CHAPTER II
LITERATURE REVIEW

1. Botanical description of *Croton stellatopilosus* Ohba

As described in the web site of Flora of Thailand (www. nationalherbarium.n/thaieuph/ThSearchUse.htm), *Croton stellatopilosus* Ohba (Fig. 2) is a kind of shrub to 6 m tall, branching from base; younger parts distinctly pubescent. *Indumentum* consisting of stellate-dendritic, cream-yellowish-brown hairs with a slightly darker center. *Leaves* obovate , margin serrate, apex acute to short acuminate, brighter below, basal glands sessile. *Inflorescences* whitish, 7-10 cm long, the staminate in the apical part often still in bud when the lower pistillate part is already in flower or even in fruit. *Staminate flowers* densely pubescent throughout; pedicel 2-6 mm long; sepals and petals both 2.5 by 1 mm; stamens 10. *Pistillate flower* densely pubescent throughout; pedicel 4-6 mm long; sepals 3 by 1.5 mm; petals not seen; stigmas 3 mm long, undivided in the lower half and once divided apically. *Fruits* 5 by 6 mm, smooth to very slightly muriculate, sulcate, quite thin-walled, sparsely pubescent to subglabrous. *Seeds* 4 by 2.5 mm, with a very small caruncle. Flowering and fruiting on December to February.

Distribution areas of this plant in Thailand appear to be in Lop Buri (Khao Thungna), Sa Kaeo, Prachin Buri (Kabinburi), Chachoengsao, Chon Buri (Khao Khieo).

The name of this plant has been changed recently based on its taxonomic characteristic. It was originally identified as *Croton sublyratus* by Airy Shaw (1972) by comparing with the collections from Andaman Island of India. However, the samples of Thai collections differ in at least two characters from the Andaman plants: the basal leaf glands are sessile and flat (distinctly protruding to nearly stipitate in *C. sublyratus*), and peculiarly cone-like, dense, pyramidal inflorescence buds (not found on the Andaman Island). Also the leaves are generally slightly smaller in Thailand. These differences have been used as the major points to separate plants from the two countries. *Croton stellatopilosus* has, therefore, been described from south-eastern Thailand and is the correct name for the source plant of plaunotol in the strictest sense (Esser and Chayamarit, 2001).

Figure 2 Various parts of *Croton stellatopilosus* Ohba: (A) Leaves; (B) Flowers; (C) Fruits

2. Thai Folk medicinal uses

C. stellatopilosus (Plau-Noi) is a Thai medicinal plant that has been used traditionally as an anthelmintic and a dermatologic agent for skin diseases (จุฬาลงกรณ์ มหาวิทยาลัย, คณะเภสัชศาสตร์, ภาควิชาเภสัชพฤกษศาสตร์, 2530; Ponglux. *et al.*, 1987). The parts of stem, bark and leaf have been used as an antidiarrheal and to normalize mentruation, whereas its flower has been used as an anthelmintic (มหิดล, มหาวิทยาลัย, คณะเภสัชศาสตร์, 1990). Firewood of Plau-Noi has been used for postpartum (เปรมจิต นาคประสิทธิ์, บรรณาธิการ, 2526). In addition, Plau-Noi and Plau-Yai *(C. oblongifolius* Roxb.) have been used together for various purposes such as stomachic, anthelmintic, emmenagogue, digestant, transquilizer and carminative. They also have been used for the treatment of lymphatic, pruritic, leprosy, tumor and yaws (ประเสริฐ พรหมมณี และคณะ, 2531; นันทวัน บุณยะ ประภัศร, 2532)

3. Bioactive compounds in *Croton stellatopilosus*

Plaunotol, the antipeptic ulcer drug (Kelnac), is an acyclic diterpene alcohol which was registered with the World Health Organization (WHO) under the code CS-684 (Ogiso *et al.*, 1985). Its tradename is Kelnac which has been manufactured by Sankyo Co., Ltd., Japan. Plaunotol was first isolated by Ogiso et al. (1978) by antiulcer activity guide fractionation of the acetone extract of the stems. The yield is 53 g per 81.5 kg dried stem.

Other compounds with antiulcer activity were tested by Kitazawa et al. in 1980, they are diterpenelactone, namely, Plaunol B, plaunol C, plaunol D, plaunol E with less yield from stem (1.2 g, 1.5 g, 0.5 g, 4.1 g, respectively from 81.5 kg dried stem).

Recently, the extract from *C. stellatopilosus* combined with surfactant has been patented for the use as shampoo for animals (Mamiya and Iso, 2003a) and the topical compositions containing fungicides 1%wt bifonazole and 10%wt (*E,Z,E*)-7-hydroxymethyl-3,11,15-trimethyl-2,6,10,14-hexadecatetraen-1-ol from *C. stellato pilosus* showed treatment of onychomycosis (Mamiya and Iso, 2003b). The chemical structures of those bioactive compounds from *C. stellatopilosus* were shown in Fig. 3.
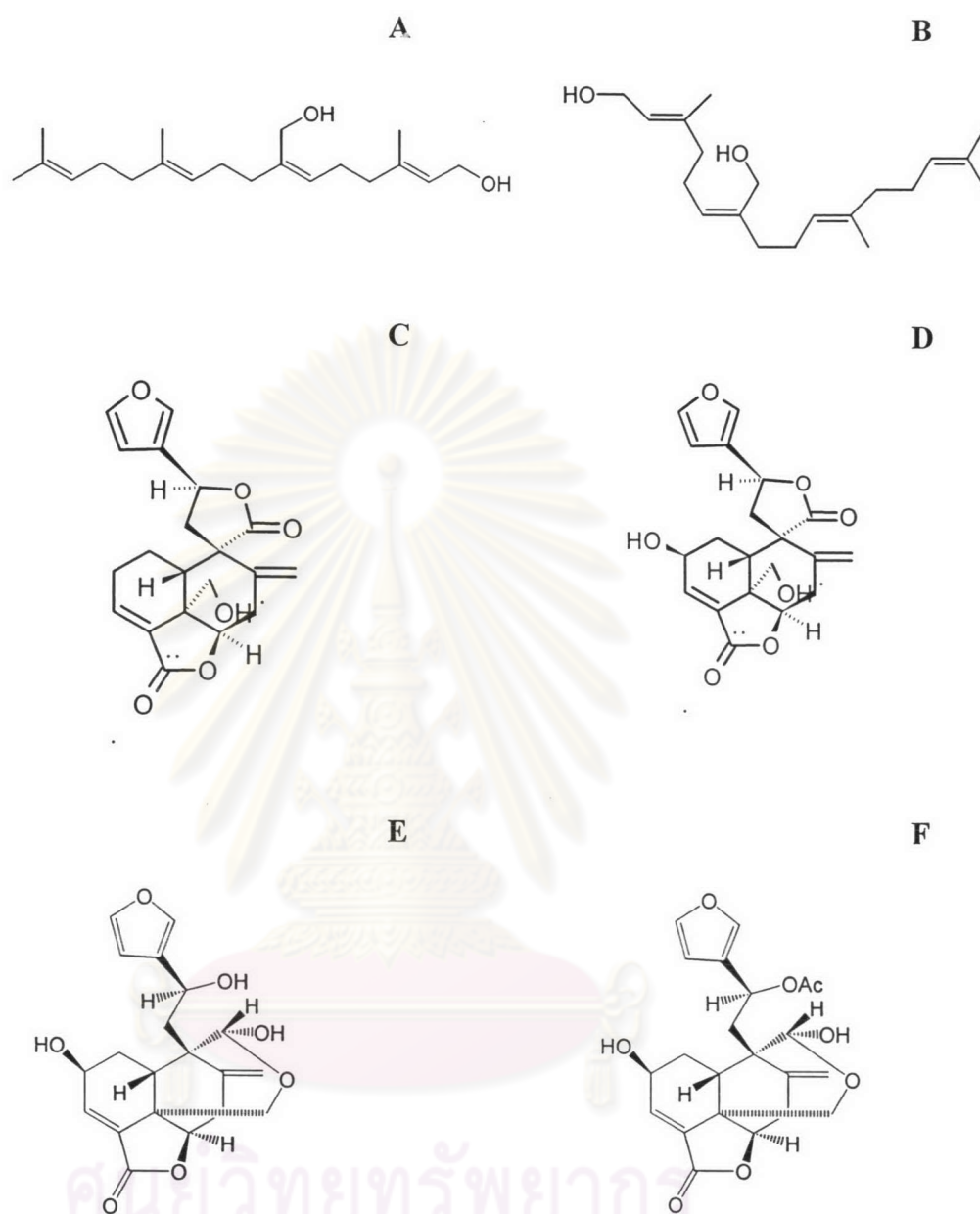
A

B

C

D

E

F

Figure 3 Bioactive compounds in *C. stellatopilosus*:
(A) Plaunotol, (B) (*E,Z,E*)-7-hydroxymethyl-3,11,15-trimethyl-2,6,10,14-hexadeca tetraen-1-ol, (C) plaunol B, (D) plaunol C, (E) plaunol D, (F) plaunol E.

4. Various techniques used for obtaining plaunotol

Among the occurrence of 15 species of *Croton* available in Thailand summarized by Tansakul 1998, it has been shown that only one species of *C. stellatopilosus* that contain plaunotol (Ogiso et al., 1981 and 1985; Vongcharoensathit

and De-Eknamkul, 1998). The quantitative analysis of plaunotol in Plau-Noi leaves in Thailand has been reported by Vongchareonsathit 1998. The content of plaunotol is in the range of 0.14-0.79 % w/w. Due to the importance of plaunotol for medicinal use, various studies have been carried out to evaluate the possibility alternative sources of plaunotol and GGOH. These studies include:

1. Cell suspension culture (Kitaoka et al., 1989). They found that no production of plaunotol but only GGOH accumulated in the lag or stationary phase of cell growth at 0.05% dry weight.

2. Callus culture (Morimoto in 1988). It was found that plaunotol is present 0.17% dry weight.

3. Chemical synthesis (by Ogiso et al., 1978; Ogiso et al., 1985; Sato et al., 1988; Takayanaki, 1994; and Tago et al., 2000).

4. Improvement of the extraction method from crude extract of *C. stellatopilosus* by Nilubol, 1993. The yield of plaunotol is 15-17 g from 6 kg dried ground leaves.

Biotechnology is one of the effective methods to improve secondary metabolite production. In the field of secondary plant products, it generally begins from the study of biosynthetic pathway in order to confirm the proposed biosynthetic pathway and to identify enzymes, and genes involved in the biosynthetic pathway. In addition, recombinant DNA techniques are capable of directly modifying the expression of genes related to biosynthesis. It is now also possible to manipulate the pathways that lead to secondary plant compounds (Bourgaud et al., 2001). The knowledge can lead to the final step of metabolic engineering to increase the product by manipulate in gene level of the enzyme limiting step (Bourgaud et al., 2001; Broun and Someville, 2001; Chen et al., 2000).

5. Methods used in biosynthetic study of secondary metabolites

Various methods used in biosynthetic studies of secondary metabolites have been given by Luckner, 1990. It can be summarized as follow:

1. Methodology of biosynthetic studies use of isotope-labeled compounds (tracer technique, feeding experiment), such as the administration (feeding) of putative precursors labeled with isotopes at one or more specific positions

to organisms producing secondary products, then isolate the secondary compounds after a suitable period of time, and determine whether they contain any of the isotopes enriched in the administered precursor. This technique could detect the intermediates and elucidate of mechanistic details of biosynthetic pathways. The isotopes used are ^{2}H, ^{3}H, ^{13}C, ^{14}C, ^{15}N, ^{18}O, ^{32}P and ^{35}S.

2. Use of enzyme preparations (cell-free preparation): In tracer experiments, administered substances have to pass many barriers before reaching the site of the biosynthesis of secondary products, so the incorporation rates are often low and sometime misleading. Most of the problems caused by compartmentalization may be avoided by using homogenates, certain fractions of homogenates, or enzyme preparations instead of whole cells. However, *in vitro* experiments with enzymes of secondary metabolism are still difficult because of small amount present and they may denature and inactivated after disintegration of their sites of storage.

Before starting enzymatic work it is necessary to have basic information about the pathway in question, i.e., about the intermediates and types of reactions. This knowledge may come from tracer experiments which are usually a prerequisite of successful enzymatic work, and from general experience on reactions of secondary product formation, i.e., on well-founded knowledge of the biochemistry of secondary metabolism.

6. Previous biosynthetic study in *C. stellatopilosus*

1. Feeding experiments

1.1 Feeding of [1-^{14}C] glucose into the leaves of *C. stellatopilosus* resulted to a low incorporation of the label into plaunotol. This may be due to the complication of site which occurs in chloroplast (Potduang, 2000).

1.2 Feeding of *C. stellatopilosus* callus with [1-^{13}C] glucose showed the incorporation of the label into phytosterols with mix origin of isoprene units from both mevalonate pathway and non-mevalonate pathway (De-Eknamkul and Potduang, 2003).

1.3 Feeding the whole leaf of *C. stellatopilosus* with [1-^{13}C] and [U-^{13}C] glucose lead to an incorporation of the label to plaunotol exclusively by the non-mevalonate pathway (Wungsintaweekul and De-Eknamkul, 2005).

2. Cell-free extract

The fraction of 20,000 g pellet from fresh leaves of *C. stellatopilosus* was studied for the enzyme activities utilizing substrate $[1-^3H]GGDP$ (Tansakul and De-Eknamkul, 1998). It showed clearly that the amount of GGOH was formed and utilized for plaunotol formation. Therefore, the biosynthetic pathway of plaunotol was proposed to be involved two enzymes, namely, phosphatases and GGOH-18-hydroxylase.

7. Compartmentalization of diterpene biosynthesis in plants and in *C. stellatopilosus*

The location of the biosynthetic site of secondary products of interest is important for choosing organ of plant and the organelle in cells to study its enzymes and genes involved in their biosynthetic pathway. Diterpene compounds in plant are believed to be biosynthesized in plastid, a double membrane-surround compartment (Trapp and Croteau, 2001) (Fig. 4). Chloroplast is commonly composed of the secondary products and enzymes of secondary metabolism. Lipophilic secondary compounds are found either as constituents of the membranes (chlorophylls) or are located in the plasm of plastids (the stroma), as lipophilic droplets-like carotenoids (Luckner, 1990). From Electron microscope of palisade mesophyll cell of *C. stellatopilosus* leaves, the oil droplets were found in plastid as shown in Fig. 5 (Sittithaworn, 2001; Potduang, 2000).
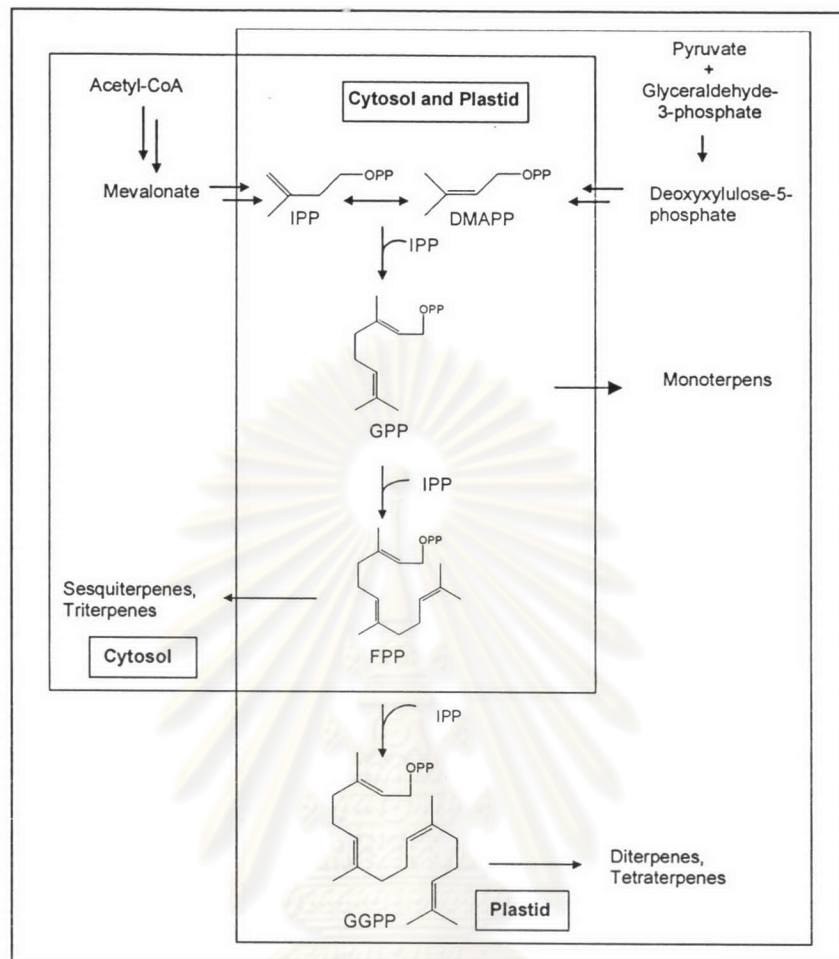
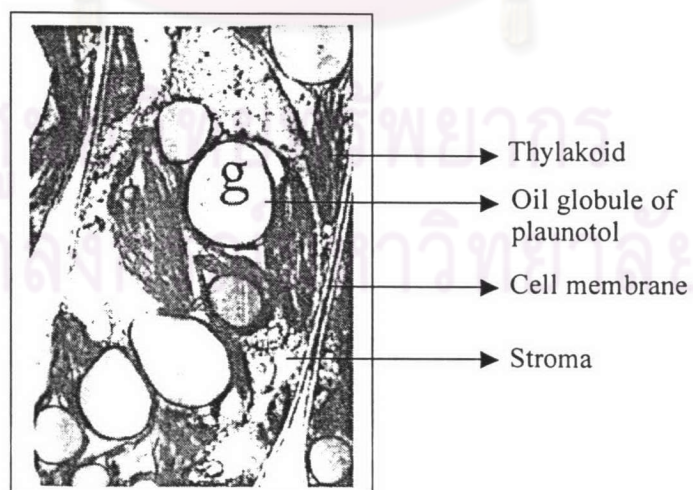**Figure 4** Compartmentation of terpenoid biosynthesis (Trapp and Croteau, 2001)

Figure 5 Electron micrograph of the palisade mesophyll cell x 3,800 time. The oil globules of plaunotol located in chloroplast (Potduang, 2000). (g = oil globule)

8. Phosphatases: Clarification

As review previously (Widlanski and Taylor, 1999), phosphatases is a group enzymes that catalyze the hydrolysis of phosphoester and phosphoanhydride bonds of a diverse set of substrates (Boyer et al., 1961). They are generally classified based on substrate specificity, mechanism of action, and pH optimum for catalytic activity (e.g. acid and alkaline phophatase).

Based on their substrate activity, phosphatases can be divided into 3 major groups;

1. Nonspecific: a mechanistically broad class of enzymes including both alkaline and acid phosphatases.

2. Phosphoprotein specific: prefer phosphoproteins or phosphopeptides as substrates such as serine/threonine phosphatases, protein tyrosine phosphatases

3. Small molecule specific: specific for one (or a related group of structurally similar) substrate(s). GGDP phosphatase from *C. stellatopilosus* can be classified in this group.

Based on mechanism, phosphatases can be divided into 2 groups:

1. Enzymes that utilize an active site nucleophile as the initial phosphoryl group acceptor. This group can be divided into two subgroups:

    - Phosphatases with active site nucleophile containing active amino acids such as histidine, cysteine, serine phosphatase.

    - Phosphatases with active site containing metal ion. Many phosphatases ultilize a two-metal ion dyad to bind phosphate ester and catalyze their subsequent hydrolysis. The metal ions such as Mg^{2+}, Mn^{2+}, Fe^{2+}, Fe^{3+}, Zn^{2+}, Co^{2+}.

2. Enzymes that transfer the phosphoryl group immediately to water.

9. Phosphatases involved in terpenoid pathways

According to Expasy database there are two groups of enzymes with phosphatase activities that use prenyldiphosphate as substrates:

1. Terpenyl diphosphatase (E.C. 3.1.7.3)

This group of phosphatases catalyze the dephosphorylation of terpenyl diphosphate compounds. Geranyl diphosphatase is an example of this type. It has been shown that partial purification of supernatant 105,000 x g from leaves of sage

(*Salvia officinalis*) from Sephadex G-150 and hydroxyapatite column can hydrolyze bornyl pyrophosphate into borneol (Croteau and Karp, 1979a; 1979b; Stafford, 1990) (Fig. 6). The phosphatase activities have shown to have two types of enzyme with difference Mr, pH optimum, and substrate specificity. The first enzyme was specific to pyrophosphate ester of borneol and pyrophosphate ester of cyclic monoterprenols: mentyl, thujyl, and terpinyl. The other type was specific to prenol pyrophosphate ester (Croteau and Karp, 1979b) (Table 1).
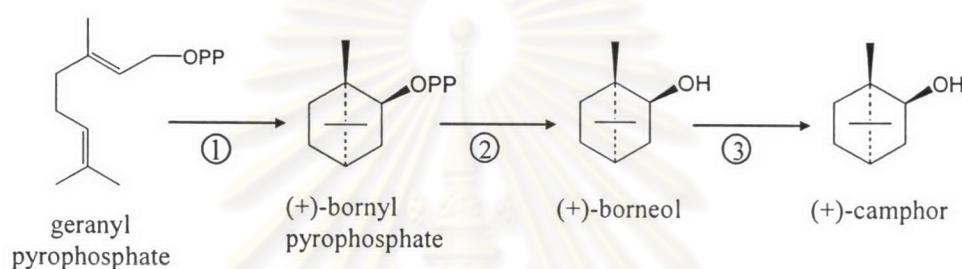

Figure 6 Pathway of (+)-camphor biosynthesis from geranyl diphosphate in *Salvia officinalis* (Stafford, 1990). Key to enzymes: (1) = (+)-bornyl pyrophosphate synthase (a cyclase); (2) = (+)-bornyl pyrophosphate phosphohydrolase; (3) = (+)-borneol dehydrogenase.

Table 1 Some properties of Bornyl pyrophosphate phosphohydrolases.

BPP phosphohydrolase	Example of substrates	pH optimum	Mr
Type 1	menthyl, thujyl, terpinyl	7.0-7.6	20 kDa
Type 2	neryl, geranyl	3.5-4.0	37, 65 kDa (overlap at 50 kDa)

2. Prenyl pyrophosphatase (EC 3.1.7.1)

(prenyl-diphosphatase, prenyl-diphosphate diphosphohydrolase, prenol pyrophosphatase, prenylphosphatase)

These enzymes catalyzed phosphoric ester hydrolysis to be prenol. They have been found in plants and animals, such as, *Citrus sinensis* (Perez et al., 1980, *Oryza sativa* (Nah et al., 2001) *Croton stellatopilosus* (Tansakul and De-

Eknamkul, 1998), rat liver (Tsai and Gaylor, 1966; Bansal and Vaidya, 1994). Most of them have been reported to be membrane-bound enzymes in microsomal fractions.

2.1 Phosphatase activities from the preparation of orange flavedo (*Citrus sinensis*)

A partial purified enzyme preparation from acetone extraction has been shown to contain a number of phosphohydrolytic activities which hydrolyzed several phosphomonoesters as well as pyrophosphate bonds as shown in Table 2 (Perez et al., 1980). The activities did not require divalent metals. It has been concluded the reaction sequence catalyzed be the enzyme is Prenyl-PP → Prenyl-P → Prenol.

Table 2 Phosphatase activities in an enzyme preparation from orange flavedo (Perez et al., 1980).

(a) Substrate 0.12 mM	Product measured	Specific activity (nkat/mg)		Purification fold
		Extract	P-cellulose	
IDP-[^{14}C]	Isopentenol-[^{14}C]	0.02	-	-
GDP-[^{3}H]	Geraniol-[^{3}H]	0.23	6.05	25
NDP-[^{3}H]	Nerol-[^{3}H]	0.27	7.26	27
FDP-[^{3}H]	Farnesol-[^{3}H]	1.12	1.75	1.6
GP-[^{3}H]	Geraniol-[^{3}H]	1.42	11.25	7.9
NP-[^{3}H]	Nerol-[^{3}H]	1.51	11.43	7.6
CiDP-[^{3}H]	Citronellol-[^{3}H]	-	2.42	-
GGDP-[^{3}H]	Geranylgeraniol-[^{3}H]	-	1.15	-
CoDP-[^{3}H]	Copalol-[^{3}H]	-	2.23	-

(b) Substrate 5 mM	Product measured	Specific activity (nkat/mg)		Purification fold
		Extract	P-cellulose	
ATP	Pi	11.22	79.9	7.1
PPi	Pi	7.14	122.4	17.0
P-nitrophenyl-phosphate	Pi	10.80	163.2	15.0
AMPi	Pi	6.80	120.7	18.0

Abbrevation: IDP = isopentenyl diphosphate; GDP = geranyl diphosphate; NDP = neryl diphosphate; FDP = farnesyl diphosphate; GP = geranyl monophosphate; NP = neryl monophosphate; CiDP = citronellol diphosphate; GGDP = geranylgeranyl diphosphate; CoDP = copalyl diphosphate; ATP = adenosine triphosphate; PPi = pyrophosphate; AMPi = adenosine monophosphate; P-cellulose = P- cellulose column chromatography.

2.2 Farnesyl diphosphatase (FDPase)

FDPase is the enzyme that hydrolyzes diphosphate group from FDP to form farnesol (FOH). The FDPase from rice (*Oryza sativa*) has shown to have a substrate specificity to FDP. It is a membrane-bound enzyme detected in the microsomal fraction. It does not require Mg^{2+} for its activity, the optimum pH is 5.5-6.3. The enzyme can be inhibited by Zn^{2+} and Mn^{2+} (Nah et al., 2001; Bansal and Vaidya, 1994). Its reaction sequence has not yet been determined whether FOH came from one step dephosphorylation or two steps as the report in orange flavedo.

It has been suggested that the physiological role of hydrolyzation of FDP into FOH by FDPase in rice seedling is to reduce the toxicity of FPP which is accumulated within the cells under stress condition (Nah et al., 2001).

2.3 Geranylgeranyl diphosphatase (GGDPase)

GGDPase is the enzyme that hydrolyzes GGDP to form GGOH. It has also been shown to be a membrane-bound enzyme of microsomal fractions *Oryza sativa*, and rat liver microsome. Its optimum pH is 7.0-7.9, and need no Mg^{2+} for the reaction (Croteau and Karp, 1979b; Nah et al., 2001; Tansakul and De-Eknamkul, 1998). Until now, there is no report on amino acid and nucleotide sequences of any GGDPase enzymes from plants.

10. Amino acid sequences of phosphatase enzymes utilizing GGDP

So far, there has been no amino acid sequences of any GGDPase in database. Therefore, conserved regions of other related enzymes that use substrates with pyrophosphate group, either in the terpenoid biosynthetic pathway or others has been of interest. These closely related enzymes can be classified as follow:

I Terpenoid biosynthetic pathway

1.1 Isoprenyl diphosphate synthases (Prenyltransferase)

1.1.1 *E*-Isoprenyl diphosphate synthase

1.1.2 *Z*-Isoprenyl diphosphate synthase

1.2 Terpenoid cyclase (Terpene synthase)

1.3 Prenyl diphosphate phosphatase

II Other phosphatases

2.1 Vacuolar H^+-pyrophosphatase (V-PPase)

2.2 Enzymes with phosphatase motif

I Terpenoid biosynthesis pathway

1.1 Isoprenyl diphosphate synthases (Prenyltransferase)

1.1.1 *E*-Isoprenyl diphosphate synthase (IPPS) is the enzyme that catalyzes the reaction of adding IDP (C5) into various prenyl diphosphate substrates to form a diverse group of terpenoids with *trans*- configuration (Fig. 7). For example, farnesyl synthase is the enzyme that add IDP into GDP to form FDP and geranylgeranyl synthase catalyzes the addition of IDP into FDP to form GGDP. Amino acid sequence of its conserved regions of *E*-isoprenyl diphosphate synthase of geranyl pyrophosphate synthase (GPPS), farnesyl pyrophosphate synthase (FPPS), and geranylgeranyl pyrophosphate synthase (GGPPS) have been shown in Fig. 8.
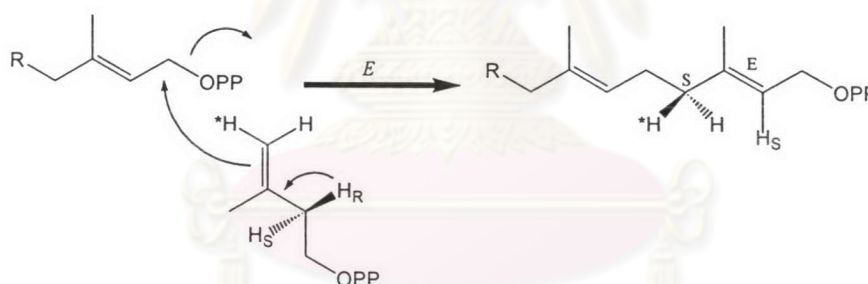

Figure 7 Prenylation by *E*-isoprenyl diphosphate synthase (Wang and Ohnuma, 2000)
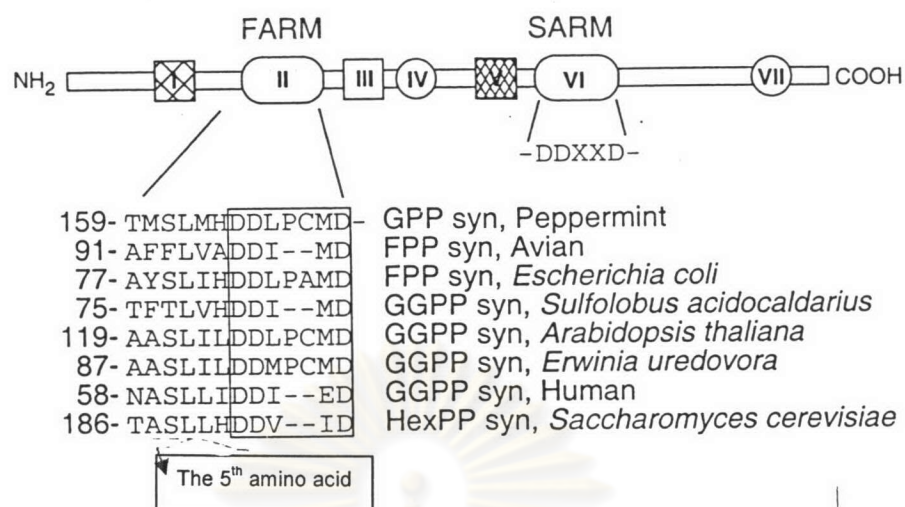
Figure 8 Structures of *E*- Isoprenyl diphosphate synthases (Wang and Ohnuma, 2000)

There are two areas of aspartate rich motif DDXXD, which is formed into substrate-binding pocket. These areas play roles in substrate binding and catalysis via chelating with Mg^{2+}, a cofactor required for enzyme activity. The first DDXXD motif (FARM) combined to prenylpyrophosphate product. The 4th and 5th amino acid before the first DDXXD motif controlled for the product elongation (Sittithaworn et al., 2001). The second DDXXD (SARM) binds to IPP (Fig. 9). The enzyme geranylgeranyl diphosphate synthase in *C. stellatopilosus* mutated at the 4th and 5th amino acids have been shown to cause the change of the length of enzymatic product from GGDP (C20) into FDP (C15) (Sitthithaworn et al., 2001).
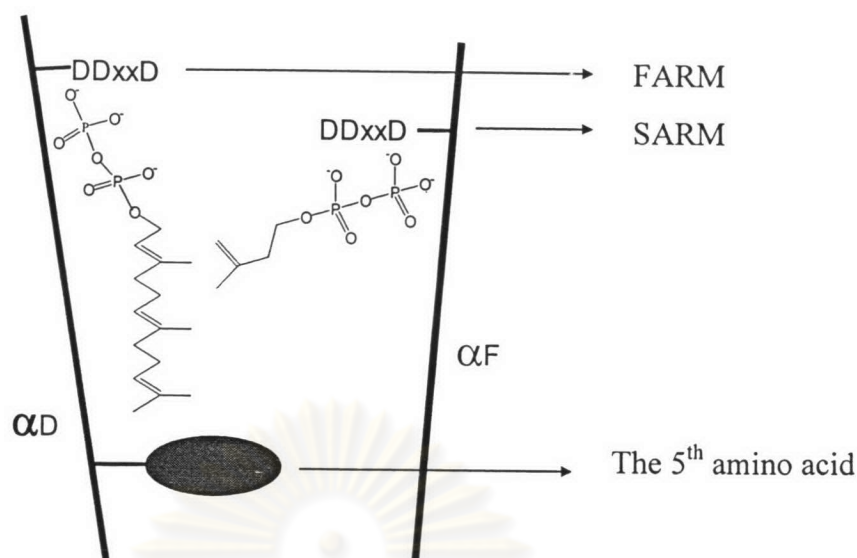
Figure 9 Schematic representation of the active site of FPPS with FDP (product) and IPD bound (Liang et al., 2002). The 5th amino acid blocks further chain elongation of the product and represents a mechanism to control the product chain length.

In addition, N-terminal transit peptide domain present in the enzymes is involved in the biosynthesis of monoterpene, diterpene, and tetraterpene which are synthesized in plastid. This area shows some identical characters such as; 1) there are 10 uncharged amino acids at N-terminal 2) central region (11-46) composed of hydroxy amino acid (serine and threonine) and positive charged amino acid (lysine, arginine and histidine).

1.1.2 *Cis*- Isoprenyl diphosphate synthases (*Z*-IPPS)

Z-IPPS is the enzyme that catalyzes the reaction of adding IDP (C5) into prenyl diphosphate substrate to form the terpenoids with *cis*-configuration (Fig. 10). The active site of Z-IPPS differs from that of E-IPPS. Z-IPPS lack of DDXXD motifs, although they require Mg^{2+} for their activity. This group of enzymes has the p-loop, a conserved motif for the pyrophosphate binding site in many phosphate binding proteins such as nucleotide triphosphate hydrolase, phosphofructokinase, sugar phosphatase. The chain length has been determined by a large amino acid in the bottom of the tunnel of active site, for example L137 (Fig. 11) (Ko et al., 2001). The multisequences alignment of these enzymes appear to have several conserved sequences.
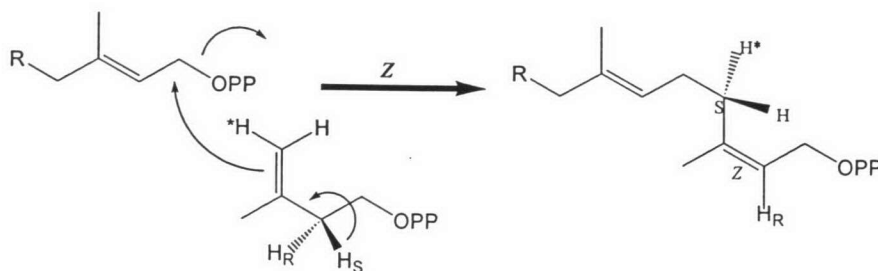
Figure 10 Prenylation by Z-isoprenyl diphosphate synthase (Wang and Ohnuma, 2000)
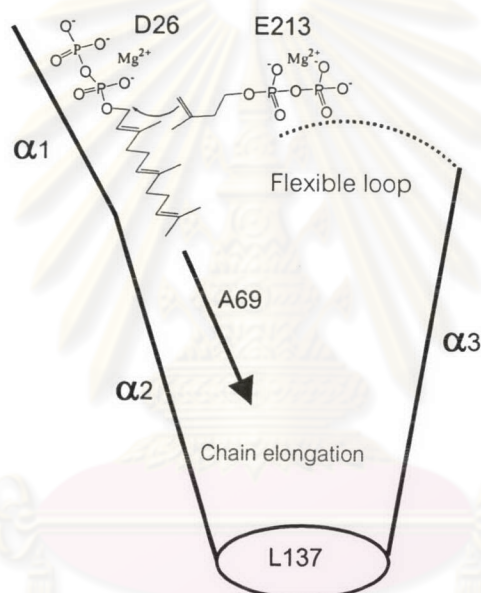
Figure 11 Schematic representation of the active site of undecaprenyl pyrophosphate synthase (UPPS) with FDP chain elongation and IDP bound (Liang et al., 2002)

1.2 Terpenoid cyclases (Terpene synthases)

Terpenoid cyclases catalyze the cyclization of substrates GDP, FDP, GGDP into various types of cyclic terpenoids. The conserve regions of terpenoid cyclases such as 1) aspartate-rich motif (I,L, or V) DDXXD motif, the binding site for metal ion that attached to diphosphate group of substrate (Mau and West, 1994; Huang et al, 1998; Bohlmann et al., 2000). 2) N-terminal transit peptide

for monoterpene synthase and diterpene synthase, which are biosynthesized in plastids. These transit peptides consisted of 50-60 amino acids with numerous of serine and threonine and few acidic amino acids (Wise et al., 1998). There is also a RRX_8W motif at the transit peptide sequence of diterpene (Bohlmann et al., 2000). 3) Serine 5 residues that are highly conserved may be involved in the binding and transferring of diphosphate group during reaction (Wise et al., 1998) 4) Highly conserved histidine and cysteine residues (Rajaomarivony et al., 1992). In Fig.12, the amino acid sequence of limone synthase (monoterpene cyclase) and casbene synthase (diterpene cyclase) show longer N-terminal sequence because of the present of plastidial targeting sequence. Vetispiradiene synthase and epi-aristolochene synthase are sesquiterpene synthase which were synthesized in cytosol, therefore they need no plastidial targeting sequence at the N-terminal.

Figure 12 The amino acid sequences comparison between terpene synthase. There are conserved regions such as aspartate-rich motif, highly conserved histidine and cysteine residues, and N-terminal plastidial targeting sequences of monoterpene and diterpene synthase (Back and Chappell, 1995). (Number within the boxes indicate the number of amino acid, percentage refers to the identity scores between the indicated domains.)

Both enzyme groups of the prenyltransferase (terpene synthase) and the terpene cyclase have analogous functions for their substrate binding and catalysis. First, both use hydrophobic interaction to bind the hydrocarbon moiety of the allylic pyrophosphate substrate. Second, both catalyze the PPi leaving by co-ordination of similar active site amino acids such as Asp, Glu, and Lys, with Mg^{2+}(Liang et al., 2002).

1.3 Prenyl diphosphate phosphatase
(phosphatidic acid phosphatase, lipid phosphate phosphatase)

So far, there have been no reports on the enzyme geranylgeranyl diphosphate phosphatases. Base on databases, however, the prenyl diphosphate phosphatases have its synonyms to phosphatidic acid phosphatase (PAP) and lipid phosphate phosphatase (LPP). They are AtLPP1 (Q9ZU49), and AtLPP2 (Q9XI60) from *Arabidopsis thaliana* (Pierrugues et al., 2001). These isoforms of enzymes are transmembrane proteins which are believed to respond to stress. Their actual substrates are believed to be both diacylglycerol pyrophosphate (DGPP) phosphatase and phosphatidate (PA) phosphatase. The sites of gene expression have been shown to be in roots, stems, leaves, buds, flowers, and siliques (Pierrugues et al., 2001). The enzymes prenyl diphosphate phosphatases: BAC41334, and BAC41335 (=Q9XI60) from *Arabidopsis thaliana* (Tokuhiro et al., 2001, unpublished) were thought to have a role in prenyl alcohol production by overexpression of prenyl diphosphate phosphatase in yeast *Saccharomyces cerevisiae*.

II Other phosphatases
2.1 Vacuolar H^+-pyrophosphatase (V-PPase)

V-PPase is the enzyme that uses PPi as substrate. It is the enzyme that transfers proton from cytosol through plasma membrane in order to produce driving force for passing metabolites to storage in plant vacuoles. The PPi originated from various cell reactions is the substrate of this enzyme, and its enzymatic product is 2Pi (Maeshima, 2000). The conserved regions of V-PPase has been reported as follows: 1) DX7KXE, the substrate binding motif (PPi binding) such as DVGADLVGKVE; 2) acidic 1 DXXXDXXXD motif such as DNVGDNVGD;

3) acidic 2 DXXXDXXXD motif such as DTXGDPXKD (Fig. 13). The aspartate residues at the PPi binding site, and at both acidic motifs played a role in binding and hydrolysis of substrate (Nakanishi et al., 2001).
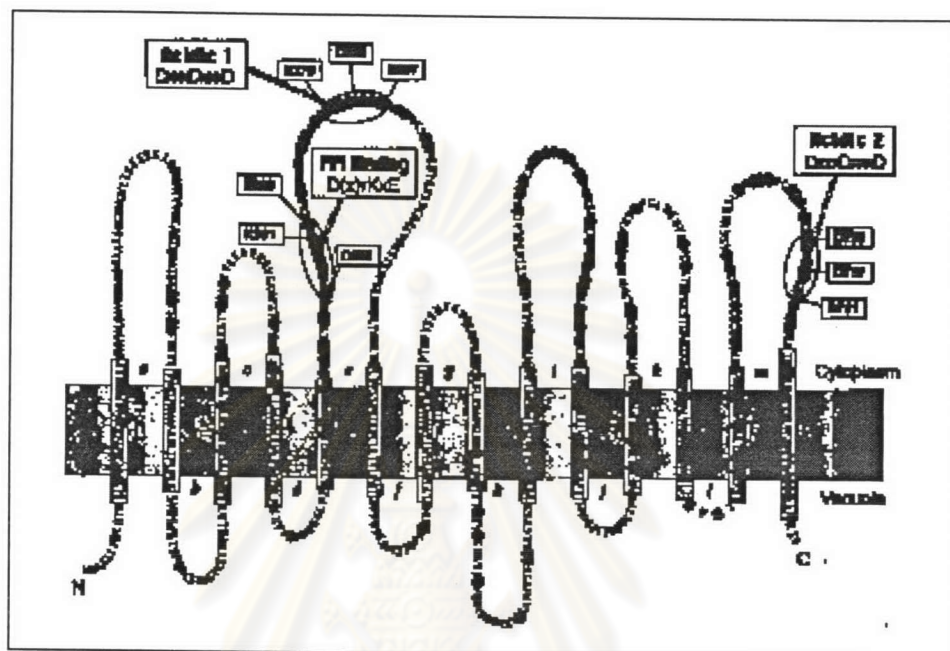

Figure 13 Transmembrane model of mung bean V-PPase (Nakanishi et al., 2001)

2.2 Enzymes with phosphatase sequence motif

Stukey and Carman, 1997 have identified the conserved phosphatase sequence motif that shares among several lipid phosphatases, the mammalian glucose-6-phosphatases, and a collection of bacterial nonspecific acid phosphatases. The conserved region is KXXXXXXRP-(X12-54)-PSGH-(X31-54)-SRXXXXXHXXXD, as shown in Fig. 14.

Protein (source, accession #)	Domain 1[bc]	Domain 2	Domain 3[d]
PGP[f] phosphatase (*E. coli*, P18201)	96-KDKVQEPRP-	54-PSGH-	36-SRLLLGMHWPRD-254
PGP phosphatase (*H. influenzae*, P44570)	93-KALFEEPRP-	54-PSGH-	41-SRVRLGMHYPID-241
PA[g] phosphatase (*M. musculus*, D84376)	119-KYTIGSLRP-	39-YSGH-	44-SRVSDYKHHWSD-283
Glucose-6-phosphatase (*H. sapiens*, P35575)	75-KWILFGQRP-	31-PSGH-	49-SRIYLAAHFPHQ-357
Glucose-6-phosphatase (*R. norvegicus*, L37333)	71-KWILFGQRP-	31-PSGH-	49-SRIYLAAHFPHQ-353
Glucose-6-phosphatase (*M. musculus*, P35576)	75-KWILFGQRP-	31-PSGH-	49-SRIYLAAHFPHQ-357
Phosphatase (*T. denticola*, L25421)	82-KRILKIPRP-	17-PSGH-	54-SRVYLGVHYPTD-341
Apyrase (*S. flexneri*, U04539)	123-KEYYKRVRP-	23-PSGH-	31-SRVICGAHWQSD-246
NS-phosphatase[h] (*S. typhimurium*, S14515)	122-KKYYMRTRP-	23-PSGH-	31-SRVICGAHWQSD-232
NS-phosphatase (*M. morganii*, P28581)	132-KEHYMRIRP-	23-PSGH-	31-SRVICGYHWQSD-249
NS-phosphatase (*P. stuartii*, P26975)	132-KEKYMRIRP-	23-PSGH-	31-SRVICGYHWQSD-248
Acid phosphatase (*Z. mobilis*, P14924)	131-KNNWNRKRP-	23-PSGH-	31-SRIVCGAHWFSD-264
Hypothetical (*S. cerevisiae*, U51031)[i]	117-KNWIGRLRP-	39-PSGH-	46-SRTQDYRHHFVD-289
Hypothetical (*S. cerevisiae*, U33057)	135-KLIIGNLRP-	41-PSGH-	38-SRVIDHRHHWYD-275
Hypothetical (*S. cerevisiae*, X87371)	127-KDYWCLPRP-	20-PSSH-	42-GRIYCGMHGILD-409
Hypothetical (*S. cerevisiae*, Z28278)	128-KDYWCLPRP-	20-PSSH-	42-GRVYCGMHGMLD-404
Hypothetical (*C. elegans*, Z68105)	160-KCYVGRLRP-	44-PSGH-	47-TRVTDNWHFPTD-318
Hypothetical (*C. elegans*, U28738)	143-KHVVGRLRP-	41-YSGH-	45-SRITDNKHHWSD-341
Hypothetical (*C. elegans*, U39648)	82-KPYFHRERP-	17-PSGH-	31-SRVALGRHYITD-345
Hypothetical (*Synechocystis sp.*, D64003)	133-KPFFNRTRP-	12-PSGH-	39-ASMYCRVHWATD-240
Hypothetical (*D. mobilis*, S01073)	77-KHLFNTPRP-	12-PSGH-	33-SRLYLRAHYPID-225
Hypothetical (*B. megaterium*, S32217)	106-KLFFQRARP-	13-PSGH-	41-SRIYLGVHYPSD-216
Chloroperoxidase (*C. inaequalis*, X85369)	352-KWEFEFWRP-	39-PSGH-	84-SRIFLGVHWRFD-609
CONSENSUS SEQUENCE	XXXXXXXRP	PSGH	SRXXXXXHXXXD

Figure 14 Consensus sequence of various phosphatases with phosphatase sequence motif (Stukey and Carman, 1997).

The features of this model include;

1. Nucleophilic attack of the substrate's phosphoryl group by the histidine of domain 3 and production of a phosphoenzyme catalytic intermediate

2. Conserved arginine residues of domains 1 and 3 hydrogen binding to the equatorial phosphoryl oxygens

3. The histidine of domain 2 participating in the protonation of the substrate leaving group

Phosphatases in which phosphatase sequence motif include prenyl diphosphate phosphatase from *Arabidopsis thaliana*, diacylglycerol pyrophosphate phosphatase (DGPP phosphatase), phosphatidic acid phosphatase (PA phosphatase), dolichyl pyrophosphate phosphatase (Dol-PP phosphatase). They are

transmembrane protein with 6 transmembrane regions. Their signal peptides have not been mentioned. The deduced amino acid sequences of lipid phosphate phosphatases from A. thaliana in Fig. 15 represent transmembrane regions and three domains of phosphatase motif.
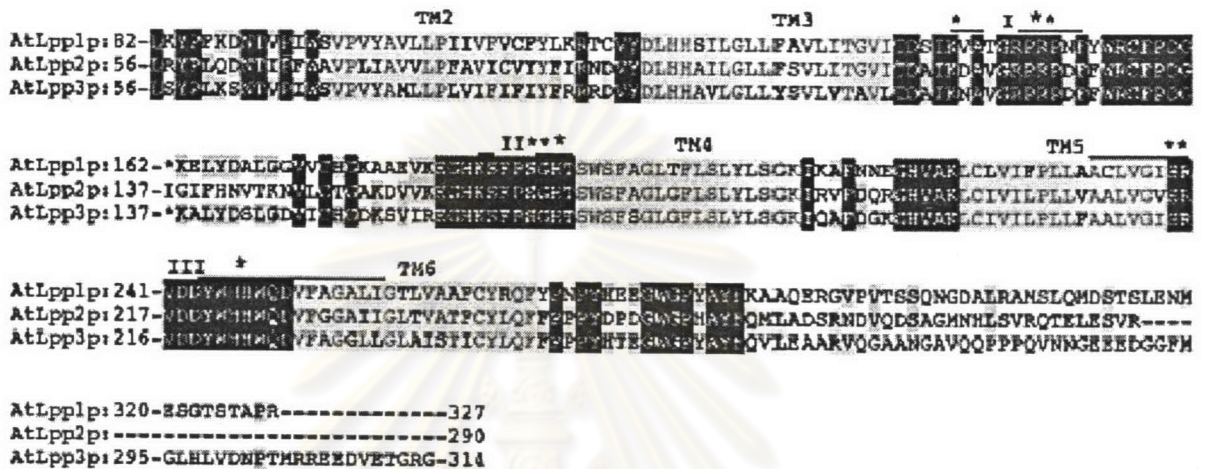

Figure 15 Comparison of amino acid sequence between the deduced AtLpp1p, AtLpp2p, and AtLpp3p proteins (Pierrugues et al., 2001). (TM = transmembrane region; I, II, III = phosphatase motif domain 1, 2, 3, respectively)

DGPP phosphatase is a vacuolar membrane protein of 34 kD. It catalyzed the removal of the β-phosphate from DGPP to form PA, and then removes the phosphate from PA to form diacylglycerol as shown in Fig. 16 (Han et al., 2004). PA phosphatase is an enzyme that catalyzes phosphatidic acid into diacylglycerol, and Dol-PP phosphatase is an enzyme that plays a role in glycosylation of lipid.
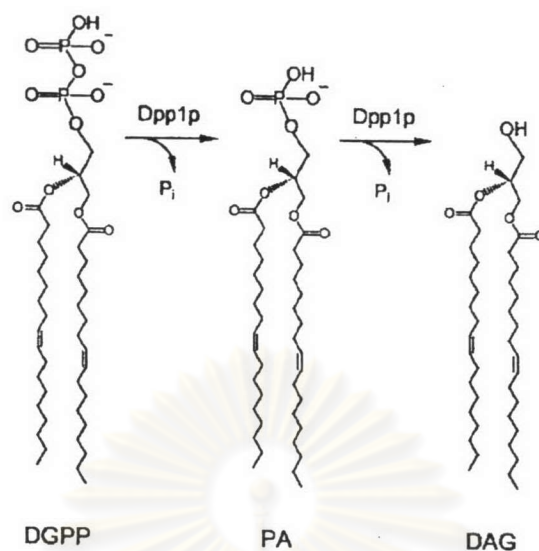
Figure 16 Dephosphorylation of DGPP from DGPP (Han et al., 2004).

11. N-terminal targeting sequences

Most proteins in a eukaryotic cell are encoded in the nuclear genome and synthesized in the cytosol. When the final destination is either mitochondrion, the chloroplast or the secretary pathway. N-terminal targeting sequence that is recognized by translocation machinery is the key part involving in the sorting (Buchanan et al., 2000; Emanuelsson et al., 2000).

Both mitochondrial and chloroplast import most of their proteins from the cytoplasm in the form of preprotein with transit peptides. The targeting sequencing will be proteolytically removed during or after the entering.

Targeting peptides are short peptide or amino acid motifs located in N-terminal end (or C-terminal in some proteins). Each compartment and membrane system requires a different targeting domain and sorting machinery as shown in Table 3.

Table 3 Peptide targeting domains for transport to different organelles.

Organelle	Targeting domain
ER	Signal peptide (SP)
Chloroplast	Transit peptide (cTSs)
Mitochondrion	Presequence or mitochondrial targeting peptides (mTPS)
Nucleus	Nuclear localization signal (NLS)
Peroxisome	Peroxisome targeting signal (PTS)
Vacuole	Vacuolar sorting signal (VSS)

Signal peptides (SPs) are responsible for targeting proteins to the ER for subsequence transport thorough the secretary pathway. Mitochondrial targeting peptides (mTPS), will be removed by proteolytic in mitochondrial matrix, rich of basic, hydroxylated, hydrophobic amino acid, whereas Chloroplast transit peptides (cTPS) consisted of numerous serine and threonine.

Prediction of targeting sequences

There are programs available for predicting the localization of protein which has targeting sequences, and also predict the cleavage site of these proteins. These programs such as, ChloroP which is the predictor of cTP (Emmanuelsson *et al.*, 2000), SignalP, the predictor to identify SPs (Nielsen et al., 1997), and TargetP, the programme to predict subcellular localization in chloroplast, mitochondria, ER/golgi/secreted, and others with a success rate of 85% in plant. TargetP also predicts potential cleavage site for presequence removal with correctly around 40% in chloroplast and 50% in mitochondria. It can distinguish between cTPs ad mTPs better than ChloroP (Emanuelsson et al., 2000). It is available as a webserver at http://www.cbc.dtu.dk/services/TargetP/.

Diterpene has the biosynthesis location at plastid, such as in chloroplasts. The amino acid sequences of diterpene synthases normally consisted of chloroplast transit peptides (Mau and West, 1994; Fischbach et al., 2001; Engprasert et al., 2004). However, the secondary structure of cTPs is not well characterized, and the sequence conservation around the stromal processing peptidase cleavage site is not strong (Gavel and Heijne, 1990).

Targeting proteins to plastids

Transport of proteins into chloroplasts involves a removable transit peptide. Chloroplast proteins encoded in nuclear DNA are synthesized in the cytosol. These proteins are translated as precursor with an N-teriminal transit peptide of 40-50 amino acid that targets the polypeptides to the chloroplast into the stroma. After that, a peptidase removes the transit peptides of stromal precursor proteins. Proteins that lack of transit peptide cannot be imported.

Pathways for transport across thylakoid membrane

There are 3 pathways involved in the transportation of protein across thylakoid membrane in chloroplast. The first two pathways will be mentioned here in detail.

1. SEC pathway (Secretory translocation system) required ATP, stimulated by pH gradient between chloroplast stroma and thylakoid lumen (Fig.17).

2. pH gradient pathway, required only the pH gradient, the twin-arginine motif is essential for protein transport via this pathway (Fig. 17).

3. SRP pathway (signal recognition particle), required GTP rather than ATP, stimulated by pH gradient, and required a stromal factor (SRP).
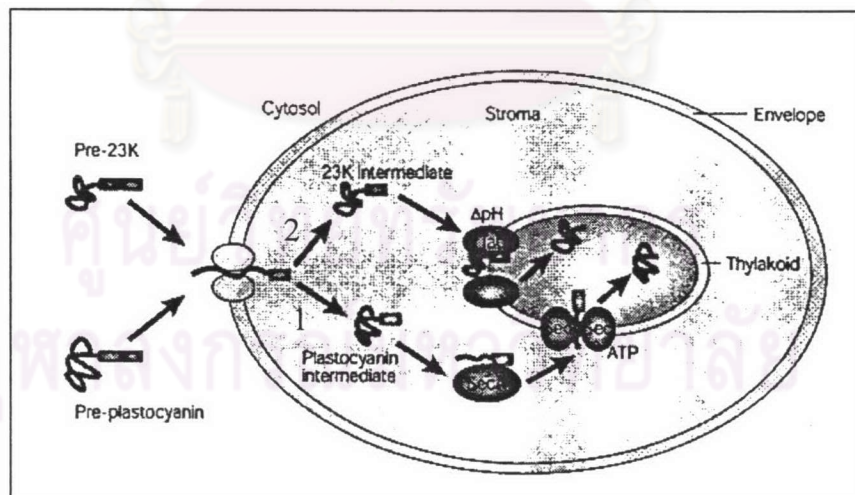

Figure 17 Pathways for protein targeting into thylakoids. (1=SEC pathway; 2=pH gradient pathway (Robinson and Bolhuis, 2001).

The twin-arginine translocation pathway encoded by tat genes, so called Tat system (Robinson and Bolhuis, 2001). Tat system operates in the thylakoid membrane of chloroplast, and in the plasma membrane of most free-living bacteria. It functions to transport **fully folded proteins** across the membrane. The transportation required no nucleoside triphosphate for their import into thylakoids, but it is totally dependent on pH gradient between the lumen of thylakoid and the stroma (Mould and Robinson, 1991)

Sec-type system (Sec pathway) found in chloroplast and bacteria. The system consists of SecA, SecY, and SecE homologues. Sec pathway transport thylakoid lumen proteins. The mechanism dependent on ATP.

Consensus sequences of chloroplast transit peptide

There are many attempts to characterize the chloroplast transit peptide. The consensus of RR motif for Tat system was identified in many proteins in *Arabidopsis thaliana*. The proteins from this plant also showed the Sec type system as shown in Fig. 18. The RRX_8W motif in diterpene synthases such as geranylgeranyl pyrophosphate synthase were suggested to be the cleavage site for chloroplast transit peptide (Engprasert et al., 2004). The RR motif involved in Tat system. So it is possible that diterpene synthases and monoterpene synthases were transported into chloroplast by Tat system.

TAT pathway

```
sp|O82660|H136_ARATH         SFSRRELLYQSAAVSLSLSSIVGPARA------
sp|Q42029|PSP1_ARATH         AVSRRLALTLLVGAAAVGSKVSPADA-------
sp|Q9XFT3|PSQ1_ARATH         ETSRRSVIGLVAAGLAGGSFVQAVLA------
sp|Q41932|PSQ2_ARATH         ESSRRSVIGLVAAGLAGGSFVKAVFA-------
sp|P82715|35.8_kDa_protein   GLSRRDLVLIGLSSPLSMFLPLSSPVTHA----
sp|O49292|TL30_ARATH         VLSRRSVMASGLVSSTTALAFPREGLA------
sp|P82538|TL26_ARATH         KCQRRLIVTFGVVAPWISLLSRAPLSFA-----
sp|O23403|T215_ARATH         AVGRRKSMMMGLLMSGLIVSQANLPTAFAP---
sp|P82281|TL29_ARATH         AFHRRDVLKLAGTAVGMELIGNGFINNVGDAKA
tr|Q9LU10|36_kDa_protein     STTRRILLTSLPMNLCFNPSRYLSALA------
tr|Q9LYR5|18_kDa_protein     EFDRRKLLVSSVGLLIGALSYDSKDGDFASA-
sp|O22870|17.5_kDa_protein   LSSRREAMLLVLGVSGGLSMSSLAAYA------
tr|Q9M222|16.9_kDa_protein   SLSRRSLVYVLVASPCLLLPALSSSA-------
tr|Q9SCY2|14.7_kDa_protein   SCGRREAIIGFGFSIGLLDNVSALA--------
tr|Q9S720|15.9_kDa_protein   GMKRRDVMLQIASSVFFLPLAISPAPA------
                                     **    :
```

```
tr|O22773|TL16_ARATH         LWKRRELSLGFMSSLVAIGLVSNDRRRHDANA
tr|Q9LM71|17.8_kDa_protein   PISRRDASIILLSSIPLTSFFVLTPSSSEA
tr|Q9SEL7|HhoA_precursor     DRGRRIMIFGSSLALTSSLLGSNQQRLPMZSAIA
tr|Q9LXX5|20_kDa_protein     QPRRRELLLKSAVAIPAILQLKEAPISAA
                                     **
```

Sec pathway

```
tr|Q39249|Violaxanthin_de    --LKELTAPLLLKLVGVLACAFLIVPSADA-----------
sp|P23321|PSO1_ARATH         ----GKCSDAVKIAGFALATSALVVSGASA-----------
sp|Q9S841|PSO2_ARATH         ----GKCSDAAKIAGFALATSALVVSGAGA-----------
sp|P11490|PLAS_ARATH         ---LKSSLKDFGVIAVATAASIVLAGNAMA-----------
sp|P42699|PLAT_ARATH         ---VKSSLKNFGVAAVAAAASIALAGNAMA-----------
sp|O22609|DEGP_ARATH         -PFSAVKPFFLLCTSVALSFSLFAASPAVESASA-------
tr|Q9ZP02|D1_processing      -MKSSVNFRQNLSVALVRIVSVLLVSSISVVTTDSPPSWG-
tr|P82869|38_kDa_protein     KNLEKLVATILIFVQVWSFLPLFGLDSAYISPAEA------
tr|Q9ASS6|16.5_kDa_protein   --TKSSFDSISFSSSTPFSASSSLLLHTSYTKRNHRCFSVQS
sp|P81760|TL17_ARATH         ---FPLKELGSIACAALCACTLTIASPVIA-----------
tr|O22160|11.6_kDa_protein   --------VSKRSLFALVSASLFFVDPALA-----------
tr|Q9ZVL6|18.3_kDa_protein   ---LIDAKQGLAALALSLTLTFSPVGTALA-----------
tr|Q9S33|17.9_kDa_protein    -------SLLPKLISFALAISLTSFSPALA-----------
tr|Q9LVV5|15_kDa_protein     -RFRSKSLSLVFSGALALGLSLSGVGFADA-----------
tr|Q9FL23|proteinase_D1      ---LKKSVIGTLTGALSLTVFSSPIS-SVA-----------
tr|Q9SSA5|40_kDa_protein     -LKECAISLALSVGLMVSVPSIALPPNAHAVANPVIPDVS-
```

Figure 18 Transit peptides of luminal proteins from *Arabidopsis thaliana* (Schubert et al., 2002).