MOLECULAR CLONING, KINETIC AND STRUCTURAL PROPERTIES OF FAMILY VII CARBOXYL ESTERASES

BY

MATSOBANE GODFREY TLOU

Submitted in fulfillment of the requirements for the degree

MAGISTER SCIENTIAE

In the faculty of Natural Sciences,
Department of Microbial, Biochemical and Food Biotechnology
University of the Free state,
Bloemfontein

January 2006

Promoters: Dr M.B Nthangeni
Dr E van Heerden
Prof. D Litthauer

ACKNOWLEDGEMENTS

Dr. MB Nthangeni, you took me under your wings and unselfishly showed me the ropes and for that I'm forever indebted to you. This work is evidence that tough love pays off.

Dr. E van Heerden, thank you for your undivided attention and assistance.

Prof. D. Litthauer, thank you for your sacrifices, time and advice.

Special thanks to Landi, Sibongile, Maria, Elsabe, Christelle, Sandile and the entire Extreme Biochemistry group. Its true, a problem shared is a problem solved.

To all the friends I have gathered in the department over the years. Thank you for your ears, smiles, chats and moral support.

To my parents and my siblings, I'm grateful for the encouragements and emotional support. Its true, trying and tough situations/times make us stronger.

Thank you, to the National Research Fund for the financial assistance.

I would not be where I am today if it wasn't for your grace and unconditional love. All this is through you my heavenly Father.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Literature review

1.1 General introduction

Lipases (E.C. 3.1.1.3) and esterases (E.C. 3.1.1.1) are hydrolases acting on the carboxyl ester bonds present in acylglycerols. Lipases and esterases show a fundamental difference in kinetics based on the properties of the substrate they hydrolyse (Jeager *et al*., 1994). Esterases catalyse the cleavage of ester bonds of short chain length fatty acids while true lipases have marked preference for longchain fatty acid substrates (Jeager *et al*., 1999). These enzymes contain a catalytic triad that consists of serine, histidine and aspartic acid, with the serine embedded in the consensus sequence Gly-X-**Ser**-X-Gly (where X represents any amino acid) at the active site (Wang and Hartsuck, 1993). Esterases show a wide substrate tolerance which led to the assumption that they have evolved to enable access to carbon sources or to be involved in catabolic pathways (Dalrymple *et al*., 1996; Ferreira *et al.,* 1993). These enzymes also display high regio- and stereospecificity, which make them attractive biocatalysts for the production of optically pure compounds in fine-chemical synthesis (Drauz and Waldmann, 1995; Bornscheuer and Kazlaukas, 1999).

1.2 Classification of lipolytic enzymes

Lipolytic enzymes are widely distributed in nature being found in plants, animals and micro-organisms (Villeneuve *et al.,* 2000). Classification of these enzymes is facilitated by a comparison of the substrate specificities of the enzymes, alignment of their amino acid sequences, comparison of their structural properties or on the bais of their biochemical and physiological properties (Bornscheuer, 2002). A classification scheme for esterases was proposed by Whitaker (1972), based on the specificity of the enzymes for the acid moiety of the substrate, such as the carboxylic ester hydrolases which catalyses the cleavage of the carboxylic acid esters. In addition to the carboxyl esterases, aryl esterases, acetyl esterases, cholin esterases, cholesterol esterases and lipases also belong to this group of hydrolytic enzymes. Classification of these enzymes by substrate specificity required that the enzymes to be compared be assayed with the same or related substrates under the same reaction conditions (Jeager *et al*., 1994).

Classification of lipolytic enzymes based on physiological properties is difficult due to the reason that the physiological functions of many esterases are not clear. This is attributed to the fact that many of them display a wide substrate specificity (Jeager *et al*., 1994), as a result it becomes difficult to assign them a specific physiological function. It has however been speculated that several classes of esterases exist, those that have evolved to enable access to

carbon sources (Dalrymple *et al*., 1996), those that are involved in catabolic pathways (Ferreira *et al.,* 1993), some that display biocide detoxification activity (Pohlenz *et al*., 1992) and those that play a pathogenic role (McQueen and Schottel, 1987) etc.

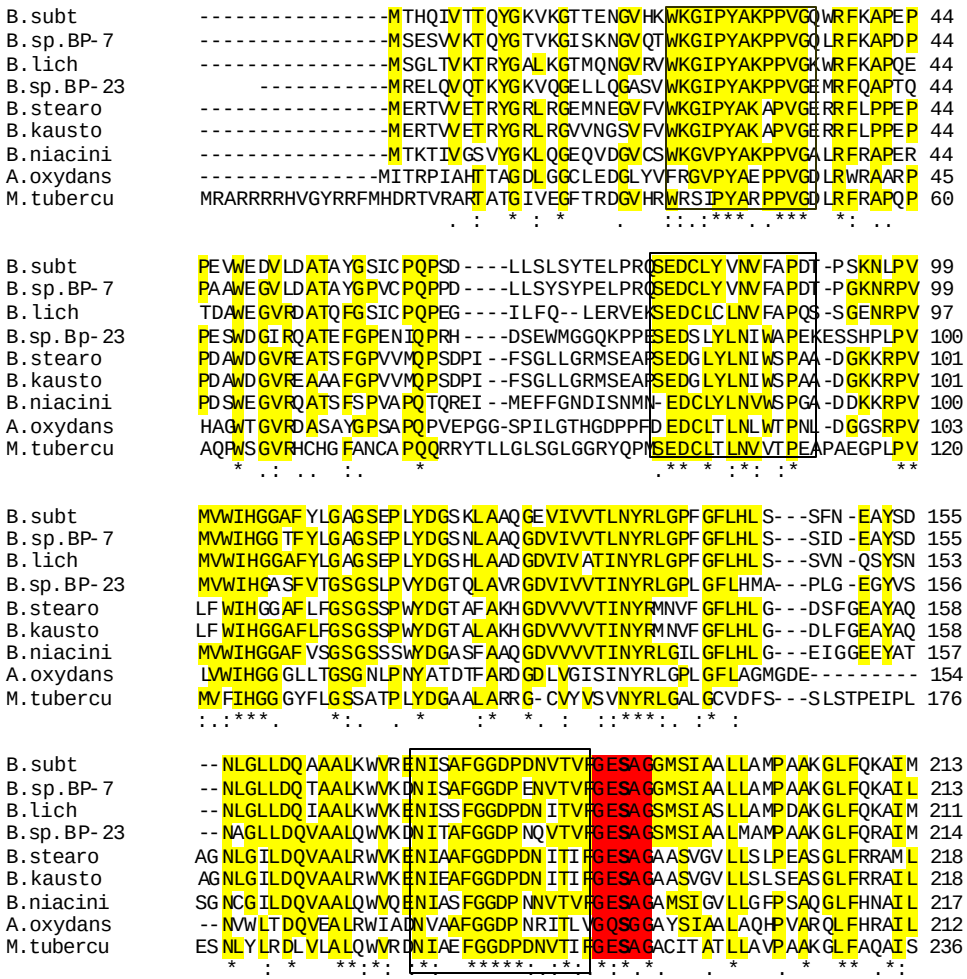
Classification of lipolytic enzymes by sequence comparison is facilitated by the increasing amount of sequence information on the public nucleotide databases. Comparison of amino acid sequence gives an indication of the evolutionary relationships between enzymes from different origins (Arpigny and Jeager, 1999) and reveals conserved sequence motifs which become characteristic features on which the classifications are based (Fiedler and Simons, 1995; Henikoff *et al*., 1997, Jaeger *et al*., 1999). In some cases, comparing enzyme amino acid sequences complements other forms of classification (i.e. classification by physiological role) by revealing conserved sequence motifs that suggest the ability of an enzyme to carry out a particular physiological function. As an example, the comparison of type B carboxyl esterase from *Peanibacillus* sp. BP-23 (Prim *et al*., 2000) to the phenidipham hydrolase from *Arthrobacter oxydans* P52 (Pohlenz *et al*., 1992), revealed the presence of a ß-lactamase signature S-X-X-K (Oefner *et al*., 1990), that suggested that the type B carboxyl esterase could also display biocide detoxification activity. However, high sequence homology cannot be related to enzyme properties such as, substrate specificity, stereoselectivity, pH, temperature optima, and in some cases completely different reactions are catalysed (Pelletier and Altenbuchner, 1995). As an example, a bromoperoxidase from *Streptomyces aureofaciens* (Hetch *et al*., 1994) shares

55% sequence identity to an esterase from *Pseudomonas fluorescens* (Pelletier and Altenbuchner, 1995) but they share very low substrate specificity.

Arpigny and Jeager (1999) collected the increased amount of information available from nucleotide, protein and crystal structures of bacterial lipolytic enzymes and proposed a comprehensive classification based mainly on the amino acid sequences and biochemical properties. This resulted in the identification of 8 different families with the largest being further divided into 6 subfamilies. Family I, which is subdivided into 6 subfamilies, contains the so called 'true' lipases: *Pseudomonas* lipases, lipases from gram positive bacteria, such as *Bacillus* and *Staphylococcus*, and other lipases, such as lipases from *Propionibacterium* and *Streptomyces.* The enzymes of Family II lack the classical pentapeptide Gly-X-Ser-X-Gly, but have a Gly-Asp-Ser-Leu motif instead. Within this family, esterases of *Strep. scabies, P. aeruginosa*, *Salmonella typhimurium, Photorbabdus luminescens* and *Aeromonas hydrophila* are found. In Family III, the extracellular lipases of *Streptomyces* and *Moraxella* are included, while Family IV comprises the enzymes similar to mammalian hormone-sensitive lipases. Enzymes originating from mesophilic bacteria (e.g *Pseudomonas oleovorans*, *Haemophilus influenza, Acetobacter pasteurianas*), from cold-adapted organisms (e.g. *Moraxella* species, *Psychrobacter immobilis*) and heat adapted organisms (*Sulfolobus acidocaldarius*) are grouped in Family V. Family VI, contains the smallest esterases known, having a molecular mass of 23-26 kDa. Enzymes found in this family include an esterase from *Pseudomonas fluorescens*, of which the

structure is known (Kim *et al*., 1997). The esterase is active as a dimer, has a typical Ser-Asp-His catalytic triad and hydrolyses small substrates and not long–chain triglycerides.

Family VII bacterial lipolytic enzymes are large enzymes (approximately 55 kDa) and share significant homology to eukaryotic acetylcholine esterases and intestine or liver carboxyl esterases (e.g. pig liver esterase). The family comprises of biotechnologically significant esterases such as, ofloxacin ester-hydrolysing esterase from *Bacillus niacini*, *p*-nitrobenzyl esterase from *Bacillus subtilis* and an esterase from *Arthrobacter oxydans* active against phenylcarbamate herbicides. The family also includes carboxyl esterases from thermophilic *Geobacillus kaustophilus* (Takami *et al*., 2004) and *Geobacillus stearothermophilus* and the carboxyl esterase from *Geobacillus stearothermophilus* has been demonstrated to be thermostable (Ewis *et al*., 2004). Several conserved motifs could be identified in the aligned amino acid sequences of Family VII lipolytic proteins (Figure 1.1).


B.subt
B.sp.BP-7
B.lich
B.sp.BP-23
B.stearo
B.kausto
B.niacini
A.oxydans
M.tubercu

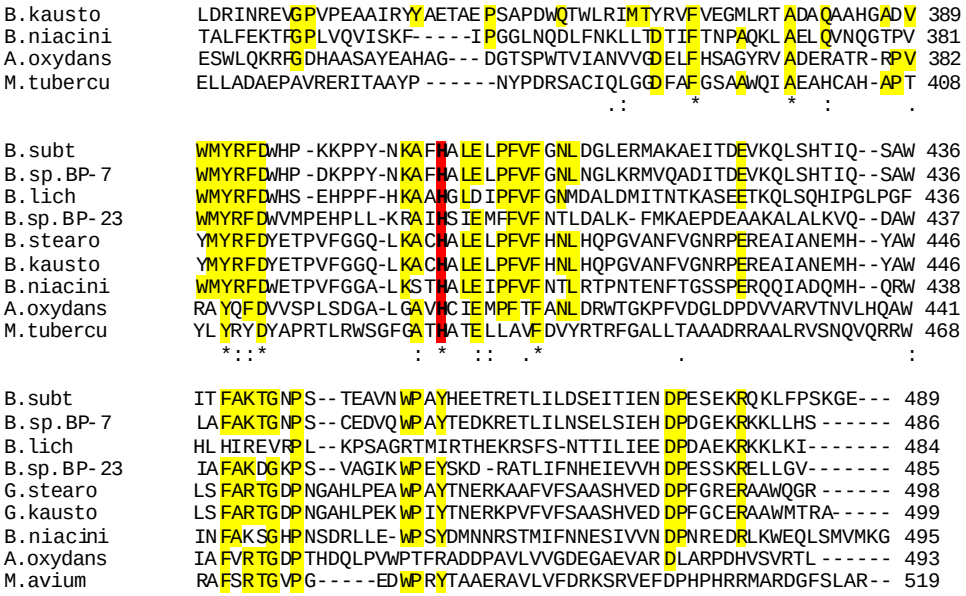


Figure 1.1: Amino acid alignment of Family VII carboxyl esterases. The consensus sequence, and the amino acids making up the catalytic triad are shown in red. Boxes include the three conserved signature patterns of Family VII carboxyl esterases. Amino acid identical in at least five sequences are shadowed in yellow. The sequences shown are: the paranitrobenzyl esterase from *B. subtilis* (B. subt, accession number U06089) carboxyl esterase from *B. lichernifromis* (B. lich, AJ315954), thermostable esterases from *Geobacillus. kaustophilus* (G. kausto, BA000043) and *G. stearothermophilus* (G. stearo, AY186196), cell bound esterases from *Bacillus* sp. Bp-7 (B. sp. BP-7, AJ278066), and *Bacillus* sp. BP-23 (B. sp. BP-23, AJ238680), lipase T from *Mycobacterium avium* (M. avium, NC_002944), and phenmedipham hydrolase from *Arthrobacter oxydans* (A. oxydans, Q01470).

1.3 Structural classification of lipolytic enzymes

Lipolytic enzymes are classified under the a/ß hydrolase fold family originally described by Ollis *et al*., (1992), based on their structural properties (Cygler *et al*., 1993). The a/ß hydrolase fold family is a growing superfamily of proteins with a wide range of properties. The a/ß-hydrolase fold (Figure 1.2) is characterised by a ß-sheet of five to eight strands connected by a-helices to form a/ß/a sandwich (Satoh *et al.,* 2002). The members of this family diverged

from a common ancestor into a number hydrolytic enzymes with a wide range of substrate specificity such as acetylcholine esterase (Sussman *et al*., 1991), serine carboxypeptidase (Liao and Remington, 1990) and haloalkane dehalogenase (Franken *et al*., 1991), together with other proteins with no known catalytic function (Hotelier *et al*., 2004). The enzymes contain the catalytic triad residues (serine, histidine and aspartate or glutamate) on the loops, of which one commonly referred to as the nucleophilic elbow which contains the active site serine residue and it is the most conserved feature of the fold (Figure 1.2) with the general conserved consensus sequence (Gly-X-Ser-X-Gly) (Arpigny and Jeager,1999).

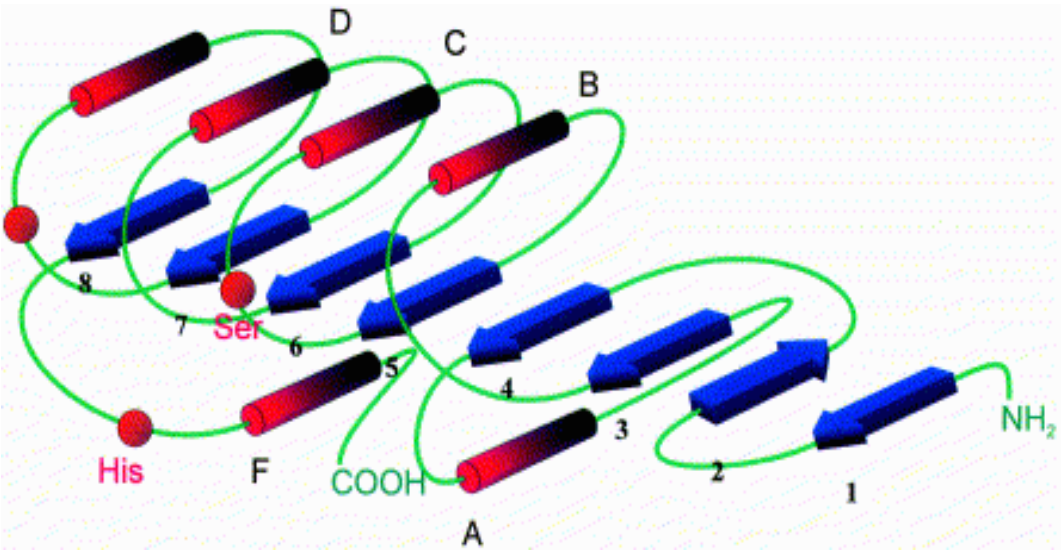


Figure 1.2: Schematic representation of the a/ß-hydrolase fold. ß-sheets (1-8) are shown as arrows, a-helices (A-F) as columns. The relative positions of the amino acids of the catalytic triad are indicated as circles (Bornscheuer, 2002).

1.4 Esterase enzyme assay methods

The most widely used plate assays for carboxl esterases assays contain triolein, olive oil or tributyrin which are emulsified mechanically in various growth media and poured into petri dishes (Jeager *et al.*, 1999). Lipolytic activity is observed by the formation of clear halos around the colonies growing on tributyrin containing agar plates (Atlas, 1996). Lipase activity on olive oil/triolein containing plates is visualized on long ultraviolet light (340 nm) as an orange-red flourescence emmission around the colonies growing on agar plates supplemented with Rhodamine B (Kouker and Jeager, 1987).

Lipase/esterase activity in bacterial culture supernatants is determined by hydrolysis of *p*-nitrophenyl esters (chromogenic substances) (Vorderwülbecke *et al*., 1992) and spectrophotomeric detection of *p*-nitrophenol at 410nm (Jeager *et al*., 1999). Lipases can be distinguished from esterases by their substrate spectra, using *p*-nitrophenyl palmitate (cleaved by lipases) vs. *p*-nitrophenyl butyrate (cleaved by esterases and sometimes also by lipases) (Jeager *et al*., 1999). However, common esterase substrates such as *p*-nitrophenyl acetate (Gutfreund and Sturtevant, 1956) lack sensitivity and stability in assay systems leading to high background signals. This prompted the development by Shan and Hammock (2001) of a reportedly sensitive esterase assay based on a-cyano-containing esters based on hydrolysis of a-

cyanopyrethreoids (derived from the dominant insecticide class pyrethroids) which produces an acid and a cyanohydrin that spontaneously rearranges into the corresponding *meta*-phenoxybenzyldehyde in aqeous solution (Shono *et al*., 1979). It was observed that this aldehyde molecule absorbs light much more strongly than the parent compound and the a-cyanophenoxybenzyl alcohol, providing the basis for a sensitive and specific assay for pyrethroid esterases (Shono *et al*., 1979). Therefore, Shan and Hammock (2001) generalised this concept by suggesting that esterase substrates containing an a-cyano group could be designed with little or no fluoresence that are transformed to a strongly fluorescing aldehyde upon ester hydrolysis. They further demonstrated this by synthesising and evaluating pyrethroid-like substrates wich yielded a fluorescent product upon hydrolysis.

1.5 Esterase/lipase catalysed reactions and the reaction mechanism

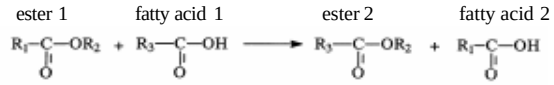
Under aqueous conditions esterases act on ester bonds present in acylglycerols to liberate free fatty acids and glycerol (hydrolysis) [Figure 1.3 (a)]. Under micro-aqueous conditions, these enzymes possess the ability to carry out the reverse reaction (esterification) [Figure 1.3 (b)], acidolysis: the exchange of acyl radicals between an ester and an acid [Figure 1.3 (c)], interesterification: the exchange of acyl radicals between an ester and ester [Figure 1.3 (d)] and alcoholysis: the exchange of acyl radicals between an ester and alcohol [Figure 1.3 (e)] (Villeneuve *et al*., 2000). Total hydrolysis of ester bonds, production of fatty acids, ester synthesis and the modification of

fats and oils (transesterification) can be carried out chemically under harsh conditions of pressure and temperature (Ebbing, 1996). However, the use of lipases/esterases is more desired since the reactions proceed under mild conditions of pressure and temperature and with specificity and reduced chemical waste (Malcata *et al*., 1990).

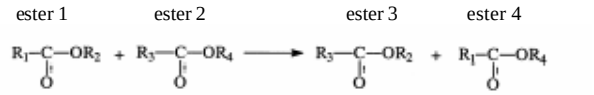
(a) Hydrolysis:

(b) Esterification:

(c) Acidolysis:



(d) Interesterification:



(e) Alcoholysis:

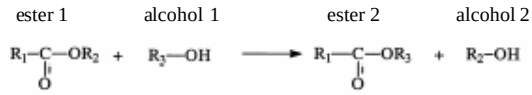


Figure 1.3: Different reactions catalysed by lipases/esterases in aqueous and non-aqueous solutions (Villeneuve *et al*., 2000).

The primary function of esterases and lipases is the hydrolysis of carboxyl ester bonds present in acyl glycerols to liberate fatty acids and glycerols. As detailed above, their active sites consists of a Ser-His-Asp/Glu catalytic triad. This catalytic triad is similar to that observed in serine proteases, and therefore catalysis by carboxyl ester hydrolases has been suggested to be similar to that of serine proteases (Jeager *et al*., 1999). Hydrolysis of the substrate takes place in four steps (Figure 1.4). It starts with an attack by the oxygen atom of the hydroxyl group of the nucleophilic serine residue on the activated carbonyl carbon of the carboxyl ester bond and a transient tetrahedral structure is formed, which is characterised by a negative charge on the carbonyl oxygen atom of the scissile ester bond and four atoms bonded to the carbonyl carbon atom are arranged as a tetrahedron (Figure 1.4). The nucleophilicity of the attacking serine is enhanced by the catalytic histidine, to which a proton from the serine hydroxyl group is transferred. This proton transfer is facilitated by the presence of the catalytic acid, which orients the imidazole ring of the histidine and partly neutralises the charge that develops on it. At this stage an acid component of the substrate is esterified to the nucleophilic serine, whereas the alcohol component diffuses away (Figure 1.4). The next stage is the deacylation step, in which a water molecule hydrolyses the covalent intermidiate. The active-site histidine activates this water molecule by drawing a proton from it. The resulting OH^- ion attacks the carbonyl carbon atom of the acyl group covalently attached to the serine (Figure 1.4). Again, a transient negatively charged tetrahedral intermediate is formed, which is stabilised by the interaction with the oxyanion hole. The histidine donates a proton to the oxygen atom of the serine residue, which

then releases the acyl component. After diffusion of the acyl product the enzyme is ready for another round of catalysis (Figure 1.4).

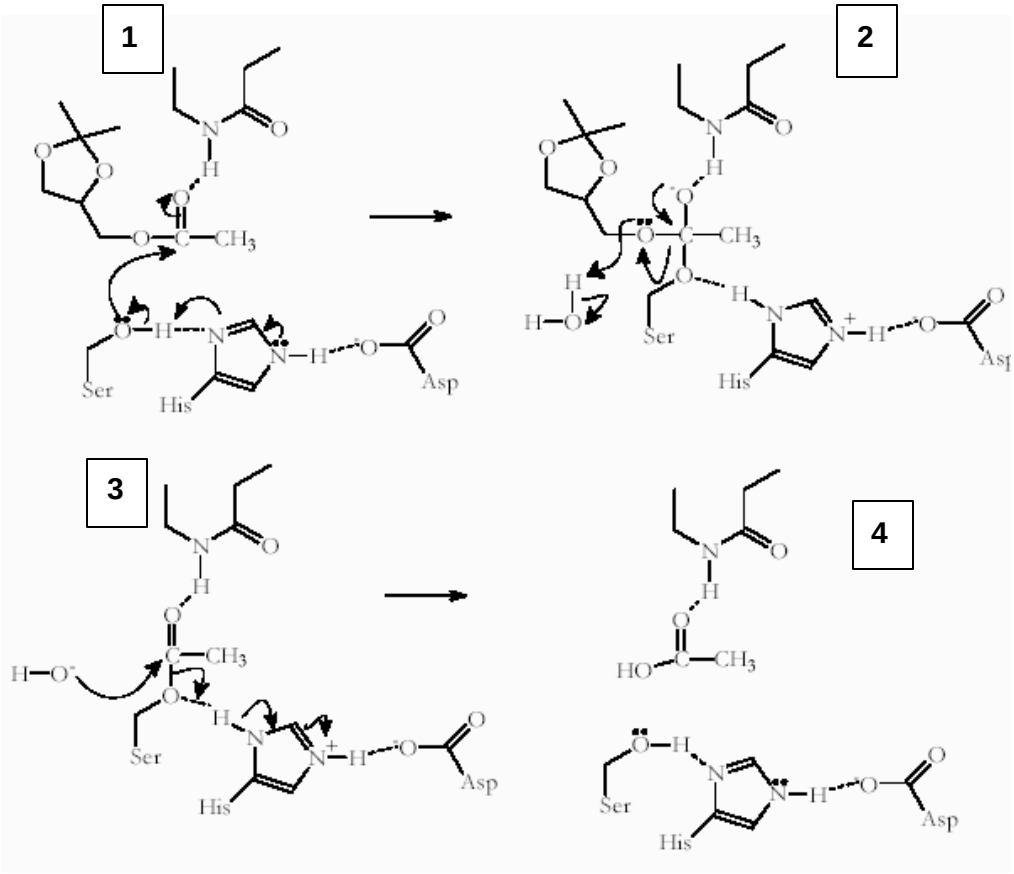


Figure 1.4: Reaction mechanism of carboxyl ester hydrolases. (1) Binding of the carboxyl ester (substrate), activation of the nucleophilic serine residue by neighboring histidine and nucleophilic attack of the substrate's carbonyl carbon atom by Ser O^-. Transient tetrahedral intermediate, with O^- stabilized by interactions with two peptide NH groups. The histidine donates a proton to the leaving alcohol component of the substrate. (3) The covalent intermediate ("acyl enzyme"), in which the acid component of the substrate is esterified to the enzyme's serine residue. The incoming water molecule is activated by the neighboring histidine residue, and the resulting hydroxyl ion performs a nucleophilic attack on the carbonyl carbon atom of the covalent intermediate. (4) The histidine residue donates a proton to the oxygen atom of the active serine residue, the ester bond between serine and acyl component is broken, and the acyl product is released (Dröge *et al*., 2000; Jeager *et al*., 1994).

Although lipases and esterases display a similar hydrolytic mechanism, there are however some differences based on their substrate specificities.

Esterases preferentially hydrolyse acyl glycerols of shorter fatty acids whereas lipases display a much broader substrate range (Malcata *et al*., 1992). It appears that the physical state of the substrate is the most likely contributing factor towards the substrate specificity. Long chain fatty acids are typically insoluble or at least poorly soluble (emulsion). Thus the lipase has to be able to identify an insoluble or heavily aggregated substrates (Jeager *et al*., 1994). Since lipases are active towards aggregated substrates, lipase activity is directly correlated to substrate area (interfacial activation), and not with the substrate concentration (Verger, 1998). However, contrary to lipases, esterase activity is not correlated to substrate area, since esterases are active against soluble substrates. Desnuelle (1961) demonstrated a fundamental difference between esterase and lipase based on their ability to be activated by interfaces. Esterase activity is a function of concentration as described by Mechaelis-Menten kinetics with the maximal rate being reached long before the solution becomes substrate saturated (Figure 1.5); the formation of the substrate/water emulsions does not change the reaction rate. In contrast, lipases showed almost no activity with the same substrate as long as the substrate was in its monomeric form. However, when the solubility limit of the substrate is exceeded, there is a sharp increase in activity as the substrate forms an emulsion (Figure 1.5).

The dependence of lipase activity on the presence of an interface led to their definition as carboxyl esterases acting on emulsified substrates (Jeager *et al*., 1994). Elucidation of the first lipase three-dimensional structures led to an explanation of the interfacial activation phenomenon, displayed by lipases. It

was found that the active site of lipases was covered by a lid-like polypeptide chain which rendered the active site inaccessible to substrate molecules, thereby causing the enzyme to be inactive on monomeric substrate molecules (Brady *et al*., 1990). However, when the enzyme was bound to a lipid interface, a conformational change took place causing the lid to move away whereby the active site of the active site became fully accesible. As a result, the hydrophobic side of the lid became exposed to the lipid phase, thus enhancing hydrophobic interactions between the enzyme and the lipid surface (Verger, 1998). The interfacial activation of lipases as a function of the lid has been used to distinguish them from esterases which lack the lid like polypeptide. However, this form of differentiation should be used with care since lipases that lacked the lid but displayed interfacial have been identified (van Pouderoyen *et al*., 2001), and some lipases that contained lid-like structures but did not display interfacial activation have also been reported (Noble *et al*., 1993, Uppenberg *et al*, 1990).

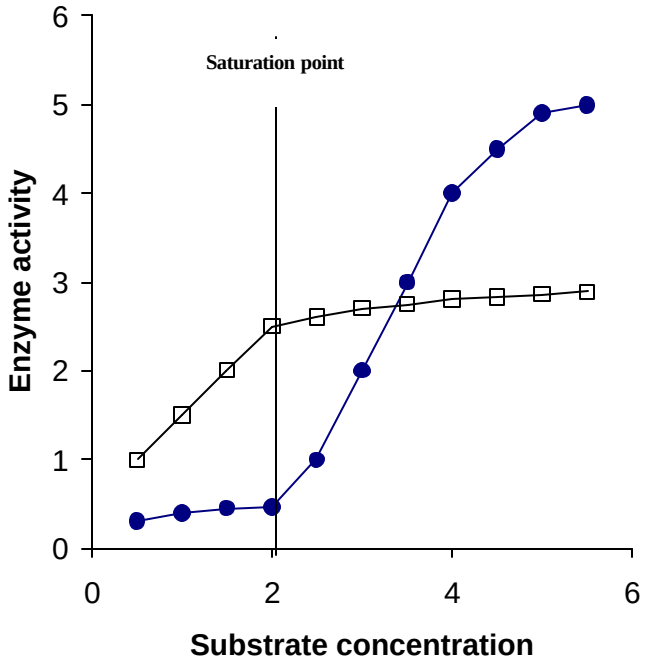


Figure 1.5: Classical activity profile of a pancreatic lipase (?) and a horse liver esterase (?) exceeding the saturation point (adapted from Jeager *et al*., 1994).

1.6 Physicochemical properties of esterases

The surface of a protein constitutes the interface through which the protein interacts with its surrounding environment. The molecular basis for the interactions between an enzyme and its substrate or inhibitor or any other type of molecular recognition is facilitated by surface contacts (Petersen *et al*., 2002). The surface is a complex steric arrangement of residues, where any residue can be found. However, the type and exact position of the different residues has a crucial impact on the functional parameters such as thermal stability, substrate specificity, activity and pH optima (Petersen *et al*., 2002).

1.6.1 pH

The molecular understanding of the interactions between a protein and its substrate, or inhibitor, is an understanding of the role of electrostatics in intermolecular interactions, such as molecular recognition. The distribution of the electrostatic potentials on the molecular surface of an enzyme is a function of pH and determines to a large extent the pH activity profile (Petersen *et al*., 2001). Other important factors include the presence and the distribution of titratable residues (residues that carry a charge above or below a certain pH) in the active cleft (Fojan *et al*., 2000). Some amino acid residues near the binding site are also titratable therefore, in order for an enzyme to function; these residues and residues within the active cleft should display an appropriate ionic state at an appropriate pH (Fojan *et al*., 2000). Since many substrates also have an ionic character, the active site of the enzyme may require particular ionic species of the substrate for optimum acitivity (Petersen *et al*., 1997).

The effect of pH on the reaction rate of an enzyme can suggest which ionisable residues are in the active site. Sensitivity to pH usually reflects an alteration in the ionisation state of one or more residues involved in catalysis and occasionally substrate binding (Horton *et al*., 1996). Usually, the catalytic activity of esterases and lipases changes with pH in a bell shaped fashion, thus yielding a maximum rate in the stability range (Zaks and Klibanov, 1985; Henke and Bornscheuer, 2002). Generally, bacterial lipases/esterases have neutral (Henke and Bornscheuer, 2002, Lee *et al*., 1999) and alkaline pH

optima (Alvarez-Macarie *et al*., 1999; Ewis *et al*., 2004; Eggert *et al*., 2000), with an exception of a lipase from *Pseudomonas fluorescens* SIK W1, which had maximum activity at pH 4.8 (Andersson *et al*., 1979).

1.6.2 Temperature

The effect of temperature on chemical reactions is basically described by the Arrhenius equation :

$$k = Ae^{-EA/RT}$$

Where k is the kinetic rate constant, A, the Arrhenius constant, Ea, the activation energy, R, gas constant and T, the absolute temperature (Ebbing, 1996). According to the Arrhenius equation, an increase in the temperature will induce an exponential increase in the reaction rate, while a decrease in the temperature induces a decrease in the reaction rate (Ebbing, 1996). However, with respect to enzymes the Arrhenius equation applies within a relatively small temperature range. This is due to the fact that enzymes are proteins and undergo inactivation or denaturation at temperatures below or above those to which they are ordinarily exposed in the environment (Nofziger, 2005).

The temperature range in which biological activity has been detected extends from -20 °C, the temperature recorded in the brine veins of the Arctic and

Antarctic sea ice (Deming, 2002), to 113 °C, the temperature at which the archea *Pyrolobus fumarii* is still able to grow (Blöchl *et al.,* 1997). Among the extremophilic microorganisms, those living at extreme temperatures have attracted much attention. Thermophiles have revealed the unsuspected upper temperature for life (Blöchl *et al.,* 1997). Their enzymes have also demonstrated a considerable potential such as the various thermostable DNA polymerases which are routinely used in PCR techniques. Psychrotrophic bacteria which face thermodynamic challenges to maintain enzyme-catalyzed reactions and metabolic rates compatible with sustained growth near or below freezing point of pure water have also been detected (D'Amico *et al.,* 2002). The use of cold active enzymes can be the key to the success of various applications such as the organic synthesis of compounds unstable at high temperatures (Gerday *et al*., 2000).

Biological activity detected at extreme temperatures is believed to be due to molecular adaptation that enzymes have undergone in response to environmental temperatures (D'Amico *et al*., 2002; D'Amico *et al*., 2003; Georlette *et al.,* 2003; Deming 2002). Molecular adaptation which is in essence a natural evolution of the enzymes occurred in response to two distinct selective pressures i.e. the requirement for stable protein structure and activity in thermophiles and the requirement of high enzymatic activity in psychrophiles (D'Amico *et al.,* 2003). Although cold active enzymes display comparable structures with those of their meso- and thermophilic homologues, there are however, some underlying differences in their flexibility and stability which are crucial points in enzyme adaptation to temperature

(Gianese *et al*., 2002). There is a clear increase in the number and strength of all known weak interactions and structural factors involved in protein stability from psychrophiles to mesophiles to thermophiles (Gianese *et al*., 2002). It was found that all the structural factors known to stabilize mesophilic and thermophilic poteins could be attenuated in strength and number in cold active enzymes. This could possibly involve an increased number and clustering of glycine residues; a decrease in proline residues in loops; a reduction in arginine residues, capable of forming multiple electrostatic interactions and bonds, and eventually a lowering of the number of ion pairs, aromatic interactions, hydrophobic interactions or hydrogen bonds in comparison with mesophilic and thermophilic homologues (D'Amico *et al*., 2002). Each protein uses a few of these structural modifications to acquire the required flexibility to be more or less adapted to the environment. As an example, the molecular activities of cold-adapted enzymes are much higher at low temperatures; this has been proposed to be due to the molecular flexibility of active sites of these enzymes (Gerday *et al*., 2000).

A wide range of esterases, from cold active to mesophilic to (hyper-) thermopilic have been identified. These enzymes are currently the centre of attraction due to their potential biotechnological applications in medicine, synthetic chemistry and food processing (Bornscheuer, 2002). The most attractive mesophilic esterases seem to originate from *Bacillus* and *Pseudomonas* species (Bornscheuer and Kazlaukas, 1999). For example, the naproxen esterase of *B. subtilis* Thai I-8 (Quax and Broekhuizen, 1994) was characterized as a very efficient enantioselective biocatalyst for the kinetic

resolution of non-steroidal anti-inflammatory drug (NSAID) esters, such as naproxen and ibuprofen methyl esters. A thermostable esterase from a hyper-thermophilic archeon, *Pyrobaculum calidifontis* VA1 that remained remarkably stable after incubation at 100 to 110 °C for 2hours has been identified by Hotta *et al*., (2002). Morana and co-workers (2002) also cloned and characterized a carboxyl esterase from an archeaon *Sulfolobus solfataricus* which was observed to retain 100% activity at 70 °C for 24hours, 50% activity at 80°C, and 15% residual activity at 90 °C. Numerous other esterases from thermophilic *Bacillus* species capable of remaining stable at temperatures above 60 °C have also been cloned and characterized (Ewis *et al*., 2004; Henke and Bornscheuer, 2002). The stability of these enzymes is necessary for industrial biotransformations because from an industrial point of view, the more thermostable the enzyme the more interesting it is for the development of new applications (Alavares-Macarie *et al*., 1999). On the other end of the temperature scale, several cold active esterases have been identified in cold-adapted microorganisms. As examples, esterases from *Moraxella* sp. Strain TA144 (Feller *et al*., 1991) and *Acinetobacter* sp. no. 6 (Suzuki *et al.,* 2002) which exhibited high specific activity at 4 ℃ and an esterase from *Pseudomonas citronellolis* which shares amino acid sequence similarity with HSL (mammalian hormone sensitive lipase, known to retain high activity at low temperature) has also been found to retain maximal activity at 15°C (Chao *et al*., 2003).

1.6.3 Enantioselectivity of esterases

Enzymes catalyse a broad spectrum of reactions with high selectivity and efficiency under mild and environmentally friendly conditions. Enzymatic transformations of organic substrates have often been used for the preparation of optically active compounds (Koeller and Wong, 2001). Biocatalytic resolution is one way to obtain enantiomerically enriched compounds by exploiting the selectivity of enzymes towards one form of the enantiomers of a racemic molecule (enantioselectivity) (Thomas *et al.,* 2002).

The mechanism behind enzyme enantioselectivity is not fully understood at present. Numerous speculations have been made in an attempt to describe the mechanism of enzyme steroselectivity (Sundaresan and Abrol, 2002; Kafri and Lancet, 2004). However, the most widely accepted model that best describes the mechanism of enantioselectivity is the three-point attachment (TPA) model (Copeland 2000; Ahn *et al.,* 2001), according to which, one enantiomer of a chiral substrate binds to a protein simultaneously at three sites, while the opposite enantiomer cannot bind to the same three sites.

There is a large body of literature on the use of microbial lipases and esterases in the synthesis of optically pure compounds (Jeager *et al*., 1999; Bornscheuer, 2002). However, due to the moderate enantioselectivity displayed by many of these enzymes, improvement of esterase/lipase enantioselectivity by directed evolution is an active field of research (Bornscheuer *et al.,* 1997; Bornscheuer, 2002, Dröge *et al*., 2000).

1.7 Biotechnological properties of esterases

Carboxyl esterases have been perceived by research scientists as one of the most important classes of industrial biocatalysts. In terms of their versatility, activity, stability in organic solvents and their high regio- and enantio-selectivity, this perception has been justified. The annual sales of lipases once accounted for 20 million US dollars which was less than 4% of the worldwide enzyme market, which was estimated at 600 million dollars (Arbige and Pitcher, 1989). Over the years, commercial use of lipases and esterases has turned into a billion dollar business, which comprises of a wide variety of different applications in the area of detergents, and the production of food ingredients and enantiopure pharmaceuticals (Table 1.1) (Villeneuve *et al*., 2000).

Carboxyl esterases are important in the metabolism of many exogenous compounds including pesticides and pharmaceuticals. Carboxyl esterase mediated hydrolysis is used in the design of many drugs and pharmacophores (Wheelock *et al*., 2001). For example, the hydrolysis of ester containing prodrugs has been employed in the development of chematherapeutic agents CPT-11 (Senter *et al*., 1996) and 10-hydroxycamptothecin fatty acid esters (Takayama *et al*., 1998). Some carboxyl esterases are being employed in the synthesis of optically pure compounds. Probably the best studied esterase for this purpose is the so-called carboxyl esterase NP (NP from naproxen, a non-steroidal anti-inflammatory drug) originating from *Bacillus subtilis* (Quax and

Broekhizen, 1994). Besides naproxen, this enzyme has been found to produce various other 2-arylpropionic acids with high enantioselectivity (Azzolina *et al*., 1995). Carboxyl esterases are also important in the hydrolysis and subsequent detoxification of pyrethoid (Casida *et al*., 1983) and carbamate (Gupta and Dettbarn, 1993; Pohlenz *et al*., 1992). However, an application that is of considerable industrial interest is the mild removal of protecting group as shown for a *p*-nitrobenzyl esterase from *B. subtilis*, which specifically remove this residue from the antibiotic Loracarbef (Zock *et al*., 1994).

Esterases can also be used in the synthesis of major flavour compounds such as vanallin (Lesage-Meessen *et al*., 1996). Vanillin is produced as a result of the activity of carboxyl esterases on plant cell wall polysaccharides such as pectin and xylan. The carboxyl esterases produce ferulic acids from the polysaccharides which is then enzymatically converted to vanillin. Feruloyl esterases have been isolated and characterised from a number of organisms. (Christov and Prior, 1999; Donaghy and Mckay, 1997; De Vries and Visser, 1999).

A considerable number of microbial carboxyl esterases is known, however, only a few of them have been used for biotechnological purposes. The major reason for this are their limited commercial availability and their frequently observed moderate enantioselectivity (Bornscheuer and Pohl, 2001). Researchers are currently trying to overcome such shortcomings by altering substrates (substrate engineering), by modifying reaction systems (medium

engineering) or by enzyme engineering (Adamczak and Krishna, 2004). Most of the enzyme engineering techniques utilise mutagenesis to introduce a limited number of alterations. The choice of residue to be altered may be based on protein design, sequence similarity, structural modelling or determined through screening or selection-based forced evolution (random mutagenesis) approaches or a combination of these techniques (Davis, 2003). However, these methods are limited to just the 20 primary proteinogenic amino acids. Ingenious molecular biological techniques have been developed for the introduction of non-natural amino acids into proteins. The incorporation of non-coded residues is most successful for those that resemble their coded counterparts, and the use of more complex amino acids in such techniques can result in poor levels of incorporation (Davis, 2003). Other forms of enzyme modification exist such as physical immobilisation of the enzyme to an insoluble support (Villeneuve *et al*., 2000). Physical immobilisation is done to overcome lack of enzyme stability under process conditions and also the difficulties in enzyme recovery and recycling (Adamczak and Krishna, 2004). The use of such modified enzyme also results in improved operational thermal stability (Montero *et al*., 1993).

Table 1.1

Examples of industrial applications of lipases/esterases (Villeneuve *et al*., 2000)

Field of industry	Application	Products
Hydrolysis		
Food (dairy)	Hydrolysis of milk fat	Flavoring agents for dairy products
Chemical (oil processing)	Hydrolysis of oils and fats	Fatty acids, mono/di -glycerides
		Reagents for lipid analysis
Chemical (detergent)	Removal of oil stains	L aundry and house hold detergents
Medical	Blood tryglyceride assays	Diagnostic kits
Esterification		
Chemical (fine chemical)	Synthesis of fine chemicals	Chiral intermidiates
		Esters, emulsifiers
Food (chemical and pharmaceutical)	Transesterification of natural oils	Oils or fats (cocoa butter).

1.8 Conclusions

Carboxyl esterases are ubiquitous in nature, being found in plants and microorganisms. Microbial craboxyl esterases are very diverse in their physiological properties, suggestive of their potential biotechnological importance. They are also very diverse in their enzymatic properties and substrate specificities which makes them attractive biocatalysts for biotechnological applications.

A large number of carboxyl esterase encoding genes have been cloned from various bacteria. However, the most attractive for biotechnological applications are from *Bacillus* and *Pseudomonas* species. This is because, esterases from these species have been shown to be capable of catalysing reactions of biotechnological significance, such as the synthesis of optically pure compounds (Kim *et al*., 2003, 2004, Quax and Broekhuizen, 1994). However, limited commercial availibility of esterases in general and their frequently observed selectivity, limits applications of these enzymes at industrial level. Directed evolution studies (aimed at designing enzymes with improved properties), microbial screening and development of improved gene cloning strategies with the aim of isolating novel enzymes are currently active fields of research.

Microbial carboxyl esterases that have been cloned thus far have grouped into various classes and families based on conserved amino acid sequence motifs and biochemical properties. The sequence information on carboxyl esterases on the public nucleotide database facilitates the classification of

newly cloned genes and the development of sequence based gene cloning strategies. Therefore, the objective of the study was to identify conserved sequences within Family VII carboxyl esterases and to use these as templates for the designation of universal primers to detect the presence of Family VII members of carboxyl esterases within genetic materials from different sources, to clone, over-express and characterize enzymes belonging to this family.

CHAPTER 2

Molecular detection of Family VII carboxyl esterase genes and cloning by the improved cassette ligation-mediated PCR of a complete gene encoding *Bacillus pumilus* carboxyl esterase

2.1 Introduction

Esterases (EC 3.1.1.3) and lipases (EC 3.1.1.1) are carboxyl ester hydrolases that catalyse the formation or cleavage of carboxyl ester bonds of acylglycerols. Lipases and esterases show a fundamental difference in kinetics, based on the substrates they hydrolyse. Esterases can be

distinguished from lipases by their preference for short chain acylglycerols substrates and the lack of requirement of interfacial activation (Jeager *et al*., 1999). Interfacial activation refers to an increase in activity displayed by lipases when acting at a lipid water interface of micellar or emulsified substrates (Verger, 1998). The active site of esterases and related enzymes consists of three catalytic residues: a nucleophilic residue (serine), a catalytic acid residue (aspartate or glutamate), and a histidine residue, always in this order in the amino acid sequence (Ollis *et al*., 1992). The nucleophilic serine residue is predominant in lipases and esterases and it is embedded within the not so conserved "consensus" sequence Gly-X-Ser-X-Gly at the active site (Wang and Hartsuck, 1993). Carboxyl ester hydrolases with a "consensus" sequence Ala-X-Ser-X-Gly around the catalytic serine have also been identified (Kim *et al*., 2002).

The physiological functions of many esterases are not clear. Some of these enzymes are known to be involved in metabolic pathways that provide access to carbon sources; such enzymes include the acetyl- and cinnamoyl esterases that are involved in the degradation of hemicellulose (Dalrymple *et al*., 1996; Ferreira *et al.,* 1993). In some plant pathogenic bacterial and fungal strains these cell wall degrading esterases are believed to be pathogenic factors (McQueen and Schottel, 1987). Detoxification of biocides or insecticide resistance is often a result of esterases that hydrolyse insecticides (Blackman *et al*., 1998). As an example, an esterase from *Bacillus subtilis* that hydrolyse the phytotoxin brefeldin A has been described (Wie *et al*., 1996).

Arpigny and Jeager (1999) proposed a classification of bacterial lipolytic proteins based on the comparison of their amino acid sequences and fundamental biochemical properties, which resulted in the identification of 8 different families. Family VII bacterial lipolytic proteins share significant amino acid sequence homology (30% identity and 40% similarity) with eukaryotic carboxyl esterases (e.g. pig liver esterase) (Arpigny and Jeager, 1999). Moreover, the family is represented by esterases that have been identified from different microbial genera, from mesophilic to thermophilic organisms. The biochemical properties of esterases from this family have also been found to be diverse depending on their sources, with thermophilic properties associated with enzymes isolated from thermophilic organisms (Ewis *et al.,* 2004; Takami *et al*., 2004), and alkalophilic properties associated with enzymes from both mesophilic (Kim *et al*., 2004) and thermophilic sources (Ewis *et al.,* 2004). The increase in the availability of protein sequences has enabled us to identify by protein alignments, conserved amino acid sequences of the family which could be used to detect the presence of the Family VII carboxylesterase encoding genes from genetic materials. We report in this study the design of degenerate primers , the amplification of gene fragments encoding members of Family VII bacterial lipolytic enzymes within genetic materials from different sources and the development of an improved cassette ligation-mediated PCR for genome walking which was used to clone the complete carboxyl esterase gene from *Bacillus pumilus* .

2.2 Materials and methods

2.2.1 Materials

Oligonucleotides (Table 2.1) were purchased from Integrated DNA Technologies (USA). GFX PCR and Gel Band Purification and GFX Micro Plasmid Prep kits were purchased from (Armesham, UK). The cloning (pGemT-Easy) and expression (pET 28a) vectors were purchased from Promega and Novagen Madison, USA, respectively. Tryptone, yeast extract and agar bacteriological were purchased from Biolabs (Johannesburg, RSA). Tributyrin and Gum arabic were purchased from Sigma (Steinheim, Germany). Kanamycin and ampicilin were respectively purchased from Roche (Manheim, Germany) and Sigma (Steinheim, Germany). IPTG, X-gal, restriction and modification enzymes were purchased from MBI Fermentas (Burlington Ontario, Canada) or New England Biolabs Inc, (Hertfordshire, UK). Thermostable DNA polymerases for PCR, were purchased from Southern Cross Biotechnology (Cape Town, RSA), Promega (Madison, USA) or Roche (Mainheim, Germany).

2.2.2 Bacterial strains and culture conditions

Bacillus licheniformis DSM12369, *Bacillus licheniformis* MBB01, *Bacillus pumilus* MBB02 (the latter strains are obtainable from the University of the Free State microbial culture collection), garden soil (obtained from the garden outside the Department of Microbial, Biochemical and Food Biotechnology,

University of the Free State, South Africa) and mine borehole biofilms denoted S1 and S2 (obtained from the Merriespruit goldmine, Welkom, RSA) were used as sources of genetic material. *Escherichia coli* JM109 and JM109 (DE3) strains (Promega, Madison, USA) were used as cloning and expression hosts, respectively. The bacterial strains were grown in LB media (10g tryptone, 5g yeast extract and 5g sodium chloride in 1L of water) at 37 °C with shaking in 50-250 ml shake flasks. TLB media contained a sonicated emulsion of 10 ml trybutyrin and 10g Gum arabic in 1L of LB medium. When solid growth media was required agar (15 g/L) was added and the incubation was done at 37 °C. When required the media was supplemented with kanamycin (50 µg/ml) or ampicillin (100 µg/ml) for the selection of plasmid carrying strains.

2.2.3 DNA preparation, manipulation and transformation

Genomic DNA isolation was done according to the method described by Shyamala and Ames (1993). Recombinant techniques using commercially available molecular grade enzymes and reagents were followed as described by Sambrook *et al*., (1989). $CaCl_2$ competent *Escherichia coli* cells were prepared and transformed with plasmid DNA as described by Sambrook *et al*., (1989).

2.2.4 Primers for the detection of Family VII bacterial lipolytic genes

Selected protein sequences encoding carboxyl esterases that belong to Family VII of bacterial lipolytic enzymes (Arpigny and Jaeger, 1999) were aligned using the CLUSTAL W tool (Thompson *et al*, 1997) to identify conserved amino acid sequences. The conserved amino acid sequences were used to design degenerate primers (Table 2.1) corresponding to conserved Blocks 1, 3 and 4. The degenerate primers were used to amplify by PCR gene fragments between Blocks 1 and 4 (primer pair BuCest519F and BuCest1605R) Blocks 1 and 3 (primer pair BuCest519F and BuCest1025R) and Blocks 3 and 4 (primer pair BuCest999F and BuCest1605R). The genomic DNA from *Bacillus licheniformis* DSM12369 (control), *Bacillus pumilus* MBB02, *Bacillus licheniformis* MBB01, garden soil (G1) and goldmine borehole biofilms (S1 and S2) were used as templates during the PCR with the designed degenerate primers under the following conditions: 1 denaturation cycle (94°C, 2 min) and 30 cycles of amplification (94°C, 30 sec; 55°C, 30 sec; 72°C, 2min).

The gene fragment amplified from the biofilm (S1) DNA extracts was ligated into pGemT-Easy vector and used to transform *Escherichia coli* JM109 cells. Plasmid minipreps and restriction analysis were conducted, 50 single clones were selected and subjected to re-amplification using T7 and Sp6 universal promoter primers that flank the multiple cloning site within the pGemT-Easy vector under the following PCR conditions: 1 denaturation cycle (94°C, 2 min) and 30 cycles of amplification (94°C, 30 sec; 55°C, 30 sec; 72°C, 1 min). The

PCR products were subjected to Restriction Fragment Length Polymorphism (RFLP) using restriction enzymes *Sau*3AI and *Hae* III.

2.2.5 Cloning of the complete *Bacillus pumilus* carboxyl esterase gene

The 500 and 600 bp gene fragments obtained with *Bacillus pumilus* were used as templates for designing locus specific primers for genome walking. The cassette ligation mediated PCR principle (Figure 2.1) was used for genome walking experiments as detailed in the published report (Nthangeni *et al,* 2005, see Appendix 1). The ligation cassette was released by digesting the pLigCas plasmid with *Bam*HI and *Xba* I, followed by excision and purification of the ~200 bp from the agarose gel with the GFX PCR and gel band purification kit. To clone the downstream region of the *Bacillus pumilus* carboxyl esterase gene, the genomic DNA was digested to completion with *Spe*I, cleaned with GFX PCR and gel band purification kit and ligated to the *Bam*HI/*Xba*I restricted cassette. Locus specific primer BpCest1154F (P1, Figure 2.1) was used as the lone primer during single strand amplification PCR (SSA-PCR) with the ligation mixture as the template. The obtained SSA-PCR product was used as the template in second round PCR with cassette specific primer CSP-F2 (C3, Figure 2.1) and nested locus specific primer BpCest1398F (P2, Figure 2.1). The upstream region of *Bacillus pumilus* carboxyl esterase gene was cloned firstly by digesting to completion the genomic DNA of *Bacillus pumilus* with *Bgl*II, followed by ligation to the

*Bam*HI/*Xba*I restricted cassette. The SSA-PCR was performed with locus specific primer BpCest1876R (P4, Figure 2.1) as the lone primer while CSP-R3 (C2, Figure 2.1) and BpCest540R (P3, Figure 2.1) were used in the second round of nested PCR.

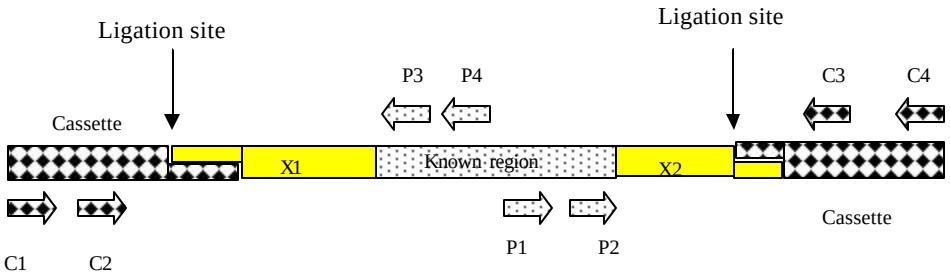


Figure 2.1: Schematic representation of the cassette ligation mediated PCR principle. The known region or loci serves as a template for the design of locus specific primers (LSP) (P1, P2, P3 and P4), which facilitate genome walking towards uknown regions of the genes (X1 and X2). The cassette binds on both ends of the gene forming direct repeats which provides binding sites for cassette specific primers (CSP). An initial round of PCR with LSP (as a lone primer) and a subsequent round of nested PCR with both LSP and CSP enhances the specificity of the method.

2.2.6 DNA sequence determination

The DNA fragments obtained by PCR amplifications were ligated into pGemT-Easy and sequenced using T7 and Sp6 promoter primers at Inqaba Biotech (Pretoria, RSA).

2.2.7 Functional expression of the lipolytic gene

The primers Bsub-Cest1F and Bsub-Cest1470R were used to amplify the open reading frame (ORF) encoding the carboxyl esterase gene with *Bacillus pumilus* genomic DNA as a template. The PCR product was digested with *Nco*I and *Sal*I and ligated into pET 28a cut with *Nco*I and *Xho*I. The ligation mixture was used to transform *Escherichia coli* host cells and selected on LB agar plates containing kanamycin. The transformed colonies were patched on TLB plates containing kanamycin and IPTG (0.02mM) and grown at 37 °C for 1-2 days. Colonies that became surrounded by zones of clearance were selected and grown in 5ml LB kanamycin followed by plasmid isolation and restriction analysis.

Table 2.1

The list of primers used in this study

Primer	Primer sequence (5'-3')	Position[2]
Cest Blocks[1]		
BuCest519F	TGGAARGGCATYCCBTAYGCVAAGCCKCCHGG	
BuCest1025R	GACTCATNWHRCCDGCBGCBGATTCSCCAAAGT	
BuCest999F	TTTGGWGAATCVGCVGCHGGYDSNATGAGC	
BuCest1605R	GCCARTCRAABCGGTACATCTABACVGG	
MBB02 Cest		
BpCest1154F	**G**GGATTAACGAGGCGCCAATTGGATAAATGC	1045
BpCest1398F	**C**GGATTCAG AGCTTCATTCTCAGGAAACGC	1283

BpCest1876R	GCC**A**CTAGTGTCCCGTAAACATACCACGGC	1895
BpCest540R	**T**GGCGTAGGGGCGGCCGCTCCACTGATGTAC	424
BsubCest1F	TGACA**G**ATCTATGAAAAATCAGCGGCGGTG	1
BsubCest1470R	CGCATCGTCGAC**T**TCTCCTTTTGAA	1470

[1] Blocks are blocks of conserved amino acid sequences as identified in Figure 2.2, MBB02 is *Bacillus pumilus* MBB02. Cest refers to carboxyl esterase genes. The F and R refer to the forward and reverse primers, respectively. [2] refers to the corresponding position of the underlined bold nucleotide of the primer within the gene sequence as submitted to the nucleotide sequence database.

2.3 Results

2.3.1 Primers for the detection of Family VII bacterial lipolytic genes

The alignment of protein sequences belonging to Family VII bacterial lipolytic enzymes resulted in the identification of four blocks showing significant amino acid sequence conservation (Figure 2.2). Degenerate primers based on the conserved Blocks were used to detect the presence of homologous genes in the genomes of *Bacillus licheniformis*, *Bacillus pumilus* and mine borehole biofilms. The largest possible DNA fragment of about 1100 bp as specified by degenerate primers derived from Blocks 1 and 4, could only be obtained with genomic DNA samples from *Bacillus licheniformis* DSM12369 (control), while *Bacillus pumilus* MBB02 and *Bacillus licheniformis* MBB01, and garden soil

and mine borehole biofilms failed to yield any PCR products (Figure 2.3 A, B, C). When Blocks 1 and 3 degenerate primers were used PCR products of about 500 bp were obtained with all the genomic DNA samples except *Bacillus licheniformis* MBB01 (Figure 2.3 A, B, C). The PCR amplification with Blocks 3 and 4 degenerate primers resulted in DNA fragments of about 600 bp with all the DNA samples and again with the exception of *Bacillus licheniformis* MBB01, garden soil and mine borehole biofilms genomic DNA (Figure 2.3 A, B, C). The identified conserved amino acid sequence Blocks are candidate templates for the designation of universal degenerate primers for PCR amplification of Family VII bacterial lipolytic genes. The failure of the conserved Blocks primer pairs to amplify any DNA fragment with *Bacillus licheniformis* MBB01 genomic DNA as the template is a possible indication that the strain lacked the homologue of Family VII carboxyl esterase genes. The 500 and 600 bp DNA fragments obtained with the genomic DNA templates of *Bacillus pumilus* were sequenced and used as templates from which locus-specific primers for genome walking into the regions flanking the carboxyl esterase were designed.

RFLP analysis of the 50 clones derived from the 500 bp fragment amplified from biofilm genetic material (S1) with *Sau*3AI and *Hae*III revealed four distinct patterns (Figure 2.4 A, B). It was observed that clones 1 and 3 of the distinct restriction patterns shared a similar *Sau*3AI pattern but displayed a different *Hae*III pattern (Figure 2.4 A, B). Similarly, clones 2 and 4 also displayed a similar restriction pattern with *Sau*3AI but a slightly different

pattern was observed with *Hae*III (Figure 2.4 A, B). Sequencing of clones 1-4, nucleotide and protein sequence analysis revealed the cloning of DNA fragments encoding Family VII carboxyl esterases (Figure 2.4. C1). Further protein sequence analysis revealed that clones 1 and 3 display 98% sequence similarity to the *Bacillus subtilis p*-nitrobenzyl esterase (Figure 2.4 C2) and clones 2 and 4 display 55% sequence similarity to an esterase from *Bacillus niacini* (Figure 2.4 C3).

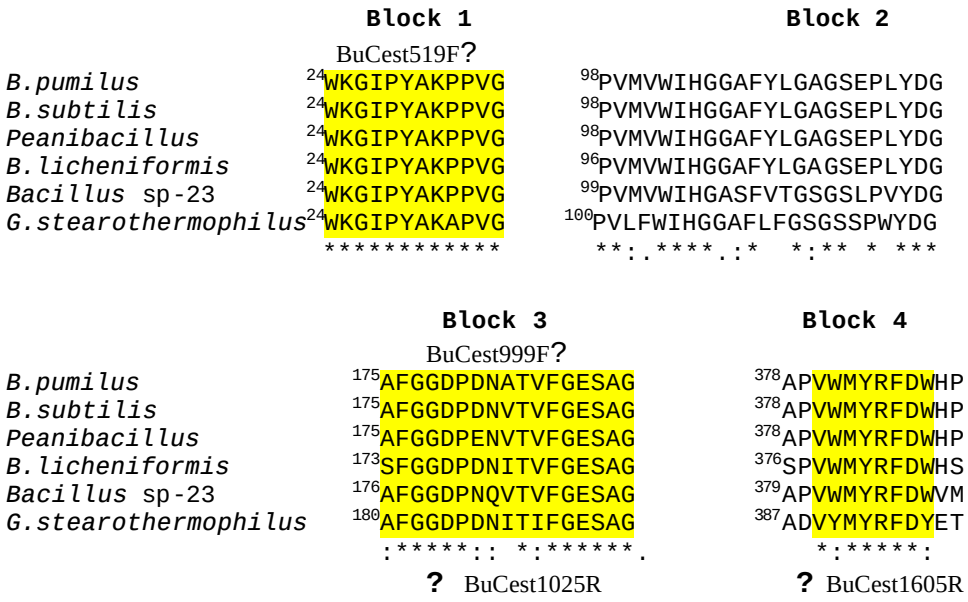


Figure 2.2: Conserved amino acid blocks of members of Family VII bacterial lipolytic proteins as deduced from nucleotide sequences submitted to the nucleotide databases. The names of the degenerate primers deduced from the sequences of the conserved amino acids are indicated above or below the block sequences. The superscripted numbers show the positions of the given

amino acids within the respective ORFs as deduced from the nucleotide databases. The arrows indicate the forward (F) and reverse (R) primers.

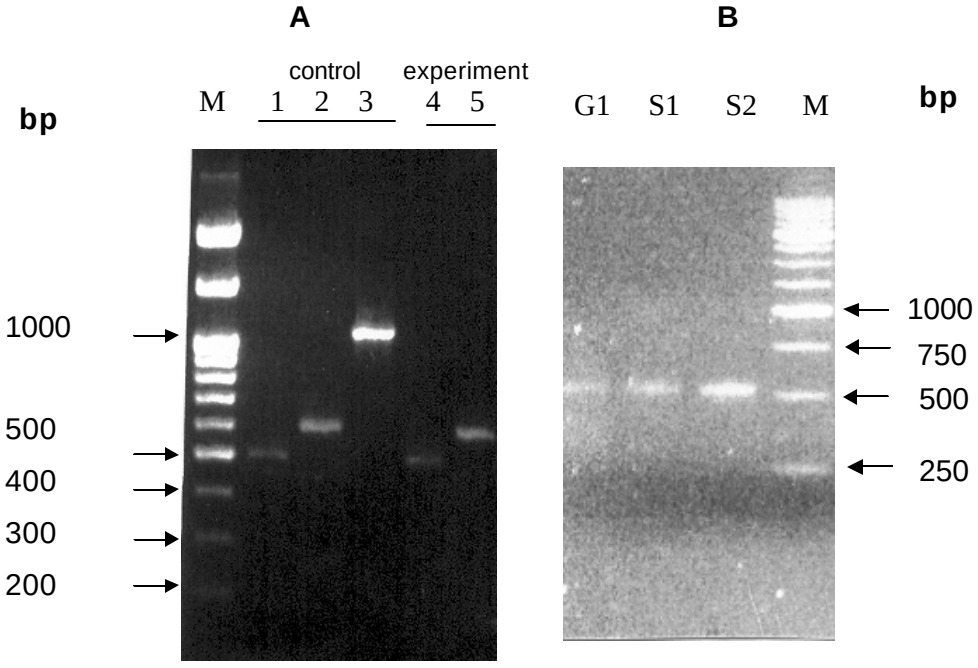


Figure 2.3: **A,** PCR with *Bacillus licheniformis* DSM 12369 (control) and *Bacillus pumilus* MBB02 (experiment) genomic DNA as templates using primer sets

BuCest519F and BuCest1025R lane (1 and 4), BuCest999F and BuCest1605R (lane 2 and 5), BuCest519F and BuCest1605R. The primer sets respectively specify the amplification of 500, 600 and 1100 bp fragments. Lane M is a 100 bp molecular weight marker. **B**, PCR with garden soil (G1), mine borehole biofilm (S1 and S2) genetic material as templates, using primer set BuCest519F and BuCest1025R. Lane M is a 250 bp molecular weight marker.

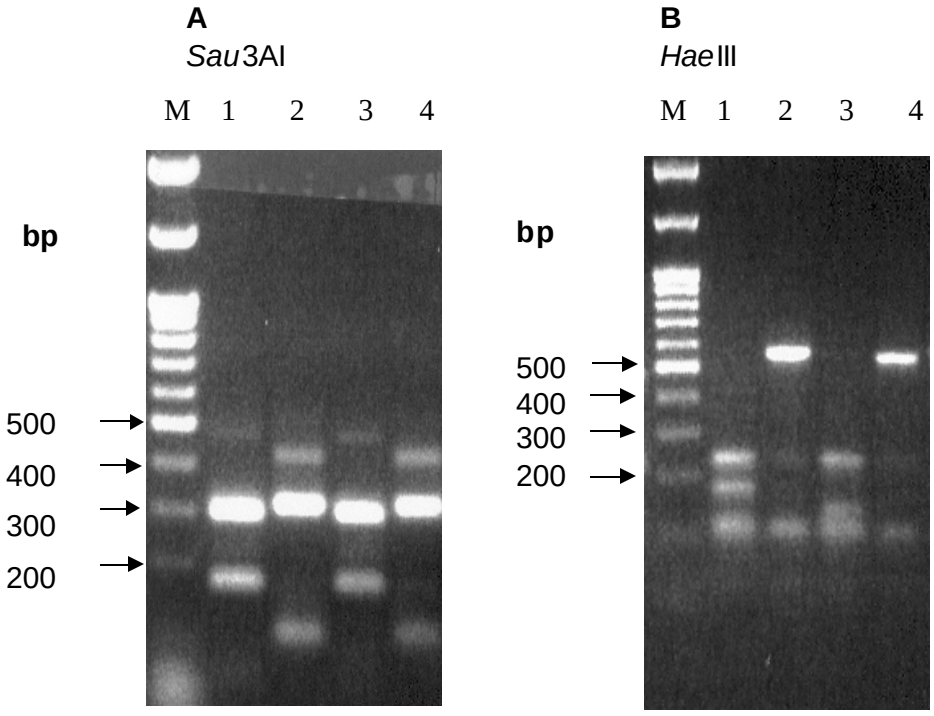


Figure 2.4: **A**, *Sau*3AI restriction patterns displayed by clones 1-4. **B**, *Hae*III restriction pattern displayed by clones 1-4. M: 100 bp molecular weight marker.

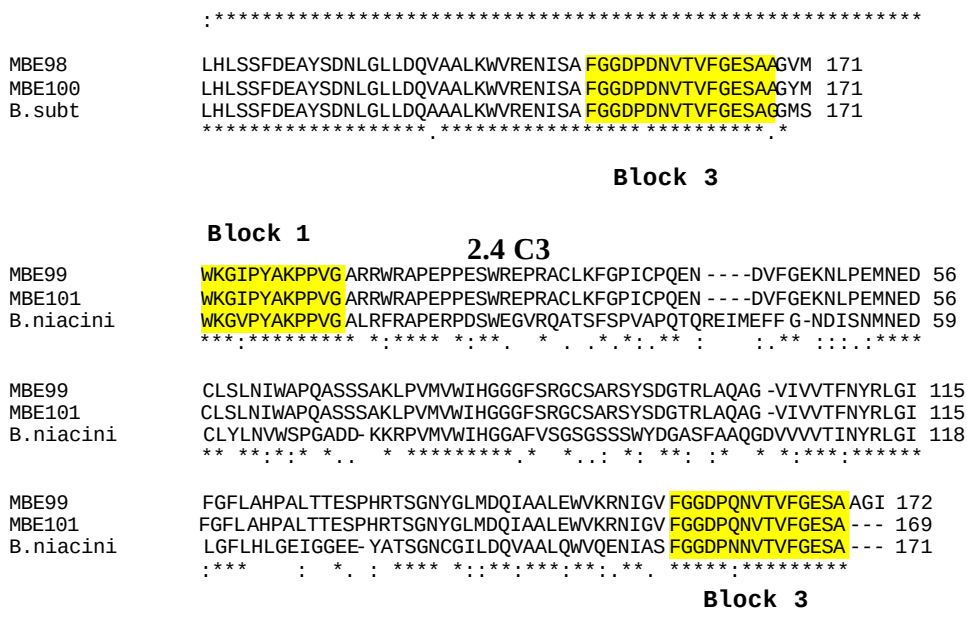


Figure 2.4: **C1**, Alignment of amino acid sequences encoding fragments of carboxyl esterases corresponding to clones 1-4 (MBE 98-101), and an indication of blocks 1 and 3 which are characteristic to Family VII bacterial lipolytic enzymes. **C2**, Alignment of clones 1,3 (MBE 98, 100) and *B. subtilis* p-NB carboxyl esterase amino acids sequences, to reveal percentage sequence similarity. **C3**, Alignment of clone 2, 4 (MBE 99, 101) and *B. niacini* carboxyl esterase amino acid sequences, to reveal percentage sequence similarity.

2.3.2 Cloning of the complete *Bacillus pumilus* carboxyl esterase gene

We have described in the accompanying paper (Appendix 1) the construction of a DNA cassette for genome walking studies. The various restriction enzyme recognition sites that flank the cassette generate sticky ends which improve the efficiency of ligation to compatibly restricted targeted DNA fragments. The availability of the various restriction sites reduces the dependence of the technique on the availability of a favourable restriction map around the known locus. Genome walking with the cassette involved an enrichment step by single strand amplification PCR (SSA-PCR) that results in the selective extension of template strand of the known locus site, through the ligation site into the cassette DNA sequence. The end result of the SSA-PCR

is the synthesis of the complementary strand of the cassette strand that becomes the template for the second round PCR. Therefore, it is only after the locus specific primer has been extended in the first round PCR (SSA-PCR) that the cassette-specific primer can find a suitable template to bind and be successfully extended in the amplification cycles of PCR.

PCR fragments of about 400 and 600 base pairs respectively corresponding to the promoter- and terminator-containing regions of *Bacillus pumilus* carboxyl esterase gene were successfully amplified during the genome walking experiments (Figure 2.5). The sequencing of the upstream and downstream DNA fragments enabled the design of specific primers at the *Bgl*II and *Spe*1 borders that specified the amplification of the 1895 bp DNA fragment (Figure 2.8 A). The nucleotide sequence was deposited in the GenBank databases under the accession number AY692083. The major open reading frame (nucleotides 334-1803) of 489 amino acid encoded a carboxyl esterase protein, which is 93% identical to the *p*-nitrobenzyl esterase from *Bacillus subtilis* (Zock *et al.*, 1994) (Figure 2.9). Upstream to the carboxyl esterase gene a partial sequence encoding the 3'-end of the putative transcriptional regulator gene was also found (data not shown).

bp M 1 2

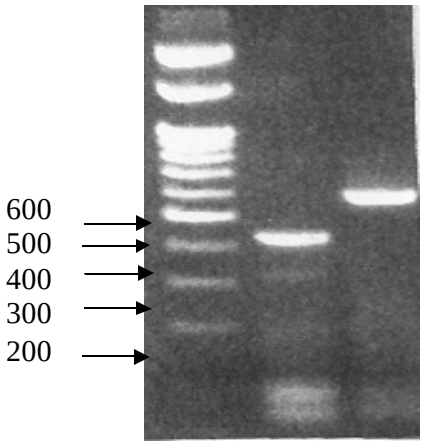


Figure 2.5: Cassette ligation mediated genome walking PCR products using the cassette/genomic DNA ligation mixture as template. Lane 1 and 2 represent represents the 400 and 600 bp fragments that respectively correspond to the up- and down-stream regions of the gene. Lane M is the 100 bp molecular weight marker.

2.3.3 Functional expression of *Bacillus pumilus* carboxyl esterase

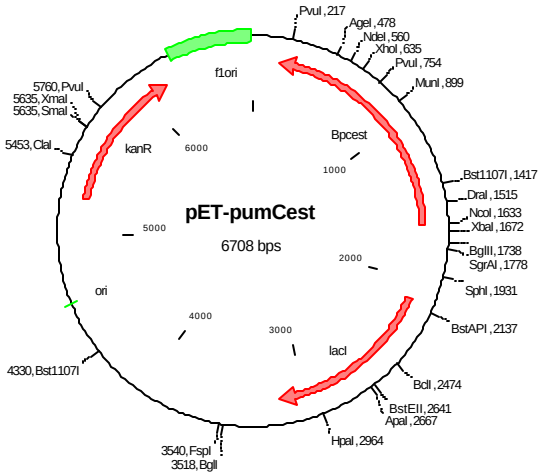
The PCR with primers Bsub-Cest1F and Bsub-Cest1470R led to the amplification of a ~1500 bp coding sequence within the *Bacillus pumilus* genomic DNA (Figure 2.6 A). Ligation of the PCR product to pET 28a resulted in the construct denoted pET-pumCest (Figure 2.6 B). The resultant transformants (carrying the appropriate recombinant plasmid) obtained upon transformation of *Escherichia coli* JM09 DE3 host cells were patched on TLB agar plates and zones of clearance could be observed after 1-2 days

incubation at 37 °C (Figure 2.7). The zones of clearances around the colonies are an indication of the hydrolysis of tributyrin by an esterase or a lipase.

A

bp 1 2

B



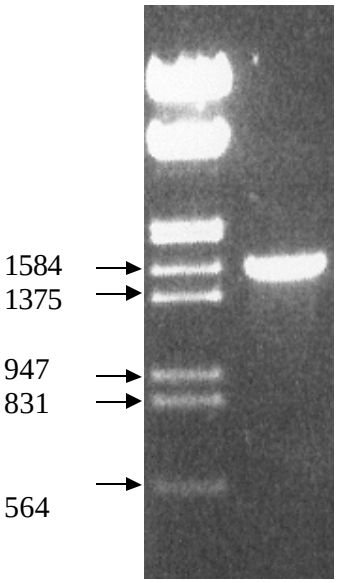


Figure 2.6: **A**, PCR amplified ~1500bp coding sequence for *Bacillus pumilus* carboxyl esterase (lane 2). Lane 1 molecular weight marker, ?DNA digested with *Eco*RI and *Hind*III (?III). **B**, Schematic representation of the expression plasmid pET-pumCest, constructed by subcloning the PCR amplified *Bacillus pumilus* carboxyl esterase gene (Bpcest) into pET 28a.

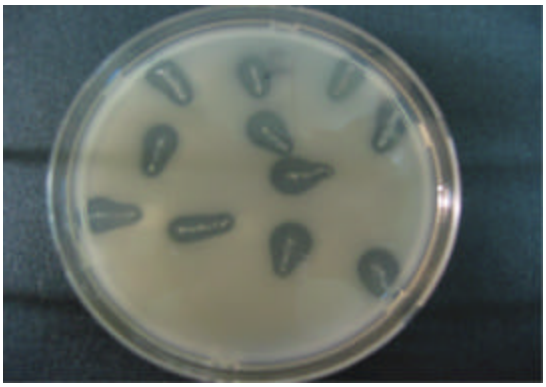


Figure 2.7: *Escherichia coli* host cells carrying pET-pumCest patched on TLB agar plates, showing zones of clearance around the growing colonies of *Escherichia coli* cells.

2.3.4 Promoter and terminator sequences analysis

Potential promoter sequences were observed at nucleotide sequences immediately upstream to putative translation start codon of the reported open reading frame (ORF) (Figure 2.8 A). The ATG start codon is separated from the putative –10 region (TATAAT) by AG-rich regions, presumably representing the ribosome-binding sequences (Figure 2.8 A). Inverted repeats could be located after the stop codon of the *Bacillus pumilus* carboxyl esterase gene (nucleotides 1830-1849, -10.7 kcals) (Figure 2.8 B). These nucleotides have the ability to form base pairs with one another and may therefore be involved in the formation of stem-loop structures that function as transcription termination signals (Figure 2.8 B).

A

```
  1 gatctatgaa aaaatcagcg gcggtgaaga agaatggcgc gtacatttag tgcaagccgt
 61 acaagccggg gtggaaaagg aagaattgtt cacttttcg aacaggctaa agaaagaaca
121 gcctgaaact gcctcttacc gcaaccgcaa actgacggaa tccaatatag aagaatggaa
181 agcgctgatg gcggaagcaa gagaactcgg tctgactgtc catgaagtca aatccttttt
241 aaaaacaatg ggaagatgat gatcggttaa gggctaacgt tcgcatctat agggtaaatc
```

-10 RBS Start codon

```
 301 tgtttataat gggggaaaaa gggagagaac acaatgactc atcaaatagt aacgactcaa
 361 tacggcaaag taaaaggcac aacggaaaac ggcgtacatc agtggaaagg catcccctac
 421 gccaagccgc ctgtcggaca atggcgtttt aaagcacctg agccgcctga agtgtgggaa
 481 gatgtccttg atgccacagc gtacggccct gtttgcccgc agccgtctga tttgctgtca
 541 ctttcgtata ctgagctgcc ccgccagtcc gaggattgct tgtttgtcaa tgtatttgcg
 601 cctgacactc caagccaaaa cctgcctgtc atggtttgga ttcacggagg cgctttttat
 661 ctcggcgcgg gcagtgagcc attgtatgac ggatcaaaac ttgcggcgca gggagaggtc
 721 attgtcgtta cactgaacta tcggctgggg ccgtttggct ttttgcactt gtcttcgttt
 781 gatgaggcgt attccgataa ccttgggctt ttagaccaag tcgccgcact gaaatgggtg
 841 cgggagaata tttcagcgtt tggcggtgat cccgataacg caacagtatt tggagaatcc
 901 gccggcggga tgagcattgc cgcgcttctc gctatgcctg cggcaaaagg cctgttccag
 961 aaagcaatca tggaaagcgg cgcttctcga acgatgacga aagaacaagc ggcgagcacc
1021 tcggcagcct ttttacaggt ccttgggatt aacgagggcc aattggataa attgcatacg
1081 gtttctgcgg aagatttgct aaaagcggcc gatcagcttc ggatcgcaga aaaagaaaat
1141 atctttcagc tgttcttcca gcccgccctt gatccaaaaa cgctgcctgc tgaaccagaa
1201 aaagcgatcg cagaaggggc cgcttccggc attccgctat taatcggaac aacccgtgat
1261 gaaggatatt tatttttcac cccggattca gacgttcatt ctcaggaaac gcttgatgca
1321 gcgctcgagt atttactagg gaagccgctg gcagagaaag ctgccgattt gtatccgcgt
1381 tctctggaaa gccaaattca tatgatgact gatttgttat tttggcgccc cgctgtcgcc
1441 tatgcatccg cacattccca ttatgcccct gtctggatgt accggttcga ttggcacccg
1501 gagaagccgc cgtacaataa aacgtttcac gcattagagc ttcctttttgt ctttggaaat
1561 ctggacggat tagaacgaat ggcaagagcg gaggttacgg atgaggcgaa acagctttct
1621 cactcgatac aatcagcatg gatcacattc gccaaaacag gaaacccaag caccgaagct
1681 gtgaattggc cggcgtatca tgaagaaacg agagagacgc tgattttaga ctcagagatt
1741 acgatcgaaa acgatcccga atctgaaaaa aggcagaagc tattcccttc aaaaggagaa
1801 taaatatggg gaaaacagga tacattggtg ctgccattgt tgtggcagct tgtatcatta
```

Stop codon **Terminator**

```
1861 ttttgtcggc cgtggtatgt ttacgggaca ctagt
```

B

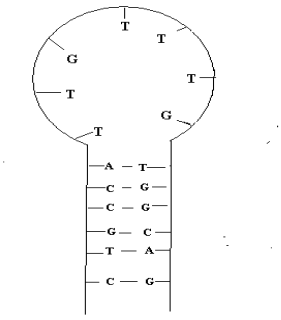


Figure 2.8: **A**, Complete nucleotide sequence of *Bacillus pumilus* carboxyl esterase gene (1895 bp). **B**, stem loop structure (-10.7 kcal), which is due to inverted nucleotide, repeats characteristic to the transcription terminator sequence. -10 refers to the -10 transcription signal, RBS is the putative ribosome binding site.

```
                                                    Block 1
B.pumilus           MTHQIVTTQYGKVKGTTENGVHQWKGIPYAKPPVGQWRFKAPEPPEVWEDVLDATAYGPV 60
B.subtilis          MTHQIVTTQYGKVKGTTENGVHKWKGIPYAKPPVGQWRFKAPEPPEVWEDVLDATAYGPI 60
Peanibacillus       MSESVVKTQYGTVKGISKNGVQTWKGIPYAKPPVGQLRFKAPDPPAAWEGVLDATAYGPV 60
B.licheniformis     MYDTTVETRFGKLKGRAENGVRIFKGVPYAKPPVGDLRFREPQRMEAWEGELDAFQFGPV 60
Bacillus            MRELQVQTKYGKVQGELLQGASVWKGIPYAKPPVGEMRFQAPTQPESWDGIRQATEFGPE 60
G.kaustophilus      MERTVVETRYGRLRGVVNGSVFVWKGIPYAKAPVGERRFLPPEPPDAWDGVREAAAFGPV 60
G.stearothermophilus MERTVVETRYGRLRGEMNEGVFVWKGIPYAKAPVGERRFLPPEPPDAWDGVREATSFGPV 60
                    *    * *::* ::*    ..  :**:****.***: **  *     *:.  :*  :**

B.pumilus           CPQPSD--LLSLSYTELPRQSEDCLFVNVFAPDTPSQN-LPVMVWIHGGAFYLGAGSEPL 117
B.subtilis          CPQPSD--LLSLSYTELPRQSEDCLYVNVFAPDTPSQN-LPVMVWIHGGAFYLGAGSEPL 117
Peanibacillus       CPQPPD--LLSYSYPELPRQSEDCLYVNVFAPDTPGKN-RPVMVWIHGGTFYLGAGSEPL 117
```

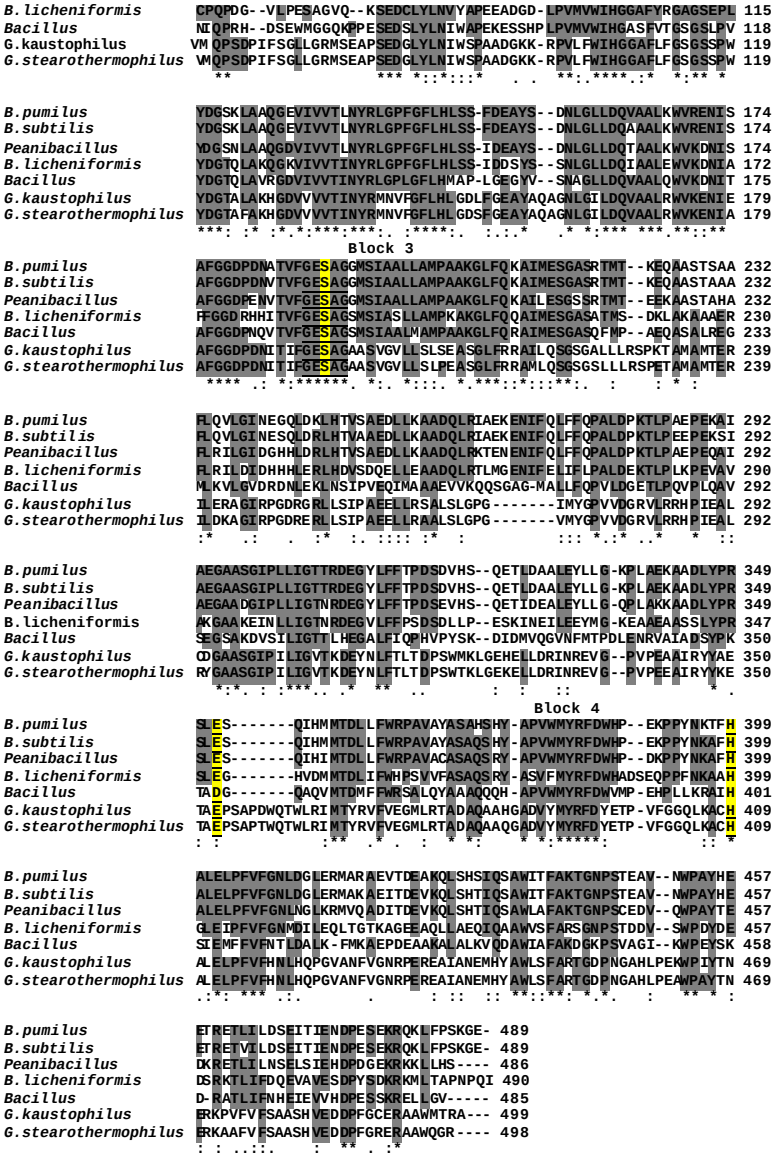


Figure 2.9: Alignment of the *Bacillus* carboxyl esterase amino acid sequences with amino acid sequences of selected Family VII lipolytic enzymes to reveal percentage sequence similarity. Blocks 1, 3 and 4 (shaded) represent conserved sequence blocks that are a characteristic of Family VII lipolytic enzymes. The consensus sequence G-X-S-X-G, and the amino acids proposed as the catalytic apparatus (S ^{187}, E/D ^{353} and H ^{399}) are underlined.

2.4 Discussion

Alignment of selected protein sequences encoding Family VII bacterial lipolytic enzymes enabled the design of degenerate primers that could be

used to detect members of Family VII lipolytic genes within bacterial genomes. We used the degenerate primers to amplify partial gene sequences encoding homologues of Family VII lipolytic enzymes in *Bacillus* genomes and genetic materials isolated from soil and mine water samples (biofilms). Nucleotide and protein sequence analysis of clones derived from the gene fragments amplified from the biofilm genetic material revealed the cloning of diverse gene fragments encoding Family VII carboxyl esterases. The fact that the degenerate primers amplified a gene fragment displaying only 55% sequence similarity to *Bacillus niacini* esterase (Kim *et al*., 2004), suggested the cloning of a novel Family VII carboxyl esterase homologue. From the partial gene sequences obtained with *Bacillus pumilus* genomic DNA, we successfully used the cassette ligation-mediated PCR to amplify a complete gene sequence (together with its promoter and terminator regions) encoding a carboxyl esterase from *Bacillus pumilus*, which was previously not known to posses such a gene. *Escherichia coli* host cells harbouring the complete carboxyl esterase gene subcloned into pET 28a became surrounded by zones of clearance when plated on TLB agar plates, an indication that a complete functional gene from *Bacillus pumilus* was amplified. Comparison of the derived amino acid sequence of *Bacillus pumilus* carboxyl esterase with those from the nucleotide database showed that it is strikingly identical (93%) to the industrially important *p*-nitrobenzyl esterase from *Bacillus subtilis* (Moore and Arnold, 1996), 82% identical to an esterase from *Peanibacillus* sp. BP-7 (Prim *et al*., 2001), 49 and 39 % identical to carboxyl esterases from *Bacillus* sp BP-23 (Prim *et al*., 2000) and *Geobacillus stearothermophilus* (Ewis *et al*., 2004), respectively. This suggested that *Bacillus pumilus* carboxyl esterase is a

member of the bacterial subclass of type B carboxyl esterases (Prim *et al*., 2000), which are a group of evolutionary related lipolytic proteins belonging to Family VII class of bacterial lipolytic proteins (Arpigny and Jeager, 1999). The fact that *Bacillus pumilus* carboxyl esterase exhibits high sequence similarity to *Bacillus subtilis p*-nitrobenzyl esterase also suggests that the coding genes of both esterases are evolutionarily conserved amongst close *Bacillus* strains (López *et al*., 1998). The motif GES^{187}AG, corresponding to the consensus sequence –G-X-S-X-G- for the active site of most serine esterases, and the highly conserved sequence around the active site of the type B carboxyl esterase family (Wang and Hartsuck, 1993) were found in *Bacillus pumilus* carboxyl esterase (Figure 2.9). The catalytic apparatus of type B carboxyl esterases involving the triad serine, glutamate or aspartate, and histidine (Wang and Hartsuck, 1993) was predicted for *Bacillus pumilus* carboxyl esterase to be 187 (S), 353 (E/D) and 399 (H) (Figure 2.9) based on protein alignment comparisons. Further sequence analysis suggested that carboxyl esterase from *Bacillus pumilus* is destined to the intracellular as no amino terminal extension corresponding to the signal peptide (which is responsible for translocation of proteins to the extracellular) could not be found. The putative TATAAT region could be identified within the promoter region of the gene but the -35 region could not be predicted with ease (Figure 2.8). The putative –10 (TATAAT) region of *Bacillus pumilus* carboxyl esterase is identical to the consensus –10-promoter region recognised by vegetative-cell sigma factors, d^A (Handelwang, 1995). This suggests that, the *Bacillus pumilus* carboxyl esterase is constitutively expressed.

We observed upon alignment of amino acid sequences encoding the *Bacillus pumilus* carboxyl esterases and selected members of Family VII of bacterial lipolytic enzymes that the N-terminal domain is relatively conserved as compared to the C-terminal domain (Figure 2.9 A). The catalytic triad is interestingly located in the variable C-terminus (using the catalytic Ser as the central amino acid) tempting to suggest that the C-termini regions play catalytic roles and are responsible for the different biochemical properties displayed by Family VII carboxyl esterase enzymes.

A significant number of proteins are currently being classified into families based on conserved amino acid sequences, creating an opportunity to design degenerate primers that could be used to screen and clone partial gene fragments from genetic materials from different sources. This has been demonstrated in this study, where partial gene fragments encoding Family VII carboxyl esterases were amplified from different sources using degenerate primers designed based on conserved sequence blocks that are characteristic to the family. Subsequent cloning of a complete and functional Family VII carboxyl esterase gene from the genome of *Bacillus pumilus* by cassette ligation-mediated PCR suggested that, a combination of protein similarity alignments which aid in the amplification of gene fragments from genetic materials and the cassette ligation-mediated genome walking principle could be useful tools in the cloning of complete genes in PCR-based bioprospecting studies for novel genes.

The contents of this chapter have been partly published: Nthangeni MB, Ramagoma F, **Tlou MG**, Litthauer D (2005). Development of a versatile cassette for directional genome walking using cassette ligation-mediated PCR and its application in the cloning of complete lipolytic genes from *Bacillus* species. *J. Microbiol. Methods* **61**, 225-234. **See Appendix 1.**

CHAPTER 3

Molecular cloning, over-expression, structure - function relationships of native and chimeric Family VII carboxyl esterases from *Bacillus* species

3.1 Introduction

The application of enzymes in organic synthesis is now a routine alternative for the organic chemist and process engineer. This is because, native or engineered enzymes catalyse reactions under mild conditions of temperature and pressure with selectivity (substrate, regio- or enantioselectivity) and reduced waste (Ebbing, 1996). However, prior to the application of any enzyme at industrial level, properties that makes it suitable for a particular application should be determined. To achieve this, high level production/over-expression of proteins (usually in the recombinant form) as a prerequisite for subsequent purifications, biochemical characterization, and three dimensional structure analysis has become standard techniques (Schumann and Ferreira, 2004).

The first step to the production of recombinant proteins involves cloning of the appropriate gene into an expression vector under the control of an inducible promoter. A variety of expression systems designed for various applications and compatibilities are available [e.g. bacterial, yeast, insect and mammalian expression systems] (Marino, 1989). However, approximately 80% of the

proteins used to solve three-dimensional structures submitted to the protein database (PDB) in 2003 were prepared in *Escherichia coli* expression systems (Sørensen and Mortessen, 2005). The T7 based pET expression system (commercialised by Novagen), is by far the most used in recombinant protein preparation (Sørensen and Mortessen, 2005). Despite the extensive use of *Escherichia coli* expression systems, there is no guarantee that every gene can be expressed efficiently in this Gram negative bacterium. Factors influencing the expression level include unique and subtle structural features on the gene sequence, the stability and efficiency of mRNA, correct and efficient protein folding, codon usage, and degradation of recombinant protein by proteases and toxicity of the protein (Baneyx, 1999, Schumann and Ferreira, 2004; Sørensen and Mortessen, 2005).

It has been reported that lipases from different *Pseudomonas* and *Burkholderia* species which are used for a variety of biotransformations cannot be over-produced using *Escherichia coli* expression systems (Jeager and Reetz, 1998). This is due to the fact that lipases from *Pseudomonas* species require the functional assistance of 30 different cellular proteins before they can be recovered from the culture supernatant in an enzymatically active state, indicating that folding and secretion are highly specific processes that normally do not function properly in heterologous hosts (Rosenau and Jeager, 2000). There is however, a large body of literature on the successful functional and over-expression of lipases and esterases of various *Bacillus* species in *Escherichia coli* using conventional over-expression systems (Nthangeni *et al*., 2001; Eggert *et al*., 2000; Ewis *et al*., 2004; Kim *et al*.,

2004). In this chapter we describe the over-expression and chimera construction aided determination of the structure-function relationships of Family VII carboxyl esterases.

Chimera construction or domain shuffling studies aimed at relating the structural to the functional properties of Family VII carboxyl esterases were prompted by an observation upon alignment and analysis of protein sequences encoding selected members of Family VII carboxyl esterases that the N-terminal regions (with the catalytic Ser embedded within the GE**S**AG motif as the reference point) of these enzymes were relatively conserved while the C-terminal regions were found to be variable in terms of amino acid conservation (Chapter 2). We therefore, hypothesised that the relatively conserved N-termini were responsible for the general configuration of the structure of the enzymes while the variable C-termini confer biochemical specificities. To test this hypothesis, hybrid carboxyl esterases were constructed by exchanging domains of *B. pumilus* and *B. licheniformis* carboxyl esterases. The ability to engineer hybrid enzymes, which contain elements from two or more enzymes, has been shown by other researchers to be useful in generating proteins with new properties by alteration of non-enzymatic and enzymatic properties (Villbrandt *et al*., 2000) and as tools for understanding structure-function relationships (Brock and Waterman, 2000).

3.2 Materials and methods

3.2.1 Materials

Oligonucleotides (Table 3.1) were purchased from Integrated DNA Technologies (USA). GFX PCR and Gel Band Purification and GFX Micro Plasmid Prep kits were purchased from Armesham (UK). The cloning (pGemT-Easy) and expression (pET 28a) vectors were purchased from Promega and Novagen (Madison, USA), respectively. Tryptone, yeast extract and bacteriological agar were purchased from Biolabs (Johannesburg, RSA). Tributyrin and Gum arabic were purchased from Sigma (Steinheim, Germany). Kanamycin and ampicillin were respectively purchased from Roche (Manheim, Germany) and Sigma (Steinheim, Germany). IPTG , X-gal, restriction and modification enzymes were purchased from MBI Fermentas (Burlington Ontario, Canada), Promega, Madison, USA) and Roche (Maiheim, Germany). Super-Therm and Pfu DNA polymerases for PCR were respectively purchased from Southern Cross Biotechnology (Cape Town, RSA) and MBI Fermentas (Burlington Ontario, Canada).

3.2.2 Bacterial strains and culture conditions

Bacillus licheniformis DSM12369 and *B. pumilus* MBB01 (obtainable from the University of the Free State microbial culture collection) were used as sources of genomic DNA. *Escherichia coli* JM109 and JM109 (DE3) strains (Promega, Madison, USA) were used as cloning and expression hosts, respectively. The bacterial strains were grown in LB media (10g tryptone, 5g yeast extract and 5g sodium chloride in 1L of water) at 37 °C with shaking in 50-250 ml shake flasks. TLB media contained a sonicated emulsion of 10 ml tributyrin and 10g

Gum arabic in 1L of LB medium. When solid growth media was required agar (15 g/L) was added and the incubation was done at 37 °C. When required the media was supplemented with kanamycin (50 µg/ml) or ampicillin (100 µg/ml) for the selection of plasmid carrying strains.

3.2.3 DNA preparation, manipulation and transformation

Genomic DNA isolation was done according to the method described by Shyamala and Ames (1993). Recombinant techniques using commercially available molecular grade enzymes and reagents were followed as described by Sambrook *et al*., (1989). $CaCl_2$ competent *Escherichia coli* cells were prepared and transformed with plasmid DNA as described by Sambrook *et al*., (1989).

3.2.4 Functional expression of the *B. pumilus* and *B. licheniformis* carboxyl esterases

Primer sets BpCEST1F, BpCEST1470R and BlCEST1F, BlCEST1452R (Table 3.1) were used to amplify the open reading frames (ORFs) encoding the carboxyl esterases from *B. pumilus* and *B. licheniformis* genomic DNA, respectively (see Appendix 1, for the cloning of the complete *B. licheniformis* carboxyl esterase gene). The PCR products were digested with *Nco*I and *Sal*I and ligated to pET 28a cut with *Nco*I and *Xho*I resulting in constructs pET-pumCEST and pET-lichCEST. The constructed primers enabled the

subcloning of carboxyl esterase ORFs as recombinant proteins containing a 6X His tag at the C-termini. The ligation mixtures were respectively used in the transformation of *Escherichia coli* host cells and selected on LB agar plates containing kanamycin. The transformed colonies were patched on TLB plates containing kanamycin and IPTG (0.02mM) and grown at 37°C for 1-2 days. Colonies that became surrounded by zones of clearance were selected and grown in 5ml LB kanamycin followed by plasmid isolation and restriction analysis.

3.2.4 Construction of chimeric Family VII carboxyl esterases

Protein sequences encoding the *B. pumilus* carboxyl esterase and *B. licheniformis* carboxyl esterase were aligned (Figure 3.1 A) using the ClustaI W tool (Thompson *et al*, 1997), the conserved region around the characteristic Gly-X-Ser-X-Gly motif (Figure 3.1 A) was identified as the region of overlap. Overlap primers BpCEST565F, BlCEST565R (Figure 3.1 B) and BlCEST565R, BpCEST565R (Figure 3.1 C) were designed based on the conserved region. Primer sets T7 promoter (A, Figure 3.2) , BlCest565R (B, Figure 3.2) and BpCest565F (C, Figure 3.2), BpCEST1470R (D, Figure 3.2) were respectively used in the amplification of the *B. licheniformis* carboxyl esterase (BlCEST) N-terminus and *B. pumilus* (BpCEST) C-terminus using constructs pET-lichCEST and pET-pumCEST as templates, under the following PCR conditions: 1 denaturation cycle (94°C, 2 min) and 30 cycles of amplification (94°C, 30 sec; 55°C, 30 sec; 72°C, 2min). The amplified BlCEST

N-terminus and BpCEST C-terminus were mixed and cleaned with the GFX PCR and Gel Band Purification kit. The purified PCR products were used as templates in an overlap PCR using T7 promoter (A, Figure 3.2) and BpCEST1470R (D, Figure 3.2). The PCR product (Hybrid 1) was digested with *Nco*I and *Sal*I and ligated into pET 28a cut with *Nco*I and *Xho*I resulting in construct pET-hybrid1CEST. Hybrid 2 was constructed by amplifying the BpCEST N-terminus and the BlCEST C-terminus with the aid of primer sets T7 promoter, BpCEST565R and BlCEST565F, BlCEST1452R respectively, using constructs pET-pumCEST and pET-lichCEST as tempates. The amplified BpCEST N-terminus and the BlCEST C-terminus were mixed and cleaned with the GFX PCR and Gel Band purification kit. The purified PCR products were used as a template in an overlap PCR with T7 promoter and BlCEST1452R. The amplified Hybrid 2 was digested with *Nco*I and *Sal*I and ligated to pET 28a cut with *Nco*I and *Xho*I resulting in expression plasmid pET-Hybrid2CEST. The two hybrid carboxyl esterase ORFs were also subcloned into pET 28a such that they contain a 6X His tag at the C-termini.

The constructs pET-Hybrid1 and pET-Hybrid2 were respectively used to transform *Escherichia coli* host cells. The transformants were selected on LB agar plates supplemented with kanamycin. The transformned colonies were patched on TLB media containing kanamycin and IPTG (0.02mM) and incubated at 37°C for 1-2 days. Colonies that became surrounded by zones of clearance were selected and grown in 5ml LB kanamycin followed by plasmid isolation and restriction analysis. The selected clones were subjected to sequencing using T7 promoter and T7 terminator primers.

A

```
B.lich          MSGLTVKTRYGALKGTMQNGVRVWKGIPYAKPPVGKWRFKAPQETDAWEGVRDATQFGSI 60
B.pum           MTHQIVTTQYGKVKGTTENGVHQWKGIPYAKPPVGQWRFKAPEPPEVWEDVLDATAYGPV 60
                *:   *.*:** :*** :***: ***********:******: .:.**.* *** :*.:

B.lich          CPQPEGILFQ--LERVEKSEDCLCLNVFAPQSSGENRPVMVWIHGGAFYLGAGSEPLYDG 118
B.pum           CPQPSDLLSLSYTELPRQSEDCLFVNVFAPDTPSQNLPVMVWIHGGAFYLGAGSEPLYDG 120
                ****..:*     *  .:***** :*****::..:* **********************

B.lich          SHLAADGDVIVATINYRLGPFGFLHLSSVNQSYSNNLGLLDQIAALKWVKENISSFGGDP 178
B.pum           SKLAAQGEVIVVTLNYRLGPFGFLHLSSFDEAYSDNLGLLDQVAALKWVRENISAFGGDP 180
                *:***:*:***.*:**************.:::**:*******:******:****:*****
```

Region of overlap

```
B.lich          DNITVFGESAGSMSIASLLAMPDAKGLFQKAIMQSGASETMPKEKAETAAETFLHILNID 238
B.pum           DNATVFGESAGGMSIAALLAMPAAKGLFQKAIMESGASRTMTKEQAASTSAAFLQVLGIN 240
                ** ********.****:***** **********:****.**.**:* ::: :**::*.*:

B.lich          PDHSEQLHDVSAKELLEAADELRDVMGENIFQLLFLPVVDRETLPLEPVTAVAQGAADDI 298
B.pum           EGQLDKLHTVSAEDLLKAADQLRIAEKENIFQLFFQPALDPKTLPAEPEKAIAEGAASGI 300
                 .: ::** ***::**:***:** .  ******:* *.:* :*** ** .*:*:***..*

B.lich          KLLIGTNRDEGVLFFTPESELLPEQKKAEILREHVGGELAKTAAELYPGSLEGQINMMTD 358
B.pum           PLLIGTTRDEGYLFFTPDSDVHSQETLDAALEYLLGKPLAEKAADLYPRSLESQIHMMTD 360
                 *****.**** *****:*:: .::.    *.  :*  **:.**:*** ***.**:****

B.lich          ILFWRPAVAFAAGQSAHSPVWMYRFDWHSEHPPFHKAAHGLDIPFVFGNMDALDMITNTK 418
B.pum           LLFWRPAVAYASAHSHYAPVWMYRFDWHPEKPPYNKTFHALELPFVFGNLDGLERMARAE 420
                :********:*:.:* ::**********.*:**::*: *.*::******:*.*: ::.::

B.lich          ASEETKQLSQHIPGLPGFHLHIREVRPLKPSAGRTMIRTHEKRSFS-NTTILIEEDPDAE 477
B.pum           VTDEAKQLSHSIQ--SAWITFAKTGNPSTEAVNWPAYHEETRETLILDSEITIENDPESE 478
                .::*:****: * . ..:  . :  .* . :.. .  : . :.::  :: * **:**::*

B.lich          KRKKLKI---- 484
B.pum           KRQKLFPSKGE 489
                **:**
```

B

```
BpCEST565F      ----------------ACAGTATTTGGAGAATCCGCCGGCGGAATGAGCATTGC 38
BlCEST565R      GGATCCTGACAACATCACGGTT TTTGGAGAATCGGCCGG--------------- 39
                          ** ** ********** *****
```

C

```
BlCEST565F      ----------------ACGGTTTTTGGAGAATCGGCCGGTTCGATGAGCATCGCC 39
BpCEST565R      TGATCCCGATAACGTAACAGTATTTGGAGAATCCGCCGG---------------- 39
                          ** ** ********** *****
```

Figure 3.1: Alignment of the BpCEST and BlCEST amino acid sequences, to compare the degree of sequence conservation of the N-terminus and C-terminus of the respective proteins. The underlined, bolded and shaded block of amino acids represents the region of overlap which served as a reference point when defining the N and the C-terminal domains (A). Alignment of the nucleotide sequences of overlap primers BpCEST565F and BlCEST565R which were used for the construction of hybrid1 (B). Alignment of nucleotide sequences of overlap primers BlCEST565F and BpCEST565R used in the construction of hybrid 2 (C).

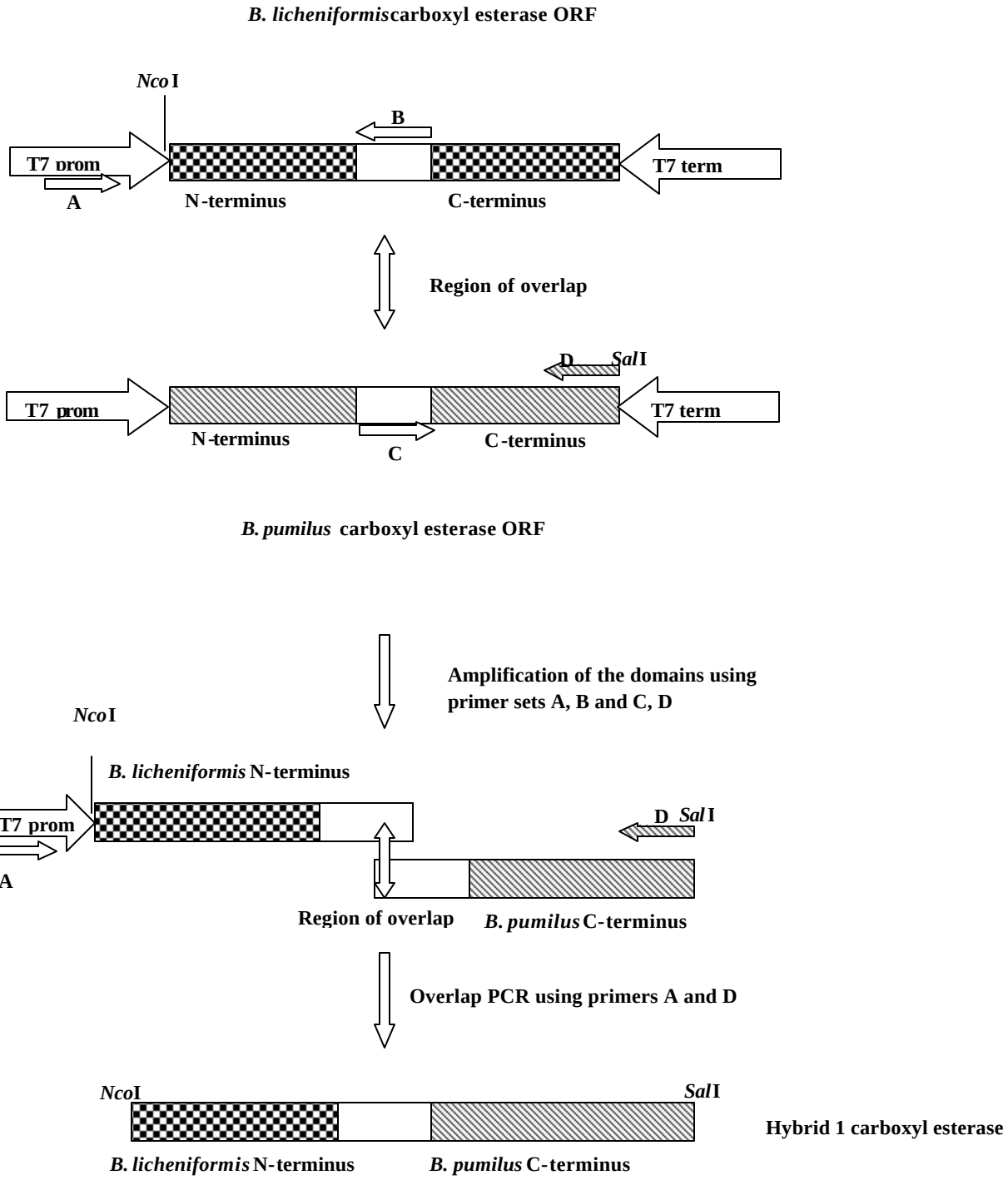


Figure 3.2: Schematic representation of the construction of the hybrid carboxyl esterases by overlap PCR aided domain exchange. Arrows B and C represent the overlap primers.

Table 3.1 List of primers used in this study

Gene[1] Primer	Primer sequence(5'-3')	Position[2]
MBB02 CEST		
BpCEST1F	GACTCG***Nco*I** **CC<u>A</u>TGG**CTCATCAAATAGTAACGA	1
BpCEST1470R	CGCATC ***Sal*I** **GTCGAC<u>T</u>**TCTCCTGAAGGGAATAGC	1470
BpCEST565F	**<u>A</u>**CAGTATTTGGAGAATCCGCCGGCGGAATGA	548
BpCEST565R	**<u>C</u>**CGGCGGATTCTCCAAATACTGTTACGTTATC	572
DSM12369 CEST		
BlCEST1F	GTATACCTA***Nco*I** **CC<u>A</u>TGG**CTGGTCTCACAGTGA	1
BlCEST1452R	GGTCAA ***Sal*I** **GTCGAC<u>G</u>**ATCTTCAGTTTTTTTCTC	1452
BlCEST565F	**<u>A</u>**CAGTTTTTGGAGAATCGGCCGGTTCGATG	544
BlCEST565R	**<u>C</u>**CGGCCGATTCTCCAAAAACCGTGATGTTG	565

[1]MBB02 is *B. pumilus* MBB02., DSM12369 is *B. licheniformis* DSM12369. CEST refers to carboxyl esterase genes. The F and R refer to the forward and reverse primers, respectively. [2] The corresponding position of the underlined bold nucleotide of the primer within the gene sequence as submitted to the nucleotide sequence database.

3.2.5 Over-expression of the carboxyl esterases

Expression levels and localisation of the *B. pumilus*, *B. lichenifromis*, hybrid 1 and hybrid 2 carboxyl esterases were monitored by inoculating 100 ml LB media contained in 250 ml shake flasks with 1 ml overnight culture corresponding to clones harbouring each carboxyl esterase construct (pET-pumCEST, pET-lichCEST, pET-hybrid1CEST or pET-hybrid2CEST). The cultures were incubated in a 30°C shaker incubator and cell growth monitored until the optical density reading (OD_{600nm}) of ~0.7 was reached. Enzyme production was induced by the addition of IPTG to each flask to a final concentration of 1 mM. Samples (500 µl) were collected at 2 hour intervals for 10 hours. The cells were separated from the supernatant (extracellular fraction) by centrifugation at 14 000 *Xg* for 2 min and washed 2X in 50 mM Tris-Cl buffer (500 µl), pH 8.0, before being resuspended in 500 µl, 50 mM Tris-Cl buffer, pH 8.0, supplemented with 1 mg/ml lysozyme and left on ice for 20 min. This was followed by sonication of the cells using a Branson Cell disruptor B30 (SmithKline Company) at 60% work outputs for 4 repeats of 10 sec sonication and 10 sec resting periods. The cell debris and membrane fractions were separated from the periplasmic and cytoplasmic proteins (cell free extract) by centrifugation of the samples at 14 000 X *g* for 10 min. The extracellular fractions and the cell free extracts were assayed for esterase activity using *p*-nitrophenyl ester (*p*NPE) assay.

3.2.6 Esterase enzyme assay/*p*-NPE assay

The assay method is based on the method used by Winkler and Stuckmann (1979) and determines both lipase and esterase activity. An artificial substrate was used to assay for esterase activity, namely *p*-nitrophenyl butyrate. (*p*NPB). This assay measures the release of *p*-nitrophenol (*p*-NP) from *p*-NPB. A stock solution (150 mM *p*-NPB) was prepared by mixing 26 µl *p*NPB with propan-2-ol to a final volume of 500 µl. Tris-Cl buffer, 50 mM, pH 8 was used as the assay buffer. The assay mixture was prepared by mixing 1 µl of the *p*NPB stock solution with the assay buffer to a final volume of 600 µl. The enzyme solution (50 µl) was added to 600 µl of the assay mixture in a 0.5 ml plastic cuvette. The liberation of *p*-NP was spectrophotomerically measured at a wavelength of 410 nm at 30 °C over a period of 10 min in a Beckman DU 650 spectrophotometer fitted with an Auto 6-sampler (water regulated) connected to a Haake D1-L heating bath / circulator (Fisons). The enzyme activity was calculated using the kinetic software supplied by Beckman. Enzyme activity, U/ml (µmole fatty acid released min^{-1} ml^{-1} enzyme), was calculated using the following formula:

$$U / ml = \frac{V}{v \times E \times d} \times \frac{A_{410}}{t}$$

V = Substrate volume (ml)

v = Enzyme volume (ml)

E = Extinction coefficient of *p*-NP at 410 nm (= 15 ml µmole^{-1} cm^{-1}) (Vorderwülbecke *et al.*, 1992)

t = Incubation time of assay (min)

d = Light path of cuvette (1 cm)

A_{440} = Absorbance wavelength

3.2.8 Protein determination

Protein concentrations were estimated by using the bicinchoninic acid (BCA) (Smith *et al.*, 1985) which was supplied as a kit by Pierce, (Rockford, IL, USA). A set of protein standards was prepared with bovine serum albumin (BSA), provided by Pierce as part of the BCA protein assay kit in the range of 25-2000 ug/ml. A standard curve was prepared and used to determine the concentrations of the unknown protein samples (Figure 3.3).

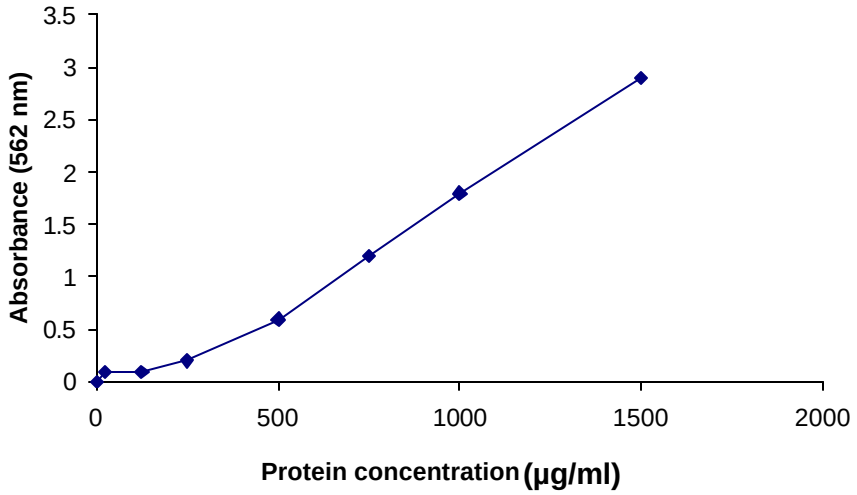


Figure 3.3: Standard curve for the Micro BCA protein assay with BSA as a protein standard. Averages of triplicate determinations are shown.

3.2.9 Enzyme purification

Cultures (100 ml, OD_{600nm} ~ 0.7) harbouring pET-pumCEST, pET-lichCEST, pET-Hybrid1CEST, and pET-Hybrid2CEST were induced for 3 hours with 1 mM IPTG and cells harvested by centrifugation at 14000 X *g* for 15 minutes at 4 °C. The cells were resuspended in 10 ml 50 mM sodium dihydrogen orthophosphate, 20 mM imidazole, pH 8.0, and lysed as described in section 3.2.6. The cytoplasmic proteins were separated from the cell debris and membrane fractions by centrifugation at 14000 *Xg* for 30 min at 4 °C. The cell free extracts (4 ml) were loaded onto the Protino® Ni 2000 prepacked columns (MACHERY-NAGEL, Germany) for the purification of polyhistidine-tagged proteins. The columns were washed 3X with 4 ml, 50mM sodium dihydrogen orthophosphate, 20mM imidazole, pH 8.0. The protein were eluted using 50 mM sodium dihydrogen orthophosphate, 250 mM imidazole, pH 8.0, and three, 3 ml fractions collected and assayed for carboxyl esterase activity (data not shown). From the fractions displaying maximum carboxyl esterase activity, 50 µl (~100 µg of protein) was collected and analysed using SDS-PAGE.

3.2.10 SDS-PAGE

Electrophoresis in 10% polyacrylamide resolving and 4% stacking gels in the presence of the anionic detergent SDS was used to monitor the purification process, to assess the homogeneity of the purified fractions and to estimate the relative molecular mass of the enzymes by comparing the electrophoretic mobility with those of standard proteins of known molecular masses.

The SDS-PAGE was performed according to Laemmli (1970). Protein bands were detected with silver staining (Switzer *et al.*, 1979).

3.2.11 Optimum pH and temperature

The optimum pH of the carboxyl esterases was determined over a range of pH 6-10. The pH range was constructed by adjusting the pH of the assay buffer cocktail used, consisting of an equimolar mixture sodium dihydrogen orthophosphate, Tris and glycine (50 mM). The enzyme was added and the reaction was allowed to proceed for 10 min at 30 °C. The optimum temperature for all the carboxyl esterases was determined over a range of 30-50 °C. In each case the assay buffer (50 mM Tris-Cl) was equilibrated at required temperature before addition of the enzyme.

3.2.12 Homology modeling of the carboxyl esterases

Homology modelling was performed with the *Bacillus subtilis* *p*-nitrobenzyl esterase (pNBE) (PDB accession number 1C7J) as a template using WHATIF (Vriend, 1990) and the models were optimised by stimulated annealing using YASARA software (Di Nola *et al*., 1991).

3.3 Results

3.3.1 Functional expression of the *B. pumilus* and *B. licheniformis* carboxyl esterases

Carboxyl esterase ORFs of ~1500 bp were successfully amplified from the genomes of *B. pumilus* and *B. licheniformis* (Figure 3.4 A). Ligation of the amplified ORFs into pET 28a respectively resulted in constructs pET-pumCEST (Figure 3.4 B) and pET-lichCEST (Figure 3.4 C) respectively corresponding to ORFs of *B. pumilus* and *B. licheniformis* carboxyl esterases. The carboxyl esterase ORFs were subcloned into pET 28a, such that the recombinant proteins contain a 6X His-tag at the C-terminal domain. The constructs were used to transform *Escherichia coli* (JM109 DE3) host cells. The resultant transformants (carrying the correct recombinant plasmids) were replica plated on TLB agar plates and zones of clearance could be observed after 1-2 days incubation at 37°C (Figure 3.4 D). The zones of clearance observed around the colonies are an indication of hydrolysis of the tributyrin by the carboxyl esterases.

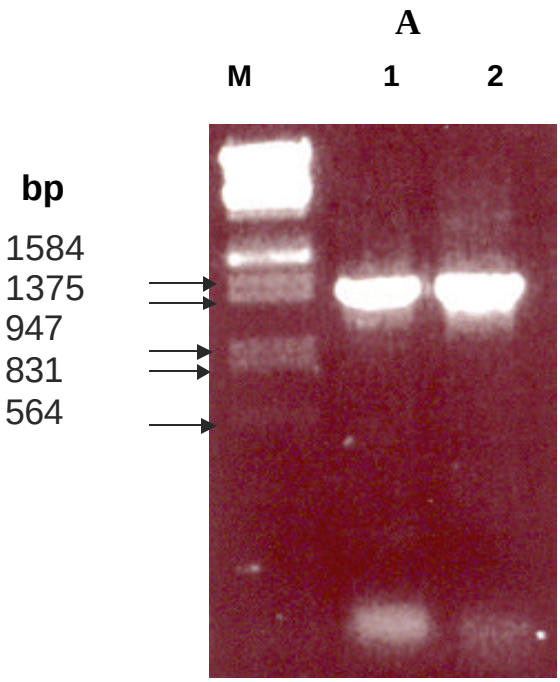


Figure 3.4: **A**, PCR amplified *B. pumilus* and *B. licheniformis* carboxyl esterase ORFs (lane 1 and 2, respectively). M, represents the molecular weight marker, ? DNA digested with *Eco*RI and *Hind*III.

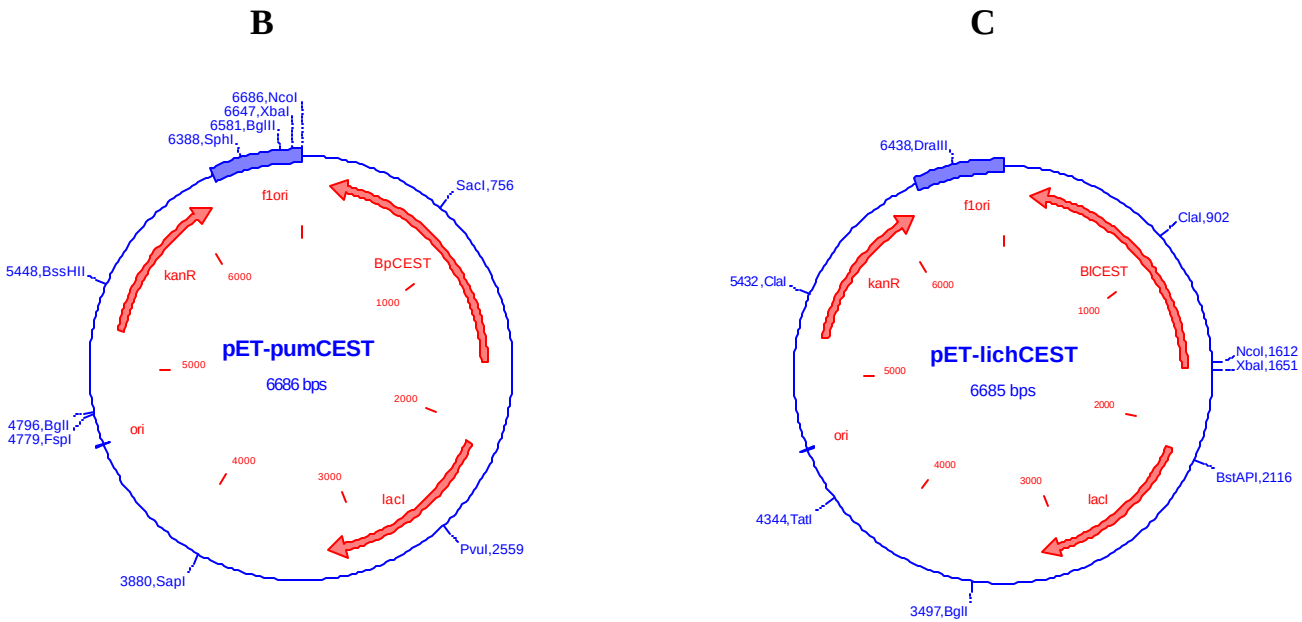


Figure 3.4 : **B**, Schematic representation of the expression plasmid pET-pumCEST. **C**, expression plasmid pET-lichCEST

D

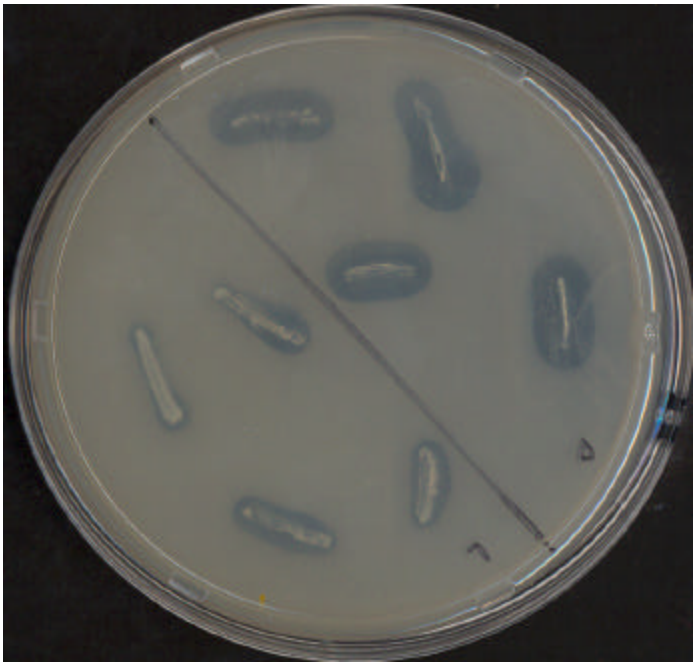


Figure 3.4: **D,** *Escherichia coli* host cells carrying pET-pumCest (P) and pET-lichCEST (L) patched on TLB agar plates, showing zones of clearance around the growing colonies of *Escherichia coli* cells.

3.3.2 Construction of chimeric Family VII carboxyl esterases

The N and C-terminal domains of about 600 and 900 bp (Figure 3.5) respectively were successfully amplified from both the *B. pumilus* carboxyl esterase (BpCEST) and *Bacillus* licheniformis carboxyl esterase (BlCEST) ORFs with the aid of the constructed overlapping primers. The amplified BpCEST N-terminus (~ 600 bp) and BlCEST C-terminus (~ 900 bp) were pooled together and used as a template in an overlap PCR which resulted in the successful amplification of a ~1500 bp hybrid 1 carboxyl esterase ORF (Figure 3.5 A). Hybrid 2 carboxyl esterase ORF (~1500 bp) (Figure 3.6 A) was

constructed by pooling together the PCR amplified BlCEST N-terminus (~600 bp) and BpCEST C- terminus (~900 bp) and using the mixture as a template in an overlap PCR. The amplified hybrid 1 and 2 carboxyl esterase ORFs were ligated into pET 28a resulting in constructs pET-hybrid1 (Figure 3.6 B) and pET-hybrid2 (Figure 3.6 C), respectively. The constructs were used to transform *Escherichia coli* host cells and selected on LB kanamycin. The resultant transformants were replica plated onto TLB agar plates and zones of clearance (Figure 3.6 D) could be observed after 1-2 days of incubation at 37°C. The resultant zones of clearance are an indication of the hydrolysis of tributyrin by the carboxyl esterases being by produced by the host *Escherichia coli* cells. Sequencing and sequence analysis of the hybrid carboxyl esterases (Figure 3.7) revealed similarities in the amino acid composition of the hybrid carboxyl esterases domains to the corresponding domains of their derivatives.

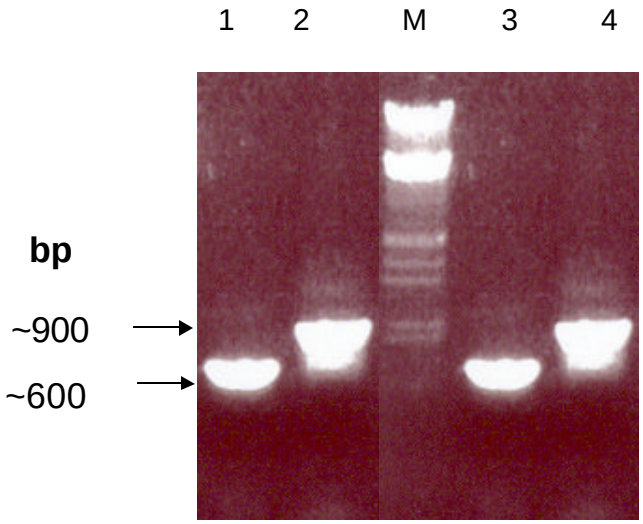


Figure 3.5: PCR amplified BpCEST N- and C-terminus (lane 1,2), BlCEST N- and C-terminus (lane 3 and 4). M, molecular weight marker, ?DNA digested with *Eco*RI and *Hind*III.

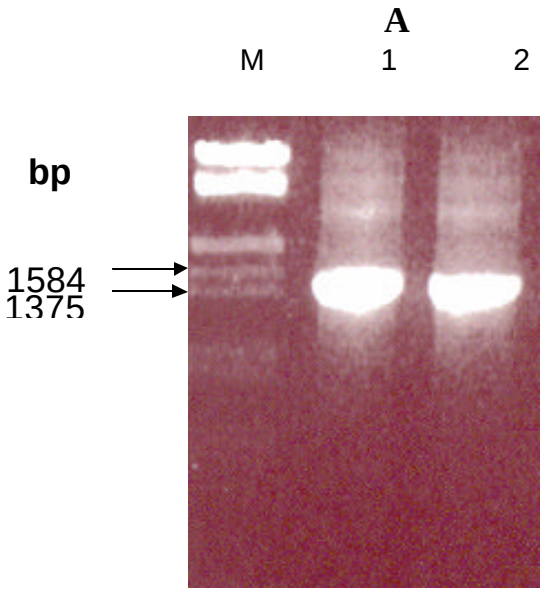


Figure 3.6: **A**, PCR amplified hybrid 1 and 2 carboxyl esterase ORFs (lane 1 and 2, respectively). M, represents the molecular weight marker, ? DNA digested with *Eco*RI and *Hind*III.

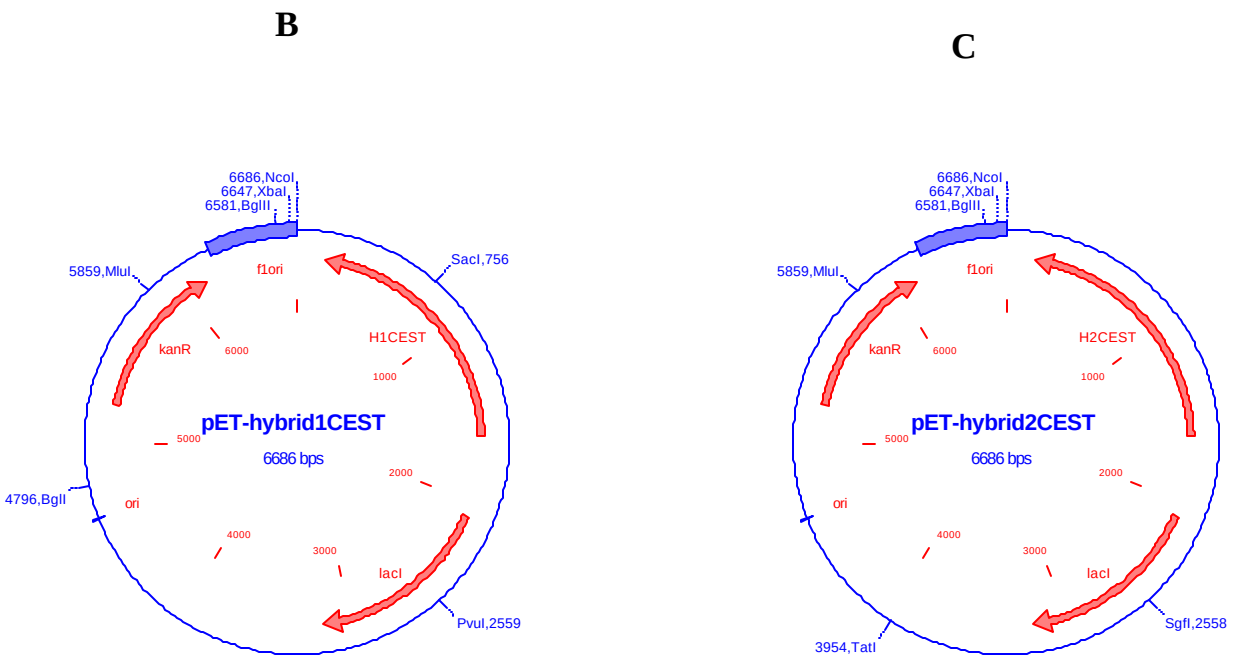


Figure 3.6: **B**, Schematic representation of the expression plasmids pET-hybrid1CEST, **C**, pET-hybrid2CEST.

D

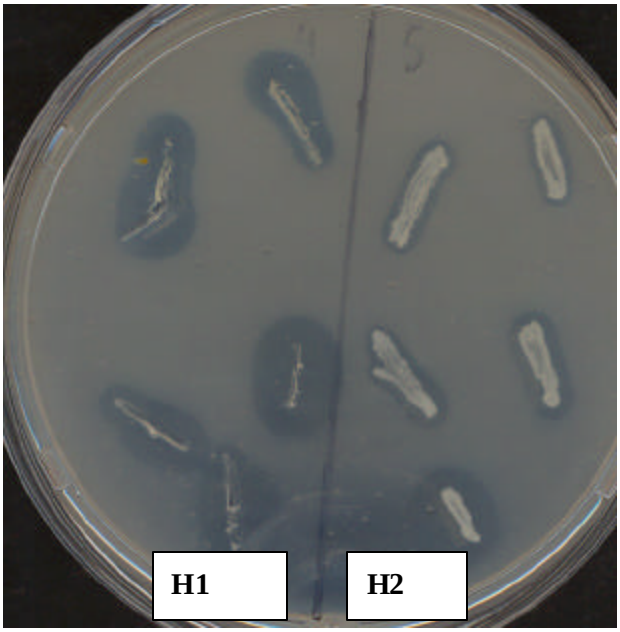


Figure 3.6: **D**, *Escherichia coli* host cells carrying pET-hybrid1Cest (H1) and pET-hybrid2CEST (H2) patched on TLB agar plates, showing zones of clearance around the growing colonies of *Escherichia coli* cells.

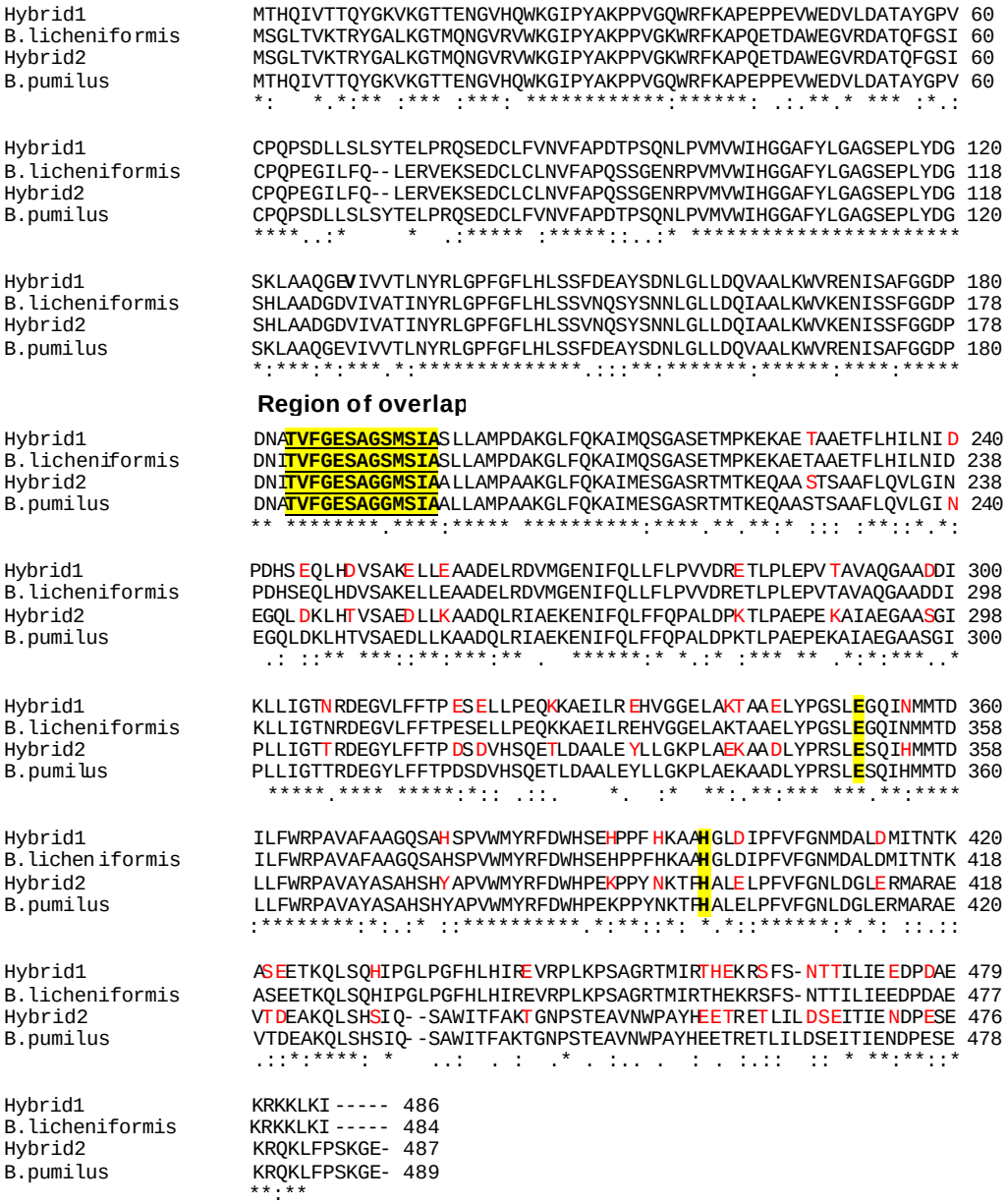
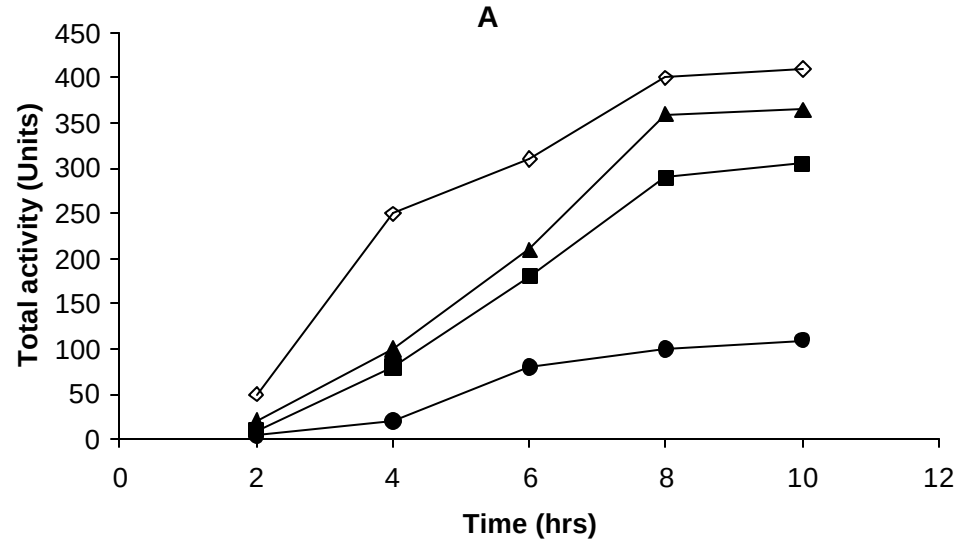


Figure 3.7: Alignment of amino acid sequences encoding *B. pumilus*, *B. licheniformis* and hybrid carboxyl esterases to reveal the degree of amino acid sequence similarity between the domains of the hybrid carboxyl esterases and the corresponding domains of their derivatives. The amino acid residues making up the catalytic triad are bolded, shaded and enlarged. The differences in the ionisable amino acid residues are indicated in red.

3.3.3 Over-expression of the carboxyl esterases

Monitoring the expression levels of the carboxyl esterases revealed a similar profile of enzyme production with the enzyme production reaching a maximum after 10 hours post induction (Figure 3.8 A). Although a similar expression profile was observed for all the carboxyl esterases, there was however a significant difference in the total activity of the enzymes. Monitoring enzyme localisation by assaying the supernatant and cell free extracts collected at 2 hours intervals post induction revealed that the enzymes accumulate in the intracellular for 2 hours after which they become secreted to the extracellular (Figure 3.8 B).



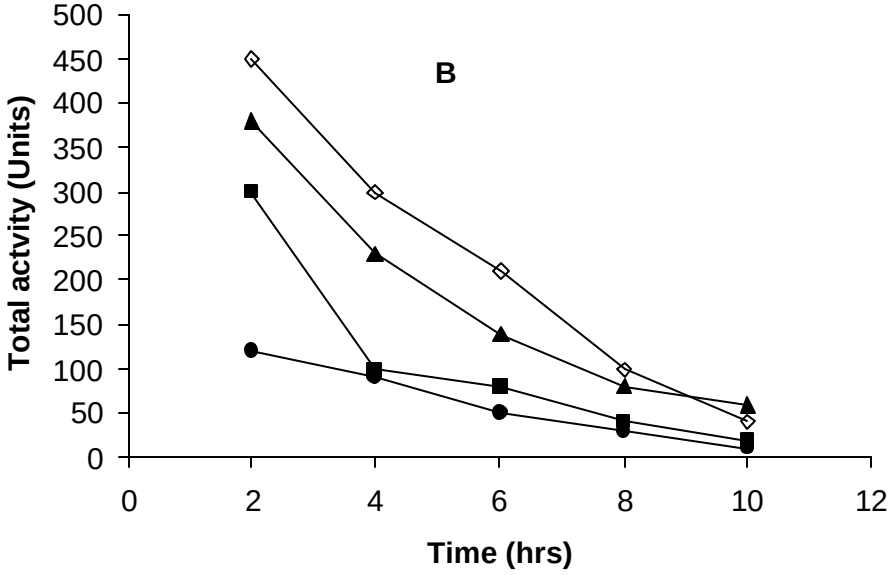


Figure 3.8: Extracellular (A) and intracellular over-expression profiles for the *B. pumilus* (?), *B. licheniformis* (¦) , hybrid 1(?) and hybrid 2 (?) carboxyl esterases.

3.3.4 Enzyme purification

SDS-PAGE analysis of the elution fractions collected after the purification steps of the *B. pumilus*, *B. licheniformis*, hybrid 1, and hybrid 2 carboxyl esterases is shown in Figure 3.9. Lanes 2, 4, 6, and 8 (Figure 3.9) which respectively correspond to the elution fractions collected with the *B. pumilus, B. licheniformis,* hybrid 1 and hybrid 2 carboxyl esterases, show 1 major distinct band (target protein) with a relative molecular mass (M_r) of ~ 50 kilo Daltons (kDa), with several other contaminating bands of differing M_r. Theoretically, proteins that are unspecifically bound to the Protino® Ni 2000 prepacked column should wash off at low imidazole concentrations, while polyhistidine-tagged proteins should be selectively eluted at high imidazole concentrations. The size of the major protein band (~ 50 kD) which can also be observed with the crude proteins (lanes 1, 3, 5, and 7, Figure 3.9) corresponds to the M_r of Family VII carboxyl esterases that have been reported in literature (Arpigny and Jeager, 1999) which suggest that it might be the targeted protein. The presence of several other protein bands is an indication that the wash steps (during purification) did not efficiently wash off the unspecifically bound proteins, which resulted in the partial purification of the carboxyl esterases.

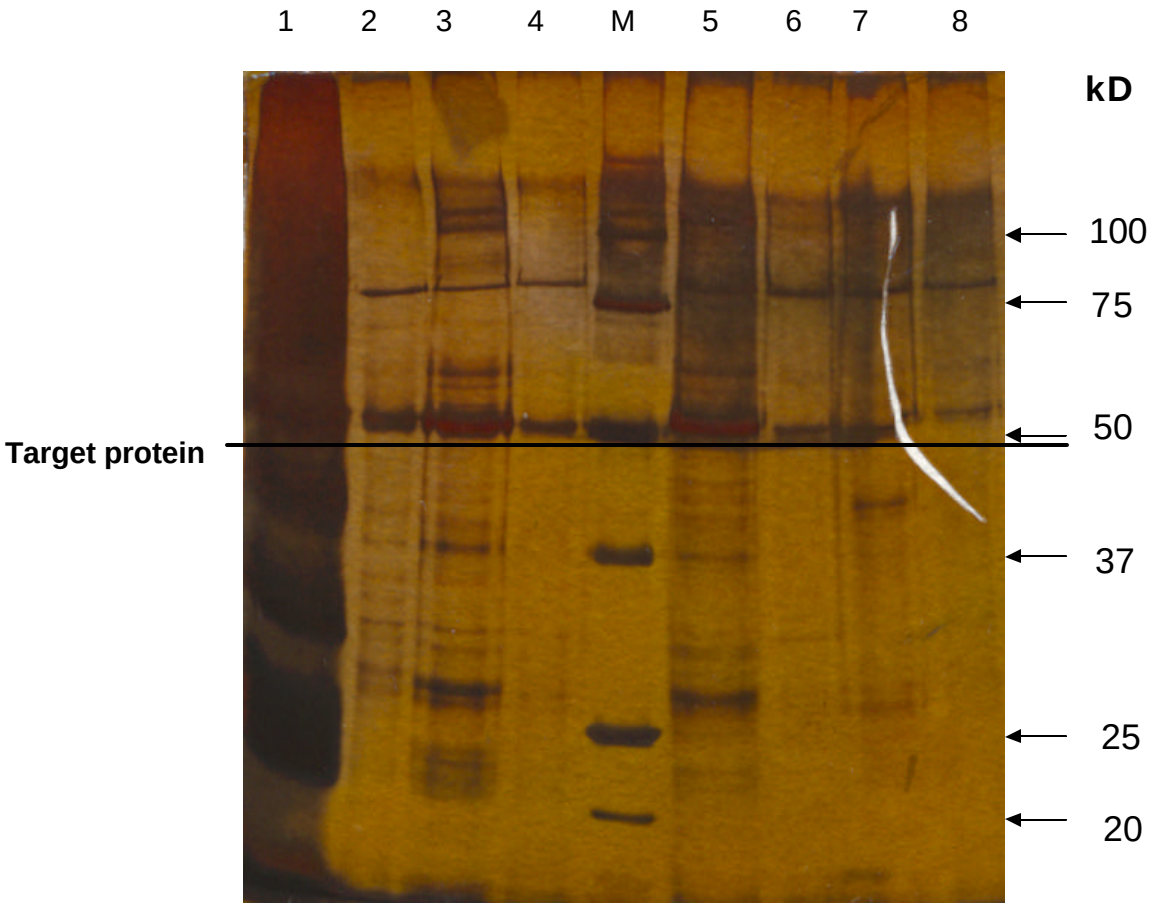


Figure 3.9: SDS-PAGE of the *B. pumilus*, *B. licheniformis,* hybrid 1, and 2 carboxyl esterases. Lanes 1, 3, 5, and 7, represent the crude carboxyl esterases, respectively. Lanes 2, 4, 6, and 8, respectively represents the partially pure carboxyl esterases. M is the protein molecular weight marker. The solid line across the the gel indicates the position of the target protein.

3.3.5 Optimum pH and temperature

The effect of pH on the activities of the crude *B. pumilus*, *B. licheniformis*, hybrid 1 and 2 carboxyl esterases is depicted in Figure 3.10 A. All the carboxyl esterases displayed activity from neutral to alkaline pH (7-9). Broad peaks of pH activity profiles with maximum activity at pH 8.5 and 9, were

observed for the *B. licheniformis* and hybrid 1 carboxyl esterases, respectively. While sharp peaks of pH activity profiles with maximum activity at pH 7.5 and 8, were observed for hybrid 2 and the *B. pumilus* carboxyl esterase, respectively. Similar temperature activity profiles (Figure 3.10 B) with maximum activity at 35 °C were observed for all the carboxyl esterases.

Table 3.2: pH optima of the native and hybrid carboxyl esterases.

	B. licheniformis	Hybrid 1	*B. pumilus*	Hybrid 2
pH optima	8.5	9	8	7.5

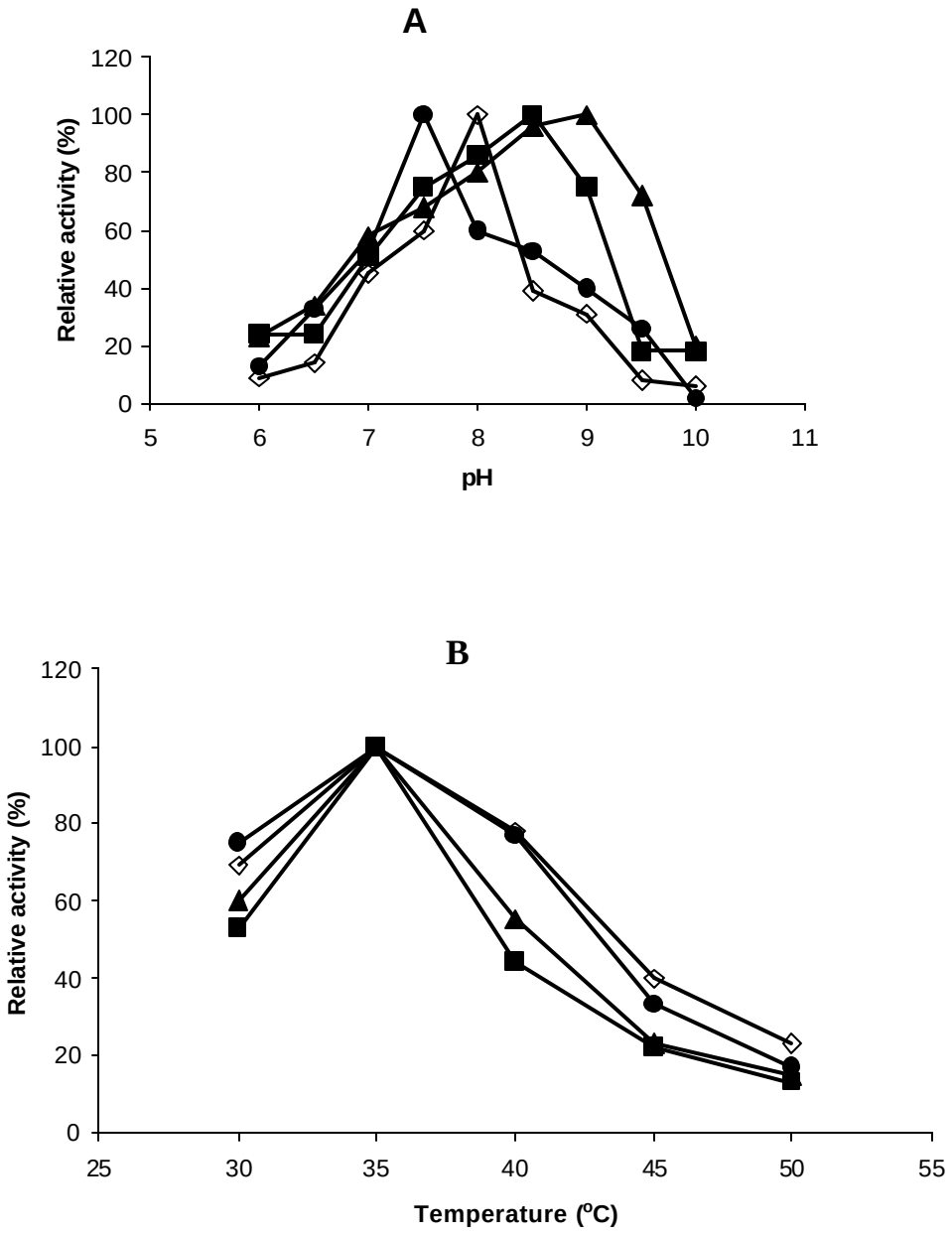


Figure 3.10: pH activity profile (A) and temperature activity profile (B) of BpCEST (?), BlCEST (¡), hybrid 1(?) and hybrid 2 (?) carboxyl esterases.

3.3.6 Homology modelling of the carboxyl esterases

The general structure of the hybrid carboxyl esterases made up of the N-terminal domain from one *Bacillus* carboxyl esterase (green) and the C-terminal domain (yellow) from the other *Bacillus* carboxyl esterases is depicted in Figure 3.11. It is also evident from the general structure of the hybrid that the walls of the substrate binding cleft is as a result of a combination of the N- and the C-terminal domains from different donors. The N-terminal domain provides the one face of the binding cleft while the C-terminal domain forms the other (Figure 3.11).

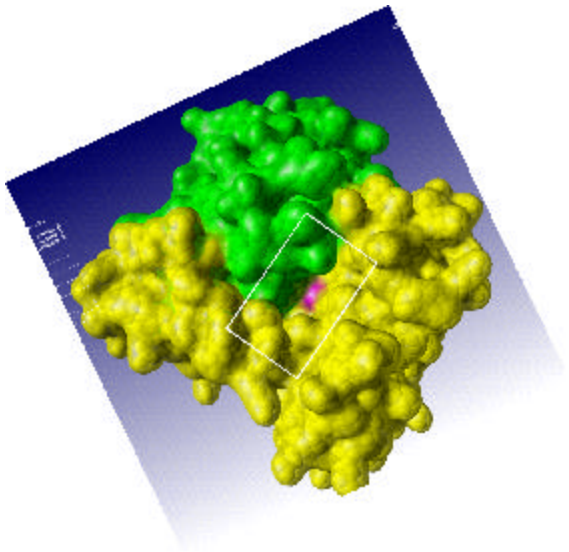


Figure 3.11: Structure of the general hybrid carboxylesterase. Position of the Ser residue (pink), the N- and C-terminal domains from the different *Bacillus* carboxyl esterases indicated as green and yellow, respecitvely

3.4 Discussion

Successful over-expression of other Family VII carboxyl esterases from other *Bacillus* species using *Escherichia coli* expression systems has been reported by Kim *et al*., (2004) and Ewis *et al*., (2004). In this study we successfully over-expressed the *B. pumilus* and *B. licheniformis* carboxyl esterase using the T7 based pET expression system in *Escherichia coli*. Monitoring the localisation and expression levels of these esterases revealed that they accumulate in the cytoplasm for 2 hours post enzyme production induction followed by secretion to the extracellular which reaches a maximum after 10 hours. Since the open reading frames (ORFs) encoding the *B. pumilus* and *B. licheniformis* carboxyl esterases were subcloned into pET 28a such that the recombinant proteins contain a 6X His-tag at their C-termini, purification was attempted using nickel affinity chromatography. The carboxyl esterases were partially purified. This is because a major protein band (target protein) of a molecular weight (~50 kDa) comparable to that of other Family VII carboxyl esterases (Zock *et al*., 1994; Prim *et al*., 2000, 2001; Kim *et al*., 2004; Ewis *et al*., 2004) and numerous other bands of varying molecular weights could be observed upon SDS-PAGE analysis of the fractions collected after the purification steps. As a result, the crude proteins were used for the determination of the pH and temperature activity profiles of the carboxyl esterases. Similar temperature profiles were observed for both carboxyl esterases with maximum activity at 35 °C. However, different pH activity profiles were observed for the esterases. A sharp peak of pH activity with maximum activity at pH 8 was observed for the *B. pumilus* carboxyl esterase,

while a broad peak of pH activity with maximum activity at pH 8.5 was observed with the *B. licheniformis* carboxyl esterase.

The *B. pumilus* and *B. licheniformis* carboxylesterases displayed activity from neutral to alkaline pH. This has been observed for other members of the family (Kim *et al.,* 2004; Ewis *et al*., 2004) and seems to be characteristic to Family VII carboxyl esterases from *Bacillus species.* The discrepancies in the peaks of pH activity profiles (from broad to sharp) is also a common feature amongst members of this family from *Bacillus* species (Kim *et al.,* 2004; Ewis *et al*., 2004). The differences in the pH activity profiles of these two esterases were initially speculated to be species specific, because, a similar observation was made on lipases from *B. pumilus* and *B. licheniformis* which displayed very high sequence similarity, but different pH activity profiles (sharp and broad pH activity profiles, respectively) (Mabizela *et al*., 2005).

To further investigate the determinants of the biochemical specificities of these enzymes, the protein sequences of the *B. pumilus and B. licheniformis* carboxyl esterases were analysed and compared to selected Family VII carboxyl esterases. It was observed upon protein sequence analysis that the N-terminal regions (with the catalytic Ser residue embedded within the GE**S**AG as reference) were relatively conserved, while the C-terminal regions were more variable. It was hypothesised upon this observation that, the N-termini are responsible for the general configuration of the structure of the proteins while the C-termini confer biochemical specificity. To investigate this hypothesis, hybrid Family VII carboxyl esterases were constructed by overlap

PCR aided domain exchange techniques. Hybrid 1 was constructed such that it contained the N-terminal domain from *B. pumilus* carboxyl esterase and the C-terminal domain from the *B. licheniformis* carboxyl esterase. Hybrid 2 contained the N- and C-terminal domains from the *B. licheniformis* and *B. pumilus* carboxyl esterases, respectively. The ORFs encoding the hybrid proteins were also subcloned into pET 28a such that their C-termini contains a 6X His-tag, over-expressed, and partially purified using Nickel affinity chromatography. The hybrid carboxyl esterases were characterised with respect to temperature and pH optima. A broad peak of pH activity profile comparable to the profile of the *B. licheniformis* carboxyl esterase was observed for hybrid 1 which contained the C-terminal domain from the *B. licheniformis* carboxyl esterase. A sharp peak of pH activity profile comparable to the profile of the *B. pumilus* carboxyl esterase was observed for hybrid 2 which contained the C-terminal domain from *B. pumilus* carboxyl esterase. These results demonstrated that the C-terminal region with the more variable amino acid sequences conferred the charateristic pH properties to the enzyme. It was also observed that the pH optima profiles of the hybrids shifted to a more alkaline value (pH 9.0) for Hybrid 1 and for Hybrid 2, the pH profile shifted to a more neutral value (pH 7.5) as compared to the pH profiles of the native donors of the C-termini domains.

What is important to note is that, the two walls of the substrate binding cleft of the hybrid carboxyl esterases (Figure 3.11) is as a result of a combination of the N- and C-terminal domains from different donors. It seems likely that specific residues in the C-terminal domain of the enzyme were mainly

responsible for conferring the pH properties to the hybrid enzymes. Unfortunately, the number of differences in ionisable amino acids (arginine, aspartic acid, glutamic acid, histidine, lysine, serine, threonine and tyrosine) are high (37 differences in total, Figure 3.7) and a simple analysis of the effect of these differences is not possible. Furthermore, the pH profile of the enzymes can be affected by mutations quite distant from the catalytic site in the 3D-structure of the enzyme (Mabizela *et al*., 2005). In future experiments, hybrids will need to be constructed in which the relative contributions toward the N- and C-terminals of the hybrid enzymes can be varied to, in this way identify a critical part or parts of the sequence which determine the pH profile of the enzyme.

The domain exchange experiments are commonly used in enzyme evolution (Coco *et al*., 2001; Suzuki *et al*., 2005; Villbrandt *et al*., 2000), therefore, the shifts in the pH optima observed for both hybrid carboxyl esterases is a possible indication of improvement of enzyme properties as a result of the domain swap. Therefore, further characterisation of the hybrid carboxyl esterases is essential, in order to establish if there were any significant improvements in the properties of the enzymes as a result of domain shuffling.

Chapter 4

Summary and concluding remarks

Carboxyl esterases are serine hydrolases that catalyse the cleavage of carboxyl ester bonds of short chain acyl esters. A number of physiological roles of carboxylesterases that have been suggested include being involved in metabolic processes that provide access to carbon sources, detoxification of compounds toxic to the organism, and as pathogenic agents that enable microorganisms to degrade the cell-wall of the host organism. Carboxylesterases are however biotechnological important with applications in the detergent, food, pharmaceutical and fine chemical industries. These enzymes are widely distributed in nature, and a number of genes encoding carboxylesterases have been cloned. This enabled the classification of bacterial lipolytic proteins into families based on conserved amino acid motifs within the protein sequences and biochemical properties. The objective of the study was to identify conserved sequences within Family VII carboxylesterases and to use these as templates for the designation of universal primers to detect the presence of Family VII members of carboxylesterases within genetic materials from different sources, to clone, over-express and characterize enzymes belonging to this family.

Amino acid sequences encoding Family VII carboxylesterases were aligned and four conserved sequence blocks named Blocks 1, 2, 3 and 4, were identified. Degenerate primers based on conserved Blocks 1,4; 1,3 and 3, 4

were used to screen the presence of homologue genes in the genomes of various *Bacillus* species including *Bacillus pumilus*. The largest possible DNA fragment of about 1100 bp as specified by degenerate primers derived from Blocks 1 and 4 could only be obtained with genomic DNA samples from *Bacillus subtilis* (control) and *Bacillus licheniformis* DSM12369, while *Bacillus pumilus* MBB02 and *Bacillus licheniformis* MBB01 genomic DNA samples failed to yield any PCR products. When Blocks 1 and 3 degenerate primers were used, PCR products of about 500 bp were obtained with all the genomic DNA samples tested except *Bacillus licheniformis* MBB01. The PCR amplification with Blocks 3 and 4 degenerate primers resulted in DNA fragments of about 600 bp with all the DNA samples with the exception of *Bacillus licheniformis* MBB01, a probable indication that the strain lacks the Family VII carboxylesterase homologue genes. Degenerate primers derived from conserved Blocks 1 and 3 specifying the amplifications of the 500 bp fragment were subsequently used to detect the presence of Family VII cest homologues in genetic materials isolated from subsurface mine boreholes biofilms. Subcloning, Restriction Fragment Length Polymorphism (RFLP) analysis and sequencing of the different DNA fragments revealed that novel members of genes encoding Family VII bacterial lipolytic enzymes could be amplified by the designed degenerate primers. The results also indicated that the identified conserved amino acid Blocks are candidate templates for the designation of universal degenerate primers for PCR amplification of Family VII bacterial lipolytic genes from genetic materials from different sources.

The study also described the improved method for genome walking based on the cassette ligation-mediated PCR principle. A 200 bp DNA cassette flanked by various restriction enzymes was introduced within the multiple cloning site of pUC18 plasmid to yield the pLigCas plasmid. Excision of the cassette with restriction enzymes located at opposite ends of the cassette resulted in the release of an efficiently annealed cassette with ends that are ligatable to a compatibly enzyme-restricted genomic DNA sample. Treatment of the excised cassette with alkaline phosphatase prevented self ligation between cassette DNA molecules and ensured preferential ligation with targeted enzyme restricted genomic DNA fragments. The PCR technique referred to as Single-Strand Amplification PCR which involves an initial amplification using a lone primer designed based on the known region of the target region and the cassette-target DNA ligation mixture as the template was employed. The single stranded DNA product obtained during the initial SSA-PCR is used as a template in the second PCR by employing a nested locus specific primer paired with a cassette specific primer in a conventional PCR which results in increased selectivity and specificity of the PCR product.

The SSA-PCR technique was used to amplify 400 and 600 bp *Bacillus pumilus* DNA fragments respectively corresponding to the promoter and terminator regions that flanked the 1100 (500 and 600 fragments) bp region of the CEST gene. The subcloning of the entire open reading frame contained within the amplified *Bacillus pumilus* carboxylesterase in pET 28a expression vector and subsequent plating on nutrient agar medium containing tributyrin resulted in the production of zones of clearence surrounding the growing

colonies, an indication that a lipolytic gene from *Bacillus pumilus* was indeed cloned. Nucleotide sequencing and comparative amino acid alignment revealed that the cloned lipolytic gene belonged to Family VII class of bacterial lipolytic enzymes. Protein sequences alignment and analysis revealed that the N-terminal regions (with the catalytic Ser embedded within the GE**S**AG motif as the reference) of Family VII carboxylestrases were relatively conserved while the C-terminal regions were found to be variable in terms of amino acid conservation.

The ORFs encoding carboxylesterases from *Bacillus pumilus* and *Bacillus licheniformis* were subcloned into pET 28a such that the expressed proteins contained the 6X His tag at the C-termini. A hypothesis was made that the conserved N-termini of Family VII carboxylesterases are responsible for the general configuration of the structure of the enzymes while the variable C-termini confer biochemical specificities. To test this hypothesis, hybrid carboxylesterases were constructed by exchanging the C-termini regions of carboxylesterases from *Bacillus licheniformis* and *Bacillus pumilus*. Hybrid 1 contained the N-terminal region from *Bacillus pumilus* carboxylesterase and the C-terminal from *Bacillus licheniformis* carboxylesterase. Hybrid 2 contained the N-terminal region of *Bacillus licheniformis* carboxylesterase and the C-terminal from *Bacillus pumilus* carboxylesterase. The two hybrids were also subcloned into pET28a expression vector as recombinant proteins containing 6X His tag at the C-termini. The native carboxylesterases from *Bacillus pumilus* and *Bacillus licheniformis* together with the hybrids were over-expressed in *Escherichia coli* JM109 (DE3) cells

The over-expressed proteins were subjected to partial purification using Nickel affinity chromatography and biochemically characterized with respect to pH and temperature optima using *p*-nitrophenyl butyrate as a subtrate. A comparison of their biochemical properties revealed comparable temperature profiles with temperature optima at 35 ºC. A sharp peak of pH profile activity with maximum activity at pH 8 was obtained for *Bacillus pumilus* carboxylesterase. The carboxylesterase from *Bacillus licheniformis* showed a broad peak of pH optimum profile with the highest activity at pH 8.5. The exchanged C-termini domains influenced the biochemical properties of hybrid carboxylesterases. Hybrid 1 which contained *Bacillus licheniformis* C-terminal domain had a broad pH profile which resembled that of the native *Bacillus licheniformis* carboxylesterase. Hybrid 2 which contained the C-terminal domain from *Bacillus pumilus* had a narrow pH profile which resembled that of native *Bacillus pumilus* carboxylesterase. Taken together, the results demonstrated that the C-terminal region is the catalytic region with the variable amino acid sequences confering the charateristic pH profiles. It was also observed that the pH optima profiles of the hybrids shifted to a more alkaline value (pH 9.0) for Hybrid 1 and for Hybrid 2, the pH profile shifted to a more neutral value (pH 7.5) as compared to the pH profiles of the native donors of the C-termini domains. It would be interesting to do biochemical studies of hybrid carboxylesterases constructed by exchanging domains of Family VII carboxylesterases from thermophilic and mesophilic sources.

CHAPTER 5

Opsomming

Kaboksiel esterases is serien hidrolases wat die kliewing van karboksiel esterbinding van kort-ketting asielesters kataliseer. 'n Aantal fisiologiese rolle van karboksiel esterases wat voorgestel is, is onder meer betrokkenheid by metaboliese prosesse wat toegang tot koolstofbronne verleen, detoksifisering van stowwe wat toksies is vir die organisme en as patogeniese agente wat mikro-organismes in staat stel om die selwand van die gasheerorganisme te degradeer. Karoboksiel esterase is biotegnologies belangrik met toepassings in die wasmiddel, voedsel, farmaseutiese en spesialis chemiese industriee. Hierdie ensieme is wyd verspreid in die natuur en 'n aantal gene wat vir karboksielseterases kodeer is gekloneer. Dit het dit moontlik gemaak om bakteriese lipolitiese ensieme te klassifiseer in families gebasseer op gekonserveerde aminosuur motiewe in die proteien opeenvolgings en biochemiese eienskappe. Die doel van hierdie studie was om gekonserveerde motiewe in Familie VII van die Karoboksiel esterases te identifiseer en om dit aan te wend as template vir die ontwerp van universele priestukke om die teenwoordigheid van Familie VII Karoboksiel esterases in die genetiese materiaal van verskeie oorspronge waar te neem, om dit te kloneer, oor-uit te druk en hierdie ensieme te karakteriseer.

Aminosuur-opeenvolgings koderend vir Familie VII Karoboksiel esterases is inlyngestel en gekonserveerde blokke, genaamd Blok 1, 2, 3 en 4 is geidentifiseer. Degenererende priemstukke gebasseer op gekonserveerde Blokke 1,4; 1,3 en 3,4 is gebruik om te toets vir die teenwoordigheid van homoloe gene in die genome van verkeie Bacillus spesies insluitend Bacillus pumilus. Die grootste moontlike DNS-fragment van ongeveer 1100 bp soos gespesifiseerd deur degenererende priemstukke verky vanaf Blokke 1 en 4 kon slegs verkry word met genomiese DNS-monsters vanaf Bacillus pumilus (kontrole) en Bacillus lichenformis DSM12369, terwyl Bacillus pumilus MBB02 en Bacillus lichenformis MBB01 genomiese DNS-monsters geen PKR-produkte opgelewer het nie. Wanneer Blokke 1 en 3 degenererende priemstukke gebruik is, is PKR-produkte van ongeveer 500 bp verkry met al die genomiese DNS-monsters getoets behalwe Bacillus lichenformis MBB01. Die PKR-amplifisering met Blokke 3 en 4 degenererende priemstukke het DNS-fragmente van ongeveer 600 bp opgelewer met al die DNS-monsters met die uitsluiting van Bacillus lichenformis MBB01, 'n moontlike aanduiding dat hierdie stam nie die Familie VII Karoboksiel esterase homoloe gene bevat nie. Degenererende priemstukke gebasseer op gekonserveerde Blokke 1 en 3, wat 'n 500 bp fragment amplifiseer, is dus gebruik om die teenwoordigheid van Familie VII CEST homoloe waar te neem in die genetiese materiaal geisoleer uit ondergrondse boorgate biofilms.

Subklonering, Beperkinsensiem Fragment Lengte Polimorfisme-analise en basispaaropeenvolging-bepaling van die verskillende DNS-fragmente het gewys dat unieke lede van gene koderend vir Familie VII bakteriese lipolitiese

ensieme geamplifiseer kan word deur die ontwerpte degenererende priemstukke. Die resultate het ook gewys dat die geidentifiseerde gekonserveerde aminosuur Blokke kandidaat-template is vir die ontwerp van universele degenererende priemstukke vir PKR-amplifisering van Familie VII bakteriese lipolitiese gene van genetiese materiaal van verskeie oorspronge.

Die studie het ook die verbeterde metode vir genoom-stap gebasseer op die kaset-ligering PKR-beginsel beskryf. 'n 200 bp DNS-kaset is aangrendend deur verskeie beperkingsenieme is geplaas in die veelvoudige kloneringsgebied van die pUC18 plasmied om die pLigCas plasmied te lewer. itsnyding van die kaset met beperkingsensieme aan weerskante van die kaset het die kaset vrysgestel met punte wat ligeerbaar was met 'n ooreenstemmende ensiem-gesnyde genomiese DNS-monster. Behandeling van die uitgesnyde kaset met alkaliese fosfatase het selfligering tussen kaset DNS-molekules verhoed en het by voorkeur ligering met ensiem-gesnyde genomiese DNS-fragmente verseker. Die PKR-tegniek beskryf as Enkelstring Amplifiserende PKR behls die aanvanklike amplifisering met 'n enkele priemstuk ontwerp op die bekende gebied van die teikengebied en diekaset teiken DNS ligeringsmengsel as die templaat is bebruik. Die enkelstring DNS-produk verkry gedurende die aanvanklike EA-PKR is gebruik in 'n tweed

Die studie het ook die verbeterde metode vir genoom-stap gebasseer op die kaset-ligering PKR-beginsel beskryf. 'n 200 bp DNS-kaset aangrendend deur verskeie beperkingsenieme is geplaas in die veelvoudige kloneringsgebied van die pUC18 plasmied om die pLigCas plasmied te lewer. Uitsnyding van

die kaset met beperkingsensieme aan weerskante van die kaset het die kaset vrysgestel met punte wat ligeerbaar was met 'n ooreenstemmende ensiem-gesnyde genomiese DNS-monster. Behandeling van die uitgesnyde kaset met alkaliese fosfatase het selfligering tussen kaset DNS-molekules verhoed en het by voorkeur ligering met ensiem-gesnyde genomiese DNS-fragmente verseker. Die PKR-tegniek beskryf as Enkelstring Amplifiserende PKR behls die aanvanklike amplifisering met 'n enkele priemstuk ontwerp op die bekende gebied van die teikengebied en die kaset teiken DNS ligeringsmengsel as die templaat is gebruik. Die enkelstring DNS-produk verkry gedurende die aanvanklike EA-PKR is gebruik in 'n tweede PKR deur gebruik te maak van interne lokus-spesifieke priemstuk gepaard met 'n kaset-spesifieke priemstuk in 'n konvensionele PKR wat verhooge selektiwiteit en spesifisiteit van die PKR-produk to gevolg het.

Die EA-PKR-tegniek is gebruik om 400 en 600 bp DNS-fragmente van Bacillus pumilus te amplifiseer wat ooreenstem met die promotor en termineerder gebiede wat aangrensend is tot die 1100 (500 en 600 bp fragmente) bp van die CEST geen. Die subklonering van die hele oopleesraam bevat in die Bacillus pumilus Karoboksiel esterase in die pET 28a uitdrukkingsvektor en daaropvolgende uitplaat op voedingsagar media met tributyrien het sones van verheldering om die groeiende kolonies tot gevolg gehad, 'n aanduiding dat 'n lipolitiese geen van Bacillus pumilus wel gekloneer is. Nukleotied opeenvolgordebepaling en vergelykende aminosuur inlynstelling het gewys dat die gekloneerde lipolitiese geen aan die Familie VII-klas van bakteriese lipolitiese ensieme behoort het. Inlynstelling van

proteien basispaaropeenvolgings en analise het getoon dat die N-terminale gebiede (met die katalitiese Ser in die GESAG-motief as die verwysing) van Familie VII Karoboksiel esterases relatief gekonserveerd is, terwyl die C-terminus-gebied meer varieerbaar is in terme van aminosuur-konservering.

Die oopleesrame koderend vir Karoboksiel esterases vanaf Bacillus pumilus en Bacillus lichenformis is gesubkloneer in pET 28a sodat die uitgedrukte proteiene die 6X His tag aan die C-terminus bevat het. 'n Hipotese is daargestel dat die gekonserveerde N-terminus van Familie VII Karoboksiel esterases verantwoordelik is vir die algemene konfigurasie van die struktuur van die ensieme, terwyl die varieerbare C-terminus biochemises spesifisiteite verleen. Om hierdie hipotese te toets, is hibried Karoboksiel esterases gekonstueer deur die C-terminale gebied van karboksiesesterases vanaf Bacillus lichenformis en Bacillus pumilus uit te ruil. Hibried 1 het die N-terminale gebied van Bacillus pumilus Karoboksiel esterase en die C-terminus van Bacillus lichenformis Karoboksiel esterase bevat. Hibried 2 het die N-terminale gebied vanaf Bacillus lichenformis Karoboksiel esterase en die C-terminale gebied vanaf Bacillus pumilus Karoboksiel esterase bevat. Hierdie twee hibriede is ook in die pET 28a uitdrukkingsvektor gesubkloneer as rekombinante proeiene met die 6X His tag by die C-terminus. Die natuurlike Karoboksiel esterase vanaf Bacillus pumilus en Bacillus lichenformis tesame met die hibriede is oor-uitgedruk in Escherichia coli JM109 (DE3) selle. Affiniteitschromatografie en biochemies gekarakteriseer in terme van pH en temperatuur optimums deur p-nitrofeniel butyraat te gebruik as 'n substaat. 'n Vergelyking van die biochemiese eienskappe het vergelykbare temperatuur-

profiele getoon met 'n optimum by 35 °C. 'n Skerp piek van pH profiel aktiwiteit met maksimum aktiwiteit by pH 8 is verkry vir Bacillus pumilus Karoboksiel esterase. Die Karoboksiel esterase van Bacillus lichenformis het 'n bree piek van pH optimum profiel getoon met die hoogste aktiwiteit by pH 8.5. Die uitgeruilde C-teminus gebiede het die biochemiese eienskappe van die hibried Karoboksiel esterases beinvloed. Hibried 1 wat die Bacillus lichenformis C-terminale gebied bevat het, het 'n bree pH profiel gehad wat ooreengestem het met die natuurlike Bacillus lichenformis Karoboksiel esterase. Hibried 2 wat die C-terminale gebied vanaf Bacillus pumilus bevat het, het 'n smal pH profiel getoon soos die van die natuurlike Bacillus pumilus Karoboksiel esterase. In geheel gesien, het die resultate gedemonstreer dat die C-terminale gebied die katalitiese gebied is met die varieerbare aminosuur opeenvolgings wat die karakteristieke pH profiele tot gevolg het. Dit is ook gesien dat die pH-optimum profiele van die hibriede na 'n meer alkaliese waarde geskuif het (pH 9.0) vir Hibried 1 en vir Hibried 2 het die profiel geskuif na 'n meer neutrale pH-waarde (pH 7.5) in vergelyking met die pH-profiele van die natuurlike skenkers van die C-teminale gebiede. Dit sou interressant wees om biochemiese studies te doen vir hibried Karoboksiel esterases gekonstueer deur uitruiling van Familie VII Karoboksiel esterases vanaf termofiliese en mesofiliese oorspronge.

CHAPTER 6

References

Adamczak M, Krishna SJ (2004). Strategies for improving enzymes for efficient biocatalysts. *Food Technol. Biotechnol.* **42**, 251-264.

Ahn S, Ramirez J, Grigorean G, Lebrilla CB (2001). Chiral recognition in gas-phase cyclodextrin: amino acid complexes-Is three point interaction still valid in the gas phase? *J. Am. Soc. Mass Spectrom.* **12**, 276-287.

Alvarez-Macarie E, Acuier-Macro V, Baratti J (1999). Chatacterization of a thermostable esterase from a moderate thermophile *Bacillus licheniformis*. *Biosci. Biotechnol. Biochem.* **63**, 1865-1870.

Andersson RE, Hedlund GB, Jensson V (1979). Thermal inactivation of a heat-resistant lipase produced by the psychrotrophic bacterium *Pseudomonas fluorescens. J. Dairy Sci.* **62**, 361-367.

Arbige MV, Pitcher WH (1989). Industrial enzymolgy: a look towards the future. *Trends Biotechnol.* **7**, 330-335.

Arpigny JL and Jaeger K-E (1999). Bacterial lipolytic enzymes: classification and properties, *Biochem. J.* **343**, 177–183.

Atlas RM, ed. (1996). Handbook of microbiological media. Boca Rayon, FL:CRC 2nd ed. pp140.

Azzolina O, Verseci D, Collina S, Ghislandi V (1995). Chiral resolution of methyl 2-aryloxypropionates by biocatalytic stereospecific. *II Farmaco* **50**, 221-226.

Baneyx F (1999). Recombinant protein expression in *Escherichia coli. Curr. Opin. Biotechnol.* **10**, 411-421.

Blackman RL, Spence JM, Field LM, Devonshire (1995). Chromosomal location of the amplified esterase genes conferring resistance to insecticides in *Mycus persicae* (Homoptera:Aphidae). *Heredity* **75**, 297-302.

Blöchl E, Rachel R, Burggrat S, Hafenbradl D, Jannasch HW, Stetter KO (1997). *Pyrolobus fumarii*, gen. and sp. No., represents a novel group of archea, extending the upper temperature limit of life to 113ºC. *Extremophiles* **1**, 14-21.

Bornscheuer UT, Altenbucner J, Meyer HH (1997). Directed evolution of an esterase for the stereoselective resolution of a key intermidiate in the synthesis of epothilones. *Rapid Communication* , p554-559.

Bornscheuer UT and Kazlaukas RJ (eds) (1999). Hydrolases in organic synthesis – regio – and stereoselective biotransformations. Wiley – VCH, Weinheim.

Bornscheuer UT and Pohl M (2001). Improved biocatalysis by directed evolution and rational protein design. *Curr. Opin. Chem. Biol.* **5**, 137-143.

Bornsheuer UT (2002). Microbial carboxyl esterases: classification, properties and application in biocatalysis. *FEMS Microbiol. Rev.* **26**, 73-81.

Brady L, Brzowski AM, Derewenda ZS, Dodson E, Dodson G, Tolley S *et al*., (1990). A serine protease triad forms the center of a triacylglycerol lipases. *Nature* **343**:767-770.

Brock BJ, Waterman MR (2000). The use of random chimeragenesis to study structure/function properties of rat and human P450c17. *Arch. Biochem. Biophys.* **373**, 401-408.

Casida JE, Gammon DW, Glickman AH , Lawrence LJ (1983). Mchanism of selective action of pyrethroid insecticides. *Annu. Rev. Pharmacol. Toxicol* **23**, 413-438.

Chao Y-P, Fu H, Wang Y-L, Haung W-B, Wang J-Y (2003). Molecular cloning of the carboxyl esterase gene and biochemical characterization of the encoded protein from *Pseudomonas citronellolis* ATTC 13674. *Res Microbiol.* **154**, 521-526.

Christov LP, Prior BA (1993). Esterases of xylan degrading microorganisms: production, properties and significance. *Enzyme Microb. Technol.* **15**, 460-475.

Coco WM, Levinson WE, Crist MJ, Hektor HJ, Darzins A, Peinkos PT, Squires CH, Monticello DJ (2001). DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nature* **19**, 354-359.

Copeland RA (2000). Enzymes: a practical introduction to structure, mechanism and data analysis, 2nd edition. John wiley and Sons, New York pp149-150.

Cygler M, Shrag JD, Sussman JL, Harel M, Silman J, Gentry MK, Doctor BP (1993). Relationship between sequence conservation and three-dimensional

structure in a large family of esterases, lipases and related proteins. *Protein Sci.* **2**, 336-382.

Dalrymple BP, Swadling Y, Cybinski DH, Xue GP (1996). Cloning of a gene encoding cinnamoyl ester hydrolase from ruminal bacterium *Butyrivibrio fibrisolvens* E14 by a novel method. *FEMS Lett.* **43**, 115-120.

D'Amico S, Claverie P, Collins T, Georlette D, Gratia E, Hoyoux A, *et al,* (2002). Molecular basis of cold adaptation. *Phil. Trans. R. Soc. Lond.* **357**, 917-925.

D'Amico S, Marx J-C, Gerday C, Feller G (2003). Activity-stability relationships in extremophilic enzymes. *J. Biol. Chem.* **278**, 7891-7896.

Davis BG (2003). Chemical modification of biocatalysts. *Curr. Opin. Biotechnol.* **14**, 379-386.

Deming JW (2002). Psychrophiles and polar regions. *Curr. Opin. Microbiol.* **5**, 301-309.

Desnuelle P (1961). Pancreatic lipase. *Adv. Enzymol.* **23**, 129-161.

De Vries RP, Visser J (1999). Regulation of the feruloyl esterase (faeA) gene from *Aspergillus niger. Appl. Environ. Microbiol.* **65**, 5500-5503.

Di Nola A, Roccatano D, Berendson HK (1994). Molecular dynamics stimulation of the docking substrates of proteins. *Proteins Struct. Funct. Genet.,* **19**, 174-182.

Donaghy J, Mckay AM (1997). Purification and characterisation of a feruloyl esterase from fungus *Penicillium expansum*. *J. Appl. Microbiol* **83**, 718-726.

Drauz K and Waldmann H (2002). Enzyme catalysis in organic synthesis. Wiley-VCH.

Dröge MJ, PriesF, Quax WJ (2000). Directed evolution of Bacillus lipases. In:lipases and lipids: structure function and biotechnological applications. Eds. Kokotos G, Kokotos V, pp 169-180.

Ebbing DD (1996). General chemistry 5th edition. Houghton Mifflin Company USA, pp563-565.

Eggert T, Pencreac'h G, Douchet I, Verger R, Dijkstra BW, Jeager K-E (2000). A novel extracellular esterase from *Bacillus subtilis* and its conversion to a monoacylglycerol hydrolase. *Eur. J. Biochem.* **267**, 6459-6469.

Ewis HE, Abdelal AT, Lu C-D (2004). Molecular cloning and characterization of two thermostable esterases from *Geobacillus stearothermophilus*. *Gene* **329**, 187-195.

Feller G, Thiry M, Arpigny JL, Gerday CL (1991). Nucleotide sequence of a lipase gene Lip2 from Antarctic *Moraxella* TA144 and site specific mutagenesis of the conserved serine and histidine residues. *DNA Cell Biol.* **10**, 381-388.

Ferreira LMA, Wood TM, Williamson G, Faulds C, Hazlewood GP, Black GW, Gilbert HJ (1993). A modular esterase from *Pseudomonas fluorescens* subsp. *cellulosa* contains a non-catalytic cellulose-binding domain. *Biochem. J.* **294**, 349-355.

Fiedler K, Simons K (1995). Protein sequence motifs: a new family of lipolytic enzymes? *TIBS* **20**, 178-179.

Fojan P, Jonson PH, Petersen MTN, Petersen SB (2000). What distinguishes esterases from lipases: A novel structural approach. *Biochimie* **82**, 103-1041.

Franken SM, Rozeboom HJ, Kalk KH, Dijkstra BW (1991). Crystal structure of haloalkane dehalogenase: an enzyme to detoxify halogenated alkanes. *EMBO. J.* **10**, 1297-1302.

Georlette D, Damien B, Blaise V, Depiereux E, Uversky VN, Gerday C, Feller G (2003). Structural and functional adaptations to extreme temperatures in psychrophilic, mesophilic and thermophilic DNA ligases. *J. Biol. Chem.* **39**, 37015-37023.

Gerday C, Attaleb M, Bentahir M, Chessa JP, Clavane P, Collins T *et al*., (2000). Cold adapted enzymes: from fundamentals to biotechnology. *Trends Biotechnol.* **18**, 103-107.

Gianese G, Bossa F, Pascarella S (2002). Comparative structural analysis of psychrophilic and meso- and thermophilic enzymes. *Proteins* **47**, 236-249.

Gupta RC, Dettbarn WD (1993). Role of carboxyl esterases in the prevention of potential *N*-methyl carbamate toxicity. *Chem. Biol. Interact.* **87**, 295-303.

Gutfreund H, Sturtevant JM (1956). The mechanism of the reaction of chymotrypsin with *p*-nitrophenyl acetate. *Biochem. J.* **63**, 656-661.

Handelwang WG (1995). The sigma factors of *Bacillus subtilis*. *Microbiol. Rev.* **59**, 1-30.

Henikoff JG, Pietrokovski S, Henikoff S (1997). Recent enhancements to the Blocks Database servers. *Nucleic Acids Res.* **25**, 222-225.

Henke E and Bornscheur UT (2002). Esterases from *Bacillus subtilis* and *B. stearothermophilus* share high sequence homology but differ substantially in their properties. *Appl. Microbiol. Biotechnol.* **60**, 320-326.

Hetch HJ, Sobek H, Haag T, Pfeifer O, van Pee KH (1994). The metal-ion-free oxireductase from *Streptomyces aureofaciens,* has an α/β hydrolase fold. *Nat. Struct. Biol.* **1**, 532-537

Horton RH, Moran AM, Ochs RS, Rawn JD, Scrimgeour KG (1996). Principles of Biochemistry, 2nd edition. Prentice Hall, Inc.USA, pp 154-155.

Hotelier T, Ludovic R, Xavier C, Negre V, Marchot P, Arnaud C (2004). ESTHER, the database for a/ß-hydrolase fold superfamily of proteins. *Nucleic Acids Res.* **32**, 145-147.

Hotta Y, Satoshi E, Harayuki A, Tadayuki I (2002). Extremely stable and versatile carboxyl esterase from hyperthermophillic Archeon. *Appl. Environ. Microbiol.* **68**, 3925-3931.

Jeager KE, Ransac S, Dijkstra BW, Colson C, Heuvel M, Misset O (1994). Bacterial lipases. *FEMS Microbiol. Rev*. **15**, 29-63.

Jeager KE and Reetz MT (1998). Microbial lipases from versatile tools for biotechnology. *Trends Biochem. Biotechnol.***16**, 396-403.

Jeager KE, Dijkstra BW, Reetz MT (1999). Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications. *Annu. Rev. Microbiol*. **53**,315-351.

Kafri R, Lancet D (2004). Probability rule for chiral recognition. *Chirality* **16**, 369-378.

Kim KK, Song HK, Shin DH, Hwang KY, Choe S, Yoo OJ, Suh SW (1997). Crystal structure of carboxyl esterases from *Psuedomons fluorescens,* an a/ß hydrolase with broad substrate specificity. *Structure* **5**, 1571-1584.

Kim HK, Choi HJ, Kim MH (2002). Expression and characterization of Ca^{2+} - independent lipase from *Bacillus pumilus* B26. *Biochim. Biophys. Acta* **1583**, 205-212.

Kim G-J, Lee E-G, Gokul B, Hahm M-S, Perna D, Choi GS (2003). Identification, molecular cloning and expression of a new esterase from *Psuedomonas* sp. KCTC 10122BP with enantioselectivity towards ketoprofen ethyl ester. *J. Mol. Catal. B: Enzym.* **22**, 29-35.

Kim HK, Na HS, Park MS, Oh TK, Lee TS (2004). Occurrence of ofloxacin ester-hydrolyzing esterase from *Bacillus niacini* EM001. *J. Mol. Catal. B* **27**, 237-241.

Koeller KM, Wong C-H (2001). Enzymes for chemical synthesis. *Nature* **401**, 232-240.

Kouker G, Jeager K-E (1987). Specific and sensitive plate assay for bacterial lipases. *Appl. Env. Microbiol.* **53**, 211-191.

Laemmli UK (1970). Most commonly used discontinuous buffer system for SDS electrophoresis. *Nature* **227**, 680-688.

Lee O-W, Koh Y-S, Kim K-J, Kim B-C, Choi H-J, Kim D-S, Suhartono MT, Pyun Y-R (1999). Isolation and characterization of a thermophilic lipase from *Bacillus thermoleovorans* ID-1. *FEMS Microbiol. Lett.* **179**, 393-400.

Lesage-Meessen G, Delattre M, Haon M, Thibault JF, Colona-Ceccaldi B, Bruneire P, Asther M (1996). A two step conversion process for vanillin production from ferulic acid combining *Aspergillus niger* and *Pycnoporus cinnabarinus*. *J. Biotechnol.* **50**, 107-113.

Liao D-I, Remington S (1990). Structure of wheat serine carboxipeptidase II at 3.5Å resolution. *J. Biol. Chem.* **265**, 6528-6531.

López C, Blanco A, Pastor FIJ (1998). Xylanase production by a new alkali-tolerant isolate of *Bacillus*. *Biotechnol. Lett.* **20**, 243-246.

Mabizela NB, Litthauer D, Nthangeni MB (2005). Probing the determinants of optimal pH-activity prolfiles of Family I.4 *Bacillus* lipases. SASBMB XIXth conference 16-20 January 2005, Stellenbosch University.

Malcata XF, Reyes HR, Hugo SG, Hill CG, Amundson CH (1992). Kinetics and mechanisms of reactions catalysed by immobilized lipases. *Enz. Microb. Technol.* **14**, 426-446.

Malcata XF, Reyes HR, Garcia CS, Hill CG, Amundson CH (1990). Immobilized lipase reactors for modification of fats and oils – a review. *JAOCS* **67**, 890-899.

Marino MH (1989). Expression systems for heterologous protein production. *BioPharm* **2**, 18-33.

McQueen DAR, Schottel JL (1987). Purification and characterization of a novel extracellular esterase from *Streptomyces scabies* that is induced by zinc. *J. Bacteriol.* **169**, 1967-1971.

Montero S, Blanco A, Virto M, Landeta L, Agud I, Slozabal R , *et al* (1993). Immobilization of *Candida rugosa* lipase and some properties of the immobilized enzyme. *Enz. Microb. Technol.* **15**, 239-243.

Moore JC and Arnold FH (1996). Directed evolution of a *para*-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol.* **14**, 458–467.

Morana A, Di Prizito N, AuriliaV, Rossi M, Cannio R (2002). A carboxyl esterase from the hypothermophilic archeon *Sulfolobus solfataricus*: cloning of the gene, characterization of the protein. *Gene* **283**, 107-115.

Noble MEM, Cleasby A, Johnson LN, Egmond MR, Frenken LGL (1993). The crystal structure of the triacylglycerol lipase from *Pseudomonas glumae* reveals a partially redundant catalytic aspartate. *FEBS*. **331**, 123-128.

Nofziger DL (2005). Temperature dependent degradation of pesticides: Theory (2005). http://soilphysics.okstate.edu

Nthangeni MB, Patterton H, van Tonder A, Vergeer WP, Litthauer D (2001). Over-expression and properties of a purified recombinant *Bacillus licheniformis* lipase: a comparative report on *Bacillus* lipases. *Enzyme Microb. Technol.* **28**, 705-712.

Nthangeni MB, Ramagoma F, Tlou MG, Litthauer D (2005). Development of a versatile cassette for directional genome walking using cassette ligation-mediated PCR and its application in the cloning of complete lipolytic genes from *Bacillus* species. *J. Microbiol. Methods* **61**, 225-234.

Oefner C, D'Arcy A, Daly JJ, Gubernator K, Charnas RL, Heinze I, Hubschwerlen C, Winkler FK (1990). Refined crystal structure of ß-lactamase from *Citrobacter freudii* indicates a mechanism of ß-lactam hydrolysis. *Nature* **343**, 284-288.

Ollis DL, Cheah E, Cygler M, Dijkstra BW, Frolow F, Franken SM et al., (1992). The a/ß hydrolase fold. *Protein Eng.* **5**, 197-211.

Pelletier I and Altenbuchner P (1995). A bacterial lipase is homologous with non-haem haloperoxidases and displays brominating activity. *Microbiol.* **141**, 459-468.

Petersen MTN, Martel P, Petersen EI, Drabløs F, Petersen SB (1997). Surface and electrostatics of cutinases. *Methods Enzymol.* **284**, 130-154.

Petersen EI, Valinger G, Sölkner B, Stubenrauch G, Schwab H (2001). A novel esterase from *Burkholderia gladioli* shows high deacetylation activity on cephalosporins is related to ß-lactamases and DD-peptidases. *J. Biotechnol.* **89**, 11-25.

Petersen MTN, Fojan P, Petersen SB (2002). How do lipases work: the electrostatic contribution. *J. Biotechnol.* **85**, 115-147.

Pohlenz HD, Boidol W, Schuttke I, Streber W (1992). Purification and properties of an *Arthrobacter oxydans* P52 carbamate hydrolase specific for the herbicide phenmedipham and nucleotide sequence of the corresponding gene. *J. Bacteriol.* **174**, 6600-6607.

Prim N, Blanco A, Martinez J, Pastor FIJ, Diaz P (2000). *est*A, a gene coding for a cell-bound esterase from *Peanibacillus* sp. BP-23, is a new member of the bacterial subclass of type B carboxyl esterases. *Res. Microbiol.* **151**, 303-312.

Prim N, Pastor FIJ, Diaz P (2001). Cloning and characterization of a bacterial cell-bound type B carboxyl esterase from *Bacillus* sp. BP-7. *Curr. Microbiol.* **42**, 237-240.

Quax WJ and Broekhuizen CP (1994). Development of a new *Bacillus* carboxyl esterase for the resolution of chiral drugs. *Appl. Microbiol. Biotechnol.* **41**, 425-435

Rosenau F, Jeager K-E (2000). Bacterial lipases from *Pseudomonas*: regulation of gene expression and mechanism of secretion. *Biochimie* **82**, 1-10.

Sambrook J, Fritsch EF and Maniatis T (1989). Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, New York.

Satoh T, Taylor P, Borson WF, Sonal P, Sanghani P, Hosakawa M, La du B (2002). Current progress on esterases: from molecular structure to function. *Drug Metab. Dispos.* **30**, 488-493.

Shumann W and Ferreira LCS. Production of recombinant proteins in *Escherichia coli*. *Genetics Mol. Biol.* **27**, 442-453.

Senter PD, Marquardt H, Thomas BA, Hammock BD, Frank IS, Svensson HP (1996). The role of serum carboxyl esterase in the activation of paclitaxel and campotothecin prodrugs. *Cancer Res.* **56**, 1471-1474.

Shan G, Hammock BD (2001). Development of a sensitive esterase assays based on a-cyano-containing esters. *Anal. Biochem.* **299**, 54-62.

Shono T, Ohsawa K, Casida JE (1979). Metabolism of trans-permethrin and cis-permethrin, trans-cypermethrin and cis-cypermethrin and decamethrin by microsomal enzymes. *J. Agric. Food Chem.* **27**, 316-325.

Shyamala Vand Ames GF (1993). Genome walking by single specific primer–polymerase chain reaction, *Methods. Enzymol.* **217**, 436–446.

Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC (1985). Measurement of protein using the bicinchoninic acid. *Analytical Biochem.* **150**, 78-85.

Sørensen HP, Mortessen KK (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J. Biotechnol.* **115**, 113-128.

Sundaresan V, Abrol R (2002). Towards a general model for protein-substrate stereoselectivity. *Protein Sci.* **11**, 1330-1339.

Sussman JL, Harel M, Frolow F, Oefner C, Goldman A, Toker L, Silman I (1991). Atomic structure of acetyl choline esterase from *Torpedo californica*: A prototypic acetyl-binding protein. *Science* **253**, 872-879.

Suzuki T, Nakayama T, Kurihara T, Nishino T, Esaki N (2002). Primary structure and catalystic properties of a cold active esterase from a psychrotroph, *Acinetobacter* sp. strain no. 6. isolated from Siberian soil. *Biosci. Biotechnol. Biochem.* **66**, 1682-1690.

Suzuki Y, Kagawa N, Fujino T, Sumiya T, Andoh T, Ishikawa K, Kimura R, Kemmochi K, Ohta T, Tanaka S (2005). A novel high-throughput (HTP) cloning strategy for site-directed designed chimeragenesis and mutation using gateway cloning system. *Nucleic Acid Res.* **33**, 2-12.

Switzer RC, Merril CR, Shifrin S (1979). A highly selective silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal. Biochem.* **98**, 231-237.

Takami H, Takaki Y, Chee GJ, Nishsi S, Shimamura S, Suzuki H, Matsui S, Uchiyama I (2004). Thermoadaptation trait revealed by the genome sequence of thermophilic *Geobacillus kaustophilus*. *Nucleic Acids Res.* **32**, 6292-6303.

Takayama H, Watanabe A, Hosakawa M, Chiba K, Satoh T, Aimi N (1998). Synthesis of a new class of campotothecin derivatives, the long chain fatty acid esters of 10-hydroxycamptothecin, as a potent prodrug candidate, and their in vitro metabolic conversion by carboxyl esterases. *Bioorg. Med. Chem. Lett.* **8**, 415-418.

Thomas SM, DiCsimo R, Nagarajan V (2002). Biocatalysis: applications and potentials for the chemical industry. *Trends Biotechnol.* **20**, 238-242.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876-4882.

van Pouderoyen G, Eggert T, Jeager K-E, Djikstra BW (2001). The crystal structure of *Bacillus subtilis* lipase: a minimal a/ß-hydrolase fold enzyme. *J. Mol. Biol.* **309**, 215-226.

Uppenberg J, Hansen MG, Paktar S, Jones TA (1994). The sequence crystal structure determination and refinement of two crystal forms of lipase *Candida antarctica*. *Structure* **2**, 293-303.

Verger R (1998). Interfacial activation of lipases: facts and artifacts. *TIBTECH* **15**, 32-38.

Villbrandt B, Sobek H, Frey B, Schomburg D (2000). Domain exchnage: chimeras of *Thermus aquaticus* DNA polymerase, *Escherichia coli* DNA polymerase I and *Thermotoga neopolitana* DNA polymerase. *Protein Eng.* **13**, 645-654.

Villeneuve P, Muderhwa JM, Graille J, Haas MJ (2000). Customizing lipases for catalysis: a survey of chemical, physical and molecular approaches. *J. Mol. Catal. B: Enzym.* **9**, 113-148.

Vorderwülbecke T, Kieslich K, Erdman H (1992). Comparison of lipases by different assays. *Enzyme Microb. Technol.* **14**, 631-639.

Vriend G (1990). WHATIF: a molecular modelling and drug design programme. *J. Mol. Graph.* **8**, 52-56.

Wang CS, Hartsuck JA (1993). Bile salt activated lipase. A multiple function enzyme. *Biochim Biophys Acta* **1166**, 1-19.

Wie Y, Swenson L, Kneusel RE, Matern U, Derewenda ZS (1996). Crystallization of a novel esterase which inactivates the macrolide toxin brefeldin A. *Acta Crystallogr. Sect. D* **52**, 1194-1195.

Winkler Uk, Stuckman M (1979). Glycogen, hyaluronate, and some other polysaccharides greatly enhance the formation of exolipase by *Serratia marcescenes*. *J. Bacteriol.* **138**, 663-670.

Wheelock CE, Serveson TF, Hammock BD (2001). Synthesis of new carboxyl esterase inhibitors and evaluation of potency and water solubility. *Chem. Res. Toxicol.* **14**, 1562-1572.

Whitaker JR (1972). Principles of enzymology for the food sciences. *Marcel Decker Inc. NY,* 481-501.

Zaks A and Klibanov AM (1985). Enzyme-catalyzed processes in organic solvents. *Proc. Natl. Acad. Sci. USA* **82**, 3192-3196.

Zock J, Cantwell C, Swartling J, Hodges R, Pohl T, Sutton K, Rosteck P, McGilvray D, Queener S (1994). The *Bacillus subtilis pnbA* gene encoding *p*-nitrobenzyl esterase: cloning, sequence and high level expression in *Escherichia coli*. *Gene* **151**, 37-43.

APPENDIX 1

Attached, is a copy of the publication: Nthangeni MB, Ramagoma F, **Tlou MG**, Litthauer D (2005). Development of a versatile cassette for directional genome walking using cassette ligation-mediated PCR and its application in the cloning of complete lipolytic genes from *Bacillus* species. *J. Microbial. Methods* **61**, 225-234.