CHAPTER 1

INTRODUCTION

1.1) Lipases

1.1.1) Lipases [1]

Lipases (EC 3.1.1.3) are distributed throughout the living organisms which form two primary divisions of the phylogenetic tree, namely the bacteria and a second division branching into both the eukarya, including animals, plants and fungi, and the archaea, with the former archaebacteria.

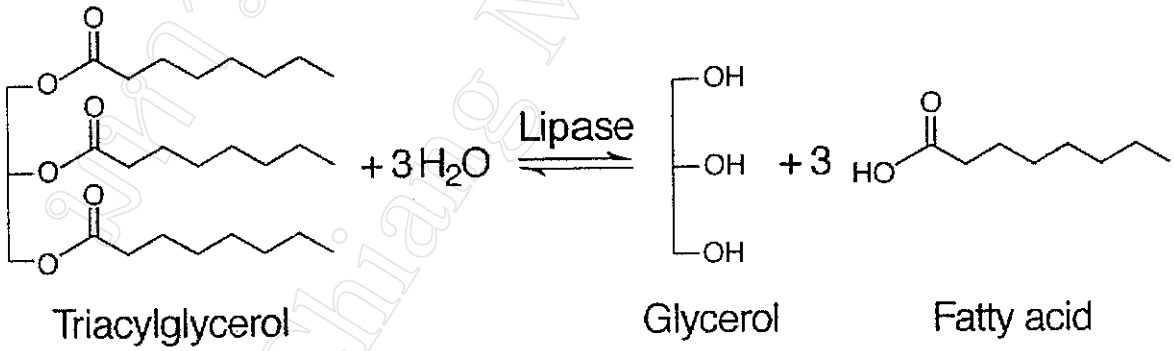


Fig. 1.1. Enzymatic reaction of a lipase catalyzing hydrolysis or synthesis of a triacylglycerol substrate.

Lipases catalyze both the hydrolysis and the synthesis of esters formed from glycerol and long-chain fatty acids (Fig 1). These reactions usually proceed with high

regio- and/or enantioselectivity, making lipases an important group of biocatalysts in organic chemistry. The reasons for the enormous biotechnological potential uses of microbial lipases include the facts that they are (1) stable in organic solvents, (2) do not require cofactors, (3) possess a broadly substrate specificity and (4) exhibit a high enantioselectivity. A number of lipases have been produced commercially, with the majority of them originating from fungi and bacteria.

Recently, two criteria have been used to classified a lipolytic enzyme as a "true" lipase: (a) It should be activated by the presence of an interface, that is, its activity should sharply increase as soon as the triglyceride substrate forms an emulsion. This phenomenon was termed "interfacial activation". (b) It should contain a "lid", which is a surface loop with the interface. However, these obviously suggestive criteria proved to be unsuitable for classification, mainly because a number of exceptions were described of enzymes having a lid but not exhibiting interfacial activation. Therefore, lipase is simply defined as carboxylesterases catalyzing the hydrolysis (and synthesis) of long-chain acylglycerols. There is no strict definition available for the term "long-chain", but glycerolesters with an acyl chain length of ≥ 10 carbon atoms can be regarded as lipase substrates, with trioleoylglycerol being the standard substrate. Hydrolysis of glycerolesters with an acyl chain length of < 10 carbon atoms with tributyrylglycerol (tributyrin) as the standard substrate usually indicates the presence of an esterase. It should be emphasized that most lipases are perfectly capable of hydrolyzing these esterase substrates.

1.1.2) Screening for lipase activity [2]

The methods for measuring the lipase activity can be classified to two main groups following the enzymatic reaction.

1.1.2.1) Hydrolysis

Microbiologists generally want to use a simple and reliable plate assay allowing the identification of lipase-producing bacteria. It has many methods to assay the lipase activity such as titrimetry, spectroscopy (photometry, fluorimetry, infra red), chromatography, radioactivity, interfacial tensiometry, turbidimetry, conductimetry, immunochemistry and microscopy. The most widely used substrates are tributyrin and poured into the petri dish. Lipase production is indicated by the formation of clear halos around the colonies grown on agar plates containing tributyrin and orange-red fluorescence visible on irradiation with a conventional UV hand lamp at $\lambda = 350$ nm on triolein plates, which additionally contain rhodamine B. Lipase activity in bacterial culture supernatants is determined by hydrolysis of *p*-nitrophenylesters of fatty acids with various chain length ($\geq$ C10) and spectrophotometer detection of *p*-nitrophenol at 410 nm. However, care must be taken to interpret the results because these fatty acid monoester substrates are also hydrolyzed by esterases. This problem can be overcome by using the triglyceride derivative 1,2-*O*-dilauryl-*rac*-glycero-3-glutaric acid resorufin ester (available from Boehringer Mannheim Roche GmbH, Germany), yielding resorufin, which can be determined spectrophotometrically at 572 nm or fluorometrically at 583 nm. A number of novel fluorogenic alkyldiacylglycerols were synthesized and used for analysis of both lipase activity and stereoselectivity. A more laboratories but reliable

method for identifying a "true" lipase is the determination of fatty acids liberated from a triglyceride, usually trioleoylglycerol, by titration. Automated equipment allows the parallel assay of a large number of samples. Determination of kinetics of lipolysis requires a tight control of the interfacial quality achieved by using the monolayer technique: A lipid film is spread at the air/way interface in a so-called "zero-order" trough consisting of a substrate reservoir and a reaction compartment. Lipase-catalyzed hydrolysis of the lipid monolayer results in changes of the surface pressure, which can be readjusted automatically by a computer-controlled barostat. It should be emphasized that this technique requires expensive equipment and experienced personnel.

1.1.2.2) Synthesis

Biotechnological applications of lipases prompt a demand for techniques to determine their activity and, if relevant, stereoselectivity. A standard reaction is the lipase-catalyzed esterification of an alcohol with a carboxylic acid, e.g. the formation of octyl laurate from lauric aicd and n-octanol reacted in an organic solvent. The initial rate of ester formation can be determined by gas chromatography. No signal method is available to determine the enantioselectivity of a lipase-catalyzed organic reaction. Generally, the enantioselectivity of product formation is determined either by gas chromatography or high performance liquid chromatography (HPLC), with chirally modified columns. Basically, two types of enantioselective lipase-catalyzed reactions are possible: (a) desymmetrization of prochiral substrates in hydrolysis or acylation reactions and (b) kinetic resolution of racemic mixtures, hydrolysis or acylation again being the two options. Recently, a number of screening methods for

lipases have been reviewed, including those, which allow the conventional determination of enantioselectivity and regionselectivity. However, without modification they are not suitable for high throughput screening.

1.1.3) Catalytic mechanism [3]

Lipases are hydrolases acting the carboxyl ester bonds present in acylglycerols to liberate fatty acids and glycerol. Most active site of lipase consists of a Ser-His-Asp/Glu catalytic triad. This catalytic triad is similar to that observed in serine proteases, and therefore catalysis by lipases is thought to proceed along a similar path as in serine proteases. Hydrolysis of the substrate takes place in two steps (Fig. 1.2).

It starts with an attack by the oxygen atom of the hydroxyl group of the nucleophilic serine residue on the activated carbonyl of the susceptible lipid ester bond. A transient tetrahedral intermediate is formed, which is characterized by a negative charge on the carbonyl oxygen atom of the scissile ester bond and four atoms bounded to the carbonyl carbon atom arranged as a tetrahedron. The intermediate is stabilized by the helix macrodipole of helix C, and hydrogen bonds between the negatively charged carbonyl oxygen atom (the "oxyanion") and at least two main-chain NH groups (the "oxyanion hole"). One of the NH groups is from the residue just behind the nucleophilic serine; the other one is from the residue at the end of strand β3. 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 precisely orients the imidazole ring of the histidine and partly neutralizes the charge that develops on it.

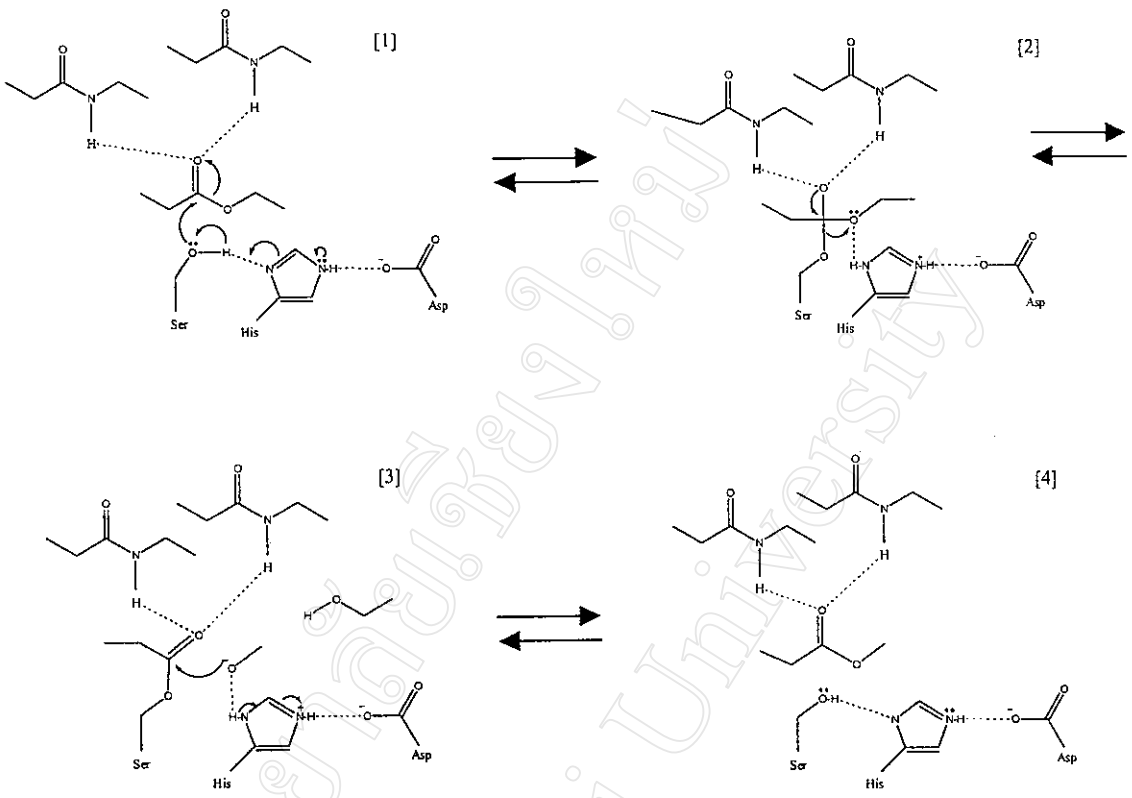


Fig 1.2. Reaction mechanism of lipases. [1] Binding of lipid, activation of nuclophilic serine residue by neighbouring histidine and nucleophilic attack of the substrate's carbonyl carbon atom by Ser O^-. [2] Transient tetrahedral intermediate, with O^- stabilized by interactions with two peptide NH 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 neighbouring histidine residue activates the incoming water molecule, 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.

Subsequently, the proton is donated to the ester oxygen of the susceptible bond, which thus is cleaved. At this stage the acid component of the substrate is esterified to the nucleophilic serine (the covalent intermediate), whereas the alcohol component diffuses away. The next stage is the deacylation step, in which a water molecule hydrolyzes the covalent intermediate. The active site histidine activates this water molecule by drawing a proton in it. The resulting OH^- ion attacks the carbonyl carbon atom of the acyl group covalently attached to the serine. Again, a transient negatively charged tetrahedral intermediate is formed, which is stabilized b interactions with the oxyanion hole. The histidine donates a proton to the oxygen atom of the active serine residue, which then releases the acyl component. After diffusion of the acyl product the enzyme is ready for another round of catalysis.

Evidence for this mechanism has come from various studies, particularly inhibitor binding to lipase and their structural analysis. In addition, a crystallographic analysis of the reaction catalyzed by haloalkane dehalogenase, another α/β hydrolase enzyme, provided definitive evidence for the occurrence of a covalent intermediate.

1.1.4) Application of lipases [3]

For several decades, lipases are already used in industry although the number of applications and therefore their importance to the enzyme manufacturing industry was rather small. It has been estimated earlier that from the worldwide enzyme market of 600 million US$, only some 20 million is accounted for by lipases. The reason for this low interest is most likely the limited availability and relatively high costs of these enzymes, especially for the potential larger applications such as the detergent industry. However, the production technology of lipases has made great progress in

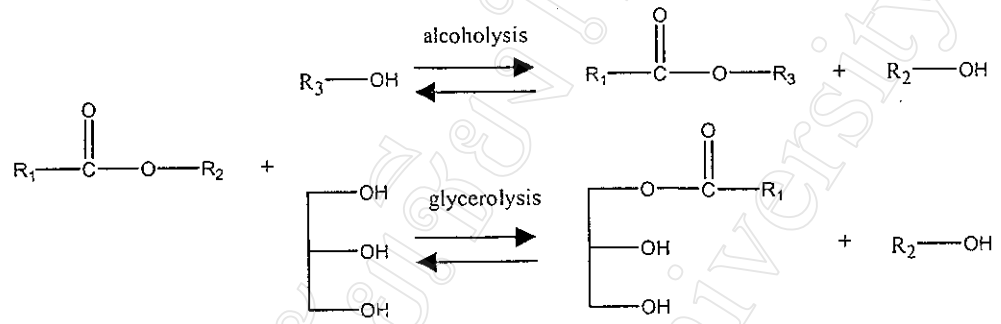
the last 5 years, mainly as a result of recombinant DNA technology by which it is possible to construct microbial strains, which produce different types of lipases in an economically attractive way. As a consequence of this development, many different lipases from as many different microorganisms are now available and especially the detergent industry did benefit from this development and is by now the largest application area of industrial lipases.

1.1.4.1) Hydrolysis *versus* synthesis

The hydrolysis of fats and oils (triacylglycerols) is an equilibrium reaction and therefore it is possible to change the direction of the reaction to ester synthesis by modifying the reaction conditions. The equilibrium between forward and reverse reactions in this case is controlled by the water content of the reaction mixture, so that in a non-aqueous environment lipases catalyze ester synthesis reactions. Different types of synthesis reactions can be distinguished: common ester synthesis from glycerol and fatty acids and the biotechnologically more important transesterification reactions in which the acyl donor is an ester (Fig. 1.3). Transesterifications involving fats and oils can be further specified depending on the type of acyl acceptor. Glycerolysis and alcoholysis refer to the transfer of an acyl group from a triglyceride to either an alcohol or glycerol. In interesterifications, the acyl group is exchanged between a (tri)glyceride and either a fatty acid (also called acidolysis) or a fatty acid ester (more specifically another (tri)glyceride). Interesterifications require a small amount of water, in addition to the amount needed for the enzyme to maintain an active hydrated state. As presence of (too much) water will decrease the amount of

ester synthesis products, the water content should be carefully adjusted to achieve accumulation of desired reaction products.

Transesterification



Interesterification

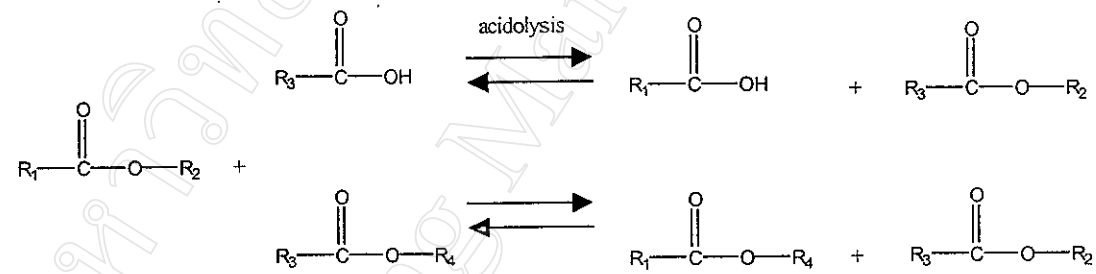


Fig. 1.3. Industrially important reactions catalyzed by a lipase. Tranesterification involves the transfer of an acyl group to an alcohol (alcoholysis) or glycerol (glycerolysis); interesterification describes the transfer of an acyl group to a fatty acid (acidolysis) or a fatty acid ester.

1.1.4.2) Detergents

The great breakthrough in the application of lipase in household detergents came when NOVO/Nordisk launched the product Lipolase™ in 1998 (Table 1.1).

This product contains the extracellular lipase from the fungus *Humicola lanuginosa*, which is produced on industrial scale using *Aspergillus niger* as a host organism. Presently, this product is globally used by at least the two largest detergent manufacturers Procter&Gamble and Unilever.

Table 1.1. Microbial lipases used as additives in household detergent.

Origin of lipase	**Product name**	**Year of introduction**	**Company (location)**
Fungal			
Humicola lanuginosa	Lipolase	1998	NOVO-Nordisk (Denmark)
Bacterial			
Pseudomonas mendocina	Lumafast	1992	Genencor (USA)
Pseudomonas alcaligenes	Lipomax	1995	Gist-brocades (The Netherlands)
Pseudomonas glumae	n.a.	n.a.	Unilever (The Netherlands)
Pseudomonas species	n.a.	n.a.	Solvay (Belgium)
Bacillus pumilus	n.a.	n.a.	Solvay (Belgium)

n.a., no annotation

The second commercially available product is Lumafast™ (genencor International) and contains a bacterial lipase from *P. mendocina* (formerly known as *P. putida*). It differs from other known *Pseudomonas* enzymes as suggested by the lack of any amino acid sequence homology, but seems to be related to the cutinase and hydrolyzes short-chain triglycerides.

Another lipase product for detergents was developed by Gist-brocades (Lipomax™) and is expected to enter the enzyme market in 1995. This product contains the extracellular lipase from *P. alcaligens* (formerly designated *P. pseudoalcaligenes*), the properties of which perfectly match with the conditions of the washing process. The enzyme is active at pH 7-11 and at temperatures up to 60°C hydrolyzing triglycerides with chain lengths varying from C2 to C18. However, the highest activity is observed with longer chains (>C12).

Finally, the patent literature describes several other bacterial lipases, which were screened for detergent applications such as those from *Pseudomonas* and *Bacillus* species.

1.1.4.3) Processing of fats and oils

In the processing of fats and oils, both hydrolysis and synthesis take place. The number of applications in which lipase is used (see Table 1.2); however, it is still limited when compared with the cheaper chemical processes.

1.1.4.4) Hydrolysis

The conventional chemical fat-splitting processes require rather harsh conditions with respect to temperature (240-260°C) and pressure (60 bar). This inevitably produces undesirable side effects, like product discoloration and degradation of some fatty acids. However, due to the cheapness and efficiency of the chemical process, enzyme applications may be economically competitive only in some special cases. Thus far, only in once case application of lipase for common fat or oil hydrolysis on small industrial scale has been reported. The Japanese company

Miyoshi Oil and Fat Co. has used a fungal lipase for the manufacturing of soap. However, fungal lipases are not necessarily the best choice, since *Pseudomonas* lipase(s) were shown to be superior to various fungal lipases in lab scale hydrolysis of beef tallow and castor oil.

Table 1.2. Biotechnological applications of bacterial lipases.

Type of reaction	Origin of lipases	Product (application)	References
Hydrolysis of fats and oils	*Pseudomonas*		(4,5)
Glycerolysis of fats and oils	*Pseudomonas*	Monoacylglycerols (surfactants)	(6-11)
Esterification to glycerol	*Chromobacterium viscosum*		(12,13)
	Pseudomonas fluorescens		(12)
(Trans)esterification to immobilized glycerol	*Chromobacterium viscosum*		(14-15)
Acylation of sugar alcohols	*Chromobacterium viscosum*	Sugar monoacylesters (surfactant) Enrichment of PUFAs	(16)
Acidolysis/Alcoholysis of fish oils	*Pseudomonas*		(17-18)
Resolution of racemic alcohol/esters	*Arthrobacter* *Pseudomonas cepacis*	Building blocks for insecticided/chiral gruds	(19,20)
Polytransesterification of diesters with diols	*Chromobacterium*	Oligomers	(21)
	Pseudomonas	Alkyds (polyester intermediates)	(22,23)
	Pseudomonas	Macrocyclic lactones	(24)
Transesterification of diesters with diols	*Pseudomonas (cepacia)*	Acrylate esters (polyacrylate intermediate)	(25-27)
Intramolecular esterification	*Pseudomonas*	Macrocyclic lactones	(28-29)

1.1.4.5) Glycerolysis

Chemical glycerolysis of fats and oils is used for the commercial production of monoglycerides, which are applied as emulsifiers in a wide range of foods and

cosmetic and pharmaceutical products. It is not likely that this process will be replaced by enzymatic process in the near future although lab scale enzymatic glycerolysis of various fats and oils has been described. Mixtures of mono- and diacylglycerols were formed; high yields of monoglyceride were obtained with *Pseudomonas* lipases. The reaction temperature should be below a certain critical value, which was also dependent on the type of fat or oil. In addition, the water content of the reaction mixture was shown to be critical factor. Several fungal lipases were shown to be inactive.

1.1.4.6) Esterification

Glycerides can also be obtained by direct esterification of free fatty acids to glycerol. However, esterification catalyzed by various microbial lipases always resulted in mixtures of glycerides with yield and composition of the mixture depending on the source of lipase. A process resulting in regioisomerically pure glycerides has been developed comprising as essential step the adsorption of glycerol onto a solid support. Lipase-catalyzed glyceride synthesis with the immobilized glycerol and various acyl donors (e.g. free fatty acids, fatty acid alkylesters, natural fats and oils) yielded milligram quantities of regio-isomerically pure di- and monoacylglycerols. *C. viscosum* was one of the 1,3-selective lipases producing the desires glycerides with high yield. Monoglycerides were separated from the other reactants in a separate vessel and the undesired products were fed back to the reactor.

1.1.4.7) Alcoholysis / acidolysis

Using acylacceptors other than glycerol additional mono-acylcompounds can be synthesized. For example, alcoholysis of sugar alcohols with various plant and animal oils has been shown to yield sugar monoesters of fatty acids. Among various lipases, the enzyme from *C. viscosum* showed good catalytic properties. However, as is the case for fat hydrolysis and glycerolysis, chemical synthesis of sugar esters is far cheaper than enzyme technology, hampering commercial application of enzymes also in this field. On the contrary, refinement of oils containing highly unsaturated fatty acids (PUFAs, poly-unsaturated fatty acids) may be a process with prospects for enzyme application on a commercial scale, because PUFAs are easily subject to decomposition in the chemical process, yielding undesirable oxidation products and polymers. Since fish oils possess poly-unsaturated fatty acids predominantly at the 2-position and this bond is relatively resistant to lipase attack, 1,3-specific lipases can be particularly useful in the concentration of poly-unsaturated fatty acids in monoglycerides. This has been shown in the case of enzymatic alcoholysis of cod liver oil and acidolysis of sardine oil, using various microbial lipases. In both cases, *Pseudomonas* lipase gave the best results.

1.1.4.8) Interesterification

In addition to the Miyoshi enzymatic oil hydrolysis process, a few lipase-catalyzed synthesis reactions in low-water environment have found (limited) application on commercial scale. An example is the transformation of low-value oils, like the palm oil mid fraction, into high-value cocoa butter triglycerides by interesterification. However, since this process is carried out using the immobilized

lipase from *Rhizomucor miehei* it will not be further discussed here. Another class of so-called structured triglycerides which may be a potential product of a lipase-catalyzed interesterification reaction is glycerides containing medium-chain fatty acids on the 1- and 3- positions and essential fatty acids on the 2-position. They can form an alternative for the medium chain triglycerides which are currently used to meet the nutritional needs of patients with maladsorption problems, because shortage of essential fatty acids in these patients can easily occur due to sole consumption of medium-chain lipids.

1.1.4.9) Application in organic synthesis

A literature survey on the application of lipases in organic synthesis reaction reveals an enormous increase in publications during the past few years, especially reactions in non-aqueous media. In nearly all cases, reactions were described on laboratory scale, with commercial applications seldom mentioned. It is not within the scope of this review to mention all these reaction, but the main areas of interest will be discussed.

1.1.4.10) Biocatalytic resolution

So far the most important application of lipases in organic chemistry is the production of optically active compounds. Most frequently, these compounds are produced through the resolution of racemic mixtures of alcohols or carboxylic esters, although stereospecific synthesis reactions are employed as well. Lipase-catalyzed resolution of racemic mixtures can occur through asymmetric hydrolysis of the corresponding esters while in non-aqueous media this approach can be extended to

stereospecific (trans) esterification reactions. In this way, optically active building blocks for insecticide have been obtained by an ester-hydrolysis reaction using *Arthrobacter* lipase. In the synthesis of various chiral drugs such as α-blockers, *P. cepacia* lipase (Lipase PS from Amano) is frequently used enzyme for racemic mixture resolution, via both hydrolysis and acylation reactions. An important factor for the economic feasibility of biocatalytic resolution of racemic mixtures is the recovery of the unwanted enantiomer. Although enantiomer recovery is a commonly applied step after classical racemate separation, this issue has been addressed by only one group in lipase catalytic processes, who described chemical inversion of the unwanted (R)-alcohol into the (S)-form.

1.1.4.11) Polymer synthesis

If, instead of a racemic ester and alcohol (or vice versa), a diester and a diol are used, stereoselective polycondensations occur in organic media. In this way, the formation of optically active trimers and pentamers was observed, using among others a lipase from *Chromobacterium* species.

For the enzymatic synthesis of alkyds, unsaturated diesters are combined with aliphatic or aromatic diols in a polytransesterification reaction using a *Pseudomonas* lipase. No isomerization of the double bond was observed under the mild conditions of the lipase-catalyzed reaction, in contrast to the extensive isomerization found during chemical polycondensation. In a subsequent cross-linking reaction, alkyds can be polymerized to industrially applicable 'general-purpose polyesters'. Several chemoenzymatic processes have been described for the preparation of various polyacylates. After a stereoselective reaction of a racemic alcohol with a (meth)

acrylate ester as acylating agent using a *Pseudomonas* lipase, (meth) acrylate polymers of higher molecular mass could be obtained employing an additional chemical polymerization step. *P. cepacia* lipase or a lipoprotein lipase from *Pseudomonas* species catalyzed the transesterification of various monosaccharides with vinylacrylates, whereupon the resulting sugar-acrylate esters were chemically polymerized. The use of the resulting polymers for biomedical applications and membranes was suggested.

1.1.4.12) Intramolecular esterification

If hydroxyl and ester moieties are present in one molecule, intramolecular esterification occurs, resulting in the synthesis of macrocyclic lactones. C14-C16 macrocyclic lactones are high-grade and expensive substances with a musky fragrance, which are used in perfumes. Upon intramolecular esterification of several hydroxy acids, the yield and ratio of mono- to oligolactone was found to depend on the lipase and on the chain length of the substrate used. In addition, macrocyclic lactones can be synthesized by direct condensation of diacids with diols.

1.1.4.13) Flavor development in food

Traditionally, bacterial lipases produced *in situ* in various food systems have been involved in development of flavor. Lipase from several bacterial species presenting in raw milk (mainly *Pseudomonas*, but also some other like *Alcaligenes* and *Achromobacter*) are known to withstand the pasteurization process and affect flavor development during cheese ripening. In addition, lipases produced by bacterial starters play a role in this process. Other examples of the involvement of lipolytic

lactic acid bacteria in flavor development are vegetable fermentations and ripening of some Italian sausages.

1.2) Recombinant DNA technology [30]

Recombinant DNA technology, which is also called gene cloning or molecular cloning, is an umbrella term that encompasses a number of experimental protocols leading to the transfer of genetic information (deoxyribonucleic acid, DNA) from one organism to another. There is no single set of methods that can be used to meet this objective; however, a recombinant DNA experiment often follows by 1) The DNA (cloned DNA, insert DNA, target DNA, foreign DNA) from a donor organism is extracted, enzymatically cleaved (cut, digested), and joined (ligated) to another DNA entity (cloning vector) to form a new, recombined DNA molecule (cloning vector insert DNA construct, DNA construct). 2) This cloning vector-insert DNA construct is transferred into and maintained within a host cell. The introduction of DNA into a bacterial host cell is called transformation. 3) Those host cells that take up the DNA construct (transformed cells) are identified and selected (separated, isolated) from those that do not. 4) If required, the DNA construct can be manipulated to ensure that the protein product that is encoded by the cloned DNA sequence is produced by the host cell. Although recombinant DNA technology developed from discoveries in molecular biology, nucleic acid enzymology, and the molecular genetics of both bacterial viruses and bacterial extrachromosomal DNA elements (plasmids), the fundamental knowledge that made recombinant DNA technology possible stems from an understanding of the structure and function of DNA.

1.2.1) Plasmid cloning vectors

Plasmids are self-replicating, double-stranded, circular DNA molecules that are maintained in bacteria as independent extrachromosomal entities. Virtually all bacterial genera have plasmids. Some plasmids carry information for their own transfer from one cell to another (F plasmids); others encode resistance to antibiotics (R plasmids); others carry specific sets of genes for the utilization of unusual metabolites (degradative plasmids); and some have no apparent functional coding genes ("cryptic" plasmids). Plasmids can range in size from less than 1 to more than 500 kb. Each plasmid has a sequence that functions as an origin of DNA replication; without this site, it cannot replicate in a host cell.

Some plasmids are represented by 10 to 100 copies per host cell; these are designated as high-copy-number plasmids. Others maintain 1 to 4 copies per cell and are called low-copy-number plasmids. Seldom does the population of plasmids in a bacterium make up more than approximately 0.1-5.0% of the total DNA. When two or more types of plasmids cannot coexist in the same host cell, they are said to belong to single incompatibility group. But plasmids from different incompatibility groups can be maintained together in the same cell. This existence is independent of the copy numbers of the individual plasmids. Some microorganisms have been found to contain as many as 8 to 10 different plasmids. In these instances, each plasmid can carry out different functions and have its own uniques copy number, and each belongs to a different incompatibility group. Some plasmids, because of the specificity of their origin of replication, can only replicate in one specific species of host cell. Other plasmids have less specific origins of replication and can replicate in a number of

bacterial species. These plasmids are called narrow- and broad-host-range plasmids, respectively.

1.2.2) Transformation and selection

The next step in a recombinant DNA experiment requires the uptake by *E. coli* of the cloned DNA. In bacteria the process of introducing purified DNA into a cell is called transformation. For *E. coli*, one of the methods of transformation requires that the cells be treated with a regimen high temperature and calcium chloride. But transformation is an inefficient process, with typically no more than one cell in a thousand being transformed. In other words, following transformation, most of the cells have not acquired a new plasmid. Furthermore, a few cells are transformed by recircularized plasmid DNA that escaped dephosphorylation by alkaline phosphatease; other acquires nonplasmid DNA, while a few are transformed by the plasmid-insert DNA construct.

The extrachromosomal DNA that lacks an origin of replication cannot be maintained with a bacterial cell. Thus, the uptake of nonplasmid DNA is of no consequence in a recombinant DNA experiment. To ensure that a plasmid-cloned DNA construct is perpetuated in its original form, the *E. coli* host cells that are used must lack the genes for restriction endonucleases and are usually not capable of carrying out exchanges between DNA molecules-because they have been made recombination negative ($RecA^-$). A plasmid-DNA insert construct is less likely to be degraded in a bacterium that does not produce restriction endonucleases. And the insert DNA will not become altered due to homologous recombination events in a recombinant deficient host cell.

After the transformation step, it is necessary to identify, as easily as possible, those cells that contain plasmid-cloned DNA constructs.

1.2.3) Polymerase Chain Reaction (PCR) [31-33]

Polymerase Chain Reaction (PCR) is an ingenious new tool for molecular biology that has had an effect on research similar to that of the discovery of restriction enzymes and the Southern blot. PCR is so sensitive that a single DNA molecule has been amplified, and single-copy genes are routinely extracted out of complex mixtures of genomic sequences and visualized as distinct bands on agarose gels. PCR can also be utilized for rapid screening and/or sequencing of inserts directly from aliquots of individual phage plaques or bacterial colonies. PCR is an effective procedure for generating quantities of a specific DNA sequence in vitro. This amplification, which can be more than a million-fold, is achieved by a three step cycling process. The essential requirements for PCR are (1) two synthetic oligonucleotide primers (~20 nucleotides each) that are complementary to regions on opposite strands that flank the target DNA sequence and that, after annealing to the source DNA, have their 3'-hydroxyl ends oriented toward each other; (2) a target sequence in a DNA sample that lies between the pair of primers which can be from 100 to 5,000 bp in length; (3) a thermostable DNA polymerase that can withstand heating to 95°C or higher; and (4) the four deoxyribonucleotides.

A typical PCR process entails a number of cycles for amplifying a specific DNA sequence; each cycle has three successive steps.

1. Denaturation. The first step in the PCR amplification system is the thermal denaturation of the DNA sample by raising the temperature within a reaction tube to

95°C. In addition to the source DNA, this reaction tube contains a vast molar excess of the two oligonucleotide primers, a thermostable DNA polymerase (e.g., *Taq* DNA polymerase, isolated from the bacterium *Thermus aquaticus*), and four deoxyribonucleotides. The temperature is maintained for about one min.

2. Annealing. For the second step, the temperature of the mixture is slowly cooled to ~55°C. During this step, the primers base pair with their complementary sequences in the source DNA.

3. Extension. In the third step, the temperature is raised to ~75°C, which us optimum for the catalytic functioning of *Taq* DNA polymerase. DNA synthesis is initiated at the 3'-hydroxyl end of each primer.

There are many ways that PCR can be used. For example, PCR is an effective procedure for detecting the presence of a known DNA sequence in very small, crude samples. Purification of the target DNA is unnecessary. For this reason, PCR can be used to determine whether a particular illness is due to a viral infection. When the sequence of the viral DNA has previously been determined, the investigators can synthesize a pair of primers that anneal to sites in the targeted viral DNA that are a specified number of bases apart. After PCR cycling, a DNA fragment of a specific size (a length equivalent to two primer lengths plus the length of the viral DNA between them) will be amplified only if the viral DNA is in the sample.

1.2.4) Mutagenesis [34-37]

In vitro mutagenesis is used to change the base sequence of a segment of DNA. The changes may be localized or general, random or targeted. More catholic and less specific methods of mutagenesis are better suited to analysis of regulatory

regions of genes, whereas more precise types of mutagenesis are used to understand the contribution of individual amino acids, or groups of amino acids, to the structure and function of a target protein. Both methods share the virtue of generating mutants in vitro, without phenotypic selection.

For this research, the PCR-mediated site-directed mutagenesis was used for testing the role of particular residues in catalytic site of a protein. The following are important advantages of PCR-based methods for site-directed mutagenesis; for example, high rates of recovery of mutants, ability to use double-stranded DNA templates, use of high temperatures, development of methods, availability of commercial kits, speed and ease, relatively high rate of errors in PCR products, introduction of unwanted nucleotides at the 3′ termini of amplified DNAs, large number of primers and amplification reactions required for each mutagenesis experiment by some PCR-based methods, requirement to optimize the conditions for PCR for each new set of primers and/or template, high frequency of unmutagenized clones resulting from contamination of the PCR-amplified and inefficiency in amplifying DNA fragments longer than 2-3 kb by standard PCR. Of the many published variants of PCR-based mutagenesis, two stand out for their durability and robustness: overlap extension mutagenesis and megaprimer mutagenesis. In overlap extension mutagenesis, two overlapping DNA fragments are amplified in separate PCRs. The mutation of interest is constructed in the region of overlap and is present in both amplified fragments. The overlapping fragments are mixed and, in a third PCR are amplified into a full-length DNA using two primers that bind to the extremes of the two initial fragments. The method is surprisingly effective, but it requires two mutagenic primers, two flanking oligonuclotides, and three PCRs to construct a

mutation. For the megaprimer method, it is the simplest and most cost-effective method of PCR-based mutagenesis currently available. The method involves two rounds of PCR that employ two flanking primers and one internal mutagenic primer containing the desired base substitutions. The flanking primers can be complementary to sequences in the cloned gene or to adjacent vector sequences. The mutagenic primer can, in theory, be oriented toward either of the flanking primers. In practice, however, the mutagenic primer is always oriented toward the nearer of the two flanking primers so that the length of the megaprimer is kept to a minimum.

1.3) Protein Purification [38]

The first step in the protein purification is the preparation of an extract containing the protein in a soluble form by using clarification or extraction techniques. The two most commonly used techniques for clarification are centrifugation and microfiltration. The precipitation is usually used as a fairly crude separation step often during the early stages of a purification procedure and it can be used as a method of concentration proteins prior to purification step as the same ultrafiltration and lyophilization method. Most purification schemes involve some from of chromatography. Frequently, it is necessary to remove salts or change the buffer after one step in the purification for the next step to work efficiently. This is often achieved by dialysis. The separation of protein on the basis of their charge is ion-exchange chromatography that the separation is based on the reversible interaction between a charged protein and an oppositely charged chromatographic medium while gel filtration is a form of partition chromatography used for separating molecules of different sizes. Furthermore, a striking characteristic of many proteins is

their ability to bind specific molecules tightly but not covalently by affinity chromatography.

A range of electrophoretic techniques are available which can separate protein on the basis of one or a combination of their three major properties, size, net charge and relative hydrophobicity. The western blotting procedure is application techniques from electrophoresis.

1.4) Literature reviews

Lipases are a family of enzymes that catalyze the hydrolysis of the ester bonds of triglycerides and under certain conditions, the synthesis of ester bonds via transesterification. Therefore, lipases are useful for many industries. Research on thermostable lipase is very interesting and they are found in many microorganisms, such as *Bacillus* sp. [39-41], *Pseudomonas* sp. [42-47], *Chromobacterium viscosum* [48], *Rhizopus rhizopodiformis* [49], *Humicola lanuginose* [50-51], *Thermomyces lauginosus* [52] and *Trichosporon asteroids* [53]. However, the production of the lipases is not enough for the industrial applications. So, the protein engineering approach will be used to optimize the lipases. For example, Dannert *et* al. (1997) [39] reported that the lipase gene from *Bacillus thermocatenulatus* in *E. coli* DH5α host cell by the use of the expression vector pUC18 with *Sau* 3A digestion yielded a cloned coding for a molecular mass of 43.09 kDa containing nucleotides 2.9 kb and 388 amino acids. When the lipase gene was changed the expression vectors and host cells, it was found that these expression levels represented a 12-15 fold improvement over the lipase expression under the control of the native promoter. The OmpA presequence upstream of the lipase gene resulted in a 3-fold increase in lipase expression when

compared with pT-BTL and pT-preBTL. A further 20-fold increase in the expression of soluble lipase was obtained when pT-OmpABTL was introduced in *E. coli* JM 105. Tan and Miller (1992) [54] reported that the lipase gene from *Pseudomonas fluorescens* B52 was subcloned in various plasmids such as pUC19, pYPT1, pYPT2, pYPT3 and pYPT4, and expressed in *E. coli* DH5α. The *E. coli* DH5α transformant containing plasmid pYPT4 produced the lipase activity at 37°C for 24 h more than the others containing plasmids pUC19, pYPT1, YPT2 and pYPT3 about 72, 58, 12 and 1.5 fold, respectively, and it produced the highest lipase activity for 96 h. The DNA sequence of the lipase structural gene contained 1,428 nucleotides and 476 amino acids with a molecular mass of 50,241 Da. The substrate-binding region of the lipase contained a Gly-X-Ser-X-Gly sequence, which was similar to the other lipase from many microorganisms such as *Staphylococcus hyicus* [55], *Staphylococcus aureus* [56], *Pseudomonas* nov.sp.109 [57], *Pseudomonas cepacia* [58], *Rhizopus delemar* [59], *Rhizopus miehei* [60] and *Geotrichum candidum* [61]. Kugimiya *et* al. (1992) [62] reported that the NH_2-terminal amino acid sequence of *Rhizopus niveus* lipase (RNL) was found by Edman degradation to be Asp-Asp-Asn-Leu-Val-Gly-Gly-Met-Thr-Leu-Asp-Leu-Pro-Ser-Asp-Ala-Pro-Pro-Ile-Ser-, and its molecular mass was estimated to be about 34 kDa from the mobility relative to markers by SDS-PAGE. Complementary DNA encoding RNL was isolated from *Rhizopus niveus* IFO4759. The transformant, *E. coli* JM109 containing pRL1068B showed the lipase activity on a tributyrin agar plate. It consisted of 958 nucleotides and 297 amino acids, which the amino acid sequence of RNL showed 56% homology to that of *Rhizomucor miehei* lipase (RML). The NH_2- and COOH- terminal positions were found at Asp and Leu, respectively. Furthermore, several thermostable lipases have been studied the cloning

and expression from thermophilic microorganisms such as *Pseudomonas fluorescens* SIK W1 [43], *Pseudomonas glumae* [64], and *Pseudomonas* sp. KWI-56 [65].

After cloning and expression of the lipase genes, the purification and characterization of lipases are very interesting to study extensively. Kim *et* al. (1994) [41] reported that the thermostable lipase from thermophilic *Bacillus* sp. 398 was purified by ammonium sulfate precipitation, DEAE-sepharose column, Butyl-Toyopearl column and DEAE-sepharose column. The purified lipase was 10,300-folds increase in the specific activity, which had a specific activity of 25,300 units/mg protein when assayed at 55°C with olive oil as the substrate. The purified lipase gave a single band by SDS-PAGE and its molecular mass was estimated to be 50 kDa. The optimum pH and temperature for its activity of the purified lipase were 8.2 and 65°C, respectively. It was fairly stable in the range of pH from 4.0 to 11.0, and it did not lose activity upon incubation at 60°C for 30 min. Sugihara *et* al. (1991) [65] reported that a thermostbale lipase from *Bacillus* sp. has been purified to homogeneity as judged by disc-PAGE, SDS-PAGE, and isoelectric focusing. The purification included ammonium sulfate fractionation, treatment with acrinol, and sequential column chromatographies on DEAE-sephadex A-50, Toyopearl HW-55F, and Butyl Toyopearl 650M. The enzyme was purified 7,760-folds starting form the culture supernatant with a yield of 9%. It was found to be a monomeric protein with a molecular mass of 22,000 Da, and pI of 5.1. The optimum pH at 30°C, and optimum temperature at pH 5.6 were 5.5-7.2, and 60°C, respectively, when olive oil was used as the substrate. The lipase was incubated with Cu^{2+}, Zn^{2+} and Hg^{2+} and showed 70% inhibition. The addition of acetone to the assay mixture in the range of 0-60%(v/v) stimulated the enzyme remarkably, whereas n-hexane had an inhibitory effect. It

consisted of 20 amino acids with the maximum threonine content (10.4%) and 27.4% of hydrophobic residues (Val, Ile, Leu, Tyr and Trp). Kordel *et* al. (1991) [47] reported that the extracellular lipase from *Pseudomonas* sp. ATCC21808 was purified by Q-sepharose column chromatography in the presence of n-octyl-β-D-glucopyranoside, Ca^{2+} precipitation of fatty acids, and Octyl-sepharose chromatography which had a purity 260-fold to a yield of 35%. The molecular mass was determined as 35,000 Da; a polyacrylamide gel under nondenaturing conditions revealed a band at 111,000. In hydrolysis of triglycerides, the lipase showed substrate specificity for saturated fatty acids from C6-C12 and unsaturated long-chain fatty acids. The NH_2-terminal sequence of the lipase contained 24 amino acids. For crystallization of the lipase by using the difference salts, 2-methyl-2,4-pentanediol, and polyethylene glycols as precipitants, the crystals appeared within 24 to 48 h at room temperature as hexagonal plates. The crystal diffracted to a resolution of about 0.25 nm. Procession photographs revealed that it belong to space group C2 with lattice constants of a = 9.27 nm, b = 4.74 nm, c = 8.65 nm, and β = 122.3°, indicating a cell constant of one molecule per asymmetric unit of the crystal. Sugihara et al. (1992) [46] reported that a thermostable lipase from *Pseudomonas cepacia* was purified by acrinol treatment, Macro-Prep methyl hydrophobic interaction chromatography and gel filtration (Sephacryl S-100HR). The specific activity was 5,040 U/mg that increases 1,120-fold from the culture supernatant with a yield of 54%. The purified lipase was a monomeric protein with a molecular mass of 36,500 Da and pI of 5.1. The optimum pH at 50°C and optimum temperature at pH 6.5 were 5.5-6.5 and 55-60°C, respectively, when olive oil was used as the substrate. Simple triglycerides of short and middle chain fatty acids (C≤12) were the preferred substrates over those of long chain fatty acids. The

enzyme cleaved all the ester bonds of triolein, with some preference for the 1,3-ester bonds. The enzyme retained all its activity even after incubation at 75°C (pH 6.5) for 30 min. In addition, the activity was not impaired during 21 h storage at pH 6.5 in 40% water-miscible solvents including methanol, ethanol, acetone, acetonitrile, dimethylformamide, dimethylsulfoxide and dioxane. The addition of dimethylsulfoxide or acetone to the assay mixture in the range of 0-35% stimulated the enzyme, whereas benzene or n-hexane had an inhibitory effect. For the amino acid sequence determination, the enzyme consisted of 20 amino acids, which had the maximum alanine content (12.7%), and the NH_2-terminal sequence of *Pseudomonas cepacia* lipase, together with those of lipases from the other *Pseudomonas* species [43,47,57,58,67-69].

Computer is an essential role to determine the three dimensional structure of the enzyme due to the molecule of protein consists of many amino acids. So, the determination of amino acid sequence is derived from a lot of small polypeptides. Computer can arrange in order to obtain the amino acid sequence quickly and reduces the partially experimental steps. It is widely used; for example, Vasel *et* al. (1993) [70] used the computer to study a model of a complex of lipase from *Rhizomucor miehei* with trilaurylglycerol. The amphiphilic helix was moved towards a position where the polar groups point into a groove of the original molecule and the non-polar groups of the helix was exposed, or changed the accessible surface of some residues in the environment of the lid and the catalytic triad from inactive form to active form. Addition to, computer is used for the comparison of the molecular structure of the enzyme, amino acid sequence, the function and characterization of the similar enzyme groups that indicated the function of each amino acid and the application to modify

the molecular structure of the enzyme, which the modification are exported in a two-methods that are chemical modification and genetic engineering. For example, Kawase *et* al. (1990) [71] reported that in order to improve the heat stability of the purified lipase OF360 from *Candida cylindracea*, lysine residues of the lipase were modified with a heterobifunctional photogenerated reagent, N-hydroxysuccinimide ester of 4-azidobenzoic (4-N_3-BzONSu), and photolyzed subsequently in the presence of decanol, which was proved to be a suitable heat stabilizer. The modified lipase showed excellent heat stability, even in this absence of decanol in the reaction system, compared to the native lipase. The activities of the modified lipase were retained about 70% after heating at 500°C for 15 min, while the native lipase lost about 80% of the original activities. The stabilization thus observed might be mainly caused by *in situ* formation of a cross-link between decanol and the lipase, and partly attributed to intra-molecular cross-linking in the lipase. It was found that 8.2 of lysine residues per one molecule of the lipase were modified with 4-N_3-BzONSu. Thus, the modified lipase OF360 showed excellent heat stability even in the absence of stabilizers such as decanol in the reaction system. On the other hand, the genetic engineering is limited to use with the protein that had been known the amino acid sequence and the three dimensional structure. The example of genetic engineering is the site-directed mutagenesis that is an one-method of protein engineering and used to modify protein or enzyme followed the requirement such as the thermostability, denature stability, or changing the substrate specificity and increasing the reaction system. For this example, Kok *et* al. (1995) [72] changed the active site Ser-99 of the lipase from *Acinetobacter calcoaceticus* BD413 to an Ala-99, via site-directed mutagenesis, resulting in production of an in active extracellular lipase.

In the preliminary research, Boonsinthai *et* al. (1997) [73] found that the thermostable lipases from the thermophilic bacteria strain P1, TP404, TP811 and TLS63 were isolated from hot spring in Chiang Mai, Thailand. The crude lipases from these bacteria had the optimum pH and temperature 6.0-8.0 and 50-80°C, respectively, with the maximum activities at pH 7.2 and 65°C. The thermostabilities of the lipases were determined over a pH range of 4.0-10.0 at 37°C and assessed at 70°C after pre-incubated over a range of temperature for 1 h. However, the productions of the lipases were limited. So, the biochemical research group at Mahidol University had cloned the thermostable lipase genes from the thermophilic bacteria strain P1, TP404 and TP811 in plasmid pUC19 and expressed in *E. coli* DH5α as host cell. The cloned lipase genes from these bacteria produced the lipase activities more than the native lipases about 2 times but the lipase productions were low quantities and were not enough to continue study [74]. Therefore, this research emphasizes in optimization of the thermostable lipase from the thermophilic bacteria *Bacillus* sp. strain P1 starting from the determination of the DNA sequence of the cloned lipase gene to find the suitable restriction site, and then changing the plasmid or host cell to high expression of the recombinant lipase and high production of lipase. After that, the recombinant lipase was purified and characterized the biochemical properties. In addition, the amino acid sequence was determined, studied the relationship between the molecular structure and the function of the enzyme, and modified the molecular structure of the enzyme to increase the enzyme production or improve the properties of the enzyme for industrial applications.

1.5) Purpose of the study

This study is to optimize the expression of cloned lipase gene from a thermophilic bacterium strain P1 in suitable host cell, purify the recombinant lipase to high purity and yield, characterize the purified lipase, study the amino acid sequence and correspond to understand the relationship between the molecular structure and the function of the enzyme, and modify the molecular structure of the enzyme for suitable industrial applications.