beldman Wageningen Universiteit

Molecular studies on protein- and carbohydrate-converting enzymes from thermophilic bacteria

Leon D. Kluskens

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
prof. dr. ir. L. Speelman,
in het openbaar te verdedigen
op vrijdag 9 januari 2004
des namiddags te vier uur in de Aula

Molecular studies on protein- and carbohydrate-converting enzymes from thermophilic bacteria
Leon D. Kluskens

In Dutch: Moleculaire studie naar eiwit- en suikerafbrekende enzymen uit thermofiele bacteriën

Ph.D. thesis Wageningen, Wageningen, The Netherlands 2004
138 p., with summary in Dutch

ISBN: 90-5808-971-1

Voor mijn ouders

BEDANKT!

Met het schrijven van dit dankwoord kan ik eindelijk beginnen aan het laatste, maar ook belangrijkste gedeelte van mijn proefschrift. Want zonder de hulp van een groot aantal mensen zou de verwezenlijking van dit boekje een stuk moeizamer, zoniet onmogelijk zijn geweest, en die personen wil hieronder graag bedanken. John en Willem, het tweejarig project kwam voor mij als een geschenk uit de hemel, aangezien ik nog niet wist wat ik wilde gaan doen, en omdat de stap naar een volledig AIO-project op dat moment nog zo groot leek.

John, als copromotor, bedankt voor het vertrouwen dat je in me uitsprak door het contract een jaar later 'gewoon' te verlengen tot 4 jaar. Grotendeels dankzij de motiverende discussies en je relativeringsvermogen werden tegenslagen tijdens het onderzoek steeds weer overwonnen.

Willem, als promotor, ondanks je drukke agenda was er altijd wel een afspraak te maken. Gesprekken met jou waren enorm prikkelend en dankzij jou werden de onderzoeksdoelen nooit uit het oog verloren.

Paranimfen en vrienden Frank en Helmut, bedankt voor jullie steun. Frank, omdat je tijdens je onderzoek, het verhuizen, en het grootbrengen van je pasgeboren dochter nog tijd vond om mij te helpen met de vormgeving van dit boekje en al het andere wat er bij promoveren komt kijken. Helmut, bedankt voor het dienen als elektronische Amsterdamse praatpaal wanneer de onderzoeksfrustraties om een relativerende of ironische oneliner vroegen.

Thijs Kaper, al vanaf 1995 kruisen onze wegen regelmatig, via afstudeervak en promotieonderzoek bij Bacteriële Genetica, tot nu in Groningen. Ontzettend bedankt voor het tijdelijk onderdak, maar ook voor de Murmeltiere- en elandenjacht, het ontbrekende bergverzet, de vele whisky's en die ene Jägermeister.

Alle 'BacGenootjes' bedankt: Alexandra, Ana, als Sister-in-arms tijdens de schrijfuurtjes in de avond; Ans, voor de vele etentjes, lab-aria's en al die kleine dingen die het labwerk aangenamer maakten; Arjen, voor de kritische discussies over Bassie en Adriaan, het verschil tussen enkel- en dubbelgebeid en de gedeelde passie voor luchtgitaarmuziek; non-canonical Corné, dank voor de ontelbare practical jokes; Don, thanks for the enjoyable bike tours and your unlimited desire to get at least one drink; Gaël, Hauke, dank voor het zeer toffe weekendje Seattle; Ineke, voor de aanstekelijke golvende lachsalvo's; Johan, ofwel de Beul van Bakel, voor de geselingen tijdens de Brug-tot-Brug tochtjes; Joyce, voor de vezelrijke discussies aan de lunchtafel; Judith, Ken, Krisztina, Nina, Odette, Pino, thanks for the hospitality during the Extremophiles meeting in Naples; Ronnie, Servé, kamergenoten Stan en Thijs Ettema (in één adem), dank voor het beheren van de Sudden-death schatkist, het triestighedenbord, de vlijmscherpe, rake en oergeestige toewensingen en het bijhouden van de thermofielengenoomsequentieteller; Valerie en Wilfried. Allen bedankt voor de talrijke 'buitenschoolse activiteiten', (post-)congres uitstapjes, kartwedstrijden, waaierkoersen, broodnodige lunchgesprekken, cafésessies en uiteraard ook de wetenschappelijke discussies.

Wim, ontzettend bedankt voor het bij-de-tijd houden van mijn computer en voor het opleuken van mijn laptop. Ria (door geen koffieautomaat te vervangen), Nees, Francis, en Renée, bedankt voor het orde scheppen in alle administratieve bijkomstigheden. Iedereen die werkzaam is op het Laboratorium voor Microbiologie en die enigerwijze heeft bijgedragen tot dit proefschrift, van harte bedankt.

Een aantal studenten hebben in de afgelopen jaren een wezenlijke bijdrage geleverd aan mijn onderzoek. Teun, Fokko, Jasper en Jurrit, zeer bedankt voor jullie inzet.

De contacten zijn niet beperkt gebleven tot het Laboratorium voor Microbiologie alleen. Gert-Jan, de beslissing om met jou samen te gaan werken is van wezenlijk belang geweest voor mijn onderzoek. Ik heb veel opgestoken van je doelgerichte aanpak en kennis van zaken op het gebied van pectinolytische enzymen, en dit is maximaal ten goede gekomen van twee hoofdstukken. Professor Antranikian and Ruth Schwerdtfeger, thank you both for the nice collaboration and your contribution to the characterization of the keratinase; Jeong-Sun Kim, thank you for its successful crystallization.

I am very grateful to Professor Kelly and the people from the Hyperthermophile Group. Thanks to you my stay in Raleigh has been both fruitful and enjoyable. Don, many thanks for letting me stay at your place, not to mention for getting me hooked on ice hockey and Krispy Kremes. I'm sure we'll meet again somewhere.

Mijn (voormalig) huisgenoten van Asterstraat 205, Always-vrienden Ingrid, Henriette en Marc als bewoners van het eerste uur, en verder Lea, Sander, Thijs E., en Krisztina, hartelijk dank voor de nodige uurtjes ontspanning naast mijn werk. Erbijblijfbijen, ontstaan in Wageningen en inmiddels bijna allemaal uitgevlogen, bedankt voor de afleiding en de blijvende contacten.

Mijn huidige werkgever BioMaDe en in het bijzonder Gert Moll wil ik bedanken voor het in de gelegenheid stellen mijn proefschrift af te kunnen ronden.

Alcina docinha, agradeço-te imenso pelo teu apoio moral, pelos teus conselhos e pelo teu carinho embora por vezes tivesse que ser à distância. Dentro em breve vamos estar finalmente juntos.

Tot slot wil ik mijn ouders en zus Marion bedanken voor de morele steun en interesse die jullie altijd getoond hebben, ook al was het soms moeilijk om precies uit te leggen wat ik allemaal aan het doen was. Pap en mam, bedankt voor alle mogelijkheden die ik dankzij jullie heb kunnen benutten: Dit boekje is voor jullie.

TABLE OF CONTENTS

1

General Introduction

Life at high temperatures

Life as we know it has been observed in all existing ecological niches, including extreme conditions, such as high and low temperatures (thermophiles vs. psychrophiles), acidic and alkaline pH conditions (acidophiles vs. alkaliphiles), high pressure (barophiles), and high salt concentrations (halophiles)(Table 1.1). Of all these so-called extremophiles, the organisms isolated most frequently and studied most extensively over the last two decades, are the thermophilic microorganisms. Whereas certain animals and plants can cope with temperatures up to 50°C, some prokaryotes can exhibit optimal growth near the boiling point of water, or even beyond. In particular, the group of thermophilic and hyperthermophilic microorganisms has caught the attention of many scientists. In general, thermophiles have their temperature optimum between 50 and 80°C, whereas the hyperthermophiles exhibit growth beyond 80°C (121)(Table 1.1). Currently topping the list is the crenarchaeon "Strain 121", most closely related to *Pyrodictium occultum* and *Pyrobaculum aerophilum*, which is able to grow at 121°C with doubling time of 24 hours (52).

For a long time the subdivision of organisms was based on their cellular architecture, which resulted in two domains: prokarya and eukarya (20). This traditional concept came to an end with the discovery of a third domain by Woese and coworkers (126). Based on comparison of 16S/18S rRNA sequences, present in all living organisms, the prokaryotes were subdivided in bacteria and archaea. It is mainly this last domain that concludes the extremophiles; with the exception of the bacterial orders of *Thermotogales*, and *Aquificales*, all thermophilic and hyperthermophilic microorganisms belong to the Archaea. Together they form the deepest branches in the phylogenetic tree of all three domains (126), and this division has lead to the hypothesis that a heat-loving organism stood at the basis of the origin and early evolution of life (31, 32, 111, 112). This theory however has been contested by other researchers (48, 86, 127).

Table 1.1 Extremophiles and their optimal growth conditions

Classification	Optimal growth conditions
Hyperthermophiles	80-121°C
Thermophiles	55-80°C
Mesophiles	20-55°C
Psychrophiles	-2-20°C
Halophiles	2-5 M NaCl
Acidophiles	pH < 4
Alkaliphiles	pH > 9

adapted from Hough & Danson (44) and Stetter (110)

Thermophiles and hyperthermophiles have been isolated from many natural habitats, including continental mud-pools (Solfatara), hot springs and geysers (Yellowstone National Park), and deep-sea sediments or vents, such as black smokers (Mid-Atlantic Ridge). In addition, they have been isolated in heated industrial environments, like the outflows of geothermal power plants and sewage sludge systems (121).

Most of the thermophiles and hyperthermophiles are chemolithoautotrophs and fix CO_2 by chemosynthesis. Besides H_2 as an important electron donor, sulfide, sulfur and ferrous iron can be used. The majority grows anaerobically, although some of them are facultative, or even strictly aerobic, yet at reduced oxygen concentrations (110). Several thermophiles and hyperthermophiles are facultative heterotrophs, being able to switch from autotrophic to heterotrophic growth when organic material is presented (110). Most of them are well equipped to degrade peptides, and many are even capable of growing on a variety of simple and complex carbohydrates (Table 1.2) (26, 58). Proteinaceous growth involves intracellular, extracellular and cell wall-associated proteases. Comparison at a genomic level showed that, although there is considerable conservation of certain proteases across all domains of life, some proteases appear to be distinctive for thermophiles, many of which still need characterization (124). The preference for peptides instead of single amino acids as carbon source, observed for some hyperthermophiles, may suggest a difference in peptide catabolism with mesophiles (7). Many hyperthermophiles grow well on both α- and β-linked carbohydrates (110, 111). Dependent on their habitat, it is provided as plant material (starch, cellulose, xylan, pectin), or part of animal cells or other microorganisms (glycogen). Like for peptides, carbohydrates are preferably used as di- or polysaccharides, because of the thermal lability of monosaccharides, as was observed in the fermentation of glucose versus cellobiose by *Pyrococcus furiosus* (33). Regarding their classification on amino acid level in conjunction with their mesophilic counterparts (http://afmb.cnrs-mrs.fr/~cazy/CAZY/, see below), thermostable glycoside hydrolases do not appear to be unique or uncommon. However, family 57 appears to have a higher number of amylolytic enzymes originating from thermophiles, mostly archaea, indicating some ancestral connection.

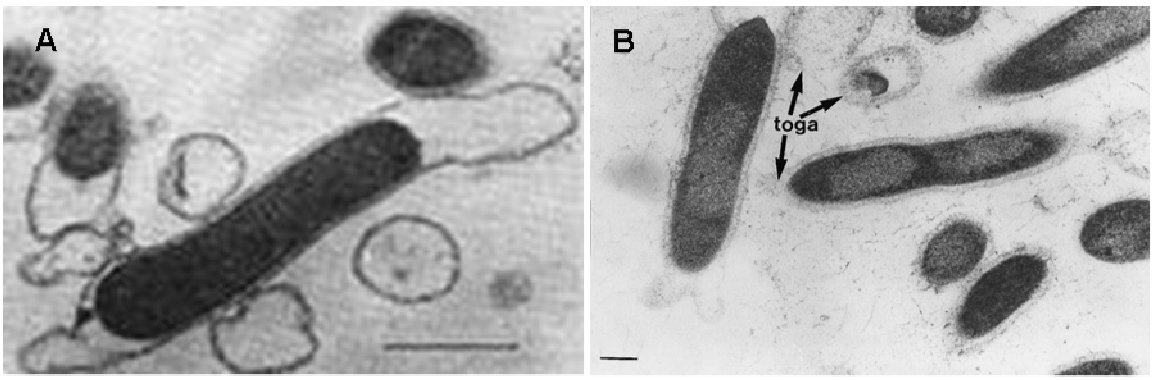


Figure 1.1 Transmission electron micrograph of thin section of *T. maritima* (A) and *F. pennivorans* (B), from http://www.biologie.uni-regensburg.de/Mikrobio/Stetter/Bilderhtml/thermotoga.html and (38), respectively.

Table 1.2 Taxonomy of thermophiles and hyperthermophiles able to grow on peptides and/or carbohydrates and their basic growth characteristics

Order	Genus	T_{opt} range (C°)	anaerobic/aerobic (an/ae)	peptides/ carbohydrates (p/c)[a]
Bacteria				
Thermotogales	*Thermotoga*	65-85	an	p, c
	Fervidobacterium	65-70	an	p, c
	Marinitoga	55-65	an	p, c
	Petrotoga	55-60	an	p, c
	Thermopallium	70	an	p, c
	Thermosipho	65-75	an	p, c
Aquificales	*Aquifex*	85	ae	-
	Hydrogenobacter	65-76	an*	-
	Hydrogenobaculum	65	an	-
	Hydrogenothermus	65	an*	-
	Persephonella	70-73	an*	-
	Sulfurihydrogenibium	60-65	fac. an	-
Thermoanaerobacteriales	*Thermoanaerobacter*	55-75	an	p, c
	Thermoanaerobacterium	55-65	an	c
	Carboxydibrachium	70	an	p, c
	Thermovenabulum	60	an	p, c
	Thermanaeromonas	70	an	c
	Thermacetogenium	55	an	-
	Moorella	55-65	an	c
	Gelria	55	an	c
	Coprothermobacter	55-60	an	p, c
Sphingobacteriales	*Rhodothermus*	65	ae	p, c
Dictyoglomales	*Dictyoglomus*	70-75	an	c
Thermales	*Thermus*	50-80	ae	p, c
	Marinithermus	70	ae	p, c
	Oceanithermus	60	ae	p, c
	Vulcanithermus	70	ae	p, c
Archaea				
Sulfolobales	*Acidianus*	70-88	fac. an	p
	Metallosphaera	65-75	ae	p
	Stygiolobus	80	an	-
	Sulfolobus	65-85	ae	p, c
	Sulfurisphaera	85	fac. an	p
Thermoproteales	*Thermofilum*	85-88	an	p, c
	Caldivirga	83	an	p, c
	Pyrobaculum	90-100	fac. an	p
	Thermocladium	75	an	p, c
	Thermoproteus	85	an	p
	Vulcanisaeta	85-90	an	p, c
Desulfurococcales	*Acidilobus*	85	an	p, c
	Aeropyrum	90-95	ae	p
	Desulfurococcus	85-90	an	p, c
	Ignicoccus	90	an	p, c
	Staphylothermus	85-90	an	p, c
	Stetteria	95	an	p
	Sulfophobococcus	90	an	
	Thermodiscus	88	an	p, c
	Thermosphaera	85	an	p, c
	Hyperthermus	99	an	p
	Pyrodictium	85-105	an	p
	Pyrolobus	103	fac. an	c
Thermococcales	*Palaeococcus*	80-83	an	p
	Pyrococcus	95-100	an	p, c
	Thermococcus	75-90	an	p, c
Archaeoglobales	*Archaeoglobus*	75-85	an	p
	Ferroglobus	85	an	p
	Geoglobus	88	an	p
	Strain 121	103	an	-

*known to accept low traces of O_2

[a]With 'peptides' complex compounds like peptone, tryptone, yeast extract are indicated. No symbol means growth on either of those substrates was not reported.

As an introduction to the work presented in this thesis, the hyperthermophilic and thermophilic microorganisms that were used will be described in more detail. *Thermotoga maritima* as well as *Fervidobacterium pennivorans* and *F. gondwanense* belong to the order of *Thermotogales*. All three organisms are rod-shaped and contain an unusual sheath-like structure, called toga (Figure 1.1). Together with the *Aquificales*, rRNA phylogeny suggests they represent the deepest phylogenetic branch within the bacterial domain (126).

T. maritima was isolated from a geothermally heated marine sediment at Vulcano, Italy (Figure 1.2). It is strictly anaerobic, has an optimal growth temperature of 80°C and is able to metabolize a range of sugars and peptides (46). Genome sequencing pointed out that almost 25% of the genes are highly similar to archaeal genes, suggesting lateral gene transfer between thermophilic bacteria and archaea (84).

F. pennivorans, an anaerobic thermophilic bacterium, was isolated from a hot spring of the Azores islands. It grows optimally at 70°C, pH 6.5, and has the remarkable ability to efficiently degrade native chicken feathers (38). Moreover, it is able to use many simple and complex carbohydrates as growth substrates. The anaerobe *F. gondwanense* was isolated from the geothermally heated well of the Great Artesian Basin of Australia. It grows optimally around 65°C and pH 7.0, fermenting a range of carbohydrates, including glucose, xylose, cellobiose and starch (2).

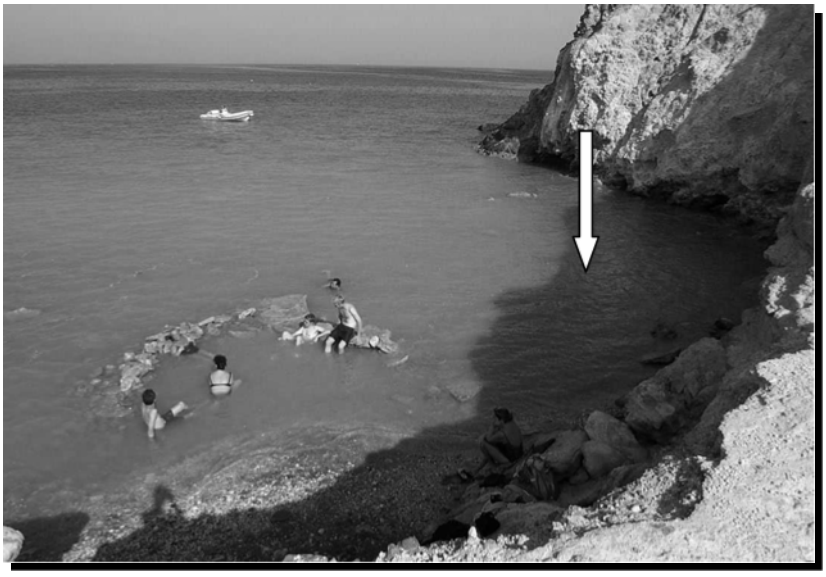


Figure 1.2 Geothermally heated marine sediment in Vulcano, Italy, habitat of the hyperthermophilic bacterium *Thermotoga maritima* (see arrow)(courtesy of Thijs Kaper).

Protein stability in hyperthermophiles

Metabolism in hyperthermophiles is accomplished by enzymes that demonstrate an extreme thermal stability. They maintain their native form at conditions, which will cause their mesophilic counterparts to denature and inactivate. Functional folding of proteins and enzymes is initially determined by their amino acid sequence (primary structure), and the generated, mutually attracting and repulsing forces result in the energetically most favorable, native structure (secondary, tertiary and quaternary) (93). While it was initially presumed that protein thermostability was a result of only additional hydrogen bonds and salt bridges (91), detailed comparisons between homologs from mesophilic and thermophilic microorganisms throughout the years showed it to be due to different combinations of several stabilizing features (68, 69, 121). The core of homologs of mesophilic and thermophilic microorganisms is often observed to be similar at the amino acid level, suggesting that gain in stabilization is achieved in regions that are less conserved (69, 121). Direct comparison of the amino acid composition of thermostable proteins to homologs from mesophiles reveals a higher contents of Ala, Arg and Tyr (69, 121). In addition, the residues Asn, Gln and Cys occur less frequently in the sequences of enzymes from thermophiles, because of their sensitivity to chemical deamination or oxidation at high temperatures (49, 69). Structural factors, such as surface loop deletion, increased occurrence of hydrophobic residues with branched side chains, and an increased proportion of charged residues also contribute to a higher intrinsic thermostability of proteins (68). Other stabilizing mechanisms include an increased number of ion-pair networks and salt bridges in thermostable proteins, compared to their mesophilic counterparts (1, 72, 128). Furthermore, an increased number of aromatic-aromatic interactions has been observed enzymes of hyperthermophilic origin, compared to homologs from mesophiles (79, 123). An abundance of disulfide bonds in intracellular proteins of thermophiles, in particular the archaeal *Pyrobaculum aerophilum* and *Aeropyrum pernix*, underlines its role in protein thermostability (80, 116). For enzymes from hyperthermophiles an increased number of subunits has been observed, compared to counterparts from mesophiles. Multimerization is believed to be a thermostabilizing factor (121). Besides enhancing enzyme activity, metal ions, like Co^{2+}, Mg^{2+}, $Ca^{2+}$and Mn^{2+}, are known to play a role in protein stabilization (19, 119).

Thermostabilization can also be achieved by external factors. Thermophiles, in contrast to mesophilic organisms, are known to accumulate compatible solutes, such as mannosylglycerate and di-*myo*-inositol-phosphate, which are believed to play a stabilizing role intracellularly (99). Likewise, chaperones are proteins that assist in the folding of other proteins, preventing them from aggregation. These so-called thermosomes have been observed in increased quantities in heat-shocked thermophiles (70, 105).

Together with the increase knowledge about the mechanisms that enable protein stability, the awareness has grown that thermostability of proteins is not realized by a single factor, but above all a combination of the mechanisms described above.

Enzymes from thermophiles as biocatalysts in industry

The application of enzymes as biocatalysts in industry has increased enormously over the last few decades, and nearly all classes are represented in industrial processes (Table 1.4). Following the discovery of the first thermophilic microorganisms, their enzymes with extreme thermal stability gained interest as a biocatalyst with great industrial potential, also because of their higher overall resistance to extreme conditions, such high or low pH and the presence of detergents and organic solvents (17, 30, 85). Their potential use in biotechnological processes has rendered them the name 'thermozymes' (118). Running industrial processes at elevated temperatures has numerous advantages: (i) reduction of the risk of contamination as compared to low temperatures, (ii) decreased viscosity and increased diffusion coefficient, leading to a higher bioavailability and solubility of organic compounds, and thereby providing a more efficient bioremediation (8), (iii) improved hydrolysis of hardly degradable or non-degradable polymers. In addition, due to their thermostability, a simple heat incubation step will suffice to remove the largest part of protein contamination, when expressed in mesophilic hosts, like *E. coli*.

With the arrival of genome sequencing projects, the complete array of genes of a number of organisms has been identified and mapped. Because of their proposed key position in phylogeny, thermophilic and hyperthermophilic microorganisms too have been involved in these projects. The first elucidated genome sequence from a hyperthermophilic microorganism was that from the methanogenic archaeon *Methanococcus jannaschii* in 1996 (18). So far, 16 genomes from thermophiles and hyperthermophiles have been sequenced (Table 1.3). Their complete genome sequences has given us information on metabolism at high temperatures, and with the mapping and identification of all genes cloning and expression of putative hydrolytic genes has become more accessible.

Table 1.3 Thermophilic and hyperthermophilic microorganisms, archaeal (A) or bacterial (B), and their completed genome sequence

Organism	Domain	Genome size (Mbase)	Reference
Methanococcus jannaschii	A	1.7	(18)
Archaeoglobus fulgidus	A	2.2	(62)
Methanobacterium thermoautotrophicum	A	1.8	(109)
Aquifex aeolicus	B	1.6	(28)
Pyrococcus horikoshii OT3	A	1.7	(56)
Aeropyrum pernix K1	A	1.7	(55)
Thermotoga maritima	B	1.9	(84)
Thermoplasma acidophilum	A	1.6	(98)
Thermoplasma volcanium	A	1.6	(57)
Pyrococcus furiosus	A	1.9	(97)
Sulfolobus solfataricus	A	3.0	(104)
Sulfolobus tokodaii	A	2.7	(54)
Pyrobaculum aerophilum	A	2.2	(37)
Methanopyrus kandleri	A	1.7	(108)
Thermoanaerobacter tengcongensis	B	2.7	(4)
Thermosynechococcus elongatus BP-1	B	2.6	(82)
Pyrococcus abyssi	A	1.7	(24)

Below the group of glycolytic and proteolytic enzymes will be discussed in more detail, with specific attention for the hydrolytic enzymes of bacterial origin. In addition, several other major industrial and special purpose enzymes will be reviewed, as well as their present and future industrial applications.

Protein and peptide degradation

By far the most important industrial enzymes are the proteases. Proteases account for about 60% of the total worldwide sale of enzymes, and almost half of this amount is represented by the alkaline proteases, which are used as additives in household detergents for laundering (Table 1.4) (95). Proteolytic enzymes catalyze the hydrolytic cleavage of peptide bonds, occurring in proteins and peptides. They are either exo-active, cleaving the peptide bond proximal to the amino- or carboxy-termini of the substrate, or endo-active, in which cleavage takes place distant from the termini. Based on their active site, a classification into four groups has been made, namely aspartic proteases, cysteine proteases, metalloproteases and serine proteases (40). Aspartic, or acidic proteases have two aspartic acid residues for their catalytic activity. Cysteine proteases have a catalytic dyad formed by cysteine and histidine residues, and most of them are highly specific. Metalloproteases require a divalent metal ion (usually Zn^{2+}) at the catalytic site, where it forms a ligand to histidine residues, with glutamate as the catalytic base. The serine proteases can be subdivided into clans, with trypsin-like and subtilisin-like serine proteases (subtilases) being the largest (for a review of the latter clan, see (106)). Both have a common reaction mechanism, involving the three residues aspartate, histidine and serine in the catalytic triad, albeit with a different orientation. Nearly all of the subtilases are synthesized as precursors, containing pre-pro-peptides (107). The pre-part (signal peptide) allows translocation over the cell membrane, the pro-part (also referred to as an intramolecular chaperone) mediates proper folding of the protease, resulting in an active enzyme after removal by autoproteolytic cleavage (47).

The majority of the proteases from thermophiles are of the serine type (for a review, see (27)). The first thermostable protease was isolated from an archaeal *Desulfurococcus* strain, and was optimally active at 95°C (25). One of the largest proteases was the subtilisin-like serine protease pyrolysin, isolated from *Pyrococcus furiosus*. With a half-life of 20 min at 105°C it is the most thermostable and thermoactive protease known to date (34). Pyrolysin appears to consist of two proteolytically active fractions, with different molecular weight (122).

Most of the characterized heat-stable proteases are derived from hyperthermophilic and thermophilic bacteria (Table 1.5). Two serine proteases were identified from the related thermophilic bacteria *Fervidobacterium pennivorans* and *F. islandicum*. Both organisms were shown to degrade native chicken feathers completely and their isolated proteases were able to degrade feather meal (38, 83). Purification of both proteases pointed out that they were membrane-bound. The characterization of the *F. pennivorans* keratinase has been described (65). A homo-multimeric protease was purified from *T. maritima*. It was found to have structural and gene sequence homology to a bacteriocin, yet its function remains unclear (43).

Table 1.4 Industrial enzymes and their applications

Industry	Enzyme class	Application
detergents	protease	protein stain removal
	amylase	starch stain removal
	cellulase	cleaning, color clarification
	mannanase	mannanan stain removal
	lipase	lipid stain removal
starch	amylase	starch liquefaction and saccharification
	amyloglucosidase	saccharification
	pullulanase	saccharification
	xylanase	viscosity reduction
	glucose isomerase	conversion glucose to fructose
	cyclodextrin-glycosyl transferase	cyclodextrin production
food	protease	milk clotting, infant formulas (low allergenic), flavor
	pectinase	fruit-based products (juices)
	pectin methylesterase	firming fruit-based products
	lipase	cheese flavor
	lactase	lactose removal (milk)
	transglutaminase	modify visco-elastic properties
baking	amylase	bread softness and volume, flour adjustment
	xylanase	dough conditioning
	lipase	dough stability and conditioning (emulsifier)
	phospholipase	dough stability and conditioning (emulsifier)
	glucose oxidase	dough strengthening
	lipoxygenase	dough strengthening, bread whitening
	protease	biscuits, cookies
	transglutaminase	laminated dough strengths
animal feed	phytase	phytate digestibility – phosphorous release
	xylanase	digestibility
	β-glucanase	digestibility
beverage	pectinase	de-pectination, mashing
	amylase	juice treatment, low calorie beer
	b-glucanase	mashing
	acetolactate decarboxylase	maturation (beer)
	laccase	clarification (juice), flavor (beer), cork stopper treatment
textile	cellulase	denim finishing, cotton softening
	amylase	de-sizing
	pectate lyase	scouring
	catalase	bleach termination
	laccase	bleaching
	peroxidase	excess dye removal
pulp and paper	lipase	pitch control, contaminant control
	protease	biofilm removal
	amylase	starch-coating, de-inking, drainage improvement
	xylanase	bleach boosting
	cellulase	de-inking, drainage improvement, fiber modification
	pectinase	bioscouring
fats and oils	lipase	transesterification
	phospholipase	de-gumming, lyso-lecithin production
leather	protease	unhearing, bating
	lipase	de-pickling

Adapted from Kirk et al. 2002 (61)

Table 1.5 Thermostable proteolytic enzymes of bacterial origin, all of which are serine proteases, except for *T. maritima* protease

Organism	T_{opt} (°C)	pH_{opt}	Details	References
F. islandicum	100	9.0	R	(83)
Aquifex pyrophylus	85	9.0	R/T	(23)
Thermus sp. Rt41A	90	8.0	T	(90)
F. pennivorans	80	10.0	T	(38)
Thermoanaerobacter keratinophilus	85	8.0	T	(96)
T. aquaticus	80	10.0	T	(81)
T. maritima	90	7.1	T	(43)

R.: purified, following overproduction in *E. coli*
T.: purified from thermophilic/hyperthermophilic host

From *Aquifex pyrophilus*, belonging to the order *Aquificales*, a serine protease was located and heterologously produced. The protease however, is not considered to be involved in the degradation of extracellular (poly)peptides, due to the inability of the organism to use it as a carbon source (23).

Polysaccharide degradation

Polysaccharides are chains of carbohydrate moieties, which can be either α- or β-linked. Both linkages require distinct glycolytic enzymes for their cleavage. The breakdown of these linkages can be achieved via β-elimination, performed by lyases, or hydrolysis, carried out by glycoside hydrolases. Like for polypeptides, polysaccharides can be degraded by endo- or exo-active enzymes, the latter generally attacking from the non-reducing end. The presence of this range of enzymes allows an organism to completely degrade carbohydrate polymers into oligomers and, eventually, monosaccharide units.

Based on their amino acid sequence homologies, the glycoside hydrolases have been classified into over 90 different families (41, 42) (URL server: http://afmb.cnrs-mrs.fr/~cazy/CAZY/), of which many contain glycosidases of thermophilic and hyperthermophilic origin. They degrade a range of polymeric substrates, which are abundantly present in nature, such as starch, (hemi)cellulose and pectin (for a review, see (6, 85)). These can be grouped according to their substrate preference, and are discussed below.

Starch-degrading enzymes

Starch (or glycogen in animals) is composed of α-glucose units, linked by α-1,4- and α-1,6-glycosidic bonds, thereby forming the two high-molecular-weight components amylose (20%) and amylopectin (80%). The starch-conversion is an important industrial process in which thermostable enzymes are applied. One of the most important applications is the enzymatic production of High Fructose Corn Syrup (HFCS), which is added to several softdrinks as a sweetener. Both the liquefaction and the saccharification steps in starch-conversion are preferably carried out at elevated temperatures. The liquefaction step is done by endo-active α-amylases, at present

primarily still originating from moderate thermophiles, like *B. licheniformis* and *B. amyloliquefaciens* (17). Many α-amylases from thermophiles have been purified and characterized, and most of them are of archaeal origin. All are highly thermostable, even in the absence of calcium ions, while many are optimally active at the desirable pH of 6.0-6.5 (for a review, see (35). The saccharolytic bacterium *Rhodothermus marinus* was found to have high levels of α-amylase activity when grown on starch (39). An α-amylase from *T. maritima* with high specific activity was produced in *E. coli*, although addition of calcium ions for its activity was required (77). Pullulanases (type I and II) are capable of hydrolyzing the α-1,6-bonds in starch. Whereas type I only debranches, type II is also able to degrade α-1,4-linkages. Both types have been identified in *T. maritima* and characterized (12, 67). The first glycosidase that has been isolated from the genus *Fervidobacterium* is a type I pullulanase from *F. pennavorans* Ven5. It was produced in *E. coli* and did not require calcium ions, like most of the reported pullulanases (10). The remaining oligosaccharides, like maltose and maltotriose, are attacked by the exo-active α-glucosidases (α-1,4) and glucoamylases (α-1,4 and α-1,6), although the latter is rare in thermophiles. Thermostable bacterial α-glucosidases have been identified in *T. ethanolicus* and *T. maritima* (9, 92).

Cellulose-degrading enzymes

Cellulose consists of glucose units linked via β-1,4-glycosidic bonds and is nature's most abundant plant polymer. Together with hemicellulose and lignin it forms a network in the cell wall of plants. Due to this incorporation and since a great part of the cellulose is composed of hydrolysis-resistant crystalline regions, alkaline pretreatment at high temperature is requisite. Consequently, thermostable cellulases have gained considerable interest. Complete hydrolysis entails the synergistic action of endo- and exoglucanases, and β-glucosidases. Most cellulose-degrading thermozymes have been isolated from bacterial thermophiles, although several glucanases of archaeal origin have been characterized (5, 59). Two thermostable endoglucanases from *T. maritima*, CelA and CelB, have been characterized and their optimal activity lies at 95°C (76). In sequence comparison and enzymatic features they appeared to be analogous to two cellulases from the related genus *T. neapolitana* (13). An endoglucanase, isolated from *A. aeolicus*, was found to be active on carboxymethylcellulose. Similar to *T. maritima* and *T. neapolitana*, growth of the organism on β-1,4-linked polymers has not been reported yet (60). In recent years, the possibilities have been explored to produce thermostable cellulases in plants for an efficient degradation of endosperm cell walls of barley. Examples of these transgenic plants are a thermostable (1,3-1,4)-β-glucanase, a hybrid from parental enzymes from *Bacillus macerans* and *B. amyloliquefaciens* and produced in barley, and a 1,4-β-glucanase from *Acidothermus cellulolyticus* produced in chloroplasts of tobacco plants (50, 51).

Xylan-degrading enzymes

Like cellulose, hemicellulose forms a fraction of the cell wall, where it is mainly composed of the heterogeneous polymer xylan. The main chain of xylan consists of β-1,4-linked xylose residues, which can be degraded by endo-active xylanases and exo-active β-xylosidases. The application of thermostable xylan-degrading enzymes in the paper industry, involving the bleaching process, is promising. The initial step in paper production involves a heat-treatment to open up the cell wall, removing 90% of the lignin from the wood pulp. The specific removal of the 10% residual lignin by thermostable xylanases instead of the traditional chemical treatment could reduce the amount of chlorine derivatives presently used (21). Up till now, not many thermostable versions of these enzymes have been characterized, since the number of thermophiles or hyperthermophiles which are known to be able to grow on this substrate is rather limited (114). Most of the xylanases described so far have been isolated from *Thermotoga* spp. that have the ability to grow on xylan, such as *T. maritima*, *T. neapolitana*, and *T. thermarum*, where they are tightly bound to the toga (113, 125, 129). With a high thermostability and a pH optimum of around 6.0 they meet the prerequisites to be used in the paper pulp bleaching process, as was already tested for one of the *T. maritima* xylanases (21). Nearly all xylanase-containing *Thermotoga* spp. also contain the corresponding β-xylosidases, to realize complete xylan degradation (114).

Pectin-degrading enzymes

Pectin is a branched heteropolysaccharide present in cell walls, where it forms a matrix, intertwining the other polymers, celluloses and hemicelluloses. The largest part of the pectin molecule consists of α-1,4-linked D-polygalacturonate, which can be partially methylated. Degradation of this polymer can occur via hydrolysis (hydrolases, belonging to family 28 of the glycoside hydrolases) or β-elimination (lyases, classified in family 1, 2, 3, 9, and 10 of the polysaccharide lyases). Pectinases are widely applied in the fruit juice extraction and clarification processes (acidic pectinases), as well as in the textile industry, where they are used for retting and degumming of fiber crops (alkaline pectinases) (53). The information on growth on pectin at high temperatures is rather limited.

Table 1.6 Thermostable pectinolytic enzymes of bacterial origin

Type	Organism	T_{opt} (°C)	pH_{opt}	Details	References
hydrolase	*T. maritima*	80	6.5	R	(64, 88)
lyase	*T. ethanolicus*	80	9.0	T	(66)
	T. maritima	90	9.0	R	(63, 87)
	Bacillus sp. TS47	70	8.0	T	(115)

R.: purified, following overproduction in *E. coli*
T.: purified from thermophilic/hyperthermophilic host

The only thermophiles for which growth on pectin has been reported are *Caldicellulosiruptor* and *Thermoanaerobacterium* strains (15, 45, 94, 101). Recently, it was found that nearly all *Thermotoga* species degrade pectin (Melike Balk, personal communication). Only one archaeon, *Desulfurococcus amyloliticus*, has been reported to grow on pectin (14). As a consequence, relatively few thermostable pectinolytic enzymes have been characterized (Table 1.6). Hence, their industrial potential may have been overlooked, since most attention has been given to other thermozymes, such as the cellulases. Potential applications could be in the fruit juice industry, where the clarification and color extraction steps are often carried out at elevated temperatures. In addition, an increase in temperature might reduce the viscosity caused by pectin. Other options might be the application in the retting of jute or flax fibers (53). Recently, the use of thermostable pectinases in bioscouring (removal of non-cellulosic compounds) of paper or cotton has been reported (71, 117).

The glycoside hydrolase family 28 only contains two thermostable galacturonases, one of which is originating from *Thermoanaerobacterium thermosulfurogenes*. The second is an exo-active polygalacturonase from *T. maritima*. It releases mono- instead of digalacturonate units from the non-reducing end, therefore its mode of action is more fungal-like (64, 88). Two pectate lyases from *Thermoanaerobacter italicus* have been purified, showing maximal activity at 80°C and pH 9.0 (66). Recently, a family 1 pectate lyase from *T. maritima* has been characterized in detail, displaying unique exocleaving properties by releasing trigalacturonate units (63, 87).

Xylose isomerases

Xylose (or glucose) isomerases catalyze the reversible isomerization of D-xylose to D-xylulose, but more importantly also of D-glucose to D-fructose. Over the last decades, their market has expanded tremendously, because of the application in the production of HCFS. Since the equilibrium of the isomerization reaction can be moved towards fructose by increasing the temperature (up to 55% of fructose at 90°C), the use of thermostable isomerases in this process is favored (11). Highly thermostable isomerases have been isolated from *T. maritima* and *T. neapolitana*, displaying optimal activity around 100°C (16, 120)(Table 1.7). Glucose isomerases require divalent metal ions, usually Mg^{2+} or Co^{2+}, for their activity and stability (11).

Current applications of thermostable enzymes and future scope

Although the amount of thermostable enzymes characterized has increased exponentially, they are applied in only few industrial processes. The best examples are special-purpose enzymes used in molecular biology, such as the *Taq* and *Pfu* DNA polymerases from *T. aquaticus* and *P. furiosus*, respectively. Coincidently, it were these very enzymes that largely expanded the knowledge of thermostable enzymes, with the introduction of PCR technology. However, a number of thermostable enzymes are on the brink of being applied in industry, thereby either replacing existing mesophilic enzymes, or adding new qualities to existing processes.

Table 1.7 Thermostable bacterial xylose isomerases

Organism	T_{opt} (°C)	pH_{opt}	Details	References
Thermoanaerobacterium thermosulfurogenes	80	7.5	T	(73)
Thermoanaerobacterium strain	80	6.8	R/T	(78)
T. saccharolyticum	80	7.0-7.5	R	(74)
T. maritima	105-110	6.5-7.5	T	(16)
T. neapolitana	97	7.1	R	(120)
T. thermophilus	95	7.0	R	(29)
T. aquaticus	85	7.0	T	(75)
T. flavus	90	7.0	R	(89)
T. caldophilus	90	-	R	(19)

R.: purified, following overproduction in *E. coli*
T.: purified from thermophilic/hyperthermophilic host

The isolation of new thermophilic and hyperthermophilic microorganisms will most likely lead to the discovery of more thermozymes with novel biocatalytic properties. Moreover, the existing thermozymes can be optimized in terms of stability and activity by using molecular approaches, such as site-directed mutagenesis and directed evolution. The increasing sequence information arriving from numerous genome projects combined with structural data, allows us to shape existing thermozymes to our interest (3, 36).

DNA microarray analysis as a biomolecular tool

DNA microarray analysis allows us to monitor the expression of many genes in parallel by measuring the corresponding RNA levels at different conditions via hybridization to DNA probes (100). The technique has profited optimally from the ongoing genome sequence projects that started off around the same time, in 1995. Its development it has advanced functional genomics on numerous (micro)organisms. The first organism of thermophilic origin subjected to this technique was the archaeon *P. furiosus*, for which the role of elementary sulfur on its metabolism was investigated by using a targeted approach (103). The same approach was recently used to examine the organism's response to thermal stress (105). The first complete-genome microarray for a thermophile followed shortly after (102).

The ability of the hyperthermophile *T. maritima* to grow on a variety of carbohydrates brought out the question which enzymes are involved in the catabolism of these sugars. By using a targeted cDNA microarray analysis the differential expression of genes encoding carbohydrate-active proteins was followed and several genes involved in sugar transport and regulation were identified (22).

Aims and outline of the thesis

The aims of the research described this thesis were (*i*) to examine the feasibility to clone and functionally express genes coding for thermostable enzymes of industrial interest in heterologous microbial production systems, (ii) to characterize these and gain insight in their mode of action, related to their structure, and (*iii*) to examine their role in the metabolic pathway of polymer degradation by studying the expression level under defined conditions. The attention was focused on both proteolytic and glycolytic enzymes, as well as a glucose isomerase. These all were produced in a heterologous expression host, purified and characterized in detail, in one case even at the structural level. Moreover, their role in catabolic pathways was examined, using a variety of approaches, including DNA microarrays.

Chapter 1 gives an overview of thermophiles and hyperthermophiles in general, their ability to use a range of simple and complex carbohydrates and peptides, and the enzymes involved the metabolism of these carbon sources. Furthermore, their thermostability and concomitant appeal to be used as biocatalysts in industry is discussed. Finally, the thermophilic and hyperthermophilic microorganisms that have been subject of this study are introduced.

The ability of *F. pennivorans* to grow on native chicken feathers prompted us to isolate the gene, encoding the protease responsible for this hydrolytic activity. Chapter 2 describes the cloning, identification and characterization of this gene encoding the protease, fervidolysin, that is responsible for keratinolysis by *F. pennivorans*. In Chapter 3, the 1.7-Å resolution crystal structure of the inactive form of fervidolysin is described in detail. This form was obtained via site-directed mutagenesis. The role of the four distinctive domains in catalysis, processing and substrate binding is discussed.

Chapter 4 describes the isolation of a xylose isomerase in *F. gondwanense*, using the highly homologous *xylA* gene from the phylogenetically related hyperthermophilic bacterium *T. maritima* as a probe in hybridization experiments. XylA was purified and characterized by examining the isomerization of glucose into fructose at high temperatures.

The genome sequence of *T. maritima* revealed two genes, which coded for enzymes involved in the degradation of polygalacturonic acid, a major component of pectin. Chapters 5 and 6 focus on the characterization of both enzymes, which belong to the first pectinases isolated and described from a thermophilic bacterium. In Chapter 5, the mode of action of an exopectate lyase is analyzed, using HPLC analysis to measure the length of the oligomers formed. Chapter 6 is concentrating on the analysis of the hydrolytic exopolygalacturonase, which appears to have a more fungal-like mode of action. The production of the two pectinases appeared to be induced by growth of *T. maritima* on pectin. To get more insight in the enzymes that are involved in pectin catabolism of this hyperthermophile, a targeted DNA microarray analysis was carried out. Besides high expression levels of the genes, encoding the two previously mentioned pectinases, the involvement of other genes could be confirmed with this functional genomics approach, and is described in Chapter 7.

Chapter 8 finally summarizes the results of the previous chapters, provides a general discussion, and concludes the main results of this thesis.

References

1. **Aguilar, C. F., I. Sanderson, M. Moracci, M. Ciaramella, R. Nucci, M. Rossi, and L. H. Pearl.** 1997. Crystal structure of the β-glycosidase from the hyperthermophilic Archaeon *Sulfolobus solfataricus*: resilience as a key factor in themostability. J. Mol. Biol. **271:**789-802.
2. **Andrews, K. T., and B. K. C. Patel.** 1996. *Fervidobacterium gondwanense* sp. nov., a new thermophilic anaerobic bacterium isolated from nonvolcanically heated geothermal waters of the Great Artesian Basin of Australia. Int. J. Syst. Bacteriol. **46:**265-269.
3. **Arnold, F. H., P. L. Wintrode, K. Miyazaki, and A. Gershenson.** 2001. How enzymes adapt: lessons from directed evolution. Trends Biochem. Sci. **26:**100-106.
4. **Bao, Q., Y. Tian, W. Li, Z. Xu, Z. Xuan, S. Hu, W. Dong, J. Yang, Y. Chen, Y. Xue, Y. Xu, X. Lai, L. Huang, X. Dong, Y. Ma, L. Ling, H. Tan, R. Chen, J. Wang, J. Yu, and H. Yang.** 2002. A complete sequence of the *T. tengcongensis* genome. Genome Res. **12:**689-700.
5. **Bauer, M. W., L. E. Driskill, W. Callen, M. A. Snead, E. J. Mathur, and R. M. Kelly.** 1999. An endoglucanase, EglA, from the hyperthermophilic archaeon *Pyrococcus furiosus* hydrolyzes β-1,4 bonds in mixed-linkage (1-3),(1-4)-β-D-glucans and cellulose. J. Bacteriol. **181:**284-290.
6. **Bauer, M. W., L. E. Driskill, and R. M. Kelly.** 1998. Glycosyl hydrolases from hyperthermophilic microorganisms. Curr. Opin. Biotechnol. **9:**141-145.
7. **Bauer, M. W., S. B. Halio, and R. M. Kelly.** 1996. Proteases and glycosyl hydrolases from hyperthermophilic microorganisms. Adv. Prot. Chem. **48:**271-310.
8. **Becker, P., I. Abu-Reesh, S. Markossian, G. Antranikian, and H. Märkl.** 1997. Determination of the kinetic parameters during continuous cultivation of the lipase producing thermophile *Bacillus* sp. IHI-91 on olive oil. Appl. Microbiol. Biotechnol. **48:**184-190.
9. **Bertoldo, C., and G. Antranikian.** 2001. Amylolytic enzymes from hyperthermophiles. Methods Enzymol. **330:**269-289.
10. **Bertoldo, C., F. Duffner, P. L. Jorgensen, and G. Antranikian.** 1999. Pullulanase type I from *Fervidobacterium pennavorans* Ven5: cloning, sequencing, and expression of the gene and biochemical characterization of the recombinant enzyme. Appl. Environ. Microbiol. **65:**2084-2091.
11. **Bhosale, S. H., M. B. Rao, and V. V. Deshpande.** 1996. Molecular and industrial aspects of glucose isomerase. Microbiol. Rev. **60:**280-300.
12. **Bibel, M., C. Brettl, U. Gosslar, G. Kriegshäuser, and W. Liebl.** 1998. Isolation and analysis of genes for amylolytic enzymes of the hyperthermophilic bacterium *Thermotoga maritima*. FEMS Microbiol. Lett. **158:**9-15.
13. **Bok, J.-D., D. A. Yernool, and D. E. Eveleigh.** 1998. Purification, characterization, and molecular analysis of thermostable cellulases CelA and CelB from *Thermotoga neapolitana*. Appl. Environ. Microbiol. **64:**4774-4781.
14. **Bonch-Osmolovskaya, E. A., A. I. Slesarev, M. L. Miroshnichenko, T. P. Svetlichnaya, and V. A. Alekseev.** 1988. Characteristics of *Desulfurococcus amyloliticus* sp. nov., a new extremely thermophilic archaebacterium isolated from thermal springs of Kamchatka and Kunashir Island. Mikrobiologiya. **57:**94-101.
15. **Bredholt, S., J. Sonne-Hansen, P. Nielsen, I. M. Mathrani, and B. K. Ahring.** 1999. *Caldicellulosiruptor kristjanssonii* sp. nov., a cellulolytic, extremely thermophilic, anaerobic bacterium. Int. J. Syst. Bacteriol. **49:**991-996.

16. **Brown, S. H., C. Sjöholm, and R. M. Kelly.** 1993. Purification and characterization of a highly thermostable glucose isomerase produced by the extremely thermophilic eubacterium, *Thermotoga maritima.* Biotech. Bioeng. **41:**878-886.
17. **Bruins, M. E., A. E. M. Janssen, and R. M. Boom.** 2001. Thermozymes and their applications. Appl. Biochem. Biotechnol. **90:**155-186.
18. **Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. Geoghagen, and J. C. Venter.** 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii.* Science. **273:**1058-1073.
19. **Chang, C., B. C. Park, D.-S. Lee, and S. W. Suh.** 1999. Crystal structures of thermostable xylose isomerases from *Thermus caldophilus* and *Thermus thermophilus*: Possible structural determinants of thermostability. J. Mol. Biol. **288:**623-634.
20. **Chatton, E.** 1938. Titres et Travaux Scientifiques (1906-1937), Sète, France.
21. **Chen, C.-C., R. Adolphson, J. F. D. Dean, K.-E. L. Eriksson, M. W. Adams, and J. Westphelling.** 1997. Release of lignin from kraft pulp by a hyperthermophilic xylanase from *Thermotoga maritima.* Enzyme Micr. Technol. **20:**39-45.
22. **Chhabra, S. R., K. R. Shockley, S. B. Conners, K. E. Scott, R. D. Wolfinger, and R. M. Kelly.** 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima.* J. Biol. Chem. **278:**7540-7552.
23. **Choi, I.-G., W.-G. Bang, S.-H. Kim, and Y. G. Yu.** 1999. Extremely thermostable serine-type protease from *Aquifex pyrophilus.* J. Biol. Chem. **274:**881-888.
24. **Cohen, G. N., V. Barbe, d. Flament, M. Galperin, R. Heilig, O. Lecompte, O. Poch, D. Prieur, J. Querellou, R. Ripp, J. C. Thierry, J. van der Oost, J. Weissenbach, Y. Zivanovic, and P. Forterre.** 2003. An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi.* Mol. Microbiol. **47:**1495-1512.
25. **Cowan, D. A., K. A. Smolenski, R. M. Daniel, and H. W. Morgan.** 1987. An extremely thermostable extracellular proteinase from a strain of the archaebacterium *Desulfurococcus* growing at 88 degrees C. Biochem. J. **247:**21-33.
26. **de Vos, W. M., S. W. Kengen, W. G. B. Voorhorst, and J. van der Oost.** 1998. Sugar utilization and its control in hyperthermophiles. Extremophiles. **2:**201-205.
27. **de Vos, W. M., W. G. B. Voorhorst, M. Dijkgraaf, L. D. Kluskens, J. van der Oost, and R. J. Siezen.** 2001. Purification, characterization, and molecular modelling of pyrolysin and other extracellular thermostable serine proteases from hyperthermophilic microorganisms. Methods Enzymol. **330:**383-393.
28. **Deckert, G., P. V. Warren, T. Gaasterland, W. G. Young, A. L. Lenox, D. E. Graham, R. Overbeek, M. A. Snead, M. Keller, M. Aujay, R. Huber, R. A. Feldman, J. M. Short, G. J. Olsen, and R. V. Swanson.** 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus.* Nature. **392:**353-358.
29. **Dekker, K., A. Sugiura, H. Yamagata, K. Sakaguchi, and S. Udaka.** 1992. Efficient production of thermostable *Thermus thermophilus* xylose isomerase in *Escherichia coli* and *Bacillus brevis.* Appl. Microbiol. Biotechnol. **36:**727-732.
30. **Demirjian, D. C., F. Moris-Varas, and C. S. Cassidy.** 2001. Enzymes from extremophiles. Curr. Opin. Chem. Biol. **5:**144-151.

31. **Di Giulio, M.** 2003. The ancestor of the Bacteria domain was a hyperthermophile. J. Theor. Biol. **224:**277-283.
32. **Di Giulio, M.** 2001. The universal ancestor was a thermophile or a hyperthermophile. Gene. **281:**11-17.
33. **Driskill, L. E., K. Kusy, M. W. Bauer, and R. M. Kelly.** 1999. Relationship between glycosyl hydrolase inventory and growth physiology of the hyperthemophile *Pyrococcus furiosus* on carbohydrate-based media. Appl. Environ. Microbiol. **65:**893-897.
34. **Eggen, R. I. L., A. C. M. Geerling, J. Watts, and W. M. de Vos.** 1990. Characterization of pyrolysin, a hyperthermoactive serine protease from the archaebacterium *Pyrococcus furiosus.* FEMS Microbiol. Lett. **71:**17-20.
35. **Eichler, J.** 2001. Biotechnological uses of archaeal extremozymes. Biotechnol. Adv. **19:**261-278.
36. **Farinas, E. T., T. Bulter, and F. H. Arnold.** 2001. Directed enzyme evolution. Curr. Opin. Biotechnol. **12:**545-551.
37. **Fitz-Gibbon, S. T., H. Ladner, U. J. Kim, K. O. Stetter, M. I. Simon, and J. H. Miller.** 2002. Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum.* Proc. Natl. Acad. Sci. U.S.A. **99:**984-989.
38. **Friedrich, A. B., and G. Antranikian.** 1996. Keratin degradation by *Fervidobacterium pennavorans*, a novel thermophilic anaerobic species of the order *Thermotogales.* Appl. Environ. Microbiol. **62:**2875-2882.
39. **Gomes, I., J. Gomes, and W. Steiner.** 2003. Highly thermostable amylase and pullulanase of the extreme thermophilic eubacterium *Rhodothermus marinus*: production and partial characterization. Bioresource Technol. **90:**207-214.
40. **Hartley, B. S.** 1960. Proteolytic enzymes. Annu. Rev. Biochem. **29:**45-72.
41. **Henrissat, B.** 1991. A classification of glycosyl hydrolases based on amino acid sequence similarity. Biochem. J. **280:**309-316.
42. **Henrissat, B., and A. Bairoch.** 1996. Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. **316:**695-696.
43. **Hicks, P. M., K. D. Rinker, J. R. Baker, and R. M. Kelly.** 1998. Homomultimeric protease in the hyperthermophilic bacterium *Thermotoga maritima* has structural and amino acid sequence homology to bacteriocins in mesophilic bacteria. FEBS Lett. **440:**393-398.
44. **Hough, D. W., and M. J. Danson.** 1999. Extremozymes Curr. Opin. Chem. Biol. **3:**39-46.
45. **Huang, C.-Y., B. K. Patel, R. A. Mah, and L. Baresi.** 1998. *Caldicellulosiruptor owensensis* sp. nov., an anaerobic, extremely thermophilic, xylanolytic bacterium. Int. J. Syst. Bacteriol. **48:**91-97.
46. **Huber, R., T. A. Langworthy, H. König, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. Arch. Microbiol. **144:**324-333.
47. **Ikemura, H., H. Takagi, and M. Inouye.** 1987. Requirement of pro-sequence for the production of active subtilisin E in *Escherichia coli.* J. Biol. Chem. **262:**7859-7864.
48. **Islas, S., A. M. Velasco, A. Becerra, L. Delaye, and A. Lazcano.** 2003. Hyperthermophily and the origin and earliest evolution of life. Int. Microbiol. **6:**87-94.
49. **Jaenicke, R.** 2000. Stability and stabilization of globular proteins in solution. J. Biotechnol. **79:**193-203.
50. **Jensen, L. G., O. Olsen, O. Kops, N. Wolf, K. K. Thomsen, and D. von Wettstein.** 1996. Transgenic barley expressing a protein-engineered, thermostable (1,3-1,4)-beta-glucanase during germination. Proc. Natl. Acad. Sci. U.S.A. **93:**3487-3491.

51. **Jin, R., S. Richter, R. Zhong, and G. K. Lamppa.** 2003. Expression and import of an active cellulase from a thermophilic bacterium into the chloroplast both in vitro and in vivo. Plant Mol. Biol. **51:**493-507.
52. **Kashefi, K., and D. R. Lovley.** 2003. Extending the upper temperature limit for life. Science. **301:**934.
53. **Kashyap, D. R., P. K. Vohra, S. Chopra, and R. Tewari.** 2001. Applications of pectinases in the commercial sector: a review. Bioresource Technol. **77:**215-227.
54. **Kawarabayasi, Y., Y. Hino, H. Horikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, A. Ankai, H. Kosugi, A. Hosoyama, S. Fukui, Y. Nagai, K. Nishijima, R. Otsuka, H. Nakazawa, M. Takayima, Y. Kato, T. Yoshizawa, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, K. Aoki, S. Masuda, M. Yanagii, M. Nishimura, A. Yamagishi, T. Oshima, and H. Kikuchi.** 2001. Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7. DNA Res. **8:**123-140.
55. **Kawarabayasi, Y., Y. Hino, H. Horikawa, S. Yamazaki, Y. Haikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, A. Ankai, H. Kosugi, A. Hosoyama, S. Fukui, Y. Nagai, K. Nishijima, H. Nakazawa, M. Takamiya, S. Masuda, T. Funahashi, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, H. Kikuchi, et. al.** 1999. Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1. DNA Res. **6:**145-152.
56. **Kawarabayasi, Y., M. Sawada, H. Horikawa, Y. Haikawa, Y. Hino, S. Yamamoto, M. Sekine, S. Baba, H. Kosugi, A. Hosoyama, Y. Nagai, M. Sakai, K. Ogura, R. Otsuka, H. Nakazawa, M. Takamiya, Y. Ohfuku, T. Funahashi, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, K. Aoki, and H. Kikuchi.** 1998. Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. DNA Res. **5:**55-76.
57. **Kawashima, T., N. Amano, H. Koike, S. Makino, S. Higiuchi, Y. Kawashima-Ohya, K. Watanabe, M. Yamazaki, K. Kanehori, T. Kawamoto, T. Nunoshiba, Y. Yamamoto, H. Aramaki, K. Makino, and M. Suzuki.** 2000. Archaeal adaptation to higher temperatures revealed by genomic sequence of *Thermoplasma volcanium.* Proc. Natl. Acad. Sci. U.S.A. **97:**14257-14262.
58. **Kelly, R. M., and M. W. Adams.** 1994. Metabolism in hyperthermophilic microorganisms. Antonie van Leeuwenhoek. **66:**247-270.
59. **Kengen, S. W., E. J. Luesink, A. J. Stams, and A. J. Zehnder.** 1993. Purification and characterization of an extremely thermostable β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus.* Eur. J. Biochem. **213:**305-312.
60. **Kim, J. O., S. R. Park, W. J. Lim, S. K. Ryu, M. K. Kim, C. L. An, S. J. Cho, Y. W. Park, J. H. Kim, and H. D. Yun.** 2000. Cloning and characterization of thermostable endoglucanase (Cel8Y) from the hyperthermophilic *Aquifex aeolicus* VF5. Biochem. Biophys. Res. Commun. **279:**420-426.
61. **Kirk, O., T. V. Borchert, and C. C. Fuglsang.** 2002. Industrial enzyme applications. Curr. Opin. Biotechnol. **13:**345-351.
62. **Klenk, H. P., R. A. Clayton, J. F. Tomb, O. White, K. E. Nelson, K. A. Ketchum, R. J. Dodson, M. L. Gwinn, E. K. Hickey, J. D. Peterson, D. L. Richardson, A. R. Kerlavage, D. E. Graham, N. C. Kyrpides, R. D. Fleischmann, J. Quakenbush, N. H. Lee, G. G. Sutton, S. R. Gill, E. F. Kirkness, B. A. Dougherty, K. McKenney, M. D. Adams, B. Loftus, J. C. Venter, et. al.** 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus.* Nature. **390:**364-370.

63. **Kluskens, L. D., G.-J. W. M. van Alebeek, A. G. J. Voragen, W. M. de Vos, and J. van der Oost.** 2003. Molecular and biochemical characterisation of the thermo-active family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima.* Biochem. J. **370:**651-659.
64. **Kluskens, L. D., G.-J. W. M. van Alebeek, J. Walther, A. G. J. Voragen, W. M. de Vos, and J. van der Oost.** 2003. PelB from the hyperthermophilic bacterium *Thermotoga maritima* encodes a fungal-like intracellular exopolygalacturonase. submitted.
65. **Kluskens, L. D., W. G. B. Voorhorst, R. J. Siezen, R. M. Schwerdtfeger, G. Antranikian, J. van der Oost, and W. M. de Vos.** 2002. Molecular characterization of fervidolysin, a subtilisin-like serine protease from the thermophilic bacterium *Fervidobacterium pennivorans.* Extremophiles. **6:**185-194.
66. **Kozianowski, G., F. Canganella, F. A. Rainey, H. Hippe, and G. Antranikian.** 1997. Purification and characterization of thermostable pectate-lyases from a newly isolated thermophilic bacterium, *Thermoanaerobacter italicus* sp. nov. Extremophiles. **1:**171-182.
67. **Kriegshäuser, G., and W. Liebl.** 2000. Pullulanase from the hyperthermophilic bacterium *Thermotoga maritima*: purification by β-cyclodextrin affinity chromatography. J. Chromatography B. **737:**245-251.
68. **Kumar, S., and R. Nussinov.** 2001. How do thermophilic proteins deal with heat? Cell. Mol. Life Sci. **58:**1216-1233.
69. **Kumar, S., C.-J. Tsai, and R. Nussinov.** 2000. Factors enhancing protein thermostability. Protein Eng. **13:**179-191.
70. **Ladenstein, R., and G. Antranikian.** 1998. Proteins from hyperthermophiles. Adv. Biochem. Eng/Biotechnol. **61:**37-85.
71. **Lange, N. E. C., L. Kongsbak, M. Shulein, M. Eskelund, and P. A. Husain** 1998. Biopreparation of textiles at high temperatures.
72. **Lebbink, J. H. G.** 1999. Molecular characterisation of the thermostability and catalytic properties of enzymes from hyperthermophiles. PhD Thesis. Wageningen University, Wageningen.
73. **Lee, C. Y., and G. J. Zeikus.** 1991. Purification and characterization of thermostable glucose isomerase from *Clostridium thermosulfurogenes* and *Thermoanaerobacter* strain B6A. Biochem. J. **273:**565-571.
74. **Lee, Y. E., M. V. Ramesh, and G. J. Zeikus.** 1993. Cloning, sequencing and biochemical characterization of xylose isomerase from *Thermoanaerobacterium saccharolyticum* strain B6A-RI J. Gen. Microbiol. **139:**1227-1234.
75. **Lehmacher, A., and H. Bisswanger.** 1990. Comparative kinetics of D-xylose and D-glucose isomerase activities of the D-xylose isomerase from *Thermus aquaticus* HB8. Biol. Chem. Hoppe-Seyler. **371:**527-536.
76. **Liebl, W., P. Ruile, K. Bronnenmeier, K. Riedel, F. Lottspeich, and I. Greif.** 1996. Analysis of a *Thermotoga maritima* DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. Microbiology. **142:**2533-2542.
77. **Liebl, W., I. Stemplinger, and P. Ruile.** 1997. Properties and gene structure of the *Thermotoga maritima* α-amylase AmyA, a putative lipoprotein of a hyperthermophilic bacterium. J. Bacteriol. **179:**941-948.
78. **Liu, S. Y., J. Wiegel, and F. C. Gherardini.** 1996. Purification and cloning of a thermostable xylose (glucose) isomerase with an acidic pH optimum from *Thermoanaerobacterium* strain JW/SL-YS 489. J. Bacteriol. **178:**5938-5945.

79. **Machius, M., N. Declerck, R. Huber, and G. Wiegand.** 2003. Kinetic stabilization of *Bacillus licheniformis* α-amylase through introduction of hydrophobic residues at the surface. J. Biol. Chem. **278:**11546-11553.
80. **Mallick, P., D. R. Boutz, D. Eisenberg, and T. O. Yeates.** 2002. Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds. Proc. Natl. Acad. Sci. U.S.A. **99:**9679-9684.
81. **Matsuzawa, H., K. Tokugawa, M. Hamaoki, M. Mizoguchi, H. Taguchi, I. Terada, S. T. Kwon, and T. Ohta.** 1988. Purification and characterization of aqualysin I (a thermophilic alkaline serine protease) produced by *Thermus aquaticus* YT-1. Eur. J. Biochem. **171:**441-447.
82. **Nakamura, Y., T. Kaneko, S. Sato, M. Ikeuchi, H. Katoh, S. Sasamoto, A. Watanabe, M. Iriguchi, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, N. Nakazaki, S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada, and S. Tabata.** 2002. Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. DNA Res. **9:**123-130.
83. **Nam, G.-W., D.-W. Lee, H.-S. Lee, N.-J. Lee, B.-C. Kim, E.-A. Choe, J.-K. Hwang, M. T. Suhartono, and Y.-R. Pyun.** 2002. Native-feather degradation by *Fervidobacterium islandicum* AW-1, a newly isolated keratinase-producing thermophilic anaerobe. Arch. Microbiol. **178:**538-547.
84. **Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, L. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, O. White, S. L. Salzberg, H. O. Smith, J. C. Venter, and C. M. Fraser.** 1999. Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima.* Nature. **399:**323-329.
85. **Niehaus, F., C. Bertoldo, M. Kähler, and G. Antranikian.** 1999. Extremophiles as a source of novel enzymes for industrial application. Appl. Microbiol. Biotechnol. **51:**711-729.
86. **Nisbet, E. G., and N. H. Sleep.** 2001. The habitat and nature of early life. Nature. **409:**1083-1091.
87. **Parisot, J., A. Ghochikyan, V. Langlois, V. Sakanyan, and C. Rabiller.** 2002. Exopolygalacturonate lyase from *Thermotoga maritima*: cloning, characterization and organic synthesis application. Carbohydrate Res. **337:**1427-1433.
88. **Parisot, J., V. Langlois, V. Sakanyan, and C. Rabiller.** 2003. Cloning expression and characterization of a thermostable exopolygalacturonase from *Thermotoga maritima.* Carbohydrate Res. **338:**1333-1337.
89. **Park, B. C., S. Koh, C. Chang, S. W. Suh, D.-S. Lee, and S. M. Byun.** 1997. Cloning and expression of the gene for xylose isomerase from *Thermus flavus* AT62 in *Escherichia coli* Appl. Biochem. Biotechnol. **62:**15-27.
90. **Peek, K., R. M. Daniel, C. Monk, L. Parker, and T. Coolbear.** 1992. Purification and characterization of a thermostable proteinase isolated from *Thermus* sp. strain Rt41A. Eur. J. Biochem. **207:**1035-1044.
91. **Perutz, M. F., and H. Raidt.** 1975. Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. Nature. **255:**256-259.
92. **Raasch, C., W. Streit, J. Schanzer, M. Bibel, U. Gosslar, and W. Liebl.** 2000. *Thermotoga maritima* AglA, an extremely thermostable NAD^+-, Mn^{2+}-, and thiol-dependent α-glucosidase. Extremophiles. **4:**189-200.
93. **Radford, S. E.** 2000. Protein folding: progress made and promises ahead. Trends Biochem. Sci. **25:**611-618.

94. **Rainey, F. A., A. M. Donnison, P. H. Janssen, D. Saul, A. Rodrigo, P. L. Bergquist, R. M. Daniel, E. Stackebrandt, and H. W. Morgan.** 1994. Description of *Caldicellulosiruptor saccharolyticus* gen. nov., sp. nov: An obligately anaerobic, extremely thermophilic, cellulolytic bacterium. FEMS Microbiol. Lett. **120:**263-266.

95. **Rao, M. B., A. M. Tanksale, M. S. Ghatge, and V. V. Deshpande.** 1998. Molecular and biotechnological aspects of microbial proteases. Microbiol. Mol. Biol. Rev. **62:**597-635.

96. **Riessen, S., and G. Antranikian.** 2001. Isolation of T*hermoanaerobacter keratinophilus* sp. nov., a novel thermophilic, anaerobic bacterium with keratinolytic activity. Extremophiles. **5:**399-408.

97. **Robb, F. T., D. Maeder, J. R. Brown, J. DiRuggiero, M. D. Stump, R. K. Yeh, R. B. Weiss, and D. M. Dunn.** 2001. Genomic sequence of hyperthermophile, *Pyrococcus furiosus*: implications for physiology and enzymology. Methods Enzymol. **330:**134-157.

98. **Ruepp, A., W. Graml, M. L. Santos-Martinez, K. K. Koretke, C. Volker, H. W. Mewes, D. Frishman, S. Stocker, A. N. Lupas, and W. Baumeister.** 2000. The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum.* Nature. **407:**508-513.

99. **Santos, H., and M. S. da Costa.** 2002. Compatible solutes of organisms that live in hot saline environments. Environ.Microbiol. **4:**501-509.

100. **Schena, M., D. Shalon, R. W. Davis, and P. O. Brown.** 1995. Quantitative monitoring of gene expression patterns with complementary DNA microarray. Science. **270:**467-470.

101. **Schink, B., and J. G. Zeikus.** 1983. Characterization of pectinolytic enzymes of *Clostridium thermosulfurogenes.* FEMS Microbiol. Lett. **17:**295-298.

102. **Schut, G. J., S. D. Brehm, S. Datta, and M. W. W. Adams.** 2003. Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: *Pyrococcus furiosus* grown on carbohydrates or peptides. J. Bacteriol. **185:**3935-3947.

103. **Schut, G. J., J. Zhou, and M. W. W. Adams.** 2001. DNA microarray analysis of the hyperthermophilic archaeon *Pyrococcus furiosus*: Evidence for a new type of sulfur-reducing enzyme complex. J. Bacteriol. **183:**7027-7036.

104. **She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Chan-Weiher, I. G. Clausen, B. A. Curtis, A. De Moors, G. Erauso, C. Fletcher, P. M. Gordon, I. Heikamp-de Jong, A. C. Jeffries, C. J. Kozera, N. Medina, X. Peng, H. P. Thi-Ngoc, P. Redder, M. E. Schenk, C. Theriault, N. Tolstrup, R. L. Charlebois, W. F. Doolittle, M. Duguet, T. Gaasterland, R. A. Garrett, M. A. Ragan, C. W. Sensen, and J. van der Oost.** 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc. Natl. Acad. Sci. U.S.A. **98:**7835-7840.

105. **Shockley, K. R., D. E. Ward, S. R. Chhabra, S. B. Conners, C. I. Montero, and R. M. Kelly.** 2003. Heat shock response by the hyperthermophilic Archaeon *Pyrococcus furiosus* Appl. Environ. Microbiol. **69:**2365-2371.

106. **Siezen, R. J., and J. A. M. Leunissen.** 1997. Subtilases: The superfamily of subtilisin-like serine proteases. Protein Sci. **6:**501-523.

107. **Siezen, R. J., J. A. M. Leunissen, and U. Shinde.** 1995. Homology analysis of the propeptides of subtilisin-like serine proteases (subtilases), p. 231-253. *In* U. Shinde (ed.), Intramolecular chaperones and folding. Landes Company, Austin.

108. **Slesarev, A. I., K. V. Mezhevaya, K. S. Makarova, N. N. Polushin, O. V. Shcherbinina, V. V. Shakhova, G. L. Belova, L. Aravind, D. A. Natale, I. B. Rogozin, R. L. Tatusov, Y. I. Wolf, K. O. Stetter, A. G. Malykh, E. V. Koonin, and S. A. Kozyavkin.** 2002. The complete genome of

hyperthermophile *Methanopyrus kandleri* AV19 and monophyly of archaeal methanogens. Proc. Natl. Acad. Sci. U.S.A. **99:**4644-4649.

109. **Smith, D. R., L. A. Doucette-Stamm, C. Deloughery, H. Lee, J. Dubois, T. Aldredge, R. Bashirzadeh, D. Blakely, R. Cook, K. Gilbert, D. Harrison, L. Hoang, P. Keagle, W. Lumm, B. Pothier, D. Qiu, R. Spadafora, R. Vicaire, Y. Wang, J. Wierzbowski, R. Gibson, N. Jiwani, A. Caruso, D. Bush, J. N. Reeve, et. al.** 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* delta H: functional analysis and comparative genomics. J. Bacteriol. **179:**7135-7155.
110. **Stetter, K. O.** 1999. Extremophiles and their adaptation to hot environments. FEBS Lett. **452:**22-25.
111. **Stetter, K. O.** 1996. Hyperthermophilic prokaryotes. FEMS Microbiol. Lett. **18:**149-158.
112. **Stetter, K. O.** 1994. The lesson of Archaebacteria. Nobel symposium, New York, Columbia University Press**:**143-151.
113. **Sunna, A., and G. Antranikian.** 1996. Growth and production of xylanolytic enzymes by the extreme thermophilic anaerobic bacterium *Thermotoga maritima.* Appl. Microbiol. Biotechnol. **45:**671-676.
114. **Sunna, A., M. Moracci, M. Rossi, and G. Antranikian.** 1997. Glycosyl hydrolases from hyperthermophiles. Extremophiles. **1:**2-13.
115. **Takao, M., T. Nakaniwa, K. Yoshikawa, T. Terashita, and T. Sakai.** 2000. Purification and characterization of thermostable pectate lyase with protopectinase activity from thermophilic *Bacillus* sp. TS47. Biosci. Biotechnol. Biochem. **64:**2360-2367.
116. **Toth, E. A., C. Worby, J. E. Dixon, E. R. Goedken, S. Marqusee, and T. O. Yeates.** 2000. The crystal structure of adenylosuccinate lyase from *Pyrobaculum aerophilum* reveals an intracellular protein with three disulfide bonds. J. Mol. Biol. **301:**433-450.
117. **Tzanov, T., M. Calafell, G. M. Guebitz, and A. Cavaco-Paulo.** 2001. Bio-preparation of cotton fabrics. Enzyme Micr. Technol. **29:**357-362.
118. **Vieille, C., D. S. Burdette, and G. J. Zeikus.** 1996. Thermozymes. Biotechnol. Annu. Rev. **2:**1-83.
119. **Vieille, C., K. L. Epting, R. M. Kelly, and G. J. Zeikus.** 2001. Bivalent cations and amino-acid composition contribute to the thermostability of *Bacillus licheniformis* xylose isomerase. Eur. J. Biochem. **268:**6291-6301.
120. **Vieille, C., J. M. Hess, R. M. Kelly, and G. J. Zeikus.** 1995. xylA cloning and sequencing and biochemical characterization of xylose isomerase from *Thermotoga neapolitana*. Appl. Environ. Microbiol. **61:**1867-1875.
121. **Vieille, C., and G. J. Zeikus.** 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol. Mol. Biol. Rev. **65:**1-43.
122. **Voorhorst, W. G. B., R. I. L. Eggen, A. C. M. Geerling, C. Platteeuw, R. J. Siezen, and W. M. de Vos.** 1996. Isolation and characterization of the hyperthermostable serine protease, pyrolysin, and its gene from the hyperthermophilic archaeon *Pyrococcus furiosus.* J. Biol. Chem. **271:**20426-20431.
123. **Voorhorst, W. G. B., A. Warner, W. M. de Vos, and R. J. Siezen.** 1997. Homology modelling of two subtilisin-like serine proteases from the hyperthermophilic archaea *Pyrococcus furiosus* and *Thermococcus stetteri.* Protein Eng. **10:**905-914.
124. **Ward, D. E., K. R. Shockley, L. S. Chang, R. D. Levy, J. K. Michel, S. B. Conners, and R. M. Kelly.** 2002. Proteolysis in hyperthermophilic microorganisms. Archaea. **1:**63-74.
125. **Winterhalter, C., and W. Liebl.** 1995. Two extremely thermostable xylanases of the hyperthermophilic bacterium *Thermotoga maritima* MSB8. Appl. Environ. Microbiol. **61:**1810-1815.

126. **Woese, C. R., O. Kandler, and M. Wheelis.** 1990. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria* and *Eucarya*. Proc. Natl. Acad. Sci. U.S.A. **87:**4576-4579.

127. **Xu, Y., and N. Glansdorff.** 2002. Was our ancestor a hyperthermophilic procaryote? Comp. Biochem. Physiol. part A. **133:**677-688.

128. **Yip, K. S., L. Britton, T. J. Stillman, J. H. G. Lebbink, W. M. de Vos, F. T. Robb, C. Vetriani, D. Maeder, and D. W. Rice.** 1998. Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. Eur. J. Biochem. **255:**336-346.

129. **Zverlov, V., K. Piotukh, O. Dakhova, G. Velikodvorskaya, and R. Borriss.** 1996. The multidomain xylanase A of the hyperthermophilic bacterium *Thermotoga neapolitana* is extremely thermoresistant. Appl. Microbiol. Biotechnol. **45:**245-247.

2

Molecular characterization of fervidolysin, a subtilisin-like serine protease from the thermophilic bacterium *Fervidobacterium pennivorans*

Leon D. Kluskens, Wilfried G.B. Voorhorst, Roland J. Siezen, Ruth M. Schwerdtfeger, Garabed Antranikian, John van der Oost, and Willem M. de Vos

Reprinted with permission from Extremophiles (2002) 6: 185-194.

Abstract

The *fls* gene encoding fervidolysin, a keratin-degrading proteolytic enzyme from the thermophilic bacterium *Fervidobacterium pennivorans*, was isolated using degenerate primers, combined with Southern hybridization and inverse PCR. Further sequence characterization demonstrated that the 2.1-kb *fls* gene encoded a 699-amino acid prepro-enzyme, showing high homology with the subtilisin family of the serine proteases. It was cloned into a pET9d vector, without its signal sequence, and expressed in *Escherichia coli*. The heterologously produced fervidolysin was purified by heat incubation followed by ion exchange chromatography and emerged in the soluble fraction as three distinct protein bands, as judged upon SDS-PAGE. Amino-terminal sequence analysis of these bands and their comparison with that determined from the biochemically purified keratinase and its predicted protein sequence, identified them as the 73-kDa fervidolysin precursor, the 58-kDa mature fervidolysin, and the 14-kDa fervidolysin pro-peptide. Using site-directed mutagenesis, the active-site residue histidine at position 79 was substituted for an alanine residue. The resulting fervidolysin showed a single protein band corresponding to the size of the 73-kDa fervidolysin precursor, indicating that its proteolytic cleavage is a result of an autoproteolytic process. Knowledge-based modeling experiments showed a distinctive binding region for subtilases, in which binding of the pro-peptide can take place prior to autoproteolysis. Assays using keratin and other proteinaceous substrates did not display activity of fervidolysin, which may be the result of the tight binding of the pro-peptide in the substrate binding site, where it then could function as an inhibitor.

Introduction

Thermophilic microorganisms are adapted to thrive at temperatures above 60°C. They are a source of interesting enzymes that are both thermoactive and thermostable (Niehaus et al. 1999; Vieille et al. 1996). Thermophiles are able to grow on a range of proteinaceous substrates, such as casein, tryptone, peptone and casamino acids. For this purpose, these organisms possess an array of enzymes that enables the degradation, uptake and further metabolism of these substrates under these extreme conditions. Within this process of proteolysis an important role is played by extracellular proteases. Initiated by a signal peptide these proteases are targeted across the cell membrane, where they hydrolyze protein polymers into peptides (Wandersman 1989).

The serine proteases comprise the majority of all the thermostable proteases produced and characterized to date (Niehaus et al. 1999). In general, most of the serine proteases belong to the subtilisin-like serine proteases, also referred to as subtilases (Barrett and Rawlings 1995; Rawlings and Barrett 1994; Siezen et al. 1991). In total, over 200 members of the subtilase family originating from all domains of life are known and classified (Siezen and Leunissen 1997). Up to now, only a limited number of thermophilic subtilases from archaea (Klingeberg et al. 1995; Völkl et al. 1994; Voorhorst et al. 1997; Voorhorst et al. 1996) and bacteria (Choi et al. 1999; Jang et al. 1992; Terada et al. 1990) have been characterized at the biochemical and genetic level (de Vos et al. 2000).

Recently, a novel thermophilic bacterium, named *Fervidobacterium pennivorans*, was isolated from a hot spring on the Azores islands. This organism has the ability to degrade native chicken feathers at high temperatures. An extracellular protease capable of hydrolyzing these keratins has been purified from *F. pennivorans* and its enzymatic properties have been described previously (Friedrich and Antranikian 1996). Until recently, keratinolytic activity was mainly attributed to actinomycetes (Böckle et al. 1995; Bressollier et al. 1999; Mukhopadhyay and Chandra 1990; Noval and Nickerson 1959), a few mesophilic fungi (Kunert 1972; Siesenop and Bohm 1995; Takiuchi et al. 1982), and some *Bacillus* species (Lin et al. 1995; Takami et al. 1989; Varela et al. 1996). Except for the thermophilic actinomycete *Thermoactinomyces candidus*, which grows at 65°C (Ignatova et al. 1999), no other thermophilic organism is reported to digest keratinaceous materials. However, no molecular features on the keratinolytic enzyme from *T. candidus* have been provided.

In this study we describe the cloning, identification and characterization of the *fls* gene encoding the protease, further referred to as fervidolysin, that is responsible for keratinolysis by *F. pennivorans*. Moreover, we expressed the *fls* gene in *Escherichia coli* and showed by site-directed mutagenesis that the produced fervidolysin is subject to autoproteolytic processing as expected from predictions based on knowledge-based modeling.

Materials and methods

Organisms, growth conditions and plasmids

F. pennivorans was cultivated at conditions described previously (Friedrich and Antranikian 1996). The bacterial strain used for initial cloning experiments was *Escherichia coli* TG1 [*supE hsd* Δ5 *thi* Δ(*lac-proAB*) F' (*tra*D35 *proAB*$^{+}$ *lacI*^{q} *lac*Z ΔM15) (Gibson 1984). *E. coli* BL21(DE3*) (hsdS gal* (λcl*ts* 857 *ind*1 *S*am7 *nin*5 lacUV5-T7 gene 1) was used for heterologous expression. Cultivation of these strains was carried out at 37°C in L broth (Sambrook et al. 1989), supplemented with either 50 μg of kanamycin or 10 μg of ampicillin per ml. The plasmids used for recombinant work were pGEF+ (Schanstra et al. 1993), pET9d from Novagen, pGEM-T supplied by Promega and pUC18 from Amersham Pharmacia Biotech Inc.

Recombinant DNA techniques

Genomic DNA of *F. pennivorans* was isolated by using an established protocol (Ramakrishnan and Adams 1995). Small-scale plasmid DNA was isolated by using the QIAGEN purification kit (Qiagen). DNA was digested with restriction endonucleases from Life Technologies, used as specified by the manufacturer. Ligation was carried out with T4 DNA ligase following the manufacturer's specifications (Life Technologies). DNA fragments were purified from agarose by QiaexII or from a PCR mix by using the PCR purification kit (Qiagen Inc.). Transformations of *E. coli* TG1 and BL21(DE3) were carried out using established procedures (Sambrook et al. 1989).

Cloning of the fls gene encoding fervidolysin

Two oppositely oriented degenerated primers (Eurogentec) were designed on the basis of the conserved regions, including either the aspartic acid or the serine active-site residue from subtilisin-like serine proteases (Figure 2.1). BG95: 5'-GTIGCIGTIMTIGAYACIGG-3' (I: inosine; M: A or C; Y: T or C). BG97: 5'-GGIGWIGSCATISWIGTICC-3' (I: inosine; W: A or T; S: C or G). PCR reactions were performed using 100 ng of both oligonucleotides and 500 ng *F. pennivorans* genomic DNA as template in a final volume of 100 µl. After an initial denaturation step of 5 min at 95°C the DNA thermal cycler (Perkin Elmer) was programmed for a PCR of 35 cycles with *Taq* polymerase (at 1 U per reaction): 1 min at 95°C (denaturation), 2 min at 35°C (annealing), and 3 min at 72°C (extension), followed by a final extension of 7 min at 72°C. The resulting 0.65-kb fragment was cloned into a pGEM-T vector (Promega) and named pLUW754 (Figure 2.1). Southern hybridization of genomic *F. pennivorans* DNA with the 0.65-kb PCR fragment was performed as described below. Hybridizing fragments were isolated from gel and subcloned in pUC18 and pGEF+, resulting in pLUW751 and pLUW752 (Figure 2.1). The 3'-end of the gene was identified and characterized by using the inverse PCR technique described below. Based on the overall sequence two primers were designed to amplify the *fls* gene. The sense primer was designed excluding a putative signal sequence, adding a methionine upstream of the initial asparagine: BG361, 5'-CGCGC**TCATGA**ATCCGAGTTTTGAGCCAAGG, (*Bsp*HI restriction site in boldface) The antisense primer was designed as follows (*Bam*HI restriction site in boldface): BG574, 5'-GCCGTTAAA**GGATCC**CTATCACTGT-CCG. PCR was executed using *Pfu* polymerase in the method described above. A product with the expected size of 2.0 kb was digested with *Bsp*HI and *Bam*HI and cloned in a pET9d expression vector, resulting in pLUW761, which was introduced into *E. coli* TG1 and BL21(DE3).

Southern hybridization

Genomic DNA of *F. pennivorans* was digested with *Hin*cII/*Hin*dIII, *Bam*HI/*Nco*I and *Pvu*I/*Hin*dIII and separated on a 1% agarose gel. Following gel electrophoresis the DNA was transferred by capillary blotting to a Hybond N^+ filter (Amersham Pharmacia Biotech Inc.) (Sambrook et al. 1989). Hybridization was performed overnight at 65°C with the 0.65-kb PCR product that was [α-P^{32}]ATP labeled (nick-translation). Subsequently, the filter was washed, twice with 2xSSC (0.3 M NaCl, 0.03 M sodium citrate) containing 1 mM EDTA and twice with 2xSSC containing 0.1% SDS at room temperature, followed by an overnight exposure of the filter to a phosphorimager screen.

DNA amplification by inverse PCR

Inverse PCR (Triglia et al. 1988) was carried out with 50 ng of genomic *F. pennivorans* DNA, digested with *Sca*I and self-ligated to obtain circular DNA. Five ng of the ligation mix was used in a PCR with 1 unit of Expand™ Long Template enzyme mixture (Roche Diagnostics). After preheating to 94°C for 5 min, 30 cycles were performed, consisting of denaturation at 94°C for 1 min, primer annealing at 55°C for 2 min, and elongation at 72°C for 3 min. The primers used were

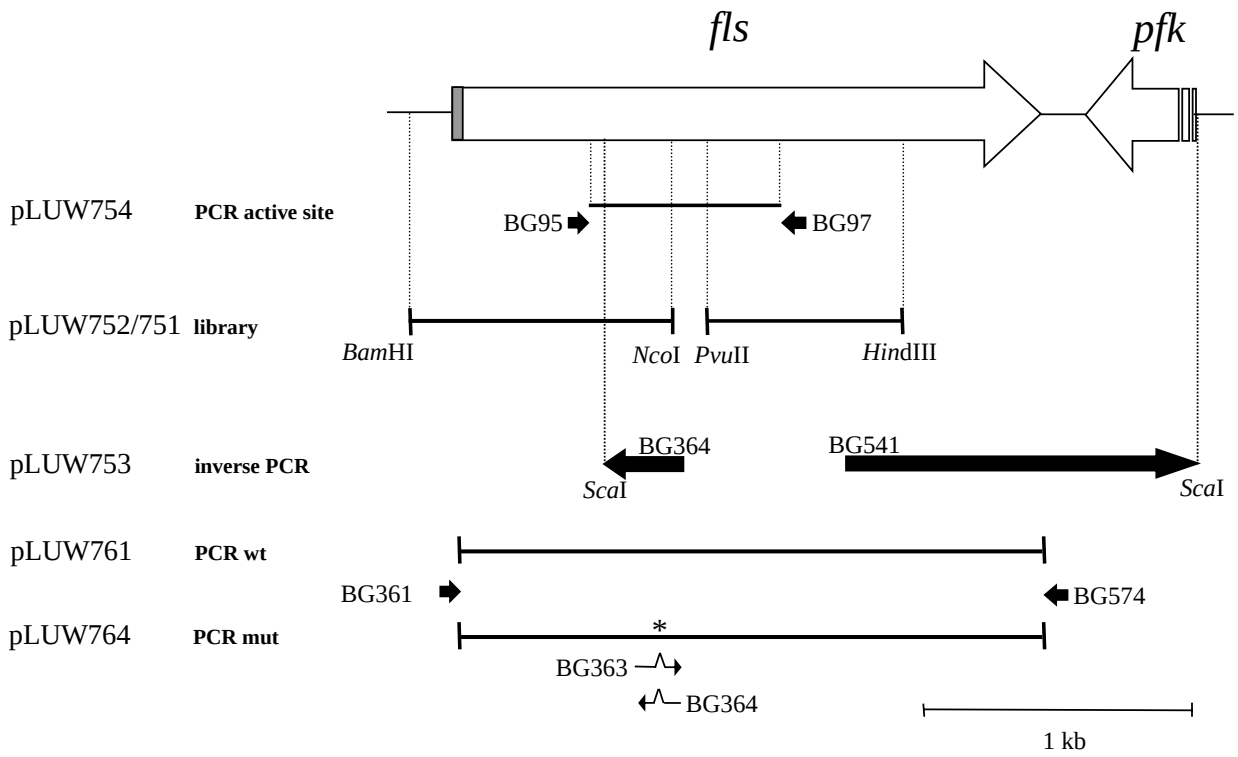


Figure 2.1 Diagrammatic overview of the cloning strategy followed to obtain the fervidolysin (*fls)* gene (top arrow). The signal sequence is depicted in grey. Downstream the *fls* gene in the opposite transcription direction an orf is depicted coding for a putative phosphofructokinase (*pfk*). Essential restriction enzymes are depicted and primers used are represented by their BG-number. All constructs are represented with their assigned pLUW number. The expression constructs are described as PCR wt and PCR mut. The star represents the mutation His → Ala. Abbreviations: wt, wildtype; mut, mutant.

BG364 and BG541. The resulting 1.2-kb PCR product, containing the 3'-end of the gene, was cloned in a pGEM-T vector and named pLUW753 (Figure 2.1).

DNA sequencing

Cloned PCR products were sequenced by the dideoxynucleotide chain termination method (Sanger et al. 1977) with a Li-Cor automatic sequencing system (model 4000L, Westburg). DNA and protein sequencing data were analyzed with DNASTAR and compared to the GenBank Data Base by BLAST (Altschul et al. 1997).

3D-protein modeling of the catalytic domain of fervidolysin

3D-modeling was carried out using QUANTA 3.2.3 (Molecular Simulations, Cambridge, UK) (Sali and Blundell 1993) and CHARMm 22 (Brooks et al. 1983) was run on a Silicon Graphics 4D25TG workstation. The model is based on the known X-ray structure of subtilisin BPN. Insertions of more than six residues relative to subtilisin were not included. After addition of H-atoms the active-site residues (Asp32, His64 and Ser221), the Asn155 (subtilisin BPN numbering (Siezen et al. 1991))

and the two β-sheet strand backbones eI and eIII involved in the substrate binding were constrained and the entire molecule was subjected to energy minimization.

Overexpression of the fls gene and purification of fervidolysin

E. coli BL21 (DE3) harboring pLUW761 was grown overnight and cells were harvested by centrifugation. The obtained cell pellet was resuspended in 20 mM Tris buffer (pH 8.0) and sonicated using a Branson sonifier. Cell debris was removed by centrifugation and the resulting supernatant was subjected to a heat incubation of 20 min at 70°C. Precipitated proteins were removed by an additional centrifugation step. The supernatant was loaded onto a Q Sepharose column (Amersham Pharmacia Biotech Inc.), which was equilibrated with the same buffer. Bound proteins were eluted by a linear gradient from 0 to 1 M NaCl in 20 mM Tris buffer (pH 8.0). The pooled fractions were concentrated (Filtron Technology Corp., 30 kDa molecular weight cutoff). Protein samples were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) in 10% (w/v) gels using the method of Laemmli (Laemmli 1970). The samples were prepared by heating for 5 min at 100°C in an equal volume of sample buffer, containing 5% SDS.

Zymogram staining

For zymogram analysis samples were loaded onto a standard SDS-PAGE gel (Laemmli 1970), after a 5 min heat incubation step at 100°C. Following electrophoresis the gel was washed overnight at room temperature in a 50 mM Tris-HCl buffer, pH 10, containing 10 mM $CaCl_2$ and soaked for 1 hour in a 1% (w/v) casein Hammarsten solution (Serva) in water, pH 10. Subsequently, the gel was incubated at 55°C for 1 hour in the washing buffer. Finally, the gel was stained with Coomassie blue and destained to reveal the zones of casein hydrolysis.

N-terminal amino acid sequence analysis

The keratinolytic enzyme purified from *F. pennivorans* (Friedrich and Antranikian 1996) was digested with thermolysin before analysis of the N-terminal sequence of the generated fragments by Edman degradation. Purified *E. coli*-produced fervidolysin was separated by SDS-PAGE (Laemmli 1970) and transferred onto a PVDF membrane (Bio-rad Laboratories) by electroblotting at room temperature in 10 mM CAPS blotting buffer (pH 11.0), containing 10 % methanol (Sambrook et al. 1989). Transferred protein bands were visualized by staining with Ponceau S. The N-termini of the generated fragments of the purified keratinolytic enzyme were analyzed by Edman degradation at a Chromatec in Greifswald, the fragments resulting from the recombinant gene were examined by Applied Biosystems 477A Protein Sequencer Chromatogram Report at the Institute for Organic Chemistry and Biochemistry in Freiburg, Germany.

Determination of N-glycosylation sites

N-glycosylation was examined by using the periodic acid-Schiff (PAS) technique, as described by Zacharius and coworkers (Zacharius et al. 1969). Alternatively, the purified keratinolytic enzyme from *F. pennivorans* was incubated overnight at 37°C with 1U of N-glycanase F (Roche

Diagnostics). The degree of *N*-glycosylation was examined by comparing the mass of the protease to its untreated state by SDS-PAGE.

Mutagenesis of the fls gene

F. pennivorans DNA was used as a template for mutagenesis of the *fls* gene, in which the active-site mutation His79Ala was introduced in a two step PCR using overlap extension (Higuchi et al. 1988) (Figure 2.1). The first step involved the two primers introducing the mutation and the two flanking primers, described earlier (BG361 and BG574). An amount of 300 ng for each primer was used. The two mutagenesis primers were: BG363 (5'-GGTTCA**GCCGGC**ACA CACGTTGCTGG, sense) and BG364 (5'-CGTGTGT-**GCCGGC**TGAACCACC GTAGG, antisense). In the second step, the two overlapping PCR products of the first step were used as a template (200 ng each) which, in combination with the flanking primers BG361 and BG574, resulted in the amplification of the complete *fls* gene carrying the mutation. The verification of the mutation was simplified by introducing a *Nae*I site at the mutation site (boldface in the primers). Nucleotides changed by the introduction of the *Nae*I site are underlined. After sequence analysis, the amplified DNA was cloned as previously described and *E. coli* BL21(DE3) was transformed with the resulting plasmid (pLUW764).

Nucleotide sequence accession number

The nucleotide sequence has been submitted to the GenBank/EMBL Data Bank with the accession number AY035311.

Results

Isolation and characterization of the fls gene

With degenerate primers BG95 and BG97 (Figure 2.1) the active site region of a subtilisin-like serine protease was amplified (de Vos et al. 2000; Voorhorst et al. 1997), as was confirmed by a database search. Southern blot analysis resulted in the isolation of a 1.1-kb *Nco*I/*Bam*HI fragment, which was subcloned in pGEF+ (pLUW752) and included the sequence of the 5'-end of the *fls* gene. A second hybridizing fragment of 0.7 kb (*Pvu*II/*Hin*dIII) covering a large part of the 3'-end of *fls*, was subcloned in pUC18 (pLUW751)(Figure 2.1). Ultimately, the 3'-end was identified by inverse PCR techniques**.** Nucleotide sequence analysis of the subclones (Figure 2.1) revealed the complete 2103-bp *fls* gene, encoding a 699-amino acid protein that showed significant similarity with the subtilase family of the serine proteases (Siezen and Leunissen 1997). The gene is flanked by cis-elements like a consensus Shine-Dalgarno sequence and a 13-nt inverted repeat (Figure 2.2). No obvious promoter regions could be identified. Following the Von Heijne rules a putative signal peptide of 21 amino acids was predicted (Figure 2.3) (von Heijne 1985; von Heijne 1986). Downstream of this signal peptide a potential pro-peptide sequence of 128 amino acids was identified based on its homology to other propeptides that include conserved hydrophobic residues

(Siezen et al. 1995) and confirmed by N-terminal analysis (see below). The mature protein contains the catalytic triad Asp41, His79 and Ser260 in this order (numbering begins at start mature protease), as well as a high degree of homology at amino acid level around these residues when comparing to other subtilases (Figure 2.3). No hydrophobic region that could serve as a membrane anchor was identified at the C-terminus. The sequence harbors two cysteine residues, one of which however is positioned in the signal peptide and therefore unlikely to be involved in disulfide bridge formation. A putative cleavage site upstream of the first cysteine residue, which is characteristic for lipoproteins, was excluded by the Von Heijne rules (von Heijne 1985; von Heijne 1986; von Heijne 1989). The predicted molecular mass of the fervidolysin precursor is 75 kDa, and 58 kDa of the predicted mature protein. BLASTP analysis of the complete amino acid sequence (Altschul et al. 1997) gave the highest hit (approximately 35% identity) with two subtilases from *Deinococcus radiodurans* (White et al. 1999). No homologues were identified in the genome sequence from the phylogenetically related bacterium *Thermotoga maritima* (Nelson et al. 1999). 152 nt downstream of the *fls* gene the 3'-end of a second orf was identified, appearing in the opposite transcription direction, which shows the highest homology with a putative carbohydrate kinase (belonging to the *pfk*B family) from *Archaeoglobus fulgidus* (Figure 2.1).

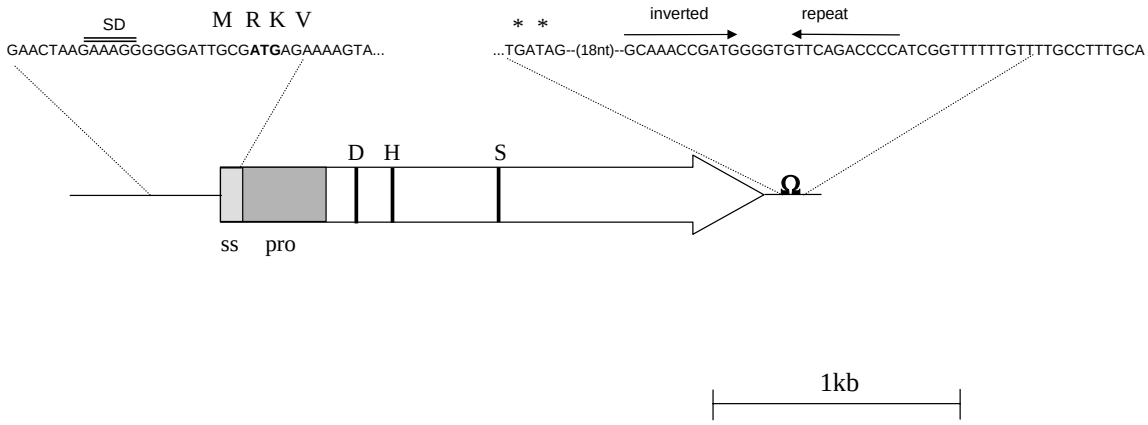


Figure 2.2 Schematic overview of the genetic organization of the *fls* gene. Upstream the *fls* gene a Shine-Dalgarno region (SD) is marked. Downstream a possible terminator (**Ω**) is identified by an inverted repeat, followed by a T-stretch. The ATG start codon is in boldface. The first 4 amino acids of fervidolysin are depicted above the sequence, stars illustrate the double stop codon (TGATAG). Ss = signal sequence, pro = pro-sequence, D = Asp, H = His, S = Ser.

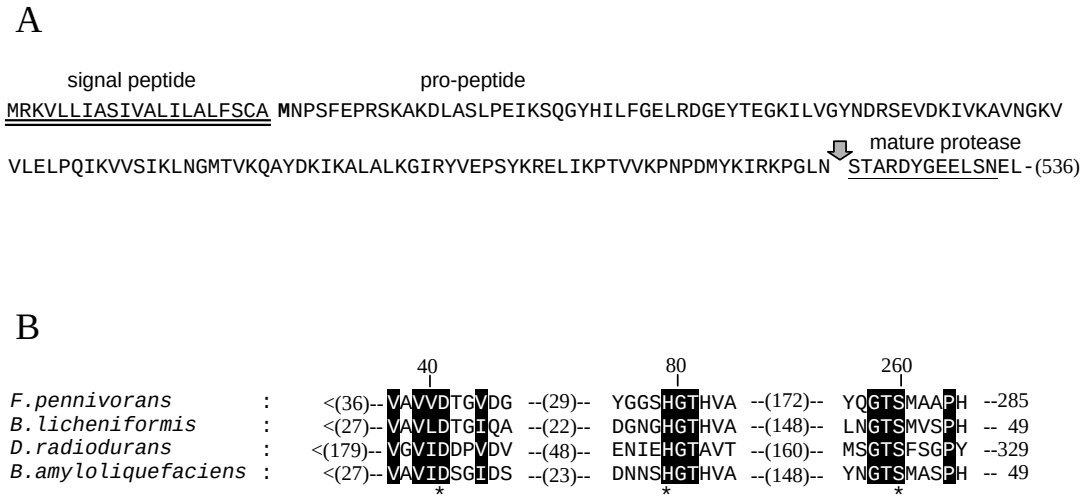


Figure 2.3 A. Overview of the N-terminal amino acid sequence of the fervidolysin precursor. The signal peptide (double underlined), the introduced methionine at the start of the pro-peptide (boldface) and the cleavage site between the pro-peptide and the mature fervidolysin (arrow) have been indicated. The N-terminal sequence, as determined by Edman degradation of the mature fervidolysin, is underlined, indicating the start of the mature protease.
B. Sequence alignment of fervidolysin with related subtilisin-like serine proteases. Partial sequences are depicted around the active-site residues Asp, His and Ser (marked by stars). Shading of the residues indicates the level of homology. The depicted sequences are available under the following GenBank accession numbers: *D. radiodurans* subtilase (AAF12593); *Bacillus licheniformis* keratinase (S78160); *Bacillus amyloliquefaciens* subtilisin BPN (X00165).

Expression of the fls gene in E. coli and purification of fervidolysin

The *fls* gene encoding fervidolysin was amplified by PCR, excluding its signal sequence and including an additional methionine at position -1, and cloned in a pET9d vector under control of the T7 promoter. The resulting plasmid was named pLUW761 and contains the putative catalytic domain and a likely pro-peptide sequence. The gene was significantly overexpressed in *E. coli* BL21(DE3), even in the absence of IPTG (Figure 2.4A, lane 1). The produced protein was initially purified from a cell-free extract by a heat incubation step of 20 min at 70°C. Fractions were analyzed on an SDS-PAGE gel, revealing three bands of approximately 73 kDa, 58 kDa and 14 kDa (Figure 2.4A). All three bands were absent in the extract of *E. coli* BL21(DE3), carrying the pET9d plasmid. No overexpression was detected when using an *fls*-construct devoid of the pro-peptide sequence (not shown). Based on the calculated mass of all three bands the upper 73-kDa band corresponded to the pro-protease precursor, whereas the second band matched the size of the predicted mature protein. The lowest band indicated the pro-peptide according to its size of 14 kDa, which was determined in a high percentage polyacrylamide gel (Figure 2.4B, lane 5). This heat-stable cell-free extract was further purified by anion exchange chromatography (Q Sepharose), during which the protein eluted around 0.45 M NaCl. All three bands co-eluted as one single peak during both ion exchange chromatography and gel filtration, showing a similar heterogeneity upon SDS-PAGE (Figure 2.4B, lane 5). To elucidate the identity of the three polypeptides the N-terminal amino acid sequence of all three protein bands was determined.

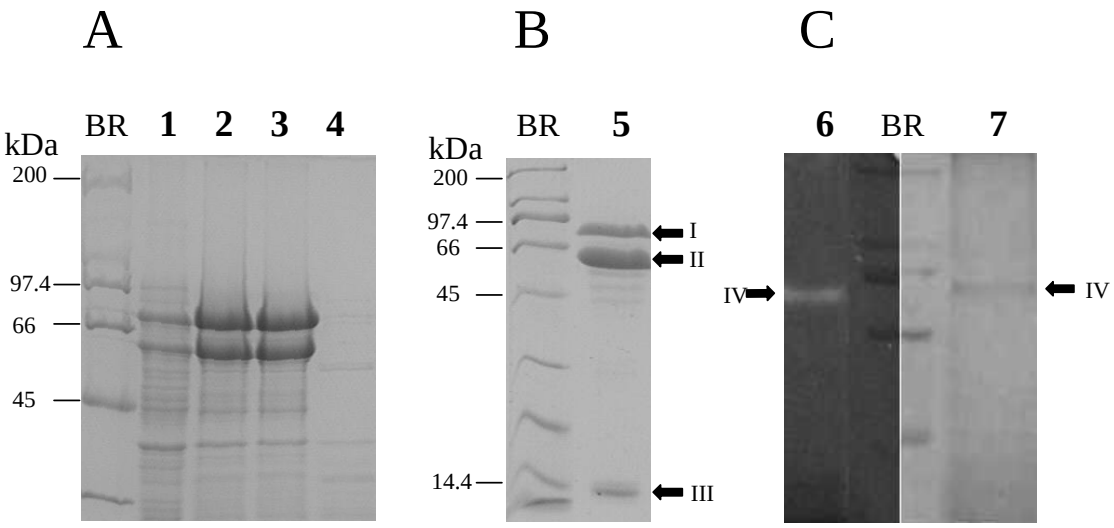


Figure 2.4 SDS-PAGE analysis of protein extracts from *E. coli* BL21(DE3), overexpressing the *fls* gene on pLUW761, at a 10% gel (**A**) and a 12% gel (**B**). Numbers adjacent to the gel indicate the type of polypeptide: I, fervidolysin precursor (73 kDa); II, mature fervidolysin (58 kDa); III, pro-peptide (14 kDa); IV, *purified F. pennivorans* keratinase (90 kDa). BR, Broad Range Marker; lane 1, crude cell extract; lane 2, soluble fraction; lane 3, soluble fraction after heat incubation; lane 4, *E. coli* harboring pET9d supernatant after heat incubation; lane 5, purified fraction, after anion-exchange chromatography. (**C**). SDS-PAGE analysis of purified *F. pennivorans* keratinase juxtaposed to zymogram staining of caseinolytic activity of *F. pennivorans* crude extract. Lane 6, crude *F. pennivorans* extract; lane 7, purified *F. pennivorans* keratinase.

The N-terminal sequence of the second, 58-kDa polypeptide was as follows: X-Thr-Ala-Arg-Asp-Tyr-Gly-Glu-Glu-Leu-Ser-Asn (X represents an unidentified residue). This matched the amino acid sequence that was expected after removal of the potential 128-residue pro-peptide (Figure 2.3). In addition, the N-termini of the 73-kDa and the 14-kDa band, corresponding to the predicted size of the pro-peptide precursor and the pro-peptide respectively, were determined. Both propeptide proteins displayed an identical N-terminal amino acid sequence: X-Ser-Lys-Ala-Lys-Asp-Leu-Ala-Ser-Leu. The experiments though surprisingly revealed that the *E. coli*-produced fervidolysin lacked the first 8 N-terminal amino acids of the propeptide (including the additional methionine), when compared to the Fls sequence (Figure 2.3). Regarding the mature fervidolysin, exactly the same N-terminal sequence was found for the keratinolytic protease that was purified from the toga of *F. pennivorans* by detergent treatment and SDS-PAGE (Friedrich and Antranikian 1996). Since initial sequence attempts failed, this batch of purified protease was first cleaved with thermolysin prior to Edman degradation. To analyze and compare the molecular mass of both the *E. coli*- and the *F. pennivorans*-produced enzyme, a zymogram staining was carried out (Figure 2.4C). SDS-PAGE analysis of crude cell extract from *F. pennivorans* revealed a clear casein-hydrolyzing band with an apparent molecular weight of around 90 kDa (Figure 2.4C, lane 6), corresponding to the most recently purified native keratinase (Figure 2.4C, lane 7). The bands in the cell-free extracts of fervidolysin expressed in *E. coli* clearly migrated faster in the gel, demonstrating an apparent difference in molecular mass of about 30 kDa. To examine whether this difference was caused by

glycosylation, the purified keratinase from *F. pennivorans* was subjected to a periodic acid-Schiff staining. However, no signs of glycosylation were found on the *F. pennivorans* purified keratinase. Alternatively, the keratinase was incubated with N-glycanase, a glycosidase that hydrolyzes the linkage between the glycan and an asparagine residue. Subsequently, its molecular mass was compared to its untreated state. Although the *N*-glycosylation motif Asn-X-(Ser/Thr) (Marshall 1972) occurs 7 times in the mature protease sequence, the results showed that there was no difference in mass before and after N-glycanase treatment, as judged from SDS-PAGE analysis (not shown).

Site-directed mutagenesis of the F. pennivorans fervidolysin

To determine whether the appearance of the two lowest bands on the SDS-PAGE was a result of either autoprocessing or proteolytic activity by the host organism, an active-site mutant was constructed in plasmid pLUW761. Following the overlap extension method, the active-site residue histidine at position 79 (Figure 2.3) was changed into an alanine residue by creating a *Nae*I restriction site at this position. The construct was called pLUW764 (Figure 2.1). The effect on the expression level was determined by analyzing heat-incubated cell-free extract on an SDS-PAGE (Figure 2.5). The 58-kDa and 14-kDa band, which were present in the cell-free extract of the wild-type construct, were absent in the extract obtained from an *E.coli* BL21(DE3) culture, containing the mutant construct.

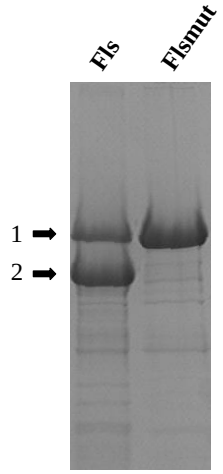


Figure 2.5 SDS-PAGE of the purified recombinant fervidolysin and the active-site mutant. Numbers adjacent to the gel indicate the type of polypeptide: 1, fervidolysin precursor (73 kDa); 2, mature fervidolysin (58 kDa) (pro-peptide not shown in this gel).

Activity of recombinant fervidolysin

The characterization of the keratinase, isolated from a feather-degrading culture of *F. pennivorans*, was described previously (Friedrich and Antranikian 1996). Unfortunately, we were not able to detect significant activity of the purified recombinant fervidolysin when using dye-labeled substrates (collagen, casein), chromogenic substrates, casein hydrolysate and keratin powder (results not shown). Attempts to initiate activity by denaturing/renaturing techniques using 6 M urea were also not successful. Alternatively, we attempted to functionally express the *fls* gene in *Bacillus subtilis* BCL1050 (Dartois et al. 1994), using a customized pBCH3 vector (Dartois et al. 1994; Stanssens et al. 1989). Although the gene was cloned successfully, as was confirmed by nucleotide sequencing and restriction analysis, neither intracellular nor extracellular production of a new protease was detected, and no activity was observed with casein as a substrate (results not shown).

Homology modeling

Fervidolysin was modeled against the crystal structure of subtilisin BPN from *Bacillus amyloliquefaciens*, of which the structure and coordinates are known (Bott et al. 1988). A schematic model of the substrate-binding region is shown in Figure 2.6. The substrate specificity is determined by the interaction between side chain residues of the substrates and the enzyme residues at the binding sites (Grøn et al. 1992).

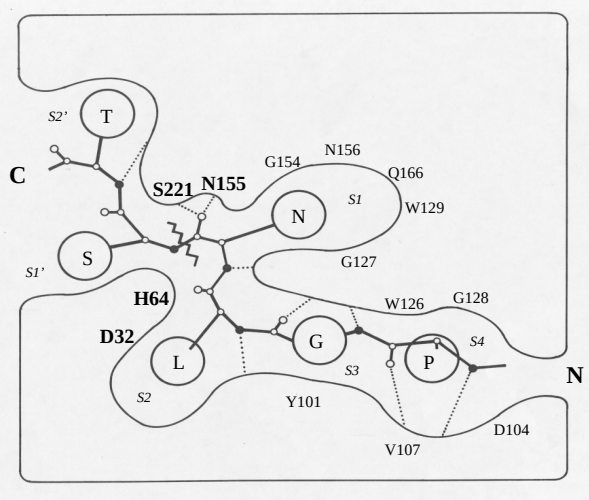


Figure 2.6 Schematic representation of the substrate binding region of fervidolysin. The catalytic residues (D32, H64 and S221) and the oxyanion (N155) are marked in boldface (Subtilisin BPN annotation). C and N denote the C-terminus and the N-terminus of the substrate, respectively. The binding subsite nomenclature *S4-S2'* is according to Schechter and Berger (Schechter and Berger 1967). The pro-peptide of the fervidolysin precursor is modeled in the substrate-binding site of the enzyme. A zigzag line indicates the cleavage site after the P1 position. Substrate backbone atoms are shown as solid spheres (nitrogen), small open spheres (carbon) and large open spheres (oxygen). Potential backbone hydrogen bonds between the pro-peptide and fervidolysin are dotted.

In general, the pockets *S1* and *S4* are the most distinctive in subtilases. Modeling of the cleavage site of the pro-peptide in these binding sites of fervidolysin revealed that the actual residues involved in this subsite interaction did not distinguish for a clear substrate preference. The *S1* pocket remains neutral of charge and average in its size, leaving room for several amino acid residues at position P1, including the asparagine in the fervidolysin pro-peptide. Also, the slightly hydrophobic *S4* pocket does not discriminate for specific residues at position P4, leaving space for the P4 proline in subsite *S4*. The 3D-modeling experiments resulted in the observation of a weak Ca^{2+}-binding site (result not shown).

Discussion

We have identified the active-site region of a subtilisin-like serine protease from *F. pennivorans* via PCR. Subtilases are characterized by conservation around the three catalytic residues aspartate, histidine and serine (Figure 2.3), and the oxyanion-hole asparagine (Siezen and Leunissen 1997). Southern hybridization and inverse PCR respectively revealed an open reading frame of 2.1 kb, flanked by regions that can be recognized as cis-elements, such as a Shine-Dalgarno sequence and a transcription terminator.

The deduced primary sequence shows high homology with the subtilisin-like serine proteases, having the highest degree of similarity with a subtilase from *D. radiodurans* (White et al. 1999). Following the Von Heijne rules (von Heijne 1986), a signal sequence of 21 amino acids could be recognized, which is *in vivo* essential for initiating the export of the protein towards the cell wall. The signal peptide is followed by a pro-peptide, which is the domain that mediates the correct folding of the protein into a biologically active state (Ikemura and Inouye 1988; Ikemura et al. 1987). The size of the pro-peptide was estimated to be 128 residues, judged from sequence similarities among hydrophobic regions in the pro-peptide (Siezen et al. 1995) and alignment comparisons (Siezen and Leunissen 1997), and determination of the N-terminus of the mature protease.

The purified fervidolysin, heterologously produced by *E. coli*, did not show activity towards casein, collagen and keratin powder and chromogenic substrates, designed for the determination of the substrate specificity of subtilases. Additional attempts to functionally express the *fls* gene in a *B. subtilis* expression system did not result in an active keratinase either. Comparison of the N-terminal sequence of the 58-kDa polypeptide with that of the thermolysin-treated keratinase indicated that the N-terminus of the *E. coli*-produced protease is identical with the purified keratinase (Figure 2.3A). As a result it could be concluded that the *fls* gene codes for the same keratinolytic protease as described earlier (Friedrich and Antranikian 1996). Since the purified keratinase was pretreated with thermolysin, N-terminus experiments were carried out on fragmented keratinase. Given that the sequence shows complete identity with the first 12 amino acids of fervidolysin, it can be considered as the N-terminus of the enzyme.

In addition, the N-terminal analyses of all three bands show that the occurrence of the mature fervidolysin is a result of autoproteolytic modification of the protease precursor, present as the 73-

kDa-protein band (Figure 2.4B). This autoproteolytic process is necessary for proteins to fold into their active conformation, however not sufficient for activation of the protease. This could imply that post-translational modification performed by the original host organism *F. pennivorans* is required. An alternative explanation could be that the protease improperly folds in both *E. coli* and *B. subtilis*, although the protease remained in the soluble fraction when expressed in *E. coli*. Addition of crude cell extract of *F. pennivorans* to the heterologously produced enzyme did not result in a restoration of activity (unpublished data). Although this might exclude possible involvement of targeting factors, it does not rule out the involvement of chaperones or other factors that act co-translationally.

Zymogram staining combined with SDS-PAGE analysis (Figure 2.4C) revealed that the apparent molecular mass of the biochemically-purified keratinase is about 30 kDa higher than its recombinant counterpart (Figure 2.4B). This difference could be due to *N*-glycosylation. Although the *N*-glycosylation motif Asn-X-(Ser/Thr) (Marshall 1972) occurs 7 times in the mature protease sequence, actual *N*-glycosylation could not be shown by treatment with N-glycanase or periodic acid-Schiff staining. Moreover, the ultimate effect on the protein mass is difficult to estimate, since the degree of glycosylation and the length of the glycan part is hard to predict.

Extracellular proteases are able to become fully active only when their pro-peptide is further degraded (Ikemura and Inouye 1988). N-terminus experiments on the mature protease showed that the cleavage of the pro-peptide, which is the first step towards activity, has taken place. In general, the pro-peptide remains noncovalently bound within the catalytic domain of the subtilase after being cleaved and functions as an inhibitor. This inhibition is then relieved by internal cleavage, after which the enzyme is able to use other substrates (Shinde and Inouye 1995a; Shinde and Inouye 1995b). This phenomenon was extensively described for kexin, a subtilase from *Schizosaccharomyces pombe* (Powner and Davey 1998). In the case of fervidolysin the pro-peptide remains tightly bound during the purification, as judged from SDS-PAGE analyses (Figure 2.4B, lane 5). Therefore the pro-peptide is likely to operate as an inhibitor. Increasing the salt concentration to weaken the pro-peptide/mature protease complex did not give rise to any keratinolytic activity. Fervidolysin contains two sites that may serve as a secondary cleavage site, both towards the C-terminal end of the pro-peptide (Siezen et al. 1995). Although the pro-peptide is lacking 8 amino acids it does not seem to be the result of the additional cleavage step, being required for rendering activity. Perhaps a degradation of the pro-peptide is instigated by external proteolytic enzymes, which might be present *in vivo*. This degradation process has been shown before, although only in eukaryal proteases that are processed at a vacuolar level (Hiraiwa et al. 1997; Naik and Jones 1998). In general both the cleavage and the removal of the pro-peptide is an autocatalytic process (Shinde and Inouye 1995a).

Mutagenesis of the active-site residue histidine into an alanine showed that the proteolytic cleavage of the fervidolysin precursor was a result of autoproteolytic activity, and not caused by proteolysis carried out by the expression host *E. coli*. Obviously, these results illustrate the importance of the catalytic histidine residue. Besides playing a crucial role in substrate catalysis (Perona and Craik 1995), the histidine residue is also involved in the autocatalytic cleavage process,

in which the pro-peptide is removed via an intramolecular autoprocessing mechanism (Shinde and Inouye 1995a). Regarding fervidolysin however, the pro-peptide appears to remain tightly bound to the mature subtilase. This clearly proves that the pro-peptide cleavage is carried out within the pro-fervidolysin molecule itself. Similar results were obtained by Li and coworkers in their work on subtilisin E from *B. subtilis* (Li and Inouye 1994).

The pro-peptide sequence plays an essential role in guiding the correct folding of the protein into an active protease (Ikemura and Inouye 1988; Ikemura et al. 1987). Examination of the N-termini of the pro-protease precursor and the pro-peptide sequence showed that the pro-peptide was lacking 8 amino acids (additional methionine included) compared to the predicted sequence. This might very well be a result of proteolytic activity carried out by the host organism, since the cleavage sequence does not resemble the N-terminal sequence of the mature protease. The prosequences are generally indispensable for the proper folding of proteins (Siezen et al. 1995) and it was shown for Fls that it could not be expressed when the pro-peptide was omitted. The importance of an intact pro-sequence was mentioned earlier by Ikemura and coworkers (Ikemura and Inouye 1988; Ikemura et al. 1987) who described the role of the pro-peptide in the expression of subtilisin E. In addition, Lee and colleagues (Lee et al. 1991) showed that partial deletions in the N-terminal pro-sequence gave rise to an inactive aqualysin I, a thermophilic subtilase from *Thermus aquaticus*, and the precursor protein was found to be unstable in *E. coli*. Chang and coworkers also showed the inactivating effect of a truncated form on the Npr protease of *Streptomyces cacaoi* (Chang et al. 1994). It is however hard to predict whether the eight lacking amino acids at the N-terminal end of the pro-peptide mainly results in an incorrect folding of fervidolysin, or that this deletion blocks further processing.

Preliminary experiments with the biochemically purified keratinase on chromogenic substrates showed that the enzyme prefers to cut between acidic amino acids rather than alkaline amino acids. While the substrate specificity for keratinases is well defined, not much is known about the amino acid sequence they prefer to cleave.

Although we were not able to produce fervidolysin in an active form, it still remains an interesting, industrial enzyme, considering its thermal origin and its substrate preference. Possibly, expression in a host that is phylogenetically closely related to *F. pennivorans*, such as *T. maritima* or even the original host itself, might be a better solution for the production of the active form of this enzyme.

Acknowledgments

We are grateful to Dr. Schiltz (Institute for Organic Chemistry and Biochemistry, Freiburg) and Dr. Hippe (ChromaTec, Greifswald) for analyzing the N-termini. We also thank Bernadette Renckens for assistance in the 3D-modeling of the subtilase. Finally, we thank Gijs Gerritse from Genencor in Leiden for assistance in the *B. subtilis* expression work. This work was partly funded by contract BIOT-CT96-0488 of the European Union.

References

Altschul SF, Madden TL, Schäffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402

Barrett A, Rawlings N (1995) Families and clans of serine peptidases. Arch Biochem Biophys 318:247-250

Böckle B, Galunsky B, Müller R (1995) Characterization of a keratinolytic serine proteinase from *Streptomyces pactum* DSM 40530. Appl Environ Microbiol 61:3705-3710

Bott R, Ultsch M, Kossiakoff A, Graycar T, Katz B, Power S (1988) The three-dimensional structure of *Bacillus amyloliquefaciens* subtilisin at 1.8 Å and an analysis of the structural consequences of peroxide inactivation. J Biol Chem 263:7895-7906

Bressollier P, Letourneau F, Urdaci M, Verneuil B (1999) Purification and characterization of a keratinolytic serine proteinase from *Streptomyces albidoflavus*. Appl Environ Microbiol 65:2570-2576

Brooks B, Bruccoleri R, Olafson B, States D, Swaninathan S, Karplus M (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4:187-217

Chang S-C, Chang P-C, Lee Y-HW (1994) The roles of propeptide in maturation and secretion of Npr protease from *Streptomyces*. J Biol Chem 269:3548-3554

Choi IG, Bang WG, Kim SH, Yu YG (1999) Extremely thermostable serine-type protease from *Aquifex pyrophilus*. Molecular cloning, expression, and characterization. J Biol Chem 274:881-888

Dartois V, Coppée J-Y, Colson C, Baulard A (1994) Genetic analysis and overexpression of lipolytic activity in *Bacillus subtilis*. Appl Environ Microbiol 60:1670-1673

de Vos W, Voorhorst W, Dijkgraaf M, Kluskens L, van der Oost J, Siezen R (2001) Purification, characterization, and molecular modelling of pyrolysin and other extracellular thermostable serine proteases from hyperthermophilic microorganisms. Meth Enzymol, 330:383-393

Friedrich AB, Antranikian G (1996) Keratin degradation by *Fervidobacterium pennavorans*, a novel thermophilic anaerobic species of the order *Thermotogales*. Appl Environ Microbiol 62:2875-2882

Gibson TJ (1984) Studies on the Epstein-Barr-Virus genome. Cambridge University, Cambridge

Grøn H, Meldal M, Breddam K (1992) Extensive comparison of the substrate preferences of two subtilisins as determined with peptide substrates which are based on the principle of intramolecular quenching. Biochemistry 31:6011-6018

Higuchi R, Krummel B, Saiki RK (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res 16:7351-7367

Hiraiwa N, Kondo M, Nishimura M, Hara-Nishimura I (1997) An aspartic endopeptidase is involved in the breakdown of propeptides of storage proteins in protein-storage vacuoles of plants. Eur J Biochem 246:133-141

Ignatova Z, Gousterova A, Spassov G, Nedkov P (1999) Isolation and partial characterisation of extracellular keratinase from a wool degrading thermophilic actinomycete strain *Thermoactinomyces candidus*. Can J Microbiol 45:217-222

Ikemura H, Inouye M (1988) In vitro processing of pro-subtilisin produced in *Escherichia coli*. J Biol Chem 263:12959-12963

Ikemura H, Takagi H, Inouye M (1987) Requirement of pro-sequence for the production of active subtilisin E in *Escherichia coli*. J Biol Chem 262:7859-7864

Jang JS, Kang DO, Chun MJ, Byun SM (1992) Molecular cloning of a subtilisin J gene from *Bacillus stearothermophilus* and its expression in *Bacillus subtilis*. Biochem Biophys Res Commun 184:277-282

Klingeberg M, Galunsky B, Sjøholm C, Kasche V, Antranikian G (1995) Purification and properties of highly thermostable, SDS resistant and stereospecific proteinase from the extreme thermophilic archaeon *Thermococcus stetteri*. Appl Environ Microbiol 61:3098-3104

Kunert J (1972) Keratin decomposition by dermatophytes: evidence of the sulphitolysis of the protein. Experientia 28:1025-1026

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685

Lee Y, Miyata Y, Terada I, Ohta T, Matsuzawa H (1991) Involvement of NH2-terminal pro-sequence in the production of active aqualysin I (a thermophilic serine protease) in *Escherichia coli*. Agric Biol Chem 55:3027-3032

Li Y, Inouye M (1994) Autoprocessing of prothiolsubtilisin E in which active-site serine 221 is altered to cysteine. J Biol Chem 269:4169-4174

Lin X, Kelemen DW, Miller ES, Shih JCH (1995) Nucleotide sequence and expression of kerA, the gene encoding a keratinolytic protease of *Bacillus licheniformis* PWD-1. Appl Environ Microbiol 61:1469-1474

Marshall R (1972) Glycoproteins. Ann Rev Biochem 41:673-703

Mukhopadhyay RP, Chandra AL (1990) Keratinase of a streptomycete. Indian J Exp Biol 28:575-577

Naik RR, Jones EW (1998) The PBN1 gene of *Saccharomyces cerevisiae*: an essential gene that is required for the post-translational processing of the protease B precursor. Genetics 149:1277-1292

Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson LD, Nelson WC, Ketchum KA, McDonald L, Utterback TR, Malek JA, Linher KD, Garrett MM, Stewart AM, Cotton MD, Pratt MS, Phillips CA, Richardson D, Heidelberg J, Sutton GG, Fleischmann RD, Eisen JA, White O, Salzberg SL, Smith HO, Venter JC, Fraser CM (1999) Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. Nature 399:323-329

Niehaus F, Bertoldo C, Kähler M, Antranikian G (1999) Extremophiles as a source of novel enzymes for industrial application. Appl Microbiol Biotechnol 51:711-729

Noval JJ, Nickerson WJ (1959) Decomposition of native keratin by *Streptomyces fradiae*. J Bacteriol 77:251-263

Perona JJ, Craik CS (1995) Structural basis of substrate specificity in the serine proteases. Protein Sci 4:337-360

Powner D, Davey J (1998) Activation of the kexin from *Schizosaccharomyces pombe* requires internal cleavage of its initially cleaved prosequence. Mol Cell Biol 18:400-408

Ramakrishnan V, Adams MWW (1995) *In*: Archaea, A Laboratory Manual: Thermophiles. In: Robb FT, Place AR (eds) Preparation of genomic DNA from sulfur-dependent hyperthermophilic Archaea. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 95-96

Rawlings N, Barrett A (1994) Families of serine peptidases. Meth Enzymol 244:19-61

Sali A, Blundell T (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779-815

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. In: Nolan C (ed). Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74:5463-5467

Schanstra JP, Rink R, Pries F, Janssen DB (1993) Construction of an expression and site-directed mutagenesis system of haloalkane dehalogenase in *Escherichia coli*. Protein Expr Purif 4:479-489

Schechter I, Berger A (1967) On the size of the active site in proteases. I Papain. Biochem Biophys Res Commun 27:157-162

Shinde U, Inouye M (1995a) Folding mediated by an intramolecular chaperone: autoprocessing pathway of the precursor resolved via a substrate assisted catalysis mechanism. J Mol Biol 247:390-395

Shinde U, Inouye M (1995b) Folding pathway mediated by an intramolecular chaperone: characterization of the structural changes in pro-subtilisin E coincident with autoprocessing. J Mol Biol 252:25-30

Siesenop U, Böhm KH (1995) Comparative studies on keratinase production of *Trichophyton mentagrophytes* strains of animal origin. Mycoses 38:205-209

Siezen R, de Vos W, Leunissen J, Dijkstra B (1991) Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases. Protein Eng 4:719-737

Siezen RJ, Leunissen JAM (1997) Subtilases: The superfamily of subtilisin-like serine proteases. Protein Sci 6:501-523

Siezen RJ, Leunissen JAM, Shinde U (1995) Homology analysis of the propeptides of subtilisin-like serine proteases (subtilases). In: Shinde U, Inouye M (eds) Intramolecular chaperones and protein folding. R.G. Landes Company, Austin, pp 233-255

Stanssens P, Opsomer C, McKeown Y, Kramer W, Zabeau M, Fritz H-J (1989) Efficient oligo-nucleotide-directed construction of mutations in expression vectors by the gapped duplex DNA method using alternating selectable markers. Nucleic Acids Res 17:4441-4454

Takami H, Akiba T, Horikoshi K (1989) Production of extremely thermostable alkaline protease from *Bacillus* sp. no. AH-101. Appl Microbiol Biotechnol 30:120-124

Takiuchi I, Higuchi D, Sei Y, Koga M (1982) Isolation of an extracellular proteinase (keratinase) from *Microsporum canis*. Sabouraudia 20:281-288

Terada I, Kwon ST, Miyata Y, Matsuzawa H, Ohta T (1990) Unique precursor structure of an extracellular protease, aqualysin I, with NH2- and COOH-terminal pro-sequences and its processing in *Escherichia coli*. J Biol Chem 265:6576-6581

Triglia T, Peterson M, Kemp D (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res 16:8186

Varela H, Ferrari MD, Belobrajdic L, Weyrauch R, Loperena L (1996) Effect of medium composition on the production by a new *Bacillus subtilis* isolate of protease with promising unhairing activity. World J Microbiol Biotechnol 12:643-645

Vieille C, Burdette DS, Zeikus JG (1996) Thermozymes. Biotechnology Annual Review 2:1-83

Völkl P, Markiewicz P, Stetter KO, Miller JH (1994) The sequence of a subtilisin-type protease (aerolysin) from the hyperthermophilic archaeum *Pyrobaculum aerophilum* reveals sites important to thermostability. Protein Sci 3:1329-1340

von Heijne G (1985) Signal sequences. The limits of variation. J Mol Biol 184:99-105

von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14:4683-4690

von Heijne G (1989) The structure of signal peptides from bacterial lipoproteins. Protein Eng 2:531-534

Voorhorst W, Warner A, de Vos W, Siezen R (1997) Homology modelling of two subtilisin-like serine proteases from the hyperthermophilic archaea *Pyrococcus furiosus* and *Thermococcus stetteri*. Protein Eng 10:905-914

Voorhorst WGB, Eggen RIL, Geerling ACM, Platteeuw C, Siezen RJ, de Vos WM (1996) Isolation and characterization of the hyperthermostable serine protease, pyrolysin, and its gene from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Biol Chem 271:20426-20431

Wandersman C (1989) Secretion, processing and activation of bacterial extracellular proteases. Mol Microbiol 3:1825-1831

White O, Eisen JA, Heidelberg JF, Hickey EK, Peterson JD, Dodson RJ, Haft DH, Gwinn ML, Nelson WC, Richardson DL, Moffat KS, Qin HY, Jiang LX, Pamphile W, Crosby M, Shen M, Vamathevan JJ, Lam P, McDonald L, Utterback T, Zalewski C, Makarova KS, Aravind L, Daly MJ, Minton KW, Fleischmann RD, Ketchum KA, Nelson KE, Salzberg S, Smith HO, Venter JC, Fraser CM (1999) Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. Science 286:1571-1577

Zacharius RM, Zell TE, Morrison JH, Woodlock JJ (1969) Glycoprotein staining following electrophoresis on acrylamide gels. Anal Biochem 30:148-152

3

Crystal structure of fervidolysin from *Fervidobacterium pennivorans*, a keratinolytic enzyme related to subtilisin

Jeong-Sun Kim, Leon D. Kluskens, Willem M. de Vos, Robert Huber and John van der Oost

J. Mol. Biol., in press (2003)

Abstract

Structure-forming fibrous proteins like keratins, gelatins and collagens are degraded only by few proteases as their tight packing limits access to the potential cleavage sites. To understand the keratin degradation in detail, we describe the first crystal structure of a keratin-degrading enzyme (keratinase), fervidolysin, from *Fervidobacterium pennivorans* as an immature form with propeptide (PD) bound. The 1.7 Å resolution crystal structure shows that the protease is composed of four domains: a catalytic domain (CD), two β-sandwich domains (SDs), and the PD domain. A structural alignment shows a distant relationship between the PD-CD substructure of fervidolysin and pro-subtilisin E. Tight binding of PD to the remaining part of the protease is mediated by hydrogen bonds along the domain surfaces and around the active cleft, and by the clamps to SD1 and SD2. The crystal structure of this multi-domain protein fervidolysin provides insights into proenzyme activation and the role of non-catalytic domains, suggesting a functional relationship to the fibronectin (FN)-like domains of the human promatrix metalloprotease-2 (proMMP-2) that degrades the fibrous polymeric substrate gelatin.

Introduction

Keratin is a major component in keratinous tissue (e.g. hair, feathers) serving as a mechanical support by formation of stiff bundles of fibers. In the waste products of agricultural livestock and poultry activities, more than ten thousand tons of insoluble keratin polypeptide is produced annually worldwide. It is classified as α- and β-keratins based on its highly regular, repetitive secondary structure components. The tightly packed polypeptide chains are further stabilized by numerous disulfide bridges, resulting in polypeptide fibers extremely resistant to common proteases.[1] A limited number of microorganisms have the capacity to use keratin as a substrate for growth by secreted specific keratinases that are able to degrade the keratin bundles. Different types of keratinases have been described, ranging in size from 16 to 440 kDa and with distinct catalytic centers: serine proteases[2, 3] and metalloproteases.[4-6] Keratinases have potential in biotechnological applications like, for example, the conversion of keratin into peptides and rare amino acids.[7-9] Moreover, these proteases deserve attention because of their role in skin physiology, both formation and degradation.[10-12]

In recent years, two novel keratinolytic enzymes have been characterized from the thermophilic *Fervidobacterium spp.*, bacteria originating from hot springs.[13-15] Fervidolysin from *Fervidobacterium pennivorans* is an extracellular subtilisin-like serine protease, which is composed of a signal peptide and a proteolytic part (a PD and a catalytic region). The proteolytic part (73 kDa) was cloned and characterized, and the identity of one of its predicted active site residues was verified by site-directed mutagenesis. Even though the purified recombinant protein partially processed its PD and its N-terminus corresponds to the secreted mature keratinase that was directly isolated from the organism[13], it was inactive.[14]

To date, no detailed structural information on any keratinase exists. We here present the crystal structure of fervidolysin, a subtilisin-like keratinase, as a proenzyme form at 1.7 Å resolution,

which provides insight into the activation mechanism and the mode of action of this important class of enzymes.

Materials and methods

Cloning and expression of the proenzyme form of fervidolysin

The wild-type construct cloned into the pET9d expression vector without the signal peptide yielded only mixtures of the PD-unprocessed protein, the processed PD and the mature part.[14] Furthermore, it was not possible to overexpress the mature form in the *E.coli*. For further work, the enzyme was produced in a mutant form in which His208 was replaced by alanine. This mutant construct was transformed into *E.coli* strain BL21(DE3) for expression. Cells were grown at 310 K in LB with kanamycin and harvested after 5 hours induction at 310 K.

Expression of selenomethionine-incorporated fervidolysin

The seleno-L-methionine (Se-Met) isoform of fervidolysin was expressed in *E.coli* B834(DE3). A single colony of *E.coli* B834(DE3) carrying the expression vector, pET9d-fervidolysin (H208A), was used to inoculate 100 ml of New Minimal Media (NMM) with normal methionine.[35] This culture was grown at 310 K until the exponential phase was reached. Bacterial cells were harvested by centrifugation. The pellet was washed two times by and resuspended in 10 ml of NMM with Se-Met. One-liter culture of supplemented NMM containing Se-Met was inoculated with 10 ml of bacteria suspension and grown at 310 K. Protein expression was induced at 310 K. After growing overnight, cells were harvested by pelleting.

Purification and crystallization

Expressed protein was purified by applying the heat treatment and sequential chromatographic steps, respectively. Harvested cells were resuspended into the buffer A (20 mM Tris·HCl, pH 8.0 and 2 mM β-mercaptoethanol), and disrupted by sonication. Heat treatment at 343 K was done against supernatant for 30 min. Denatured proteins were removed by centrifugation. The cleared solution was applied onto the Q-Sepharose column and protein was eluted by NaCl gradient from 0 to 1 M in buffer A. Protease was pooled against the buffer B (buffer A and 100 mM NaCl), concentrated and loaded onto the Superose 12 gel filtration column.

The purified fervidolysin was concentrated to 15 mg/ml in buffer B for the crystallization. Initial screening was performed at 293 K by vapor-diffusion using Cryschem sitting drop plates, in which 2 μl of native protein was mixed with equal volume of reservoir solutions. Thin plate-clustered crystals were obtained in 20 % polyethyleneglycol 4000, 10 % isopropanol and 0.1 M Tris·HCl, pH 8.5. Slight modification of this initial condition gave crystals suitable for the X-ray diffraction within 2 weeks with a maximum dimension of $0.1 \times 0.1 \times 0.1$ mm^3.

Table 3.1 Data collection and refinement statistics

Parameters	L1	L2	L3	L4
Wavelength (Å)	0.9794	0.97855	0.9790	0.9500
Space group	P1			
Cell parameters	49.59 Å, 50.99 Å, 79.20 Å, 93.90 °, 104.78 °, 115.40 °			
Resolution (Å) (last shell)	20.0-1.70 (1.73-1.70)	20.0-2.50 (2.54-2.50)	20.0-2.50 (2.54-2.50)	20.0-2.50 (2.54-2.50)
Completeness (%) overall/last shell	94.1/91.1	98.4/96.4	98.4/96.5	98.3/96.2
R*sym*[1] (%) overall/last shell	5.0/43.2	5.2/8.3	4.4/7.8	4.5/6.9
Reflections observed/unique	216842/68665	89259/22608	88653/22584	88876/22569
I/Sigma	12.2	20.4	20.2	21.7
Figure of merit[2] Solve/Resolve	0.61/0.74			
R_{factor}[3] (%)	19.1 (24.8)			
R_{free}[4] (%)	21.7 (25.6)			
No of atoms Protein/water/Ca	5013/568/1			
Rmsd				
bonds (Å)	0.007			
angles (°)	1.39			
Geometry (%)				
most favored	90.4			
additionally allowed	9.6			

rmsd: root-mean-square-deviation

[1] $Rsym = \Sigma_{hkl} \Sigma_j |I_j - \langle I \rangle| / \Sigma_{hkl} \Sigma_j I_j$, where $\langle I \rangle$ is the mean intensity of reflection *hkl*.

[2] Figure of merit = $|\sum P(\alpha) e^{i\alpha} / \sum P(\alpha)|$, where $P(\alpha)$ is the phase probability distribution and α is the phase (20.0-2.5 Å).

[3] $R_{factor} = \Sigma_{hkl} ||Fobs| - |Fcalc|| / \Sigma_{hkl} |Fobs|$; where Fobs and Fcalc are respectively, the observed and calculated structure factor amplitude for reflections hkl included in the refinement.

[4] R_{free} is the same as R_{factor} but calculated over a randomly selected fraction (9.3 %) of reflection data not included in the refinement.

Data collection and structure determination

Crystals belonged to the space group P1 with unit cell dimensions of a = 49.59 Å, b = 50.99 Å, c = 79.20 Å, α = 93.90 °, β = 104.78 °, γ = 115.40 °, and they contain one molecule in the asymmetric unit. Before cryocooling, crystals were briefly immersed in the same precipitant solution containing 10–15 % glycerol. All diffraction data were collected at 100 K on the wiggler beamline BW6 of DORIS (DESY Hamburg, Germany). An MAD data set was measured at four different wavelengths comprising the Se K-absorption peak, inflection points as well as two remote points and was processed, merged by DENZO and SCALEPACK package.[36] Ten selen sites of 11 were identified using the program SOLVE and phases were improved by density modification with the program RESOLVE[37]. All residues except some in the disordered regions were built using program O[38] with the initial electron density map calculated from the MAD data set and refined by CNS package.[39] Water molecules were automatically added and optimized further during the final cycles of model inspection. The quality of the final model was analyzed with PROCHECK[16] and is summarized in the Table 3.1. The coordinates were deposited with the Protein Data Bank (PDB accession No. 1R6V).

Results

Overall structure

The crystal structure was elucidated by multi-wavelength anomalous dispersion (MAD) phasing with selenium as an anomalous scatterer. The overall three-dimensional structure of fervidolysin has an elongated shape with approximate dimensions of 100 × 50 × 36 $Å^3$ (Figure 3.1). Except for the N-terminal 3 residues, residues Ile123-Ala132 of the linker region between PD and CD, residues Asn305-His312 of the CD, and Asp518-Ser527 of the SD1, 640 amino acids out of 671 were defined by electron density. The final model consists of four independent domains. An N-terminal PD (Ser9-Asn128) is characterized by a globular structure with extended N- and C-terminal segments running in opposite directions. The central CD (Ser129-Pro448) is similar to a subtilisin-like serine protease with an αβ mixed folding. The two C-terminal SDs (Gly454-Asn563 and Thr565-Gln679 for the SD1 and the SD2, respectively) have β-sandwich folds (Figure 3.1A). Analysis of the Ramachandran plot by PROCHECK[16] showed that all modeled residues for the protein were found either in most favored regions (90.4 %) or in additionally allowed regions (9.6 %) (Table 3.1).

The PD domain

The PD of fervidolysin forms a globular domain consisting of a four-stranded antiparallel β-sheet flanked by two short two-turn helices with an extra strand parallel to the helices (Figure 3.2A). The globular region is stabilized by hydrophobic interactions and by a number of hydrogen bonds. Hydrophobic residues Ile27, Phe29, Ile42, Val44, Ile55, Val59, Ile75, Ile89, Ile97, and Val100 form

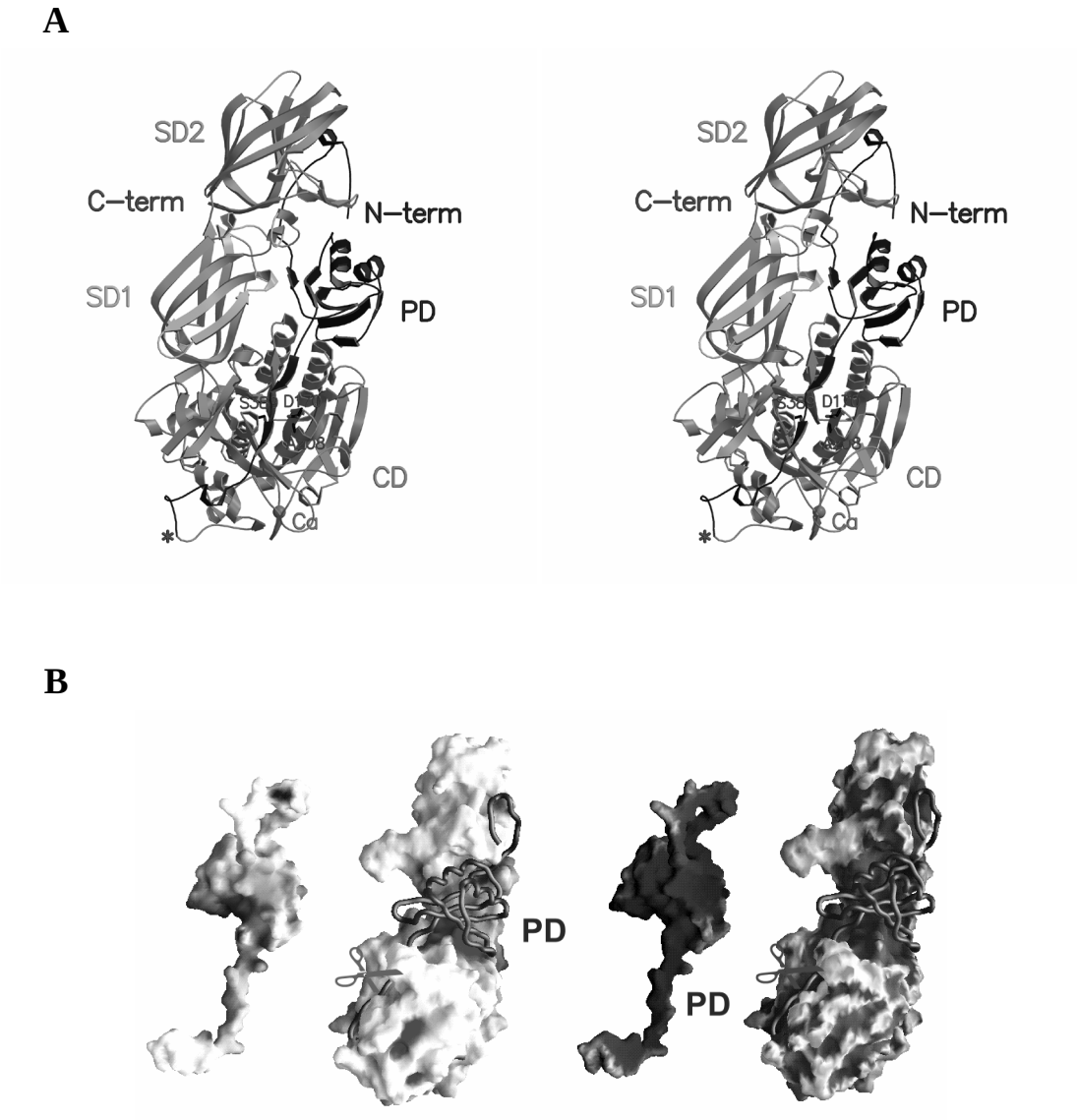


Figure 3.1 (**A**) Stereopresentation of the overall shape of fervidolysin. The PD, the CD and two SDs are shown in alternating colors and labeled. The PD-processing site at the boundary of PD and CD is indicated by red asterisks. The active site residues were drawn by stick models, calcium-binding sites by balls, and labeled, respectively. The figures were prepared by Molscript and rendered with Raster3D.[40-41] (**B**) The mature part of fervidolysin embraces the PD, which fits into the curved surface of CD. While the mature part was drawn by the potential surface (left) and by the surface curvature (right), the PD was drawn by coil. The positively charged region of right surface representation was colored in blue and the negatively charged one in red. The active site was marked by scissors. The molecules are rotated by 90° compared with Figure 3.1A. The figures were prepared by GRASP.[42]

the core of this module. The central structural motif of two helices walled by four β-strands is similar to the elongation factor G, [17] the PD of subtilisin E-PD[18] and the PD of procarboxypeptidase A[19] as the first representative of those structures (Figure 3.2A). Whereas the N-terminal extended region forms a one-turn helix and interacts with both SDs, the C-terminal part spans the whole active site of and is connected to the CD (Figure 3.1).

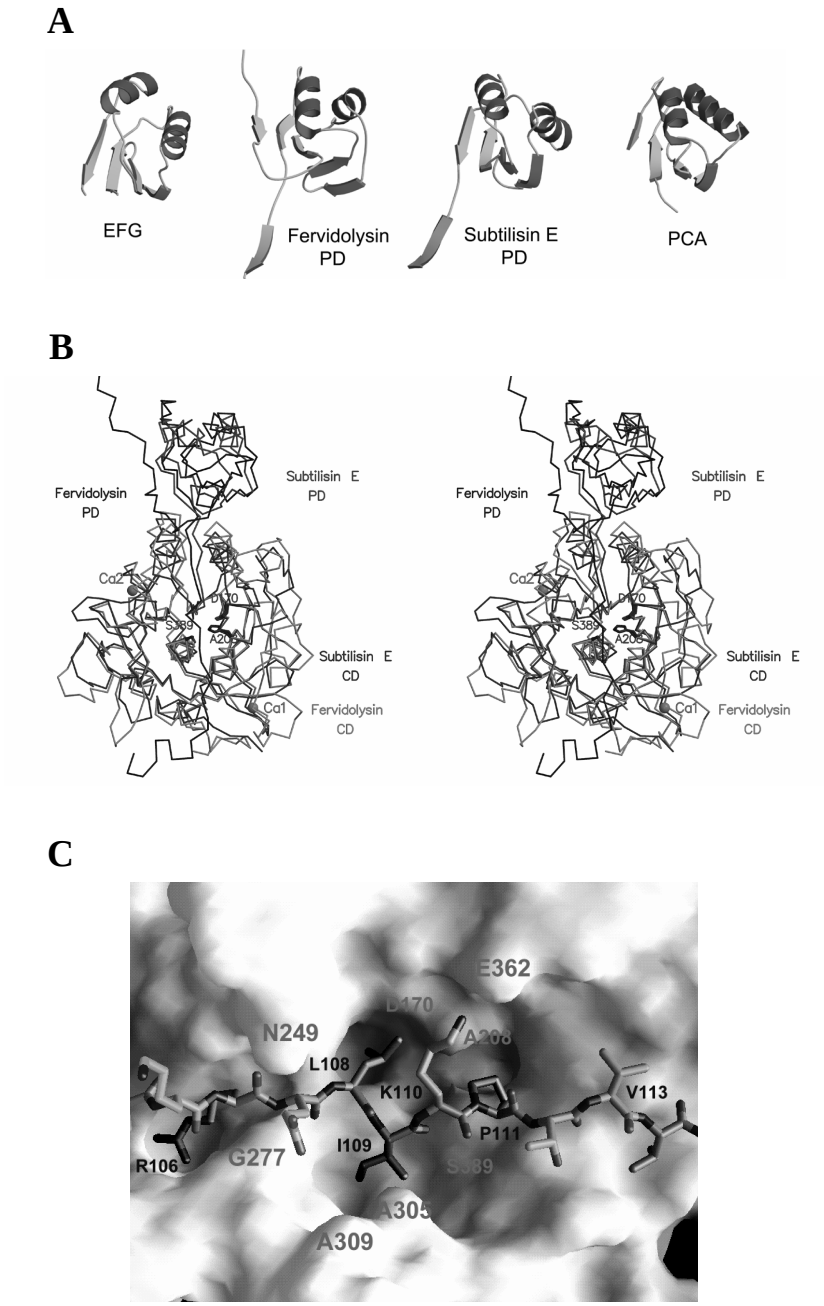


Figure 3.2 (**A**) Comparison of PD with the elongation factor G (EFG, PDB 1DAR), with the PD of subtilisin E-PD complex (subtilisin E, PDB 1SCJ), and with procarboxypeptidase A (PCA, PDB 1PCA). The fervidolysin PD has stretched and elongated regions on both termini. (**B**) Stereopresentation of PD-CD subcomplex of fervidolysin superposed with subtilisin E-PD complex. The structures of the PD and CD of fervidolysin superpose well with the corresponding domains of the subtilisin E-PD complex including the catalytic residues. The labeling of catalytic residues appears only on fervidolysin. The calcium ions were taken from subtilisin E and presented by balls with labels of Ca1 and Ca2. (**C**) Presentation of PD binding at the active site. The PD is displayed as a stick model and the CD by the electrostatic potential surface. The C-terminal arm of the PD crosses and binds to the active site intercalating between two short anti-parallel β-strands of CD (Arg106-Leu108 interacting with Ser275-Gly277 and Thr112-Val114 with Tyr384-Gln386). This binding mode probably defines the substrate binding at the unprimed and the primed sites, respectively. Arg106 of the PD appears to define a specific binding site by interacting with the Gly280 and the Asp284 using its Nε, and with the Asn249 using its carbonyl oxygen.

The CD domain

The architecture of CD closely resembles that of a subtilisin indicated also by the high amino acid sequence conservation in this part (41.7 %), except for some insertions: Phe192-Asp201, Ala218-Val226, Asp242-Gly250, Ala307-Gln313, and Glu362-Gly379. The subcomplex structure of CD and PD compares well with that of the subtilisin E-PD complex, where the catalytic Ser221 had been mutated to Cys221 to block maturation (Figure 3.2B).[18] The catalytic triad in the wild type of fervidolysin consists of Ser389, His208, and Asp170 and corresponds to Ser221, His64, and Asp32 in subtilisin (Figure 3.2B). In fervidolysin, His208 has been mutated to Ala208, in order to block autoproteolysis. A water molecule occupies the imidazole side chain position in the mutant. Probably as a consequence of this mutation, the Ser389 adopts an alternate rotamer conformation, compared with subtilisin. Like the subtilisin E-PD complex, two nearly parallel α-helices (residues Gly253-Gly268 and Ser282-His295) on the surface of CD form the main site of interaction with the PD (Figure 3.2B). Two insertions, Asp242-Gly250 and Glu362-Gly379 in fervidolysin, invoke local structural changes at the binding cleft, compared with the subtilisin E-PD complex. Two peptide nitrogen atoms of Thr388 and Ser389 form the oxyanion hole, where a water molecule mimics the carbonyl oxygen of the scissile peptide bond. The S2 pocket (nomenclature compared to subtilisin) is constructed by residues Asp170-Thr171, Tyr204-Gly206 and Ile239-Asp242, but the S1 pocket is only partially defined by Ser275-Gly277, Ser302-His312 (Ala312) and Thr388 because of the flexible loop of Asn305-His312 with a weak electron density. The S1' pocket is formed by the residues of Ser207, His208 (Ala208), Tyr385 and Glu368, and capped by the inserted loop of Glu362-Gly379 (Figure 3.2B). Each subsite is occupied by Leu108 of PD, Ile109, Lys110, and Pro111 from S2 to S2', respectively. However, the scissile peptide bond Ile109-Lys110 is bound in a distorted manner and the adjacent interactions may not exactly model a substrate (Figure 3.2C).

A calcium-binding site was observed within CD (Figures 3.1A, 3.2B), which exactly matches that of the subtilisin E-PD complex.[18] The calcium ion is coordinated by six oxygen atoms, three from the side chains of acidic residues (Glu137, Asp179, Asp221) and three from the peptide oxygens (Lys219, Lys223, Ile225) (Figure 3.3A). The second calcium-binding site of subtilisin is not occupied in fervidolysin at the contact region with SD1 (Figure 3.2B), because two ligands of subtilisin E (side chain oxygens of Thr174 and Asp197), which are coordinating the calcium ion together with a water molecule, are not conserved in fervidolysin (Val321 and Ser347). Instead, five water molecules at the interface mediate the tight interactions with the SD1, together with direct polar contacts between residues of the two domains (Figure 3.3B).

The SD1 domain

The SD1 segment emerges from CD and is connected to SD2. It is composed of a pair of β-sheets, which forms a frequently observed β-sandwich fold with four-stranded (strand F, E, B, C) and the three-stranded β-sheets (strand G, A, D) (Figure 3.4A). A hydrophobic core is built by a number of hydrophobic residues, Phe458, Val460, Val462, Val471, Val474, Val476, Met478, Phe498, Ile501, Ile508, Val510, Phe549, Leu558, and Val560. This fold is topologically related to the immuno-

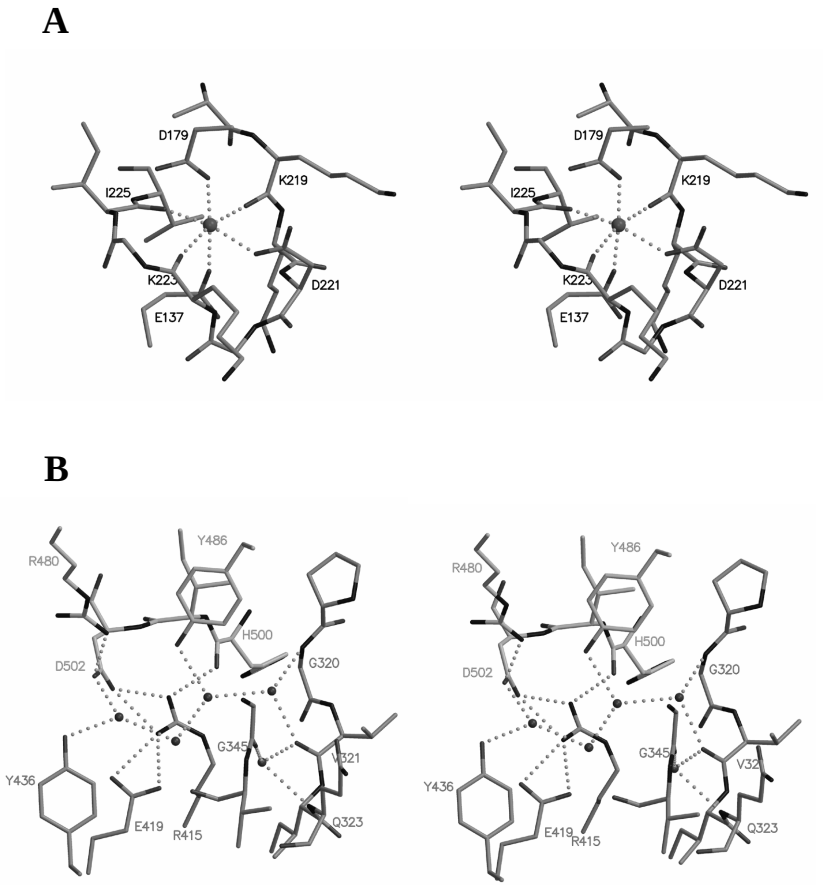


Figure 3.3 (**A**) Stereopresentation of the Ca1-binding site. Calcium ion is hexa-coordinated and forms an octahedral geometry. Peptide oxygen atoms of Lys219, Lys223 and Ile225, and OD1 of Asp221 accomplish a plane while two oxygens from the acidic side chains of Glu137 and Asp179 occupy both axial positions. (**B**) Stereopresentation of the Ca2-binding site. The second calcium-binding site is not conserved in fervidolysin, and a calcium ion is not observed either. Residues of CD and SD1 were differentiated by different colors of carbon atoms. Five water molecules were displayed with red balls and polar interactions, which are mediated by intervening water molecules, displayed by dotted lines, along with direct interaction between two domains.

globulin (Ig) fold seen in the FN domain,[20] in domain II of carboxypeptidase D,[21] in the central SD of protocatechuate 3,4-dioxygenase,[22] and in the Ig constant domains[23], except for the long insertion between strands E and F (Asp518-Ser527, Figure 3.4A) that has weak electron density and was modeled as a polyalanine. When superimposed with other FN-like structures,[20] this loop matches with the RGD sequence-containing loop, while the KGN sequence-containing loop (Lys466, Gly467 and Asn468) is located at the opposite side of this domain (Figure 3.4A). The disulfide bridge observed in the other FN-like structures does not exist in SD1. The interactions with neighboring domains are established mainly by hydrophobic residues, except for some polar interactions with the CD (the Nε of Lys489 with the OH of Tyr281 and the interactions at the second putative calcium-binding site) and the SD2 (the carbonyl oxygen of Ile534 with the Nε2 of Gln612, the Oε2 of Glu540 with the peptide nitrogen of Thr675, the NH1 of Arg537 with the

carbonyl oxygen of Trp673 and the peptide nitrogen of Ala465 with the carbonyl oxygen of Phe677).

The SD2 domain

The C-terminal module SD2 is also made up of a pair of β-sheets with an added long C-terminal region (Asp656-Gln679). This addition forms a β-sheet (strand h, h', h'', figure 3.4B) and generates the contact regions with SD1 and PD (Figure 3.1A). Unlike in SD1, both β-sheets consist of four strands (strand d, a, f, g and c, b, e, g', respectively). A hydrophobic core is constructed by residues of Leu566, Val568, Phe570, Leu574, Leu577, Pro580, Phe582, Val584, Val596, Leu606, Leu609, Ile613, Phe615, Ile617, Ile627, I629, Val633, Leu635, Val642, Leu646, and Ile655.

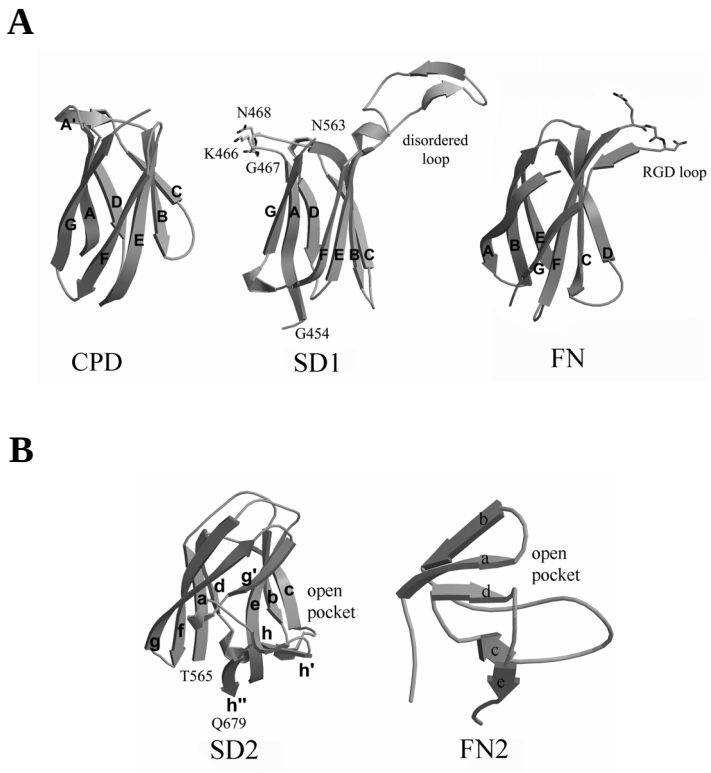


Figure 3.4 Comparison of two SDs with other FN domains. (**A**) The SD1. The β-sandwich folding of the FN-like domain is conserved on the SD1 of fervidolysin except for a long insertion between strand E and strand F. This invisible loop exactly matches the RGD sequence-containing loop of an FN domain (FN, PDB 1FNF), with the short exposed KGN sequence-containing loop located on the opposite direction. But their topologies are not exactly identical. One domain of the carboxypeptidase D (CPD) from duck shows topologically identical folding motif to the SD1 of fervidolysin. However, this domain lacks RGD sequence-containing loop. (**B**) The SD2. Unlike SD1, this domain is composed of two four-stranded β-sheets. The addition of one β-strand to the second sheet separates an additional open hydrophobic pocket from the hydrophobic core, along with the C-terminal adding that forms three β-strands and penetrates the domain interface of the SD1 and the SD2. One of the FN-type II domains of proMMP-2 displays a similar role in PD binding by the formation of an open hydrophobic pocket (FN2, PDB 1CK7).

The central sandwich structure of this domain is similar to SD1 except for a disordered loop in SD1 and for the additional C-terminal β-strands in SD2. The last strand of the second β-sheet (strand g') and the C-terminal addition (strand h, h', h'') in SD2 separates an open hydrophobic pocket from the main hydrophobic core (Figure 3.4B). This extra hydrophobic pocket is formed by contiguous hydrophobic residues of Pro582, Val583, Val585, Leu589, Val592, Ile617, Gln618, Trp652, Ile654, Ile655, and Tyr671 and has a groove towards the PD (Figures 3.1A, 3.5A).

The PD binding

The PD binding to the remaining part of the molecule features hydrophobic interactions and hydrogen bonds. Leu14 and Leu17 from the helix turn of the PD at the N-terminus seal the open hydrophobic pocket of SD2 (Figures 3.1A, 3.5A). In the second hydrophobic pocket by Ala465, Pro469, Val471, Pro513, Arg542 and Phe562 of SD1 and by Ser610, Gln612 and Phe677 of SD2, Tyr25 of PD is trapped and its OH is hydrogen-bonded to the carbonyl oxygen of Ile676 (Figure 3.5B). In addition, Leu32 binds to a hydrophobic niche formed between CD (Tyr283) and SD1 (Pro492 and Thr473). Several hydrogen bonds are also observed along the domain surfaces and around the active site. A PD loop (Pro18-Phe29) is interacting with both SDs by forming six hydrogen bonds, among which three are mediated by water molecules. Like two acidic residues of PD in subtilisin E (Glu63 and Asp65) contributing to binding to CD by capping the peptide nitrogens of two parallel α-helices (Thr133-Ala134 and Qln103-Tyr104),[24] the Glu101 of PD in fervidolysin interacts with two peptide nitrogen atoms (Tyr283-Thr284) of a helix of CD (Ser282-His295).

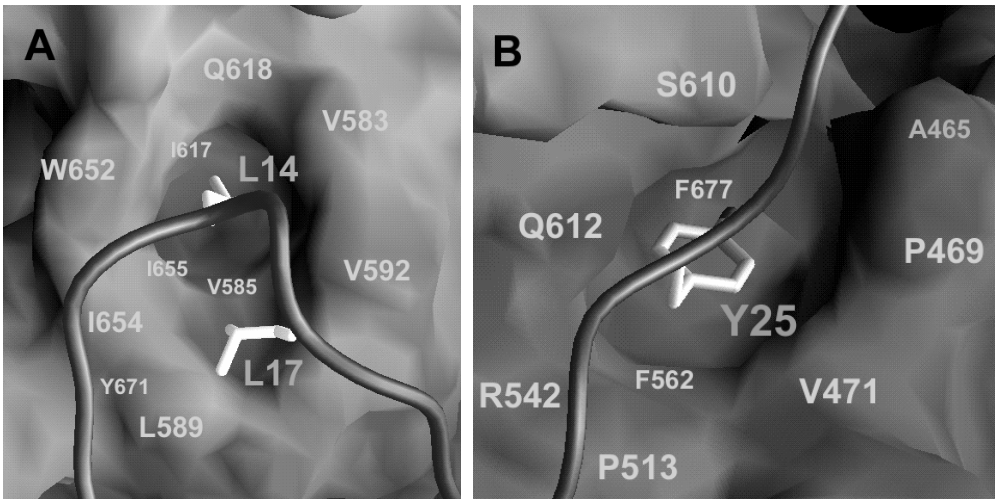


Figure 3.5 Surface presentation of hydrophobic interactions and hydrogen bonds at the active site. (**A**) Leu14 and Leu17 of the PD seal the open hydrophobic pocket of the SD2. The short PD chain carrying both hydrophobic residues is displayed as blue line and the open hydrophobic pocket of SD2 is presented as a solid surface. (**B**) Tyr25 is another hydrophobic anchor at the domain interface of the SD1 (Ala465, Pro469, Val471, Pro513, Arg542, and Phe562) and the SD2 (Ser610, Gln612, and Phe677).

However, the second acidic residue in subtilisin E (Asp65) is not conserved in fervidolysin (Ser103). A straight chain, Pro102-Lys115 of PD, fills the substrate-binding site, where two PD strands (Arg106-Leu108 and Thr112-Val114) interact with Ser275-Gly277 and Tyr384-Gln386 of the CD to form a local β-sheet, respectively. These interactions probably define the general binding mode of a substrate at the unprimed and the primed sites, respectively, with the peptide bond between Ile109 and Lys110 marking the scissile peptide bond (Figure 3.2C). In addition to the main chain hydrogen bonds, side chain interactions are observed in this region. The side chain of P1' residue (Lys110) interacts with the side chains of Glu368 and Ser207. Arg106 inserts its side chain into a pocket of CD, and makes hydrogen bonds with the carbonyl oxygen of Gly280 and with the OD1 of Asp254 (Figure 3.2C). Towards the C-terminal part of PD, Pro116 interacts with Tyr384 and the side chain of Arg133 is hydrogen-bonded to the carbonyl oxygen of Met120, whose side chain again makes hydrophobic interactions with the CD.

Discussion

Tight control is required for proteases that degrade structural proteins, in order to prevent damage to the cells. Commonly, they are secreted as proenzymes and activated by the removal of PD regions through autoproteolysis or by other proteases.[25-26]

A significant structural feature that distinguishes fervidolysin from the other subtilisin-like serine proteases is the presence of two SDs at the C-terminus. The folding of SD2 is unique in that both β-sheets are comprised of four strands. The additional β-strand (strand g') of the second β-sheet and the C-terminal insertion (strands h, h', h") create an open hydrophobic pocket in SD2, which is separated from the hydrophobic core (Figure 3.4B). Similar to the open hydrophobic pocket within a pair of β-sheets of FN2 in the crystal structure of proMMP-2, which contributes to fixing the N-terminal elongated chain of PD by residue Phe37,[27] this open hydrophobic pocket of SD2 holds two hydrophobic residues Leu14 and Leu17 of PD (Figure 3.5A). The PD is additionally anchored to SD2 and SD1 by the hydrophobic pocket at the interface, where Tyr25 of PD binds to the carbonyl oxygen of Ile676 with its OH (Figure 3.5B). Also Leu32 of PD binds in a hydrophobic niche formed by the CD and the SD1.

The two SDs in fervidolysin show structural similarities to the FN or the Ig fold, which is known to be involved in cell adhesion[28-33] or in protein-protein interactions[34] through their RGD sequence-containing loop. Fervidolysin has also been reported to bind to the cell surface.[13] Even though the RGD is lacking in the primary sequences, SD1 has a structurally equivalent long disordered loop (Asp517-Ser527), which may be involved in protein interactions and thereby becoming stabilized. In addition, there is an exposed KGN sequence-containing loop of the SD1 on the opposite site, which may also be the cell-binding motif (Figure 3.4A). We suggest that the role of the SDs is to mediate the interaction with β-keratin and to assist in disassembly of the layers of β-structures so that individual strands may bind to the mature protease, similar to the observed interaction with the C-terminal extended region of PD (Ser103-Lys115) at the active cleft and will subsequently be hydrolyzed.

An overall substrate-binding mode in the catalytic cleft of fervidolysin is suggested from the PD region crossing the whole active site. In the present structure as shown, the C-terminal extended region (Ser103-Lys115) of PD binds to the active cleft by forming a β-sheet with the CD. This interaction is likely to define the substrate binding at the unprimed (Arg106-Leu108) and the primed regions (Thr112-Val114), respectively (Figure 3.2C). Unlike the subtilisin E-PD complex where a β-strand (Val73-Tyr77) of PD intercalates between the two parallel β-strands of Gly100-Gln103 and Ser125-Gly128 of CD at the unprimed site,[18] only one CD strand (Ser275-Gly277) in fervidolysin contributes to the PD binding. The insertion of Asp242-Gly250 into CD of fervidolysin prevents the formation of another strand here (Figure 3.2B). Therefore, the binding cleft at the unprimed site is wider than that of subtilisin and the PD binding in fervidolysin may not be as tight as subtilisin E. But the loss of interactions here is partly compensated by Asn249, which interacts with its ND2 with the carbonyl oxygen of Arg106. On the other hand, the insertion of Glu362-Gly380 in fervidolysin sterically hinders PD binding along the strand of Thr381-Gln386, corresponding to the substrate recognition strand in subtilisin (Lys213-Ser218) at the primed site and caps the S1' pocket. As a result, the PD crosses over this strand with forming three hydrogen bonds. In conclusion, the wider substrate-binding cleft of fervidolysin at the unprimed site is designed to accept bulkier substrates and the cleft at the primed site requires a straight chain. Overall the putative substrate-binding cleft of CD in fervidolysin favors an extended protein chain (Figures 3.1B, 3.2C), in accord with the known preference for β-keratin.[13]

The detailed substrate-binding mode of fervidolysin at the active site can be modeled by comparison with the subtilisin E-PD complex. The catalytic triad and the putative substrate binding pockets are formed in fervidolysin. However, the presence of a bulky side chain at the S2' site (Pro111) rather distorts the overall environment of the active site. Therefore, the PD residues do not fit exactly into each corresponding pocket including the residues of P1 (Ile109) and P1' (Lys110) (Figure 3.2C). Ser62 in the S2 pocket of subtilisin E, which confers the substrate specificity by interacting with a negatively charged side chain (Glu76 of PD),[24] is not conserved in fervidolysin (Gly206). The substitution of Thr171 for Ser39 of subtilisin E in the S2 pocket and the insertion of two residues (Tyr204-Gly205) as well as the inserted loop (Asp241-Gly250) construct a shallow pocket, which is wider and more hydrophobic than that of subtilisin E (Figure 3.2B). The chain Asn305-His312 of fervidolysin with a weak electron density participates in forming an S1 pocket and is connected to the putative second calcium-binding site. However, this chain is shorter than that of subtilisin E (Asn155-Gly166) and might be located more closely to the catalytic residues than the second putative calcium-binding site, to construct a smaller S1 pocket than that of subtilisin E (Figure 3.2B). The S1' pocket in fervidolysin seems to define the enzyme specificity and is designed to accommodate residues with a long hydrophilic side chain, which can interact with Oγ of Ser207 or with Glu368 (Figures 3.2B, C).

A previous molecular study on wild type recombinant fervidolysin revealed that no activity was detectable, even though the PD had been partly processed. On the other hand, an active site mutant of fervidolysin (H208A) showed the unprocessed PD, suggesting that the partial maturation was a result of autocatalytic processing.[14] The structure helps to explain why the cleaved PD in the wild

type enzyme does not dissociate from the remaining part of the protease and how the maturation can be accomplished. Prominent differences in the PD sequences of fervidolysin compared with those of subtilisin E, are the additional N-terminal elongated region (Ser9-Thr38) (Figure 3.2B) and the Arg106 of fervidolysin (Ala74 in subtilisin E-PD complex). The former includes all hydrophobic anchors (Leu14, Leu17, Tyr25 and Leu32) observed in PD binding to the mature protease and the side chain of the latter accomplishes several hydrogen bonds with CD as was described above. Unlike in subtilisin that is activated through the removal of the whole PD by an autolytic single cleavage,[24] the physiological activation of fervidolysin is proposed to be achieved through the stepwise removal of PD. The successive elimination of the PD in fervidolysin is accomplished by several proteolytic cleavages that may not be autolytic.

Feather-degrading keratinases have been found in a limited number of microorganisms and show diversity in molecular weight distributions. In this perspective, fervidolysin (73 kDa) is an enzyme of medium size. The multi-domain arrangements in the crystal structure of fervidolysin suggest that small molecular weight keratinases (less than 35 kDa) may contain only the catalytic part, whereas keratinases of high molecular weight may consist of multi-domain structures or oligomeric proteases, which will contribute to their substrate specificities. Because of their robust nature, and catalysis under conditions that favor substrate solubility, thermostable enzymes are candidates for biotechnological applications. Since fervidolysin is highly thermostable[13-14], the structural information described here might contribute to the design of an efficient protease for the degradation of persistent fibrous polypeptides.

Acknowledgements

We thank Gleb Bourenkov, Hans Bartunik of BW6 at DESY (Hamburg, Germany), and Takashi Tomizaki, Clemens Schulze-Briese of the SLS beamline X06SA at PSI (Villigen, Switzerland) for helping with synchrotron data collection.

References

1. **Parry, D. A. D. & North, A. C. T.** (1998) Hard α-keratin intermediate filament chains: substructure of the N-and C-terminal domains and the predicted structure and function of the C-terminal domains of type I and type II chains. *J. Struct. Biol.* **122** 67–75.
2. **Lin, X., Lee, C. -G., Casale, E. S. & Shih, J. C. H.** (1997). Purification and characterization of a keratinase from a feather-degrading *Bacillus licheniformis* strain. *Appl. Environ. Microbiol.* **58**, 3271-3275.
3. **Evans, K. L., Crowder, J. & Miller, E. S.** (2000). Subtilisins of *Bacillus spp*. hydrolyse keratin and allow growth on feathers. *Can. J. Microbiol.* **46**, 1004-1011.
4. **Brouta, F., Descamps, F., Fett, T., Losson, B., Gerday, C. & Mignon, B.** (2001). Purification and characterization of a 43.5 kDa keratinolytic metalloprotease from *Microsporum canis*. *Med Mycol.* **39**, 269-275.

5. **Allpress, J. D., Mountain, G. & Gowland, P. C.** (2002). Production, purification and characterization of an extracellular keratinase from Lysobacter NCIMB 9497. *Lett. Appl. Microbiol.* **34**, 337-342.
6. **Lee, H., Suh, D. B., Hwang, J. H. & Suh H. J.** (2002). Characterization of a keratinolytic metalloprotease from *Bacillus sp.* SCB-3. *Appl. Biochem. Biotechnol.* **97**, 123-133.
7. **Böckle, B., Galunsky, B. & Müller, R.** (1995). Characterization of a keratinolytic serine protease from *Streptomyces pactum* DSM 40530. *Appl. Environ. Microbiol.* **62**, 3705-3710.
8. **Cheng, S. W., Hu, H. M., Shen, S. W., Takagi, H., Asano, M. & Tsai, Y. C.** (1995). Production and characterization of keratinase of a feather-degrading *Bacillus licheniformis* PWD-1. *Biosci. Biotechnol. Biochem.* **59**, 2239-2243.
9. **Santos, R. M. D. B., Firmino, A. A. P., de S, C. M. & Felix, C. R.** (1996). Keratinolytic activity of *Aspergillus fumigatus fresenius*. *Curr. Microbiol.* **33**, 364-370.
10. **Tsuboi, R., Ko, I. J., Matsuda, K. & Ogawa, H.** (1987). A new keratinolytic proteinase from clinical isolates of *Trichophyton mentagrophytes*. *J. Dermatol.* **14**, 506-508.
11. **Qin, L. M., Dekio, S. & Jidoi, J.** (1992). Some biochemical characteristics of a partially purified extracellular keratinase from *Trichophyton schoenleinii*. *Zentralbl. Bakteriol.* **277**, 236-244.
12. **Mignon, B., Swinnen, M., Bouchara, J. P., Hofinger, M., Nikkels, A. & Pierard, G.** (1998). Purification and characterization of a 315 kDa keratinolytic subtilisin-like serine protease from *Microsporum canis* and evidence of its secretion in naturally infected cats. *Med. Mycol.* **36**, 395-404.
13. **Friedlich, A. B. & Antranikian, G.** (1996). Keratin degradation by *Fervidobacterium pennivorans*, a novel thermophilic anaerobic species of the order Thermotogales. *Appl. Environ. Microbiol.* **62**, 2875-2882.
14. **Kluskens, L. D., Voorhorst, W. G., Siezen, R. J., Schwerdtfeger, R. M., Antranikian, G., van der Oost, J. & de Vos, W. M.** (2002). Molecular characterization of fervidolysin, a subtilisin-like serine protease from the thermophilic bacterium *Fervidobacterium pennivorans*. *Extremophiles* **6**, 185-194.
15. **Nam, G. W., Lee, D. W., Lee, H. S., Lee, N. J., Kim, B. C., Choe, E. A., Hwang, J. K. *et al.*** (2002). Native-feather degradation by *Fervidobacterium islandicum* AW-1, a newly isolated keratinase-producing thermophilic anaerobe. *Arch. Microbiol.* **178**, 538-547.
16. **Laskowski, R., MacArthur, M., Hutchinson, E. & Thorton, J.** (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291.
17. **al-Karadaghi, S., Aevarsson, A., Garber, M., Zheltonosova, J. & Liljas, A.** (1996). The structure of elongation factor G in complex with GDP: conformational flexibility and nucleotide exchange. *Structure* **4**, 555-565.
18. **Jain, S. C., Shinde, U., Li, Y., Inouye, M. & Berman, H. M.** (1998). The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0 Å resolution. *J. Mol. Biol.* **284**, 137-144.
19. **Guasch, A., Coll, M., Aviles, F. & Huber, R.** (1992). Three-dimensional structure of porcine pancreatic procarboxypeptidase A. A comparison of the A and B zymogens and their determinants for inhibition and activation. *J. Mol. Biol.* **224**, 141-157.
20. **Leahy, D. J., Hendrickson, W. A., Aukhil, I. & Erickson, H. P.** (1992). Structure of a fibronectin type III domain from tenasin phased by MAD analysis of the selenomethionyl protein. *Science* **258**, 987-991.
21. **Gomis-Rueth, F. X., Companys, V., Qian, Y., Fricker, L. D., Vendrell, J., Aviles, F. X. & Coll, M.** (1999). Crystal Structure of Avian Carboxypeptidase D Domain II: A Prototype for the Regulatory Metallocarboxypeptidase Subfamily. *EMBO J.* **18**, 5817-5826.

22. **Orville, A. M., Elango, N., Lipscomb, J. D. & Ohlendorf, D. H.** (1997). Structures of competitive inhibitor complexes of protocatechuate 3,4-dioxygenase: multiple exogenous ligand binding orientations within the active site. *Biochemistry* **36**, 10039-10051.
23. **Sondermann, P., Huber, R., Oosthuizen, V. & Jacob, U.** (2000). The 3.2-A crystal structure of the human IgG1 Fc fragment-Fc gamma RIII complex. *Nature* **406**, 267-273.
24. **Gallagher, T., Gilliland, G., Wang, L. & Bryan, P.** (1995). The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'. *Structure* **3(9)**, 907-914.
25. **Birkedal-Hansen, H.** (1995). Proteolytic remodeling of extracellular matrix. *Curr. Opin. Cell. Biol.* **7(5)**, 723-735.
26. **Kohn, E. C. & Liotta, L. A.** (1995). Molecular insights into cancer invasion-strategies for prevention and intervention. *Cancer Res.* **55**, 1856-1862.
27. **Morgunova, E., Tuuttila, A., Bergmann, U., Isupov, M., Lindqvist, Y., Schneider, G., & Tryggvason, K.** (1999). Structure of human pro-matrix metalloproteinase-2: activation mechanism revealed. *Science* **284**, 1667-1670.
28. **Pierschbacher, M. D. & Ruoslahti, E.** (1984). Varients of the cell recognition site of fibronectin that retain attachment-promoting activity. *Proc. Natl. Acad. Sci.* **81**, 5985-5988.
29. **Yamada, K. M. & Kennedy, D. W.** (1984). Dualistic nature of adhesive protein function: Fibronectin and its biologically active peptide fragments can autoinhibit fibronectin function. *J. Cell. Biol.* **99**, 29-36.
30. **Gardner, J. M. & Hynes, R. O.** (1985). Interaction of fibronectin with its receptor platelets. *Cell* **42**, 439-448.
31. **Gartner, T. K. & Bennet, J. S.** (1985). The tetrapeptide analog of the cell attachment site of fibronectin inhibits platelet aggregation and fibronogen binding to activated platelets. *J. Biol. Chem.* **260**, 11891-11894.
32. **Plow, E. F., Pierschbacher, M. D., Ruoslahti, E., Marguerie, G. A. & Ginsberg, M. H.** (1985). The effect of Arg-Gly-Asp containing peptides on fibrinogen and von Willebrandfactor binding to platelets. *Proc. Natl. Acad. Sci.* **82**, 8057-8061.
33. **Suzuki, S., Oldberg, A., Hayman, E. G., Pierschbacher, M. D. & Ruoslahti, E.** (1985). Complete amino acid sequence of human vitronectin deduced cDNA. Similarity of cell attachment sites in vitronectin and fibronectin. *EMBO J.* **4**, 2519-2524.
34. **Pytela, R., Pierschbacher, M. D. & Ruoslahti, E.** (1985). Identification and isolation of a 140 kilodalton cell surface glycoprotein with properties of a fibronectin receptor. *Cell* **40**, 191-198.
35. **Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. & Huber, R.** (1995). High-Level Biosynthetic Substitution of Methionine in Proteins By Its Analogs 2-Aminohexanoic Acid, Selenomethionine, Telluromethionine And Ethionine In *Escherichia Coli*. *Eur. J. Biochem.* **230**, 788-796.
36. **Otwinowski, Z. & Minor, W.** (1997). Procession of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
37. **Terwilliger, T. C. & Berendzen, J.** (1999). Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 849-861.
38. **Jones, T. A., Zou, Z. -Y., Cowan, S. W. & Kjeldgaard, M.** (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogrogr. A* **47**, 110-117.
39. **Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S. *et al.*** (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905-921.

40. **Kraulis, P. J.** (1991). A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.
41. **Merritt, E. A. & Bacon, D. J.** (1997). Raster3D: Photorealistic Molecular Graphics. *Meth. Enzymol.* **227**, 505-524.
42. **Nicholls, A., Bharadwaj, R. & Honig, B.** (1993). GRASP-graphical representation and analysis of surface properties. *Biophys. J.* **64**, A166.

4

Isolation, expression and characterization of a glucose isomerase produced by the thermophilic bacterium *Fervidobacterium gondwanense*

Leon D. Kluskens, Ans C.M. Geerling, Jurrit Zeilstra, Willem M. de Vos, and John van der Oost

Submitted

Abstract

The *xylA* gene, coding for xylose isomerase from the thermophilic bacterium *Fervidobacterium gondwanense*, was cloned and overexpressed in *Escherichia coli*. The produced xylose isomerase, which is highly homologous to homologs from *Thermotoga maritima* and *T. neapolitana*, was purified and characterized. It is optimally active at 70°C, pH 7.3, with a specific activity of 15.0 U/mg for the interconversion of glucose to fructose. When compared to *T. maritima* XylA at 85°C, a higher catalytic efficiency was observed. Divalent metal ions Co^{2+} and Mg^{2+} were found to enhance the thermostability.

Introduction

The demand for cost-effective production of sweeteners has led to a tremendous increase in the production of High Fructose Corn Sirup (HFCS) over the last decades. In this process, starch undergoes enzymatic liquefaction via an α-amylase, subsequent saccharification by a glucoamylase, and finally the isomerization of glucose into the much sweeter fructose. The latter conversion can be catalyzed by the enzyme D-xylose isomerase (D-xylose keto-isomerase, EC 5.3.1.5). The physiological role of xylose isomerases (XIs) is the isomerization of xylose to xylulose. However, they also catalyze the commercially attractive conversion of glucose in fructose, and are therefore commonly referred to as glucose isomerases (for a review, see (4)). In addition XIs are of industrial interest since their production in yeast would allow for the fermentation of xylose to ethanol.

Despite their presence in a broad range of prokaryotes, commercially applied glucose isomerases are xylose isomerases obtained from mesophilic bacteria such as *Actinoplanes* sp., *Bacillus* sp., *Streptomyces* sp., and *Arthrobacter* sp. (4). However, at alkaline conditions and elevated temperatures unwanted by-product formation and side reactions (e.g. Maillard reactions) can occur (6). Subsequently, a XI is needed that can withstand higher operating temperatures and a slightly acidic pH. A number of XIs from organisms that are able to ferment xylose at high temperatures have been isolated and characterized, ranging from the moderately thermophilic *Thermus* sp. and *Thermoanaerobacter* sp. (8, 14), to the hyperthermophilic bacteria *Thermotoga neapolitana* (19) and *T. maritima* (5). To date, no archaea containing XIs have been described. The thermophilic bacterium *Fervidobacterium gondwanense*, isolated from heated geothermal waters, also belongs to the order *Thermotogales* (2). Its optimal growth at 68°C and its ability to use xylose as a carbon source prompted us to screen the organism for the presence of a *xylA* gene, potentially encoding a XI with optimal properties for fructose production. We here report on the isolation of the *xylA* gene from *F. gondwanense*, using the *T. maritima xylA* gene as a probe, its functional expression in *Escherichia coli*, and the biochemical characterization of the overproduced enzyme.

Results and Discussion

T. maritima MSB8 was grown at 80°C in 120-ml bottles, using 30 ml Marine Medium Thermotoga (MMT) (11). *F. gondwanense* (DSM13020) was grown as described previously (2). *T. maritima* and *F. gondwanense* genomic DNA was isolated according to established protocols (16). Small-scale plasmid DNA isolation was carried out using the Qiagen purification kit (Qiagen). Based on the *T. maritima xylA* sequence (TM1667, (5)), two primers were designed for PCR amplification (following an initial denaturation step of 5 min at 95°C: 60 s at 95°C, 30 s at 55°C, 60 to 120 s at 72°C, DNA thermal cycler Perkin Elmer), using primers BG1034 (5'-GGGAGGTGTTT**TCATGA**CAGAATTTTTCCC, *Bsp*HI site in bold) and BG1035 (5'-GGGTCTCAAAC**GAATTC**TCACCTCAGTTCTGC, *Eco*RI site in bold). The 1.4-kb PCR product was [α-P^{32}]ATP labeled by nick-translation, to be used as a probe. *F. gondwanense* genomic DNA (500 μg DNA) was digested with a range of restriction enzymes and run overnight on a 0.8% agarose gel. DNA fragments were subsequently transferred to a Hybond N^+ filter (Amersham Pharmacia Biotech Inc.) by capillary blotting (17). Hybridization of the Southern Blot was carried out overnight at 65°C with the *T. maritima xylA* probe. Subsequently, the filter was washed at low stringency: at room temperature with 2xSSC (0.3 M NaCl, 0.03 M sodium citrate) containing 1 mM EDTA, and with 2xSSC containing 0.1% SDS. Subsequently, the filter was exposed overnight to a phospho-imager screen, and the resulting images were scanned using a Storm-imager (Molecular Dynamics). Specific hybridization signals were obtained with a 3-kb *Eco*RI/*Nco*I-fragment, which was recovered and cloned in an *Eco*RI/*Nco*I digested pUC28 (3), and followed by *E. coli* XL1-Blue transformation (17). Ampicillin-resistant (10 μg/ml) colonies were transferred to NEF-978 filters (Biotechnology Systems), and colony hybridization was performed with the aforementioned *T. maritima xylA* probe (17). Five colonies showed a positive signal and were found to contain the 3-kb insert. Plasmid DNA from one positive transformant (called pWUR34) was isolated, and the insert was sequenced with a Li-Cor automatic sequencing system (model 4000L). DNA and protein sequencing data were analyzed with DNASTAR and compared to the GenBank Data Base by BLAST (1).

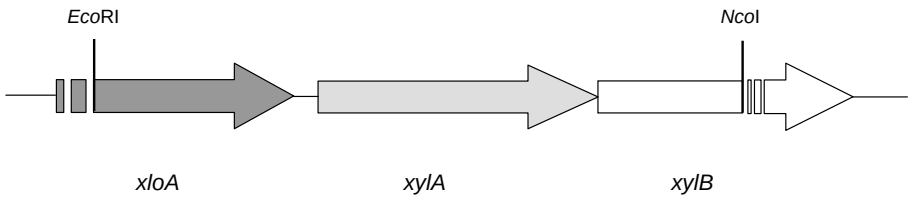


Figure 4.1 Schematic organization and flanking regions of the *F. gondwanense xylA* gene, the 3'-end of the xylosidase gene *xloA* (upstream), and the 5'-end of the putative xylulokinase *xylB* (downstream). The 3-kb hybridization fragment is indicated by the flanking *Eco*RI and *Nco*I restriction sites.

Nucleotide sequence analysis of the 3-kb insert of pWUR34 revealed an open reading frame, termed *xylA*, of 1,302-bp, which could encode a protein of 433 amino acids (*F. gondwanense* XylA; see Figure 4.1). The nucleotide sequence has been submitted to the GenBank/EMBL Data Bank (nr. AY431100). Homology with established xylose isomerases was found, predominantly from (hyper)thermophiles, with 62% identity to *T. maritima* and *T. neapolitana* XylA (Figure 4.2). The *F. gondwanense* XylA, like all XIs originating from *Thermotoga* species, belongs to class II of the xylose isomerases, and it distinguishes from class I by an extended N-terminal region (22). Flanking the *xylA* gene, two truncated ORFs were identified. Upstream an 800-bp fragment, which showed homology to a xylosidase (*xloA*) was identified when translated. Downstream from *xylA* the first 750 bp of a sugar kinase-like protein (putative xylulokinase) were recognized.

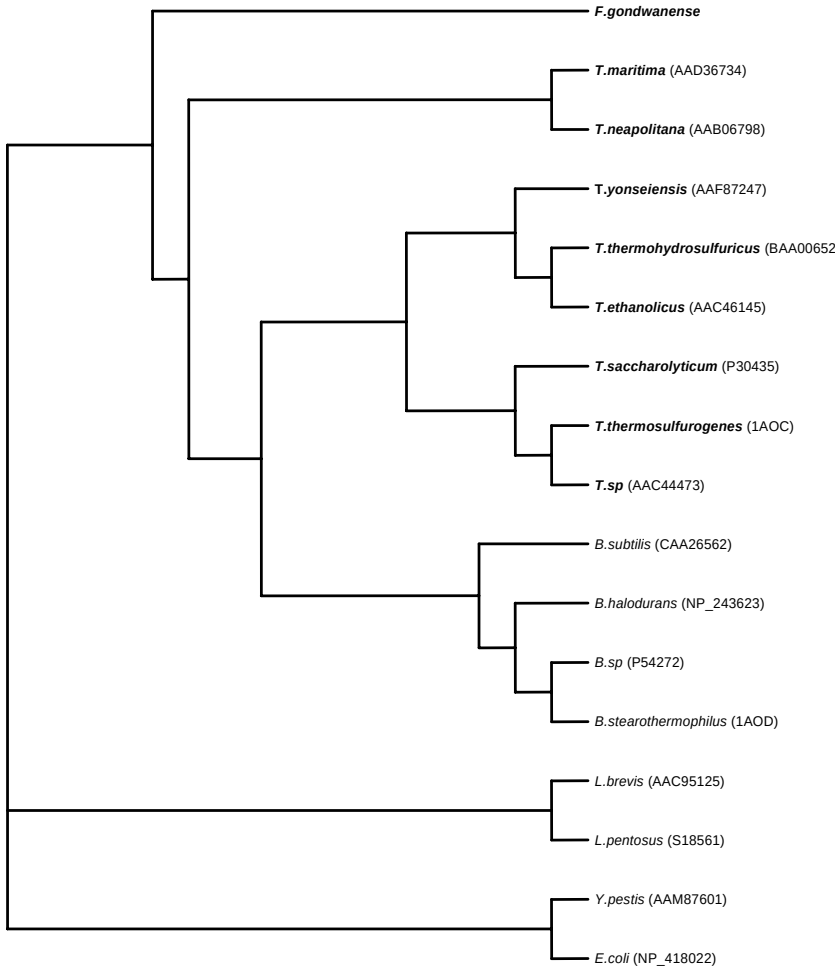


Figure 4.2 Dendogram of a selection of bacterial xylose isomerases. XIs originating from (hyper)thermophiles are shown in bold. GenBank accession numbers are added between brackets, except for *B. stearothermophilus* and *T. thermosulfurogenes* XI, for which the PDB accession numbers is given.

All three genes are positioned in the same transcriptional direction and part of a conserved gene cluster. Similar genetic organizations have been observed in the xylose isomerase-containing loci of *Bacillus subtilis* and *Lactobacillus pentosus* (15, 21) (Figure 4.1). For production of the *F. gondwanense* and *T. maritima* XylA the *E. coli* T7 expression system was used. Two primers were designed for amplification of the *F. gondwanense xylA* gene, introducing an *Nco*I site (bold) in the sense primer (BG1105; 5'- GAGGTGAAGT**CCATGG**CATACTTCAATGTTG) and a *Bam*HI site (bold) in the antisense primer (BG1106; 5'- CATATCTCGCTC**GGATCC**TTACCTCCCGCCG). The amplified 1.4-kb fragment was cloned into pET24d and named pWUR51. Likewise, the *T. maritima xylA* gene was PCR amplified with the aforementioned primers BG1034 and BG1035, and cloned as *Bsp*HI-*Eco*RI fragment in pET24d (pWUR05). The pET24d-derived plasmids were used to transform BL21(DE3) to kanamycin resistance (50 μg/ml). Single colonies of *E. coli* BL21 (DE3) harboring either pWUR05 or pWUR51 were used to inoculate 1-L cultures, grown 16 to 18 h in a baffled 2-L Erlenmeyer flask, at 150 rpm and 37°C. Cells were harvested by centrifugation (8,500 × g, 15 min, 4°C), the obtained cell pellet was resuspended in 20 mM Tris buffer (pH 8.0), and cells were disrupted by French Press. Cell debris was removed by centrifugation (8,500 × g, 15 min, 4°C) and the resulting supernatant was subjected to heat incubation, 30 min at 65°C for *F. gondwanense* XylA, 20 min at 80°C for *T. maritima* XylA. Precipitated proteins were removed by an additional centrifugation step (16,000 × *g*, 10 min, room temperature). The supernatant was loaded onto a Q-Sepharose anion-exchange column (Amersham Pharmacia Biotech Inc.), which was equilibrated with the same buffer. Bound proteins were eluted by a linear gradient from 0 to 1 M NaCl in 20 mM Tris buffer (pH 8.0). Peak fractions were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) in 10% (w/v) gels using the method of Laemmli (12). XylA-containing fractions were pooled and further separated on a Superdex 200 gel filtration column (Amersham Pharmacia Biotech Inc.), using 50 mM NaPi buffer, 100 mM NaCl, (pH 7.5). The protein concentration was spectrophotometrically determined at 280 nm. Analysis of purified fractions on SDS-PAGE revealed a band of approximately 45-kDa (results not shown), corresponding to the calculated size of 50 kDa for *F. gondwanense* XylA. As judged by SDS-PAGE analysis the enzyme was purified to homogeneity.

F. gondwanense XylA activity was measured by determining the formation of D-glucose using an established protocol (Glucose Trinder kit; Sigma) from D-fructose, or measuring the formation of D-fructose or D-xylulose (9), from D-glucose or D-xylose, respectively. Standard assays were carried out at 70°C for 1 h in 0.5 ml of 50 HEPES mM buffer (pH 7.4), using 0.5 M of glucose, fructose or xylose, in the presence of 1 mM $CoCl_2$ and 10 mM $MgCl_2$. The spectroscopic assays were performed in a microtiter plate reader (Labsystems iEMF reader MF). One unit of isomerase activity is defined as the amount of enzyme needed to produce 1 μmol of product per min under the assay conditions.

Kinetic parameters of *F. gondwanense* XylA were determined under standard conditions, using xylose, fructose, and glucose. The equilibrium of glucose isomerization generally can be driven towards a higher percentage of fructose when the processing temperature is increased. Kinetics was carried out at 85°C, at which a fructose fraction of almost 55% can be reached, without the

occurrence of undesired side reactions (4, 20). As is shown in Table 4.1, the most optimal parameters for *F. gondwanense* XylA were obtained with its natural substrate xylose, showing the lowest K_m and the highest V_{max} of all. In comparison with values obtained for *T. maritima* XylA, *F. gondwanense* XylA distinguishes by lower K_m values at 85°C, whereas the V_{max} is slightly higher. This results in a higher catalytic efficiency (k_{cat}/K_m), especially for xylose, but also for glucose (Table 4.1). In relation to kinetic parameters from *T. maritima* XylA at 90°C, for which optimal V_{max} values of 68.4 (xylose), 20.5 (fructose), and 16.2 μmol/min/mg (glucose) are reported (5), its values found for 85°C are proportional (Table 4.1).

The requirement of *F. gondwanense* XylA for divalent cations such as Co^{2+}, Mg^{2+}, and Mn^{2+} for activity and/or stability was examined using concentrations of 1 to 10 mM on EDTA-treated and dialysed *F. gondwanense* XylA. In the presence of 1 mM $CoCl_2$ and 10 mM $MgCl_2$ *F. gondwanense* XylA rendered the highest activity. The pH optimum of the isomerase activity was determined in the following range of buffers: acetate (50 mM, pH 4.0 to 5.4), PIPES (50 mM, pH 5.4 to 7.0), HEPES (50 mM, pH 7.2 to 8.2), and HEPBS (50 mM, pH 8.4 to 9.0). The pH values of all buffers were adjusted at 70°C. The pH optimum of *F. gondwanense* XylA was 7.3, in agreement with values determined for *T. maritima* XylA (6.5-7.5, (5)) and most other thermostable XIs (7.0-7.6, (20)).

Temperature optimum was determined using 50 mM HEPES, pH 7.4. The temperature optimum of *F. gondwanense* XylA following standard assay conditions was determined at 70°C using glucose and fructose as a substrate (not shown). The specific activity was 15.0 and 24.9 U/mg on glucose and fructose, respectively. These values are comparable with values found for XylA from *Thermus aquaticus* and *Streptomyces flavogriseus* at the same temperature (7, 13).

Table 4.1 Kinetic parameters of *F. gondwanense* and *T. maritima* XylA, determined at 85°C, on xylose, fructose, and glucose

F. gondwanense XylA	Xylose	Fructose	Glucose
K_m (mM)	6.3 ± 1.7	25.0 ± 2.9	28.5 ± 4.5
V_{max} (U/mg)	16.0 ± 0.9	15.3 ± 0.5	11.4 ± 0.5
k_{cat} (s^{-1})	13.3	12.7	9.4
k_{cat}/K_m ($mM^{-1}\,s^{-1}$)	2.1	0.51	0.33
T. maritima XylA			
K_m (mM)	76.0 ± 2.0	35.0 ± 9.9	33.7 ± 8.7
V_{max} (U/mg)	19.9 ± 2.0	12.2 ± 1.0	9.0 ± 0.6
k_{cat} (s^{-1})	16.9	10.3	7.6
k_{cat}/K_m ($mM^{-1}\,s^{-1}$)	0.22	0.30	0.23

The enzyme's thermostability was investigated by incubating the enzyme at temperatures ranging from 65 to 85°C at various periods of time, followed by measurement of the residual activity. The half-life for each temperature was the incubation time, after which 50% of the activity was left. Half-life values, representing the thermostability of *F. gondwanense* XylA, are shown in Table 4.2. In the presence of the divalent metal ions Co^{2+} and Mg^{2+} the half-life of *F. gondwanense* XylA increased considerably, with an increase from 15 min to over 5 h at 70°C. The stabilization of *F. gondwanense* XylA is mainly increased by Co^{2+}, as can be derived from the values determined for 85°C. Here, more stability is gained with 1 mM of $CoCl_2$ than with 10 mM of $MgCl_2$ (Table 4.2).

Thermal unfolding experiments (Differential Scanning Calorimetry) were carried out on a MicroCal VP-DSC in a temperature range from 50 to 125°C at a heating rate of 0.5°C/min. *F. gondwanense* XylA (0.3 mg/ml) was dialyzed against 20 mM HEPES, pH 7.4, prior to the experiment. The stabilizing effect of divalent metal ions was validated by DSC, with Co^{2+} having a more stabilizing effect than Mg^{2+}. The melting temperature of *F. gondwanense* XylA increased by more than 10°C when 1 mM $CoCl_2$ was added (79.5 vs. 90.4°C, not shown). Although Co^{2+} appears to be slightly biased in XIs, Mg^{2+} and Mn^{2+} too are reported provide stability (4, 10, 18). Compared to its thermophilic orthologs from *T. maritima* and *T. neapolitana*, the xylose isomerase from *F. gondwanense* is more moderately thermoactive and thermostable.

In conclusion, a thermostable xylose isomerase from *F. gondwanense* was *E. coli*-produced and characterized. Considering the observed catalytic efficiency at 85°C, *F. gondwanense* XylA might be a good candidate to be used in the glucose conversion processes.

Table 4.2 Half-life values of *F. gondwanense* XylA incubated at 65 to 85°C for a period of time

Addition	65°C	70°C	75°C	85°C
none	13 h	15 min	5 min	<1 min
10 mM Mg^{2+}	nd	nd	nd	4 min
1 mM Co^{2+}	nd	nd	nd	15 min
10 mM Mg^{2+}/ 1 mM Co^{2+}	18 h	>5 h	90 min	15 min

nd: not determined

References

1. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. H. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389-3402.
2. **Andrews, K. T., and B. K. C. Patel.** 1996. *Fervidobacterium gondwanense* sp. nov., a new thermophilic anaerobic bacterium isolated from nonvolcanically heated geothermal waters of the Great Artesian Basin of Australia. Int. J. Syst. Bacteriol. **46:**265-269.
3. **Benes, V., Z. Hostomsky, L. Arnold, and V. Paces.** 1993. M13 and pUC vectors with new unique restriction sites for cloning. Gene. **130:**151-152.
4. **Bhosale, S. H., M. B. Rao, and V. V. Deshpande.** 1996. Molecular and industrial aspects of glucose isomerase. Microbiol. Rev. **60:**280-300.
5. **Brown, S. H., C. Sjöholm, and R. M. Kelly.** 1993. Purification and characterization of a highly thermostable glucose isomerase produced by the extremely thermophilic eubacterium, *Thermotoga maritima.* Biotech. Bioeng. **41:**878-886.
6. **Bucke, C.** 1983. Glucose transforming enzymes. p. 93-127. *In* W. Fogarty (ed.), Microbial enzymes and biotechnology. Applied Science Publishers, London.
7. **Chen, W. P., A. W. Anderson, and Y. W. Han.** 1979. Extraction of glucose isomerase from *Streptomyces flavogriseus.* Appl. Environ. Microbiol. **37:**785-787.
8. **Dekker, K. A., A. Sugiura, H. Yamagata, K. Sakaguchi, and S. Udaka.** 1992. Efficient production of thermostable *Thermus thermophilus* xylose isomerase in *Escherichia coli* and *Bacillus brevis*. Appl. Microbiol. Biotechnol. **36:**727-732.
9. **Dische, Z., and E. Borenfreund.** 1951. A new spectrophotometric method for the detection and determination of keto sugars and trioses. J. Biol. Chem. **192:**583-587.
10. **Hartley, B. S., N. Hanlon, R. J. Jackson, and M. Rangarajan.** 2000. Glucose isomerase: insights into protein engineering for increased thermostability. Biochim. Biophys. Acta. **1543:**294-335.
11. **Kluskens, L. D., G.-J. W. M. van Alebeek, A. G. J. Voragen, W. M. de Vos, and J. van der Oost.** 2003. Molecular and biochemical characterisation of the thermo-active family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima.* Biochem. J. **370:**651-659.
12. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. **227:**680-685.
13. **Lehmacher, A., and H. Bisswanger.** 1990. Comparative kinetics of D-xylose and D-glucose isomerase activities of the D-xylose isomerase from *Thermus aquaticus* HB8. Biol. Chem. Hoppe-Seyler. **371:**527-536.
14. **Liu, S.-Y., J. Wiegel, and F. C. Gherardini.** 1996. Purification and cloning of a thermostable xylose (glucose) isomerase with an acidic pH optimum from *Thermoanaerobacterium* strain JW/SL-YS 489. J. Bacteriol. **178:**5938-5945.
15. **Lokman, B. C., P. van Santen, J. C. Verdoes, J. Kruse, R. J. Leer, M. Posno, and P. H. Pouwels.** 1991. Organization and characterization of three genes involved in xylose catabolism in *Lactobacillus pentosus*. Mol. Gen. Genet. **230:**161-169.
16. **Ramakrishnan, V., and M. W. W. Adams.** 1995. *In*: Archaea, A Laboratory Manual: Thermophiles, p. 95-96. *In* F. T. Robb, and A. R. Place (eds), Preparation of genomic DNA from sulfur-dependent hyperthermophilic Archaea. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
17. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** (eds.) 1989 Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

18. **Vieille, C., K. L. Epting, R. M. Kelly, and G. J. Zeikus.** 2001. Bivalent cations and amino-acid composition contribute to the thermostability of *Bacillus licheniformis* xylose isomerase. Eur. J. Biochem. **268:**6291-6301.
19. **Vieille, C., J. M. Hess, R. M. Kelly, and G. J. Zeikus.** 1995. xylA cloning and sequencing and biochemical characterization of xylose isomerase from *Thermotoga neapolitana*. Appl. Environ. Microbiol. **61:**1867-1875.
20. **Vieille, C., and G. J. Zeikus.** 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol. Mol. Biol. Rev. **65:**1-43.
21. **Wilhelm, M., and C. P. Hollenberg.** 1984. Selective cloning of *Bacillus subtilis* xylose isomerase and xylulokinase in *Escherichia coli* genes by IS5-mediated expression. EMBO J. **3:**2555-2560.
22. **Zeikus, G. J.** 1996. Molecular determinants of thermozyme activity and stability: analysis of xylose isomerase and amylopullulanases. *In:* K. H. Park, J. F. Robyt, and Y.-D. Choi (eds), Enzymes for Carbohydrate Engineering. Elsevier Science B.V., Amsterdam.

5

Molecular and biochemical characterisation of the thermo-active family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima*

Leon D. Kluskens, Gert-Jan W.M. van Alebeek, Alphons G.J. Voragen, Willem M. de Vos, and John van der Oost

Reprinted with permission from Biochem. J. (2003) 370: 651-659. Copyright 2003 Biochemical Society

Abstract

The ability of the hyperthermophilic bacterium *Thermotoga maritima* to grow on pectin as a sole carbon source coincides with the secretion of a pectate lyase (PelA) in the extracellular medium. The *pel*A gene of *T. maritima* was functionally expressed in *Escherichia coli* as the first heterologously produced thermophilic pectinase, and purified to homogeneity. Gel filtration indicated that the native form of PelA is tetrameric. Highest activity (422 U/mg, with a K_m of 0.06 mM) was demonstrated on polygalacturonic acid (PGA), whereas pectins with an increasing degree of methylation (DM) were degraded at a decreasing rate. In the tradition of pectate lyases, PelA demonstrated full dependency on Ca^{2+} for stability and activity. The enzyme is highly thermoactive and thermostable, operating optimally at 90°C and pH 9.0, with a half-life for thermal inactivation of almost 2 h at 95°C, and an apparent melting temperature of 102.5°C. Detailed characterization of the product formation with PGA indicated that PelA has a unique eliminative exo-cleavage pattern liberating unsaturated trigalacturonate (uG_3) as the major product, as opposed to unsaturated digalacturonate for other exopectate lyases known. The unique exo-acting mode of action was supported by progression profiles of PelA on oligogalacturonides (DP of 3 – 8) and the examination of the bond cleavage frequencies.

Introduction

Pectin is an abundant structure component in plant cell walls, and functions as a matrix holding cellulose and hemicellulose fibers. It is composed of a main chain of (partly) methyl-esterified α-1→4-D-polygalacturonate, and highly branched rhamnogalacturonan (rhamnose-galacturonate stretches), in which the latter can serve as a binding site for 1→4-linked side chains like arabinans, galactans or arabinogalactans [1].

Enzymatic modification of the pectin polymer is catalyzed by a variety of pectinolytic enzymes. They can be classified into (*i*) esterases that remove the methyl and/or acetyl groups from pectin, and (*ii*) depolymerases that cleave the backbone, either by hydrolysis (hydrolases) or nonhydrolytic β-elimination (lyases) [2]. Almost all bacteria and fungi able to grow on pectin possess pectate lyases that degrade non- or low methylated pectin (pectate) in either an exo- or endo-active way. They are generally located extracellularly, occasionally membrane-bound, and rarely present in the cytoplasm. Pectate lyases belong to the class of polysaccharide lyases, where they are classified in five out of twelve families (1, 2, 3, 9 and 10) [3]. With the exception of family 2 and 9, all of these have a three-dimensional structure available.

Enzymatic pectin degradation finds its way in several industrial processes, like the fruit-juice industry, and textile and paper treatment [4]. To date, attention has been mainly focused on pectinolytic enzymes from mesophilic fungi (e.g. *Aspergillus niger*) and bacteria (e.g. *Erwinia* and *Bacillus* species). The only thermophiles that have been reported to be capable of growing on pectin are *Caldicellulosiruptor* strains [5-7], a few *Clostridia* [8,9] and one archaeon, *Desulfurococcus amylolyticus* [10]. In addition, only a few thermo-active pectate lyases, originating from *Thermoanaerobacter italicus* [11] and *Bacillus* sp.[12], have been biochemically characterized.

Thermotoga maritima, an anaerobic, hyperthermophilic bacterium growing optimally at 80°C, can metabolize a wide variety of carbohydrates, including polymers such as xylan, starch and cellulose [13-15]. With the recent completion of the genome sequence of this organism [16], numerous enzymes that play an important role in displaying its heterotrophic potential were listed, including gene products putatively involved in pectin utilization. In this paper we report on the first heterologously expressed thermoactive pectate lyase (PelA)[1] from the hyperthermophile *T. maritima*, its biochemical characterization and provide a detailed analysis of its unique exocleaving properties, making PelA the first, extensively characterized, exopectate lyase generating trigalacturonate units.

Materials and methods

Organisms, growth conditions and plasmids

Thermotoga maritima strain MSB8 (DSM 3109) was routinely grown under anaerobic conditions at 80°C, pH 6.5, in 120-ml bottles, using 30 ml MMT medium (Marine Medium *Thermotoga*), containing (per liter): 40 g sea salt (Sigma), 3.1 g PIPES, 1.0 g yeast extract, 4.0 g tryptone, 2 ml trace elements [17], 0.5 g resazurin, 1 mM $CaCl_2$, 2 mM $MgSO_4$, 1 mM $Na_2S_2O_3$. Flasks were repeatedly flushed with N_2 and sterilized at 120°C. Pectin as a substrate with a DM of 65% and 89% (Sigma) was prepared separately as a 1% (w/v) stock and after autoclaving added to a 2x MMT medium to reach a final concentration of 0.5% (w/v). Glucose (20 mM) was added separately as filtered solutions. *T. maritima* was grown under mildly shaken conditions (100 rpm) to reduce inhibition by hydrogen formation. Growth was monitored by measuring the amount of hydrogen produced. To overcome the problem of underestimation when measuring the H_2 yield at ambient temperature, methane was used as an internal standard according to Kengen et al. [18].

The bacterial strain used for initial cloning experiments was *Escherichia coli* TG1 [*supE hsd* Δ5 *thi* Δ(*lac-proAB*) F' (*tra*D35 *proAB*$^+$ *lacI*^{q} *lac*Z ΔM15). *E. coli* BL21(DE3*) (hsdS gal* (λcl*ts* 857 *ind*1 *S*am7 *nin*5 *lac*UV5-T7 gene 1) was used for heterologous expression. The plasmid used for recombinant work was pET24d from Novagen.

Recombinant DNA techniques

Genomic DNA of *T. maritima* was isolated by using an established protocol [19]. Small-scale plasmid DNA isolation was carried out using the QIAGEN purification kit (Qiagen). DNA was digested with restriction endonucleases and ligated with T4 DNA ligase, according to the manufacturer's specifications (Life Technologies). DNA fragments were purified from agarose by QiaexII or from a PCR mix by using the PCR purification kit (Qiagen). Chemical transformation of *E. coli* TG1 and BL21(DE3) was carried out using established procedures [20].

The gene encoding the pectate lyase (*pel*A, TM0433) was identified in the course of the *T. maritima* genome sequence project [16]. Primers including appropriate restriction sites for cloning purposes (thereby changing serine into an alanine downstream the initial methionine) were designed to

exclude a putative signal sequence of 27 amino acids: BG906 (sense), 5'-GCGC**CCATGG**CTCTCAATGACAAAC CTGTGGG (*Nco*I restriction site in bold). BG887 (anti-sense), 5'-CGGATTGGTTGA**GGATCC**TTACTGAGCC (*Bam*HI restriction site in bold). A PCR reaction was carried out using 300 ng of both oligonucleotides and 100 ng *T. maritima* genomic DNA as a template in a final volume of 50 µl. After an initial denaturation step of 5 min at 95°C the DNA thermal cycler (Perkin Elmer) was programmed for a PCR with *Pfu* polymerase (1 U): 1 min at 95°C, 2 min at 50°C, and 3 min at 68°C, finished after 35 cycles by a final extension of 7 min at 68°C. The obtained PCR-product was digested with *Nco*I and *Bam*HI and cloned in a pET24d expression vector (Novagen), resulting in pLUW742, which was introduced into *E. coli* TG1 and BL21(DE3).

DNA and amino acid sequence analysis

Cloned PCR products were sequenced by the dideoxynucleotide chain termination method [21] with a Li-Cor automatic sequencing system (model 4000L). DNA and protein sequence data were analyzed with the DNASTAR package and compared to the GenBank Data Base by BLAST [22]. ClustalX was used to create an alignment of the *T. maritima* pectate lyase with homologs from the lyase family. GeneDoc was applied to manually adjust the alignment.

Purification of PelA

E.coli BL21(DE3) harboring pLUW742 was grown overnight (37°C, 150 rpm) in a 5 ml TYK (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 50 µg/ml kanamycin) preculture, after being isolated as a single colony from plate (TYK + 1.5% granulated agar). One ml was used to inoculate 1 liter of TYK in a baffled 2-L Erlenmeyer. Following an overnight growth at 37°C at 120 rpm, the culture was centrifuged for 15 min at 8,500 × *g* at 4°C and the cells were resuspended in 10 ml Tris-HCl (20 mM, pH 8.0). The cell suspension was sonified (Branson sonifier, 3 × 15'') and cell debris was removed by centrifugation at 16,000 × *g* for 10 min. The resulting supernatant was incubated for 20 min at 80°C and precipitated proteins were removed by an additional centrifugation step. The heat-stable cell-free extract was loaded onto a Q Sepharose column (Amersham Pharmacia Biotech, Inc.), which was equilibrated with Tris-HCl (20 mM, pH 8.0). Bound proteins were eluted by a linear gradient from 0 to 1 M of NaCl in the same buffer. Fractions containing PelA were pooled and concentrated (Filtron Technology Corp., 30 kDa cut-off). To determine its native conformation PelA was run over a gel filtration column (Superdex 200, Amersham Pharmacia Biotech, Inc.) and compared to a set of marker proteins, using 20 mM Tris-HCl, 100 mM NaCl, pH 8.0 as elution buffer.

Isolation of extracellular proteins of T. maritima

A 30-ml culture of *T. maritima* grown on glucose or pectin was spun down for 10 min at 10,000 × *g*. Proteins were recovered from the medium fraction by an 85% (w/v) ammonium sulfate precipitation [23] and a subsequent centrifugation step (10.000 × *g*, 10 min). The pellet was resuspended in 50 mM MOPS buffer (pH 7.7) and ammonium sulfate was removed by washing using Centricon filters (Millipore, 10 kDa cut-off).

Gel electrophoresis and zymogram staining

Protein samples were boiled in 5 × sample buffer for 5 min and analyzed by SDS-PAGE in 10% (w/v) gels [24]. Following migration, protein bands were stained with Coomassie Brilliant Blue. For zymogram staining, samples were boiled for 2.5 min before applying onto a 10% SDS-PAGE gel, which contained 0.05 % (w/v) polygalacturonic acid (the boiling time was reduced to increase the efficiency of refolding into an active enzyme). Denaturing agent SDS was removed by washing gels overnight at 4°C in 0.1 M MOPS buffer (pH 7.7), 1% (w/v) Triton X-100, 1 mM DTT. Subsequently, gels were incubated in 0.1 M MOPS buffer (pH 7.0), 1 mM $CaCl_2$, at 80°C for 1 h and stained with 0.05% (w/v) Ruthenium Red (Sigma) for 5 min. Zymograms were washed with water until pectate lyase activity bands appeared.

Enzyme assay

Initial pectate lyase activity was determined spectrophotometrically at 230 nm by measuring the formation of Δ4,5 unsaturated products, using a standard assay adapted from Collmer et al. ($\varepsilon_{232\ nm}$ = 4,600 M^{-1} ·cm^{-1}) [25]. Ten µl of enzyme (0.1-1µg/assay) was added to a closed 1-ml quartz cuvette containing 990 µl of substrate stock solution (0.25% (w/v) PGA, 50 mM Tris-HCl pH 8.0, 0.6 mM $CaCl_2$). The subsequent increase in absorbance at 230 nm was monitored as a function of time with a spectrophotometer (Hitachi U-2010). One unit of enzyme forms 1 µmol of Δ4,5 unsaturated product/min. Assays were carried out at 80°C and pH 8.0 unless otherwise stated. The pH of all buffers was equilibrated at the temperatures used. Enzyme thermostability was studied by determining the residual activity after incubation of PelA (0.05 mg/ml) in 50 mM Tris buffer pH 8.0, with or without 0.6 mM $CaCl_2$, in sealed glass tubes.

Alternatively, assays for HPLC purposes were carried out at 90°C in 20 mM MOPS buffer and 0.8 mM $CaCl_2$, pH 7.0, to approach the physiological pH. When using 0.5% (w/v) polygalacturonic acid (ICN) or pectin (DM of 30% and 74%) as a substrate 1 µg of PelA in 1 ml reaction volume was used. Activity assays on saturated oligoGal*p*A (2 mM) with a degree of polymerization (DP) of 3 to 8, generated as described previously [26], were carried out with an appropriate amount of enzyme for each oligogalacturonide. The pH of MOPS was adjusted for use at 90°C [27]. Samples of 100 µl were taken in time and the reaction was stopped by addition of 10 µl acetic acid (10% v/v) and chilling on ice. Samples were stored at –20°C until analyzed by HPSEC and HPAEC analysis.

HPSEC analysis

High performance size exclusion chromatography (HPSEC) analyses were done on three TSKgel columns (7.8 mm ID × 30 cm per column) in series (G4000 PWXL, G3000 PWXL, G2500 PWXL; Tosohaas) in combination with a PWX-guard column (Tosohaas). Elution was carried out at 30°C with 0.2 M sodium nitrate at 0.8 ml/min. The eluate was monitored using a refractive index detector. Calibration was performed using dextrans, pectins and oligoGal*p*A.

HPAEC analysis

High performance anion exchange chromatography (HPAEC) analyses at pH 12 were performed as described previously [28]. Δ4,5 unsaturated oligoGal*p*A were selectively detected with UV detection at 235 nm (UV1000, Thermo Separation Products), whereas both saturated and Δ4,5 unsaturated oligoGal*p*A were detected using a pulsed amperometric detector (PAD, Electrochemical Detector ED40, Dionex, Sunnyvale, CA). Pure saturated oligoGal*p*A (DP 1-7) and Δ4,5 unsaturated oligoGal*p*A (DP 2-7) were used as standards for external calibration of the system. The amount of product formation was quantified by peak integration (Chromquest). The specific activity (nmoles of product/min/mg protein) was calculated from the formation of unsaturated oligoGal*p*A in time (initial 4h). Activity on pectins was corrected for unsaturated oligoGal*p*A, produced by chemical β-elimination.

Differential scanning calorimetry

Thermal unfolding experiments were carried out on a MicroCal VP-DSC in a temperature range from 50 to 125°C at a heating rate of 0.5°C/min. All experiments were performed in buffer solutions with an ionic strength of 50 mM. The protein (0.3 mg/ml) was dialyzed against the buffer solution prior to the experiment.

Results

Genetic characterization of the pelA *gene from T. maritima*

The *pel*A gene (1,104 base pairs encoding 367 amino acids) was annotated in the *T. maritima* genome as pectate lyase (TM0433) [16]. PelA belongs to family 1 of the polysaccharide lyases, a group of carbohydrate-active enzymes that cleave polysaccharide chains by β-elimination [3]. A database search [22] indicated highest similarity with a pectate lyase from *Bacillus halodurans* (51%) [29]. The first 27 amino acids of PelA were predicted to be the signal sequence (Figure 5.1) [30]. Furthermore no transmembrane domains could be identified (SMART) [31,32], altogether suggesting an extracellular localization.

The results of the database search were used for a multiple sequence alignment at amino acid level between pectate and pectin lyases (Figure 5.1). The invariant residues located in the substrate binding cleft of pectin and pectate lyases (the active-site related residues R279 and D184 - *Bacillus subtilis* BsPel numbering is used throughout the manuscript) are conserved in all sequences.

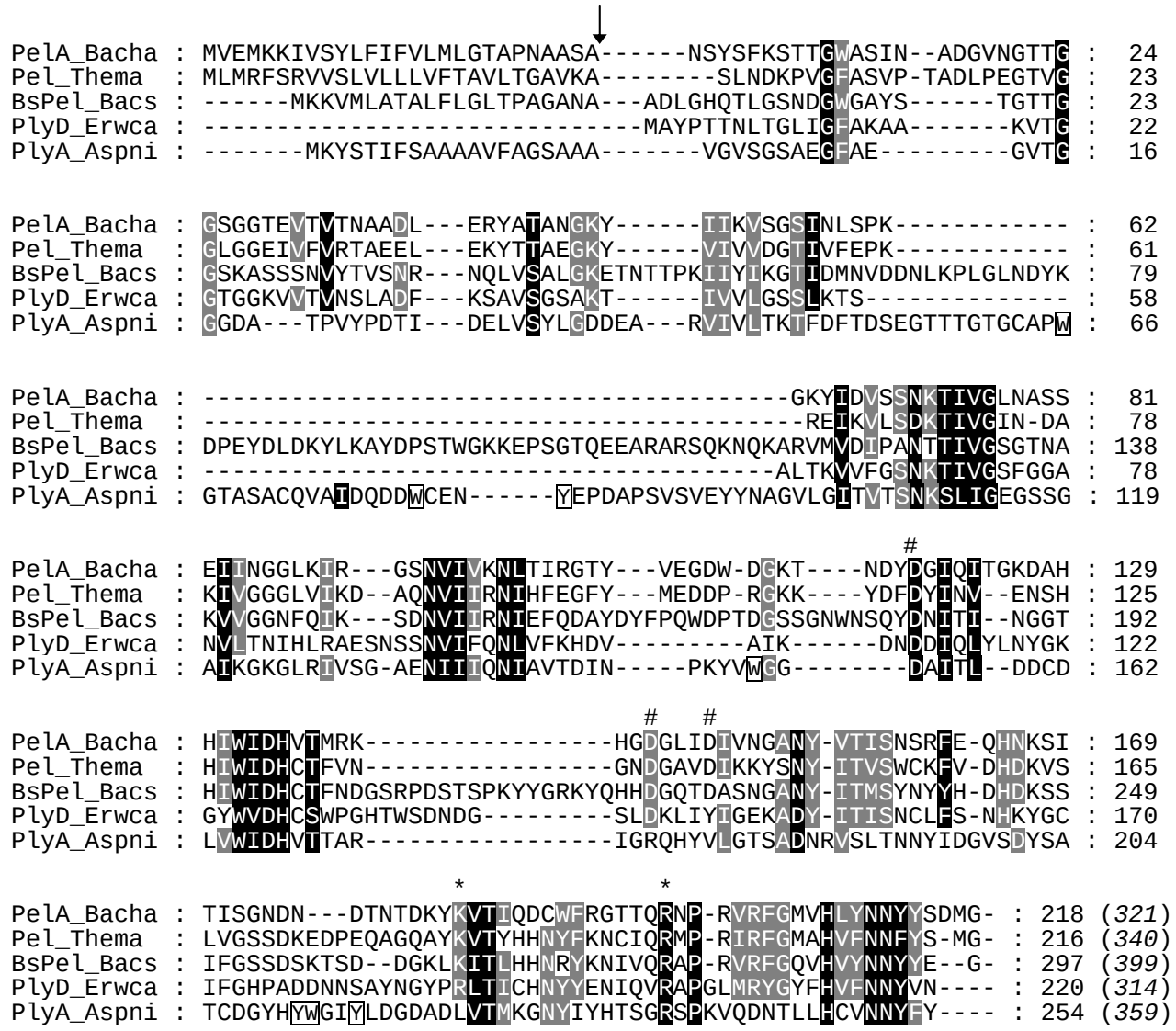


Figure 5.1 Multiple sequence alignment of pectate and pectin lyases belonging to family 1 of the polysaccharide lyases. Where necessary, the alignment was adapted as described previously [46]. PelA_Bacha (*Bacillus halodurans*, GenBank accession number AB041769), Pel_Thema (*Thermotoga maritima*, AE001722) and BsPel (*Bacillus subtilis*, D26349) are pectate lyases. PlyD_Erwca (*Erwinia carotovora*, M65057) and PlyA_Aspni (*Aspergillus niger*, X60724) are pectin lyases. An arrow indicates the predicted cleavage site between the signal peptide and the mature enzyme. Proposed Ca^{2+}-binding aspartates (#) and catalytic residues (*) are marked. Overall homology is highlighted in black, partial homology is highlighted in grey. Aromatic residues expected to contribute to the substrate binding by PlyA_Aspni [33] are boxed.

All three aspartate residues (D184, D223 and D227) that were shown to form a complex between pectate lyases and Ca^{2+} [33,34] are present in *T. maritima* PelA (Figure 5.1). Aromatic residues believed to be present in the substrate-binding cleft and expected to contribute to the affinity for noncharged, highly esterified pectin, are evident in the *Aspergillus niger* pectin lyase (PlyA) [33], but could not be identified in PelA (Figure 5.1).

Cloning, overproduction and purification of PelA

The *pel*A gene without the signal peptide-encoding sequence was amplified by PCR, including a methionine at position –1 of the predicted signal peptidase cleavage site and changing the serine at position +1 into an alanine. The resulting 1,026-bp fragment was cloned into a pET24d vector under the control of a T7 promoter, resulting in pLUW742. The gene was functionally overexpressed in *E. coli* BL21(DE3) by overnight incubation at 37°C. The protein was purified from a cell-free extract by heat incubation, followed by anion-exchange chromatography (Q Sepharose) during which the protein eluted at 0.6 M NaCl. PelA was considered pure after the last purification step, as judged by SDS-PAGE analysis (Figure 5.2, *lane* 1).

Growth of T. maritima on pectin

T. maritima was grown on glucose for 2 days, at which point the cells reached the stationary phase. Besides using these monosaccharides the organism also showed the capacity to grow well on citrus pectin with a DM of 65%. The maximal doubling time when grown on pectin was calculated to be 260 min, compared to 82 min when grown on glucose under the same conditions (not shown). Cultures grown on pectin were tested on PGA for pectate lyase activity. High levels of activity were found in the medium, confirming the extracellular localization of PelA.

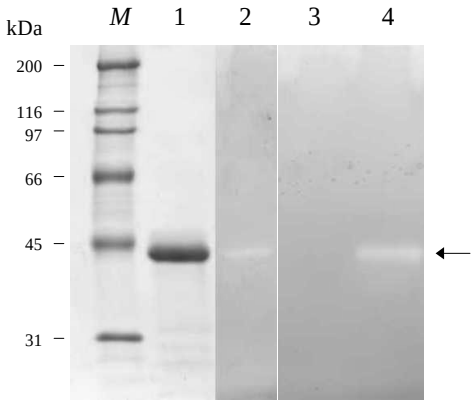


Figure 5.2 SDS-gel electrophoresis of recombinantly produced PelA juxtaposed to a PGA-zymogram analysis of recombinant PelA and the medium fraction of *T. maritima* grown on glucose and pectin. Samples were prepared as described under 'Experimental Procedures'. The sizes of the marker proteins (*M*) are shown at the left in kDa. The position of PelA is indicated with an arrow on the right. *Lane 1*, purified recombinant PelA, Coomassie-stained; *lane 2*, purified recombinant PelA, zymogrammically-stained; *lane 3*, medium fraction *T. maritima* culture, grown on glucose; *lane 4*, medium fraction *T. maritima* culture, grown on pectin.

Enzyme properties

PelA activity in the medium fraction was furthermore confirmed by zymogram experiments. Zymograms of the recombinant PelA displayed a prominent PGA-degrading activity band at around 40 kDa (Figure 5.2, *lane* 2). Similar analysis of the medium fraction of pectin-grown *T. maritima* demonstrated an activity band of identical molecular weight (Figure 5.2, *lane* 4), whilst no extracellular activity was detected in glucose-grown cultures (Figure 5.3, *lane* 3). It is therefore very likely that the *E. coli*-produced PelA and the secreted pectate lyase are identical. When analyzed by gel filtration PelA eluted, similarly to rabbit muscle aldolase, at 158 kDa, indicating that PelA most likely is a 151.2-kDa tetramer (the calculated molecular mass of the mature monomer is 37.8 kDa) (result not shown).Initial activity of PelA was assayed by monitoring the increasing absorbance at 230 nm of Δ4,5-unsaturated reaction products. Following the standard assay by Collmer et al. [25] the activity of PelA was 422 U/mg. The K_m for PGA was found to be 0.06 mM. The effect of pH and temperature was examined. Optimal catalysis was observed over a broad temperature range, with an optimum at 90°C. Activity increased moderately with an increasing pH until pH 9.0, after which it dropped rapidly (not shown).

Since Ca^{2+} is known to be essential for the catalytic activity of pectate lyases, PelA activity was examined on PGA (0.25% (w/v)) in the presence of the chelator EGTA. Treatment with 1 mM of EGTA completely abolished PelA activity. The effect of Ca^{2+} was examined at different $CaCl_2$ concentrations. An increase in activity was detected until apparent saturation was reached at 0.6 mM $CaCl_2$. Further increase, up to 1 mM, had no significant effect on the activity of the enzyme (not shown). The low level of activity detected in the absence of $CaCl_2$ might be attributed to residual trace amounts of Ca^{2+}, still present after purification.

Given that *T. maritima* is of marine origin, the effect of NaCl on the PelA activity was examined. When considering the activity of PelA without NaCl as 100%, PelA displays its highest activity at 200 mM NaCl, exceeding the initial activity more than five times (not shown). Addition of more NaCl results in a reduction of the water activity, which in its turn leads to gel formation between Ca^{2+} ions and PGA, causing the drop in activity.

Thermostability of PelA

Purified PelA (0.05 mg/ml) was incubated in a 50 mM Tris buffer, pH 9.0, in the presence of 0.6 mM $CaCl_2$ at 90, 95 and 100°C, respectively, and without $CaCl_2$ at 90°C. Samples taken at timed intervals were analyzed at 80°C for residual PelA activity. At 95°C the half-life of PelA was found to be 110 min, whereas at 100°C PelA lost half of its activity after 15 min. The absence of Ca^{2+} caused a dramatic decrease in residual activity of PelA, losing half of its activity within five minutes of incubation at 90°C (Figure 5.3). The presence of NaCl, up to a concentration of 1M, appeared to have no effect on the thermostability (not shown).

Additionally, the melting temperature of PelA (0.3 mg/ml) was examined by DSC in the presence and absence of Ca^{2+}. These studies showed that in addition of 0.6 mM $CaCl_2$ PelA thermally unfolded at a temperature of 102.5°C. This melting temperature is over 20°C higher than during

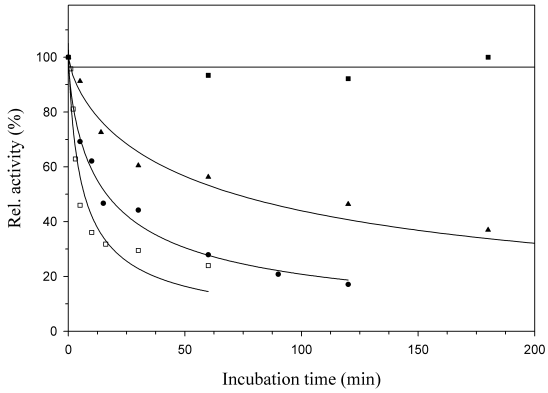


Figure 5.3 Thermostability of PelA examined by measuring the residual activity after incubation under standard conditions at 90°C (■), 95°C (▲), 100°C (●), and at 90°C in the absence of $CaCl_2$ (□). The activity determined under standard conditions at 90°C was defined as 100%. All data represent the means of duplicate assays.

incubation without Ca^{2+} (81.5°C), thereby corroborating the importance of Ca^{2+} for the thermostability of PelA (results not shown).

Mode of action of PelA on substrate polymers

Degradation of PGA by PelA was monitored by HPSEC and HPAEC (Figure 5.4). HPSEC columns were calibrated with pectins and oligoGal*p*A, allowing the identification of the substrate polymer and produced oligogalacturonides. Activity upon PGA resulted in the direct formation of an oligomer, which was identified as trigalacturonate when related to a standard set of oligoGal*p*A. Trigalacturonate in its turn is further processed into a dimer and monomer, both of which are occurring after 24 h (Figure 5.4A). The repositioning of the large polymer peak, in relation to the polymer in the blank reaction (dotted line), indicates that PelA displays exocleaving activity towards PGA.

The oligomers generated by PelA were examined in more detail on HPAEC using pulsed amperometric detection (PAD, (un)saturated oligoGal*p*A) and UV detection (unsaturated oligoGal*p*A). The peak appearing in Figure 5.4A was clearly identified via UV detection as an unsaturated trigalacturonate (uG_3). Already, a slight formation of uG_2 can be identified (Figure 5.4B). Figure 5.4C shows the progression profiles of the amount of Δ4,5 unsaturated oligoGal*p*A (UV) generated in time, corrected for chemical β-elimination. It can be seen that uG_3 is immediately produced from the start of the reaction. These results illustrate that PelA attacks from the reducing end, since only then uG_3 can be generated. This mode of action is then followed by degradation of uG_3 into unsaturated digalacturonate (uG_2), as can be deduced from the increasing amount of released uG_2. The corresponding uG_1 (5-keto-4-deoxyuronate) cannot be detected by UV as a result of its open ring structure, and is also relatively unstable [35], nevertheless it was earlier demonstrated by HPSEC. From the progression profile the specific activity for the generation of uG_3 from PGA over the first 4h was calculated to be 9.9 U/mg.

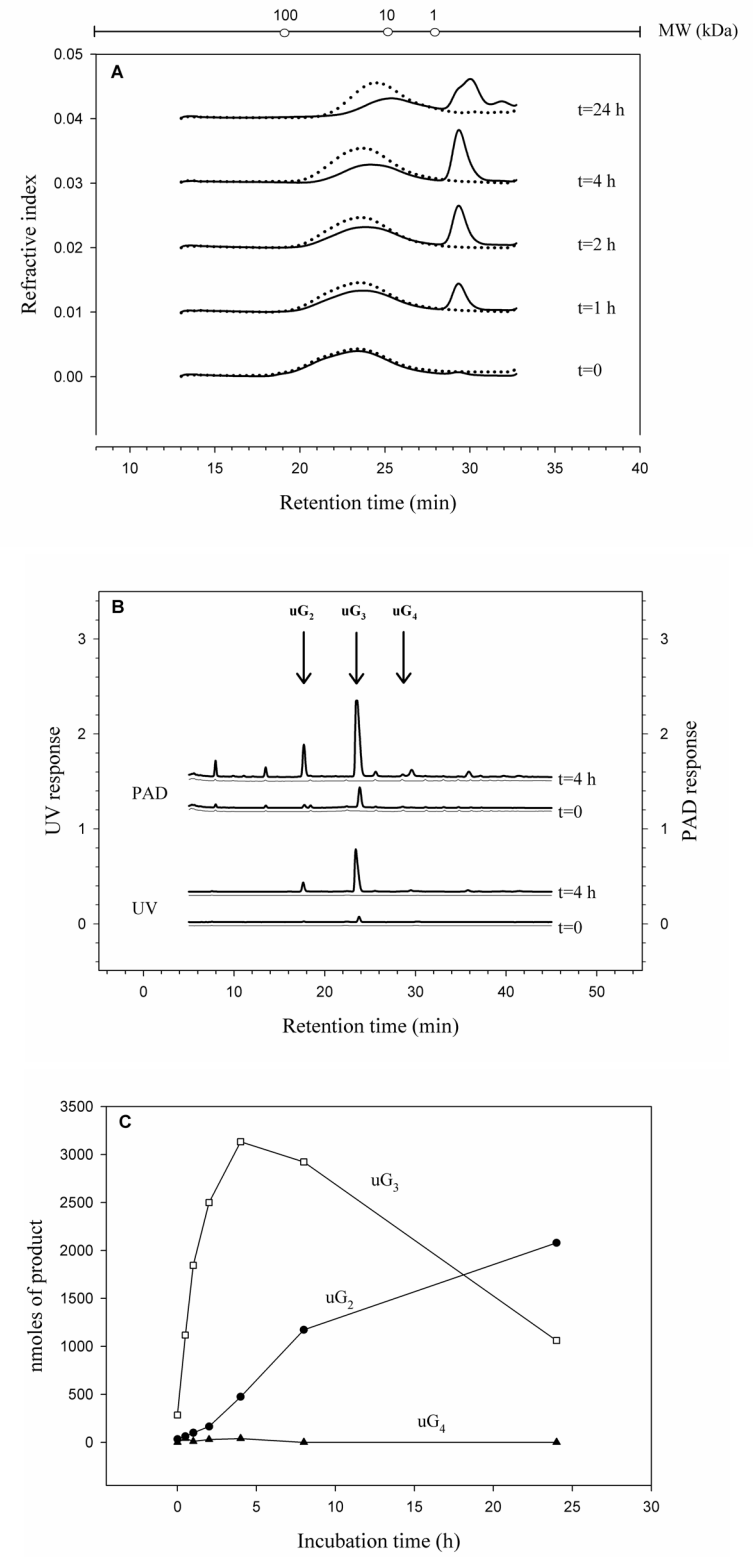


Figure 5.4 Degradation of PGA by PelA monitored in time by (**A**) HPSEC, (**B**) HPAEC (PAD and UV), and (**C**) progression profiles, derived from UV detection pattern. Dashed lines indicate blank experiments (PGA incubation without PelA). The molecular weight distribution (kDa) derived from a standard curve is stated at the top. Unsaturated oligogalacturonides are indicated as follows: uG_2 (●), uG_3 (□), and uG_4 (▲).

When assayed on pectin with a different degree of methylesterification (DM), the product formation indicates that a substantial amount of uG_3 is produced, despite the high degree of methylation (Figure 5.5). However, the specific activity of PelA diminished considerably, retaining only 41% and 2% of its activity on DM30% and DM74%, respectively (4.1 and 0.21 U/mg). Assays for HPLC purposes were carried out at the more physiological pH of 7.0, to minimize non-enzymatic demethylation and β-elimination when using pectin as a substrate [36]. Demethylation of pectin under growth conditions of *T. maritima* was not observed, however, non-enzymatic formation of unsaturated oligomers evidently occurred under HPLC assay conditions (Figure 5.5). In time, oligomers with a DP up to 8 were generated as a result of temperature-induced β-elimination of pectin with a degree of methylation of 74%. Clear fragmentation of the substrate as a direct result of the assay conditions could be identified by a drastic repositioning of the polymer peak (Figure 5.5A). The generated oligogalacturonides were shown to be unsaturated as detected via UV (Figure 5.5B).

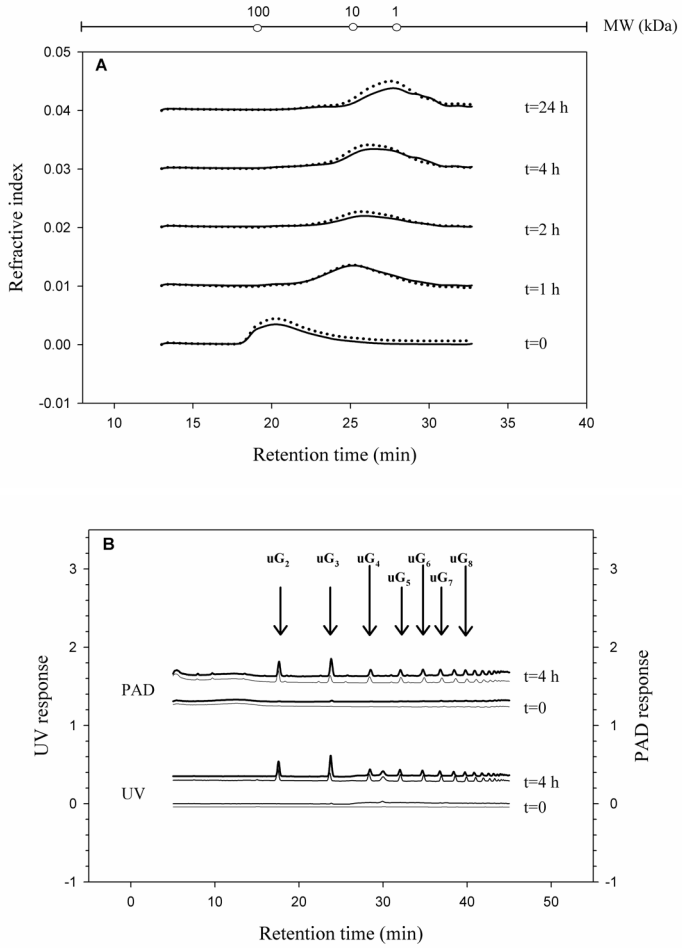


Figure 5.5 Degradation of pectin (DM = 74%) by PelA monitored in time by HPSEC (**A**) and at times t=0 and t=4 h by HPAEC (PAD and UV) (**B**). Dashed lines indicate blank experiments (pectin DM74% incubated without PelA). The molecular weight distribution (kDa) derived from a standard curve is stated at the top.

Table 5.1 Bond cleavage frequencies and rate of β-elimination of PelA on saturated oligogalacturonides. BCFs are calculated from the Δ4,5 unsaturated products as detected by UV. Activity rates were determined from the produced Δ4,5 unsaturated oligogalacturonides (UV) and the corresponding saturated oligogalacturonides produced (PAD). The reducing end of the oligoGal*p*A is shown in boldface (**G**).

DP																							Activity (mU/mg)	
																							UV	PAD
3													G	— 100	G	—	**G**						34.1	36.2
4										G	— 99	G		— 1	G	—	**G**						3326	3818
5							G	—	G	— 99	G		— 1	G	—	**G**							8258	8492
6					G	—	G	— 1	G	— 98	G		— 1	G	—	**G**							6905	6399
7			G	—	G	—	G	— 2	G	— 97	G		— 1	G	—	**G**							5217	4478
8	G	—	G	—	G	—	G	— 3	G	— 96	G		— 1	G	—	**G**							3584	3356

Mode of action on oligoGalpA. The mode of action of PelA on oligoGal*p*A (DP 3-8) was examined by HPAEC, analogous to PGA as a substrate. Progression profiles were constructed and showed a linear formation of both the saturated and the corresponding unsaturated oligogalacturonides (hence avoiding secondary product formation). From these linear areas both the bond cleavage frequency (BCF) and the activity rate were calculated (Table 5.1). Besides confirming its obvious preference for releasing Δ4,5 unsaturated trimers it shows that the specific activity rates calculated from the saturated and unsaturated product formation correspond well. Noteworthy is the leap in activity rate when the oligogalacturonide length is increased from $(Gal pA)_3$ to $(Gal pA)_4$, which is a hundred fold. Analysis of the intermediate products demonstrated that incubation of $(Gal pA)_n$ results in the parallel production of uG_3 and $G_{(n-3)}$. The fact that the saturated counterproduct $G_{(n-3)}$ is formed in equivalent amounts indicates that PelA follows a multiple chain attack, as opposed to a single chain-multiple attack mode that would result in the immediate degradation of the saturated intermediate. An increase in activity rates was found with

increasing DP up to 5, with a significant elevation from (Gal*p*A)$_3$ to (Gal*p*A)$_4$ of over 100 times. Surprisingly, both rates drop again when the DP is further augmented. From the BCFs and the activity rates on the subsequent oligogalacturonides, we deduced that PelA most likely contains five subsites.

Discussion

T. maritima has the ability to grow on a variety of polysaccharides, such as cellulose, starch and xylan, for which it possesses a huge amount of carbohydrate-active enzymes, mainly glycoside hydrolases [13-15]. In the present study we have demonstrated that *T. maritima* is also able to grow on pectin, the methylated α-1,4-linked polymer of galacturonic acid. Activity of pectate lyase, one of the key enzymes in the degradation process, was detected in the medium fraction of a pectin-grown culture. Activity was displayed as one single band, which shows that the production of PelA is induced by growth on pectin. Also, it strongly implies that PelA is the only secreted pectinolytic enzyme.

A candidate gene (TM0433), designated PelA, for the pectinolytic degradation has been identified in the *T. maritima* genome [16], based on a high sequence similarity with a family 1 pectate lyase from *B. halodurans* [3]. PelA contains a signal sequence and lacks additional transmembrane helices, which coincides with its presence in the medium fraction.

Subsequent characterization of the purified, *E. coli*-produced PelA validated its classification as a pectate lyase, displaying the highest activity on nonmethylated polygalacturonic acid (PGA) with an absolute requirement for Ca^{2+}. PelA cleaves the substrate by means of β-elimination optimally under alkaline conditions (pH 9.0). Detailed examination of the reaction products demonstrated that PelA is an exo-cleaving pectate lyase (EC 4.2.2.9) following multiple chain attack, making it one of the few exopectate lyases from family 1; others are classified in family 2 and 9. Based on the activity rates the substrate binding cleft most probably consists of five subsites, three of which interact with the reducing end of the substrate. A large increase in activity rate was observed with an increase in DP from 3 to 4. This emphasizes the high affinity that subsite +3 has for the galacturonic acid monomer and relates well to the predominantly formed product uG_3. PelA generates unsaturated trigalacturonate as the first product, thereby showing that the enzyme attacks the polymer from the reducing end. To the best of our knowledge this is the only exopectate lyase able to cleave off trimers (trigalacturonate) as opposed to digalacturonate what is usually observed in bacteria and fungi [37,38]. Unquestionably, the recent crystallization of *T. maritima* PelA [39] will give more insight into the exact mode of action of the enzyme.

When PGA is replaced by pectin, the activity of PelA strongly decreases as a direct result of a higher DM. Nevertheless, the organism is able to use highly methylated pectin (up to 89%) as a carbon source. No pectin methylesterases could be identified in the genome sequence from *T. maritima* and also no esterase activity was detected in the medium fraction of pectin-grown cells. Unsaturated oligogalacturonides, generated by chemical β-elimination of highly methylated pectin, might serve as a substrate that is more ideal for PelA under growth situations. It is obvious that, to a

certain extent, the elevated temperatures play a role in enabling growth of *T. maritima* on pectin. It is therefore tempting to speculate on the absolute necessity of PelA under the extreme *in vivo* conditions, since the oligogalacturonides liberated by PelA could equally be generated via chemical β-elimination. In order to do so, we would need to know whether these oligogalacturonides still contain methyl groups or not and to what extent PelA could accept them. Up to now, the rate of tolerance for methylesters in the substrate binding cleft of PelA is not known.

T. maritima PelA is the first heterologously expressed pectate lyase originating from a hyperthermophilic organism. With its temperature optimum of 90°C it is the most thermoactive pectate lyase known to date. Compared to other extracellular polymer-degrading enzymes from *T. maritima* like xylanases, mannanases and cellulases, PelA shows comparable properties with respect to thermoactivity and –stability [40]. Based upon gel filtration, we found PelA to be a tetramer. With the exception of pectate lyase b from the thermophilic bacterium *Thermoanaerobacter italicus* (dimer) [11], all pectate lyases characterized to date are monomeric. A higher degree of oligomerization is observed more frequently in thermophilic enzymes than in their mesophilic counterparts, and is regarded as a potential stabilization factor [41]. The large effect that Ca^{2+} ions have on the thermostability of pectate lyases has been described frequently, for pectate lyases from both hyperthermophilic and mesophilic origin.

Enhancement of pectate lyase activity by NaCl has been stated before. Some reports describe total dependency on NaCl for activity [42,43]; other pectate lyases are only stimulated [44] or not influenced at all. Benen and coworkers have proposed that the positive effect of NaCl on the activity is caused by a discharge of loosely bound Ca^{2+} from pectate chains, leading to an increase in availability of substrate and free Ca^{2+} [45]. Since $CaCl_2$ concentrations higher than 0.6 mM did not significantly stimulate PelA activity the NaCl-induced enhancement of the activity is probably due to an increase in substrate accessibility. The drop in activity when the NaCl concentration exceeds 250 mM can be ascribed to gel formation, caused by a reduction of the water activity. The NaCl concentration at which PelA is most active is much lower than the salt concentration of seawater, the natural environment of *T. maritima*, which contains about 500 mM of NaCl. However, pectate in nature is generally methylated, which reduces its capacities as a Ca^{2+}-pectate gel.

Growth on pectin of the hyperthermophilic bacterium *T. maritima* appears to be initiated by the secretion of PelA, an extracellular exopectate lyase. Together with a putative exopolygalacturonosidase [16] the organism holds at least two pectinases, both active in the initial process of pectin catabolism. With the genome sequence available, the road towards the elucidation of additional genes related to pectin degradation appears to be wide open. For this reason, DNA microarray analyses on the *T. maritima* genome, with the intention to unravel the pectin catabolism, are currently under investigation.

Acknowledgments

The authors would like to thank Ing. Ans Geerling for technical assistance and Dr. Jacques Benen for helpful discussions.

References

1. **Carpita, N. C., and Gibeaut, D. M.** (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant Rev. **3,** 1-30
2. **Sakai, T., Sakamoto, T., Hallaert, J., and Vandamme, E. J.** (1993) Pectin, pectinase and protopectinase: production, properties, and applications. Adv. Appl. Microbiol. **39,** 213-294
3. **Coutinho, P. M., and Henrissat, B.** (1999) http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html
4. **Kashyap, D. R., Vohra, P. K., Chopra, S., and Tewari, R.** (2001) Applications of pectinases in the commercial sector: a review. Bioresource Technol. **77,** 215-227
5. **Huang, C.-Y., Patel, B. K., Mah, R. A., and Baresi, L.** (1998) *Caldicellulosiruptor owensensis* sp. nov., an anaerobic, extremely thermophilic, xylanolytic bacterium. Int. J. Syst. Bacteriol. **48,** 91-97
6. **Bredholt, S., Sonne-Hansen, J., Nielsen, P., Mathrani, I. M., and Ahring, B. K.** (1999) *Caldicellulosiruptor kristjanssonii* sp. nov., a cellulolytic, extremely thermophilic, anaerobic bacterium. Int. J. Syst. Bacteriol. **49,** 991-996
7. **Rainey, F. A., Donnison, A. M., Janssen, P. H., Saul, D., Rodrigo, A., Bergquist, P. L., Daniel, R. M., Stackebrandt, E., and Morgan, H. W.** (1994) Description of *Caldicellulosiruptor saccharolyticus* gen. nov., sp. nov: An obligately anaerobic, extremely thermophilic, cellulolytic bacterium. FEMS Microbiol. Lett. **120,** 263-266
8. **Wiegel, J., Ljungdahl, L. G., and Rawson, J. R.** (1979) Isolation from soil and properties of the extreme thermophile *Clostridium thermohydrosulfuricum*. J. Bacteriol. **139,** 800-810
9. **Schink, B., and Zeikus, J. G.** (1983) Characterization of pectinolytic enzymes of *Clostridium thermosulfurogenes.* FEMS Microbiol. Lett. **17,** 295-298
10. **Bonch-Osmolovskaya, E. A., Slesarev, A. I., Miroshnichenko, M. L., Svetlichnaya, T. P., and Alekseev, V. A.** (1988) Characteristics of *Desulfurococcus amyloliticus* sp. nov., a new extremely thermophilic archaebacterium isolated from thermal springs of Kamchatka and Kunashir Island. Mikrobiologiya **57,** 94-101
11. **Kozianowski, G., Canganella, F., Rainey, F. A., Hippe, H., and Antranikian, G.** (1997) Purification and characterization of thermostable pectate-lyases from a newly isolated thermophilic bacterium, *Thermoanaerobacter italicus* sp. nov. Extremophiles **1,** 171-182
12. **Takao, M., Nakaniwa, T., Yoshikawa, K., Terashita, T., and Sakai, T.** (2000) Purification and characterization of thermostable pectate lyase with protopectinase activity from thermophilic *Bacillus* sp. TS 47. Biosci. Biotechnol. Biochem. **64,** 2360-2367
13. **Chhabra, S. R., Shockley, K. R., Ward, D. E., and Kelly, R. M.** (2002) Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. Appl. Environ. Microbiol. **68,** 545-554
14. **Huber, R., Langworthy, T. A., König, H., Thomm, M., Woese, C. R., Sleytr, U. B., and Stetter, K. O.** (1986) *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. Arch. Microbiol. **144,** 324-333
15. **Huber, R., and Stetter, K. O.** (1992) The order Thermotogales *in: "The Prokaryotes" (A. Balows, et al., eds.)* (Springer, ed), pp. 3809-3815, Berlin, Heidelberg, New York
16. **Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, L. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. A., White, O., Salzberg,**

S. L., Smith, H. O., Venter, J. C., and Fraser, C. M. (1999) Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. Nature **399,** 323-329

17. **Zehnder, A. J. B., Huser, B. A., Brock, T. D., and Wuhrmann, K.** (1980) Characterization of an acetate-decarboxylating, non-hydrogen-oxidizing methane bacterium. Arch. Microbiol. **124,** 1-11
18. **Kengen, S. W. M., Luesink, E. J., Stams, A. J. M., and Zehnder, A. J. B.** (1993) Purification and characterization of an extremely thermostable β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus.* Eur. J. Biochem. **213,** 305-312
19. **Ramakrishnan, V., and Adams, M. W. W.** (1995) Preparation of genomic DNA from sulfur-dependent hyperthermophilic Archaea *in*: Archaea, A Laboratory Manual: Thermophiles (Robb, F. T., and Place, A. R., eds), pp. 95-96, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
20. **Sambrook, J., Fritsch, E. F., and Maniatis, T.** (eds) (1989) *Molecular cloning: a laboratory manual*, 2nd Ed. Edited by Nolan, C., Cold Spring Harbor Laboratory Press, Cold Spring Harbor
21. **Sanger, F., Nicklen, S., and Coulson, A. R.** (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74,** 5463-5467
22. **Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J.** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25,** 3389-3402
23. **Scopes, R. K.** (1982) *Protein purification. Principles and practice. Springer-Verlag, New York*
24. **Laemmli, U. K.** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227,** 680-685
25. **Collmer, A., Ried, J. L., and Mount, M. S.** (1988) Assay methods for pectic enzymes. Methods Enzymol. **161,** 329-399
26. **van Alebeek, G.-J. W. M., Christensen, T. M. I. E., Schols, H. A., Mikkelsen, J. D., and Voragen, A. G. J.** (2002) Mode of action of pectin lyase A of *Aspergillus niger* on differently C_6-substituted oligogalacturonides. J. Biol. Chem. **277,** 25929-25936
27. **Daniel, R. M., and Danson, M. J.** (2001) Assaying activity and assessing thermostability of hyperthermophilic enzymes. Methods Enzymol. **334,** 283-293
28. **van Alebeek, G.-J. W. M., Zabotina, O., Beldman, G., Schols, H. A., and Voragen, A. G. J.** (2000) Esterification and glycosydation of oligogalacturonides: examination of the reaction products using MALDI-TOF MS and HPAEC. Carbohydr. Polym. **43,** 39-46
29. **Takami, H., and Horikoshi, K.** (2000) Analysis of the genome of an alkaliphilic *Bacillus* strain from an industrial point of view. Extremophiles **4,** 99-108
30. **von Heijne, G.** (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14,** 4683-4690
31. **Letunic, I., Goodstadt, L., Dickens, N. J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R. R., Ponting, C. P., and Bork, P.** (2002) Recent improvements to the SMART domain-based sequence annotation resource. Nucleic Acids Res. **30,** 242-244
32. **Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P.** (1998) SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. U.S.A. **95,** 5857-5864
33. **Mayans, O., Scott, M., Connerton, I., Gravesen, T., Benen, J., Visser, J., Pickersgill, R., and Jenkins, J.** (1997) Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate-binding clefts of pectin and pectate lyases. Structure **5,** 677-689

34. **Pickersgill, R., Jenkins, J., Harris, G., Nasser, W., and Robert-Badouy, J.** (1994) The structure of *Bacillus subtilis* pectate lyase in complex with calcium. Struct. Biol. **1,** 717-723

35. **Voragen, A. G. J.** (1972) Characterization of pectin lyases on pectins and methyl oligogalacturonates. Ph.D. Thesis, Wageningen, Landbouwhogeschool Wageningen, The Netherlands

36. **Kravtchenko, T. P., Arnould, I., Voragen, A. G. J., and Pilnik, W.** (1992) Improvement of the selective depolymerization of pectic substances by chemical -elimination in aqueous solution. Carbohydr. Polym. **19,** 237-242

37. **Shevchik, V. E., Kester, H. C. M., Benen, J. A. E., Visser, J., Robert-Baudouy, J., and Hugouvieux-Cotte-Pattat, N.** (1999) Characterization of the exopolygalacturonate lyase PelX of *Erwinia chrysanthemi* 3937. J. Bacteriol. **18,** 1652-1663

38. **Shevchik, V. E., Condemine, G., Robert-Baudouy, J., and Hugouvieux-Cotte-Pattat, N.** (1999) The exopolygalacturonate lyase PelW and the oligogalacturonate lyase Ogl, two cytoplasmic enzymes of pectin catabolism in *Erwinia chrysanthemi* 3937. J. Bacteriol. **181,** 3912-3919

39. **McDonough, M. A., Ryttersgaard, C., Bjørnvad, M. E., Lo Leggio, L., Schülein, M., Schrøder Glad, S. O., and Larsen, S.** (2002) Crystallization and preliminary X-ray characterization of a thermostable pectate lyase from *Thermotoga maritima.* Acta Crystallogr. Sect. D **58,** 709-711

40. **Adams, M. W. W., and Kelly, R. M.** (2001) *Hyperthermophilic enzymes (Part A)*. Methods Enzymol. **330**, Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto

41. **Veille, C., and Zeikus, G. J.** (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microb. Mol. Biol. Rev. **65,** 1-43

42. **Tardy, F., Nasser, W., Robert-Baudouy, J., and Hugouvieux-Cotte-Pattat, N.** (1997) Comparative analysis of the five major *Erwinia chrysanthemi* pectate lyases: Enzyme characteristics and potential inhibitors. J. Bacteriol. **179,** 2503-2511

43. **Truong, L. V., Tuyen, H., Helmke, E., Binh, L. T., and 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

44. **Kobayashi, T., Hatada, Y., Suzumatsu, A., Saeki, K., Hakamada, Y., and Ito, S.** (2000) Highly alkaline pectate lyase Pel-4A from alkaliphilic *Bacillus* sp. strain P-4-N: its catalytic properties and deduced amino acid sequence. Extremophiles **4,** 377-383

45. **Benen, J. A. E., Kester, H. C. M., Pařenicová, L., and Visser, J.** (2000) Characterization of *Aspergillus niger* pectate lyase A. Biochemistry **39,** 15563-15569

46. **Henrissat, B., Heffron, S. E., Yoder, M. D., Lietzke, S. E., and Jurnak, F.** (1995) Functional implications of structure-based sequence alignment of proteins in the extracellular pectate lyase superfamily. Plant Physiol. **107,** 963-976

6

PelB from the hyperthermophilic bacterium *Thermotoga maritima* encodes a fungal-like intracellular exopolygalacturonase

Leon D. Kluskens, Gert-Jan W.M. van Alebeek, Jasper Walther, Alphons G.J. Voragen, Willem M. de Vos, and John van der Oost

Submitted

Abstract

An intracellular pectinolytic enzyme, PelB, from the hyperthermophilic bacterium *Thermotoga maritima*, was functionally produced in *Escherichia coli* and purified to homogeneity. PelB belongs to family 28 of the glycoside hydrolases, consisting of pectin-hydrolyzing enzymes. As one of the few bacterial exopolygalacturonases, it is able to remove monogalacturonate units from the non-reducing end of polygalacturonate. Detailed characterization of the enzyme showed that PelB is highly thermo-active and thermostable, with a melting temperature of 105°C and the highest temperature optimum described to date for hydrolytic pectinases. PelB showed an increasing activity on oligosaccharides with an increasing degree of polymerization. The highest activity was found on the pentamer (1000 U/mg). In addition, the affinity increased in conjunction with the length of oligoGal*p*A chain. PelB displayed specificity for saturated oligoGal*p*A and was unable to degrade unsaturated or methyl-esterified oligoGal*p*A. Analogous to the exopolygalacturonase from *Aspergillus tubingensis* it demonstrated low activity on xylogalacturonan. Calculations on the subsite affinity displayed the presence of 4 subsites and a high affinity for Gal*p*A at subsite +1, which is typical for exo-active enzymes. The physiological role of PelB and the previously characterized exopectate lyase (PelA) is discussed.

Introduction

Pectin is a complex polysaccharide present in the cell wall of higher plants, where it forms a network by embedding the other cell wall polysaccharides cellulose and hemicellulose. The backbone of the pectin molecule mainly consists of (partly methylated) homogalacturonan, interspersed with rhamnogalacturonan units, which often contain sugar side chains composed of arabinan and galactan (5).

Degradation of the pectin polymer occurs via a set of pectinolytic enzymes, which can roughly be divided into esterases, which remove ferulic acid, methyl or acetyl groups, and depolymerases. The latter can be classified into lyases (β-elimination) and hydrolases (30). All hydrolases involved in degradation of pectin are classified as members of family 28 of the glycoside hydrolases, including the endopolygalacturonases, exopolygalacturonases and rhamnogalacturonases (8, 23). Although a handful of endopolygalacturonases, generally of fungal origin (7, 9, 22, 27, 35, 40), and a single rhamnogalacturonase (26) have been the object of crystallization experiments, a 3-dimensional structure of an exopolygalacturonase is not yet available.

Exo-acting polygalacturonases generally cleave the homogalacturonan part of pectin from the non-reducing end. Exopolygalacturonases (EC 3.2.1.67) are produced by fungi and plants and catalyze the hydrolytic release of monogalacturonic acid. The mostly bacterial exo-poly-α-galacturonosidases (EC 3.2.1.82) liberate digalacturonic acid residues from galacturonan (23, 30).

In recent years, many (hyper)thermophilic organisms have been described with the main emphasis on their starch and cellulose-degrading capacities (4, 41). Although sufficiently present in nature, pectin-degrading (hyper)thermophiles have received relatively little attention (2, 3, 15, 28, 33, 43). Considering their thermostability and –activity, as well as their slightly acidic pH optimum,

galacturonases from these organisms are believed to have potential in the application in clarification processes of fruit juices. Up to now, only a few thermostable pectinolytic enzymes have been characterized in detail (20, 21, 37).

The hyperthermophilic bacterium *Thermotoga maritima* is able to grow on a large variety of simple and complex carbohydrates, such as glucose, maltose, starch, laminarin, xylan and cellulose (6, 16). In addition, we recently reported on its ability to use pectin as a carbon source (20). The *T. maritima* genome sequence revealed the presence of at least two pectinolytic genes (24), one of which, an exopectate lyase (PelA), has been characterized in detail (20). We here report on the overproduction, purification and characterization of an exopolygalacturonase from *T. maritima*, hereafter referred to as PelB[1]. Furthermore, the physiological role and the anticipated synergy between both pectinolytic enzymes of *T. maritima* will be discussed.

Materials and methods

Organisms, growth conditions and plasmids

Thermotoga maritima strain MSB8 (DSM 3109) was grown at 80°C and pH 6.5 as described previously (20). The bacterial strain used for initial cloning experiments was *Escherichia coli* TG1 [*supE hsd* Δ5 *thi* Δ(*lac-proAB*) F' (*tra*D35 *proAB*$^+$ *lacI*^{q} *lac*Z ΔM15). *E. coli* BL21(DE3*) (hsdS gal* (λcl*ts* 857 *ind*1 *S*am7 *nin*5 *lac*UV5-T7 gene 1) was used for heterologous expression. The plasmid used for recombinant work was pET24d from Novagen.

Polygalacturonic acid (PGA) was obtained from ICN. Saturated oligoGal*p*A (DP 2-8) and unsaturated oligoGal*p*A (DP 3-7) were prepared and purified from polygalacturonase and pectin lyase digestions as described previously (38). Methyl-esterification of saturated oligoGal*p*A ((6-O-CH_3-Gal*p*A)$_{4\text{-}6}$) was carried out with anhydrous acidic methanol (39). Modified hairy regions (MHR) were isolated from apple, saponified and used as a source of xylogalacturonan (34).

Recombinant DNA techniques

Genomic DNA of *T. maritima* was isolated by using an established protocol (29). Small-scale plasmid DNA isolation was carried out using the QIAGEN purification kit (Qiagen). DNA was digested with restriction endonucleases and ligated with T4 DNA ligase, according to the manufacturer's specifications (Life Technologies). DNA fragments were purified from agarose by QiaexII or from a PCR mix by using the PCR purification kit (Qiagen). Chemical transformation of *E. coli* TG1 and BL21(DE3) was carried out using established procedures (31).

The gene encoding an exopolygalacturonase (TM0437) was identified in the course of the analysis of the *T. maritima* genome. (24) Primers for gene amplification were designed: BG888 (sense), 5' –CCGGAGGGATGA**CCATGG**AAGAAC (*Nco*I site in bold), and BG889 (antisense), 5' –GCGTCACCTC**GGATCC**TTATTTCAGC (*Bam*HI site in bold). A PCR was carried out on 100 ng genomic DNA of *T. maritima*, following the procedure as described earlier (20). After

digestion with *Nco*I and *Bam*HI, the gene product was cloned in a pET24d expression vector (Novagen). The resulting plasmid, pLUW741, was introduced in *E. coli* TG1 and BL21(DE3).

DNA and amino acid sequence analysis

Cloned PCR products were sequenced by the dideoxynucleotide chain termination method (32) with a Li-Cor automatic sequencing system (model 4000L). DNA and protein sequencing data were analyzed with the DNASTAR package and compared to the GenBank Data Base by BLAST (1). ClustalX and GeneDoc were used for multiple alignment and subsequent adjustment of the exopolygalacturonase amino acid sequence, respectively.

Purification of PelB

E.coli BL21(DE3) harboring pLUW741 was grown overnight (37°C, 150 rpm) in a 5-ml TYK (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 50 µg/ml kanamycin) preculture. One ml was used to inoculate 1 L of TYK in a baffled 2-L Erlenmeyer. Following an overnight growth at 37°C at 120 rpm, the culture was centrifuged for 15 min at 8,500 × *g* at 4°C, medium was discarded and the cells were resuspended in 10 ml Tris-HCl, 20 mM, pH 8.0. The cell suspension was sonicated (3 × 15'') and cell debris was removed by centrifugation at 16,000 × *g* for 10 min. The resulting supernatant was incubated for 20 min at 80°C and precipitated proteins were removed by an additional centrifugation step (16,000 × *g*, 10 min). The heat-stable cell-free extract was loaded onto an ion-exchange chromatography column (Q Sepharose, Amersham Pharmacia Biotech, Inc.), which was equilibrated with Tris-HCl, 20 mM, pH 8.0. Bound proteins were eluted by a linear gradient from 0 to 1 M of NaCl. Fractions containing PelB were pooled and concentrated (Filtron Technology Corp., 30 kDa cutoff). Protein concentrations were spectrophotometrically calculated using the absorption coefficient. Its native conformation was determined by running PelB over a gel filtration column (Superdex 200, Amersham Pharmacia Biotech, Inc.) and comparing it to a set of marker proteins, using 20 mM Tris-HCl, 100 mM NaCl, and pH 8.0 as elution buffer.

Enzyme assays and kinetics

PelB activity was measured by determining the formation of reducing sugar end groups, using the Nelson-Somogyi assay (36). Standard assays were carried out at 80°C in 1 ml of phosphate buffer (100 mM, pH 6.5), containing 0.25% (w/v) polygalacturonic acid (PGA). The reaction was started by the addition of an appropriate amount of PelB and samples were taken at regular time intervals. The reaction was stopped by adding 200 µl of the sample to a Somogyi reagent mix and treated according to the protocol. Finally, the sample was analyzed at 520 nm. 1 U is defined as 1 µmol of reducing end groups released per minute. A 100 mM McIlvaine buffer was used for determining the pH optimum of PelB.

Kinetic constants were measured under optimal enzyme conditions (80°C, pH 6.5) in a 30 mM phosphate buffer using saturated oligogalacturonic acids with a degree of polymerization (DP) of 2 to 8 ((Gal*p*A)$_2$ to (Gal*p*A)$_8$). Care was taken to measure initial reaction rates. K_m and V_{max} were calculated using the Michaelis-Menten fit in Table Curve (SPSS Inc., AISN Software). The

turnover rate (k_{cat}) was calculated from V_{max}, using a calculated molecular mass of 50,483 Da for PelB. The substrate specificity was examined by measuring PelB activity on 1 mM of saturated oligoGal*p*A.

Enzyme reactions used for HPLC analyses were carried out at 80°C in a 30 mM phosphate buffer (pH 6.4). PGA and xylogalacturonan (MHR) were used in 0.25% (w/v) concentrations, (un)saturated oligoGal*p*A and methylated oligoGal*p*A were used in an end concentration of 2 or 2.5 mM. PelB (4.6 ng/ml) was used in an incubation volume of 400 µl. Samples (50 or 100 µl) were taken at time intervals and reactions were stopped by cooling on ice and by addition of 0.4 sample volume of 50 mM NaOH, thereby increasing the pH to 8.0-8.5. Samples were stored at –20°C until analyzed by HPAEC.

HPAEC analysis

High performance size exclusion chromatography (HPAEC) analysis at pH 12 was performed as described previously (39). Saturated and unsaturated oligoGal*p*A were detected using a pulsed amperometric detector (PAD, Electrochemical Detector ED40, Dionex, Sunnyvale, CA). Pure saturated oligoGal*p*A (DP 1-7) were used as standards for external calibration. The amount of product formation was quantified by peak integration (Chromquest). The specific activity (nmoles of product/min/mg protein) was calculated from the formation of saturated oligoGal*p*A in time.

HPSEC analysis

High performance size exclusion chromatography (HPSEC) analyses were done on three TSKgel columns (7.8 mm ID × 30 cm per column) in series (G4000 PWXL, G3000 PWXL, G2500 PWXL; Tosohaas) in combination with a PWX-guard column (Tosohaas). Elution was carried out at 30°C with 0.2 M sodium nitrate at 0.8 ml/min. The eluate was monitored using a refractive index detector. Calibration was performed using dextrans, pectins and oligoGal*p*A.

Differential scanning calorimetry

Thermal unfolding experiments were carried out on a MicroCal VP-DSC in a temperature range from 50 to 125°C at a heating rate of 0.5°C/min. Enzyme samples were dialyzed against a phosphate buffer (50 mM, pH 6.5) prior to analysis.

Calculation of subsite affinities

Subsite affinity values were calculated by using the obtained kinetic data as described by Hiromi and coworkers (12, 13). The subsite affinity A_n was calculated according to the equation:

$$\ln (k_{cat}/Km)_{n+1} - \ln (k_{cat}/K_m)_n = A_{n+1}/RT$$

The parameter k_{cat} was derived from the maximum velocity (V), divided by the molar concentration of the enzyme (e_0, included in V_{max}). R and T are the gas constant and the temperature (in Kelvin), respectively. The values A_{-1} and k_{int} were derived from a plot of $\exp(A_{n+1}/RT)$ against $(1/k_{cat})_n$:

$$\exp(A_{n+1}/RT) = [k_{int}/(k_{cat})_n - 1].\ \exp(A_{-1}/RT)$$

in which A_{-1} and k_{int} are determined from the vertical and horizontal intercepts, respectively.

Results

Molecular characterization of PelB

The *pelB* gene (locus number TM0437) was identified in the *T. maritima* genome and was annotated as a putative exo-poly-α-D-galacturonosidase (24). *PelB* is 1,341 bp in length, which corresponds to an protein with a molecular mass of 50 kDa. The highest sequence similarity at amino acid level was found with a polygalacturonase from the bacterium *Ralstonia solanacearum* (47%) (10). The absence of a clear signal sequence consensus indicates that the enzyme's localization is most likely cytoplasmic (42). *PelB* is positioned in the same gene cluster as the previously described *pelA* gene (20)(Figure 6.1). Comparative gene analysis with the aim of examining the distribution of *pelB* homologs demonstrated no conservation in genome environment in other completely sequenced genomes. The tight clustering with seven surrounding genes in the same transcriptional direction (TM0436-443), with no or small intergenic regions, suggests that *PelB* might be transcribed polycistronically (Figure 6.1). PelB belongs to the large family 28 of the glycoside hydrolases consisting of endopolygalacturonases (EC 3.2.1.15), exopolygalacturonases (EC 3.2.1.67), exo-poly-α-galacturonosidases (EC 3.2.1.82), and rhamnogalacturonases (EC 3.2.1.-) (8). By using ClustalX a multiple sequence alignment was made for the right-handed parallel β-helix domain of a selection of family 28 members (Figure 6.2). Independently, we modeled PelB on EPG2, an endo-active polygalacturonase from *Erwinia carotovora* with low amino acid identity (23%), using the fold recognition server of 3D-PSSM (17). The 3D-structure of EPG2, a right-handed parallel β-helix, has been elucidated previously (27). Sequence conservation predominantly occurs in the regions flanking both catalytic aspartate residues (Asp239 and Asp260, PelB numbering), as well as the residues Asp261 and His296, believed to be of importance in the catalytic process, and Arg327 and Lys329, which may play a role in substrate binding (Figure 6.2) (40).

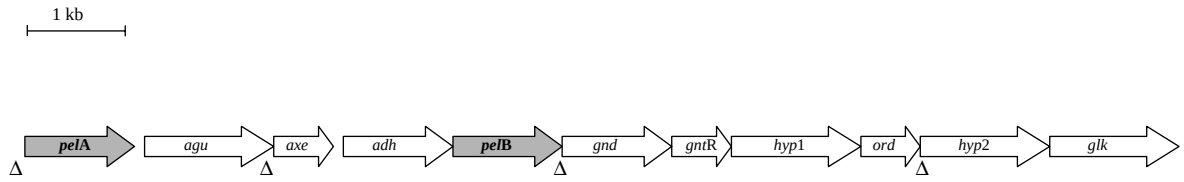


Figure 6.1 Schematic organization of the pectinase gene cluster in *T. maritima* (TM0433-0443). *pel*A and *pel*B are shown as grey arrows. Adjacent genes; α-glucuronidase (*agu*), acetyl xylan esterase (*axe*), Zn^{2+}-containing alcohol dehydrogenase (*adh*), 6-phosphogluconate dehydrogenase, decarboxylating (*gnd*), transcriptional regulator (GntR), oxidoreductase (*ord*), gluconate kinase (*glk*), two conserved hypothetical proteins (*hyp*1 and 2). Intergenic spacing with putative promoter regions is indicated by (Δ).

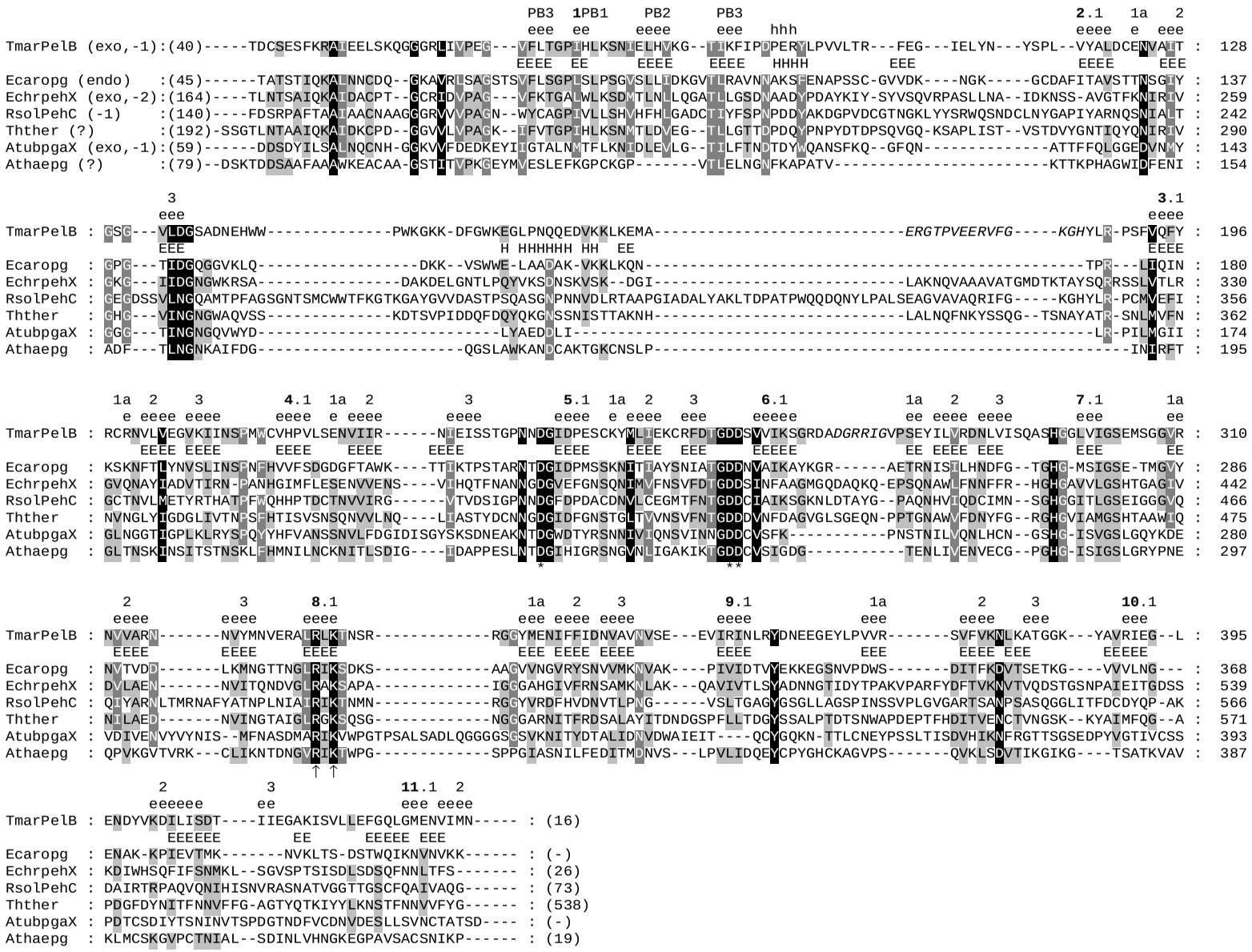


Figure 6.2 Multiple sequence alignment of parallel β-helix segment of family 28 glycoside hydrolases. Sequences (GenBank identifier): PelB *T. maritima* (AAD35522.1), EPG2 *Erwinia carotovora* (CAA35998.1), PehX *Erwinia chrysanthemi* (AAA24842.1), *Ralstonia solanacearum K60* PehC (AAL24033.1), PG *Thermoanaerobacterium thermosulfurigenes* (AAB08040.1), Pgx *Aspergillus tubingensis* (CAA68128.1), Pgx2 *Arabidopsis thaliana* (AAF21195.1). The mode of action (endo or exo) and the amount of Gal*p*A cleaved off, respectively, are annotated in parentheses. A question mark indicates an unknown activity mode. The secondary structure is depicted for *E. carotovora* PG (in capitals, using Expasy's Swiss model, entry 1BHE) and *T. maritima* (small characters, derived from model based upon *E. carotovora* 1BHE in the program 3D-PSSM), (17) for which E (e) indicates strand and H (h) helix. The parallel β-strands (PB1, 1a, 2 and 3) forming 11 coils are shown for *E. carotovora* and *T. maritima* sequences, with the coil number printed in bold. Catalytic residues are indicated by stars, residues supposedly involved in substrate-subsite interaction are highlighted with arrows. Insertions in PelB in comparison to EPG2 are printed in italics.

The predicted secondary structure of PelB corresponds well with that of *E. carotovora* EPG2, with only a few exceptions. Like EPG2, the parallel β-helix comprises of 10 complete turns. PelB contains a few inserted β-strands (.1a), and one large insert of 15 amino acids is present prior to the first β-sheet of coil 3, which is on the edge of the pronounced substrate-binding cleft in EPG2 (Figure 6.2).

Expression and purification

The 1,341-bp *pelB* gene was cloned into a pET24d vector as an *Nco*I-*Bam*HI-fragment, resulting in pLUW741. Introduction in *Escherichia coli* BL21(DE3), resulted in the overproduction of the 50-kDa PelB, which was verified by SDS-PAGE analysis. The enzyme was purified to homogeneity by heat treatment of the cell-free extract, followed by anion-exchange chromatography, where the protein eluted at 0.6 M of NaCl. Analysis of PelB by gel filtration resulted in a peak with an estimated size of 212 kDa, corresponding to results of SDS-PAGE analyses of the unboiled sample, suggesting that the conformation of pelB is a tetramer (not shown).

Enzyme characteristics

PelB was examined by incubation with PGA following standard assay conditions. The experiments showed an increase in reducing power indicating that PelB is active on PGA, the non-methylated homogalacturonan part of the pectin molecule. Hydrolysis of PGA, analyzed by high performance size exclusion chromatography (HPSEC), showed the initial formation of only monogalacturonic acid, with a simultaneous decrease in length of PGA (not shown). Therefore, PelB can be regarded as an exo-acting polygalacturonase. Highest activity using PGA as a substrate was measured at 80ºC, making it the most thermo-active polygalacturonase reported to date. Differential Scanning Calorimetry (DSC) showed that PelB has a melting temperature of 105ºC (not shown). The pH optimum of PelB was determined at 6.4, which appears to be slightly more alkalic than what has been identified for other polygalacturonases up to now. It displays a significant drop in activity when the pH is increased to 7 (not shown). Zymogram experiments were carried out with PelB and *T. maritima* medium fraction and cell extract using PGA as a substrate. This pointed out that PelB

is located intracellularly, which was displayed by a clear activity zone of the cytoplasmic fraction and was absent in equally analyzed medium fractions (not shown).

Mode of action of PelB

To examine its mode of action in more detail, hydrolysis products of oligogalacturonic acids generated by PelB were analyzed by HPAEC. Initial reaction product of all substrates tested was monogalacturonic acid (not shown), indicating that PelB is an exopolygalacturonase (EC 3.2.1.67). The activity rate on 0.25% (w/v) PGA was found to be 6.1 U/mg over the first two hours. A range of Δ4,5 unsaturated oligoGal*p*A, containing a double bond between C4 and C5 at the non-reducing end, was incubated with PelB and analyzed by HPAEC. Unsaturated $(Gal pA)_{3\text{-}5}$ were not hydrolyzed by the enzyme. Since the unsaturated bond on this range of substrates is located at the non-reducing end it can be concluded that PelB is attacking from the non-reducing end. Moreover, fully methylated $(Gal pA)_{4\text{-}6}$ were not hydrolyzed by PelB, indicating that the presence of methyl-esters prevents the enzyme from hydrolyzing oligoGal*p*A. The recent finding of Kester and coworkers that the exopolygalacturonase from *Aspergillus tubingensis* is also active on xylogalacturonan (18) prompted us to test this substrate as well. Analyzed by HPAEC the formation of a D-galacturonate peak could be observed, directly after addition of PelB, which is a result of its established galacturonase activity. Only when high concentrations of PelB were used on xylogalacturonan (25 µg/ml rather than 3.2 ng/ml when assayed on PGA), a minor amount of monoxylogalacturonan was detected in addition to D-galacturonate (not shown).

Enzyme kinetics

PelB activity was initially demonstrated by assaying on PGA as a substrate. Since it seems highly unlikely that the cytoplasmic PelB utilizes the large polymer as its natural substrate, kinetic parameters (K_m and V_{max}) were determined with saturated $(Gal pA)_n$ (n = 2 to 8). PelB (8-16 ng) and substrate (up to 8 mM) were incubated at 80°C for 10 and 15 min. Table 6.1 shows the kinetic parameters for PelB on Gal*p*A ranging from di- to octagalacturonate.

Table 6.1 Kinetic parameters of PelB from *T. maritima* on saturated oligogalacturonates (Gal*p*A) with length n = 2 to 8.

$(Gal pA)_n$	n	K_M (mM)	V_{max} (U/mg)	k_{cat} (s^{-1})	k_{cat}/K_M ($mM^{-1}\,s^{-1}$)
digalacturonate	2	0.34	216	182	534
trigalacturonate	3	0.34	816	685	2016
tetragalacturonate	4	0.29	987	829	2859
pentagalacturonate	5	0.24	1112	934	3892
hexagalacturonate	6	0.11	977	821	7461
heptagalacturonate	7	0.07	1024	860	12288
octagalacturonate	8	0.06	1003	843	14042
polygalacturonate	170	0.06	1170	936	15600

With an increasing DP the substrate affinity increased significantly, up to 0.06 mM for (Gal*p*A)$_8$. The activity rates of PelB (V_{max}) raised until moderating towards a plateau value around 1000 U/mg at (Gal*p*A)$_4$, where k_{cat} values seem to be independent when DP exceeds n = 4. The catalytic efficiency, k_{cat}/K_m, gave a constant increase with an increasing DP, with a value for (Gal*p*A)$_8$ that is almost 30-fold higher than for (Gal*p*A)$_2$ (Table 6.1).

Subsite mapping

Based on the assumptions of Hiromi (12) that the intrinsic rate of hydrolysis (k_{int}) in the productive complex is independent of the length of the substrate K_m and V_{max} values were used to calculate the subsite affinities (see equations Materials and Methods). The subsite affinity A_{n+1} (kJ/mol) was calculated for an enzyme-substrate complex from n = 2 to 5. The intrinsic rate constant k_{int} was determined by plotting exp(A_{n+1}/RT) against $(1/k_{cat})_n$, which also allowed us to calculate the binding affinity for subsite -1. k_{int} was found to be 262 s^{-1}. Affinity values are shown in Figure 6.3 as a schematic representation of the subsite binding cleft of PelB.

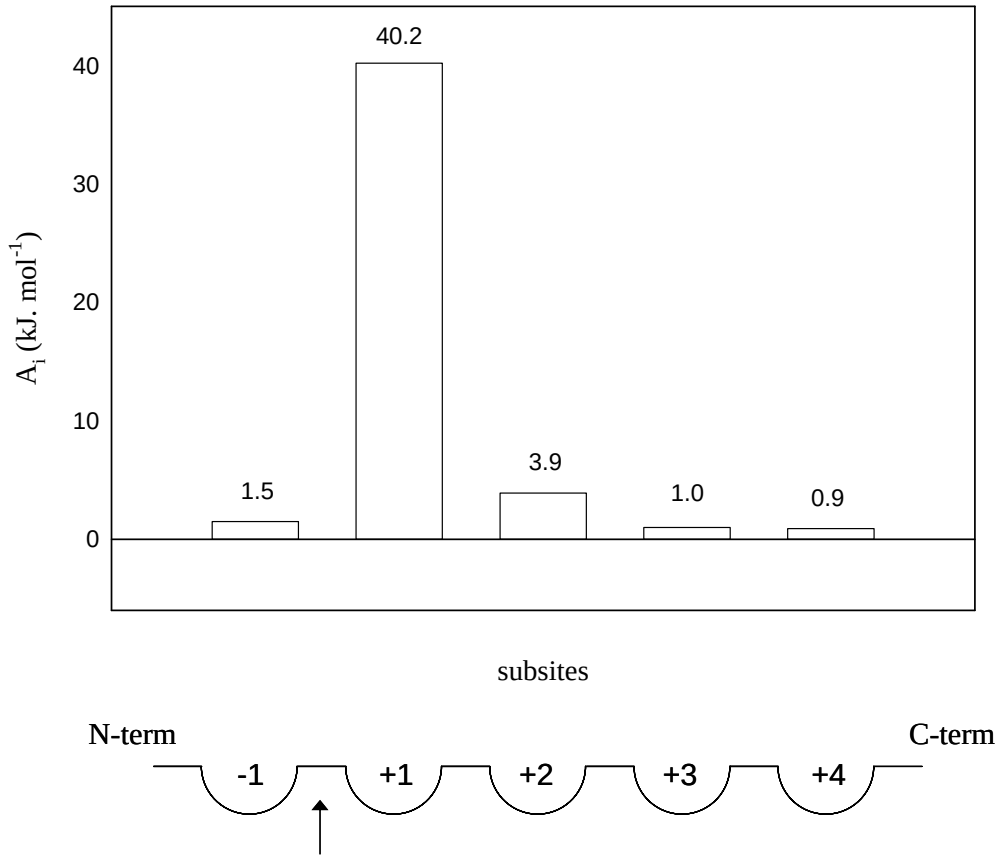


Figure 6.3 Schematic representation of the subsite map of exopolygalacturonase PelB. Subsites are numbered from -1 to +4, with the non-reducing sugar end facing the N-terminus of the enzyme. Binding affinity values are illustrated by bar diagrams. The catalytic cleavage site is indicated by an arrow.

The highest binding affinity was found for the penultimate subsite +1 (40.2 kJ.mol^{-1}), whereas the affinity lowered considerably when moving towards the reducing end of the substrate. Along with its exocleaving activity, thereby liberating monogalacturonic acid, the catalytic site of PelB should be located in between subsites -1 and +1 (Figure 6.3). Comparative modeling previously showed that the binding cleft of polygalacturonases can maximally hold eight Gal*p*A residues, resulting in a subsite order from -5 to +3 (7). Since the substrate most likely binds with the non-reducing end towards the N-terminus of the enzyme (25), this implies that PelB most probably contains 4 subsites, going from -1 to +3.

Discussion

A pectinolytic hydrolase (PelB) from the hyperthermophilic bacterium *T. maritima* was heterologously produced and purified to homogeneity. Detailed characterization of this enzyme is described in this paper, which is a continuation on the recent report of an exopectate lyase (PelA) from the same organism (20). Despite its clear exocleaving characteristics, the highest similarity at amino acid level was found with family 28 endopolygalacturonases (EC 3.2.1.15), although it should be noted that the number of available endopolygalacturonase sequences is surpassing that for exocleaving galacturonate hydrolases. The apparent absence of a signal peptide and the detection of pectinolytic activity in the cell fraction and not in the medium fraction supported our belief in PelB being cytoplasmic, in contrast to the majority of the polygalacturonases that has been examined. Optimal activity on homogalacturonic acid was observed at 80°C, making it the most thermoactive hydrolase active on this polysaccharide to date. Because of their catalytic and stability properties, thermostable pectinolytic enzymes can be of great use in industrial processes. Considering its slightly acidic pH optimum of 6.4, PelB might be applicable in the fruit-juice industry, where it can be included in the clarification or color extraction steps, which are often carried out at elevated temperatures.

Although bacterial exo-acting polygalacturonases commonly generate digalacturonate, PelB was shown to liberate monogalacturonic acid as the first and only product on PGA and oligoGal*p*A. Based on its mode of action, PelB should be classified as an exopolygalacturonase (pgx; EC 3.2.1.67). To date, no crystal structure of an exopolygalacturonase is available. Since PelB has high similarity on primary structure level with endopolygalacturonases, especially around the catalytic regions (23), we assume that the substrate binds with the non-reducing sugar end directing towards the N-terminus of the enzyme, as has been suggested for endopolygalacturonases by Pagès and coworkers (25). Perhaps the large insertion prior to coil 3 is containing residues that might play a role in obstructing the substrate-subsite -2 interaction, although this insertion seems to be absent in the exo-active *A. tubingensis* polygalacturonase. Cho and coworkers (7) described the amino acid residues in *Aspergillus aculeatus* polygalacturonase involved in hydrogen bonding interactions between the substrate binding-cleft residues and octaGal*p*A, and aligned the equivalent residues of *E. carotovora* EPG2. Two residues believed to be involved in substrate binding at subsite -2 in *E. carotovora* EPG2, namely Arg152 binding the carboxyl group and Lys229 interacting with 2-OH,

are also conserved in PelB and *A. tubingensis* Pgx. Direct obstruction of a possible Gal*p*A-interaction with its equivalent subsite -2 might therefore be brought about by adjacent residues. While phylogenetically classified amongst the bacterial endopolygalacturonases (23), PelB displays characteristics that clearly bear more resemblance to the group of fungal exopolygalacturonases (3.2.1.67). Obviously, the primary structure alone restricts us to explain PelB's mode of action in more detail. Considering the homology between exo- and endogalacturonases, the difference in mode of action is most probably depending on subtle changes in the catalytic and/or substrate binding region. Unfortunately, only a few exopolygalacturonases have been fully characterized and identified and therefore the amount of available sequences is limited.

Exopolygalacturonases that liberate monogalacturonate (EC 3.2.1.67) are generally produced by fungi and plants, with the exception of one originating from the bovine ruminal bacterium *Butyrivibrio fibrisolvens* (11). Like PelB, this enzyme too is localized intracellularly. *B. fibrisolvens* also contains an exopectate lyase that generates unsaturated trigalacturonates, similar to PelA. To the authors' knowledge, both *T. maritima* and *B. fibrisolvens* are the only two bacteria described that contain such a similar combination of pectinolytic enzymes, although the exopolygalacturonase from *B. fibrisolvens* was shown to degrade both saturated and Δ4,5 unsaturated oligoGal*p*A (44).

Kinetic analyses have shown that PelB hydrolyzes oligoGal*p*A very rapidly with an increasing affinity for longer oligoGal*p*A. The specific activity (reaching a plateau for $(Gal pA)_4$ around 1000 U/mg) is among the highest known for polygalacturonases, and the highest of all oligoGal*p*A-active exohydrolases. The highest affinity was found for the subsite +1. This high value is typical for exo-active hydrolytic enzymes, such as the exopolygalacturonase from *A. tubingensis* and a Barley β-D-glucosidase (14, 19). The absolute value however (+40.2 kJ/mol) is much higher than has been reported earlier for this subsite. The reason for this might be the thermo-active character of the enzyme, which obliges PelB to bind its substrate tightly enough at high temperatures. An affinity value closer to mesophilic values might lead to a spontaneous dissociation of the substrate-subsite complex. The intrinsic rate constant, k_{int}, is rather low compared to the highest values found for k_{cat}.

Cho and coworkers tested kinetic models of octagalacturonate, using three polygalacturonases (including *A. aculeatus* PG), and concluded that the binding clefts in polygalacturonases can accommodate maximally eight Gal*p*A residues at subsites from -5 to +3 (7). Along with the suggestions of Pagès and coworkers that the Gal*p*A is binding with the non-reducing end directing towards the N-terminus of the enzyme (25), PelB can accomodate only four subsites in total, namely from –1 to +3, which was shown by the activity rate that reached a maximum at $(Gal pA)_4$ (Table 6.1). However, the catalytic efficiency factor (k_{cat}/K_m) still increases with an increase in DP of the substrate, thereby suggesting an extended substrate-binding region. According to this model, oligoGal*p*A exceeding a DP of 4 will comprise Gal*p*A monomers at the reducing end that are presumably exposed to the solvent region. This preference for longer oligoGal*p*A seems to be in conflict with its cytoplasmic character and may perhaps be due to conformational changes of the substrate, thereby facilitating binding to the substrate-binding cleft.

T. maritima contains at least two evident pectinolytic enzymes. PelA appears to be the only extracellular enzyme in *T. maritima* able to depolymerize the homogalacturonic acid part of pectin

into, predominantly, unsaturated trigalacturonates (20). However, PelB's inability to degrade these intermediates suggests an intermediate conversion of unsaturated oligoGal*p*A. Although the unsaturated oligoGal*p*A tolerated high-temperature conditions without being degraded, other *in vivo* factors besides temperature and pH might play a, to date indefinite, role in its stability. Alternatively, unsaturated oligoGal*p*A may be saturated by another, yet unidentified, pectinolytic enzyme. To address questions like these, we are currently using DNA microarray analyses to get insight in the complete set of genes involved in pectin catabolism by *T. maritima*.

Acknowledgments

The authors are very thankful to Dr. J.A.E. Benen and Ing. H.C.M. Kester for helpful discussions and Ans Geerling for technical assistance.

References

1. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. H. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389-3402.
2. **Bonch-Osmolovskaya, E. A., A. I. Slesarev, M. L. Miroshnichenko, T. P. Svetlichnaya, and V. A. Alekseev.** 1988. Characteristics of *Desulfurococcus amyloliticus* sp. nov., a new extremely thermophilic archaebacterium isolated from thermal springs of Kamchatka and Kunashir Island. Mikrobiologiya. **57:**94-101.
3. **Bredholt, S., J. Sonne-Hansen, P. Nielsen, I. M. Mathrani, and B. K. Ahring.** 1999. *Caldicellulosiruptor kristjanssonii* sp. nov., a cellulolytic, extremely thermophilic, anaerobic bacterium. Int. J. Syst. Bacteriol. **49:**991-996.
4. **Bruins, M. E., A. E. M. Janssen, and R. M. Boom.** 2001. Thermozymes and their applications. Appl. Biochem. Biotechnol. **90:**155-186.
5. **Carpita, N. C., and D. M. Gibeaut.** 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. The Plant Rev. **3:**1-30.
6. **Chhabra, S. R., K. R. Shockley, D. E. Ward, and R. M. Kelly.** 2002. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. Appl. Environ. Microbiol. **68:**545-554.
7. **Cho, S. W., S. Lee, and W. Shin.** 2001. The X-ray structure of *Aspergillus aculeatus* polygalacturonase and a modeled structure of the polygalacturonase-octagalacturonate complex. J. Mol. Biol. **311:**863-878.
8. **Coutinho, P. M., and B. Henrissat.** 1999. Carbohydrate-active enzymes server. http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html.
9. **Federici, L., C. Caprari, B. Mattei, C. Savino, A. Di Matteo, G. De Lorenzo, F. Cervone, and D. Tsernoglou.** 2001. Structural requirements of endopolygalacturonase for the interaction with PGIP (polygalacturonase-inhibiting protein). Proc. Natl. Acad. Sci. U.S.A. **98:**13425-13430.
10. **González, E. T., and C. Allen.** 2003. Characterization of a *Ralstonia solanacearum* operon required for polygalacturonate degradation and uptake of galacturonic acid. Mol. Plant-Microbe Interact. **16:**536-544.

11. **Heinrichova, K., M. Wojciechowicz, and A. Ziolecki.** 1985. An exo-D-galacturonanase of *Butyrivibrio fibrisolvens* from the bovine rumen. J. Gen. Microbiol. **131:**2053-2058.
12. **Hiromi, K.** 1970. Interpretation of dependency of rate parameters on the degree of polymerization of substrate in enzyme-catalyzed reactions. Evaluation of subsite affinities of exo-enzyme. Biochem. Biophys. Res. Commun. **40:**1-6.
13. **Hiromi, K., Y. Nitta, C. Numata, and S. Ono.** 1973. Subsite affinities of glucoamylase: examination of the validity of the subsite theory. Biochim. Biophys. Acta. **302:**362-375.
14. **Hrmova, M., E. A. MacGregor, P. Biely, R. J. Stewart, and G. B. Fincher.** 1998. Substrate binding and catalytic mechanism of a Barley β-D-glucosidase/(1,4)-β-D-glucan exohydrolase. J. Biol. Chem. **273:**11134-11143.
15. **Huang, C.-Y., B. K. Patel, R. A. Mah, and L. Baresi.** 1998. *Caldicellulosiruptor owensensis* sp. nov., an anaerobic, extremely thermophilic, xylanolytic bacterium. Int. J. Syst. Bacteriol. **48:**91-97.
16. **Huber, R., T. A. Langworthy, H. König, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. Arch. Microbiol. **144:**324-333.
17. **Kelley, L. A., R. M. MacCallum, and M. J. E. Sternberg.** 2000. Enhanced genome annotation using structural profiles in the program 3D-PSSM. J. Mol. Biol. **299:**499-520.
18. **Kester, H. C. M., J. A. E. Benen, and J. Visser.** 1999. The exopolygalacturonase from *Aspergillus tubingensis* is also active on xylogalacturonan. Biotechnol. Appl. Biochem. **30:**53-57.
19. **Kester, H. C. M., M. A. Kusters-Van Someren, Y. Muller, and J. Visser.** 1996. Primary structure and characterization of an exopolygalacturonase from *Aspergillus tubingensis*. Eur. J. Biochem. **240:**738-746.
20. **Kluskens, L. D., G.-J. W. M. van Alebeek, A. G. J. Voragen, W. M. de Vos, and J. van der Oost.** 2003. Molecular and biochemical characterisation of the thermo-active family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima.* Biochem. J. **370:**651-659.
21. **Kozianowski, G., F. Canganella, F. A. Rainey, H. Hippe, and G. Antranikian.** 1997. Purification and characterization of thermostable pectate-lyases from a newly isolated thermophilic bacterium, *Thermoanaerobacter italicus* sp. nov. Extremophiles. **1:**171-182.
22. **Lu, C. T., T. Tada, Y. Nakamura, K. Wada, K. Nishimura, Y. Katsuya, M. Sawada, M. Takao, and T. Sakai.** 2000. Crystallization and preliminary X-ray analysis of endopolygalacturonase SE1 from *Trichosporon penicillatum.* Acta Crystallogr. D. **56:**1668-1669.
23. **Markovič, O., and Š. Janeček.** 2001. Pectin degrading glycoside hydrolases of family 28: sequence-structural features, specificities and evolution. Protein Eng. **14:**615-631.
24. **Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, L. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, O. White, S. L. Salzberg, H. O. Smith, J. C. Venter, and C. M. Fraser** 1999. Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima.* Nature. **399:**323-329.
25. **Pagès, S., W. H. Heijne, H. C. M. Kester, J. Visser, and J. A. E. Benen.** 2000. Subsite mapping of *Aspergillus niger* endopolygalacturonase II as studied by site-directed mutagenesis. J. Biol. Chem. **275:**29348-29353.
26. **Petersen, T. N., S. Kauppinen, and S. Larsen.** 1997. The crystal structure of rhamnogalacturonase A from *Aspergillus aculeatus*: a right-handed parallel beta helix. Structure. **5:**533-544.

27. **Pickersgill, R., D. Smith, K. Worboys, and J. Jenkins.** 1998. Crystal structure of polygalacturonase from *Erwinia carotovora* ssp. *carotovora.* J. Biol. Chem. **273:**24660-24664.
28. **Rainey, F. A., A. M. Donnison, P. H. Janssen, D. Saul, A. Rodrigo, P. L. Bergquist, R. M. Daniel, E. Stackebrandt, and H. W. Morgan.** 1994. Description of *Caldicellulosiruptor saccharolyticus* gen. nov., sp. nov: An obligately anaerobic, extremely thermophilic, cellulolytic bacterium. FEMS Microbiol. Lett. **120:**263-266.
29. **Ramakrishnan, V., and M. W. W. Adams.** 1995. *In*: Archaea, A Laboratory Manual: Thermophiles, p. 95-96. F. T. Robb, and A. R. Place (eds), Preparation of genomic DNA from sulfur-dependent hyperthermophilic Archaea. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. **Sakai, T., T. Sakamoto, J. Hallaert, and E. J. Vandamme.** 1993. Pectin, pectinase and protopectinase: production, properties, and applications. Adv. Appl. Microbiol. **39:**213-294.
31. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** (eds.) 1989 Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors Proc. Natl. Acad. Sci. U.S.A. **74:**5463-5467.
33. **Schink, B., and J. G. Zeikus.** 1983. Characterization of pectinolytic enzymes of *Clostridium thermosulfurogenes.* FEMS Microbiol. Lett. **17:**295-298.
34. **Schols, H. A., E. J. Bakx, D. Schipper, and A. G. J. Voragen.** 1995. A xylogalacturonan subunit present in the modified hairy regions of apple pectin. Carbohydrate Res. **279:**265-279.
35. **Shimizu, T., T. Nakatsu, K. Miyairi, T. Okuno, and H. Kato.** 2002. Active-site architecture of endopolygalacturonase I from *Stereum purpureum* revealed by crystal structures in native and ligand-bound forms at atomic resolution. Biochemistry. **41:**6651-6659.
36. **Somogyi, M.** 1952. Notes on sugar determination. J. Biol. Chem. **195:**19-23.
37. **Takao, M., T. Nakaniwa, K. Yoshikawa, T. Terashita, and T. Sakai.** 2000. Purification and characterization of thermostable pectate lyase with protopectinase activity from thermophilic *Bacillus* sp. TS 47. Biosci. Biotechnol. Biochem. **64:**2360-2367.
38. **van Alebeek, G.-J. W. M., T. M. I. E. Christensen, H. A. Schols, J. D. Mikkelsen, and A. G. J. Voragen.** 2002. Mode of action of pectin lyase A of *Aspergillus niger* on differently C_6-substituted oligogalacturonides. J. Biol. Chem. **277:**25929-25936.
39. **van Alebeek, G.-J. W. M., O. Zabotina, G. Beldman, H. A. Schols, and A. G. J. Voragen.** 2000. Esterification and glycosydation of oligogalacturonides: examination of the reaction products using MALDI-TOF MS and HPAEC. Carbohydr. Polymers. **43:**39-46.
40. **van Santen, Y., J. A. E. Benen, K.-H. Schroter, K. H. Kalk, S. Armand, J. Visser, and B. W. Dijkstra.** 1999. 1.68-A crystal structure of endopolygalacturonase II from *Aspergillus niger* and identification of active site residues by site-directed mutagenesis. J. Biol. Chem. **274:**30474-30480.
41. **Veille, C., and G. J. Zeikus.** 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol. Mol. Biol. Rev. **65:**1-43.
42. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14:**4683-4690.
43. **Wiegel, J., L. G. Ljungdahl, and J. R. Rawson.** 1979. Isolation from soil and properties of the extreme thermophile *Clostridium thermohydrosulfuricum.* J. Bacteriol. **139:**800-810.
44. **Wojciechowicz, M., K. Heinrichova, and A. Ziolecki.** 1982. An exopectate lyase of *Butyrivibrio fibrisolvens* from the bovine rumen. J. Gen. Microbiol. **128:**2661-2665.

7

Transcriptional analysis of pectin catabolic pathways in the hyperthermophile *Thermotoga maritima*

Leon D. Kluskens, Keith H. Shockley, Donald E. Ward, Shannon B. Conners, Robert M. Kelly, Willem M. de Vos, and John van der Oost

Submitted

Abstract

A careful analysis has been performed of the genome sequence of the hyperthermophilic bacterium *Thermotoga maritima*, suggesting that two gene clusters are involved in the catabolism of the sugar polymer pectin. One cluster codes for homologs of enzymes involved in pectin depolymerization, uptake and modification, while the other cluster appears to encode enzymes responsible for cytoplasmic catabolism of saturated and unsaturated galacturonate. A targeted microarray was used to examine differential expression of these and selected other *T. maritima* genes during growth on pectin and cellobiose. Expression of the genes encoding the characterized exopectate lyase PelA and exopolygalacturonase PelB, was highly elevated on pectin (23.5-fold and 12.7-fold, respectively), together with that of a putative third depolymerizing enzyme, a predicted α-glucuronidase (3.3-fold induction). Additional genes in this cluster also showed induction during growth on pectin, including an alcohol dehydrogenase and a putative methylesterase, proposed to play a role in the cytoplasmic release and conversion of methyl groups from pectin oligosaccharides. Furthermore, the expression of several genes in the galacturonate catabolic cluster was strongly upregulated during growth on pectin. A KDG kinase and a KDPG aldolase, present in the same cluster, showed similar induction levels and most likely link galacturonate catabolism with glycolysis. The KdgR homolog positioned upstream of this cluster may play a role in the regulation of the expression of at least some or all pectinolytic genes in *T. maritima*.

Introduction

The hyperthermophilic anaerobe *Thermotoga maritima* grows optimally at 80°C and is capable of transporting and utilizing a variety of simple and complex saccharides (10). The genome of this organism appears to encode the highest number of glycosidases of any sequenced hyperthermophilic genome, regardless of size (Chhabra and Kelly, unpublished observation). A number of these enzymes, involved in the degradation of xylan, cellulose and mannan, have been characterized in detail previously (1, 4). The availability of the whole genome sequence of *T. maritima* has significantly accelerated the identification and characterization of additional polysaccharide-degrading enzymes. It was recently found that *T. maritima* is capable of utilizing pectin as a carbon source as well (15), in agreement with the presence of two characterized pectinolytic enzymes in its genome, the exopectate lyase *pel*A (TM0433) and exopolygalacturonase *pel*B (TM0437) (15, 16, 28). The *pel*A and *pel*B genes are located within a larger gene cluster (TM0430 to TM0443), suggesting that additional members may encode proteins involved in the stepwise degradation of pectin. Although the regulatory mechanism for the *T. maritima* pectin regulon has not been studied, a *gnt*R-like gene is present in the *pel*A/*pel*B cluster. In addition, a homolog of the KdgR pectin repressor from *Erwinia chrysanthemi* (TM0065, (47)), is located in a distinct gene cluster with homologs to *uxa*C, *uxu*AB, and *kdg*AK, genes previously implicated in the intracellular catabolism of sugar acids like galacturonate in other organisms (36).

The utility of DNA microarrays for the analysis of sugar-specific gene expression has been demonstrated previously (3). To investigate details of pectin catabolism in *T. maritima*, predicted

on the basis of its genome content, we performed a targeted cDNA microarray analysis in which 267 *T. maritima* genes were included that potentially encode enzymes, transporters and regulators of sugar and protein metabolism. Apart from two main clusters of predicted pectin-related genes during growth on pectin, the results also reveal the induction of additional genes that may be involved in pectin degradation in *T. maritima.*

Materials and methods

Growth of T. maritima, DNA and RNA isolation

T. maritima MSB8 was grown anaerobically at 80°C, pH 6.5, in 120-ml or 1-L bottles under mildly-shaken conditions. Cellobiose was chosen as a comparison condition for pectin after observation that under otherwise identical conditions, growth of *T. maritima* was similar whether cellobiose or pectin was used as a carbon source. The medium composition (1 liter) was: 40.0 g of sea salt, 3.1 g of PIPES, 1.0 g of yeast extract, 4.0 g of tryptone, 2 ml of trace elements, and 1 ml (0.5 g/L) of resazurin. Filter sterilized (10 mM cellobiose, 20 mM glucose) or autoclaved (pectin, 0.25% (w/v), 65% DM) carbon source was added separately from a stock solution after sterilization. Medium was flushed with N_2 and reduced by adding Na_2S (0.5 g/L) prior to inoculation (1%). Growth was monitored by determining the optical density at 600 nm or by epifluorescent cell count using acridine orange stain (2). For DNA isolation *T. maritima* was grown overnight in 120-ml bottles, containing 30 ml of medium. Genomic DNA extraction was carried out as described previously (44). For RNA isolation *T. maritima* cultures were grown until the early to mid logarithmic growth phase, corresponding to a doubling time (t_D) of 54 min for cellobiose-grown cells (after 6 h) and 156 min for pectin-grown cells (after 11 h). Residual pectin was removed by passing the culture through a filter. Cells were harvested by centrifugation at $8{,}000 \times g$ for 15 min and RNA was isolated as described previously (38), for which the ethanol precipitation step was followed by purification using the RNA isolation kit from Promega (SV RNA isolation). Quality and quantity of the RNA was determined by spectrophotometrical inspection and agarose gel electrophoresis.

Primer design and PCR

The genome sequence of *T. maritima* (26) was obtained from the website of The Institute of Genome Research (http://www.tigr.org/tigr-scripts/CMR2/GenomePage3.spl?database=btm). Open reading frames (ORFs) of genes selected for DNA microarray analysis were imported into Vector NTI Suite (Informax, Inc.) and used as templates for PCR primer design to generate PCR fragments of around 600 to 800 bps. A complete overview of selected genes has been previously reported (3) and included all ORFs predicted to be involved in pectin transport, degradation, and metabolism.

Targeted T. maritima cDNA microarray and data analysis

Amplifications were performed with *Taq* polymerase on 80 ng of genomic DNA for each PCR reaction (35 rounds), with an annealing temperature of 50°C. Generated PCR products were purified (Qiagen) and resuspended in 50% (v/v) dimethyl sulfoxide, to reach a final printing concentration of 50 ng/µl. Following randomization, eight replicates of each PCR product were spotted onto CMT-GAPS aminosilane-coated microscope slides (Corning, Corning, NY) with a 417 Arrayer (Affymetrix, Santa Clara, CA). UV crosslinking was performed at 75°C for 2 h in a GS GeneLinker UV Chamber (Biorad, Hercules, CA) at 250 mJ.

cDNA preparation and hybridization

Total *T. maritima* RNA was isolated as described above and cDNA was prepared using Stratascript (Stratagene) and random hexamer primers (Invitrogen Life Technologies, Carlsbad, CA). Generated cDNA products were purified (Qiagen) and reacted with monoreactive Cy-3 and Cy-5 NHS-esters (Amersham Biosciences, Inc., Piscataway, NJ). Hybridizations and washing steps were carried out as described by Hegde and coworkers (9), although no poly-T was added to the hybridization mixture. Seven slides were used for hybridizations. These included three slides on which pectin (Cy5) and cellobiose (Cy3) were compared, three slides on which pectin with tryptone (Cy5) and cellobiose with tryptone (Cy3) were compared, and one on which pectin was labeled with Cy3 and Cy5 to assess dye effects. Results observed with cellobiose and cellobiose plus tryptone were largely similar for interesting candidate genes, so fold change data from the cellobiose-pectin comparison is represented here, although the "with tryptone" cases are utilized for clustering.

Data analysis

Slides were scanned using the Scanarray 4000 scanner (GSI Lumonics, Canada Billerica, MA) at the Genome Research Laboratory at North Carolina State University. Quantarray (GSI Lumonics) was used for the extraction of the signal intensity data. Local background values were subtracted from each corresponding intensity spot before analysis and negative values set to 1. Log_2-transformed background-corrected intensities were utilized to perform mixed model ANOVA analysis (46) using models described previously to correct for systematic biases between arrays, dyes, pins, spots, and blocks (3). A Bonferroni correction was applied to ensure an experiment-wide false positive rate of $\alpha = 0.05$, dividing $\alpha = 0.05$ by 1573, the number of comparisons performed over all four treatment combinations. This yielded a corrected p-value of 3.17×10^{-5} ($-\log_{10}$(p-value) = 4.5) associated with an experiment-wide false positive rate of $\alpha = 0.05$. The volcano plot for the cellobiose-pectin comparison was constructed using the $\log_2$-transformed fold change and the $-\log_{10}$-transformed significance for all 267 genes on pectin as compared to cellobiose (46).

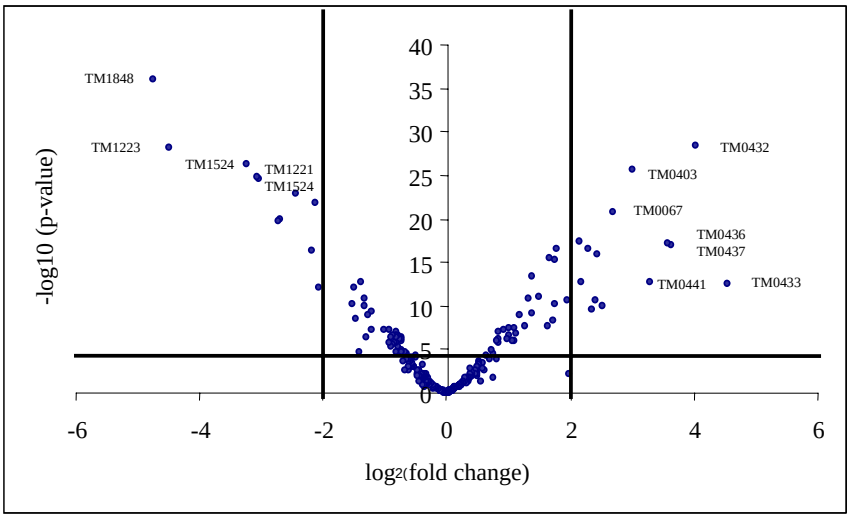


Figure 7.1 Volcano plot illustrating the range of fold changes observed for genes when comparing expression on pectin to expression on cellobiose. Log_2-transformed fold changes are plotted on the x-axis while log_{10}-transformed p-values are plotted on the y-axis. The horizontal line represents a Bonferroni-corrected significance of $\alpha = 0.05$. The vertical lines represent a fourfold change in expression in either direction. Genes in the upper left quadrant represent genes expressed more highly on cellobiose than pectin, while genes in the upper right quadrant represent genes expressed more highly on pectin than cellobiose. TM-numbers of significance for this study are explained in Table 7.1, details on the remaining ORFs can be found on http://webbie.che.ncsu.edu/extremophiles/microarray/index.html.

Results and discussion

Transcriptional analysis by targeted cDNA microarray revealed the induction of a number of predicted and known pectin-active genes during growth on pectin as compared to cellobiose. Overall patterns of difference between pectin and cellobiose are shown as a volcano plot (Figure 7.1). After a conservative Bonferroni adjustment of significance level (indicated by the horizontal line in Figure 7.1), the expression of a number of genes is significantly changed, with 11 genes showing four-fold or higher upregulation on cellobiose and 14 genes showing four-fold or higher upregulation on pectin. Associated fold changes and p-values for all ORFs relevant for the present study are listed by gene identification number (Table 7.1). Results for the remainder of genes are available as a web annex at http://webbie.che.ncsu.edu/extremophiles/microarray/ index.html. A cluster of genes was found with relatively high expression during growth on both pectin and pectin-tryptone, including a number of predicted pectin-active genes (Figure 7.2).

Table 7.1 Overview of *T. maritima* ORFs, corresponding function and expression level when carbon source is switched from cellobiose to pectin.

TM	ORF[a]	Protein description	Mean intensity ratio ($log_2 \pm SD$)	Reference
0061	AAD35155	β-endo-1,4-xylanase (*xyn*A)	0.49 ± 0.17	(45)
0062	AAD35156	carbohydrate-binding module	1.00 ± 0.16	(5)
0063	AAD35157	hyp. protein	0.57 ± 0.18	-
0064	AAD35158	D-galacturonate isomerase (*uxa*C)	0.59 ± 0.16	c
0065	AAD35159	transcr. regulator (*kdg*R)	2.14 ± 0.17	(47)
0066	AAD35160	2-keto-3-deoxy-phosphogluconate aldolase	2.36 ± 0.31	(8)
0067	AAD35161	2-keto-3-deoxy-gluconate kinase	2.69 ± 0.17	b
0068	AAD35162	D-altronate oxidoreductase (*uxa*B)	2.41 ± 0.29	c
0069	AAD35163	D-altronate dehydratase (*uxa*A)	2.28 ± 0.19	c
0070	AAD35164	xylanase (*xyn*B)	0.10 ± 0.16	(45)
0430	AAD35515	sugar ABC transporter, permease	1.09 ± 0.17	b
0431	AAD35516	sugar ABC transporter, permease	1.97 ± 0.70	b
0432	AAD35517	sugar ABC transporter, sugar-binding protein	4.04 ± 0.18	b
0433	AAD35518	pectate lyase (*pel*A)	4.55 ± 0.47	(15)
0434	AAD35519	α-glucuronidase (*agu*)	1.74 ± 0.15	(39)
0435	AAD35520	acetyl xylan esterase-rel. protein	1.95 ± 0.23	b
0436	AAD35521	Zn^{2+}-containing alcohol dehydrogenase	3.56 ± 0.28	b
0437	AAD35522	exopolygalacturonase (*pel*B)	3.63 ± 0.29	(16)
0438	AAD35523	6-phosphogluconate dehydrogenase	2.44 ± 0. 21	b
0439	AAD35524	transcr. regulator, gntR family	1.67 ± 0.14	b
0440	AAD35525	hyp. protein	1.63 ± 0.25	-
0441	AAD35526	oxidoreductase, short-chain dehydrogenase (*kdu*D)	3.28 ± 0.34	c
0442	AAD35527	cons. hyp. protein	2.52 ± 0.32	-
0443	AAD35528	gluconate kinase	2.17 ± 0.22	b
0209	AAD35301	6-phosphofructokinase	-0.40 ± 0.17	(6)
0710	AAD35792	transcr. regulator, MarR family	-0.32 ± 0.18	b
0767	AAD35849	maltodextrin glycosyl transferase	-2.11 ± 0.13	b
0816	AAD35989	transcr. regulator, put. Mar family	-0.38 ± 0.38	b
1017	AAD36094	pecM-like protein	1.75 ± 0.22	c
1201	AAD36276	arabinogalactan endo-1,4-β-galactosidase, put.	0.00 ± 0.16	b
1218	AAD36293	transcr. regulator, LacI family	-1.41 ± 0.30	b
1219	AAD36294	oligopeptide ABC transporter, ATP-binding protein	-2.43 ± 0.14	b
1220	AAD36295	oligopeptide ABC transporter, ATP-binding protein	-2.68 ± 0.18	b
1221	AAD36296	oligopeptide ABC transporter, permease	-3.07 ± 0.16	b
1222	AAD36297	oligopeptide ABC transporter, permease	-2.73 ± 0.19	b
1223	AAD36298	oligopeptide ABC transporter, periplasmic oligopeptide-binding prt.	-4.48 ± 0.20	b
1524	AAD36591	endoglucanase (*cel*A)	-3.23 ± 0.16	(18)
1525	AAD36592	endoglucanase (*cel*B)	-3.02 ± 0.16	(18)
1835	AAD36898	cyclomaltodextrinase, put.	-1.29 ± 0.22	b
1848	AAD36910	cellobiose-phosphorylase (*cep*A)	-4.75 ± 0.15	(31)

a GenBank identification number
b Nelson et al. (26)
c this study

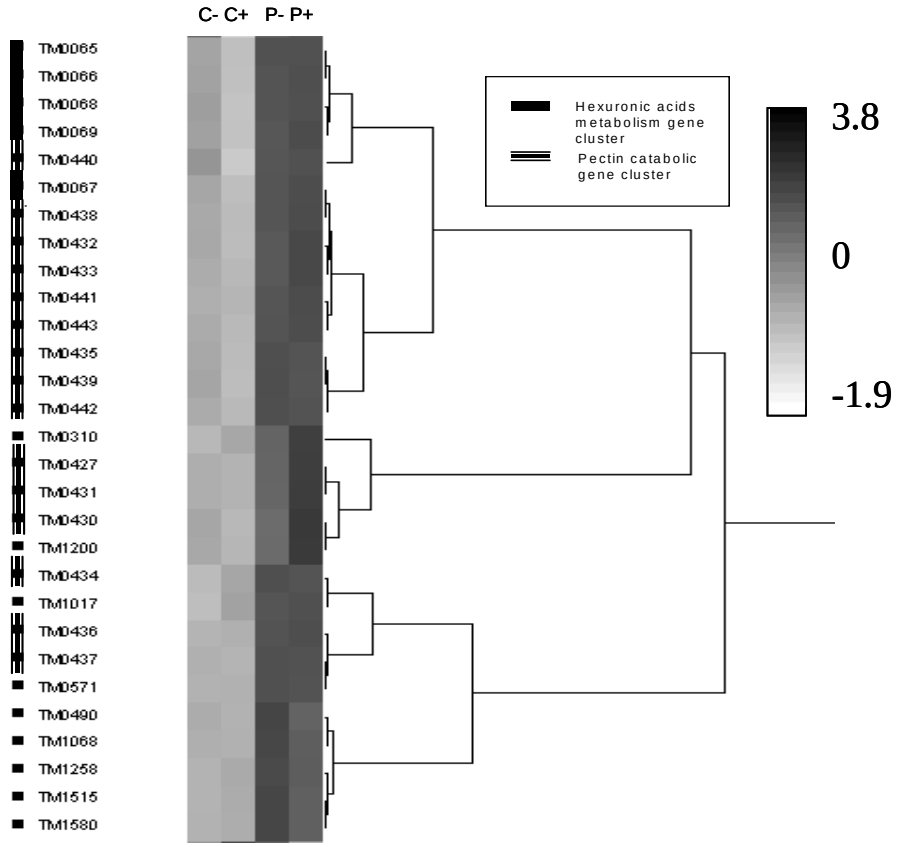
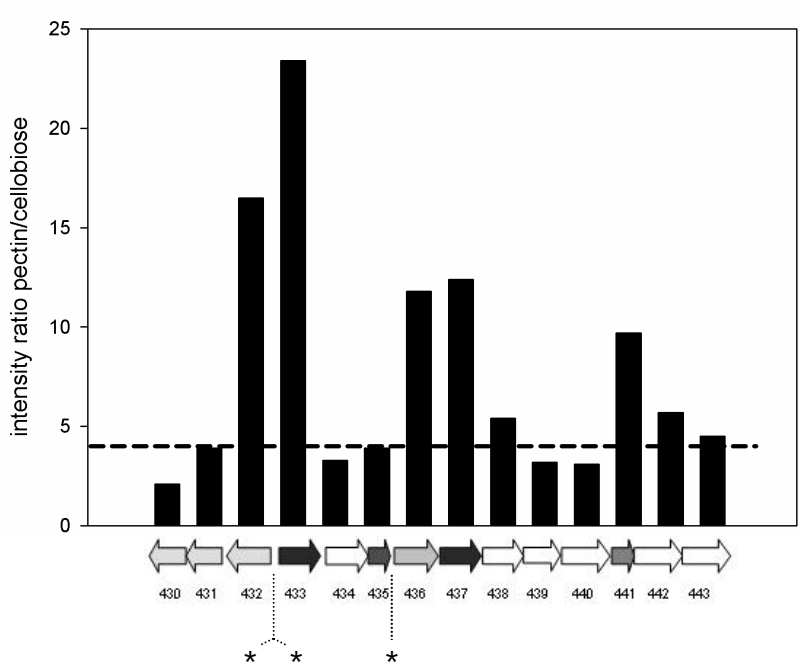


Figure 7.2 Clustering of genes with the highest relative expression using least squares mean estimates of treatment effects for growth on pectin and cellobiose, each with and without added tryptone.

Differential expression of genes involved in degradation of pectin

The exopectate lyase *pel*A (TM0433) and exopolygalacturonase *pel*B (TM0437) (15, 16) are located in a gene string with twelve additional ORFs (Figure 7.3A). Both *pel*A and *pel*B displayed significant upregulation during growth of *T. maritima* on pectin as compared to cellobiose as a carbon source, with changes of 23.5 and 12.4-fold, respectively. These genes were expressed at very high levels relative to scanner detection limits during growth on pectin; therefore, induction ratios presented here may underestimate the true differential expression of these genes.

Located upstream and in opposite transcriptional direction of *pel*A are genes that code for a sugar binding protein (TM0432) which is similar to one from *Thermoanaerobacter tengcongensis* (24% of identity in 393 aa), and two ABC-type permeases (TM0430 and TM0431) which resemble counterparts of the TogMNAB system utilized by *E. chrysanthemi* 3937 for the uptake of oligogalacturonides (sequence identity 40% in 247 aa, and 37% in 253 aa, respectively) (11). An ATP-binding subunit homologous to TogA is not encoded by the *T. maritima* gene string. The 16.5-fold upregulation of the sugar-binding protein and 3.9- and 2.1-fold changes for the two permeases, indeed suggest a role in the uptake of extracellular pectin-derived oligomers in *T. maritima*. As *T. maritima* does not appear to possess a glycoside-pentoside-hexuronides (GPH) TogT-type homolog (13), one or more ABC transporters are likely to play an important role in oligogalacturonide uptake, and co-regulation of genes encoding alternative ATP-binding proteins might allow



	(potential) KdgR box
* TM0433 (PelA)	TTTGAAGT**ATCAGA**CATCTCAG
* TM0432 (ABC transporter)	TTTGAAAC**ATCAGA**TATCAAAA
* TM0436 (Adh)	TCTGAAAC**ATCA**A-CATCTTAA
E. chrysanthemi	AATGAAACgcaGg-TTtATTT
E. carotovora	RWWGAAANNNNNN-TTTCAKKA
consensus	--tGAAacATCAg-taTCt---

Figure 7.3 (A) Schematic representation of the gene cluster containing *Pel*A and *Pel*B, juxtaposed to bar diagrams illustrating the expression ratios when grown on pectin as a carbon source. Dashed lines indicate a 4-fold upregulation (see also Table 7.1 for details on annotation and regulation levels). Stars indicate the upstream sequences of TM0432, TM0433, and TM0436, which are shown in detail in B. **(B)** Alignment of putative KdgR operators, as observed in TM0432, 0433, and 0436 (double underlined), compared to KdgR box consensus of *E. chrysanthemi* (23) and *E. carotovora* (underlined) (19).

completion of a functional Tog-like transport system. In *E. chrysanthemi*, TogB is believed to act as a chemoreceptor for oligogalacturonides, supposedly in concert with a methyl-accepting chemotaxis protein (11). A methyl-accepting chemotaxis protein (TM0429) located just downstream of the *T. maritima* Tog homologs was not differentially expressed here, which may reflect either independent regulation, non-specificity of this protein's activity, or cross-hybridization between several closely related paralogs known to exist in the genome (TM0023, TM0918).

The *agu* gene (TM0434) that is located directly downstream from *pel*A, has recently been demonstrated to encode an α-glucuronidase (39). Because of the clustering with pectin genes, we

performed an independent analysis that indeed revealed that this enzyme is also capable of degrading tetragalacturonic acid, an intermediate in the pectin degradation process (Kluskens and van der Oost, unpublished data). Unfortunately, the extremely high nucleotide identity of TM0434 with the paralogous TM1068 (99% in 1,392 bp) did not allow for detecting differences in regulation of these two genes; from the current experiment it can only be concluded that the expression of at least one paralog is induced during growth on pectin.

Many pectin-degrading organisms use extracellular methylesterases (PMEs) to remove the methyl ester group from pectin, thereby facilitating the action of pectate lyases. The *T. maritima* genome encodes nine putative esterases, including a protein with low sequence identity to the recently-characterized pectin acetylesterase *pae*X from *E. chrysanthemi* (TM1160, 23% identity in 194 aa) (37). The expression levels of all nine genes remained largely unaffected with the exception of TM0435, which was elevated 3.9 fold during growth on pectin. This gene, annotated as an acetyl xylan esterase, displays no similarity with known PMEs, which are all classified within family 8 of carbohydrate esterases (5). No signal sequence could be identified, which suggests a cytoplasmic localization for the TM0435 gene product. Recently, suppressive subtractive hybridization approaches revealed that despite highly identical rRNA sequence, the genome content of *T. maritima* MSB8 and *Thermotoga* sp. strain RQ2 may differ by at least 20% (27). BLASTX searches show that strain RQ2 contains a pectin methylesterase-like protein, highly homologous to one from *E. chrysanthemi* (27). It is unclear whether the absence of this gene in MSB8 is due to loss or that its presence in RQ2 is the result of gain. To our knowledge, demethylation of pectin oligomers in the cytoplasm, as is suggested here in the case of *T. maritima*, has never been described before. The cytoplasmic methyl removal would imply that *T. maritima* is capable of transporting methylated oligogalacturonate. An explanation for this unusual feature might be the capacity of *T. maritima* to catabolize methanol, as will be discussed below.

One of the largest changes observed in this study was the 11.8-fold induction of TM0436, encoding a putative Zn^{2+}-containing alcohol dehydrogenase (Adh) with 33% identity in 365 aa with NosE of *Nostoc* sp. GSV224 (20). Members of this broad superfamily convert several alcohols, ranging from hexanol to methanol, into aldehydes or ketones. The level of upregulation of this gene suggests that the enzyme is directly or indirectly involved in the process of pectin degradation. A plausible explanation could be that the enzyme plays a role in the detoxification or metabolism of methanol, a direct metabolite from pectin demethylation (Balk and Stams, unpublished data).

Several genes downstream of *pel*B (TM0437) that are highly similar to genes involved in gluconate catabolism, also show significantly elevated expression levels during growth on pectin (Table 7.1). TM0438 is similar to *E. coli* 6-phosphogluconate dehydrogenase (*gnd*) (48% identity in 475 aa). The putative regulatory protein TM0439 shares similarity with *Bacillus subtilis* GntR (31% identity in 209 aa), a repressor of the GntI gluconate utilization system (22), and with ExuR from *Yersinia pestis* (26% identity in 222 aa), a regulator that controls *uxa*ABC expression in galacturonate catabolism (41). TM0443 resembles the gluconate kinase of *B. subtilis* (GntK; 37% identity in 451 aa) (32). The unsaturated pectate cleavage product DKI (5-keto-4-deoxyuronate) is converted to DKII (2,5-keto-3-deoxygluconate) by a deoxyuronate isomerase (KduI) in *E.*

carotovora and *E. chrysanthemi* (12). In *T. maritima*, however, no ortholog of KduI could be detected by sequence similarity searches. It should be noted that also in some other pectinolytic organisms, e.g. *Klebsiella pneumoniae*, no *kdu*I-like gene has been detected in their genome (36). The DKI intermediate is generated in the periplasm by PelB, and has been reported to be rather unstable (43). Hence, under certain conditions DKI may be spontaneously converted to DKII; alternatively, it can not be ruled out that the isomerization of DKI to DKII is catalyzed by the galacturonate isomerase (TM0064, see below), which would resemble the promiscuity of metabolic enzymes that has recently been described by Lamble and coworkers (17). TM0441 is homologous to a keto-3-deoxygluconate oxidoreductase (*kdu*D) that catalyzed the conversion of DKII to 2-keto-3-deoxy-gluconate (KDG); its expression level is 9.8-fold elevated in the presence of pectin.

Two ORFs within the *pel*A/*pel*B cluster encode hypothetical proteins which are induced 3.1 (TM0440) and 5.7 fold (TM0442) (Table 7.1), suggesting some involvement in pectin catabolism. TM0440 is similar to a predicted protein of unknown function from *Oceanobacillus iheyensis* (41% identity in 477 aa). No TM0442 homolog was found in other organisms.

Differential expression of genes involved in sugar acids catabolism

The second cluster of pectin-related genes in the *T. maritima* genome are homologous to sugar acids catabolism genes (Figure 7.4) (29, 36). Like *B. subtilis*, *T. maritima* appears to lack an *uxa*A homolog and contains only one *uxa*B/*uxu*B homolog (26).

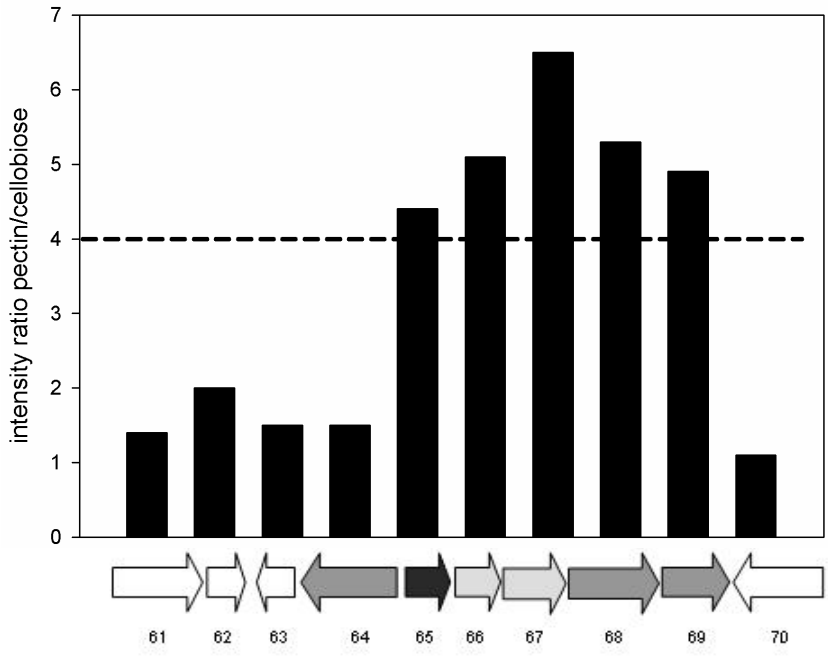


Figure 7.4 Schematic representation of the gene cluster containing *Kdg*R, juxtaposed to bar diagrams illustrating expression ratios when grown on pectin. Dashed lines indicate a 4-fold upregulation (see Table 7.1 for details on annotation and regulation levels).

The organization of these genes in *T. maritima* differs from other pectin-utilizing organisms in that *uxa*A and *uxa*B are located in the same gene string as *kdg*K, *kdg*A, and *kdg*R (TM0065) (47); however, levels of upregulation during growth on pectin suggests that these five genes are transcribed as an operon. Similarity between the *T. maritima kdg*R homolog and the pectin repressor KdgR of *E. chrysanthemi* (33% identity in 247 aa) also suggest that it may act as a local regulator of these genes (23, 24, 34, 40). TM0068 and TM0069, which share sequence identity with *uxa*B and *uxa*A, respectively, display 5.3-fold and 4.9-fold upregulation during growth on pectin. Homologs of these genes are known to be involved in isomerization of tagaturonate into altronate, and in dehydration of altronate into 2-keto-3-deoxy-gluconate (KDG), the central metabolite of pectin catabolism. Similar upregulation is displayed by the putative 2-keto-3-deoxy-phosphogluconate aldolase (*kdg*A, TM0066, 5.1 fold) and a 2-keto-3-deoxy-gluconate kinase (*kdg*K, TM0067, 6.4 fold). KDG, formed after the stepwise conversion of galacturonate, is phosphorylated into 6-phospho-KDG, followed by cleavage into glyceraldehyde-3-phosphate and pyruvate by the aldolase. A candidate gene for *edd*, coding for a (6-phospho)gluconate dehydratase which would bridge the gluconate and galacturonate metabolic pathways by formation of KDG in the Entner-Doudoroff pathway, might be TM0006. This is based on similarity with a recently established archaeal gluconate hydratase (Ettema and Van der Oost, unpublished data).

A gene encoding a homolog of the uronate isomerase UxaC (TM0064) is positioned upstream and in opposite transcriptional direction of TM0065 and shows a slight but significant upregulation during growth on pectin (1.5 fold). *E. chrysanthemi uxa*C mutants are incapable of growing on galacturonate or glucuronate (12). The hypothetical protein encoded by the gene downstream of *uxa*C also shows a 2-fold expression increase during growth on pectin, but its function remains to be identified.

Other strongly upregulated genes

Several genes that do not reside in the two main clusters discussed above showed significant changes as well. TM1017 is homologous to the C-terminus of *E. chrysanthemi* PecM (21% identity in 281 aa), a membrane-associated protein that forms a regulatory couple together with PecS in this organism. The latter protein is a repressor of the extracellular cellulase, pectinase and blue pigment production (35). PecM is required for the DNA-binding capacity of PecS in *E. chrysanthemi* (30). A search for homologs in the *T. maritima* genome resulted in two candidates sharing identity with PecS, TM0710 (28% in 281 aa) and TM0816 (26% in 106 aa), neither of which were strongly induced during growth on pectin. Strong upregulation of TM0816 has been observed upon heat shock (Pysz and Kelly, unpublished data).

TM1201 is predicted to encode a putative arabinogalactan endo-1,4-beta-galactosidase that is believed to be involved in the removal of arabinogalactan units linked to complex polysaccharides, such as pectin. However, this gene did not show changes in expression level between cellobiose and pectin. This might be related to possible differences between the composition of pectin used during this experiment, and pectin encountered in the natural habitats of *T. maritima*.

A connection between iron limitation and pectin catabolic genes has been previously noted in *E. chrysanthemi* 3937, where low iron availability was found not only to induce genes involved in iron transport, but also the pectate lyases *pel*D and *pel*E. The transcriptional repressor Fur (ferric uptake regulation) plays a key role in the regulation of these genes. Regulation of *pel*D and *pel*E expression by Fur is believed to couple virulence and iron uptake by preventing activation of these genes by Crp (7). Expression of one of the two Fur homologs in *T. maritima* (TM1515) was induced 2.0-fold and a second gene (TM1176), annotated as a metal-sensing transcriptional regulator, was upregulated 2.6 fold.

ORFs strongly upregulated by cellobiose

Similar to results observed previously during growth on carboxymethylcellulose (3), members of a predicted pathway for cellobiose utilization were induced during growth on cellobiose. These included TM1218, which encodes a transcriptional regulator belonging to the LacI family (2.7-fold), the ABC transporter components TM1219-TM1223, the genes encoding two well-characterized cellulases, CelA (TM1524, 9.4-fold) and CelB (TM1525, 8.1-fold), and cellobiose-phosphorylase (TM1848, *cepA*; 27-fold induction on cellobiose) (31). Cellobiose as a carbon source also resulted in an increase in expression level of a maltodextrin glycosyltransferase (TM0767, 4.3-fold) and a cyclomaltodextrinase (TM1835, 2.5-fold). These genes are linked to α-glucan metabolism, and a role in starch degradation has been demonstrated (3).

Regulation of pectinolytic genes in T. maritima

At least four different proteins are known to be involved in regulating pathways involved in hexuronide degradation in *E. coli* and other gram-negative bacteria (36). In general, ExuR and UxuR regulate genes involved in glucuronide, galacturonate, and glucuronate breakdown, while KdgR regulates the polygalacturonate breakdown pathway. Gluconate transport and metabolism is usually regulated by GntR. Additional regulators of related processes in *E. coli* include IdnR and UidR. Our microarray results, coupled with genomic organization, suggest possible involvement of *T. maritima* KdgR and GntR in regulation of pectin degradation genes. However, the involvement of other regulators can not be ruled out, since the array used for this study contained only a subset of regulators in the genome. Although the DNA binding specificity of the *T. maritima* KdgR protein has not yet been experimentally established, the crystal structure of this protein suggests that it binds DNA as a tetramer, perhaps interacting with a target sequence containing two 12-14 base palindromic operator sites (47). The upstream sequence of TM0065 contains two palindromic sequences separated by 7 bases beginning 86 bases upstream of the start codon, a configuration similar to that observed in the *E. coli* IclR promoter (14). The sequences show similarity to KdgR binding sites previously observed and predicted for purple bacteria (36). The regulatory regions and the upstream regions of all essential genes from both gene clusters (Figure 7.1 and 7.2) were examined for the presence of repeats, palindromic sequences and conserved motifs that potentially could operate as binding sites for regulators, such as KdgR (Figure 7.4). PelA and the sugar-binding protein of the ABC transporter (TM0433 and TM0432, respectively) are transcribed in opposite

direction, displaying an intergenic region of 246 nucleotides. The expression level of both genes is very similar. In search of consensus sequences that could serve as an indication for coordinated regulation, we found an inverted repeat with striking similarity, preceding both genes, starting 31 bases upstream of the start codon for TM0432 and 10 bases for TM0433, respectively (Figure 7.3B). The genes of the cluster TM0436-TM0443 are separated by intergenic regions of 22 nucleotides at the most. Therefore transcription is most likely initiating upstream the alcohol dehydrogenase, TM0436. The region preceding TM0436 contains a weak palindromic sequence, which has some similarity with the earlier mentioned palindromes found for TM0432 and TM0433 (Figure 7.3B). A divergent orientation of genes is also found in the cluster containing the genes for sugar acids catabolism (TM0064 and TM0065). As is the case for TM0432 and TM0433, both genes have an almost identical level of expression. Clear palindromic sequences that show consensus with palindromes from the first cluster however could not be identified, which could indicate the involvement of a different regulator.

Microarray studies have suggested that *T. maritima* regulates gene expression over a range of sugar substrates, and appears to repress genes involved in the breakdown of non-glucose sugars in the presence of glucose (3). Since the cAMP receptor protein (CRP) plays an active regulatory role in *E. chrysanthemi* (25, 33), it was analyzed whether this may be conserved in *T. maritima*. A CRP-like gene is present in the genome although a role in global catabolite repression has not been described in *T. maritima*. While the *T. maritima* genome encodes homologs of global catabolite repressors from both Gram-negative bacteria (TM1171; 30% identity in 181 aa to a CRP-like protein from *Bacteroides thetaiotaomicron*), and Gram-positive bacteria (*T. maritima* contains 5 CcpA homologs of *B. subtilis*, belonging to the same lacI regulator family; TM0299, TM0949, TM1200, TM1218, and TM1856, all around 31% identity in 330 aa), their actual physiological role is not clear. The mechanisms mediated by these proteins are expected to be substantially different in *T. maritima* due to the lack of a PTS system and adenylate cyclase homolog (26). Indeed, previous work has suggested that carbon catabolite repression in *Thermotoga neapolitana* appears to be independent of cAMP levels (42). *T. maritima* also has a glucokinase (TM1469) sharing 54% and 55% similarity respectively with the *S. coelicolor* and *B. megaterium* ROK family glucokinases, shown to be involved in PTS-independent catabolite repression (21).

Proposed pectin catabolism pathway

A schematic overview of pectin catabolism in *T. maritima* is given in Figure 7.5. The pectinolytic activity of the characterized PelA and PelB was supported by elevated induction levels in microarray experiments. Preliminary results indicate that a predicted third depolymerizing enzyme, Agu (TM0434), is also active on oligogalacturonate, although the presence of a close paralog prevented us from concluding which of the two was overexpressed. A homolog of the putative pectin methylesterase, found in *Thermotoga* strain RQ2, was not present on the MSB8 genome; however, the induction of expression of a predicted esterase (TM0435) and alcohol dehydrogenase (TM036) suggest that these gene products may be involved in the cytoplasmic demethylation of pectin. Following translocation of galacturonate oligomers into the cytoplasm, expression data

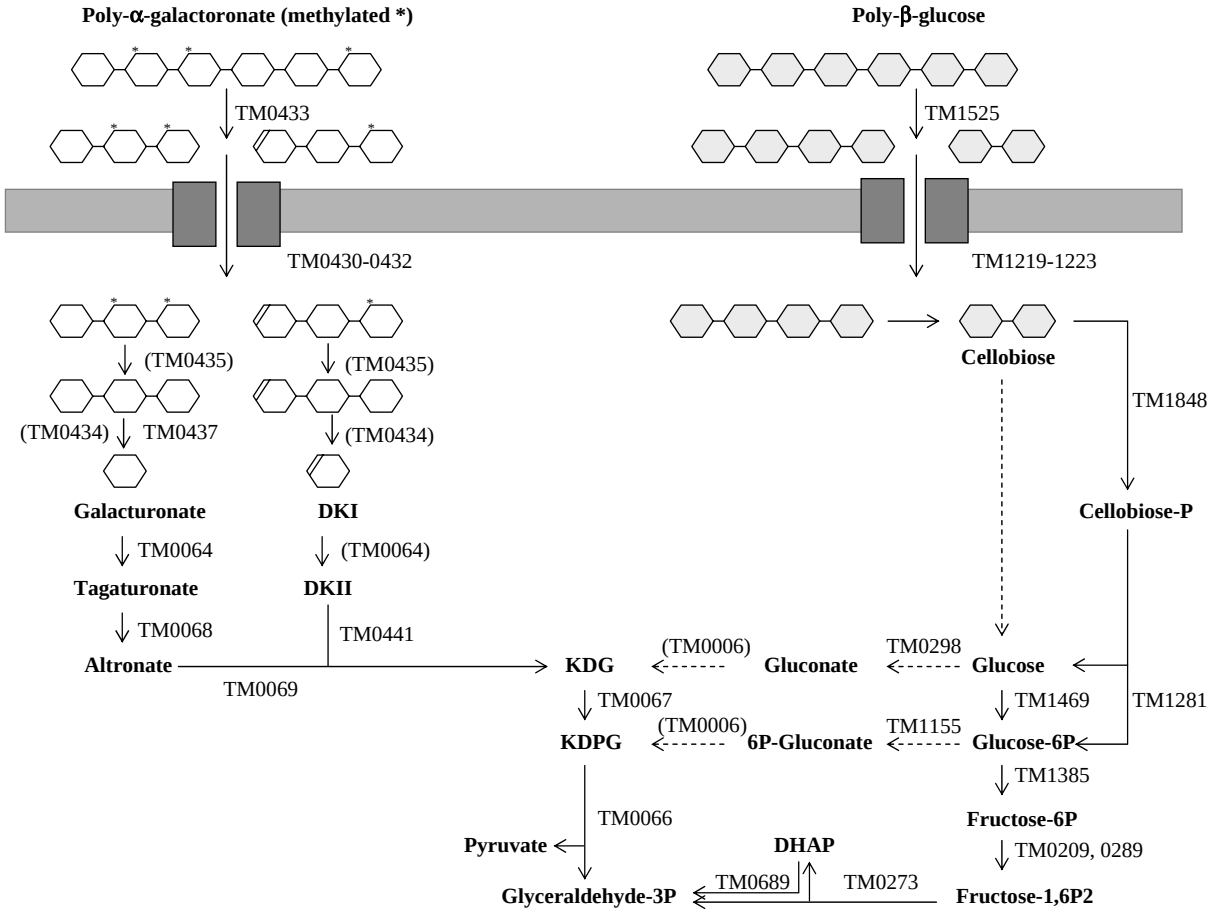


Figure 7.5 Proposed pectin catabolism in *T. maritima*. Methylation of galacturonate units is indicated by stars. Genes appointed to conversions based on homology searches are displayed between brackets. The corresponding gene function to the TM numbers is explained in Table 7.1, except for TM0006 (put. (6-phospho)gluconate kinase, *edd*), TM0273 (fructose-1,6-bisphosphate aldolase), TM0289 (phosphofructokinase,) (6), TM0298 (alcohol dehydrogenase, Zn^{2+}-containing), TM1155 (glucose-6-phosphate dehydrogenase), TM1281 (6-phospho-β-glucosidase), TM1385 (phosphoglucose isomerase), and TM1469 (hexokinase).

suggest separate pathways for the metabolism of saturated and unsaturated galacturonate. Galacturonate seems to be converted via tagaturonate and altronate into the intermediate KDG, while the unsaturated monomer 5-keto-4-deoxyuronate is converted to this intermediate via 2,5-keto-3-deoxygluconate by a yet unidentified isomerase. The action of promiscuous enzymes in this conversion step, as for example the uronate isomerase TM0064, can not be ruled out. An identified KDG kinase and KDG aldolase complete the pathway, thereby providing pyruvate and glyceraldehyde-3-phosphate for further conversion via central metabolic pathways. KdgR, a regulator protein operating as a repressor, influences the expression of most pectin catabolic genes in other species, and this role may be conserved in *T. maritima*. It is unclear whether a *PecS-PecM* regulatory couple similar to that of *E. chrysanthemi* operates in *T. maritima*, although proteins with sequence similarity to these genes have been identified in the genome. Details of catabolite repression, including regulators involved in this organism, require further research.

Conclusions

By means of a targeted cDNA microarray we have verified the induction of a number of *T. maritima* genes encoding enzymes apparently required for the successive conversion of the pectin polymer to central metabolites. This organism may follow a more rudimentary metabolic route than pectin-degraders such as *E. chrysanthemi* and *A. niger*, for which a broader spectrum of pectinolytic enzymes and regulators has been identified. The metabolic pathway of this sugar polymer appears to have only two missing links: a pectin methylesterase and the KduI isomerase. No orthologs, but enzymes with similar catalytic capacities that are encoded on the two pectinase-related clusters (TM0435 and TM0064) might be responsible for the respective activities. Our results confirm the value of targeted microarray studies for unraveling and verifying metabolic pathways.

Acknowledgments

The authors would like to thank R. Wolfinger and K. Scott, SAS Institute, Cary, NC, for help with implementing the mixed model analysis.

References

1. **Bronnenmeier, K., A. Kern, W. Liebl, and W. L. Staudenbauer.** 1995. Purification of *Thermotoga maritima* enzymes for the degradation of cellulosic materials. Appl. Environ. Microbiol. **61:**1399-1407.
2. **Chhabra, S., K. N. Parker, D. Lam, W. Callen, A. Snead, E. J. Mathur, J. M. Short, and R. M. Kelly.** 2001. β-mannanases from *Thermotoga* species. Methods Enzymol. **330:**224-238.
3. **Chhabra, S. R., K. R. Shockley, S. B. Conners, K. E. Scott, R. D. Wolfinger, and R. M. Kelly.** 2003. Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima.* J. Biol. Chem. **278:**7540-7552.
4. **Chhabra, S. R., K. R. Shockley, D. E. Ward, and R. M. Kelly.** 2002. Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. Appl. Environ. Microbiol. **68:**545-554.
5. **Coutinho, P. M., and B. Henrissat.** 1999. Carbohydrate-active enzymes server. http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html.
6. **Ding, Y.-H. R., R. S. Ronimus, and H. W. Morgan.** 2001. *Thermotoga maritima* phosphofructokinases: expression and characterization of two unique enzymes. J. Bacteriol. **183:**791-794.
7. **Franza, T., I. Michaud-Soret, P. Piquerel, and D. Expert.** 2002. Coupling of iron assimilation and pectinolysis in *Erwinia chrysanthemi* 3937. Mol. Plant Microbe Interact. **15:**1181-1191.
8. **Griffiths, J. S., N. J. Wymer, E. Njolito, S. Niranjanakumari, C. A. Fierke, and E. J. Toone.** 2002. Cloning, isolation and characterization of the *Thermotoga maritima* KDPG aldolase. Bioorg. Med. Chem. **10:**545-550.
9. **Hegde, P., R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J. E. Hughes, E. Snesrud, N. Lee, and J. Quackenbush.** 2000. A concise guide to cDNA microarray analysis. Biotechniques. **29:**548-562.

10. **Huber, R., T. A. Langworthy, H. König, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. Arch. Microbiol. **144:**324-333.
11. **Hugouvieux-Cotte-Pattat, N., N. Blot, and S. Reverchon.** 2001. Identification of TogMNAB, an ABC transporter which mediates the uptake of pectic oligomers in *Erwinia chrysanthemi* 3937. Mol. Microbiol. **41:**1113-1123.
12. **Hugouvieux-Cotte-Pattat, N., G. Condemine, W. Nasser, and S. Reverchon.** 1996. Regulation of pectinolysis in *Erwinia chrysanthemi.* Annu. Rev. Microbiol. **50:**213-257.
13. **Hugouvieux-Cotte-Pattat, N., and S. Reverchon.** 2001. Two transporters, TogT and TogMNAB, are responsible for oligogalacturonide uptake in *Erwinia chrysanthemi* 3937. Mol. Microbiol. **41:**1125-1132.
14. **Ibañez, E., E. Campos, L. Baldoma, J. Aguilar, and J. Badia.** 2000. Regulation of expression of the *yia*KLMNOPQRS operon for carbohydrate utilization in *Escherichia coli*: Involvement of the main transcriptional factors. J. Bacteriol. **182:**4617-4624.
15. **Kluskens, L. D., G.-J. W. M. van Alebeek, A. G. J. Voragen, W. M. de Vos, and J. van der Oost.** 2003. Molecular and biochemical characterisation of the thermo-active family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima.* Biochem. J. **370:**651-659.
16. **Kluskens, L. D., G.-J. W. M. van Alebeek, J. Walther, A. G. J. Voragen, W. M. de Vos, and J. van der Oost.** 2003. *Pel*B from the hyperthermophilic bacterium *Thermotoga maritima* encodes a fungal-like intracellular exopolygalacturonase. submitted.
17. **Lamble, H. J., N. I. Heyer, S. D. Bull, D. W. Hough, and M. J. Danson.** 2003. Metabolic pathway promiscuity in the Archaeon *Sulfolobus solfataricus* revealed by studies on glucose dehydrogenase and aldolase. J. Biol. Chem. **submitted.**
18. **Liebl, W., P. Ruile, K. Bronnenmeier, K. Riedel, F. Lottspeich, and I. Greif.** 1996. Analysis of a *Thermotoga maritima* DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. Microbiology. **142:**2533-2542.
19. **Liu, Y., G. Jiang, Y. Cui, A. Mukherjee, W. L. Ma, and A. K. Chatterjee.** 1999. $kdgR_{Ecc}$ negatively regulates genes for pectinases, cellulase, protease, $harpin_{Ecc}$, and a global RNA regulator in *Erwinia carotovora* subsp. *carotovora.* J. Bacteriol. **181:**2411-2422.
20. **Luesch, H., D. Hoffmann, J. M. Hevel, J. E. Becker, T. Golakoti, and R. E. Moore.** 2003. Biosynthesis of 4-methylproline in cyanobacteria: Cloning of nosE and nosF genes and biochemical characterization of the encoded dehydrogenase and reductase activities. J. Org. Chem. **68:**83-91.
21. **Mahr, K., G. P. van Wezel, C. Svensson, U. Krengel, M. J. Bibb, and F. Titgemeyer.** 2000. Glucose kinase of *Streptomyces coelicolor* A3(2): Large-scale purification and biochemical analysis. Antonie van Leeuwenhoek. **78:**253-261.
22. **Miwa, Y., and Y. Fujita.** 1988. Purification and characterization of a repressor for the *Bacillus subtilis gnt* operon. J. Biol. Chem. **263:**13252-13257.
23. **Nasser, W., S. Reverchon, G. Condemine, and J. Robert-Baudouy.** 1994. Specific interactions of *Erwinia chrysanthemi* KdgR repressor with different operators of genes involved in pectinolysis. J. Mol. Biol. **236:**427-440.
24. **Nasser, W., S. Reverchon, and J. Robert-Baudouy.** 1992. Purification and functional characterization of the KdgR protein, a major repressor of pectinolysis genes of *Erwinia chrysanthemi.* Mol. Microbiol. **6:**257-265.

25. **Nasser, W., J. Robert-Baudouy, and S. Reverchon.** 1997. Antagonistic effect of CRP and KdgR in the transcription control of the *Erwinia chrysanthemi* pectinolysis genes. Mol. Microbiol. **26:**1071-1082.
26. **Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, L. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, O. White, S. L. Salzberg, H. O. Smith, J. C. Venter, and C. M. Fraser.** 1999. Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima.* Nature. **399:**323-329.
27. **Nesbø, C. L., K. E. Nelson, and W. F. Doolittle.** 2002. Suppressive subtractive hybridization detects extensive genomic diversity in *Thermotoga maritima.* J. Bacteriol. **184:**4475-4488.
28. **Parisot, J., A. Ghochikyan, V. Langlois, V. Sakanyan, and C. Rabiller.** 2002. Exopolygalacturonate lyase from *Thermotoga maritima*: cloning, characterization and organic synthesis application. Carbohydrate Res. **337:**1427-1433.
29. **Peekhaus, N., and T. Conway.** 1998. What's for dinner? Entner-Doudoroff metabolism in *Escherichia coli.* J. Bacteriol. **180:**3495-3502.
30. **Praillet, T., S. Reverchon, J. Robert-Baudouy, and W. Nasser.** 1997. The PecM protein is necessary for the DNA-binding capacity of the PecS repressor, one of the regulators of virulence-factor synthesis in *Erwinia chrysanthemi.* FEMS Microbiol. Lett. **154:**265-270.
31. **Rajashekhara, E., M. Kitaoka, Y.-K. Kim, and K. Hayashi.** 2002. Characterization of a cellobiose phosphorylase from a hyperthermophilic eubacterium, *Thermotoga maritima* MSB8. Biosci. Biotechnol. Biochem. **66:**2578-2586.
32. **Reizer, A., J. Deutscher, M. H. Saier Jr, and J. Reizer.** 1991. Analysis of the gluconate (gnt) operon of *Bacillus subtilis*. Mol. Microbiol. **5:**1081-1089.
33. **Reverchon, S., D. Expert, J. Robert-Baudouy, and W. Nasser.** 1997. The cyclic AMP receptor protein is the main activator of pectinolysis genes in *Erwinia chrysanthemi.* J. Bacteriol. **179:**3500-3508.
34. **Reverchon, S., W. Nasser, and J. Robert-Baudouy.** 1991. Characterization of *kdgR*, a gene of *Erwinia chrysanthemi* that regulates pectin degradation. Mol. Microbiol. **5:**2203-2216.
35. **Reverchon, S., W. Nasser, and J. Robert-Baudouy.** 1994. *pecS*: a locus controlling pectinase, cellulase and blue pigment production in *Erwinia chrysanthemi.* Mol. Microbiol. **11:**1127-1139.
36. **Rodionov, D. A., A. A. Mironov, A. B. Rakhmaninova, and M. S. Gelfand.** 2000. Transcriptional regulation of transport and utilization systems for hexuronides, hexuronates and hexonates in gamma purple bacteria. Mol. Microbiol. **38:**673-683.
37. **Shevchik, V. E., and N. Hugouvieux-Cotte-Pattat.** 2003. PaeX, a second pectin acetylesterase of *Erwinia chrysanthemi* 3937. J. Bacteriol. **185:**3091-3100.
38. **Shockley, K. R., D. E. Ward, S. R. Chhabra, S. B. Conners, C. I. Montero, and R. M. Kelly.** 2003. Heat shock response by the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl. Environ. Microbiol. **69:**2365-2371.
39. **Suresh, C., A. A. Rus'd, M. Kitaoka, and K. Hayashi.** 2002. Evidence that the putative α-glucosidase of *Thermotoga maritima* MSB8 is a pNP α-D-glucuronopyranoside hydrolyzing α-glucuronidase. FEBS Letters. **517:**159-162.

40. **Thomson, N. R., W. Nasser, S. McGowan, M. Sebaihia, and G. P. C. Salmond.** 1999. *Erwinia carotovora* has two KdgR-like proteins belonging to the IclR family of transcriptional regulators: identification and characterization of the RexZ activator and the KdgR repressor of pathogenesis. Microbiology. **145:**1531-1545.
41. **Valmeekam, V., Y. L. Loh, and M. J. San Francisco.** 2001. Control of exuT activity for galacturonate transport by the negative regulator ExuR in *Erwinia chrysanthemi.* Mol. Plant Microbe Interact. **14:**816-820.
42. **Vargas, M., and K. M. Noll.** 1996. Catabolite repression in the hyperthermophilic bacterium *Thermotoga neapolitana* is independent of cAMP. Microbiology. **142:**139-144.
43. **Voragen, A. G. J.** 1972. Characterization of pectin lyases on pectins and methyl oligogalacturonates. Ph.D. Thesis, Wageningen, Landbouwhogeschool Wageningen, The Netherlands Ph.D. Landbouwhogeschool Wageningen.
44. **Ward, D. E., S. W. M. Kengen, J. van der Oost, and W. M. de Vos.** 2000. Purification and characterization of the alanine aminotransferase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its role in alanine production. J. Bacteriol. **182:**2559-2566.
45. **Winterhalter, C., and W. Liebl.** 1995. Two extremely thermostable xylanases of the hyperthermophilic bacterium *Thermotoga maritima* MSB8. Appl. Environ. Microbiol. **61:**1810-1815.
46. **Wolfinger, R. D., G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules.** 2001. Assessing gene significance from cDNA microarray expression data via mixed models. J. Comput. Biol. **8:**625-637.
47. **Zhang, R., Y. Kim, T. Skarina, S. Beasley, R. Laskowski, C. Arrowsmith, A. Edwards, A. Joachimiak, and A. Savchenko.** 2002. Crystal structure of *Thermotoga maritima* 0065, a member of the IclR transcriptional factor family. J. Biol. Chem. **277:**19183-19190.

8

Summary and concluding remarks

Thermophiles and hyperthermophiles are microorganisms that have the ability to grow at high temperatures. The majority belongs to a recently discovered domain, the archaea, and their position in the phylogenetic tree of life has put forward the suggestion that they stood at the basis of life as we now know it. Growth at high temperatures involves thermostable enzymes, and the question how this stabilization takes place at a structural level has played a central role in research of these organisms. In addition, because of this thermostability they gained interest in applications as biocatalysts in industry.

The bacterial domain holds a few members that can be considered thermophilic, since they grow optimally above 55°C. All belong to either the *Aquificales* or *Thermotogales* order. The organisms that were used in this study, *Fervidobacterium pennivorans*, *F. gondwanense* and *Thermotoga maritima*, belong to the latter. They grow optimally at temperatures from 65 to 80°C and are able to use a large variety of carbohydrates and peptides, for which they possess a large range of enzymes, mostly hydrolytic. The research describes the detailed molecular and biochemical characterization of enzymes that enable growth of these organisms on, occasionally, remarkable substrates.

Following a general introduction on this thesis in **Chapter 1**, the heterologous production of the keratinolytic protease (Fls) from *F. pennivorans* in *Escherichia coli* and its molecular characterization is described in **Chapter 2**. With degenerate primers designed against the conserved catalytic regions of known serine proteases a 2.1-kb gene could be isolated, coding for a 699-amino acid pre-pro-protease. The resulting protease, a subtilisin-like serine protease, was not found to be active on either of the substrates tested, including keratin of chicken feathers. This may be the result of the tight binding of the pro-peptide in the substrate binding site, where it could function as an inhibitor. Although inactive, analysis by SDS-PAGE showed that processing of the 14-kDa propeptide had occurred and this autoproteolytic step could be eliminated when the catalytic residue His was inactivated. Although the *E. coli*-produced Fls had a subunit size of around 60 kDa, the biochemically purified keratinase was 90 kDa. Nonetheless, N-terminal sequence analysis of both the recombinant protease and wild-type keratinase, fragmented by thermolysin, revealed complete identity. The weight difference is most probably due to association of the keratinase with the outer cell envelope (toga), also since no signs of N-glycosylation were detected. In addition, a *fls* homolog of identical length was identified in *Fervidobacterium islandicum* recently (1).

The 1.7 Å resolution crystallization structure of the Fls active-site mutant, described in **Chapter 3**, helped us to explain why the cleaved propeptide domain (PD) does not dissociate from the remaining part of the protease and how the maturation can be accomplished. Tight binding of PD to the remaining part of the protease is mediated by hydrogen bonds along the domain surfaces and around the active cleft, and by the clamps to two sandwich domains, SD1 and SD2. Most likely, the successive elimination of the PD in Fls is accomplished by several proteolytic cleavages that may not be autolytic, as has been shown for many other subtilases, such as Subtlisin E. Distinctive from other subtilases, fervidolysin has two sandwich domains, SD1 and SD2. They show structural similarities to the fibronectin or the immunoglobulin (Ig) fold of the human promatrix metalloprotease-2 (proMMP-2) that degrades the fibrous polymeric substrate gelatin. It is known that this motif is involved in cell adhesion or in protein-protein interactions. Therefore, it was

proposed that the role of the SDs is to mediate the interaction with β-keratin, the protein structure in chicken feathers, and to assist in disassembly of the layers of β-structures so that individual strands may bind to the mature protease.

In **Chapter 4** the isolation and characterization of a xylose isomerase (XylA) from *F. gondwanense* is described. Via hybridization using the *xylA* gene from the phylogenetically cognate *T. maritima* as a probe, the *F. gondwanense xylA* gene was identified from a DNA mini-library. It was located between a putative xylosidase and a sugar kinase, all in the same transcriptional direction. Thermoactivity and –stability was found to be highly dependent on the presence of divalent cations, like Co^{2+} and Mg^{2+}, which were shown to increase the enzyme's melting temperature by 10 degrees. With an optimal temperature of 70°C the *E. coli*-produced *F. gondwanense* XylA displayed a lower optimum than its homologous counterparts from *T. maritima* and *T. neapolitana*. Nevertheless, with an optimum at a temperature level that avoids the unwanted Maillard side reactions, it could be of industrial value.

Besides many other carbohydrate polymers, *T. maritima* was found to grow on pectin. This was line with the annotation of two pectinases in the genome sequence of the organism. Chapter 5 and 6 describe the detailed characterization of these two pectinolytic enzymes. **Chapter 5** focuses on the first heterologously produced thermostable pectate lyase, PelA. The enzyme was functionally produced in *E. coli* and purified to homogeneity. Activity of pectate lyase was detected in the medium fraction of a pectin-grown culture of *T. maritima*, which showed that the production of PelA is induced by pectin. In addition, zymogram assays showed that both enzymes, recombinant and native, are the same. PelA was highly active on non-methylated polygalacturonate, releasing unsaturated trigalacturonate via β-elimination, following multiple chain attack. Based on the activity rates the substrate binding cleft most probably consists of five subsites, three of which interact with the reducing end of the substrate. PelA was found to be highly thermoactive, with an optimal reaction temperature at 90°C, and needed Ca^{2+} for its activity and stability. The measured half-life for thermal inactivation is almost 2 h at 95°C. Based on gel filtration, PelA was considered to be a tetramer, which is proposed to be a potential thermostabilizing factor. Its biochemical features makes it a excellent candidate to be used the process of pectin removal in the paper and pulp industry.

The characteristics of a second pectinase, the exopolygalacturonase PelB, are described in **Chapter 6**. In contrast to PelA, this pectinolytic enzyme was characterized as an intracellular, hydrolyzing enzyme, belonging to family 28 of the glycoside hydrolases. When examined on oligogalacturonate as a substrate, the enzyme was found to remove monogalacturonic acid units from the non-reducing end, instead of digalacturonate, which is usually observed for bacterial exogalacturonases. Kinetics on substrate with an increasing degree of polymerization showed that the highest activity was observed when using pentaoligomers. These data were used to calculate the subsite affinity of PelB, and with structure information of related endogalacturonases a substrate-enzyme binding model was proposed. The model contained 4 subsites, in which the highest affinity was measured for subsite +1. PelB was found to be the most thermoactive and –stable hydrolytic

pectinase known to date. The fruit-juice industry, with its clarification and color extraction steps that are often carried out at elevated temperatures, might be a good market for this enzyme.

The ability of *T. maritima* to grow on pectin, supported by two pectinolytic enzymes, PelA and PelB, encouraged us to examine pectin catabolism in more detail. A targeted microarray was used to examine differential expression of *pelA* and *pelB* and other selected *T. maritima* genes during growth on pectin and cellobiose. The results of this research are described in **Chapter 7**. Genes showing an elevated expression level when switched to pectin were found to be located in generally two clusters. One cluster codes for homologs of enzymes involved in pectin depolymerization, uptake and modification, while the other cluster appears to encode enzymes responsible for cytoplasmic catabolism of saturated and unsaturated galacturonate. Indeed, elevated levels of expression were found for *pelA* and *pelB*, together with a third gene, believed to encode a putative galacturonase. Other genes in this cluster showing higher expression encoded an esterase and an alcohol dehydrogenase, for which a putative role in the cytoplasmic release and conversion of methyl groups from pectin oligosaccharides was proposed. In the second cluster, a KDG kinase and a KDPG aldolase showed higher induction levels on pectin and most likely link galacturonate catabolism with glycolysis. The KdgR homolog positioned upstream of this cluster may play a role in the regulation of the expression of at least some or all pectinolytic genes in *T. maritima*. The microarray results combined with biochemical data lead to a proposed pathway for pectin catabolism in *T. maritima*, which appears to be following a more rudimentary route than renowned pectin-degraders like *Erwinia chrysanthemi*.

The knowledge on thermostable enzymes as biocatalysts has increased enormously over the last years, primarily due to the growing amount of thermozymes overproduced in mesophilic hosts, and their subsequent purification and characterization. The study on the thermozymes, as described in this thesis, is an example of an integrated approach in which bioinformatics, molecular genetics, structural analysis, DNA microarray analysis, biochemistry and fermentation has been combined. In the post-genomic era, the characterization of thermozymes is anticipated to grow exponentially, and with that the diversity in application of biocatalysts. Recent developments like high-throughput random mutagenesis, including laboratory evolution, and computational design have opened up doors towards the engineering of so-called tailor-made enzymes (2). The commercialization of these protein engineering technologies is advancing rapidly, and it is expected to significantly affect the application of enzymes, also from thermophiles, in a wide range of industrial processes.

1. Gödde, C., K. Sahm, L. Kluskens, W.M. de Vos, and G. Antranikian. Cloning and expression of a serine protease encoding gene from *Fervidobacterium islandicum*. Poster presentation VAAM, Göttingen 2002.
2. Huisman, G., and S.G. Sligar. 2003. Evolution of protein technologies from voodoo to science. *Curr. Opin. Biotechnol.* **14**: 357-359

9

Samenvatting en conclusies

Thermofielen zijn micro-organismen die bij hoge temperaturen kunnen groeien. Het merendeel behoort tot een recentelijk ontdekt domein, de archaea, en hun diepe positie in de fylogenetische boom suggereert dat deze micro-organismen aan de basis hebben gestaan van het leven op aarde. Leven bij hoge temperaturen, tot soms 100°C en zelfs hoger, vereist de werking van enzymen met een activiteit en stabiliteit bij dezelfde temperatuur. De vraag hoe deze enzymen, die in veel gevallen op het oog weinig verschillen van dezelfde enzymen uit mesofielen, hun stabiliteit bewaren, speelt een centrale rol in het onderzoek naar deze organismen. Hun thermoactiviteit en stabiliteit maakt hen bovendien tot interessante biocatalysatoren voor de industrie.

Naast de archaea, bevat ook het domein der bacteriën enkele micro-organismen die als thermofiel (groei vanaf 55°C) of zelfs hyperthermofiel (groei vanaf 80°C) beschouwd kunnen worden. Ze behoren voornamelijk tot twee orders, de *Aquificales* en de *Thermotogales*. De bacteriën die in deze studie aan bod zijn gekomen, te weten *Fervidobacterium pennivorans*, *F. gondwanense*, en *Thermotoga maritima*, behoren tot de laatste order. Zij groeien optimaal bij een temperatuur rond de 65°C, in het geval van *T. maritima* zelfs 80°C, en gebruiken daarvoor een grote verscheidenheid aan suikers en eiwitten. Deze koolstof- en energiebronnen worden door hen afgebroken dankzij een scala van enzymen, waarvan de meesten hydrolytisch van aard zijn. Dit onderzoek beschrijft de gedetailleerde moleculaire en biochemische karakterisatie van enkele van deze enzymen, welke deze thermofiele micro-organismen in staat stellen om op soms opmerkelijke substraten te groeien.

Volgend op de algemene introductie in **Hoofdstuk 1**, wordt in **Hoofdstuk 2** de heterologe productie beschreven van een keratinase uit *F. pennivorans*, fervidolysine. Dankzij gedegenereerde primers, gebaseerd op geconserveerde regio's rond catalytische residuen van bekende serine proteases, kon een gen van 2100 baseparen geïsoleerd worden. Dit gen codeerde voor een serine protease met een totale lengte van 699 aminozuren, bestaand uit een signaal peptide, een propeptide en een catalytisch gedeelte. Dit protease, behorende tot de subtilase subgroep, bleek na productie in *Escherichia coli* inactief te zijn op alle geteste eiwitsubstraten, inclusief keratine en kippenveren. Zijn inactiviteit zou te wijten kunnen zijn aan een hechte binding van het propeptide in de substraat-bindingsplek, alwaar het een inhiberende rol zou kunnen spelen. Ondanks zijn inactiviteit toonde SDS-PAGE analyse van het protease dat het 14 kDa-grote propeptide afgesplitst wordt van het catalytisch gedeelte. Een mutant, waarin het catalytisch residue histidine was uitgeschakeld, vertoonde deze (auto)proteolytische eigenschap niet. Het recombinante fervidolysine van ongeveer 60 kDa week in massa significant af van de 90 kDa, die werd gevonden voor het biochemisch gezuiverde keratinase. Desalniettemin toonde N-terminale sequentieanalyses aan dat beide eiwitten identiek waren. Het massaverschil wordt zeer waarschijnlijk veroorzaakt door de sterke binding van het keratinase met delen van het buitenmembraan (toga) van het organisme. *N*-glycosylering van het eiwit niet kon worden aangetoond. Bovendien is fervidolysine in aminozuur lengte (699) identiek aan de recentelijk geïsoleerde homoloog islandisine uit *F. islandicum* (1).

Hoofdstuk 3 beschrijft de kristalstructuur van de bovengenoemde mutant van fervidolysine. De structuur, met een resolutie van 1.7 Å, helpt ons te begrijpen waarom het propeptide domein (PD) niet dissocieert van het overige gedeelte van het protease, en hoe maturatie van het protease kan worden bewerkstelligd. De hechte binding van PD wordt veroorzaakt door waterstofbruggen aan de

domeinoppervlakken en rond de holte waar de actieve residuen zich bevinden, alsmede door interacties via waterstofbruggen met de twee sandwichdomeinen, SD1 en SD2. Vermoedelijk wordt de uiteindelijke verwijdering van het PD voor fervidolysine bereikt dankzij meerdere proteolytische digesties die niet noodzakelijkerwijs autolytisch hoeven te zijn. Een dergelijk proces is eerder aangetoond voor ander subtilases, zoals Subtilisin E. Een duidelijk verschil met andere subtilases is echter de aanwezigheid van de twee sandwichdomeinen SD1 en SD2 in fervidolysine. Op structureel niveau tonen ze overeenkomsten met fibronectine en de immunoglobuline (Ig) vouwing van het humane promatrix metalloprotease-2 (proMMP-2), wat het polymere substraat gelatine afbreekt. Van dit motief is bekend dat het betrokken is bij cel-adhesie en eiwit-eiwit interacties. Daarom wordt verondersteld dat de SDs betrokken zijn bij het sturen van de interactie met β-keratine, de eiwitstructuur van kippenveren, en dat ze een rol spelen bij de ontrafeling van de β-structuurlagen, zodat de individuele lagen toegankelijk worden voor het catalytisch domein van fervidolysine.

In **Hoofdstuk 4** wordt de isolatie en karakterisatie van een xylose isomerase (XylA) uit *F. gondwanense* beschreven, wat in staat is om glucose om te zetten tot het zoetere fructose. Via hybridisaties op een mini-DNA bank, gebruikmakend van het *xylA* gen uit *T. maritima* als probe, werd het *xylA* gen van *F. gondwanense* geïdentificeerd. Het gen ligt tussen een vermoedelijk xylosidase (stroomopwaarts) en een suiker kinase (stroomafwaarts), allen in dezelfde transcriptierichting. De thermoactiviteit en –stabiliteit van het *E. coli*-geproduceerde enzym bleek zeer afhankelijk van divalente metaalionen, zoals cobalt en magnesium. Hun aanwezigheid verhoogde de smelttemperatuur van het enzym met ruim 10 graden. De optimale temperatuur van het enzym is met zijn 70°C iets lager dan de homologen uit *T. maritima* en *T. neapolitana*. Aangezien het isomerase optimaal functioneert bij temperaturen waarbij ongewenste nevenreacties als gevolg van extreem hoge temperatuur worden vermeden, kan dit enzym van industriële betekenis zijn.

T. maritima heeft de capaciteit om op vele verschillende suikerpolymeren te groeien, inclusief pectine. Voor dit laatste substraat bevat het genoom van *T. maritima* twee genen die coderen voor potentiële pectinolytische enzymen. Hoofdstuk 5 en 6 behandelt de karakterisatie van beiden. **Hoofdstuk 5** beschrijft de eerste heterologe expressie van een thermostabiel pectaat lyase, PelA. Het enzym werd functioneel geproduceerd in *E. coli* en gezuiverd. Pectaat lyase-activiteit werd gevonden in de medium fractie van *T. maritima*, gekweekt op pectine en hiermee werd aangetoond dat de productie van PelA geïnduceerd wordt door pectine. Bovendien toonde zymogrammen aan dat het natieve en het *E. coli*-geproduceerde PelA een en hetzelfde enzym zijn. Zeer hoge PelA activiteit werd gemeten op niet-gemethyleerd polygalacturonzuur, wat resulteerde in de afsplitsing van onverzadigd trigalacturonzuur via β-eliminatie. Hierbij volgt het enzym een zg. multiple-chain-attack. Gebaseerd op de initiële reactiesnelheden bij verschillende oligogalacturonzuren werd geconcludeerd dat de substraat-bindingsplek hoogstwaarschijnlijk bestaat uit 5 subsites, waarvan er 3 een interactie ondergaan met het reducerende uiteinde van het substraat. PelA was uiterst thermoactief, met een optimale omzettingstemperatuur van 90°C, en bleek voor zijn activiteit en stabiliteit afhankelijk te zijn van Ca^{2+}. De halfwaardetijd voor inactivatie bij 95°C werd gesteld op

bijna 2 uur. Uit gelfiltratie-experimenten bleek dat PelA zeer waarschijnlijk een tetrameer is. Zijn thermodynamische kenmerken maken PelA tot een zeer interessant enzym voor de papier en pulpindustrie.

De kenmerken van het tweede pectinase, het exopolygalacturonase PelB, staan beschreven in **Hoofdstuk 6**. In tegenstelling tot PelA is dit enzym cytoplasmatisch en is het hydrolytisch van aard. Het behoort tot familie 28 van de glycoside hydrolases. PelA is in staat monogalacturonzuur aan het niet-reducerende uiteinde van oligogalacturonzuren te verwijderen. Dit wijkt af van het gebruikelijke mechanisme van bacteriële exogalacturonases, welke in de meeste andere gevallen digalacturonzuur afsplitsen. Kinetiekexperimenten op oligogalacturonzuren van verschillende lengtes toonden aan dat de hoogste activiteit werd behaald op pentaoligogalacturonzuur. Deze kinetiekdata werden gebruikt om de subsite-affiniteit van PelB met het substraat te berekenen. Deze data werden gecombineerd met structuurinformatie van homologe endogalacturonases om tot een substraat-bindingsmodel te komen. Het model komt tot 4 subsites, waarin de hoogste affiniteit is weggelegd voor subsite +1. PelB is het meest thermoactieve en –stabiele hydrolytische pectinase dat tot nu toe bekend is. Een goede afzetmarkt voor dit enzym zou de vruchtensapindustrie kunnen zijn. De klarificatie en kleurextractieprocessen die daarin worden toegepast worden vaak uitgevoerd bij verhoogde temperaturen.

T. maritima’s groeicapaciteit op pectine, waarbij het wordt ondersteund door de twee pectinolytische enzymen PelA en PelB, vormden de aanzet tot een nadere bestudering van het pectine catabolisme. Een ‘gerichte’ (targeted) DNA microarray werd gebruikt om de verschillende expressieniveaus van *pelA*, *pelB* en andere geselecteerde *T. maritima* genen te volgen gedurende de groei op pectine en cellobiose. De resultaten van dit onderzoek staan weergegeven in **Hoofdstuk 7**. Genen die een verhoogd expressieniveau vertoonden wanneer werd overgeschakeld naar pectine als koolstofbron bleken zich hoofdzakelijk te bevinden in twee clusters. Het ene cluster bevat vooral homologen van enzymen die een rol spelen bij de depolymerisatie, opname en modificatie van pectine, het tweede cluster codeert voornamelijk voor enzymen die waarschijnlijk verantwoordelijk zijn voor het cytoplasmatische catabolisme van verzadigd en onverzadigd galacturonzuur. Zoals verwacht werden verhoogde expressieniveaus gevonden voor *pelA* en *PelB*, alsmede voor een gen, waarvan wordt vermoed dat het codeert voor een derde galacturonase. Andere genen in dit eerste cluster met verhoogde expressieniveaus waren een esterase en een alcohol dehydrogenase, waaraan mogelijk een rol in de cytoplasmatische omzetting van methylgroepen van pectine oligosacchariden kan worden toebedeeld. In het tweede cluster bevinden zich een KDG kinase en een KDPG aldolase met een hoger expressieniveau op pectine. Deze enzymen verbinden hoogstwaarschijnlijk het galacturonaatcatabolisme met de glycolyse. De kdgR-homoloog in dit cluster zou een rol kunnen spelen in de regulatie van de expressie van een groot aantal van deze pectinolytische genen in *T. maritima*. In combinatie met biochemische data hebben de microarray resultaten geleid tot een voorstel voor de pectine catabolisme route in *T. maritima*.

Onze kennis over thermostabiele enzymen als biocatalysatoren is in de loop van de afgelopen jaren enorm gegroeid, voornamelijk dankzij het toenemende aantal thermozymen dat geproduceerd wordt in mesofiele gastheren, wat de opwerking en karakterisatie vereenvoudigd. De studie naar thermozymen, zoals in dit proefschrift beschreven is, is een voorbeeld van een geïntegreerde benadering waarin bioinformatica, moleculaire genetica, structuuranalyses, DNA microarray analyses, biochemie en fermentatie zijn gecombineerd. Verwacht wordt dat in het post-genomisch tijdperk het aantal karakteriseringen van thermozymen exponentieel zal gaan groeien, en daarmee ook de toepassingsdiversiteit van biocatalysatoren. Recente ontwikkelingen zoals high-throughput random mutagenese, 'reageerbuis'-evolutie en het gebruik van computertechnieken, hebben het pad geëffend voor de ontwikkeling van zg. tailor-made enzymen (2). De commercialisatie van deze technologieen neemt in rasse schreden toe, en zal in de toekomst zeer waarschijnlijk van invloed zijn op de toepassing van enzymen in verscheidene industriële processen. Dit behelst ook de enzymen afkomstig van thermofielen.

1. Gödde, C., K. Sahm, L. Kluskens, W.M. de Vos, and G. Antranikian. Cloning and expression of a serine protease encoding gene from *Fervidobacterium islandicum*. Poster presentation VAAM, Göttingen 2002.
2. Huisman, G., and S.G. Sligar. 2003. Evolution of protein technologies from voodoo to science. *Curr. Opin. Biotechnol*. **14**: 357-359

CURRICULUM VITAE

Leonardus Dorothea Kluskens werd op 7 juli 1972 geboren in het St. Jans Gasthuis te Weert. Na het succesvol doorlopen van de kleuterschool en Basisschool de Bongerd in Nederweert, werd de dagelijkse gang per fiets gemaakt naar de Philips van Horne Scholengemeenschap in Weert, alwaar hij in 1991 het diploma behaalde in het Voortgezet Wetenschappelijk Onderwijs.

Leon ging verder en verhuisde in 1991 richting Wageningen, waar hij begon aan de studierichting Levensmiddelentechnologie. Naarmate deze vorderde, verschoof zijn interesse steeds meer richting de genetische, moleculair-biologische en enzymatische aspecten van levensmiddelen. Met het doorlopen van het vervolgpracticum bij het Laboratorium voor Microbiologie betrad hij de fascinerende wereld van de hitteresistente micro-organismen, en de stap naar een eerste afstudeervak bij de werkgroep Bacteriële Genetica was toen eenvoudig gemaakt. Tijdens dit afstudeervak bestudeerde hij 5 maanden lang de cofactorspecificiteit van het thermostabiele glutamaat dehydrogenase uit *Pyrococcus furiosus*. Een tweede afstudeervak van 6 maanden werd gevolgd bij Industriële Microbiologie, waar hij de heterologe expressie van schimmelgenen in de gist *Saccharomyces cerevisiae* nader onder de loep nam. Vervolgens toog hij, daags nadat het Nederlands volleybalteam de gouden medaille had veroverd op de Olympische Spelen in Atlanta, richting l'Institut des Produits de la Vigne in Montpellier voor een 6 maanden durende stage, alwaar het mooie leven gerechtvaardigd werd door nuttig onderzoek naar een gen coderend voor een β-glucosidase uit *Aspergillus oryzae*. Eenmaal terug in Nederland studeerde hij in september 1997 af, met als oriëntatie bioconversies.

Het vuur voor hyperthermofiele micro-organismen werd weer aangewakkerd toen de mogelijkheid zich voordeed om bij Bacteriele Genetica twee jaar lang aan de heterologe expressie van een thermostabiel keratinolytisch enzym te werken. Ondanks het koppige karakter van dit enzym werd zijn contract tussentijds verlengd tot de promotiewaardige periode van 4 jaar, wat hem in staat stelde onderzoek te doen naar de karakteristieken van meerdere thermostabiele eiwitten. Dit alles leidde tot een aantal wetenschapplijke manuscripten, welke gebundeld zijn in dit proefschrift.

Sinds november 2002 is Leon werkzaam als onderzoeker bij BioMaDe, in Groningen.

LIST OF PUBLICATIONS

W.M. de Vos, W. Voorhorst, M. Dijkgraaf, L.D. Kluskens, J. van der Oost, and R.J. Siezen (2001) Purification, characterisation, and molecular modelling of pyrolysin and other extracellular thermostable serine proteases from hyperthermophilic microorganisms. *Methods Enzymol* **330**: 383-393

T. Kaper, C. Verhees, J.H.G. Lebbink, J. van Lieshout, L. Kluskens, D.E. Ward, S.W.M. Kengen, W.M. de Vos, and J. van der Oost (2001) Characterization of β-glycosyl hydrolases from *Pyrococcus furiosus*. *Methods Enzymol*. **330**: 329-346

L.D. Kluskens, W.G.B. Voorhorst, R.J. Siezen, R.M. Schwerdtfeger, G. Antranikian, J. van der Oost, and W.M. de Vos (2002) Molecular characterization of fervidolysin, a Subtilisin-like serine protease from the thermophilic bacterium *Fervidobacterium pennivorans*. *Extremophiles* **6**: 185-194

L.D. Kluskens, G.-J. van Alebeek, A.G.J. Voragen, W.M. de Vos, and J. van der Oost (2003) Molecular and biochemical characterisation of the thermo-active family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima*. *Biochem. J*. **370**: 651-659

J.-S. Kim, L.D. Kluskens, W.M. de Vos, R. Huber and J. van der Oost (2003) Crystal structure of fervidolysin from *Fervidobacterium pennivorans*, a keratinolytic enzyme related to subtilisin. J. Mol. Biol. *in press*

L.D. Kluskens, K.H. Shockley, D.E. Ward, S.B. Conners, R.M. Kelly, W.M. de Vos, and J. van der Oost (2003) Transcriptional analysis of pectin catabolic pathways in the hyperthermophile *Thermotoga maritima. Submitted*

L.D. Kluskens, A.C.M. Geerling, J. Zeilstra, W.M. de Vos, and J. van der Oost (2003) Isolation, expression and characterization of a glucose isomerase produced by the thermophilic bacterium *Fervidobacterium gondwanense. Submitted*

L.D. Kluskens, G.-J.W.M. van Alebeek, J. Walther, A.G.J. Voragen, W.M. de Vos, and J. van der Oost (2003) *PelB* from the hyperthermophilic bacterium *Thermotoga maritima* encodes a fungal-like intracellular exopolygalacturonase. *Submitted*

Cover image: Acidic mud pool in Solfatara, Pozzuoli, Naples, Italy
Cover design: Frank de Bok and Leon Kluskens

Printing: Grafisch bedrijf Ponsen & Looijen, Wageningen