PURIFICATION, CRYSTALLIZATION AND CLONING OF TRIBUTYRIN ESTERASE FROM *LACTOCOCCUS*

by

Jiong Zheng

A thesis presented in partial fulfillment of the requirements for the degree of

Master of Technology in Biotechnology

in the

Institute of Molecular Biosciences

at

MASSEY UNIVERSITY, NEW ZEALAND

July, 1998

ABSTRACT

Tributyrin esterase is an enzyme that has been isolated and purified from lactococcal starter strain by research staff at the New Zealand Dairy Research Institute. It has been shown to play an important role in production and control of flavour development during cheese ripening, but little is known about its biochemical characteristics.

New studies on tributyrin esterase have been initiated, with the aim of carrying out a three dimensional structure determination to completely understand the molecular basis and the nature of its *in vivo* activity.

This thesis is divided into three main parts. In the first part, the purification of tributyrin esterase from a genetically modified strain *Lc. lactis* subsp. *cremoris* B1079 is described. The procedure investigated for optimization of the protocol and a partial study of factors affecting tributyrin esterase activity are described. In the second part, crystallization trials for tributyrin esterase are described. Several crystals were obtained, with the best ordered crystals being grown from 2.6M ammonium sulfate, these have been shown to diffract to 3.0 Å, and belong to the space group C222 with cell dimensions a=76Å, b=178Å c=179Å. In the third part, the lipase gene was ligated into 4.75kbp expression vector proEX, which contains a his-tag sequence upstream of the multiple cloning site. The ligation reaction mixture was transformed into competent *E.coli* DH5α cells. This should allow the expression of tributyrin esterase in *E.coli* and eventually provide a great yield of protein and make purification easier.

ACKNOWLEDGEMENT

I sincerely thank my supervisor, Dr Gill Norris, for her remarkable guidance, and supervision. Her encouragement, patience and enthusiasm throughout the course of this project were deeply appreciated. I would also like to gratefully thank my co-supervisors Associate Professor John Tweedie and Dr Ross Holland (NZDRI), for their support, encouragement and understanding. Work and learning from them will undoubtedly be one of my most memorable experiences.

I would also like to express my appreciation and thank to Professor M Rao Bhamidimarri, Head of Institute of Engineering and Technology, and Professor Pat Sullivan, Head of Institute of Molecular Biosciences, for initiating this project.

I would like to thank the following people for their assistance in the completion of this project:

All members of the Microbiology, Nutrition & Enzyme Science Section, New Zealand Dairy Research Institute;

All members in protein purification Lab, Massey University;

All the occupants of "Twilight Zone", particularly, Prasad, Carole Flyger, Catherine Day, Mark Patchett, Kathryn stowell for their practical assistance and invaluable advices

Dick Poll for use of SMART system

Thanks to my wife for her love, understanding, patience and encouragement.

Finally, I would like to thank the Institute of Molecular BioSciences, Massey University for providing me with the knowledge, environment and resources that enabled me to fulfil this course of study.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

LIST OF TABLES

ABBREVIATIONS

BCA	Bicinchoninic acid
BSA	Bovine serum albumin
BTP	Bis-tris propane; 1,3-bis[tris-hydroxymethyl)-methylamino]propane
DNA	Deoxyribonucleic acid
DTT	D-Dithiothreitol
EDTA	Ethylenediamine tetra-acetic acid (di-sodium salt)
EPPS	N-[2-Hydroxyethyl]piperazine-N'[3-propanesulfonic acid]
FPLC	Fast protein liquid chromatography
HEPES	N-[2-Hydroxyethyl]piperazine-N'[2-ethanesulfonic acid]
HIC	Hydrophobic Interaction Chromatography
HPLC	High performance liquid chromatography
IEX	Ion Exchange Chromatography
MES	2-[N-Morpholino]ethanesulfonic acid
MOPS	3-[N-Morpholino]propanesufonic acid
NMWC	Nominal Molecular Weight Cut-off
NaAc	Sodium acetate buffer
$(NH_4)_2SO_4$	Ammonium sulfate
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEG-mme	Polyethylene glycol monomethyl ether
PIPES	1,4-Piperazinediethanesulfonic acid
Q-	Quaternary amino
SDS	Sodium dodecyl sulfate
SEC	Size Exclusion Chromatography
TE (*x*:*y*)	*x* mM Tris/HCl ph8.0, *y* mM EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
TRIS	Tris(hydroxymethyl)aminomethane

ABBREVIATIONS FOR AMINO ACIDS

Amino Acid	Three Letter Symbol	One letter Symbol*
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Asparagine or aspartic acid	Asx	B
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glutamine or glutamic acid	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1 INTRODUCTION AND LITERATURE REVIEW

1.1 Esterases and Lipases

Lipases and esterases constitute an important group of enzymes, which are associated with fat metabolism as well as with fat degration. They are widely distributed in nature, being present in numerous tissues and fluids of animals, plants and microbes. They are physiologically important since they hydrolyze fats and oils giving rise to free fatty acids and partial glycerides. These are essential for metabolic processes, such as fatty acid transport, oxidation, and the resynthesis of glycerides and phospholipids.

There is still much confusion concerning the exact meaning of the terms lipase and esterase, with these terms often being used interchangeably. According to the International Union of Biochemistry nomenclature (Florkin and Stotz, 1965), the term esterase is assigned to enzymes that catalyze the hydrolysis of carboxylic acid ester bonds. Esterases are therefore classified as hydrolyases and belong to the class 3.1.1. A lipase on the other hand is defined as an enzyme that hydrolyzes carboxylic esters of glycerol and belongs to class 3.1.1.3 hydrolyases. This means that the term esterase is more nonspecific and includes lipases. However, it should be noted that a typical lipase may also catalyze the hydrolysis of esters other than glycerides (Brocherhoff *et al.*1974).

Since Sarda and Desnuelle (1958) found that the lipolytic activity was significantly enhanced by the presence of micelles, this so called " interfacial activation" phenomenon has been regarded as the major distinction between lipases and esterases (Borgstrom, et al, 1984 and Tsujita *et al.*, 1990). More recently, however, Martinez, *et al.* (1992) and Lesuiss, (1993) discovered some lipolytic enzymes with clear sequence homology to lipases that are not activated by interfacial interaction. They therefore suggested that the classification based solely on the interfacial activatability is too simplistic. This idea was supported by Ransac *et al.*, (1996) and Brocherhoff *et al.*, (1974) who proposed to define lipases as carboxylesterases that are able to hydrolyze long chain acylglycerols (10 or more carbons). Nevertheless, the activity of the majority of the known lipases increases in the presence of a water-lipid interface. Thus, the molecular nature of this activation has become an important aspect of lipase research.

Apart from the obvious biological importance, there are other, more specific, reasons that make the study of lipolytic enzymes interesting and rewarding, such as medical and industrial applications. The manipulation of lipolytic activities will probably play a part in future methods for treating malfunctions of fat metabolism and thus control cardiovascular diseases. Of similar interest are the practical aspects of lipases in the food and other industries. In this current project, the lipase of interest was found to have the industrial importance in cheese making. The related background is thus reviewed extensively in following sections.

1.2 Cheese Making and Esterase

1.2.1 Cheese Making Technology

The manufacture of cheese is one of the oldest examples of biotechnology known to human beings. This process has long been regarded as a means of preserving the nutritional value of milk in a safe, palatable form. Traditionally, therefore, the long maturation time of some varieties of cheese was seen as an advantage. Modern cheese making technology, (see figure 1.1) which has evolved from these principles, can generally be regarded as a two stage process. In the first stage (manufacture stage), curds are made from milk, through a series of processes (standardization, pasteurization, curdling, cutting, cooking and draining). In the second stage (ripening stage), mature cheese is developed from curds. The details of this stage are discussed in next section.

1.2.2 Cheese Ripening

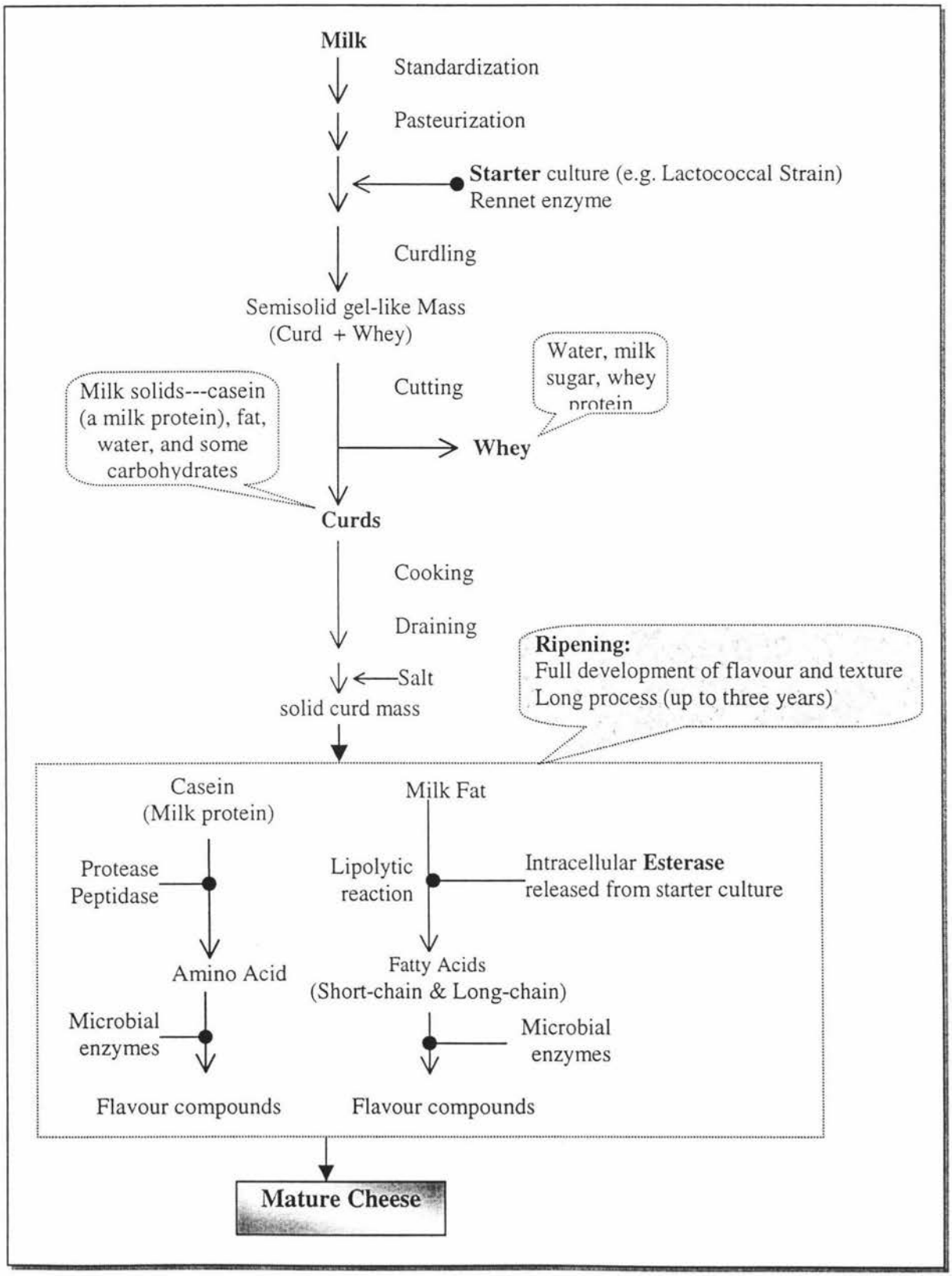
The maturation process involves the slow, controlled decomposition of the fats, proteins and carbohydrates of bland lactic curds. It is mediated by both the starter bacteria, which the cheese maker adds deliberately, and the secondary flora whose composition varies widely between different cheese types. Overall, the major objectives of a ripening system are:

1. To produce the typical flavour and texture of mature cheese;
2. To improve the reliability of mature flavour development;
3. To increase the range of products and flavours available to consumer;

In order to achieve these aims, the ripening process can take up to three years, depending on the type of cheese. In the commercial world, this prolonged period between making cheese and realising its value may be a disadvantage to a factory scale cheese maker who

needs a rapid turnover to cover the high costs of capital equipment and labour. Therefore, it is commercially advantageous to accelerate the ripening.

Figure 1.1: Scheme of Cheese Making Technology



1.2.3 Esterase in Cheese Ripening

Cheese ripening is essentially an enzymatic process involving curd breakdown by proteolysis, lipolysis and other enzyme-catalyzed reactions, resulting in cheese with typical flavours and texture. Some of these enzymes are present in milk whereas others are synthesised by starter lactic acid bacteria and liberated by the autolysis of these bacteria. While the role of proteinases in the cheese ripening process has been studied extensively (for reviews, see Fox, 1989; Olsen, 1990; Kok, 1993), the exact role of lipolytic enzymes is not yet completely understood. McNeil & Connolly (1989) suggested that after esterase was released from starter cells trapped in cheese curd, lipolysis occurred to release free fatty acids from triacylglycerols in milk fat. It has been known for some time that short-chain fatty acids such as *n*-butanoic and *n*-hexanoic acids are potent flavour compounds, at concentrations below 10 mg/kg (Brennand *et al.* 1989). Short chain fatty acids may also be esterified to compounds such as alcohols to give rise to a further range of potent flavour compounds (Bills *et al.* 1965). This evidence implies that the lipolytic activity in starter bacteria probably plays a significant role in flavor development (Crow *et al.*, 1994; Lawrence *et al.* 1976).

1.3 Esterases from *Lactococcus* and other starter bacteria

1.3.1 Lipolytic activity in starter bacteria

Lactococcus (formerly *Streptococcus*) species (and sometimes *Lactobacillus* species) are widely used as bacterial starters in cheese manufacture. Lipolytic activity due to these strains has been reported by a number of groups (Lawrence et al. 1967; Umemoto & Sato, 1975; Kamaly *et al.* 1989). The level of lipase activity in these strains is low however compared to lipolytic activity found in organisms such as *Pseudomonas* and other dairy spoilage microbes. While the lipases of different species of bacteria prefer to hydrolyze different substrates, most of the strains produce a lipase that is able to catalyze the hydrolysis of tributyrin. Umemoto *et al.* (1968) observed that when a tributyrin emulsion served as substrate, lipolysis in the presence of various dairy lactic acid bacteria was optimal between pH 6 and 8 and at 37°C. Oterholm *et al.* (1968) also tested a number of lactic acid bacteria, and found that all cell-free extracts possessed lipolytic activity against an emulsion of tributyrin, which suggested that the lipase associated with lactic acid bacteria is intracellular. In addition to lipolytic activity, extracellular esterase activity was also observed when cell free extracts of the various lactic acid bacteria were

incubated with an aqueous solution of triacetin (Lawrence, 1967a; Brockerhoff & Jensen 1974). Several esterases have been purified and partially characterised from several starter strains, and some of their features are summarised in Table1.1.

1.3.2 Esterase from *Lactococcus*

1.3.2.1 Tributyrin Esterase from *Lactococcus lactis* subsp *cremoris* E8

A tributyrin esterase from *Lactococcus lactis* subsp *cremoris* E8 (Holland & Coolbear, 1995) has been identified at the New Zealand Dairy Research Institute. The activity appeared to be located in the cytoplasm of the organism, although there was a limited degree of association with the cell membrane. The following purification protocol was used to isolate this enzyme from the bacteria. Salt fractionation (60% ammonium sulfate precipitation) was followed by hydrophobic interaction chromatography (HIC) on an alkyl Superose column. The active fractions from this step were subjected to gel filtration chromatography (SEC) using an Superose-6 column, and the final step involved ion exchange chromatography (IEX) on a Mono-Q anion exchange column. The enzyme was shown to have a monomeric molecular mass of 29 kDa by SDS-PAGE but to associate as a holoenzyme of molecular mass of approximately 109 kDa in solution, implying that it is a tetramer in its native configuration. It has optimum activity at neutral pH, on milk fat and liberates mainly C_4, C_{14}, and C_{16} fatty acids from milk fat.

1.3.2.2 Esterase from *Lactococcus lactis* subsp. *lactis* NCDO 763

Chich *et al.*, (1997) purified an intracellular esterase from *L. lactis* subsp. *lactis* NCDO 763, using polyethylene glycol (PEG) fractionation, IEX (Mono Q at pH 8.0), HIC (phenyl Sepharose) and IEX (Mono-Q at pH 5.5). This enzyme was shown to be intracellular homotrimer with a monomeric molecular mass of 29 kDa, which is very similar to that of the tributyrin esterase (in 1.3.2.1). While the enzyme hydrolyses *para*-nitrophenyl esters from C2 to C12 and *ortho*-nitrophenyl butyrate, maximum activity is however, observed with *p*-nitrophenyl butyrate at pH 8.0 and 55°C. The enzyme was found to be strongly inhibited with pefabloc, diisopropyl fluoro-phosphate and 3,4-dichloroisocoumarin, demonstrating that it belongs to the group of B-esterases that have serine at their active site (Aldridge, 1993). While this esterase preferentially hydrolyzes *p*-nitrophenyl (pNP) butyrate, it is also able to hydrolyse pNP esters longer than C8. As this enzyme does not strictly correspond to the criteria for esterase defined by Tsujita *et*

al. (1990), the author proposed that this esterase should be classified as belonging to a new esterase/lipase family.

1.3.2.3 An Esterase from *Lactococcus lactis* subsp. *lactis* ACA-DC 127

Tsakalidou *et al.* (1992) purified an esterase from *Lactococcus Lactis* subsp. *lactis* ACA-DC 127 originating in Greek feta cheese. The purification steps involved were IEX (DEAE-cellulose) and SEC (Sephadex G100). The esterase had a molecular weight of 68,000 Da, optimum activity with 4-nitrophenyl butyrate at pH 8.0 and 45 °C and was strongly inactived by PMSF. Interestingly, 1,10-phenathroline and EDTA, which are metal chelators, had no effect on enzyme activity, showing that the enzyme is not dependent on the present of metal ions for activity. In fact the bivalent cations Hg^{2+} and Cu^{2+} were shown to inhibit the esterase activity.

1.3.2.4 The Genetically Modified Lactococcal Strain used in this study

The obvious benefits of controlling the production of desirable enzymes and proteins by lactic acid bacteria (LAB) used for fermented products (*e.g.* cheese) has led to the development of some genetic tools (for review, see Kleerebezen and Kupiers, 1997). De Ruyter (1996) developed some controllable expression systems based on the autoregulatory properties of nisin biosynthesis by *Lactococcus lactis*. Using the N-terminal sequence of the esterase purified from *Lactococcus lactis* subsp *cremoris* E8 by Holland & Coolbear (1995) to design probes, the gene of this esterase was isolated, cloned and overexpressed in a lactococcal strain using a controlled nisin promoter system (Holland, 1997). The resultant strain, *Lactococcus lactis* subsp. *cremoris* B1079 (est^+) has been utilised by the New Zealand Dairy Research Institute to overproduce tributyrin esterase in amounts 50-100 fold higher than in strain E8. The characteristics of the recombinant strain have been compared with that of original strain (Holland, 1997), and been show to have basically the same characteristics as those of the original esterase, except for one difference that shows up in the HIC purification step.

1.3.3 Esterase from other starter bacteria

Gobbetti, *et al.* (1997) purified an intracellular tributyrin esterase from *Lactobacillus plantarum* 2739, a strain isolated from good quality Cheddar cheese. This enzyme is a monomer with relative molecular mass of ca. 85 kDa that is able to hydrolyze β-naphthyl

Table 1.1: Summary of Comparison of Different Microbial Esterases

Microbial source	Molecular mass (kDa)	pI	Factors Affecting Activity[C]	Substrate	Purification steps	Reference
L. lactis ssp *cremoris* E8	109 with monomer 29 (25.5[a])	5.49[a]	Protease inhibitor	tributyrin, *p*-nitrophenyl butyrate	$(NH_4)_2SO_4$ fractionation, alkyl Superose, Superose 6, Mono Q	Holland,1995
L. lactis ssp. *Lactis* NCDO 763	91 with monomer 29	4.5	DFP +, pefabloc +, 3,4-DCI +, Hg^{2+} +,	C4 > C6, C8>C2 >C10 >C12	PEG fractionation, Mono Q (pH 8.0), Phenyl Sepharose, Mono Q (pH 5.5),	Chich,1997
L. lactis ssp. *Lactis* ACA-DC 127	68 monomer	ND	EDTA -, DFP -, PMSF +, Cu^{2+} +, Hg^{2+} +, Ca^{2+} -,	<C8	DEAE-cellulose, Sephadex G-100	Tsakalidou et al. (1992)
L. *plantarum* 2739	85 monomer	ND	Hg^{2+} +, Ag^{+} +,	C2-C10, prefer β-naphthyl	DEAE-cellulose, Sephacryl 200, CM-Cellulose, Mono Q	Gobbetti,1995
Lactobacillus fermentum DT41	67 monomer	ND	Hg^{2+} +, Ag^{+} +, PMSF +,	C2-C10 Limited in C12-C14	Q Sepharose, Sephacryl 200, phenyl Superose, Mono Q	Gobbetti et al. (1997)

(a) Predicticted from computer program
(b) Complete Protease Inhibitor (from Boehringer Mannhem)
(c) + : increase activity; -: decrease activity;

(β-NA) esters of fatty acids from C_2 to C_{10}, tributyrin, tricaprylin, and, to a lesser extent, milk fat. The same authors (Gobbetti *et al.*1997), recently purified a cell surface-associated esterase from *Lactobacillus fermentum* DT41, a starter used in the production of Parmesan cheese. This esterase was shown to be a monomer with a molecular mass of 67 kDa and has optimum activity at pH 7.0 between 30-35 °C.

1.4 Structure Studies on Lipase

As mentioned in previous sections, lipases are widely used in a number of practical applications. Studies on the applied technology of these enzymes are hindered by the lack of detailed knowledge about the molecular basis for their activity. The ability to manipulate lipase activity and to construct designed enzymes is dependent on such knowledge, allowing rational site directed mutagenesis to be carried out in order to bring about changes in activity. Fundamental studies of lipolytic enzymes have been driven by these demands. In this section reported structural studies on esterases from starter bacteria and other lipases are discussed extensively.

1.4.1 Primary Amino Acid Sequences of Lipases

1.4.1.1 Consensus Sequence

The first complete amino acid sequence of a lipase was reported by De Caro (1981), and a review by Antonian (1988) listed 13 sequences; today their number probably exceeds 60. While there is little sequence similarity between the microbial lipases, they all contain the G-X-S-X-G motif, in which the Ser has been identified as being essential for catalysis. In many cases this has been confirmed either by chemical labelling with inhibitors or site-directed mutagenesis, or both (Cygler, *et al.*1997). Another common feature in the sequences of microbial lipases is a HG (His-Gly) dipeptide located about 70 residues N-terminal to the catalytic serine at the margins of the lipid-binding pocket (Langin, *et al.*, 1993).

1.4.1.2 Primary Amino Acid Sequence of Tributyrin Esterase

Comparison of all the known N-terminal sequences of esterases from starter bacteria (Table1.2) reveals a striking similarity between the esterase from *Lactococcus lactis* sp. *lactis* NCDO 763 (see 1.3.2.2) and that from *L. lactis* sp. *cremoris* E8 (1.3.2.1). In

addition, they also have same monomeric molecular weight, suggesting these two lipases might be the same protein or very closely related in their structure and kinetics.

Table 1.2: Comparison of N-Terminal sequences of lipases from *Lactococcus* species

Microbial source	N-terminal sequence	Reference
L. lactis ssp *cremoris* E8	AVINIEYYSEVLGMNRKVNVIYPESSK	Holland,1995
L. lactis ssp. *lactis* NCDO 763	AVINIEYYSEVLGMNXKVNVIYPEXRFE	Chich,1997
L. plantarum 2739	SNEHTQEVLNQTVAD	Gobbetti,1995

* X= unidentified residue

A search of the GCG database revealed that there are some sequence similarities between tributyrin esterase (EC3.1.-.-) from *L. lactis* ssp *cremoris E8*, and two xylanases, acetylxylan esterase (EC 3.1.-.-) from *Caldocellum saccharolyticum* (coded by the *Xyn C* gene) (~45% overall similarity and ~34% identity) and xylanase Z (EC 3.2.1.8) from *Clostridium thermocellum* (~38% overall similarity and ~30% identity). The comparison is shown in Figure1.2.

```
TEL    1  MAVINIEYYSEVLGMNRKVNVIY....PESSKVEDFTQTDIPVLYLLHGMSGNE  50
XynC   1  MAIMQINFYSKMLKKNTTILAILPVDKPDKKFQKDVDSENLKTLYLLHGYAGNY  54
XynZ  51    VVNISYFSTATNSTRPARVYL....PPGYSKD....KKYSVLYLLHGIGGSE  94

TEL   51  NSWIIRSGIERLIRHTNLAIVMPSTDLGFYVNT.TYGMNYF.DAIAHELPKVI  101
XynC  55  MDWLCGARIVELSMRYNVAVFLPSGENSFYLDD.EEKEEYFGEFVGNEIIEFT  106
XynZ  95  NDWFEGGGRANVIADNLIAEGKIKPLIIVTPNTNAAGPGIA.DGYENFTKDLL  146

TEL   102 NNFFPNLSTKRE...KREKNFIAGLSMGGYGAYR..LAL.GTDYFSYAASLS   147
XynC  107 RSVFP....IPQ...KREKTFIGGLSMGGYGALRNGLKY.NKNFVGIIALSS   150
XynZ  111 NSLIPYIESNYSVYTDREHRAIAGLSMGGGQSFN..IGLTNLDKFAYIGPIS   196

TEL   143 GVL..TFDGMEENFKENPA...YWGGIFGNWETFKGSDNEILSLADRKQENK   188
XynC  147 ALIIHKIAGIPKDYRNAYASYNYYRRVFGDLNSLIGSDKDINALVTKLKQEK

TEL   189 ...PKLYAWCGKQDFLFPGNEYATAELKKLG    222
XynC  197 GSIPKIYMACGRDDFLVQENRDLFNFLKNEG    233
```

Figure 1.2 Comparison of amino acid sequence

TEL (tributyrin esterase from *L. lactis* ssp. *cremoris E8*), Xyn C (acetylxylan esterase from *Caldocellum saccharolyticum* coded by Xyn C gene) and Xyn Z (Xylanase Z from *Clostridium thermocellum*)

In general, proteins with common functions sometimes have similar sequences limited to short stretches around essential catalytic residues (Drables, *et al.*, 1997). From the comparison of the above proteins, it is clearly shown that the sequence similarity among

these proteins is around the LYLLHGXXG (40~50) and GLMSGG region (114~128). Hence, Ser (124) can be postulated as the catalytic centre. However, the three-dimensional structure of acetylxylan esterase from *Caldocellum saccharolyticum*, which has similar molecular size, is still unknown. The three-dimensional structure of the catalytic domain of xylanase Z from *Clostridium thermocellum* has been determined at 2.6Å (Dominguez, *et al.*, 1995). The catalytic domain, residues 531 to 789 of Xyn Z is located in the C-terminal region of this enzyme (Souchon, *et al.* 1994 and Grepinet, *et al.* 1988), which does not contain the sequence similarity with tributyrin esterase. This comparison, however, indicates that these enzymes may have a common evolutionary origin. Furthermore, it has been shown that homologous proteins usually have the same main-chain folding and possibly related functions (Topham, *et al.*, 1990). Comparison and analysis of known three-dimensional protein structures has also clearly demonstrated that proteins with ~30% sequence identity or more, are likely to adopt similar tertiary folds (Schrag, *et al.*, 1997). It will not be unexpected therefor, if the tertiary structures of the tributyrin esterase and the non catalytic domain of xylanase Z, are shown to be similar.

1.4.2 Three-Dimensional Structures of Lipases

In 1990, the first crystal structures of two unrelated lipases were reported: one of an enzyme purified from a fungus *Rhizomucor miehei* (RmL) at 1.9 Å resolution (Brady, *et al.*, 1990) and the other of the human digestive enzyme, pancreatic lipase (hPL) at 2.3Å resolution (Winkler, 1990). To date the 3-dimensional structures of more than 10 lipases from mammalian and bacterial sources have been determined by X-ray crystallography and work on more is near completion. Several common structural features have been identified in these proteins.

1.4.2.1 Catalytic Triad in the Active Site

In all the lipase structures solved to date, the serine in G-X-S-X-G consensus is located in a sharp turn between a β-strand and a buried α-helix. The side chain of the catalytic serine is hydrogen bonded to the side chains of two other residues, usually His and Asp (or Glu), which serves to make the serine hydroxyl group strongly nucleophilic. (Derewenda, *et al.*, 1991)

1.4.2.2 Lipase Fold

Analysis of known protein structures has shown that similar folds are sometimes adopted by proteins with little sequence similarity (Bowie, *et al.*,1991 & Alon, 1995). The α/β hydrolase fold is one example of a tertiary fold that is adopted by a number of proteins displaying no sequence similarity (Schrag, *et al.*, 1997). Apart from the pancreatic lipases, all the lipases with known three-dimensional structures contain only one domain with an α,β doubly wound protein fold and are formed from a parallel β sheet and a number of helices that flank the sheet on both sides. The minimal fragment of this fold common to all lipases is a subset of the α/β-hydrolase fold as described by Ollis *et al.* (1992) that contains a five-stranded β sheet and two α helices.

1.4.2.3 Two Classes of Conformers: Closed and Open

One of the important features of an enzyme is the shape of the substrate-binding site and its accessibility to the substrate. Structures of the majority of the enzymes show that their active sites are on the surface of the molecule that are accessible to solvent. This is not always the case for lipases. The various lipase structures determined to date can be divided into two categories: those with the active site accessible to the solvent (the open form) and those with an inaccessible active site, where a "lid" closes over the active site (the closed form). For some lipases both forms have been experimentally observed (Cygler, *et al.*,1997).

1.4.2.4 Oxyanion Hole

Previous crystallographic and biochemical studies of lipases have also shown that the mechanism of hydrolysis by lipases is similar to that of the serine proteases. In both cases, an oxyanion created during hydrolysis, is located in the so-called "oxyanion hole" and is stabilised through non-covalent interaction with some electrophiles (Martinez, *et al.*, 1992).

1.5 Needs and aims of the current project

As reviewed in above section, little is known about the molecular basis for trybutyrin esterase activity in starter culture bacteria. We know from studies undertaken by Holland (1995) that tributyrin esterase hydrolyses tributyrin. But we cannot be sure however, that tributyrin is the preferred substrate *in vivo*. The primary structural homology to the non

catalytic domain of xylanase Z from *Clostridium thermocellum* is intriguing, as it is not-known whether this domain has any specific functions. However, the amino acid sequence homology has raised the possibility that the enzyme, which contains a postulated active site similar to many other hydrolases, may be able to hydrolyse other biomolecules. Many families of enzymes that are classified through function have been found to have similar or conserved structures or contain conserved structural motifs (Schrag, *et al.*, 1997). Often the structural homology is more striking than sequence homology. Therefore the determination of the three dimensional structure of tributyrin esterase and the subsequent analyses and comparison with other three dimensional structures in the protein data base should unequivocally show whether or not the enzyme is a true esterase, or whether it belongs to another enzyme class.

The three-dimensional structure determination of a protein by X-ray crystallographic techniques requires the production of large single crystals. As reviewed in the following section, the difficulties of screen searching for crystal growth conditions may require milligrams of protein for the initial crystallisation trials. Once crystals suitable for X-ray analysis can be reproducibly grown, additional protein will be needed to improve their quality and size and to prepare heavy-atom derivatives. For all these reasons it is essential to develop a suitable purification procedure in order to produce enough fresh protein of reproducible quality.

The addition of poly-His terminal extensions on target proteins using genetic engineering has for a long time been regarded as a powerful tool for use in the isolation of target proteins due to the ease of the purification procedures which can be applies (see review in section 1.6). In this technique, the gene of a target protein is cloned into a vector that contains the codons for six histidine residues and a fusion protein is produced in a high expression system. A convenient and abundant source of pure protein can thus be relatively easily obtained.

The aim of the current project is to provide a foundation for the subsequent three dimensional structural determination of tributyrin esterase. For all the reasons described above, the immediate goals addressed are first, to purify the tributyrin esterase from the genetically modified lactococcal strain (section 1.3.2.4), secondly, to crystallize this enzyme, and thirdly, to develop an expression system in *E. coli* suitable for over-

expressing the tributyrin esterase. As a longer-term goal, the development of a tributyrin esterase overexpression system in *E.coli* would not only facilitate the preparation of the tributyrin esterase necessary for the structural determination, but would also provide a foundation for site directed mutagenesis studies of the enzyme.

Having addressed the above goals of the current project, in the following sections, a brief overview of facilitating protein purification by histidine tagging will be given. The topic of protein crystallization will also be reviewed.

1.6 Facilitating Protein Purification by Histidine Tagging

In biochemistry and molecular biology, the expression of histidine-tagged fusion proteins is a well-established technique for the identification, and purification of gene products (Hochui *et al.*, 1988; Porath *et al.*, 1976). This technique is based on the affinity of five to six consecutive histidine residues for metal chelating lidands, such as N-nitrilotri-acetic acid (NTA) or imidodiacetic acid (IDA). Association and dissociation of the His-tagged protein can be triggered by competitors such as histidine and imidazole, or by changing the pH of the buffer solution.

Recently, Ledent *et al.*, (1997) reported that the addition of a His-tag to the C-terminal end of β-lactamase had an unexpected influence on both the structure and the thermal unfolding of the protein. Lee *et al.*, (1997) also proved that the direct addition of a His-tag to both the C-terminal and the N-terminal end of an aryl-esterase resulted in a significant change in the enzyme activity. These results served as a warning that His-tag addition may not always be as 'neutral' as has been assumed. Therefore, careful comparison of wild-type and His-tagged proteins is necessary.

1.7 Protein Crystallization

1.7.1 Principles of Protein Crystallization

While protein crystallization has been mainly a trial-and-error procedure in which the protein is slowly precipitated from solution, recent reports have identified the basic processes that occur when crystals form. (Ries-kautt, *et al.*, 1990 and Drenth, 1994) The crystallization of a macromolecule is directly related to its solubility characteristics, which can be represented by a phase diagram (Fig.1.3). Of the many variables affecting

solubility, two of the most important are the concentration of the protein and the concentration and type of precipitant used. In the region under the solubility curve, both nucleation and crystal growth are impossible and any crystal formed previously will dissolve. Above the solubility curve, the concentration of protein is higher and the protein is supersaturated. This region may itself be subdivided into three zones: (i) The precipitate phase, where excess protein immediately separates from the solution in an amorphous state; (ii) The labile phase (nucleation zone), where excess protein separates from the solutions in a crystalline form. Crystallization may occur both as a shower of microcrystals and an amorphous precipitate. (iii) The metastable zone , in which the supersaturated solution may not be nucleated for a long period of time, unless the solution is mechanically shocked or a seed crystal introduced. In this zone, crystals are able to grow and the nucleation of new crystals is minimal.

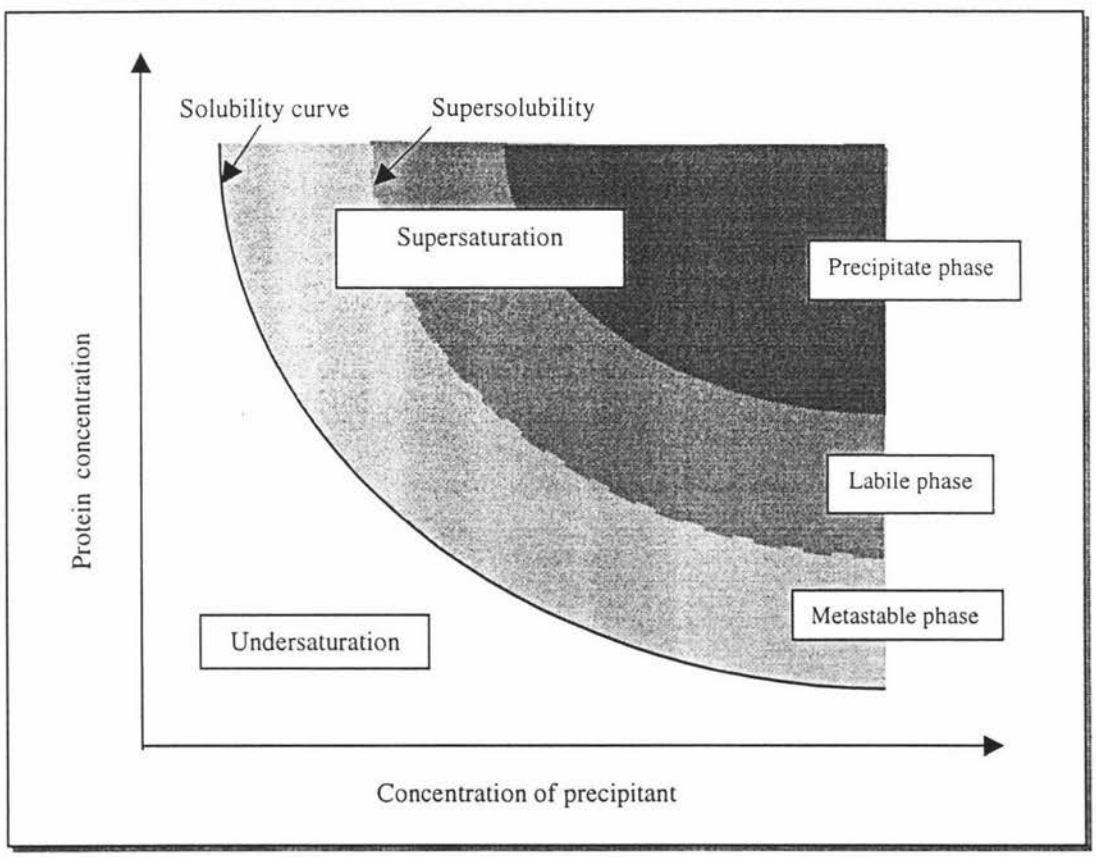


Figure 1.3 Phase diagram.

Overall, there are three major steps in crystallizing a protein:

1. The protein is dissolved in a suitable solvent from which it must be separated in crystalline form. It is important that there should be no aggregation of protein molecules in this solvent.
2. The solution is slowly brought to supersaturation, as spontaneous formation of nuclei is best achieved at a high supersatuation.
3. Once nuclei have formed, actual crystal growth can begin.

1.7.2 Screening for Crystallization Conditions

A large number of parameters influence the crystallization process: protein purity and concentration, buffer type, pH, temperature, ionic strength, and the presence of minute amounts of detergent or organic or inorganic molecules. As a general rule, the more pure the protein, the better the chances of growing crystals. In order to successfully grow crystals suitable for X-ray diffraction, proteins not only have to be pure in terms of lack of other compounds, they have also to have the same surface properties, especially the same charge distribution on their surface (Drenth, 1994).

Finding the best combination among many crystallization parameters is a multivariable optimisation problem. During the past decade, the development of incomplete factorial screening protocols and optimisation strategies have provided investigators with an extended portfolio of effective crystallization tools (Carter, 1990; Jancarik & Kim, 1991; Kingston, 1996).

The process of crystallizing a protein can be divided into three discreet stages. These are (1) screening for useful crystallization conditions or leads; (2) optimisation of one or more initial conditions to produce single crystals suitable for X-ray diffraction analysis; and (3) reproducible production of single crystals for X-ray data collection.

1.7.3 Methods of Crystallization

1.7.3.1 Vapor Diffusion Methods

One of the most commonly employed techniques for growing protein crystals is the hanging drop method. A droplet containing the protein, with buffer, crystallization agent, and additives, is equilibrated against a reservoir containing a solution of crystallizing agent at a higher concentration than in the droplet. Thus, concentrations within the

protein-containing droplet are manipulated remotely by diffusion through air. Equilibration proceeds by diffusion of the volatile species (usually H_2O) until the vapour pressure of the components in the droplet equals those of the reservoir.

1.7.3.2 Dialysis methods

The protein is separated from a large volume of solvent by a semi-permeable membrane which gives small molecules (ions, additives, buffer, and so on) free passage through the membrane, but prevents free passage of the protein. Thus, the dialysis method is uniquely suited to crystallization at low ionic strength and in the presence of volatile reagents such as alcohols. The kinetics of equilibrium depend on the ratio of the concentration of precipitant inside and outside the membrane, the temperature, and the geometry of the cell.