CHAPTER II

HISTORICAL

2.1 The diversity of triterpenes

Terpenoids and steroids are natural product groups that can be found widely in living organism and have diversely chemical structures. Natural terpenoid derivatives can be found from a small molecule as in volatile oil, which has 10 carbon atoms in the structure, to a big molecule like natural rubber. The complexity of terpenoid derivatives can be also found from linear a polymer to a complicated chemical structure for example, ginkolide from *Ginkgo bilaba* leaves or azadirachtin from *Azadirachta indica* seeds.

Triterpenes are the largest group of natural terpenoids that originated from the composition of isoprene 6 units to construct the 30-carbon structure. The biosynthesis of triterpene derives from isoprene units, dimethylallyl pyrophosphate (DMAPP) and isopentenyl pyrophosphate (IPP), assembled as 'head to tail' to give geranyl pyrophosphate then geranyl pyrophosphate connect as 'head to tail' to give farnesyl pyrophosphate (FPP). The combination of FPP as 'tail to tail' gives squalene which is oxidized at the olefin group at C-2 to be 2,3-oxidosqualene. 2,3-Oxidosqualene is the substrate of triterpene synthase. During cyclisation reaction, the intermediates occur in two preconformations at the active site of triterpene synthase (Figure II-1) as follows.

1. Pre-chair-boat-chair conformation gives an intermediate which has the main structure as protosteryl cation. The cyclisation of protosteryl cation leads to triterpenoid products which are devided to 3 groups.
    1.1 Lanosterol which is the precursor of steroid synthesis in animals and fungi.
    1.2 Cycloartenol which is the precursor of steroid synthesis in plants.
    1.3 Cucurbitadienol which is a compound that is mostly found in family Cucurbitaceae.

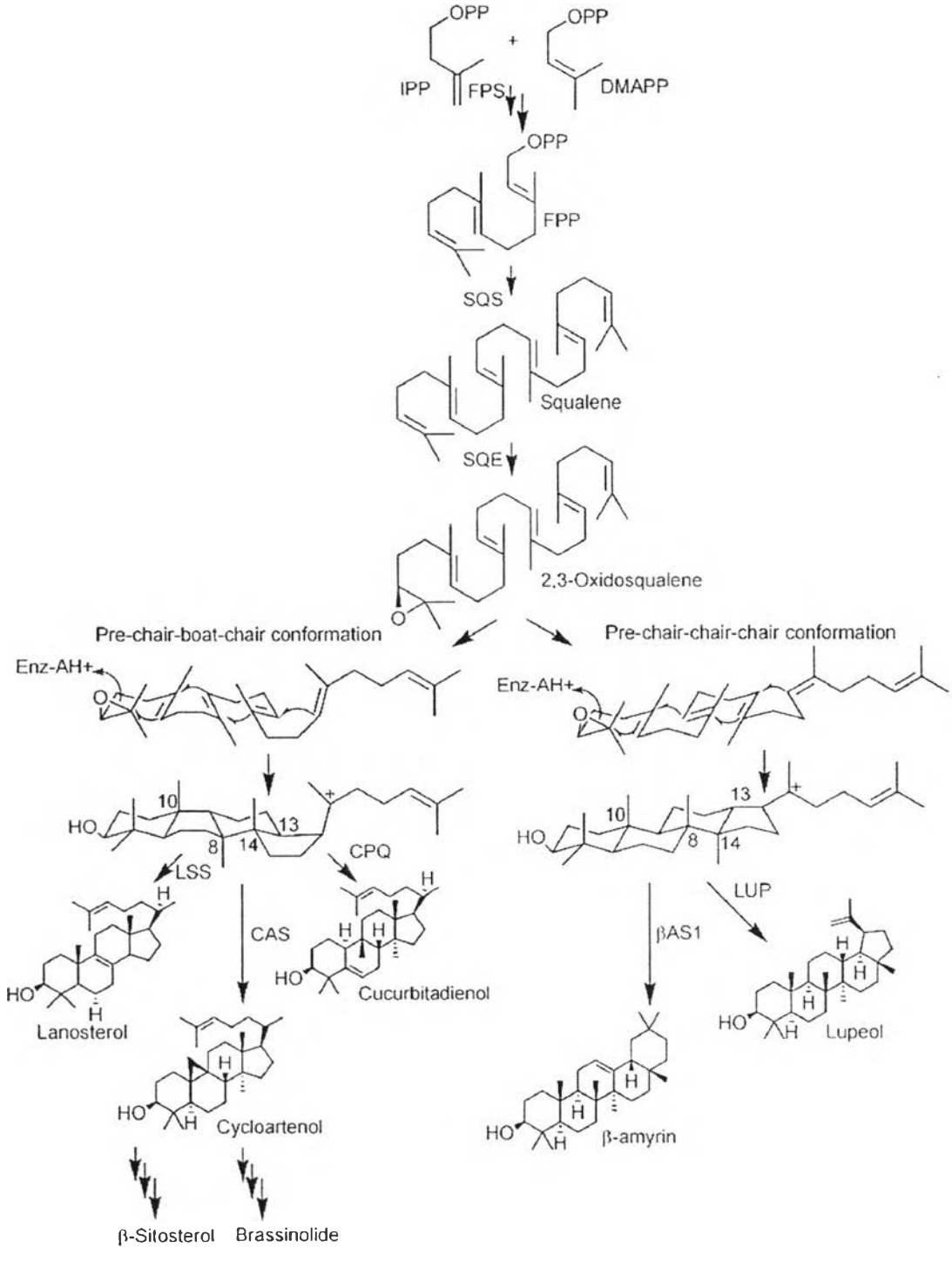

Figure II-1 Triterpene biosynthetic pathways modified from [4] and [6]. FPS, farnesyl pyrophosphate synthase; SQS, squalene synthase; SQE, squalene epoxidase; LSS, lanosterol synthase; CAS, cycloartenol synthase; CPQ, cucurbitadienol synthase, βAS, β-amyrin synthase; LUP, lupeol synthase.

2. Pre-chair-chair-chair conformation gives an intermediate which has the main structure as damaryenylcation. The cyclisation of damaryenylcation leads to triterpenoid products especially teteracyclic or pentacyclic triterpenes in plants, which are the triterpenes mostly found in nature and found to have biological activities.

Triterpenoid biosynthetic pathway in eukaryotes starts from the cyclisaton reaction of the precursor, 2,3-oxidosqualene, by triterpene synthase to give a product with molecular formula $C_{30}H_{50}O$ consisting of 1-5 rings on the structure [7]. Each triterpene synthase catalyses the precursor to a specific product. In a plant, there are plenty of triterpene synthases compete to change the precursor to steroids and pentacyclic triterpenoids [4]. In the cyclisation reaction of triterpenoids comprises of 4 steps:

1. Protonation
2. Cyclisation
3. Rearrangement
4. Deprotonation

(Figure II-2)

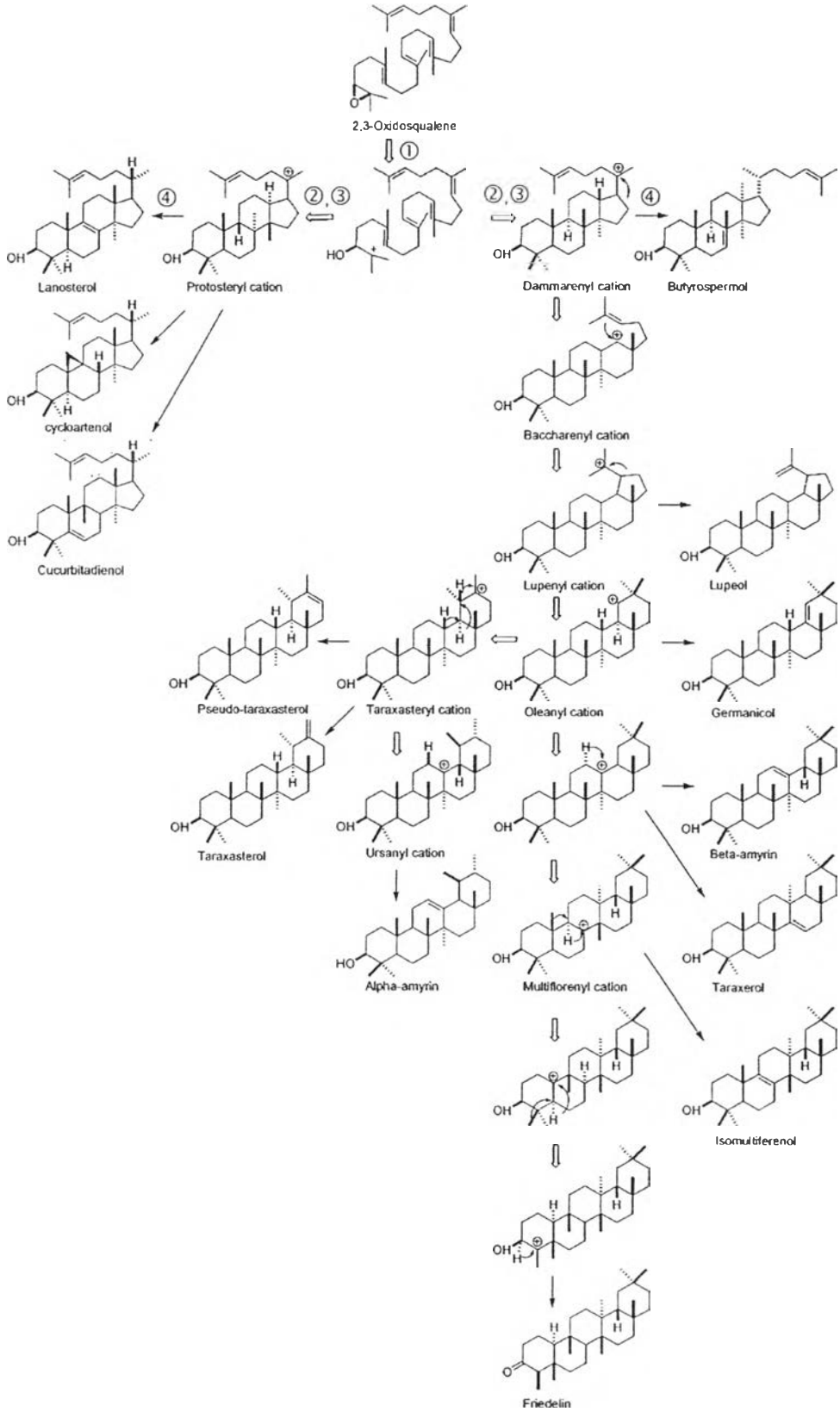

Figure II-2 The overall steps of triterpenoid cyclisation reaction modified from [8] and [9].

The diversity of triterpenoids happens from the rearrangement which depends on the type of the triterpene synthase in a plant. The diversity of triterpenoid structure in rearrangement step involves rearrangements of the tertiary carbocation, the methyl group and the hydride shift starting from E-ring, which is the last position of cyclisation, to A-ring, which is the last position of carbocation occurs. Finally, the dehydration gives a double bond replacing carbocation. The final product is up to the type of each enzyme, for example β-amyrin biosynthesis catalysed by β-amyrin synthase with 3 steps of rearrangement before deprotonated at C12 (Figure II-2) [9]. It can be concluded that the diversity of triterpenoid structures depend on the type of triterpene synthase, the steps of rearrangement and the last position of carbocation before it is deprotonated to be double bond on the structure.

The diversity of triterpenoid structures by triterpene synthases lightens an interesting point to study molecular biology of triterpene synthase in plants. Amino acid sequence on triterpene synthase gene would be the way to understand how triterpene synthases catalyse the common precursor to several different triterpenes.

Recently, there are two triterpene synthases that have been reported about the X-ray crystal structures, squalene-hopene cyclase from bacteria *Alicyclobacillus acidocaldarius* [10] and oxidosqualene cyclase from human [11]. Therefore, the mechanism of plant oxidosqualene cyclase has been studied and explained according to the enzyme models for searching the active site.

2.2 Squalene-hopenecyclase (SHC) from *Alicyclobacillus acidocaldarius*

Squalene-hopenecyclase (SHC) produces hopanoids which can be compared to sterols in eukaryotes. The biosynthetic pathway of hopanoids is similar to OSC that SHC catalyses squalene to hopene, while OSC catalyses 2,3-oxidosqualene, which is a squalene analog, to cyclic triterpenes [12] (Figure II-3).

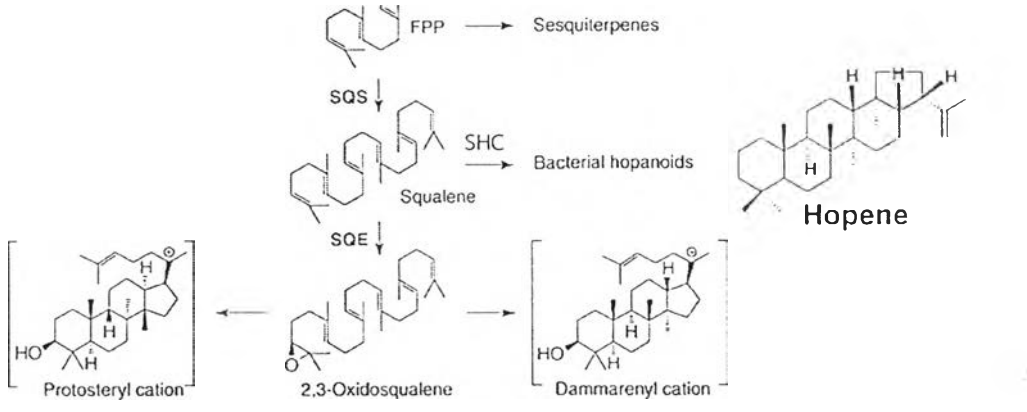

Figure II-3 The similarity of biosynthetic pathways of SHC and OSC [4]. SQS, squalene synthase; SQE, squalene epoxidase; SHC, squalene-hopene cyclase.

Although SHC has 19-25% amino acid sequence similar to oxidosqualene cyclase (OSC) and has shorter sequence, but the identity of both sequences would derive from a common ancestor due to the conserve region QW motif which would be unique to this enzyme family. The conserved region QW motif can be found spread on the amino acid sequence and reveal to stabilize and maintain integrity of the enzyme by forming H-bond to fix outer barrel helix together (arrows Figure II-4). The active site consists of the DXDD motif of which aspartic acid is the amino acid to start deprotonation reaction.

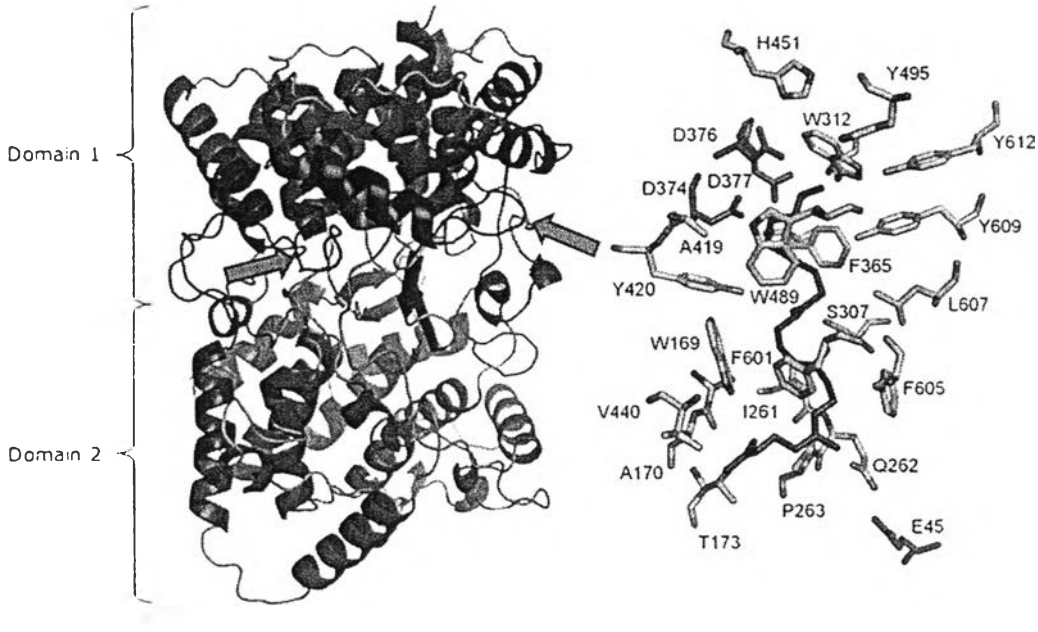

Figure II-4 The structure and active site of *alicyclobacillus acidocaldarius* SHC. The arrows point glutamine (Q) and tryptophan (W) which help to stabilize enzyme structure [12].

2.3 Human oxidosqualene cyclase

Human oxidosqualene cyclase is a lanosterol synthase. It is a monotopic memberane protein that has part of protein inserted into the membrane. The area of the inserted membrane surface forms of a plateau and a channel which consist of Tyr237, Cys233 and Ile524 that act like a close-open gate to lead the precursor to the active site (Figure II-5) [11].

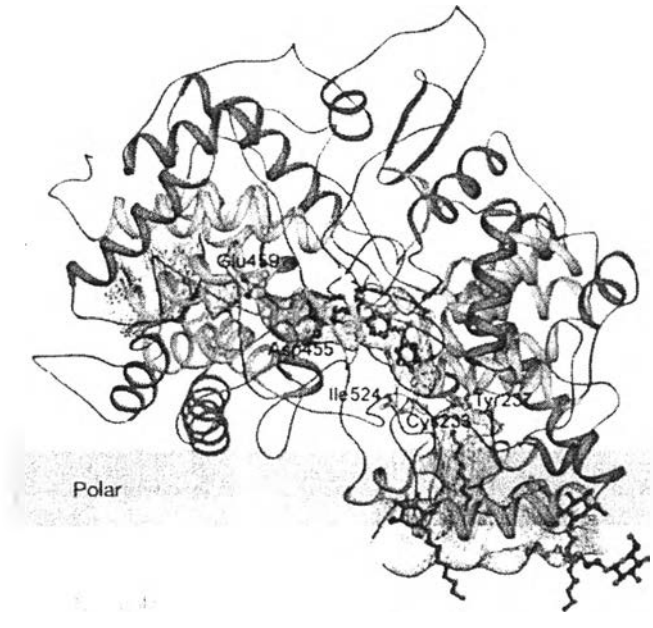

Figure II-5 The structure and active site of human OSC [11].

After the substrate occupies the active site, the substrate sets its structure in pre-conformation waiting for cyclisation. The substrate of OSC has a pre-conformation as pre-chair-boat-chair due to the stearic effect from Tyr98 that presses C10 be lower than molecular plane while the SHC substrate is in a prefold-chair-chair-chair conformation without stearic effect at the active site (Figure II-6) [11].

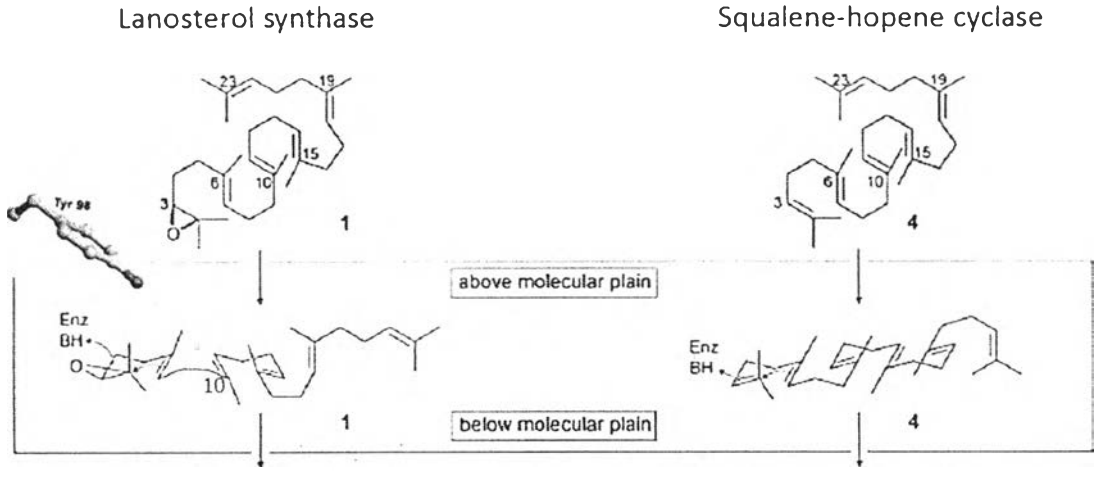

Figure II-6 Pre-conformation of the substrates at OSC and SHC at active sites modified from [13].

The active site of eukaryotic enzyme consists of amino acid residues XXDCX that aspartic acid is assumed to start cyclisaton reaction compare to the prokaryotic amino acid residue DXDD.

In the cyclisation reaction of the protosteryl cation by lanosterol synthase, Asp455 is activated by Cys456 and Cys533 causing Asp be more acidic followed transferring the proton to 2,3-oxidosqualene and lead to a cascade of cyclisation [11]. The cyclisation of A-ring causes carbocation at C-6 of oxidosqualene stabilized by Trp387 with cation-¶ interaction. This leads to the cyclisation of B-ring, and carbocation at C-10 is formed and stabilized with Phe444 and Trp581 to create the C-ring conformation. The C-ring conformation causes carbocation at C-14 which is stabilized by Phe363 and His232 with cation-¶ interaction to form the protosteryl cation. During lanosterol formation, carbocation at C-20 on the prostosteryl cation moves to C8/C9, which is the area surrounded by 7 aromatic amino acids, resulting from the rearrangements of two methyls and three hydride shifts. Finally, His232, which is a base amino acid and close to the carbocation at C8/C9, deprotonated from the intermediate to give lanosterol (Figure II-7).

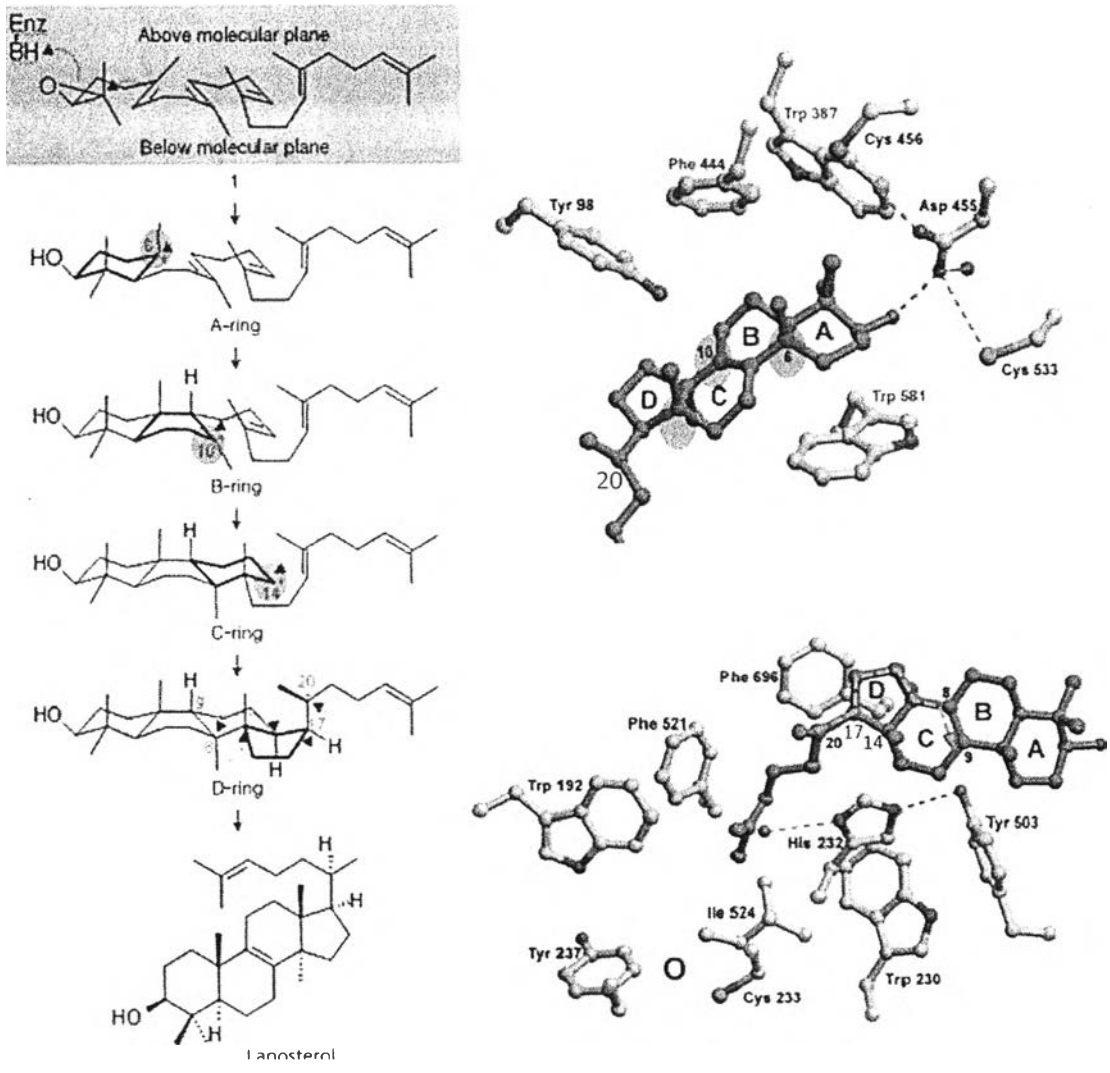

Figure II-7 A, cyclisation and rearrangement of lanosterol via the protosteryl cation; B and C, Protosteryl cation structure at the active site of human oxidosqualene cyclae [11]

2.4 Plant oxidosqualene cyclase

The molecular biology studies of plant OSCs or triterpene synthases comparing between human OSC and SHC has shown that triterpene synthases of the QW motif (Figure II-8), which help to stabilize the protein structure, and the DCTAE motif, which has aspartic acid playing an important role to protonate and start the cascade cyclisation [14]. Plant OSCs probably have the mechanism similar to the enzyme models that need aromatic amino acids at the active site to stabilize carbocation for the cascade cyclisation reaction.

Figure II-8 An example of QW motif from a part of amino acid sequence alignment of 5 oxidosqualene cyclases from *Kalanchoe daiagremontiana* [9].

ß-amyrin synthase (PNY) and lupeol synthase (OEW) have been studied for site-directed mutagenesis to identify for amino acid residues that have the role for their product specificity [15]. Trp259 has been found to help stabilizing the oleanyl cation and inducing the rearrangement to form ß-amyrin, in contrast without this effect of Leu replacement by Trp259, the reaction would be terminated to give lupeol (Figure II-9) [15].

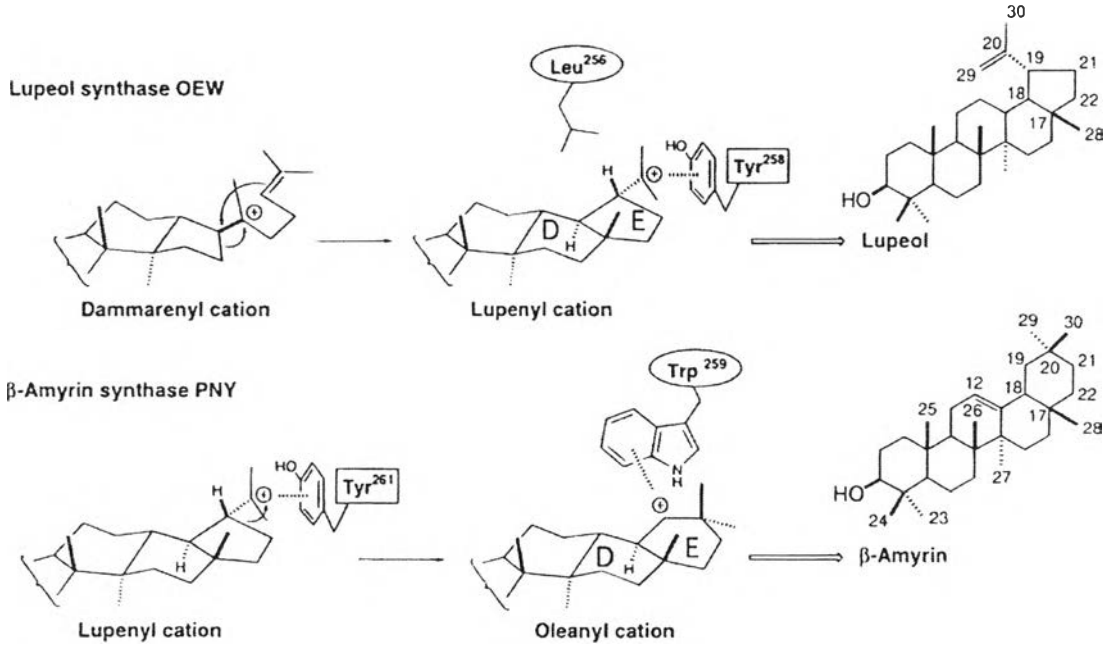

Figure II-9 Aromatic amino acid effect of the cyclisation of triterpenes at the active site of lupeol synthase and β-amyrin synthase [15].

Based on the current information, the cyclisation mechanism can be predicted via the intermediates on the pathway (Figure II-2).

2.5 Reported characterized plant OSC

In the literature, plant OSCs have been reported for molecular biology studies and genes in NCBI data base (Table II-1). Many characterized genes have been found to cyclise specific single product, but some have been shown to encode multifunctional enzymes. The characterized triterpene synthases are the data for studying amino acid involved the active site, protein structure and triterpene synthetic in plants.

Table II-1 Reported characterized plant

No.	Organism	Name	Enzyme	Accession no.	Reference
1	*Arabidopsis thaliana*	*At1g78500*	Pentacyclic triterpene synthase: β-seco-amyrin and α-seco-amyrin	AB274959	[16]
2	*Arabidopsis thaliana*	*LUP1*	multifunctional OSC: lupeol, β-amyrin	U49919	[17]
3	*Arabidopsis thaliana*	*YUP8H12R.43*	multifunctional OSC lupeol, taraxasterol, β-amyrin, ψ-taraxasterol, bauerenol, α-amyrin , multiflorenol	AC002986	[18]
4	*Arabidopsis thaliana*	*CAS1*	cycloartenol synthase	U02555	[19]
5	*Arabidopsis thaliana*	*LAS1*	lanosterol synthase	AB247155	[20]
6	*Arabidopsis thaliana*	*LSS*	lanosterol synthase	DQ508794	[21]
7	*Arabidopsis thaliana*	*At1g78950*	β-amyrin synthase	AB374428	[22]
8	*Solanum lycopersicum*	*ITTS1*	β-amyrin synthase	HQ266579	[8]
9	*Solanum lycopersicum*	*TTS2*	multifunctional OSC δ-amyrin, β-amyrin, α-amyrin, multiflorenol, ψ-taraxasterol, taraxasterol	HQ266580	[8]
10	*Taraxacum officinale*	*TRV*	putative OSC	AB025346	[23]
11	*Taraxacum officinale*	*TRW*	Lupeol synthase	AB025345	[23]

No.	Organism	Name	Enzyme	Accession no.	Reference
12	*Allium macrostemon*	*ALLOSC1*	putative OSC	AB025353	[24]
13	*Betula platyphylla*	*BPX1*	cycloartenol synthase	AB055509	[25]
14	*Betula platyphylla*	*BPX2*	cycloartenol synthase	AB055510	[25]
15	*Betula platyphylla*	*BPW*	lanosterol synthase	AB055511	[25]
16	*Betula platyphylla*	*BPY*	β-amyrin synthase	AB055512	[25]
17	*Medicago truncatula*	*βAS*	β-amyrin synthase	AJ430607	[26]
18	*Aster sedifolius*	*AsOXA1*	β-amyrin synthase	AY836006	[27]
19	*Kandelia candel*	*KcMS*	Multifunctional OSC lupeol, β-amyrin, α-amyrin	AB257507	[28]
20	*Pisum sativum*	*PSY*	β-amyrin synthase	AB034802	[29]
21	*Pisum sativum*	*PSM*	Multifunctional OSC α-amyrin, β-amyrin, δ-amyrin, ψ-taraxasterol, butyrospermol, lupeol, germanicol, taraxasterol	AB034803	[29]
22	*Pisum sativum*	*PSX*	Cycloartenol synthase	AB009029	[29]
23	*Luffacy lindrica*	*LcIMS1*	Isomultiflorenol synthase	AB058643	[30]
24	*Luffacy lindrica*	*LcCAS1*	Cycloartenol synthase	AB033334	[30]
25	*Glycyrrhiza glabra*	*GgLUS1*	Lupeol synthase	BAD08587	[31]
26	*Glycyrrhiza glabra*	*GgCAS1*	Cycloartenol synthase	AB025968	[31]
27	*Glycyrrhiza glabra*	*GgbAS1*	β-amyrin synthase	AB037203	[31]
28	*Panax ginseng*	*PNY*	β-amyrin synthase	AB009030	[32]

No.	Organism	Name	Enzyme	Accession no.	Reference
29	*Panax ginseng*	*PNX*	Cycloartenol synthase	AB009029	[32]
30	*Panax ginseng*	*PNA*	Dammarenediol-II synthase	AB265170 (in DDBJ)	[33]
30	*Kalanchoe daigremontiana*	*KdFRS*	Friedelin synthase	HM623870	[9]
31	*Kalanchoe daigremontiana*	*KdCAS*	Cycloartenol synthase	HM623872	[9]
32	*Kalanchoe daigremontiana*	*KdLUP*	Lupeol synthase	HM623871	[9]
33	*Kalanchoe daigremontiana*	*KdGLS*	Glutinol synthase	HM623869	[9]
34	*Kalanchoe daigremontiana*	*KdTAS*	Taraxerol synthase	HM623868	[9]
35	*Olea europaea*	*OEA*	Mixed amyrin synthase α-amyrin, β-amyrin, ψ-taraxasterol, butyrospermol	AB291240	[34]
36	*Olea europaea*	*OEX*	Cycloartenol synthase	AB025344	[34]
37	*Olea europaea*	*OEW*	Lupeol synthase	AB025343	[34]
38	*Lotus japonicus*	*LjAMY1*	Putative OSC	AF478454	[35]
39	*Lotus japonicus*	*LjAMY2*	β-amyrin synthase	AF478455	[35]
40	*Lotus japonicas*	*OSC6*	Putative OSC	AB244670	[36]
41	*Lotus japonicus*	*OSC7*	Lanosterol synthase	AB244671	[36]
45	*Catharanthus roseus*	*CrAS*	Amyrin synthase α-amyrin,β-amyrin	JN991165	[37]
42	*Lotus japonicas*	*OSC1* (the full –length version of *LjAMY1*)	β-amyrin synthase	AB181244	[38]
43	*Lotus japonicus*	*OSC3*	Lupeol synthase	AB181245	[38]
44	*Lotus japonicus*	*OSC5*	Cycloartenol synthase	AB181246	[38]

2.6 *Alangium lamarckii*

Alangium lamarckii Thw., which is called as "Pru (ปรู๋)" in Thai, is a member of family Alangiaceae. The synonym scientific name of *A. lamarckii* is *Alangium salviifolium* (Linn. F.) Wang. *A. lamarckii* also has other Thai synonym names such as "Pu (ปู๋)","Phlu (ผลู)", "Makluaka (มะเกลือกา)","Matapu (มะตาปู๋)". Plant description of *A. lamarckii* has been recorded [39] as a small to medium tree which grows up to a height of 5-15 m. Leaves are 2.5-7 cm. wide and 5-15 cm. long, elliptical and alternate. Flowers are 1-2.5 cm. long, white and fragrant, and grow in fascicles. Fruits are 1-1.5 cm. long, small, green, nearly globular and purplish-red when ripe (Figure II-10).

A. lamarckii can be found widely in India and countries in Indochinese peninsula. In Thailand, *A. lamarckii* is found widespread in mixed deciduous forests in every part of Thailand except the south. Various parts of *A. lamarckii* are used in traditional Thai medicine. The root bark is used as anti-diarrhea and emetic. The stem bark is used to cure diarrhea and asthma and used as expectorant. The wood is used to heal hemorrhoids. The fruit is used as anti-diarrhea and destroying parasitic worms [39, 40].

A. lamarckii is well known as an alkaloid source. Previous phytochemical research of *A. lamarckii* reported many classes of alkaloids for example 1',2'-dehydrotubulosine and alangine from dried fruits [41], alangiumkaloids A and B from stem [42] and alkorine from leaves [43]. However, previous study found 4 triterpenoid compounds in *A. lamarckii* leaves e.g. friedelin, stigmasterol, β-sitosterol and isoalangidiol [5] (Figure II-11).

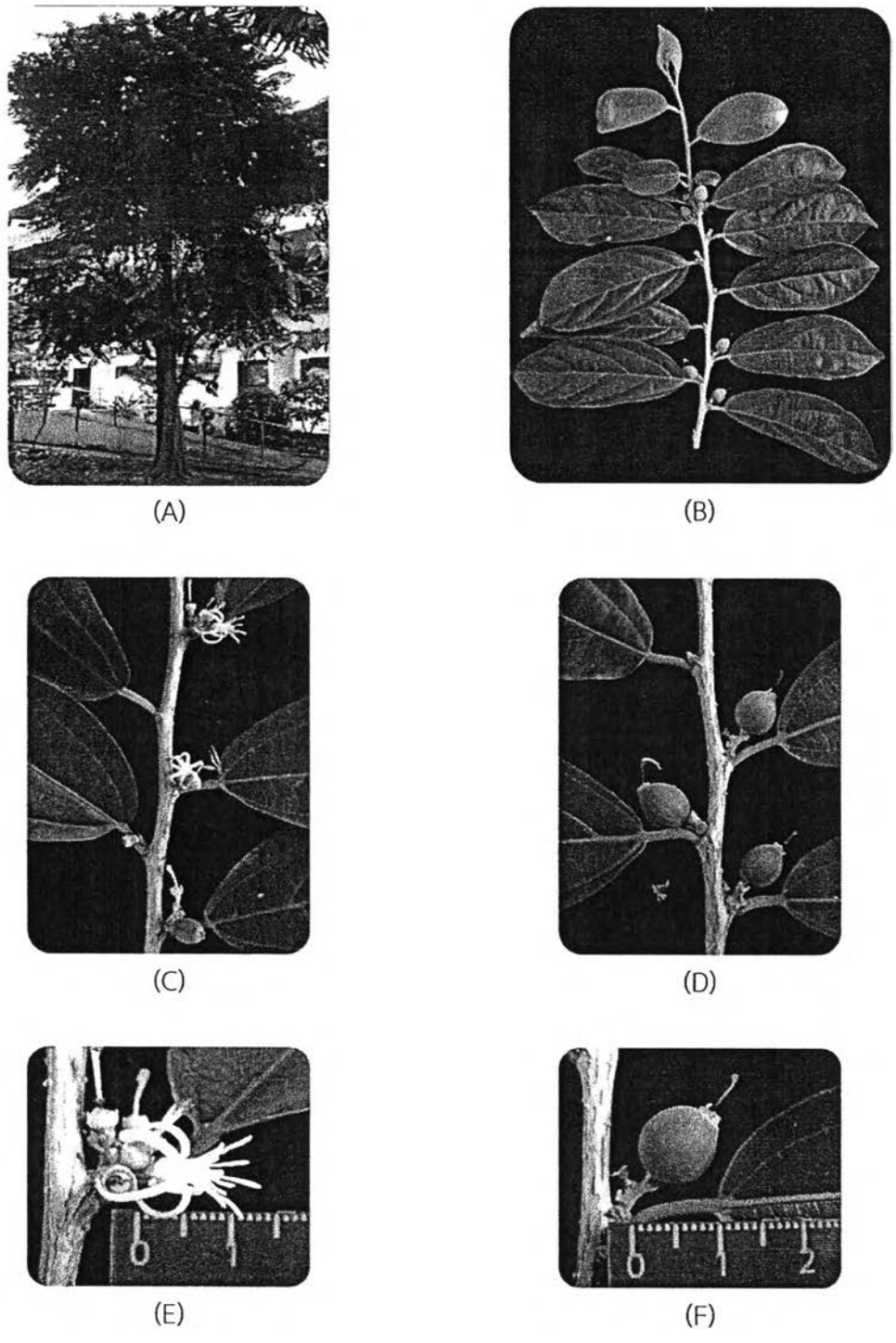

Figure II-10 *A. lamarckii* following (A) whole plant, (B) leaves, (C) and (E) flowers between leaves and branches, and (D) and (F) fruits between leaves and branches.

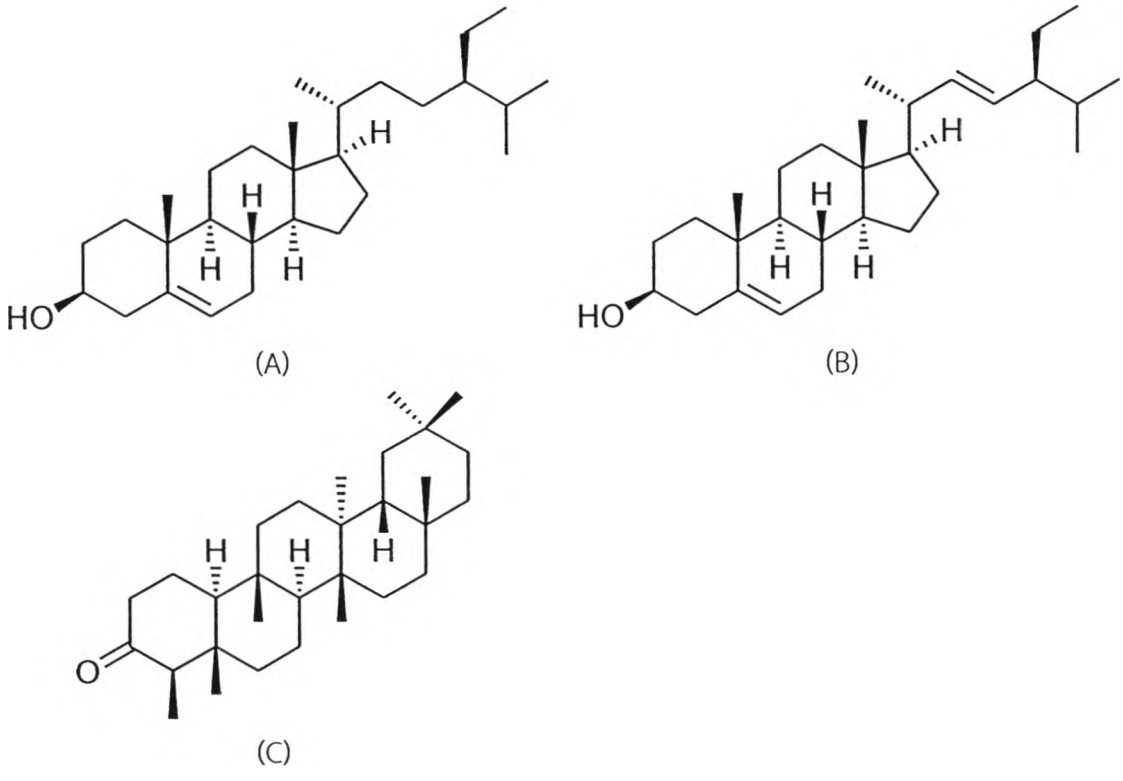

Figure II-11 The structures of triterpeoid compounds found in *A. lamarckii* leaves as following (A) β-sitosterol, (B) stigmasterol and (C) friedelin.