CLONING, EXPRESSION, AND CHARACTERIZATION OF INSECT CLASS I GLUTATHIONE S-TRANSFERASES FROM *ANOPHELES DIRUS* B

KANYA JIRAJAROENRAT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(MOLECULAR GENETICS AND GENETIC ENGINEERING)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2000

ISBN 974-664-344-4

Thesis
entitled

CLONING, EXPRESSION, AND CHARACTERIZATION OF INSECT CLASS I GLUTATHIONE S-TRANSFERASES FROM *ANOPHELES DIRUS* B

Ms. Kanya Jirajaroenrat
Candidate

Albert J. Ketterman, Ph.D.
Major Advisor

Asst. Prof. Chanan Angsuthanasombat, Ph.D.
Co-advisor

Chartchai Krittanai, Ph.D.
Co-advisor

Prof. Liangchai Limlomwongse
Ph.D.
Dean
Faculty of Graduate studies

Prof. Prapon Wilairat, Ph.D.
Chairman
Master of Science Programme in Molecular Genetics and Genetic Engineering
Institute of Molecular Biology and Genetics

Thesis
entitled

CLONING, EXPRESSION, AND CHARACTERIZATION OF INSECT CLASS I GLUTATHIONE S-TRANSFERASES FROM *ANOPHELES DIRUS* B

was submitted to Faculty of Graduate Studies, Mahidol University
for the degree of Master of Science
(Molecular Genetics and Genetic Engineering)
on
June 2, 2000

Ms. Kanya Jirajaroenrat
Candidate

Albert J. Ketterman, Ph.D.
Chairman

Asst. Prof. Chanan Angsuthanasombat, Ph.D.
Member

La-aied Prapanthadara, Ph.D.
Member

Chartchai Krittanai, Ph.D.
Member

Prof. Liangchai Limlomwongse
Ph.D.
Dean
Faculty of Graduate studies

Prof. Sakol Panyim, Ph.D.
Director
Institute of Molecular Biology and Genetics

ACKNOWLEDGEMENT

I would like to express my sincere gratitude and deep appreciation to Dr. Albert J. Ketterman, my Major Advisor for his warm guidance, invaluable advice, supervision and encouragement throughout this thesis project. He was never lacking in kindness and support. I am equally grateful to Professor Sakol Panyim, the director of the institute and Asst. Prof. Chanan Angsuthanasombat and Dr Chartchai Krittanai, my co-advisors for their valuable comments, advice, supervision and encouragement throughout this thesis project.

I would like to thank Dr. La-aied Prapanthadara for her helpfulness in giving all of valuable materials.

I wish to thank Ms. Peerada Prommeenate and Ms. Saengtong Pongjaroenkit, my seniors for their great assistance in solving technical problems and wonderful friendships. I also thank the members of BDG group and all my friends for their comments, advice, support, and friendships throughout this thesis project.

I am particularly indebted to the National Science and Technology Development Agency (NSTDA) for the scholarship, which enabled me to undertake this study.

Kanya Jirajaroenrat

4136308 MBMG/M : MAJOR: MOLECULAR GENETICS AND GENETIC ENGINEERING; M.Sc. (MOLECULAR GENETICS AND GENETIC ENGINEERING)
KEY WORDS : GLUTATHIONE S-TRANSFERASE/ ANOPHELES DIRUS/ MOSQUITO LARVAE

KANYA JIRAJAROENRAT: CLONING, EXPRESSION, AND CHARACTERIZATION OF INSECT CLASS I GLUTATHIONE S-TRANSFERASES FROM *ANOPHELES DIRUS* B. THESIS ADVISORS: ALBERT J. KETTERMAN, Ph.D., CHANAN ANGSUTHANASOMBAT, Ph.D., CHARTCHAI KRITTANAI, Ph.D. 131 p. ISBN 974-664-344-4.

Glutathione S-transferases (GSTs: E.C.2.5.1.18) are a multiple gene family of multifunctional dimeric enzymes which catalyze a board range of substrates and play an important role in detoxication of xenobiotic compounds. The GSTs in insects are of interest because they are involved in insecticide resistance. In this thesis study, three cDNA sequences of glutathione S-transferases: *adgst1-2*, *adgst1-3* and *adgst1-4*, the alternatively spliced products of the *adgst1AS1* gene, were obtained from the 4th instar larvae of *Anopheles dirus* B mosquito by RT-PCR reactions. The nucleotide sequences of these three cDNAs share >67% identity and the translated amino acid sequence share >61% identity. A comparison of the *An. dirus* to the *An. gambiae* enzymes shows the adGST1-2 versus agGST1-4, adGST1-3 versus agGST1-5 and adGST1-4 versus agGST1-3 have 85, 92 and 85% amino acid sequence identity respectively, which confirms that the orthologous isoenzymes occurred across the anopheline species. These three genes were expressed at high levels, approximately 15-20 mg from 200 ml of the *E. coli* culture. The recombinant enzymes were purified by affinity chromatography on an S-hexylglutathione agarose column. The subunit sizes of adGST1-2, adGST1-3 and adGST1-4 are 24.3, 23.9 and 25.1 kDa. The recombinant enzymes have high activities with CDNB, detectable activity with DCNB but markedly low activity with ethacrynic acid and *p*-nitrophenethyl bromide. The adGST1-3 was shown to be the most reactive enzyme from the kinetic studies. Permethrin inhibition of the three enzymes was different being uncompetitive for adGST1-2, noncompetitive for adGST1-3 and competitive for adGST1-4. Despite the enzymes being splicing products of the same gene and sharing identical sequence at the N-terminal domain, these GSTs show distinct substrate specificities, kinetic properties and inhibition properties.

4136308 MBMG/M : สาขาวิชา: อณูพันธุศาสตร์และพันธุวิศวกรรมศาสตร์; วท.ม.
(อณูพันธุศาสตร์และพันธุวิศวกรรมศาสตร์)

กัญญา จิระเจริญรัตน์: การโคลนนิ่ง การแสดงออก และ การศึกษาคุณสมบัติของโปรตีนไอโซเอนไซม์กลูตาไธโอน เอสทรานสเฟอเรส กลุ่มที่หนึ่งในแมลง จากยุงสายพันธุ์ *Anopheles dirus* กลุ่ม บี (CLONING, EXPRESSION, AND CHARACTERIZATION OF INSECT CLASS I GLUTATHIONE S-TRANSFERASES FROM *ANOPHELES DIRUS* B.). คณะกรรมการควบคุมวิทยานิพนธ์: Albert J. Ketterman, Ph.D., ชนันท์ อังศุธนสมบัติ, Ph.D., ชาติชาย กฤตนัย, Ph.D. 131 หน้า. ISBN 974-664-344-4.

กลูตาไธโอน เอสทรานสเฟอเรส (GST) เป็นเอนไซม์ที่ประกอบด้วยสองหน่วยย่อยมีการแสดงออกในหลายรูปแบบซึ่งสามารถเร่งปฏิกิริยาโดยใช้สารตั้งต้นได้หลายชนิดและมีบทบาทสำคัญในการทำลายสารพิษต่างๆ เอนไซม์นี้ที่พบในแมลงเป็นที่สนใจเนื่องจากมีส่วนเกี่ยวข้องในการดื้อต่อยาฆ่าแมลง ในงานวิจัยนี้สามารถแยกเอนไซม์ GST กลุ่มที่หนึ่ง เส้นที่สอง เส้นที่สาม และ เส้นที่สี่ (*adgst1-2*, *adgst1-3* and *adgst1-4*) ได้จากลูกน้ำยุงระยะที่สี่สายพันธุ์ *Anopheles dirus* กลุ่ม B ซึ่งเป็นผลิตผลของการตัดและต่ออย่างสลับ (alternatively spliced) ของยีน *adgst1AS1* โดยอาศัยวิธีการเพิ่มขยายชิ้น DNA โดยการถอดรหัสย้อนกลับจากเส้น mRNA (RT-PCR) เอมไซม์ที่ได้มีความเหมือนกันในระดับยีน 67% และในระดับโปรตีน 61% นอกจากนี้เอนไซม์เส้นที่สอง สาม และสี่ยังมีความเหมือนกับเอนไซม์เส้นสี่ ห้า และสาม จากยุงสายพันธุ์ *Anopheles gambiae* ในระดับโปรตีน 85%, 92% และ 85% ตามลำดับ เอนไซม์นี้สามารถแสดงออกได้ประมาณ 15-20 มิลลิกรัมจากการเลี้ยงเชื้อ*E.coli* ปริมาณ200 มิลลิลิตรและสามารถทำให้บริสุทธิ์โดยการจับกับ S-hexylglutathione agarose gel แต่ละหน่วยย่อยของเอนไซม์เส้นที่สอง สาม และสี่มีขนาด 24.3, 23.9 และ 25.1 กิโลดาลตันตามลำดับ เมื่อศึกษาคุณลักษณะของเอนไซม์ทั้งสามรูปแบบพบว่า สามารถเร่งปฏิกิริยาของสาร CDNB ได้ดีเมื่อเทียบกับสารตั้งต้นชนิดอื่นๆ โดยที่เอนไซม์เส้นที่สามสามารถเร่งปฏิกิริยาได้ดีที่สุด เมื่อใช้ยาฆ่าแมลงชนิด Permethrin เป็นสารยับยั้งพบว่า สามารถยับยั้งการเร่งปฏิกิริยาของเอนไซม์ทั้งสามในลักษณะที่แตกต่างกัน แม้ว่าเอนไซม์ทั้งสามรูปจะเป็นผลิตผลของการตัดและต่ออย่างสลับจากยีนเส้นเดียวกันและใช้ส่วนปลายด้านอะมิโนร่วมกันแต่คุณลักษณะของเอนไซม์มีความแตกต่างกัน

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BHC	Benzene hexachloride
Bis-acrylamide	*N,N'*-methylenebisacrylamide
BSA	Bovine serum albumin
BSP	Bromosulfophalein
CDNB	1-Chloro-2,4-dinitrobenzene
CTAB	Cetyltrimethylammonium bromide
DCNB	1,2-Dichloro-4-nitrobenzene
DDT	Dichlorodiphenyltrichlorethane
DEPC	Diethyl pyrocarbonate
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
GSH	Glutathione
GST	Glutathione S-transferase
IC_{50}	50 percent of inhibition
IPTG	Isopropyl-B-D thiogalactopyranoside
LB	Luria-bertani media
OP	Organophosphorus insecticide
PIPES	Piperazine-*N,N'*-bis(2-ethanesulfonic acid)
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEM	Simple and efficient method for Transformation
TB	Transformation buffer
TEMED	*N,N,N',N'*-tetramethylenediamine
Tris	2-hydroxymethyl-2-methyl-1,3-propanediol
UV	Ultraviolet
X-gal	5-bromo 4-chloro 3-indolyl β-D-galacto-side

CHAPTER I

INTRODUCTION

1.1 Mosquitoes

Mosquitoes are insects belonging to the order *Diptera*, the True Flies (1). Like all True Flies, they have two wings, but unlike other flies, their wings have scales and their mouthparts (in female mosquitoes) form a long piercing-sucking proboscis. Males differ from females by having feathery antennae and mouthparts not suitable for piercing skin. Nectar is their principal food source.

The 3000 mosquito species found in the world are divided among 28 different genera. They are classified as follow (2):

Order *Diptera*

Family *Culicidae*

Genus *Aedes, Anopheles, Culex, Culiseta, Mansonia Deinocerites, Haemagogus, Orthopodomyia, Psorophora, Toxorhynchites, Uranotaenia, Wyeomyia* etc.

The mosquito goes through four separate and distinct stages of its life cycle: Egg, Larva, Pupa, and Adult. Only female mosquitoes bite animals and drink blood. Male mosquitoes do not bite, but feed on the nectar of flowers (1).

Mosquitoes are highly developed blood-sucking insects and are the most formidable transmitters of disease in the animal kingdom. Mosquito-borne diseases are caused by human parasites that have a stage in their life cycle that enters the blood

stream. The female mosquito picks up the blood stage of the parasite when she imbibes blood to develop her eggs. The parasite generally uses the mosquito to complete a portion of their own life cycle and either multiply, change in form inside the mosquito or do both. After the mosquito lays her eggs, she seeks a second blood meal and transmits the fully developed parasites to the next host (3). Mosquitoes are capable of transmitting diseases such as malaria, yellow fever and dengue to man, encephalitis to man and horses, and heartworm to dogs (1).

1.2 Malaria

Malaria is a common and serious tropical disease. It is a protozoan infection transmitted to human beings by mosquitoes biting mainly between sunset and sunrise. Human malaria is caused by four species of the genus *Plasmodium*: *Plasmodium falcifarum, Plasmodium vivax, Plasmodium ovale* and *Plasmodium malariae* (4). Malaria parasites are transmitted from one person to another by the female anopheline mosquito. The males do not transmit the disease as they feed only on plant juices (5).

There are about 380 species of anopheline mosquito, but only 60 or so are able to transmit the parasite. The anopheline mosquitoes that act as the Indo-Chinese malaria vectors are *Anopheles nigerrimus, An. annularis, An. culicifacies, An. dirus, An. fluviatilis, An. jeyporiensis, An. maculatus* and *An. minimus* (6). Like all other mosquitoes, the anophelines breed in water, each species having its preferred breeding grounds, feeding patterns and resting place. Their sensitivity to insecticides is also highly variable.

The significance of malaria as a health problem is increasing in many parts of the world. The distribution of malaria varies greatly from country to country and

within the countries themselves. In 1990, 75% of all recorded cases outside of Africa were concentrated in nine countries: India, Brazil, Afghanistan, Sri Lanka, Thailand, Indonesia, Vietnam, Cambodia, and China (5).

To control the malaria vector, several commercially insecticides were used in controlling both larval and adult mosquitoes. However, malaria has become more widespread in recent years due to the effects of increasing drug and insecticide resistance (4).

1.3 Insecticides

1.3.1 Classification of insecticides

The insecticides are grouped according to chemical nature and origin: natural, synthetic, organic and inorganic. The major insecticides in use today are (7):

1. Chlorinated hydrocarbon insecticides

The chlorinated hydrocarbon compounds include such important insecticides as DDT, BHC, chlordane and dieldrin. All compounds, which belong to this group, are characterized by:

- The presence of carbon, chlorine, hydrogen and sometimes oxygen atoms, including a number of C-Cl bonds.
- The presence of cyclic carbon chains (including benzene rings).
- Lack of any particular active intramolecular sites.
- Apolarity and lipophilicity.
- Chemical unreactivity i.e. they are stable in the environment.

These chemicals are often considered to belong to the group of organochlorine pesticides.

2. Organophosphorus insecticides

The organophosphorus insecticides (or, more commonly, the OP compounds) are often referred to as organophosphates. Whether they are true phosphates or other types of organophosphorus compounds, these chemicals possess the common characteristic that they are, or can become, excellent inhibitors of cholinesterase. The OP group of insecticides includes such generally toxic compounds as parathion and TEPP and such selective compounds as malathion and ronnel. The OP insecticides are used as stomach and contact poisons, as fumigants, and as systemic insecticides for nearly every type of insect control.

3. Carbamate insecticides

The carbamates are the latest arrival in the field of anticholinesterase insecticides, and at present many new compounds are still in the process of being marketed. In general, carbamate insecticides are synthetic derivatives of physostigmine (commonly called eserine), which is the principal alkaloid of the plant *Physostigma venenosum* (calabar beans). Physostigmine was known to be an inhibitor of cholinesterase. The insecticides in this group include propoxur and bendiocarb.

4. Thiocyanate insecticides

The organic thiocyanates were important insecticides before DDT and the chlorinated cyclodienes came into widespread use. In general, the thiocyanates produce rapid knockdown and paralysis of insects and have been especially used for fly sprays. No insects have developed resistant to these compounds.

5. Inorganic insecticides

Inorganic insecticides are relatively nonspecific, and since they are not too toxic to insects, large quantities are required to control insect pests in the field.

Because of these limitations, organic, particularly synthetic chemicals have gradually replaced inorganic insecticides. Nevertheless, there are two groups of inorganic chemicals that are used today as insecticides: arsenicals and fluorides.

6. Pyrethoids and pyrethrins

Pyrethrum is found in the flowers of plants belonging to the family *Compositae* and the genus *Chrysanthemum*. The species, which possess a high enough toxic content to be used for manufactures are *C. cinerariaefolium* and *C. coccineum*. Pyrethum is essentially nontoxic to mammals and is very fast acting toward insects. The synthetic pyrethroids are pyrethrum analogue such as allethrin, phthalthrin and permethrin.

1.3.2 Resistance mechanisms

Resistance is a process of genetic selection of traits that favor survival of either direct treatment by insecticides or their residues. There is no hard evidence that a given insecticide can cause the appropriate genetic changes by directly interacting with the genes of an insect. The process of induction whereby treatment by a given chemical induces the expression of certain metabolic enzymes has been documented, but this is not considered true resistance. Instead, the resistance occurs by combining natural mutation in a population with selection by insecticide treatment. This process was described by evolutionary biologists as an example of Darwinian "survival of the fittest." In this case the most fit being that portion of a population most capable of withstanding a dose of insecticide (8). Adaptations are mediated by either metabolic or non-metabolic mechanisms, which can be divided into four major categories.

1. Behavioral resistance

The evolution of behavior to avoid contact with insecticides seems a likely means of resistance, especially since some insecticides, such as pyrethoids, act as irritants or repellents (9).

2. Reduced penetration

Alteration in the lipid composition of the insect cuticle can result in decreased rates of insecticide absorption resulting in resistance. Noppun *et al* (10) reported that reduced cuticular penetration appeared to be an important mechanism for resistance to fenvelerate in a resistant strain of the diamondback moth, *Plutella xylostella*. Differences in insecticide transport across the cuticle have also been demonstrated in houseflies (11).

3. Metabolic resistance

This mechanism involves a range of enzymes, which normally detoxify foreign lipophilic chemicals. Many of these enzymes are known to be inducible. The three most important systems of detoxication in insects are (12):

1) The mono-oxygenases (mixed-function oxidases, microsomal oxidases or cytochrome P-450 dependent oxidases), which degrade a wide variety of pesticides or natural toxicants.

2) The hydrolases (including esterases), which degrade carbamate, organophosphate and pyrethoid insecticides as well as the juvenile hormone and its analogues.

3) The glutathione S-transferases, which are importance in metabolism of organophosphate insecticides.

1.4 Glutathione S-transferases

Many sophisticated defense strategies have evolved in organisms enabling them to deal with the constant threat by a board spectrum of both foreign and endogenous cytotoxic and genetoxic compounds (13). Enzyme systems that transact the chemical detoxication and elimination processes of these structurally diverse molecules, many of which are highly non-polar, can be divided into three majors but interrelated groups:

Phase I enzymes, such as the cytochrome P450 superfamily, activate chemicals by forming reactive functional groups (e.g. epoxides) in them.

Phase II enzymes deactivate these reactive chemicals by appending a hydrophobic moiety (e.g. glutathionyl-, glucoronyl-, or sulphuryl-) to the functional group.

Phase III enzymes mediate the cellular elimination of the inactive and water-soluble products.

The glutathione S-transferases (GSTs: E.C.2.5.1.18) a group of multigene isoenzymes represents an integral part of the phase II detoxication mechanism. They are widely distributed in nature, being found in bacteria, yeast, molds, fungi, mollusks, crustacean, worm parasites, frogs, insects, plants, fish, birds and mammals (14). These intracellular enzymes protect cells against both xenobiotic and endogenous compound (15), and stress by catalyzing the nucleophilic addition of the thiol of reduced glutathione (γ-glutamyl-cysteinyl-glycine) to electrophilic centers in organic compounds. The resultant glutathione conjugates, more water-soluble, can be exported from animal cells by putative membrane ATP-dependent pump systems (16-17) after

which they are metabolized via the mercapturic acid pathway (**Figure 1.1**) and eventually eliminated (18).

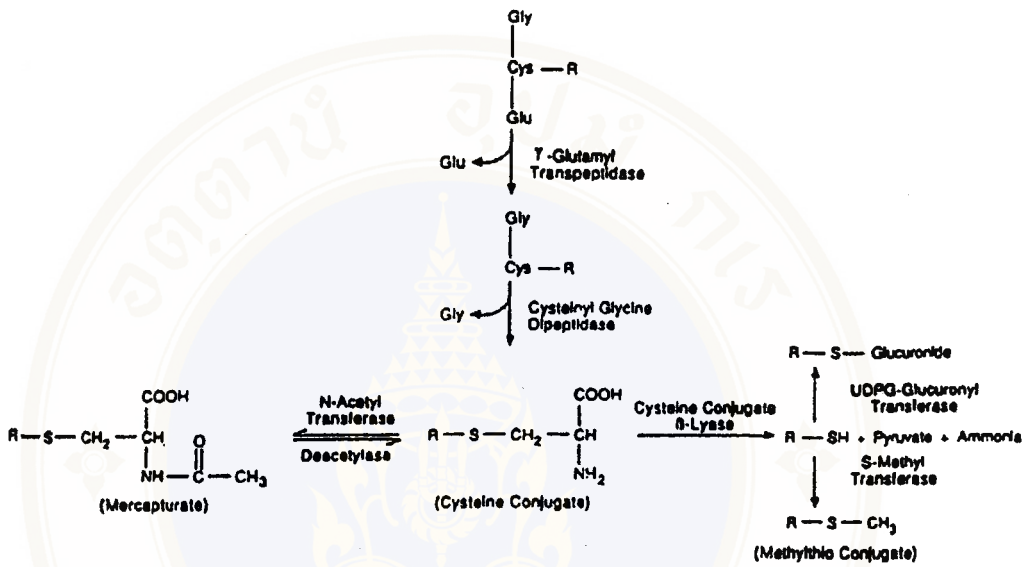


Figure 1.1 Metabolic pathways of the glutathione conjugates.

Essentially all eukaryotic species appear to possess multiple isoenzymes. Multiple forms of GSTs have been discovered in virtually every organism in which GST activity has been found. The cytosolic GSTs exists as either homo- or hetero-dimeric forms due to multiple genes and monomer hybridization (19). Heterodimers have not been identified between molecules of different classes (20). Members within any class exhibit similar monomer sizes, about 20 to 27 kDa, share high amino acid sequence identity, typically 60 to 80%, and have distinctive but overlapping substrate specificities (21). The membrane bound or microsomal GSTs are quite distinct from their soluble counterparts and will not be delt with here.

1.4.1 Physical properties

1.4.1.1 Molecular weight and subunit size

Molecular weights of GSTs have been found in general to fall within the range 36-50 kDa. Almost invariably, they consist of two subunits of reported molecular weight between 20-27 kDa. Molecular weight cited for the native enzymes cluster around a mean of 43.0 ± 0.8 kDa, and for a subunit, a mean of 24.9 ± 3.1 kDa (22).

1.4.1.2 Isoelectric point

Isoelectric points (pI) have been determined for enzymes from a variety of organisms, generally by isoelectrofocussing or by chromotofocussing. In the case of housefly GSTs, it appears that those enzymes with a pI higher than about pH 6.5 are catalytically more versatile than those of low pI, but it does not appear possible generally to establish any relationship between pI and function (23-24). Analysis of data presented by Dierickx (25) shows that the K_m with respect to CDNB appears to vary systemically with isoelectric point for these enzymes.

1.4.1.3 Binding to the affinity gels

The multifunctional nature of GSTs allowed a variety of affinity gels to be designed that can be used to isolate GSTs. These include agarose containing immobilized bromosulfothalein (BSP), choric acids, glutathione, S-hexylglutathione, S-octylglutathione, thyroxine and triazine dye (26). Among these affinity gels two matrices in particular, glutathione-agarose (27) and S-hexylglutathione-agarose (28) have been widely used to purify GSTs as they display both excellent specificity and yield of these enzymes.

1.4.2 GST structure

1.4.2.1 General structure

All known GSTs are dimeric, non-allosteric enzymes with one active site per monomer. While each active site functions independently from the other, the enzymes' quaternary structure is vital for activity since residues from one subunit complete the active site of the other. The enzymes are therefore usually homodimers (22, 29-30).

Polypeptide chain of each monomer folds into two distinct domains; domain I and domain II. They are linked with a short hexamer and consist of 48-59% alpha helix and 8-10% beta strands. Although only about <30% sequence identity exists between the different classes of GSTs, their quaternary structure is remarkably similar as is their configuration within the active sites which includes a conserved Tyr residue among others.

Domain I (residue 1-80) comprises the smaller N-terminal end and consists of 4 strands of β-pleated sheet flanked by α-helices. The glutathione binding site (G-site) is located in this domain whose most interesting feature is the "βαβαββα" topology moeity observed in domain I of other GSH-utilising enzymes such as glutathione peroxidase (31), *E.coli* thioredoxin (32), and T_4 glutaredoxin (33). Removing of the alpha-3 helix showed similarity to domain occurring in glutathione reductase, *p*-hydroxybenzoate hydroxylase, and D-glyceraldehyde 3-phosphate dehydrogenase (34).

Domain II, the larger domain from approximately residue 80 to the C-terminus, contains 5 amphipathic α-helices (α4-α8) arranged in a right-handed spiral. The hydrophobic binding site (H-site) is located here and a few of the conserved residues are involved in forming this hydrophobic site. It is the variability of the residues

within the H-site, which confers the differential reactivity and substrate specificity to the various GSTs. A conserved domain II Asp residue from the other subunit interacts with GSH's amino group (29).

1.4.2.2 Structural similarity between different GST classes

Protein folding is similar for Alpha, Mu and Pi class GSTs and their active sites are in similar positions in space (**Figure 1.2**). Structural differences are a result of changes in the C-terminus. Both Mu and Pi GSTs are shorter than Alpha GST by 4-8 residues. In the Pi and Mu enzymes the C-terminus forms a wall restricting access to the active site. The Mu class has an extended loop (mu loop) between β2 and α2 of domain I while in Alpha class the extra C-terminal residues form a helix (α9) which sits on and helps form the H-site. In the Mu and Pi class the H-site is larger and more open to solvent and has polar residues able to bind to polar substrate atoms whereas the H-site of Alpha is more hydrophobic with a pair of conserved Phe residues in the helix (15).

An active site Tyr is conserved in all the aligned sequences of mammalian, *Schitosoma japonica* and maize GSTs (34), Tyr6 for Mu, Tyr7 for Pi and Tyr9 for Alpha class GSTs. Site directed mutagenesis studies showed that in all cases where the conserved Tyr was replaced by Phe, 90% or more of wild type specific activity was lost (37).

Insect GSTs was categorized within the theta class on the basis of sequence comparisons (37-38). The theta class GSTs have been proposed as evolutionary forerunner of the alpha, mu and pi mammalian GSTs, based on the apparent distribution of the former in a diverse range of organisms including bacteria, yeast, plants and insects (39).

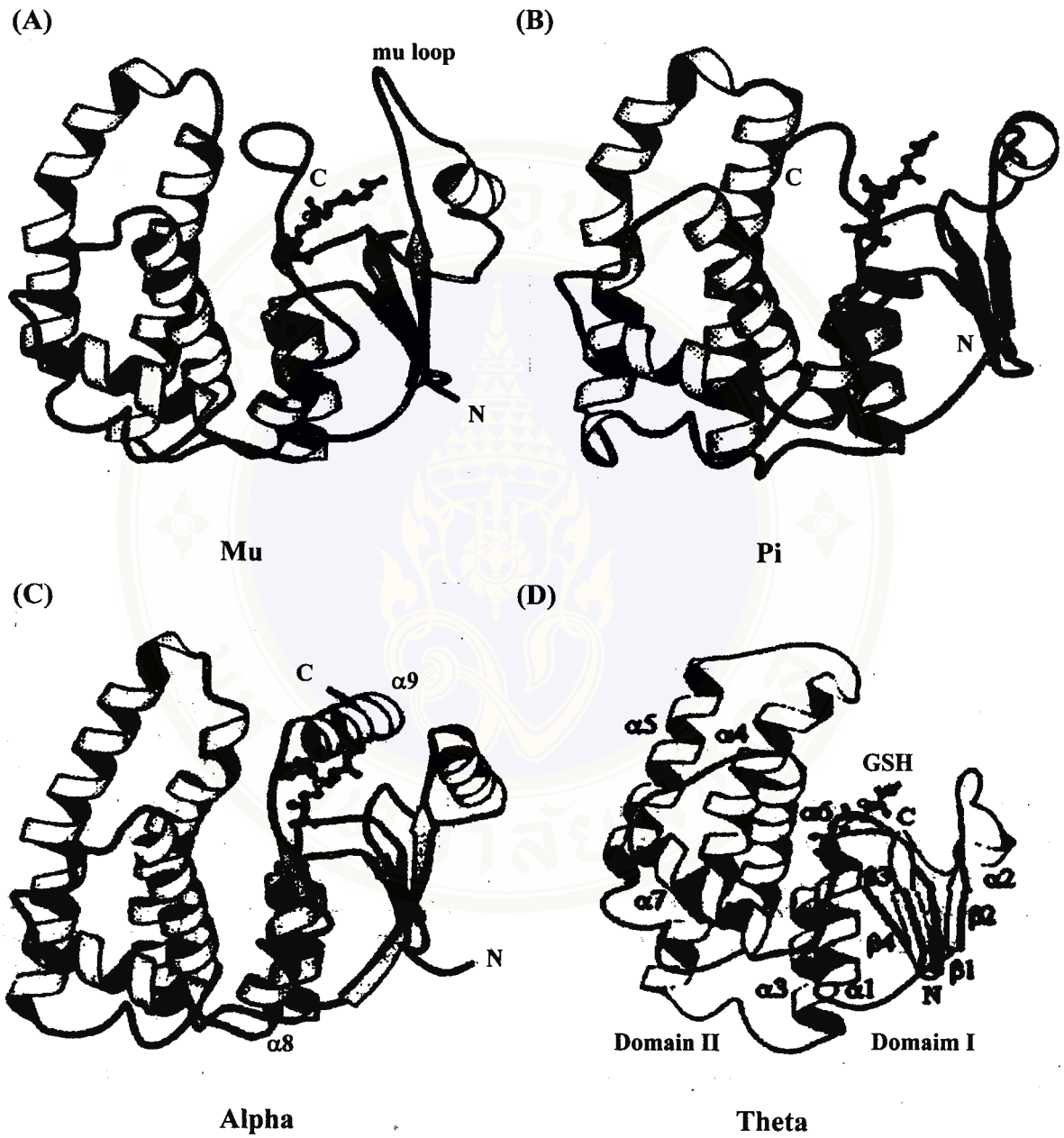


Figure 1.2 Ribbon representation of GST crystal structures. The locations of alpha-helices (denoted α) and beta-strands (denoted β) are indicated. Glutathione (GSH) is shown in ball-and-stick fashion. (A) mu-class rat GST (B) pi-class human GST (C) alpha class human GST (D) theta-class *Lucilia cuprina* GST.

A crystal structure was solved from the Australian sheep blowfly, *Lucilia cuprina* (40). There are a number of obvious differences between the *Lucilia* GST crystal structure and the mammalian GST crystal structure. The *Lucilia* structure does not have the extended mu loop structure characteristic of the mu class enzyme or the C-terminal helix characteristic of the alpha class enzyme. Unlike the mammalian enzymes, helix 5 is shorter and is not bent. The conformation of the linker region between the two domains, the loop connecting helices 4 and 5, and the absence of the helix-8 in the Lucilia structure are other obvious differences. The active site is shown to be much deeper within the protein by having a conserved Ser9 at the N-terminal region (**Figure 1.3**).

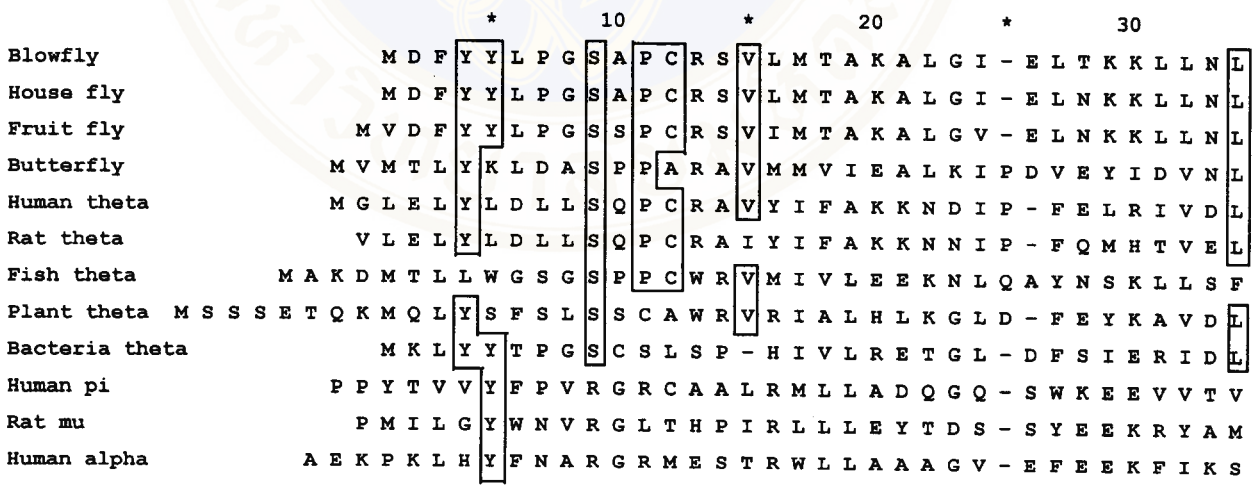


Figure 1.3 N-terminal sequence alignment. The sequence numbering refers to the *Lucilia* amino acid sequence. The conserved residues have been boxed.

1.4.3 Catalytic properties

1.4.3.1 Catalytic mechanism

The GST enzymes are a family of enzymes that catalyze a number of distinct glutathione-dependent reactions: in addition to their ability to catalyze the formation of conjugates. GST can also serve as peroxidases, isomerases and DDT-dehydrochlorinase (19). It is thought that GSH binds first since its cellular concentration (1-10 mM) is considerably higher than the dissociation constant between GSH and the enzyme (41). In general there are 4 types of the interactions with GSH in the binding site (15):

1. Stabilization and orientation of the γ-Glu of GSH
2. Alignment of the glutathione peptide backbone
3. Stabilization of the terminal carboxylate of glycine
4. Interaction with the -SH of cysteine for catalysis

The fundamental basis for all the various catalytic activities of GST is the ability of the enzyme to lower the pKa of the sulfhydryl group of reduced glutathione from 9.0 in aqueous solution to about 6.5 when bound in the active site (42). Glutathione exists as the thiolate (GS^-) anion at neutral pH when complexes with GST (43-45). A basic group on the enzyme (possibly a histidine residue) tends to increase the concentration of the thiolate anion by abstracting the proton from the GSH thiol. It is proposed that once GS^- is formed in the active site of GST, it becomes capable of reacting spontaneously, by nucleophilic attack, with electrophilic xenobiotics that are situated in close proximity (46). Thus catalysis by GST occurs through the combined ability of enzyme to promote the formation of GS^- and to bind hydrophobic electrophilic compounds at a closely adjacent site.

Some GSTs have been shown to have secondary catalytic activities including a selenium-independent peroxidase activity with organic hydroperoxides (47). In addition to their catalytic activity, GSTs may function as intracellular transporters of various non-substrate hydrophobic compounds such as bilirubin, heme, steroids, thyroid hormones and bile salts (48).

1.4.3.2 Substrate specificity

One of the fascinating aspects of GST enzymology is their ability to catalyze reactions towards a large number of structurally diverse substrates. Examples include alkyl- and arylhalides, lactones, epoxides, quinones, esters and activated alkenes (19). Although the range of substrates that GSTs recognize is board, they do share the common feature of being mostly hydrophobic and bearing an electrophilic center. The specific activities of GSTs towards certain substrates have proved useful in classifying new GSTs. For example, Alpha class GSTs are highly reactive towards cumene hydroperoxide, Mu class GSTs have a presence for epoxides, Pi class GSTs display high reactivity towards ethacrynic acid whilst Theta class GSTs are highly reactive towards *p*-nitrophenethyl bromide. However, such classifications can be misleading as, for example, a rat Alpha class GST has been shown to exhibit high reactivity towards aflatoxin B_1-8,9-epoxide (49).

Numerous electrophilic compounds may serve as substrates for GSTs. Most substrates used are products of modern chemical industry and have no biological relevance. The most important substrate used for the demonstration of multiple forms of GST in various biological species is 1-chloro-2,4-dinitrobenzene (CDNB). When conjugated with GSH, it gives a compound (**Figure 1.4**) possessing an absorbance

spectrum sufficiently different from that of CDNB to allow a simple spectrophotometric assay at 340 nm (50-51).

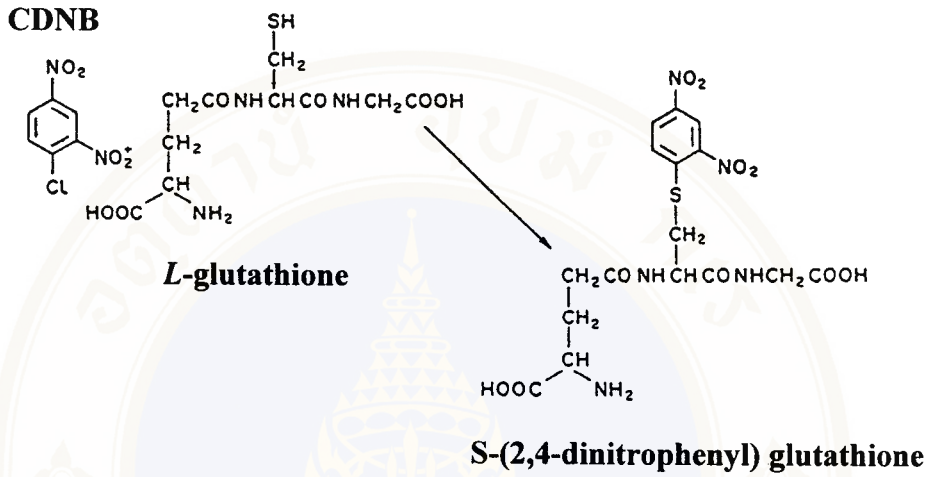


Figure 1.4 Conjugation of CDNB with glutathione by glutathione S-transferase catalysis.

Notwithstanding its significance for detection of GST activity, it should be stressed that certain forms of the enzyme express low activity with this substrate. Rat GST 5-5 (52) and two maize GSTs (53) are examples of enzymes that display low activity with these substrates.

Some of the inter-class differences can readily be described to the structural variations of the C-terminal residues observed in the crystal structures of the different GSTs. Photoaffinity labelling studies (54) and deletion mutants (55) have implicated the C-terminal region of GSTs in substrate selectivity. The C-terminal tail which forms part of substrate binding in alpha, mu and pi class GST crystal structures, is longer in the alpha and mu class enzymes. The extra extension in the alpha class enzyme adopts a helical fold with the hydrophobic face of the helix forming part of the

substrate-binding site wall. As a consequence, the substrate-binding site becomes enclosed by the wall and is smaller and more hydrophobic than the mu and pi classes (35).

The origin of the distinction intra-class substrate specificities observed in GSTs is, at least in part, due to subtle changes in the substrate-binding site. The substitutions either changed the size and shape of existing residues (e.g. Met to Leu or Ala to Val) or introduced some hydrogen-bonding functionality (e.g. Phe to Ser or Ala to Ser) can affect the substrate specificity (35). Another source of substrate specificity may lie in the degree of flexibility of the C-terminal tail which projects into the substrate binding site.

1.4.3.3 Kinetic studies

The basic observation in the analysis of the kinetics of GSTs is that the substrate rate saturation curve is non-Michaelian or non-hyperbolic (56). The kinetic evidence, obtained in studies on widely separated species, supports the basic mechanism being a bimolecular random sequential one. In several instances, the kinetics approximated closely to random order equilibrium mechanism or rapid equilibrium mechanism (57-60), in others, a more complex behavior compatible with a random order steady-state mechanism (61-63). This is important since such kinetic behavior would be generated by a covalently modified enzyme intermediate which, given the highly reactive nature of the substrates, has been a mechanistic possibility which had to be given serious consideration.

1.4.3.4 Inhibition studies

In several studies involving GSTs, inhibition data have proved useful when substrate specificities, immunological data, or other criteria have been insufficient.

Inhibitors that are substrate analogs may be used to probe the active site and the catalytic mechanism. Some inhibitors may be suitable as ligands in affinity chromatography. Other may act in vivo and be involved in normal cellular control mechanisms or be pharmacologically useful substances. Absolute values are not directly comparable since the original data cited include IC_{50} values, inhibition constants (K_i), or simply percent inhibition at a given inhibitor concentration. There are many studies investigating the inhibition kinetics, the data showed that the inhibitors used for studies the inhibition of the GST conjugating reaction appeared to act in variable types; non-competitive, competitive or uncompetitive. Furthermore, they have shown distinct inhibition constant to different isoforms of GSTs (22).

1.4.4 GST isoenzymes

The characterization of cytosolic GST has been greatly facilitated by the availability of affinity chromatography gels to which these enzymes bind. The multifunctional nature of GST has allowed a variety of affinity gels to be designed that can be used to isolate GST. The existence of multiple cytosolic GSTs has necessitated the use of high-specificity or high-resolution analytical techniques, or both to allow identification of isoenzymes and the subunits they comprise. The analytical methods that have proved valuable in identification include SDS-PAGE (64), isoelectric focusing (65-68), reversed-phase high-pressure liquid chromatography (HPLC) (69), electrospray mass spectrophotrometry (70), Western blotting (71), and immunoassay with either polyclonal antibody against purified GST, specific GST peptides, or monoclonal antibodies (72-76).

In mammals, at least 18 isoforms from rat, 15 isoforms from mouse and 20 isoforms from human have been reported (14). There have been many attempts to

classify the different forms of GSTs. The initial classification of GSTs was based on the specificity to organic substrate (77). However, later studies showed that various isoenzymes had broad overlapping patterns of substrate specificity (19). To date, there are 10 distinct classes of cytosolic GSTs on the basis of extensively characterized physiochemical properties, immunological properties and amino acid sequence identity (38, 78). These include the 7 mammalian classes; Alpha, Mu, Pi, Theta (19, 38, 47, 78-79), Kappa (80), Zeta (81) and Omega (82). The other classes include the Sigma in cephalopods and arthropods, the Phi in plants and the Delta in insects (83). Generally, GSTs within a class will have >50% amino acid sequences identity and between classes <30% identity (84). Previously the Delta class in insects has been referred to as Theta class (79).

In recent years, because of construction and characterization of cDNA clones, there has been a rapid increase in the structural information available about glutathione S-transferases and their genes. Sequence analysis of cDNAs has revealed that GST subunits are encoded by multiple gene families and has enabled evolutionary studies between different members of these families (14).

The cDNA clones complementary to the mRNAs encoding the GST subunits from both non-mammalian and mammalian have been constructed and analyzed by a number of laboratories. The nucleotide sequences of GSTs are approximately 630-670 base pairs and the derived amino acid sequences are 210-230 residues. The GST sequences among different organisms are conserved at the N-terminal part while they are divergent at the C-terminal part (14). However, GSTs are highly variable and encoded by complex gene families (18).

1.4.5 Insect GSTs

As in other organisms, insect GSTs exists as dimeric proteins, but their structures are poorly documented and often appear contradictory. Multiple forms of these enzymes have been reported for house fly (84-89), grass grub (90), and *Drosophila* (91-92). There are at least three isoforms present in the mosquitoes, *Aedes aegypti* (93-94), two forms in the sheep blowfly (95) and seven forms in *Anopheles gambiae* (96).

On the basis of N-terminal amino acid sequence comparisons, insect GSTs were classified as belonging to the class theta family (37,79). However, the multiple forms of GST in *Drosophila* can be classified into two groups belonging to distinct gene clusters (97-100). Sequence data for insect GSTs are limited, but there is evidence for at least two classes of insect GSTs; class I and class II (101-102).

The sequence of two insect class II or sigma class GST genes have been published, *aggst2-1* from *Anopheles gambiae* (103) and *DmGST2* from *Drosophila melanogaster* (101), and a GenBank search identified additional class II genes from *Musca domestica*, *MdGST2* (accession no.1170110) and *Manduca sexta, MsGST2* (accession no. 1170115). Both *aggst2-1* and *DmGST2* are single copy genes with no closely related sequences present in the genome of either species.

The insect class I or delta class GSTs, in contrast, are encoded by members of a large gene family. In *D. melanogaster*, 8 divergent intronless genes are found within a 14-kb DNA segment (104), at least 5 different class I GST genes are present in *M. domestica* (105). Furthermore, a class I GST gene was found in *An. dirus*, *adgst1-1* (106) and several class I GST genes are reported within a 15.3-kb DNA segment from *An. gambiae* including an intronless gene, *aggst1-2*, and exons coding for, *aggst1-3*,

aggst1-4, *aggst1-5*, *aggst1-6* and *aggst1-7* (107). The *aggst1-7* is arranged in opposite orientation to the other genes. Furthermore, the *aggst1-3*, *aggst1-4*, *aggst1-5* and *aggst1-6* have been found sharing an N-terminus because of alternative RNA splicing. These evidences indicate that complex gene families encode the insect GSTs.

The GSTs in insect are of interest because they play an important role in insecticide resistant. The different forms of GSTs exhibited different specificities for the insecticide studies. Several data suggests that GSTs have an important role in the acquisition of resistance to insecticides (89). First, for many insecticides, including organophosphorus compounds, cyclodienes, and 1,1,1-trichloro-2,2-bis (*p*–chlorophenyl) ethane, glutathione S-conjugation represents a detoxication pathway. Second, some resistant insects have been found to metabolize insecticides more efficiently via a GSH-dependent route. Third, high levels of GST activity have been detected in some resistant insect strains. Finally, laboratory selection for resistance has been correlated with enhanced GST activity and GST dependent insecticide metabolism. Despite the importance of GST in the resistance process, the genetic mechanism underlying it remains to be elucidated.

The GSTs in anopheline mosquitoes, malaria vectors, are of interest because they are involved in DDT resistance (108). Although malaria control programs are established in all epidemic areas, these mosquitoes can not be eliminated and malaria still exists in human beings. Fractionation of homogenates from larvae of the major African malaria vector, *An. gambiae*, and the Thai vector, *An. dirus*, through sequential column chromatography, showed broadly similar GST activity profiles as shown in **Figure 1.5** (106). Three subgroups (peak IV, V and VI), all contained multiple GST isoenzymes which were able to detoxify the DDT in both species (96,

109-110) and were found in an increased amount and specific activity for DDT dehydrochlorination in the DDT resistant *An. gambiae*. Seven recombinant *An. gambiae* GST isoenzymes were obtained through the PCR amplification of cDNA using specific primers to class I insect GSTs (107, 108, 111) while only one recombinant class I GST had been amplified from *An. dirus* (106). These data indicate that there are other isoenzymes, which can be obtained from *An. dirus*.

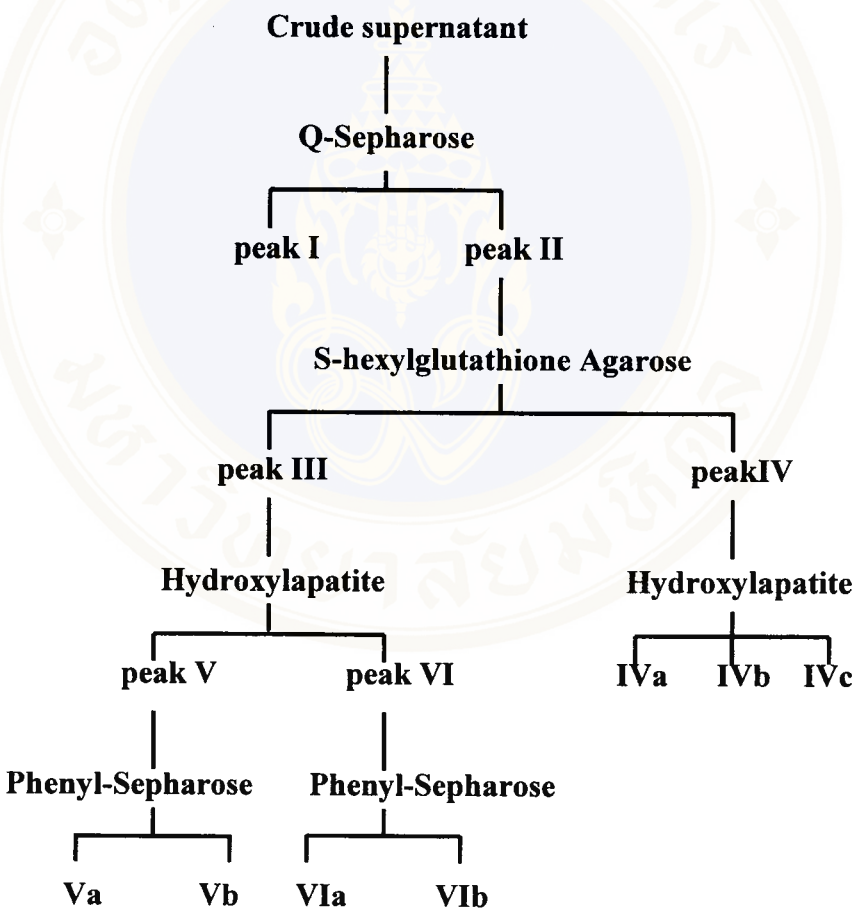


Figure 1.5 Scheme for the purification of the glutathione S-transferases from *An. gambiae* and *An. dirus*.

Recently, the 7.5 kb of the *adgst1AS1* gene (**Figure 1.6**) was identified from an *An. dirus* genomic library (112). This gene showed >78.5% nucleotidesequence identity to *aggst1α* from an *An. gambiae*. The arrangement and the nucleotide sequence identity of each exon between these two genes are similar. The alternatively spliced products of *adgst1AS1* gene are predicted to produce four mature transcripts as in the *aggst1α* gene. One of the splicing products, *adgst1-1*, which is the splicing product of exon 1,2 and exon 3D, has been characterized previously by Prapanthadara *et al* (106). However, cloning and characterization of the other three splicing products of the *adgst1AS1* gene is of interest to determine the difference between the splicing products.

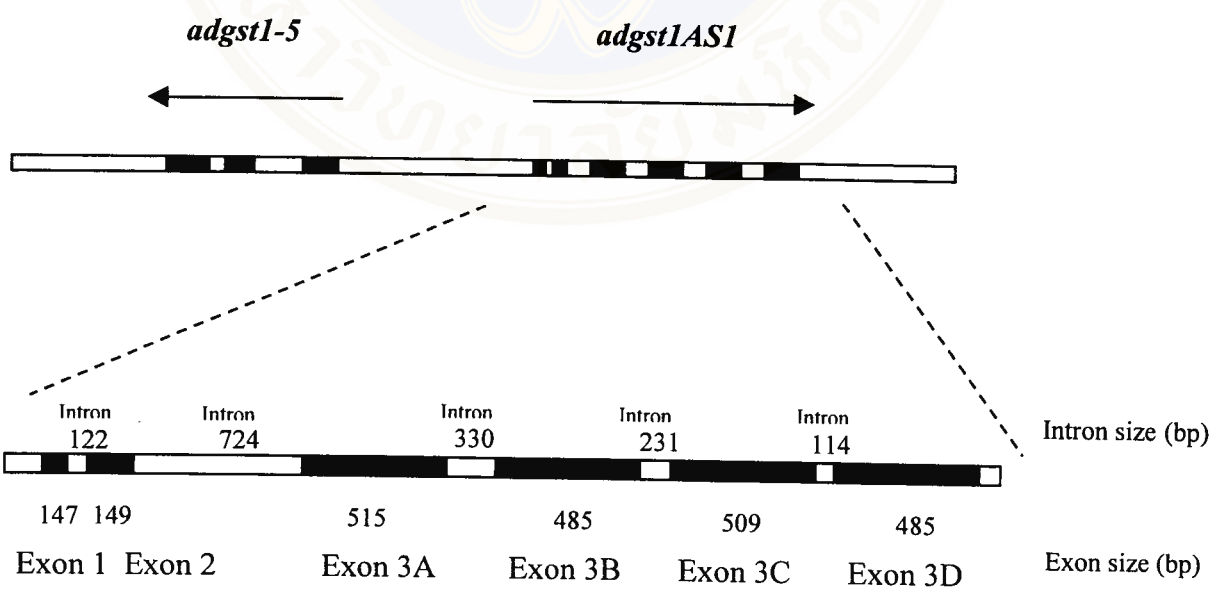


Figure 1.6. The *adgst1AS1* and *adgst1-5* genes organization of 8A.2 clone.

1.5 Objectives

In this thesis, the three remaining alternatively spliced products of the *adgst1AS1* gene were cloned to confirm whether this gene could produce all of alternatively spliced products. The recombinant proteins were further enzymatically characterized and compared between the other GST isoenzymes. The different C-termini of these GSTs are expected to be involved in the determination of their substrate specificities, kinetic properties and inhibition properties.

To achieve the above goals, it was necessary to: -

1. Obtain class I GST cDNAs from the 4th instar larvae of *Anopheles dirus* B mosquito by the reverse transcriptase polymerase chain reaction (RT-PCR) using primers designed based on the *adgst1AS1* gene.

2. Clone these GST sequences into the *NdeI* site in the pET3a expression vector to eliminate the T_7-tag leader, which may effect the activity of the GST protein.

3. Express and purify the recombinant GST proteins for further characterization.

4. Characterize the purified GST proteins in three properties;

-Substrate specificity using CDNB, DCNB, ethacrynic acid and p-nitrophenethyl bromide as substrates.

- Kinetic properties of CNDB conjugation: V_{max}, K_m, k_{cat} and k_{cat}/K_m.

- Inhibition studies using permethrin as inhibitor.

5.Compare all results obtained from this experiment with class I insect GSTs and other organism' GSTs to determine the relationship of the isoenzymes among those species.

CHAPTER II

MATERIALS AND METHODS

2.1 Mosquito strain

The Department of Parasitology, Faculty of Medicine, Chiangmai University, established the *Anopheles dirus* B colony used in this study. The colony was identified on the basis of its morphological and chromosomal characteristics. This strain was reviewed by Baimai (113).

2.2 Total RNA isolation

2.2.1 Total RNA extraction by TRIzol™ Reagent

The 4th instar larvae of *Anopheles dirus* mosquito were frozen immediately in liquid nitrogen and homogenized by a mortar and pestle under liquid nitrogen to a powder. Then 1 ml of TRIzol™ Reagent, a mono-phasic solution of phenol and guanidine isothiocyanate, was added per 50-100 mg of larvae. The homogenized sample was incubated for 5 minutes at room temperature to permit the complete dissociation of nucleoprotein complexes. Then 0.2 ml of chloroform was added per 1 ml of TRIzol™ Reagent. The sample was shaken vigorously by hand for 15 seconds and incubated at room temperature for 2-3 minutes. Then the sample was centrifuged at 12,000x *g* for 15 minutes at 4°C. Following centrifugation, the mixture was separated into a lower red, phenol-chloroform phase, an interphase, and a colorless upper aqueous phase. RNA remains exclusively in the aqueous phase. The aqueous phase was transferred to a fresh tube and the total RNA was precipitated by mixing

with 0.5 ml of absolute ethanol per 1 ml of TRIzol™ Reagent used for the initial homogenization. The sample was incubated at room temperature for 10 minutes and centrifuged at 12,000x *g* for 10 minutes at 4°C. The supernatant was removed and the RNA pellet was washed once with 75% ethanol. The sample was then mixed by vortexing and centrifuged at 7,500x *g* for 5 minutes at 4°C. The RNA pellet was briefly air dried for 5-10 minutes. The total RNA was dissolved in RNase-free water or DEPC treated water and incubated for 10 minutes at 55°C- 60°C.

The concentration and purity of the isolated total RNA were determined by spectrophotometry. The absorbances were measured at 230, 260 and 280 nm (A_{230}, A_{260} and A_{280}). The RNA concentration was calculated by assuming that a 40 µg/ml RNA solution will have an absorbance of 1.0 at 260 nm. The RNA purity was determined from A_{260}/A_{230} ratio and A_{260}/A_{280} ratio. Both ratios should be approximately 2.0. An A_{260}/A_{230} ratio less than 2.0 indicates the contamination of guanidine isothiocyanate whereas an A_{260}/A_{280} ratio less than 2.0 indicates the contamination of protein. If either case occurs, the isolated RNA should be precipitated again.

2.2.2 Total RNA clean up

To purify RNA from protein and guanidine isothiocyanate contamination and to eliminate all RNA molecules shorter than 200 nucleotides such as degraded RNAs and small RNAs, the total RNA isolated by the previous method was cleaned up using RNeasy® Mini Kits (QIAGEN) as described in the manufacturer's instructions. The purified RNA was eluted in 30-50 µl of RNase-free water per 100 µg of initially loaded RNA. The concentration and purity of RNA were determined as described previously. The RNA solution was stored at –70°C until used.

2.3 Primer design

The primers used in these experiments have been designed according to the 5' and 3' sequences of *Anopheles dirus*, *adgst1AS1* gene. The sets of primers were designed for cloning into the pUC19 cloning vector and the pET3a expression vector.

The sequences of the primers that were used for cloning into the pUC19 vector

*Bam*HI
Forward primer: **Oligo1** 5'CCGGCGGGATCCATGGATTTTTATTACCTACCC 3'

Reverse primers: **adgst1-2** 5'TCAATGCTTAATCCGATCGAAAAACG 3'

adgst1-3 5'CGCCGTCGACATATGTTACTTCTCAAAGTACTT 3'

adgst1-4 5'TCATTTTTGTGTGAAGCGCCCGA 3'

The sequences of the primers that were used for cloning into the pET3a vector

*Nde*I
Forward primer: **Jane1** 5'CCGAGAGCATATGGATTTCTACTACCTTCCC 3'

Reverse primers: **adgst1-2** 5'TCAATGCTTAATCCGATCGAAAAACG 3'

*Nde*I
adgst1-3 5'CGCCGTCGACATATGTTACTTCTCAAAGTACTT 3'

adgst1-4 5'TCATTTTTGTGTGAAGCGCCCGA 3'

2.4 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)

2.4.1 The first strand cDNA synthesis

First strand cDNA was synthesized by using SUPERSCRIPT^{TM} II RNase H^{-} Reverse Transcriptase (GIBCO BRL). For 20 µl of reaction, 1-5 µg of total RNA was mixed with 500 ng/µl of oligo$(dT)_{15}$ primer, then DEPC treated water was added to give a 12 µl reaction volume. The mixture was heated to 70°C for 10 minutes then

quick chilled on ice. The final concentration of 1X First Strand Buffer (50 mM Tris-HCl pH 8.3 at room temperature, 75 mM KCl and 3 mM $MgCl_2$), 10 mM DTT and 50 μM each of dATP, dGTP, dCTP and dTTP were added to the mixture. The mixture was mixed gently by pipetting and incubated at 42°C for 2 minutes. Then 200 units of SUPERSCRIPT™ II RNase H⁻ Reverse Transcriptase was added, mixed gently by pipetting and incubated at 42°C for 50 minutes. The reaction was inactivated with heating at 70°C for 15 minutes. To remove RNA complementary to the cDNA, 1 unit of *E.coli* Rnase H was added and incubated at 37°C for 20 minutes.

To eliminate oligo$(dT)_{15}$ primers, nucleotides, reverse transcriptase enzyme and salts, the synthesized first strand cDNAs were purified by QIAquick PCR Purification Kit (QIAGEN) as described in the manufacturer's instructions.

2.4.2 Polymerase Chain Reaction (PCR)

The PCR reactions were performed using Oligo1 primer and the specific reverse primers for *adgst1-2, adgst1-3* and *adgst1-4*. The 50 μl of reaction consisted of 500 ng of the first strand cDNA, 30 pmol of each forward and reverse primer, 1X ThermoPol Reaction Buffer (10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl pH 8.8 at 25°C, 2 mM Mg_2SO_4 and 0.1% TritonX-100), 200 μM each of dATP, dGTP, dCTP and dTTP, and 0.5 units of $Vent_R$® DNA Polymerase (BioLabs). The PCR amplification was performed using a P-E Thermal Cycler 2400. After the initial denaturing step at 94°C for 5 minutes, 35 cycles were performed consisting of the denaturing step at 94°C for 30 seconds, the annealing step specific for *adgst1-2, adgst1-3* and *adgst1-4* at 60°C, 64°C or 62°C respectively, for 30 seconds and the extension step at 72°C for 1 minutes. The final extension step was performed at 72°C

for 7 minutes. The PCR products were kept at 4°C and analyzed by agarose gel electrophoresis.

2.5 Agarose gel electrophoresis

The PCR products were analyzed by 1% agarose gel in 1X TBE buffer (0.09 M Tris-borate and 2 mM EDTA pH 8.0). Loading buffer (0.25% bromophenol blue and 30% glycerol in water) was added to the DNA sample, mixed and loaded into the wells. The electrophoresis was performed at a constant voltage, 120 volts for 1 hour. The gel was stained with 0.5 μg/ml of ethidium bromide solution in water for 10 minutes and destained with distilled water for 15 minutes. The gel was observed under ultraviolet (UV) light.

2.6 Subcloning into the pUC19 cloning vector

2.6.1 DNA digestion

The PCR products were purified from agarose gel with a GENECLEAN II® Kit as described in the manufacturer's instructions. After purification, 500 ng of PCR product were digested with 10 units of *Bam*HI (BioLabs). The 20 μl of reaction consisted of 1X NEBuffer *Bam*HI (150 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, and 1 mM DTT, pH 7.9 at 25°C) and 100 μg/ml of bovine serum albumin. The reaction was incubated at 37°C for 3 hours. The pUC19 vectors were digested with *Bam*HI (BioLabs) and then digested with 10 units of *Hin*cII (BioLabs). The 20 μl of reaction consisted of 1X NEBuffer 3 (100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, and 1 mM DTT, pH 7.9 at 25°C) and 100 μg/ml of bovine serum albumin. The

reaction was incubated at 37°C for 3 hours. The digested DNA was purified from solution with a GENECLEAN II® Kit after each enzyme digestion. The construction of the recombinant plasmid in pUC19 is shown in **Figure 2.1**.

2.6.2 DNA ligation

The purified digested PCR products and pUC19 vector were incubated in the ratio 3:1 (insert:vector). The 20 µl of reaction consisted of 1 unit of T_4 DNA ligase (GIBCO BRL), 1X DNA Ligase Reaction Buffer (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT and 5% (w/v) polyethylene glycol-8000) and distilled water. The ligation mixtures were incubated at 16°C overnight and kept at 4°C until used.

2.7 Preparation of competent cells by the Simple and Efficient Method (SEM).

About 10-12 colonies of *E. coli* DH5α host for the pUC19 vector or BL21 (DE3)pLysS host for the pET3a vector were inoculated into 250 ml of SOB media (20 g of bacto-tryptone, 5 g of bacto-yeast extract and 0.5 g of NaCl in 1 litre of distilled water) in a 2-litre flask with shaking at 18°C until the OD at 600 nm was approximately 0.6. The *E. coli* culture was incubated on ice for 10 minutes and then transferred to a 500 ml-centrifuge bottle. The culture was centrifuged at 2,500x *g* at 4°C for 10 minutes. The cell pellets were resuspended in 80 ml of ice-cold TB buffer (10 mM PIPES, 55 mM $MnCl_2$, 15 mM $CaCl_2$ and 250 mM KCl), incubated on ice for 10 minutes and centrifuged at 2,500x *g* at 4°C for 10 minutes. The cell pellets were collected and resuspended in 20 ml of ice-cold TB buffer. Then dimethyl sulfoxide (DMSO) was slowly added with stirring to give a 7% final concentration and the mixture was incubated on ice for 10 minutes. One hundred microlitres of competent

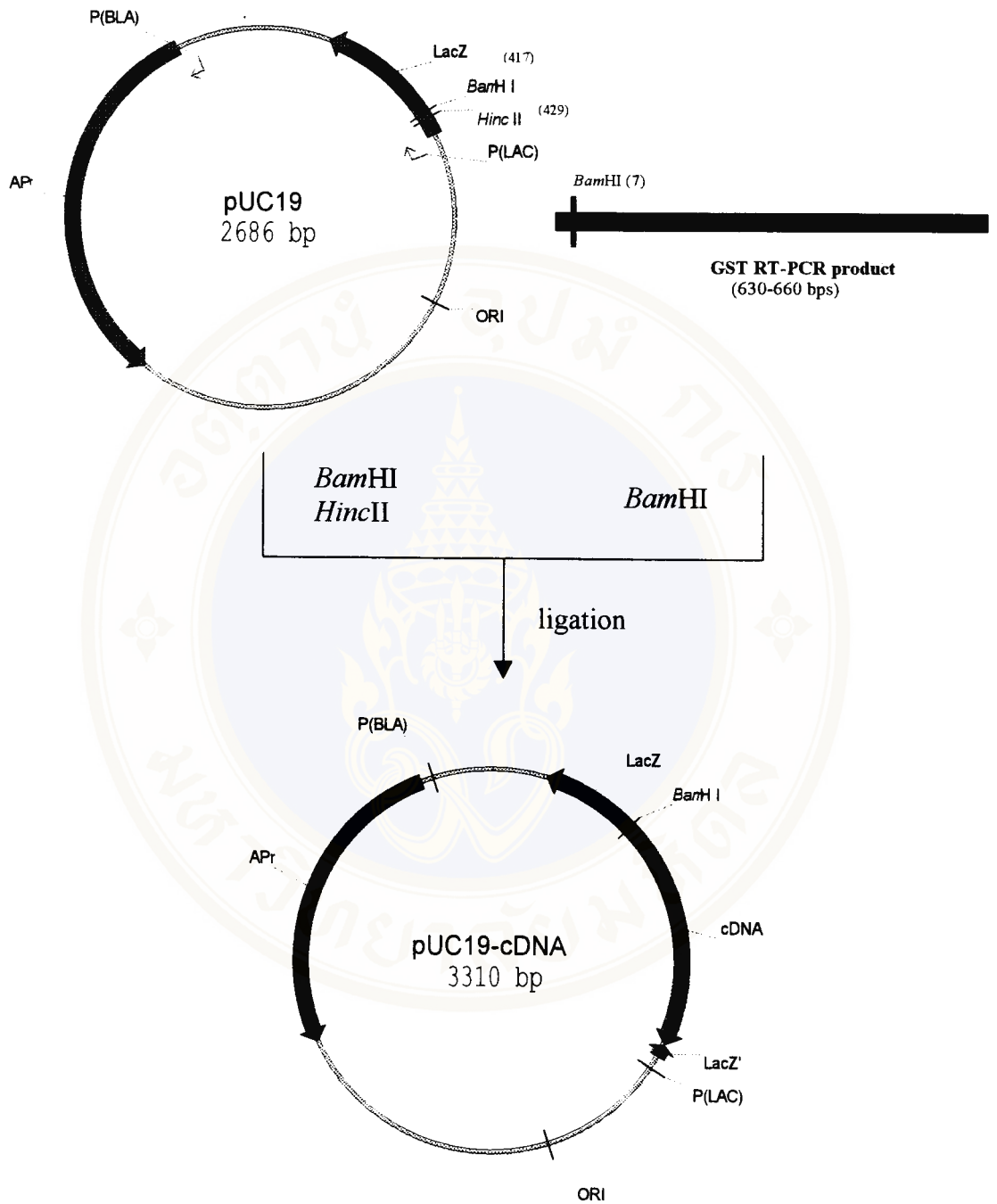


Figure 2.1 Construction of the recombinant plasmid in the pUC19 vector. The figure illustrates construction of the recombinant plasmid containing the GST gene cloned into *Bam*HI and *Hinc*II sites of the pUC19 vector. The pUC19 contains the LacZ promotor (pLac) and the ampicillin resistant gene (*bla*) which is used as the screening marker for the recombinant clones.

cells were aliquoted immediately into 1.5 ml-microtubes and kept at -70°C until used for transformation.

2.8 Transformation of competent cells

Five microlitres of ligation mixture were added to 100 µl of competent cells and incubated on ice for 30 minutes. The cells were heat-shocked at 42°C for 90 seconds and immediatedly incubated on ice for 5 minutes. The transformed cells were added to 500 µl of LB broth (10 g of bacto-tryptone, 5 g of bacto-yeast extract and 10 g of NaCl in 1 litre of distilled water) and incubated at 37°C for 1 hour. Two hundred microlitres of transformed cells were spread on LB agar plates (1 litre of LB broth, 15 g of bacto-agar) containing 100 µg/ml ampicillin, 400 µg of X-gal in dimethylformamide and 800 µg of IPTG in water for *E. coli* DH5α host or 100 µg/ml ampicillin and 34 µg/ml chloramphenicol for *E. coli* BL21(DE3)pLysS host. The LB agar plates were incubated at 37°C overnight.

2.9 Screening of recombinant clones

2.9.1 Blue-white selection

When an insert was cloned into the polylinker region of an α-complementing plasmid, pUC19, the protein coding region of the α-fragment of β-galactosidase was disrupted. Thus a pUC19 vector with an insert fails to α-complement in *E.coli* DH5α strain expressing the β-galactosidase, and the host cells remain β-gal$^-$ in phenotype. In contrast, plasmids without insert will α-complement, converting the host cell from the β-gal$^-$ to β-gal$^+$ phenotype. A mixture of colonies transformed with either vector alone

or vector-ligated to insert were grown on a LB agar plate containing X-gal, a chromogenic substrate for β-galatosidase. Colonies transformed with vector alone were blue, while those transformed with vector plus insert appeared white. This test was used for rapid screening to find a small number of potential recombinants from numerous colonies on a plate.

2.9.2 Simplified Rapid Size Screening Protocol

A single colony was incubated in 50 µl of prewarmed Lysis buffer (5 mM EDTA, 10% (w/v) sucrose, 0.25% (w/v) SDS, 100 mM NaOH, 60 mM KCl and 0.05% (w/v) bromophenol blue) at 37°C for 5 minutes. The lysate was immediately incubated on ice for 5 minutes and centrifuged at 12,000 rpm for 1 minutes. Twenty microlitres of supernatant were loaded onto a 1% agarose gel. The gel was run at 80 volts for 2 hours and observed under ultraviolet light after staining. Plasmid containing a DNA insert will show an increased band size compared to that of no insert DNA.

2.10 Cetylmethylammonium bromide (CTAB) plasmid mini preparation

A single colony of *E. coli* was incubated in 3ml of LB broth containing 100 µg/ml ampicillin for DH5α host or 100 µg/ml ampicillin and 34 µg/ml chloramphenicol for *E. coli* BL21(DE3)pLysS host with shaking at 37°C overnight. The *E. coli* culture was transferred to 1.5 ml-microtube and centrifuged at 5,000 rpm for 2 minutes. The cell pellets were resuspended in 200 µl of STET buffer (8% sucrose, 0.1% TritonX-100, 50 mM EDTA and 50 mM Tris pH 8.0) and mixed by vortexing. The resuspended cells were added with 5 µl of 50 mg/ml of lysozyme, mixed by vortexing and incubated at room temperature for 20 minutes. The mixture was boiled at 100°C for 45 seconds and centrifuged at 12,000 rpm for 10 minutes. The cell pellets were

recovered then 1/10 volume of 5% (w/v) CTAB solution was added, mixed by inversion and let stand for 10 minutes. The mixture was centrifuged at 12,000 rpm for 10 minutes then the pellet was dissolved with 300 µl of 1.2 M NaCl. Then 10 µl of 1 mg/ml RNase were added and incubated at 37°C for 15 minutes. An equal volume of chloroform was added, mixed by inversion and centrifuged at 12,000 rpm for 10 minutes. The aqueous phase was transferred to a new 1.5 ml-microtube and 2 volumes of absolute ethanol were added. After incubation at -80°C for 15 minutes, the mixture was centrifuged at 12,000 rpm for 10 minutes. The pellets were washed with 70% ethanol twice and briefly air-dried. The pellets of plasmid DNA were dissolved in 30 µl of sterile distilled water and kept at -20°C.

2.11 DNA sequencing

The DNA sequencing reactions were performed with the ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer; ddNTP-Dye Terminators labeled with dRhodamine acceptor dyes, dNTPs, AmpliTaq® DNA Polymerase, $MgCl_2$ and Tris-HCl buffer pH 9.0) as described in the manufacturer's instructions. The reactions were then analyzed on the ABI PRISM 377 (Perkin Elmer) in both forward and reverse directions. Twenty-five cycles of PCR for sequencing were performed using the following conditions; the denaturing step at 96°C for 10 seconds, the annealing step at 50°C for 5 seconds and the extension step at 60°C for 4 minutes. The products from thermal cycling were transferred to a new 0.5 ml-microtube and added to 80 µl of sterile distilled water, 1/10 volume of 3 M sodium acetate and 2.5 volume of 95% ethanol. The mixture was incubated at -20°C overnight and centrifuged at 12,000 rpm for 15 minutes. The pellet was washed with 70%

ethanol once and briefly air-dried. The pellets were resuspended with 2 µl of blue dextran solution [5 µl of deionized formamide and 50 mM EDTA pH 8.0 (50 mg of blue dextran in 1 ml of EDTA)]. The samples were heated at 90°C for 2 minutes and quickly cooled before loading onto the sequencing gel.

2.12 Subcloning into pET3a expression vector

2.12.1 PCR reaction

The subcloning step was performed by the PCR reaction using the Jane1 primer and reverse primers specifically to *adgst1-2*, *adgst1-3*, *adgst1-4*. The PCR conditions were performed as described in the RT-PCR conditions.

2.12.2 DNA digestion

The PCR products were purified from agarose gel using GENECLEAN II® Kit. After purification, 500 ng of PCR products were digested with 10 units of *Nde*I (BioLabs). The 20 µl of reaction consisting of 1X NEBuffer 4 (50 mM potassium acetate, 20 mM Tris acetate, 10 mM magnesium acetate, and 1 mM DTT, pH 7.9 at 25°C) was incubated at 37°C overnight. The digested DNA was purified by GENECLEAN II® Kit and the DNA concentration was estimated on a 1% agarose gel.

2.12.3 pET3a vector preparation for adgst1-2 and adgst1-4 subcloning

The pET3a vector was digested with *Bam*HI (BioLabs) as described for the pUC19 vector and then was filled in to make blunt ends with T_4 DNA Polymerase (BioLabs). The 100 µl of fill-in reaction consisted of 1X T_4 DNA Polymerase Buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, and 1 mM DTT, pH 7.9 at 25°C), 100 µM dNTPs, 50 µg/ml BSA, 3 units of T_4 DNA Polymerase and sterile distilled water. The mixture was incubated at 12°C for 20 minutes and immediately inactivated at

75°C for 10 minutes. The blunt end vectors were digested with 20 units of *Nde*I (BioLabs) as described for the PCR products. The digested DNA was purified with a GENECLEAN II® Kit after each step and the DNA concentration was estimated on a 1% agarose gel. The constructions of the *adgst1-2* and *adgst1-4* recombinant plasmids in the pET3a vectors are shown in **Figure 2.2 and 2.4**.

2.12.4 pET3a vector preparation for adgst1-3 subcloning

The pET3a vector was digested with 20 units of *Nde*I (BioLabs) as described for the PCR products. The digested vector was dephosphorylated to prevent recircularization by calf intestinal alkaline phosphatase or CIAP (Promega). The 100 μl of CIAP-reaction consisted of CIAP 1X Reaction Buffer (50 mM Tris-HCl pH 9.3, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$ and 1 mM spermidine), up to 20 pmol of 5'-ends and distilled water. The reaction was added with 0.1 unit of CIAP and incubated at 37°C for 30 minutes. The last step was repeated 3 times. The digested vector was purified with a GENECLEAN II® Kit after each step and the DNA concentration was estimated on a 1% agarose gel. The construction of the *adgst1-3* recombinant plasmids in the pET3a vectors is shown in **Figure 2.3**.

2.12.5 DNA ligation

The digested PCR products were ligated with the prepared pET3a vectors in the ratio 10:1 (insert:vector) and transformed into *E. coli* DH5α competent cells. The positive clones were selected using the Simplified Rapid Size Screening Protocol. The recombinant plasmids were extracted by the CTAB method and sequenced in both directions. The positive clones were then transformed into *E. coli* BL21(DE3)pLysS competent cells.

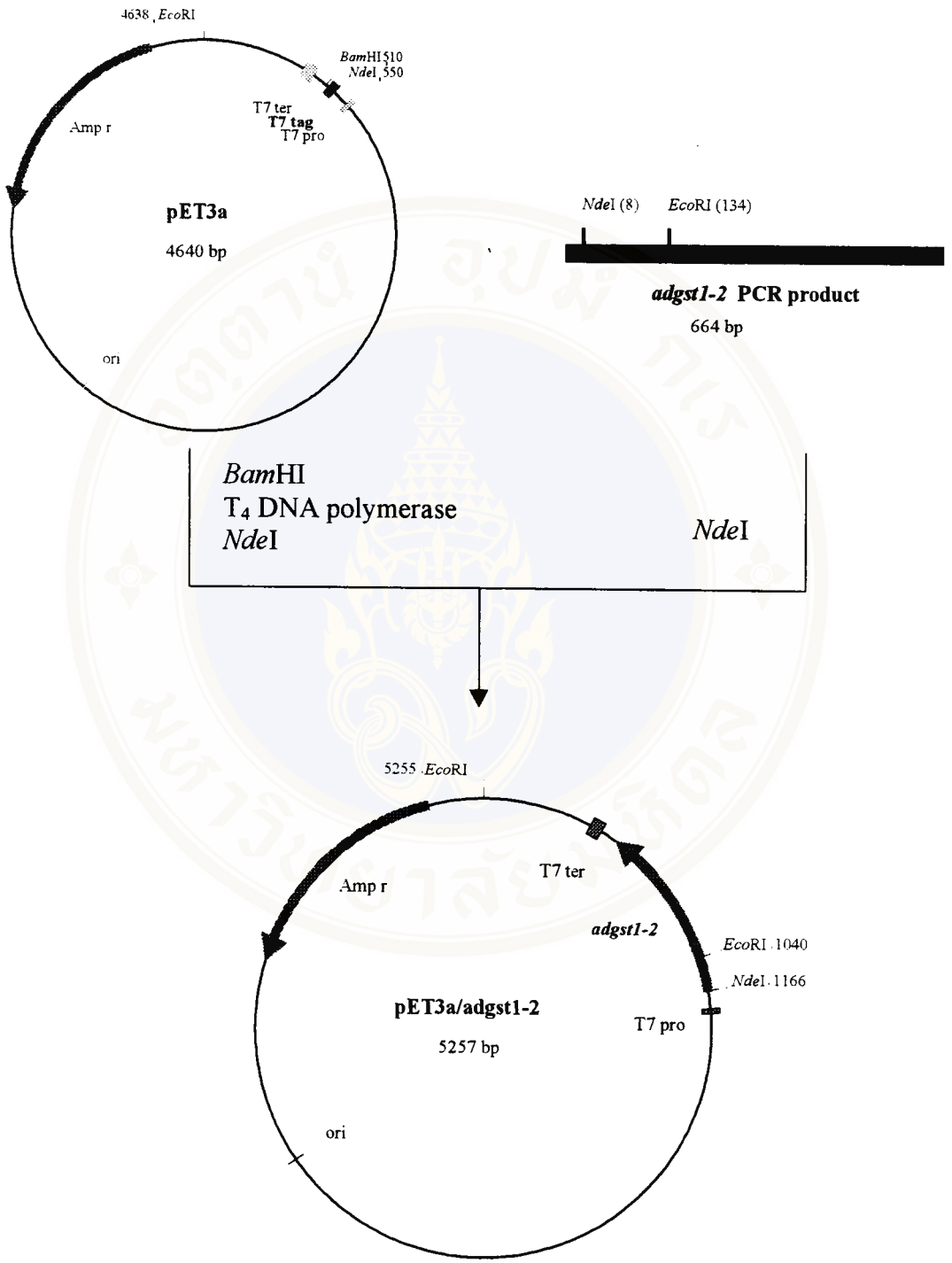


Figure 2.2 Construction of the *adgst1-2* recombinant plasmid in the pET3a vector. The figure illustrates construction of the recombinant plasmid containing the *adgst1-2* gene cloned into the pET3a vector. The pET3a contains the ampicillin resistant gene for screening and T_7promotor for expression.

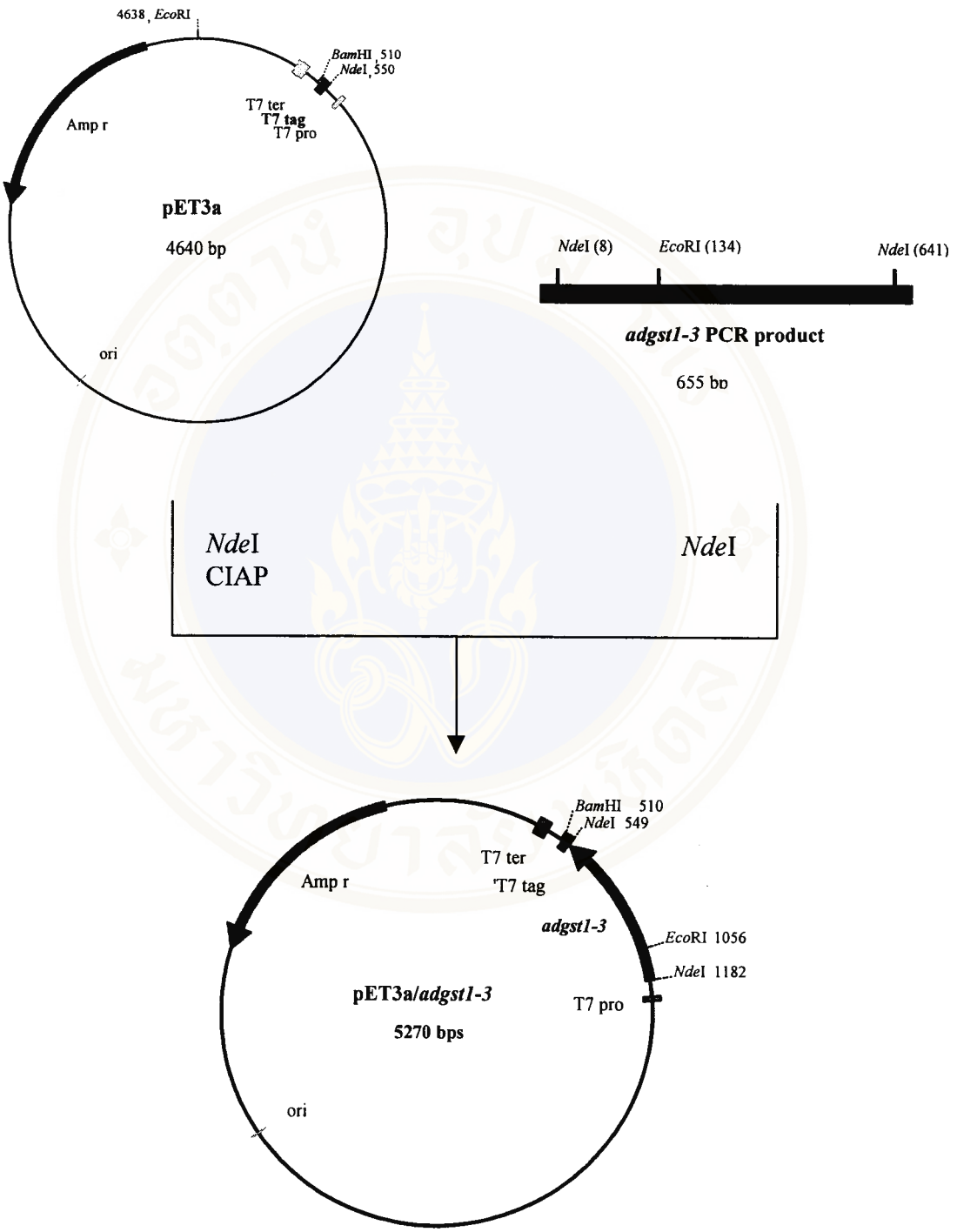


Figure 2.3 Construction of the *adgst1-3* recombinant plasmid in the pET3a vector. The figure illustrates construction of the recombinant plasmid containing the *adgst1-3* gene cloned into the pET3a vector. The pET3a contains the ampicillin resistant gene for screening and T_7promotor for expression.

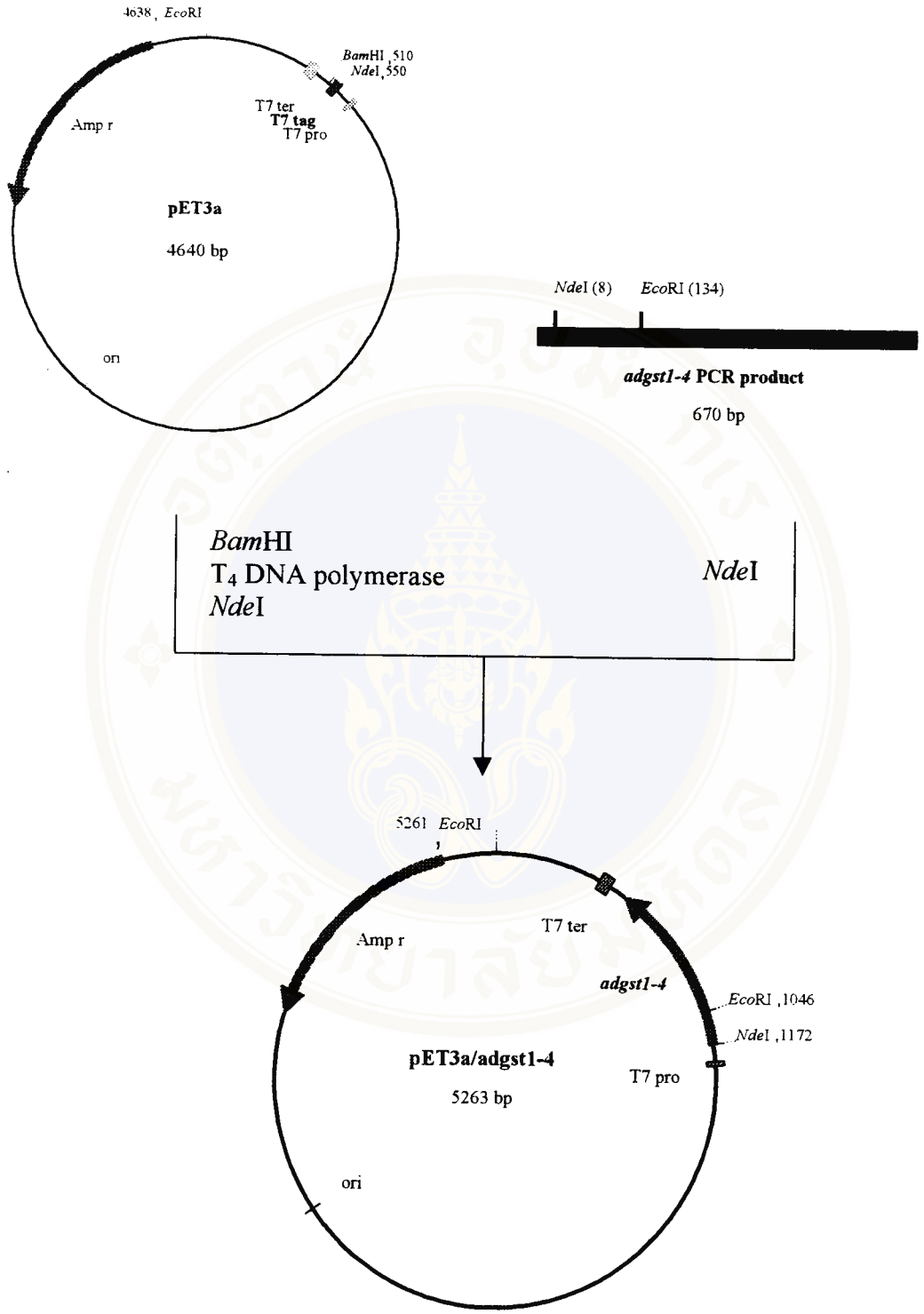


Figure 2.4 Construction of the *adgst1-4* recombinant plasmid in the pET3a vector. The figure illustrates construction of the recombinant plasmid containing the *adgst1-4* gene cloned into the pET3a vector. The pET3a contains the ampicillin resistant gene for screening and T_7-promotor for expression.

2.13 Checking the insert orientations in the pET3a vectors

The recombinant plasmids were digested with 5 units of *Eco*RI (Boehringer) to determine the orientation of the insert. The 10 µl of reaction consisted of 1X Buffer H (100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, and 1 mM DTT, pH 7.5 at 37°C) and sterile distilled water. The mixture was incubated at 37°C for 1.5 hour. The mixture was then analysed by agarose gel electrophoresis.

2.14 Expression of recombinant clones

A colony of *E. coli* BL21(DE3)pLysS which contained a recombinant plasmid was incubated in 3 ml of LB broth containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol at 37°C overnight. An aliquot of culture equal to 1% of the final induction culture was transferred to fresh LB broth containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol. The culture was incubated until the OD at 600 nm was about 0.6 and then induced with a final concentration of 0.1 mM IPTG for 3 hours. The culture was then incubated on ice for 20 minutes, transferred to a 50 ml-centrifuge tube and centrifuged at 5,000 rpm, 4°C for 10 minutes. The pellets were collected and kept at -20°C until used.

2.15 Preparation of cell lysate

The pellets of a 50 ml induced culture were resuspended by mixing with 4.8 ml of buffer (50 mM Tris-HCl pH 7.4, 1 mM EDTA), 200 µl of 100 mg/ml lysozyme and 3.6 µl of β-mercaptoethanol. The mixture was incubated at 4°C until no clumped cells were observed and then 50 µl of 1M DTT was added. The suspension was lysed by French Press and centrifuged at 10,000 rpm, 4°C for 20 minutes. The supernatant

containing the soluble form of the recombinant protein was collected and kept on ice for affinity chromatography.

2.16 Protein purification

2.16.1 S-hexylglutathione affinity chromatography

The soluble target protein was purified using S-hexylglutathione immobilized on agarose (Sigma). S-hexylglutathione agarose gels were packed in a column and equilibrated with ten bed volumes of Equilibrating buffer (50 mM Tris-HCl pH 7.4, 1 mM EDTA and 0.2 M NaCl). The supernatant sample was applied to the column while collecting the buffer washing through the column. The column was washed with five bed volumes of Washing buffer (50 mM Tris-HCl pH 7.4, 1 mM EDTA and 1 M NaCl) while collecting the buffer washing through the column. The recombinant enzyme was eluted from the column with five bed volumes of Elution buffer (5 mM S-hexylglutathione and 10 mM DTT in Equilibrating buffer). All the steps were performed at 4°C.

2.16.2 Concentration of the purified recombinant GST

The target protein was concentrated using centriprep-10 (amicon) ultrafiltration units. The 40-ml of eluent containing the target protein was added to the centriprep-10, then centrifuged at 2,500 rpm, 4°C for 2.5 hours using SORVALL 26S, SH-3000 rotor. The final volume of the concentrated eluent was about 2.5 ml at this step. All the steps were performed at 4°C.

2.16.3 Elimination of S-hexylglutathione

S-hexylglutathione that bound to the recombinant GST was eliminated using a PD-10 column (Pharmacia). After a PD-10 column was equilibrated with 50 mM

phosphate buffer pH 6.5, the concentrated eluent containing the target protein was applied to the column. The purified GST was eluted with 50 mM phosphate buffer pH 6.5 and 10 mM DTT. The elimination step was performed twice and carried out at 4°C. The purified GST was concentrated again using the centriprep-10 as described before. The purified concentrated GST was stored in 40% glycerol at -20°C.

2.17 Protein Assay

Protein concentration was determined by the method of Bradford (114) using the Bio-Rad protein reagent (Bio-Rad). The concentrated reagent was diluted 1:5 in distilled water and filtered through Whatman No.1 filter paper to remove the insoluble dye before use. A protein standard curve was determined at six concentrations of bovine serum albumin; 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 mg/ml. The assay was started by adding 300 µl of diluted reagent to 10 µl of sample in a microtiter plate. The mixture was incubated at room temperature for 5 minutes. The absorbance at 595 nm was measured and the protein concentration was calculated from the standard curve.

2.18 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-polyacrylamide gel was performed following the formula in **Table 2.1**. The protein sample was mixed with 4X SDS reducing sample buffer [100 mM DTT in stock sample buffer (4.8 ml of distilled water, 1.2 ml of 0.5M Tris-HCl pH 6.8, 1.0 ml of glycerol, 2.0 ml of 10%(w/v) SDS and 0.5 ml of 0.1%(w/v) bromophenol blue)] at a ratio 3:1, and then loaded into a well. The electrophoresis was performed using 1X Tris-glycine buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine and 0.1% SDS) in the Bio-Rad Mini-PROTEAN® II system and run first at a constant 100 volts for 15

minutes and then with 200 volts for 45 minutes. The protein bands were stained in Coomassie staining solution (0.1% Coomassie Brilliant Blue R250, 50% methanol and 10% glacial acetic acid) at room temperature for 1 hour and subsequently destained in Destaining solution (10% methanol and 10% glacial acetic acid) until the background was clear. The gel was photographed under visible white light and incubated in distilled water overnight. Then the gel was wrapped in a cellophane membrane and air-dried for a permanent record.

Table 2.1 Formula of SDS-PAGE stacking gel and separating gel (0.75 mm x 2 gels)

Solution	Stacking gel (6%) 2 ml	Separating gel (15%) 10 ml
Acrylamide solution (ml) (30% Acrylamide + 0.3% Bis-acrylamide)	0.4	5.0
1.5 M Tris-HCl pH 8.8 (ml)	-	2.5
0.5 M Tris-HCl pH 6.8 (ml)	0.5	-
Distilled water (ml)	1.07	2.29
10% SDS (µl)	20	10
10% (w/v) Ammonium persulphate (µl)	13.33	100
TEMED (µl)	6.67	10

2.19 Determination of GST activity

The method for determination of GST activity with 1-chloro-2,4-dinitrobenzene (CDNB) is modified from the method reported by Prapanthadara *et a* (110). The standard assay for GST activity is the measurement of the activity (Δ OD/min) with 1 mM CDNB and 10mM GSH at 340 nm in 0.1 M phosphate buffer pH 6.5. The substrate solution was prepared fresh before use by adding 50 µl of 42 mM CDNB (Sigma) in ethanol into 900 µl of 21 mM GSH (Sigma) in 0.1 M phosphate buffer pH 6.5. The reaction was started by adding 100 µl of 0.1 M phosphate buffer pH 6.5 with 10 µl of enzyme sample to a well in a microtitre plate and then quickly adding 100 µl of substrate solution. The rate of conjugation between GSH and CDNB was monitored by continuously measuring the change in absorbance at 340 nm for 1 minute using a SpectraMax 250 at 26-27°C. The extinction coefficient of 9.6 $mM^{-1}cm^{-1}$ (18) was used to convert the absorbance to moles.

2.20 Determination of substrate specificities

Specific activities were determined spectrophotometrically as described in section 2.19 Determination of GST activity. All measurements were performed at 26-27°C in 0.1 M potassium phosphate buffer. The GST activities were measured with glutathione and other 4 substrates; 1-chloro-2,4-dinitrobenzene or CDNB (Sigma), 1,2-dichloro-4-nitrobenzene or DCNB (Aldrich), ethacrynic acid (Sigma) and p-nitrophenethyl bromide (Aldrich). The condition for the GST assay of each substrate was performed according to **Table 2.2**.

Table 2.2 Conditions for spectrophotometric GST assay with various substrates.

Substrate	[Substrate] (mM)	[GSH] (mM)	pH	λmax (nm)	ε ($mM^{-1}cm^{-1}$)
CDNB	1.0	10.0	6.5	340	9.6
DCNB	1.0	10.0	7.5	345	8.5
Ethacrynic acid	0.2	5.0	6.5	270	5.0
p-Nitrophenethyl bromide	0.1	5.0	6.5	310	1.2

Stock solutions of GSH were prepared in 0.1 M potassium phosphate buffer at suitable pH for each substrate. All substrates were limited in water solubility and prepared in ethanol. Stock solutions of the substrates were prepared in ethanol and diluted into the assay buffer. The concentration of ethanol was kept constant at 5% (v/v) for the assays with all substrates. These ethanol concentrations did not affect the GST activity. Specific activities were calculated according to the molar extinction coefficient (ε) for each substrate. The specific activities are the means ± standard error for at least five separate assays.

2.21 Determination of kinetic parameters

CDNB was chosen as electrophilic substrate for determination of V_{max}, K_m, k_{cat}, and k_{cat} K_m values. Ten different CDNB concentrations; 0.05, 0.10, 0.25, 0.50, 0.75, 1.00, 1.50, 2.00, 2.50 and 3.00 mM were used and GSH was held constant at a

saturating concentration of 15 mM for adGST1-2, 10 mM for both adGST1-3 and adGST1-4. Kinetic parameters were also established by varying the GSH concentrations; 0.25, 0.50, 1.00, 2.50, 5.00, 7.50, 10.00, 12.00, 15.00, 20.00 mM at a saturating concentration of CDNB, 1 mM for adGST1-2 and adGST1-3, and 3 mM for adGST1-4. The initial rate of the enzymatic reaction was measured spectrophotometrically as described for the GST activity in section 2.19. Triplicates of rate determinations were made at each concentration. The steady state kinetics was studied following Michaelis-Menten kinetics. The maximal velocity (V_{max}) and the Michaelis constant (K_m) were determined by non-linear regression analysis. The catalytic constant (k_{cat}) and the catalytic efficiency (k_{cat}/K_m) were calculated on an active-site basis using the subunit molecular mass of each enzyme. The kinetic constants are the means ± standard error for at least three separate assays.

2.22 Insecticide inhibition studies

Permethrin (Chem Service), a pyrethroid insecticide, was used as inhibitor in this study. The GST assay were performed by varying CDNB concentrations; 0.025, 0.05, 0.10, 0.25, 0.50, 0.75, 1.00, 1.50, 2.00 and 2.50 mM, with GSH held constant at a saturating concentration of 15 mM for adGST1-2 and 10 mM for both adGST1-3 and adGST1-4. The inhibition assay was performed in the absence and presence of various concentrations of permethrin; 0.01, 0.05 and 0.1 mM. The initial rate of reaction was used to construct a double reciprocal plot, 1/V versus 1/S or a Hanes-Wolff plot, S/V versus S. The inhibitor constant (K_i) was determined from the values of the appropriate intercepts on both axes (115).

CHAPTER III

RESULTS: MOLECULAR BIOLOGY

3.1 Amplification of GST cDNAs

The total RNAs that used as templates in amplification are extracted from the 4th instar larvae of *Anopheles dirus* B susceptible strain, without any selection pressure. RT-PCR products were named *adgst1-2, adgst1-3* and *adgst1-4* according to the 3'-end specific primers designed from the *adgst1AS1* gene. They were amplified using the Oligo1 primers and the specific reverse primers at the optimal annealing temperatures, 60°C, 64°C and 62°C respectively. Electrophoresis of the amplified products was performed using a 1% agarose gel at 120 volt for 1 hour. The size of the products was about 650 base pairs, as shown in **Figure 3.1**.

3.2 Construction of the recombinant plasmid into pUC19 vector

The RT-PCR products were digested with *Bam*HI to generate a 5' overhang. The pUC19 plasmids were digested with *Bam*HI and *Hinc*II to generate an overhang end and one blunt end as described in the methods. The digested RT-PCR products and the digested pUC19 plasmids were purified from the gel and then ligated to give the recombinant plasmids.

The recombinant plasmids were transformed into the *E. coli* DH5α host. After incubation on LB agar plates containing 100 μg/ml ampicillin, 400 μg of X-gal in dimethylformamide and 800 μg of IPTG in water at 37°C overnight, 50 white colonies

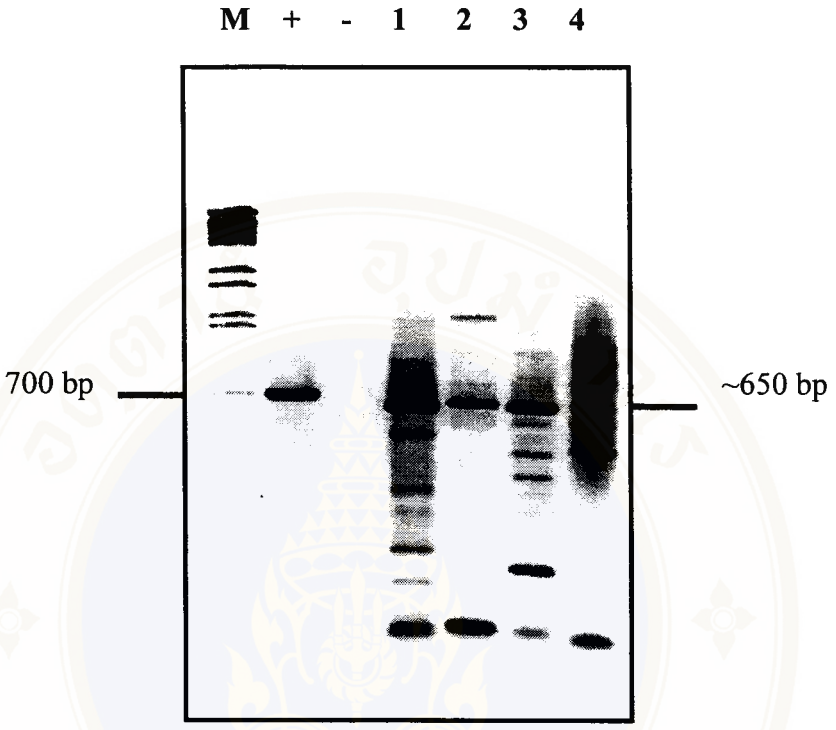


Figure 3.1 Ethidium bromide staining of RT-PCR products on a 1% agarose gel.

Lane M contains a λDNA/*Bst*EII marker.

Lane + is a positive control which used the *adgst1-1* clone10 recombinant plasmid as template.

Lane - is a negative control that contained no template.

Lane 1 contains RT-PCR products of *adgst1-1,* clone10 amplified at annealing temperature 60°C.

Lane 2 contains RT-PCR products of *adgst1-2* amplified at annealing temperature 60°C.

Lane 3 contains RT-PCR products of *adgst1-3* amplified at annealing temperature 64°C.

Lane 4 contains RT-PCR products of *adgst1-4* amplified at annealing temperature 62°C.

of each plate were randomly selected and streaked on a new plate. After agarose gel electrophoresis, 25 recombinant clones of *adgst1-2*, 10 recombinant clones of *adgst1-3* and 8 recombinant clones of *adgst1-4* were obtained for the expected size of approximately 3,320 base pairs by Simplified Rapid Size Screening Protocol as described in the methods. Three clones of each product were prepared for sequencing.

3.3 DNA sequencing in the pUC19 vector

Three recombinant pUC19 plasmids of each product were sequenced by an automated DNA sequencer, the ABI PRISM 377 (Perkin Elmer) in both forward and reverse directions. One recombinant plasmid of each product that showed the identically derived amino acids to the *adgst1AS1* gene was then sequenced again in both directions.

3.4 Subcloning into the pET3a vector

The PCR method was used for the subcloning step with the Jane1 primer and reverse primers specifically to *adgst1-2*, *adgst1-3*, *adgst1-4* to generate the *Nde*I site for ligation into the pET3a vector. The pET3a vector for *adgst1-2* and *adgst1-4* was digested with *Bam*HI, filled in with T_4 DNA Polymerase to generate blunt ends and then digested with *Nde*I to generate an overhang end. The pET3a vector for *adgst1-3* was only digested with *Nde*I. The PCR products digested with *Nde*I were subcloned without T7 tag gene for adgst1-2 and adgst1-4, and in front of T7 tag gene for adgst1-3 in the digested pET3a vectors. Then they were transformed into the *E. coli* DH5α host.

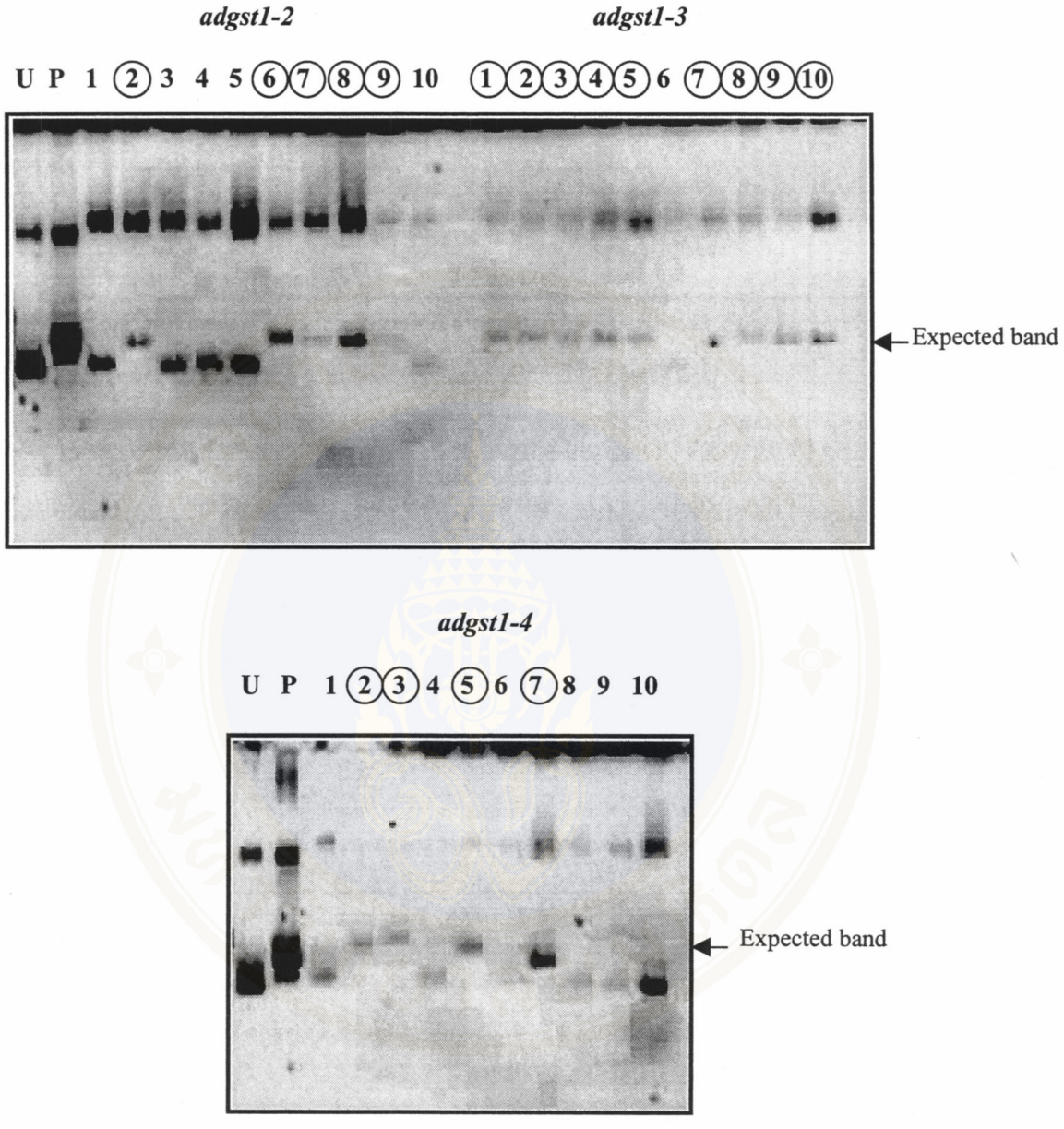


Figure 3.2 Recombinant plasmid size screening by Simplified Rapid Size Screening Protocol. A 1% agarose gel was run at 80 volt for 2 hours. Plasmid containing a DNA insert will show an increased band size compared to that of no insert DNA.

Lane U is an uncut pET3a vector.

Lane P is an uncut *adgst1-1* clone10 in pET3a vector.

Lane 1-10 is the uncut *adgst1-2*, *adgst1-3* and *adgst1-4* recombinant plasmids in the pET3a vector from clone number 1-10.

The circles are clones, which show the expected size.

After growing on LB agar plates containing 100 μg/ml ampicillin, 10 recombinant clones from each plate were randomly selected and streaked on a new plate. After agarose gel electrophoresis, 5 recombinant clones of *adgst1-2*, 9 recombinant clones of *adgst1-3* and 4 recombinant clones of *adgst1-4* showed the expected size of 5,250-5,270 base pairs by Simplified Rapid Size Screening Protocol as illustrated in **Figure 3.2**.

3.5 Checking of the insert size and orientation in the pET3a vectors

Four clones of *adgst1-2* (clone 6, 7, 8 and 9), 5 clones of *adgst1-3* (clone 1, 2, 3, 4 and 5) and 4 clones of *adgst1-4* (clone 2, 3, 5 and 7) were selected for restriction analysis with *Eco*RI to observe the correct size and orientation of insertion. The result of restriction digestion is shown in **Figure 3.3**. The recombinant plasmids that showed the expected restriction pattern; two bands of approximately 1,040 and 4,200 base pairs, were selected for sequencing.

3.6 DNA sequencing in the pET3a vector

The recombinant clones were extracted and sequenced to confirm the nucleotide sequence. Sequencing was performed using T_7-promotor and T_7-terminator as the forward and reverse primer respectively. The positive recombinant plasmid was repeatedly sequenced in both directions and subcloned into *E.coli* BL21(DE3)pLysS host for expression. The nucleotide sequence and the translated amino acid sequence of each recombinant are shown in **Figure 3.4**, **3.5** and **3.6**. The full-length nucleotide sequences of *adgst1-2*, *adgst1-3* and *adgst1-4* are 654, 630 and 660 base pairs respectively whereas that of *adgst1-1* is 630 base pairs. The translated amino acid

sequences of these genes which calculated by Vector NTI5 program are 217, 209 and 219 residues respectively whereas that of adGST1-1 is 209 residues.

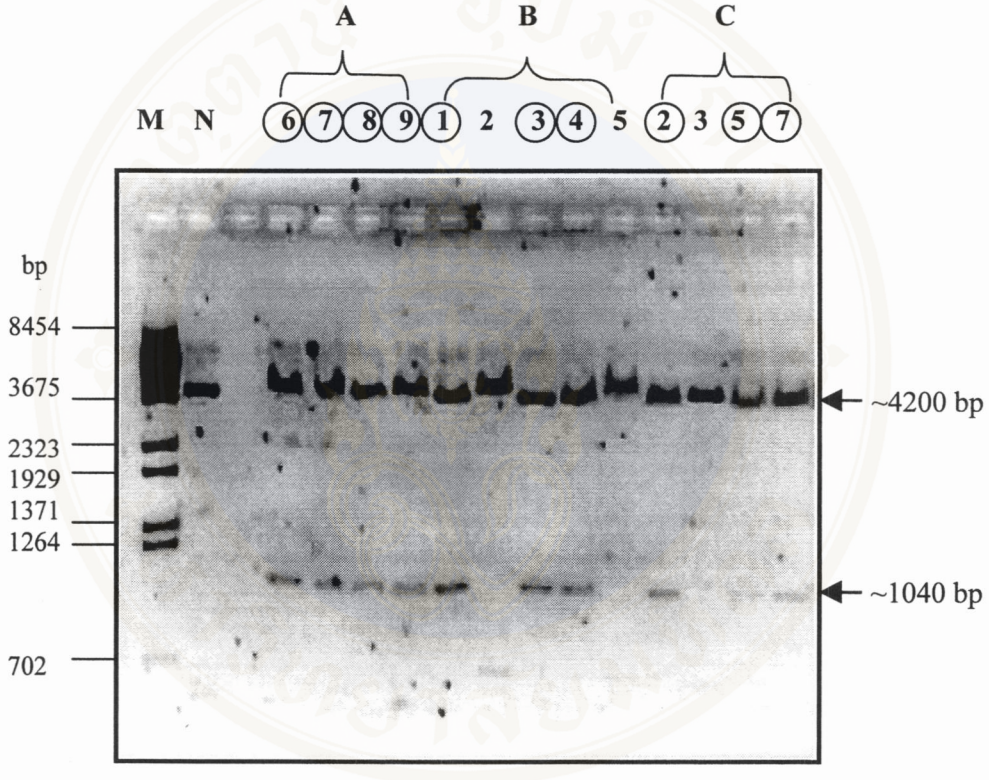


Figure 3.3 Restriction endonuclease analysis of the recombinant plasmids with *Eco*RI. The figure shows a 1% agarose gel run at 100 volt for 1.5 hours with the *Eco*RI digestion patterns of the recombinant plasmids.

Lane M contained λDNA/*Bst*EII marker.

Lane N is the *adgst1-1*clone10 recombinant plasmids in pET3a vector.

Lane A is the *adgst1-2* recombinant plasmids in pET3a vector, clone 6, 7, 8 and 9.

Lane B is the *adgst1-3* recombinant plasmid in pET3a vector, clone 1, 2, 3, 4 and 5.

Lane C is the *adgst1-4* recombinant plasmid in pET3a vector, clone 2, 3, 5 and 7.

The circles are clones, which show the expected restriction pattern.

3.7 Sequence alignment with the *adgst1AS1* gene

The nucleotide sequences and the translated amino acid sequences of *adgst1-2*, *adgst1-3* and *adgst1-4* were compared to the sequences of the *adgst1AS1* gene isolated from a genomic library with the ClustalX program. The changes in the nucleotide sequence still coded for the same amino acids as illustrated in **Figure 3.4**, **3.5** and **3.6**. In *adgst1-2*, the nucleic acid at position 609 was changed from C to T and position 624 was changed from G to A but the translated amino acids were not changed. In *adgst1-3*, the nucleic acid at position 621 was changed from C to T and at position 629 was changed from G to A, these two changes were in the specific primer designed from the 3'-end of *aggst1-5* and did not change the amino acid. Furthermore, in *adgst1-4*, the nucleic acid at position 528 was changed from G to A but the translated amino acid was not changed. Their nucleotides are different from the genomic sequence at some positions changed from either C to T or G to A but the translated amino acid sequences are not changed.

All nucleotides changed appear to be "transition mutation", which is a spontaneous mutation resulting from tautomeric shifts in the base of DNA. They involved the replacement of a purine in one strand of DNA with the other purine and the replacement of a pyrimidine in the complementary strand with the other pyrimidine (116).

***adgst1-2* cDNA** (accession no. AF273038)

```
1     ATG GAT TTC TAC TAC CTT CCC GGA TCT GCG CCG TGC CGT GCC GTT   45
1      M   D   F   Y   Y   L   P   G   S   A   P   C   R   A   V   15

46    CAG ATG ACG GCG GCC GCC GTT GGC GTC GAG CTG AAC CTG AAG CTC   90
16     Q   M   T   A   A   A   V   G   V   E   L   N   L   K   L   30
                                                  EcoRI
91    ACC AAC CTG ATG GCG GGC GAG CAC ATG AAG CCG GAA TTC CTG AAG   135
31     T   N   L   M   A   G   E   H   M   K   P   E   F   L   K   45

136   CTG AAT CCA CAG CAC TGC GTG CCG ACG CTT GTG GAC GAT GGG TTC   180
46     L   N   P   Q   H   C   V   P   T   L   V   D   D   G   F   60

181   GCC CTG TGC GAG TCA CGG GCC ATC ATG TGC TAT CTG GTG GAG AAG   225
61     A   L   C   E   S   R   A   I   M   C   Y   L   V   E   K   75

226   TAC GGC AAA CCG ATC GAG GCC GAT CGG CTC CTC CCG AGC GAT CCT   270
76     Y   G   K   P   I   E   A   D   R   L   L   P   S   D   P   90

271   CAG CGG CGG GCG ATC GTC AAC CAG CGG TTG TAC TTC GAC ATG GGC   315
91     Q   R   R   A   I   V   N   Q   R   L   Y   F   D   M   G   105

316   ACG CTG TAC CAG CGC TTC GGT GAT TAC TAC TAT CCG CAA ATC TTC   360
106    T   L   Y   Q   R   F   G   D   Y   Y   Y   P   Q   I   F   120

361   GAG GGA GCT GCG GCC AGC GAG GCT AAC TAT GCG AAA ATT GGG GAG   405
121    E   G   A   A   A   S   E   A   N   Y   A   K   I   G   E   135

406   GCT CTG ACG TTC CTC GAC ACG TTT CTG GAG GGT GAC GCG AAG TTT   450
136    A   L   T   F   L   D   T   F   L   E   G   D   A   K   F   150

451   GTG GCG GGT GGT GAT TCC TTT TCC CTG GCG GAC ATT AGT GTC TAC   495
151    V   A   G   G   D   S   F   S   L   A   D   I   S   V   Y   165

496   GCG ACG CTC ACC ACG TTC GAG GTG GCT GGG CAT GAT TTC AGT GCG   540
166    A   T   L   T   T   F   E   V   A   G   H   D   F   S   A   180

541   TAT GGC AAT GTT CTG CGG TGG TAC AAG AGT ATG GCC GGC ACG ATT   585
181    Y   G   N   V   L   R   W   Y   K   S   M   A   G   T   I   195

586   CCC GGT GCG GAC ATG AAC CGC AGT TGG GCC GAG GCG GCA AGA CCG   630
196    P   G   A   D   M   N   R   S   W   A   E   A   A   R   P   210

631   TTT TTC GAT CGG ATT AAG CAT TGA   654
211    F   F   D   R   I   K   H   *    217
```

Figure 3.4 Nucleotide and translated amino acid sequence of *adgst1-2*. The length of nucleotide sequence is 654 base pairs and translated amino acid sequence is 217 residues. Underlined residues indicate the restriction endonuclease site. Bold underlined residues indicate the nucleic acid that is different from the genomic nucleic acid sequence of the *adgst1AS1* gene.

***adgst1-3* cDNA** (accession no. AF273039)

```
1     ATG GAT TTC TAC TAC CTT CCC GGA TCT GCG CCG TGC CGT GCC GTT   45
1      M   D   F   Y   Y   L   P   G   S   A   P   C   R   A   V   15

46    CAG ATG ACG GCG GCC GCC GTT GGC GTC GAG CTG AAC CTG AAG CTC   90
16     Q   M   T   A   A   A   V   G   V   E   L   N   L   K   L   30
                                                  EcoRI
91    ACC AAC CTG ATG GCG GGC GAG CAC ATG AAG CCG GAA TTC CTG AAG   135
31     T   N   L   M   A   G   E   H   M   K   P   E   F   L   K   45

136   ATT AAC CCC CAA CAC TGT ATT CCG ACG CTG GTC GAC AAT GGC TTT   180
46     I   N   P   Q   H   C   I   P   T   L   V   D   N   G   F   60

181   GCG CTG TGG GAG TCG CGC GCC ATC TGC ACG TAC TTG GCG GAG AAG   225
61     A   L   W   E   S   R   A   I   C   T   Y   L   A   E   K   75

226   TAC GGC AAG GAC GAC AAG CTG TAC CCG AAG GAC CCG CAG AAG CGC   270
76     Y   G   K   D   D   K   L   Y   P   K   D   P   Q   K   R   90

271   GCC GTC GTC AAC CAG CGA CTG TAC TTC GAC ATG GGC ACG CTG TAC   315
91     A   V   V   N   Q   R   L   Y   F   D   M   G   T   L   Y   105

316   CAG CGC TTT GCC GAC TAC TAC TAC CCG CAG ATC TTC GCC AAG CAG   360
106    Q   R   F   A   D   Y   Y   Y   P   Q   I   F   A   K   Q   120

361   CCG GCC AAC GCG GAG AAC GAG AAG AAG ATG AAG GAT GCG GTC GAC   405
121    P   A   N   A   E   N   E   K   K   M   K   D   A   V   D   135

406   TTC CTG AAC ACC TTC CTG GAC GGG CAC AAG TAC GTG GCG GGC GAC   450
136    F   L   N   T   F   L   D   G   H   K   Y   V   A   G   D   150

451   AGT CTG ACG ATC GCG GAT CTG ACC GTG CTG GCC ACG GTT TCG ACG   495
151    S   L   T   I   A   D   L   T   V   L   A   T   V   S   T   165

496   TAC GAC GTG GCG GGC TTC GAG CTG GCC AAG TAC CCG CAC GTG GCG   540
166    Y   D   V   A   G   F   E   L   A   K   Y   P   H   V   A   180

541   GCG TGG TAC GAG CGC ACC CGC AAG GAA GCG CCC GGT GCC GCC ATC   585
181    A   W   Y   E   R   T   R   K   E   A   P   G   A   A   I   195

586   AAC GAG GCC GGC ATC GAG GAG TTC AGG AAG TAC TTT GAG AAG TAA   630
196    N   E   A   G   I   E   E   F   R   K   Y   F   E   K   *   209
```

Figure 3.5 Nucleotide and translated amino acid sequence of *adgst1-3*. The length of nucleotide sequence is 630 base pairs and translated amino acid sequence is 209 residues. Underlined residues indicate the restriction endonuclease site. Bold underlined residues indicate the nucleic acid that is different from the genomic nucleic acid sequence of the *adgst1AS1* gene.

***adgst1-4* cDNA** (accession no. AF273040)

```
1     ATG GAT TTC TAC TAC CTT CCC GGA TCT GCG CCG TGC CGT GCC GTT   45
1      M   D   F   Y   Y   L   P   G   S   A   P   C   R   A   V   15

46    CAG ATG ACG GCG GCC GCC GTT GGC GTC GAG CTG AAC CTG AAG CTC   90
16     Q   M   T   A   A   A   V   G   V   E   L   N   L   K   L   30
                                                  EcoRI
91    ACC AAC CTG ATG GCG GGC GAG CAC ATG AAG CCG GAA TTC CTG AAG   135
31     T   N   L   M   A   G   E   H   M   K   P   E   F   L   K   45

136   CTA AAC CCG CAA CAC TGC ATC CCG ACG CTG GTC GAC GAG GAC GGG   180
46     L   N   P   Q   H   C   I   P   T   L   V   D   E   D   G   60

181   TTT GTG CTG TGG GAG TCG CGC GCC ATC CAG ATC TAC CTG GTC GAG   225
61     F   V   L   W   E   S   R   A   I   Q   I   Y   L   V   E   75

226   AAG TAC GGC GCC CAC GAT GCG GAC CTG GCC GAG CGG CTT TAC CCA   270
76     K   Y   G   A   H   D   A   D   L   A   E   R   L   Y   P   90

271   AGC GAT CCG CGT CGC CGT GCC GTC GTC CAC CAG CGG CTG TTC TTC   315
91     S   D   P   R   R   R   A   V   V   H   Q   R   L   F   F   105

316   GAT GTG GCC GTC CTG TAC CAG CGC TTC GCC GAG TAC TAC TAT CCG   360
106    D   V   A   V   L   Y   Q   R   F   A   E   Y   Y   Y   P   120

361   CAG ATC TTC GGC CAG AAG GTG CCG GTG GGT GAT CCG GGC CGG TTG   405
121    Q   I   F   G   Q   K   V   P   V   G   D   P   G   R   L   135

406   CGC TCG ATG GAG CAG GCA CTC GAG TTC CTC AAC ACG TTC CTC GAG   450
136    R   S   M   E   Q   A   L   E   F   L   N   T   F   L   E   150

451   GGC GAG CAG TAC GTT GCC GGT GGT GAT GAT CCG ACG ATC GCC GAT   495
151    G   E   Q   Y   V   A   G   G   D   D   P   T   I   A   D   165

496   CTG AGC ATA CTG GCG ACG ATC GCC ACG TAC GAA GTG GCC GGC TAC   540
166    L   S   I   L   A   T   I   A   T   Y   E   V   A   G   Y   180

541   GAT CTG CGC CGG TAC GAG AAC GTT CAG CGG TGG TAC GAG CGG ACG   585
181    D   L   R   R   Y   E   N   V   Q   R   W   Y   E   R   T   195

586   AGC GCG ATC GTC CCC GGG GCT GAT AAG AAC GTG GAG GGA GCT AAG   630
196    S   A   I   V   P   G   A   D   K   N   V   E   G   A   K   210

631   GTT TTC GGG CGC TAC TTC ACA CAA AAA TGA   660
211    V   F   G   R   Y   F   T   Q   K   *   219
```

Figure 3.6 Nucleotide and translated amino acid sequence of *adgst1-4*. The length of nucleotide sequence is 660 base pairs and translated amino acid sequence is 219 residues. Underlined residues indicate the restriction endonuclease site. Bold underlined residues indicate the nucleic acid that is different from the genomic nucleic acid sequence of the *adgst1AS1* gene.

3.8 Nucleotide sequences alignment

The nucleotide sequences of the *An. dirus* three novel transcripts; *adgst1-2*, *adgst1-3* and *adgst1-4*, and the previous reported sequence of *adgst1-1* (123) were compared using ClustalX program. All of them are the alternatively spliced products of *adgst1AS1* gene isolated from a genomic library. Their nucleotide sequences share >67% identity as shown in **Table 3.1**. The *adgst1-1* has the greatest identity, 83% to *adgst*1-3, although they are located in distant positions whereas they have approximately 70% to *adgst1-2* and *adgst1-4*.

Table 3.1 Percent identity of nucleotide sequences

	adgst1-1	*adgst1-2*	*adgst1-3*	*adsgt1-4*
1.	100	68	83	73
2.		100	67	69
3.			100	73
4.				100

From the nucleotide sequence alignment in **Figure 3.7**, these four transcripts are identical at 5'-end from position 1-135 except *adgst1-1* are different at the beginning of sequence because of its different primer. The nucleotide sequences from position 1-135 of *adgst1-1*, *adgst1-2*, *adgst1-3* and *adgst1-4* are identical to exon2 of *adgst1AS1* gene while the rest of sequences are identical to each of exon3; 3D, 3C, 3B and 3A respectively. This indicates that these four adGSTs share a common 5' exon, which is spliced to one of four alternative exons to produce mature transcripts occurring in the *adgst1AS1* gene.

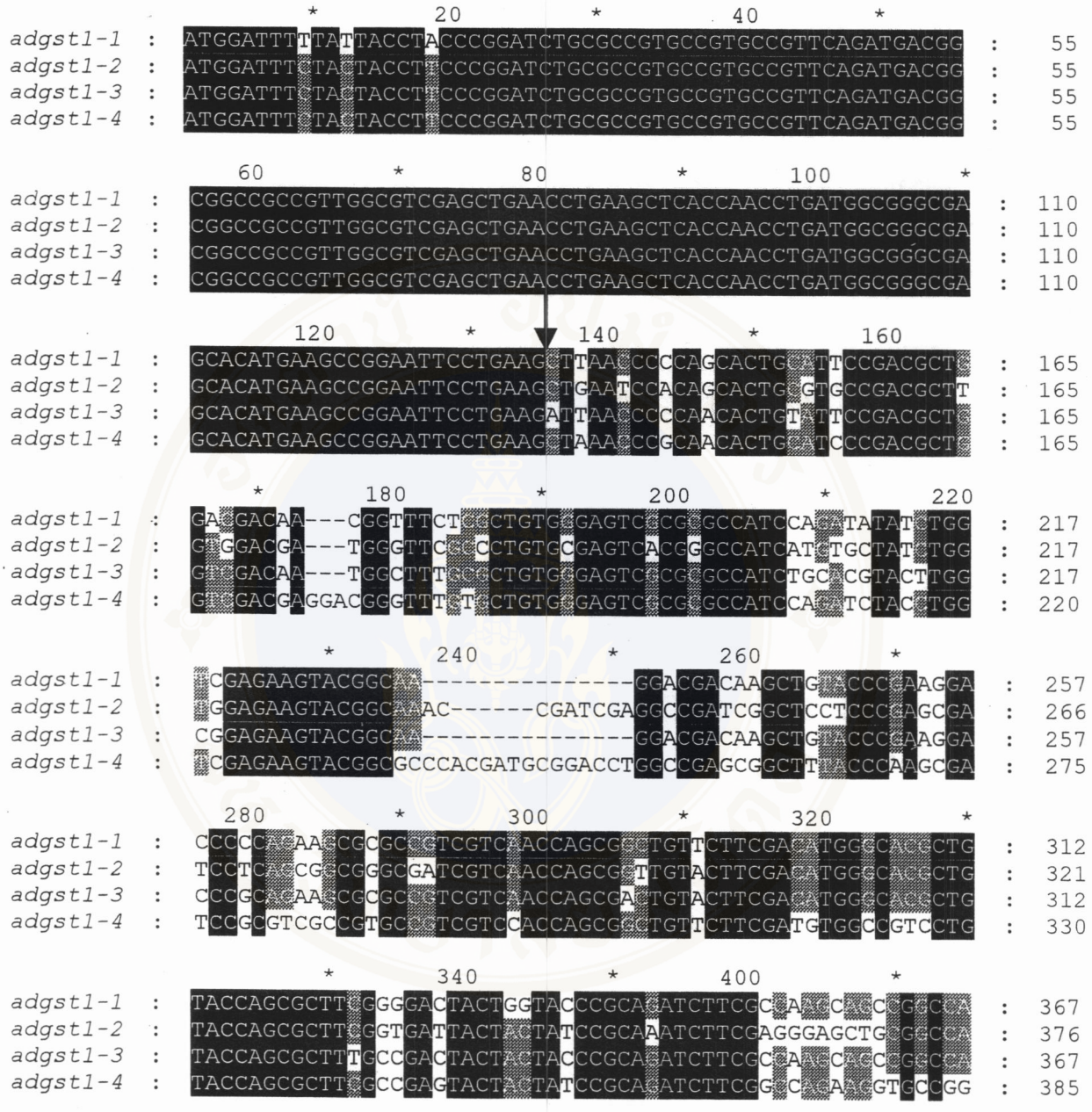


Figure 3.7 Alignment of nucleotide sequences of *adgst1-1*, *adgst1-2*, *adgst1-3*, and *adgst1-4*. The *adgst1-1*, clone10 is used in this alignment. Gaps introduced to maximize sequence similarity are shown by a horizontal dash. Black shadings represent 100% sequence similarity and gray shadings represent 80% sequence similarity. A vertical arrow marks the junction between exon 2 and exon 3 in *adgst1-1*, *adgst1-2*, *adgst1-3* and *adgst1-4*.

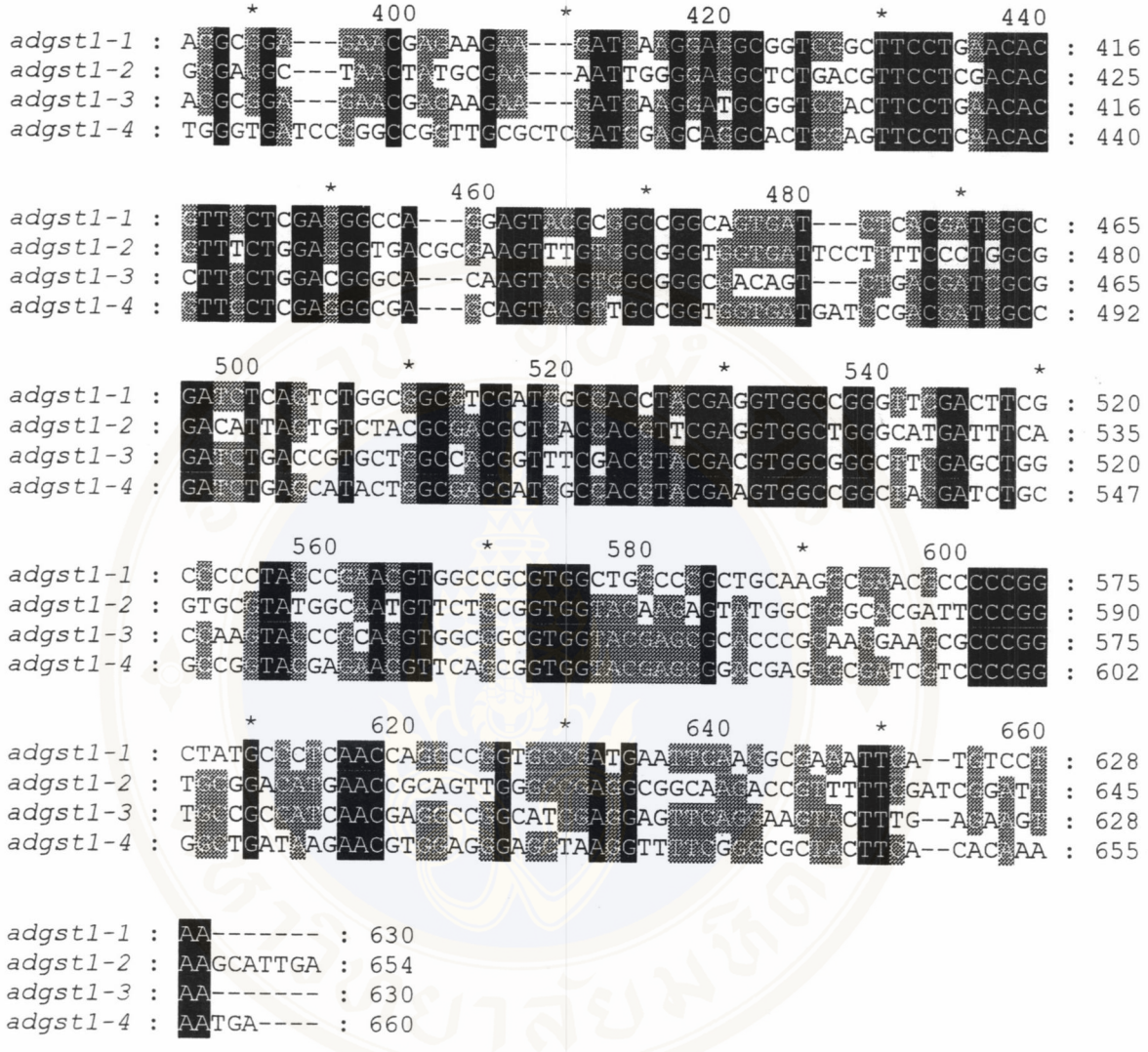


Figure 3.7 (continue)

3.9 Amino acid sequence alignment

The translated amino acid sequences of adGST1-2, adGST1-3, adGST1-4 and adGST1-1 from (117) share >61% identity. Their amino acid residues at position 1-45 in the N-termini, share 100% identity while their C-termini share >51% identity. C-

terminal sequence of adGST1-1 has greatest identity to adGST1-3, 73%, while it has only 52% and 56% to adGST1-2 and adGST1-4 as shown in **Table 3.2**.

Table 3.2 Percent identity of C-terminal sequences.

	adGST1-1	adGST1-2	adGST1-3	adGST1-4
1.	100	52	73	56
2.		100	54	51
3.			100	55
4.				100

The alignment of these translated amino acid sequences is shown in **Figure 3.8**. Their N-termini, which known as the glutathione-binding site, are very conserved, while C-termini, which known as the hydrophobic binding site are very variable. To determine which parts of C-termini may involve in the substrate specificities, their amino acid sequences were compared with the previous reported secondary structure of adGST1-1 (117). It was found that amino acid residues in all loops of C-termini are variable, especially a loop between α3 and α4 of adGST1-4 is longer. Furthermore, two α-helices, α5 and α7 are also highly variable. From significantly variation in C-termini, it indicates that the overall tertiary structure of these four alternatively spliced products of the *adgst1AS1* gene may be changed in order to increase the variable of substrate binding.

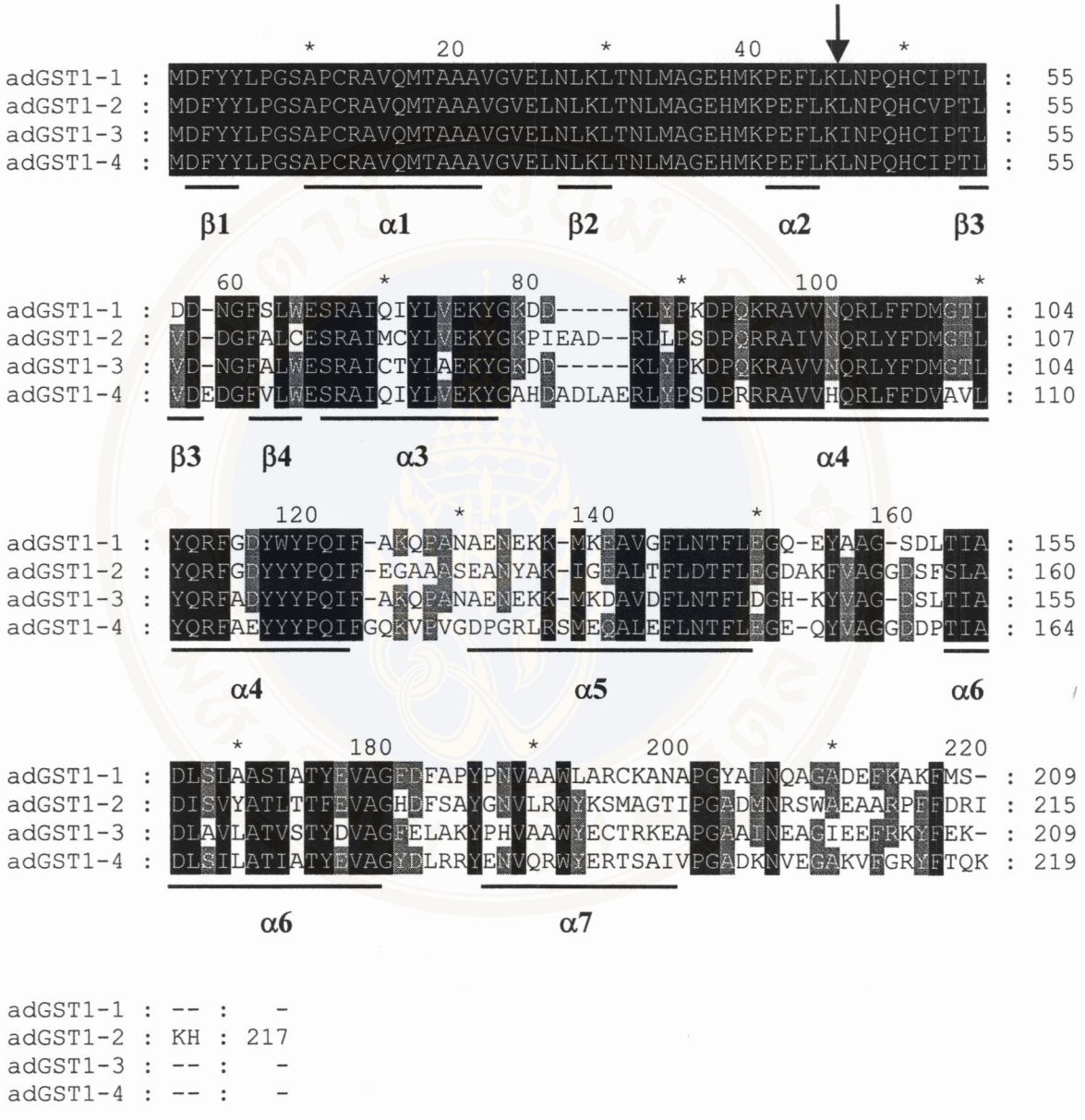


Figure 3.8 Alignment of derived amino acid sequences of adGST1-1, adGST1-2, adGST1-3, and adGST1-4. The adGST1-1, clone10 is used in this alignment. Gaps introduced to maximize sequence similarity are shown by a horizontal dash. Black shadings represent 100% sequence similarity and gray shadings represent 80% sequence similarity. Residues assigned to secondary structure elements of adGST1-1 are underlined. β refers to a β-strand and α refers to an α-helix.

3.10 Comparison of amino acid sequences within the insect class I GSTs

The amino acid sequences of adGST1-2, adGST1-3 and adGST1-4 were compared with the other insect class I GSTs using the ClustalX program and the percent identity was calculated by the Genedoc program as illustrated in **Table 3.3**.

All other insect delta class or class I GSTs used for comparison were reported in GenBank and grouped as Dipteran or Lepidopteran. The Dipteran consists of *Anopheles dirus* (Thai malaria mosquito), *Anopheles gambiae* (African malaria mosquito), *Culicoides variipennis*, *Lucilia cuprina* (Australian sheep blowfly), *Musca domestica* (housefly) and *Drosophila spp*. The Lepidopteran consists of *Bombyx mori* (silk moth), *Plutella xylostella* (diamondback moth) and *Manduca sexta* (tobacco hornworm). When amino acid sequences were compared within Dipteran insects, they showed >50% identity except agGST1-7 which had only 39-40% identity. *Anopheles dirus* GSTs showed the highest amino acid sequence identity to *Anopheles gambiae*. AdGST1-2 and agGST1-4 shared 85% identity, adGST1-3 and agGST1-5 92% identity, and adGST1-4 and agGST1-3 85% identity. When they were compared to Lepidopteran insects, they had only 32-45% amino acid sequence identity.

3.11 Comparison of amino acid sequences with other classes of GSTs.

The amino acid sequences of adGST1-2, adGST1-3 and adGST1-4 were compared with the other classes GSTs using the ClustalX program and the percent identity was calculated by the Genedoc program as illustrated in **Table 3.4**.

Insect sigma class or class II GSTs used in comparison were *Anopheles gambiae* GST2, *Drosophila melanogaster* GST2, *Musca domestica* GST2,

Table 3.3 The percent amino acid sequence identity comparison within the insect class I GSTs.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. **adGST1-2**	100	64	61	61	47	57	**85**	62	62	41	58	59	57	57	54	54	54	53	53	42	36	32
2. **adGST1-3**		100	63	77	50	60	63	**92**	80	44	67	70	70	70	66	66	67	66	66	45	38	35
3. **adGST1-4**			100	63	52	**85**	59	64	65	39	57	57	54	56	53	53	53	52	53	42	34	33
4. **adGST1-1**				100	48	58	63	76	**91**	45	72	67	68	65	64	64	63	62	63	45	38	39
5. **agGST1-1**					100	48	47	50	50	36	45	45	45	47	43	43	43	43	44	36	32	28
6. **agGST1-3**						100	59	62	61	37	53	53	52	53	50	50	51	50	50	41	31	32
7. **agGST1-4**							100	64	66	41	59	59	57	58	56	56	55	55	55	43	37	34
8. **agGST1-5**								100	82	41	64	69	69	68	65	65	65	64	65	45	36	35
9. **agGST1-6**									100	44	73	67	68	67	65	65	64	63	64	44	37	37
10. **agGST1-7**										100	43	41	42	42	39	39	39	39	40	45	31	34
11. **Culicoides1**											100	64	63	63	62	62	62	60	61	45	37	34
12. **Lucilia1**												100	92	82	80	81	81	79	80	43	36	37
13. **Musca1**													100	84	81	82	82	82	82	42	37	38
14. **dmGST1-1**														100	91	92	93	92	93	44	36	35
15. **dseGST1-1**															100	99	98	96	97	43	35	35
16. **dsiGST1-1**																100	98	97	98	43	35	35
17. **dyGST1-1**																	100	97	98	42	35	35
18. **deGST1-1**																		100	99	41	35	35
19. **dtGST1-1**																			100	42	34	35
20. **Bombyx1**																				100	32	34
21. **Plutella1**																					100	44
22. **Manduca1**																						100

The amino acid sequences were obtained from the GenBank database. The abbreviations are adGST1-1 for *Anopheles dirus* (1786091), agGST for *An. gambiae*; agGST1-1 (2842739), agGST1-3 (3549276), agGST1-4 (3549274), agGST1-5 (2842717), agGST1-6 (2842718) and agGST1-7 (3549271), Culicoides1 for *Culicoides variipennis* (2738075), Lucilia1 for *Lucilia cuprina* (1346214), Musca1 for *Musca domestica* (121696), dmGST1-1 for *Drosophila melanogaster* (121694), dseGST1-1 for *D. sechellia* (232193), dsiGST1-1 for *D. simulans* (232192), dyGST1-1 for *D. yakuba* (232195), deGST1-1 for *D. erecta* (232191), dtGST1-1 for *D. teissieri* (232194), Bombyx1 for *Bombyx mori* (3201479), Plutella1 for *Plutella xylostella* (3582502), Manduca1 for *Mandula sexta* (1170115).

Table 3.4 The percent amino acid sequence identity comparison with the other classes of GSTs.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. **adGST1-2**	100	64	61	61	8	9	11	6	3	16	15	20	17	20	22	18	23	6	6	4	8	21
2. **adGST1-3**		100	63	77	10	10	12	6	4	16	16	19	17	22	22	21	22	6	6	6	8	19
3. **adGST1-4**			100	63	9	10	12	4	5	17	17	18	17	23	20	22	21	4	4	4	8	19
4. **adGST1-1**				100	9	10	12	6	5	16	17	17	19	23	22	22	23	7	4	7	9	20
5. **agGST2**					100	50	54	30	38	6	6	8	6	8	8	8	7	19	16	12	6	8
6. **dmGST2**						100	81	22	30	1	4	5	6	8	7	8	9	15	11	11	7	5
7. **Musca2**							100	25	31	3	5	6	8	10	9	9	11	19	12	11	8	6
8. **Choristoneura**								100	27	4	5	5	3	4	5	4	4	18	18	10	8	5
9. **Manduca2**									100	3	5	4	4	7	6	7	5	23	18	14	6	2
10. **clove-theta**										100	18	13	13	16	13	14	15	3	4	2	4	25
11. **E.coli-theta**											100	11	12	14	16	13	17	5	4	2	7	16
12. **plaice-theta**												100	13	16	15	16	15	5	5	5	5	14
13. **chick-theta**													100	42	48	42	49	5	5	4	5	16
14. **rat-theta**														100	49	90	52	8	5	5	7	18
15. **mouse-theta1**															100	48	81	8	6	6	8	16
16. **mouse-theta2**																100	51	7	5	5	8	17
17. **human-theta**																	100	7	5	5	10	17
18. **human-pi**																		100	24	16	10	7
19. **human-alpha**																			100	16	7	5
20. **human-mu**																				100	6	2
21. **human-kappa**																					100	7
22. **human-zeta**																						100

The amino acid sequences were obtaind from the GenBank database. The abbreviations are adGST1-1 for *Anopheles dirus* GST1-1 (1786091), agGST2 for *An. gambiae* GST2 class sigma (1170108), dmGST2 for *Drosophila melanogaster* GST2 class sigma (729641), Musca2 for Musca domestica GST2 class sigma (1170110), Choristoneura for Choristoneura fumiferana GST class sigma (6671050), Manduca2 for *Mandula sexta* GST2 class sigma (1170115), clove-theta for Dianthus caryophyllus GST2 class theta (417094), E.coli-theta for Escherichia coli GST1 class theta (1799682), plaice-theta for Pleuronectus platessa GST-A class theta (232216), chick-theta for Chick GST-CL1 class theta (1170112), rat-theta for Rattus norvegicus GST1 class theta (6980992), Mouse-theta1 for Mus musculus GST1 class theta (6680123), Mouse-theta2 for Mus musculus GST2 class theta (6754088), human for Homo sapiens; theta for GSTT1-1 (4564185), pi for GSTP1-1 (121746), alpha for GSTA1-1 (121730), mu for GSTM1-1 (121735), kappa for mitochondrial GST13-13 (3041680) and zeta for GSTZ1 (translated from U86529).

Choristoneura fumiferana GST (spruce budworm) and *Manduca sexta* GST2. Non-mammalian theta class GSTs were *Dianthus caryophyllus* GST2 (clove pink plant), *Escherichia coli* GST1, *Pleuronectes platessa* GST-A (plaice fish) and *Gallus gallus* GST-CL1 (chick). Mammalian theta class GSTs used in comparison were *Rattus norvegicus* GST1 (rat), *Mus musculus* GST1 and GST2 (house mouse). For *Homo sapiens* GSTs, all classes choosen were class theta, pi, alpha, mu, kappa and zeta. The amino acid sequences of adGST1-2, adGST1-3 and adGST1-4 showed the highest identity to the mammalian theta class GSTs of about 20%, while they showed the lowest identity to the mammalian pi, alpha, mu, kappa, zeta and insect sigma class <12%. For non-mammalian theta class GSTs, they share about 15-20% identity.

3.12 Alignment of N-terminal domains

Based on the Crystal structure of a *Lucilia cuprina* delta class GST, a serine residue near the N-termini of the theta class GSTs, plays a vital role in the catalytic mechanism of the GSTs (40). To determine whether serine residue is conserved in glutathione binding site, the N-terminal domains of four alternatively spliced products of *adgst1AS1* gene were compared with those of insect class I and the theta class GSTs as shown in **Figure 3.9**.

From the alignment results, a serine residue in N-terminal domains are conserved at:

Ser9 in *Anopheles dirus* GST1-1, 1-2, 1-3, and 1-4, *Anopheles gambiae* GST1-3, 1-4, 1-5 and 1-6, *Plutella* GST1, *Lucilia* GST1 and *Musca* GST1, members of the Dipteran insects.

Ser10 in *A. gambiae* GST1-1 and *Drosophila melanogaster* GST1-1.

Ser11 in *Manduca* GST1, *Culicoides* GST1, chick, mouse, rat and human theta GSTs.

Ser13 in plaice fish theta GST.

Ser14 in *Bombyx* GST1.

Ser16 in *Escherichia coli* theta GST.

Ser17 in clove plant GST.

It was shown that most of insect class I or delta class GSTs contain conserved at position 9 from the N-termini.

```
                          *          20          *          40          *
Ad1-1          : --------MDFYYLPGSAPCRAVQMTAAALGV-ELNLKLTNLMA------ :  35
Ad1-2          : --------MDFYYLPGSAPCRAVQMTAAAVGV-ELNLKLTNLMA------ :  35
Ad1-3          : --------MDFYYLPGSAPCRAVQMTAAAVGV-ELNLKLTNLMA------ :  35
Ad1-4          : --------MDFYYLPGSAPCRAVQMTAAAVGV-ELNLKLTNLMA------ :  35
Ag1-1          : -------MLDFYYLPGSAPCRAVQMVAEAVHV-KLNLKYLDLMA------ :  36
Ag1-3          : --------MDFYYLPGSAPCRAVQMTAAAVGV-ELNLKLTDLMK------ :  35
Ag1-4          : --------MDFYYLPGSAPCRAVQMTAAAVGV-ELNLKLTDLMK------ :  35
Ag1-5          : --------MDFYYLPGSAPCRAVQMTAAAVGV-ELNLKLTDLMK------ :  35
Ag1-6          : --------MDFYYLPGSAPCRAVQMTAAAVGV-ELNLKLTDLMK------ :  35
Ag1-7          : -----MTPV-LYYLPPSPPCRSVLLLAKMIGV-ELELKALNVME------ :  37
Plutella1      : --------MKLYKLDMSPPARATMMVAEALGV-KVDTVDVNLMK------ :  35
Manduca1       : ------MVMTLYKLDASPPARAVMMVIEALKIPDVEYIDVNLLE------ :  38
Bombyx1        : ---MPVQPIKLYYLPPSPPCRAVMMTARVLEL-DLHLITTNIMN------ :  40
Culicoides     : ------MGLDFYYLPGSSPCRAVQMTAKAVGV-DLNLKLTNLMA------ :  37
Lucilia1       : --------MDFYYLPGSAPCRSVLMTAKALGI-ELNKKLLNLQA------ :  35
Musca1         : --------MDFYYLPGSAPCRSVLMTAKALGI-ELNKKLLNLQA------ :  35
DmeGST1-1      : -------MVDFYYLPGSSPCRSVIMTAKAVGV-ELNKKLLNLQA------ :  36
E.coli-theta   : -MSKPAITLWSDAHFPSPYVLSAWVALQEKGL-SFHIKTIDLDS------ :  42
clove-theta    : MSSSETQKMQLYSYSSSSCAWRVRIALHLKGL-DFEYKAVDLLK------ :  43
plaice-theta   : ----MAKDMTLLWGSGSPPCWRVMIVLEEKNLQAYNSKLLSFEK------ :  40
chick-theta    : ------MGLELYLDLLSQPCRSIYIFARTNNI-PFEFKHVELFKDSVLGK :  43
mouse-theta1   : ------MVLELYLDLLSQPCRAIYIFAKKNNI-PFQMHTVELRK------ :  37
rat-theta      : ------MGLELYLDLLSQPSRAVYIFAKKNGI-PFQLRTVDLLK------ :  37
human-theta    : ------MGLELYLDLLSQPCRAVYIFAKKNDI-PFELRIVDLIK------ :  37
```

Figure 3.9 Alignment of N-terminal domains between the insect class I and the theta class GSTs. All amino acid sequences are searched from GenBank database as Table 4.9 and 4.10. Residue, which is conserved among all sequences, has been boxed.

CHAPTER IV

RESULTS : ENZYMOLOGY

4.1 Expression of recombinant GSTs in *E. coli*

The recombinant plasmids of *adgst1-2*, *adgst1-3* and *adgst1-4* in pET3a vector were transformed into *E. coli* BL21(DE3)pLysS and incubated in LB broth containing 100μg/ml ampicillin and 34μg/ml chloramphenicol until the OD at 600 nm was about 0.6. The culture was then induced with a final concentration of 0.1mM IPTG for 3 hours. The recombinant proteins were expressed under the T_7 promotor system without the 13-residue leader T_7-tag attached to the N-terminus. The crude extracts were collected before induction and after induction, and run on the SDS-PAGE to show the expression level as illustrated in **Figure 4.1**. The expression level of adGST1-2, adGST1-3 and adGST1-4 after 3 hours of IPTG induction were about 5-10-fold greater than before induction. Whereas adGST1-1 clone 10 was expressed at the similar level both before and after the IPTG induction, which indicated that the T_7 promotor system is not suitable for this recombinant GST. The overall proteins were expressed at high levels, yielding approximately 50-60 mg from 200 ml of *E. coli* cultures. The adGST1-3 showed highest expression level of the expected band, about 2-fold greater than adGST1-2 and adGST1-4 and 5-fold greater than adGST1-1.

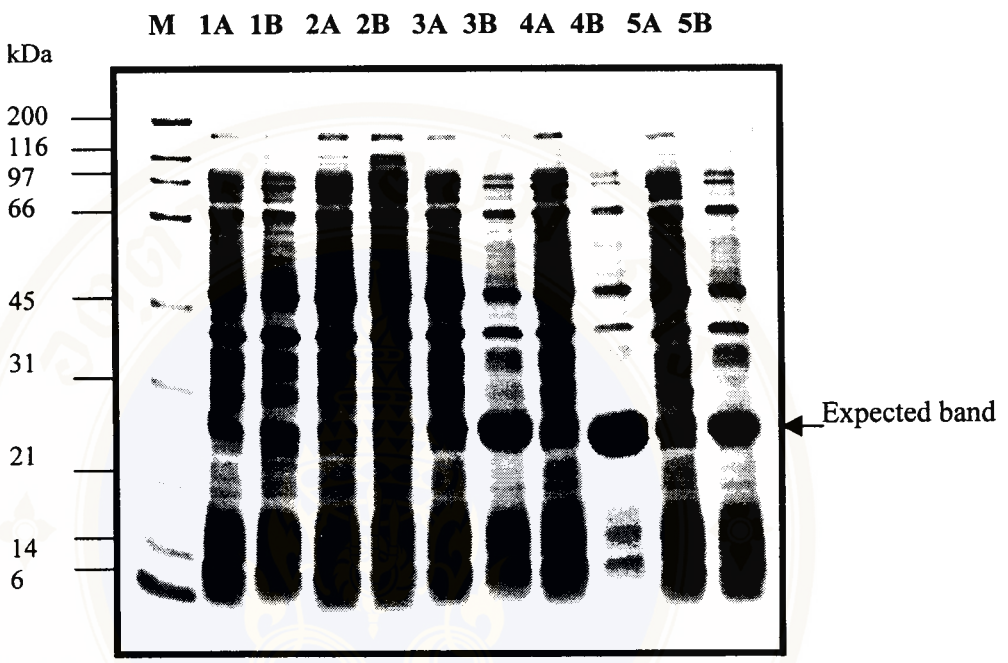


Figure 4.1 The SDS-PAGE of the crude extracts of *Anopheles dirus* GSTs before and after IPTG induction.

Lane M contained molecular weight marker as indicated at the left.

Lane A is a 0.1 OD of culture before IPTG induction.

Lane B is a 0.1 OD of culture after 3-hr IPTG induction.

Lane 1, 2, 3, 4 and 5 are pET3a vector, adGST1-1 (clone10), adGST1-2, adGST1-3 and adGST1-4 respectively.

4.2 Protein purification

The GSTs from *An. dirus* have been separated by sequential column chromatography into peaks of activity, each containing multiple enzymes. Peak 4 has been isolated based on binding to the S-hexylglutathione agarose column and separated into 3 peaks; 4a, 4b and 4c based on binding and order of elution from the hydroxylapatite column (110). These three recombinant enzymes can bind to the affinity chromatography on an *S*-hexylglutathione column but lack N-terminal homology with GST from peak 4a (APIVLY), which suggests that these three recombinant GSTs are localized in peak 4b or 4c. After the final step of purification, the enzymes approximately 15-20 mg or 40-70% of total activity were recovered from the S-hexylglutathione agarose column as illustrated in **Table 4.1**. Although *E. coli* possesses native GSTs, only trace amounts of protein from the control culture were detected+. The specific activity of unpurified control (*E. coli* host carried pET3a without insert) was 0.001 μmole/min/mg and adGST1-1 was 0.68 μmole/min/mg, whereas those of adGST1-2, adGST1-3 and adGST1-4 were 22.0, 26.0 and 18.3 μ mole/min/mg respectively. After purification step, the specific activities of these three enzymes were increased approximately 1.5-2.0-fold greater than unpurified enzymes. AdGST1-3 showed the greatest percent recovery and bound to the S-hexylglutathione affinity column more tightly than adGST1-2 and adGST1-4.

Table 4.1 Purification of the *Anopheles dirus* GST proteins.

Enzymes	Step	Total protein (mg)	Total activity (μmole/min)	Specific activity (μmole/min/mg)	Yield (%)	Fold purification
Control	Supernatant	51	0.05	0.001	-	-
	Affinity column	N.D.	N.D.	-	-	-
AdGST1-1	Supernatant	46	31	0.68	-	-
	Affinity column	N.D.	N.D.	-	-	-
AdGST1-2	Supernatant	54	1190	22.0	100	1
	Affinity column	15	650	43.3	55	2
AdGST1-3	Supernatant	61	1580	26.0	100	1
	Affinity column	21	1110	51.9	70	2
AdGST1-4	Supernatant	45	822	18.3	100	1
	Affinity column	13	350	26.9	43	1.5

Enzyme activity was followed by the standard assay using CDNB as the substrate. N.D. is not detected. Control is *E. coli* host carried the pET3a without insert.

4.3 Physical properties

Based on the amino acid compositions, adGST1-4, which its subunit size is 25.1 kDa, are the largest protein comparing to the alternatively spliced products within the *adgst1AS1* gene as illustrated in **Table 4.2**. Their isoelectric points are slightly different.

Table 4.2 Physical properties of the purified *Anopheles dirus* GSTs.

Physical properties	Recombinant GSTs			
	AdGST1-1[(117)]	AdGST1-2	AdGST1-3	AdGST1-4
Nucleotide sequence length (bp)	630	654	630	660
Amino acid sequence length (residues)	209	217	209	219
Subunit size (kDa)	23.4	24.3	23.9	25.1
Isoelectric point	5.47	5.23	6.13	5.30

The subunit size and isoelectric point were calculated by by Vector NTI5 software. AdGST1-1 is clone 10 recombinant.

The purified enzymes have shown to be homogeneous preparations, which appear single protein bands as illustrated in **Figure 4.2**. The expected bands are about 23-25 kDa relative to the board range molecular weight marker, which correspond to the calculated molecular weight of the GST subunits in **Table 4.2**.

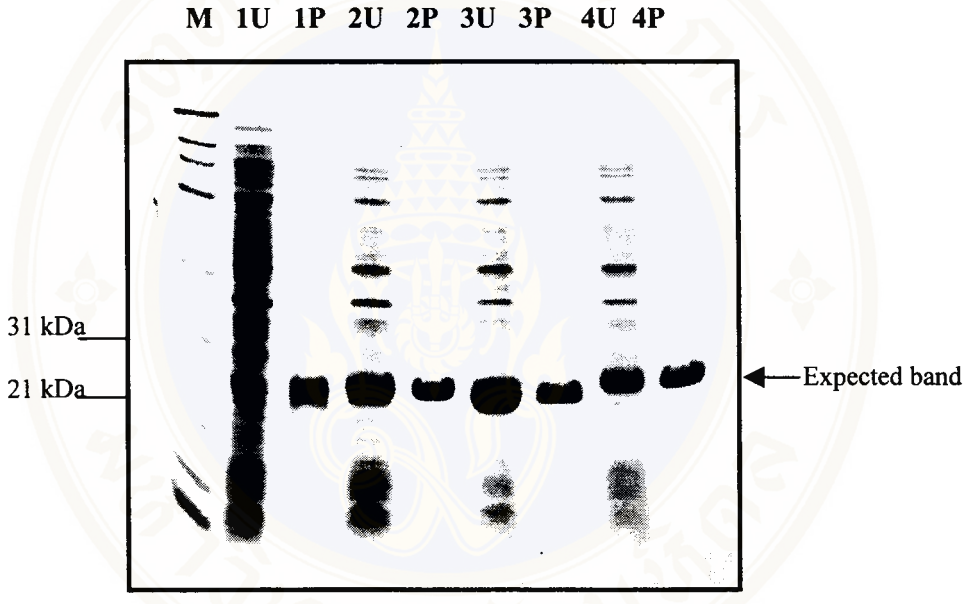


Figure 4.2 The SDS-PAGE of unpurified and purified recombinant *An. dirus* GSTs.

Lane M contains molecular weight marker.

Lane U contains a 0.1 OD of the induced *E. coli* culture

Lane P contained approximately 5 μg of purified protein.

Lane 1, 2, 3 and 4 are adGST1-1clone10, adGST1-2, adGST1-3 and adGST1-4 respectively.

4.4 Substrate specificities

The specific activities measured for the recombinant *An. dirus* GSTs with various GST substrates are shown in **Table 4.3**. They could not be compared to the previous data of the adGST1-1 (106) because that GST contains the 13-residue leader T_7 tag protein at the N-terminus, which affects the activity of the enzyme.

Table 4.3 Substrate specificities of the recombinant *Anopheles dirus* GSTs.

	Specific activity (μmole/min/mg of protein)			
Substrates	adGST1-1[117]	adGST1-2	adGST1-3	adGST1-4
CDNB	52.2 ± 5.99	43.3 ± 2.79	59.7 ± 1.83	29.1 ± 2.10
DCNB	-	0.080 ± 0.005	0.163 ± 0.012	0.027 ± 0.004
Ethacrynic acid	-	<0.001	0.027 ± 0.003	0.026 ± 0.010
p-Nitrophenethyl bromide	-	0.050 ± 0.006	0.008 ± 0.002	0.028 ± 0.002

The data are means ± standard error for at least five separate assays. The substrate concentrations used were CDNB 1mM; DCNB 1 mM; ethacrynic acid 0.2 mM; and p-nitrophenyl bromide 0.1 mM.
The adGST1-1 is clone 10 without T_7 tag leader protein.

The data indicates that CDNB appears to be the substrate giving highest activity. The adGST1-3 has the greatest activity with two general substrates, CDNB, which gives 59.7 ± 1.83 μmole/min/mg and DCNB, which gives 0.163 ± 0.012 μ mole/min/mg. Using the mammalian pi class substrate, ethacrynic acid, adGST1-2 had no detectable activity while adGST1-3 and adGST1-4 had a similar specific activity, approximately 0.026 μmole/min/mg. With the mammalian theta class substrate, *p*-

nitrophenethyl bromide, all three enzymes had also low specific activities, approximately 0.01-0.05 μmole/min/mg. It confirms that all three adGSTs are distinct from the mammalian pi and theta class GSTs. Although they are the alternative splicing products from the same gene and their N-termini share 100% identity, their substrate specificities are different, indicating that the specificities of the enzymes depend on the C-termini.

4.5 Kinetic parameters

The steady state kinetics were studied with various concentrations of GSH and CDNB. The reactions followed Michaelis-Menten kinetics and the kinetic parameters were determined by non-linear regression analysis as shown in **Table 4.4**.

Table 4.4 Kinetic parameters of *Anopheles dirus* GSTs.

Kinetic parameters	adGST1-1[117]	adGST1-2	adGST1-3	adGST1-4
V_{max}	12.9 ± 0.63	63.9 ± 3.50	67.5 ± 1.97	40.3 ± 1.89
$K_{m\ CDNB}$	0.104 ± 0.028	0.214 ± 0.025	0.100 ± 0.012	0.523 ± 0.067
$K_{m\ GSH}$	0.858 ± 0.179	1.30 ± 0.151	0.404 ± 0.054	0.833 ± 0.084
k_{cat}	5.03	25.9	26.9	16.9
$k_{cat}/K_{m\ CDNB}$	48.4	121	269	32.2
$k_{cat}/K_{m\ GSH}$	5.86	20.0	66.5	20.2

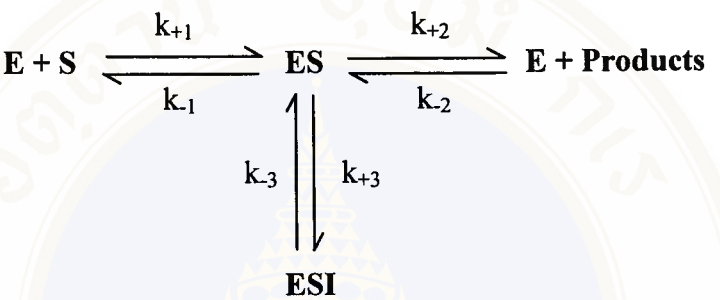
The units are: V_{max} : μmole/min/mg, K_m : mM, k_{cat} : s^{-1}, k_{cat}/K_m : $mM^{-1}s^{-1}$.
The data are the mean ± standard error of at least three separate experiments.

The Maximal velocity or V_{max} of adGST1-2 and adGST1-3 are similar, approximately 60 μmole/min/mg, 1.5-fold higher than adGST1-4, 40 μmole/min/mg and 5-fold higher than adGST1-1, 13 μmole/min/mg. The Michaelis constant or K_m, which indicate the binding affinity of the enzyme, is also important in enzyme characterization. Among all the adGST enzymes, adGST1-3 has the greatest affinity for GSH and CDNB substrates; K_m = 0.1 mM for CDNB and 0.4 mM for GSH, whereas the other enzymes have detectable affinity to CDNB, K_m = 0.1-0.5 mM, but low affinity to GSH, K_m = 0.8-1.2 mM. All adGST enzymes have higher affinity to CDNB than GSH. To determine the catalytic properties, the turn over number for CDNB, k_{cat} and the catalytic efficiency, k_{cat}/K_m with respect to CDNB and GSH were calculated. It was found that all catalytic results of adGST1-3 are significantly greater than the other enzymes. Therefore, among the alternatively spliced products of *adgst1*AS1 gene, adGST1-3 is the most reactive in catalyzing CDNB conjugation while adGST1-1 is the least.

4.6 Insecticide inhibition

In this thesis, permethrin, the pyrithroid insecticide that common used to control the insect in the present day, was used as the inhibitor to study the inhibition kinetics of this insecticide to the three alternatively spliced products of *adgst1*AS1 gene; adGST1-2, adGST1-3 and adGST1-4. There is no report that study about the kinetics of the permethrin inhibition. From double reciprocal plots, 1/V versus 1/S and Hanes-Wolff plots, S/V versus S, permethrin inhibition displays different inhibition kinetics among these three adGST enzymes.

For adGST1-2, the 1/V versus 1/S plot shows parallel straight lines while the S/V versus S plot shows an intercept on the vertical axis as shown in **Figure 4.3**. The plots indicate that the permethrin is an uncompetitive inhibitor for adGST1-2. This type of inhibition can be described by the following reaction scheme:



The permethrin (I) can only combine with the adGST1-2 /CDNB (ES) complex but not with the free adGST1-2 (E) hence both K_m and V_{max} of the adGST1-2 reaction are decreased.

According to the 1/V versus 1/S plot, the inhibitor constant (K_i) is calculated from the following equation:

$$\text{y-intercept} = \frac{(1+i/K_i)}{V}$$

Hence, K_i for adGST1-2 = 0.019 ± 0.005 mM

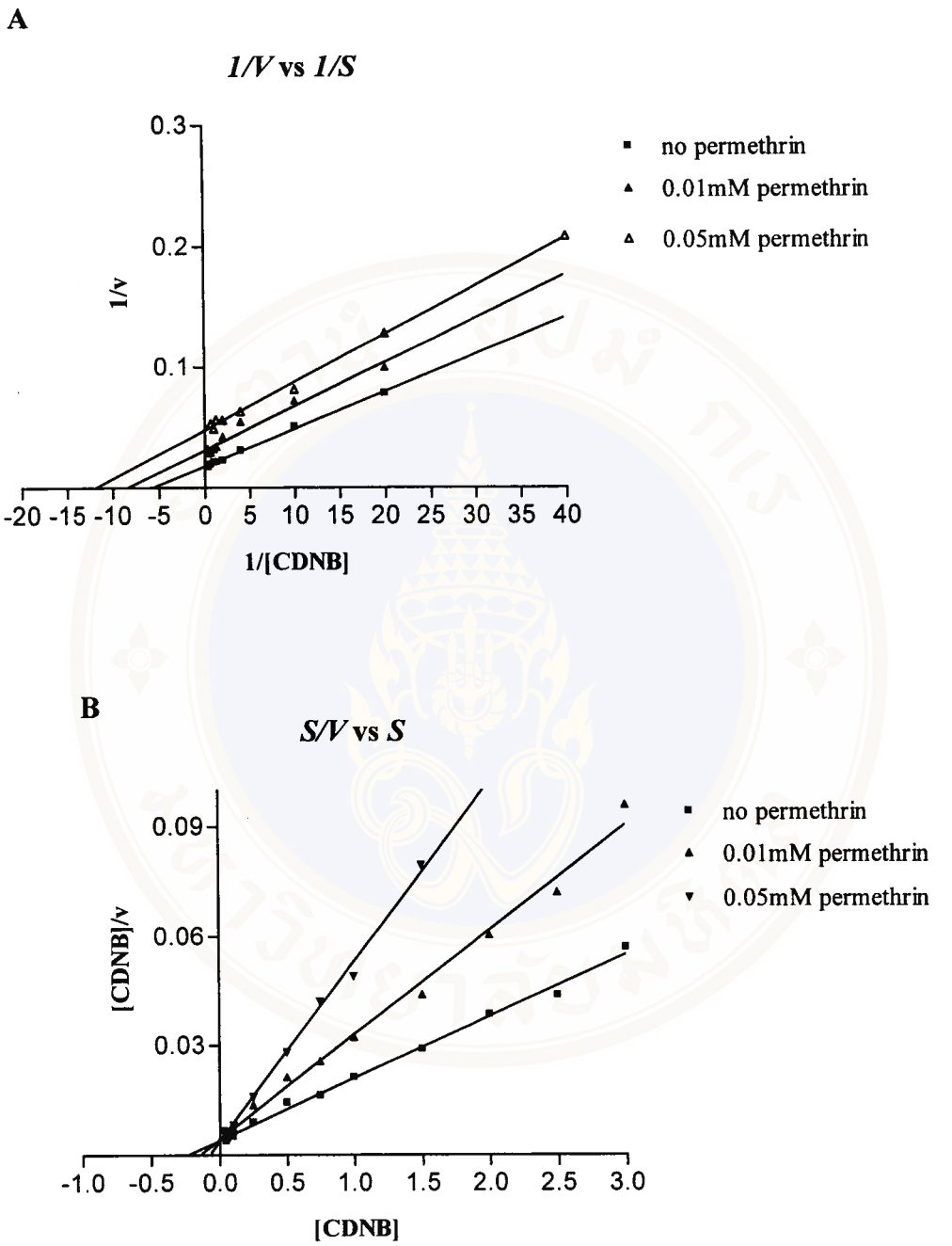
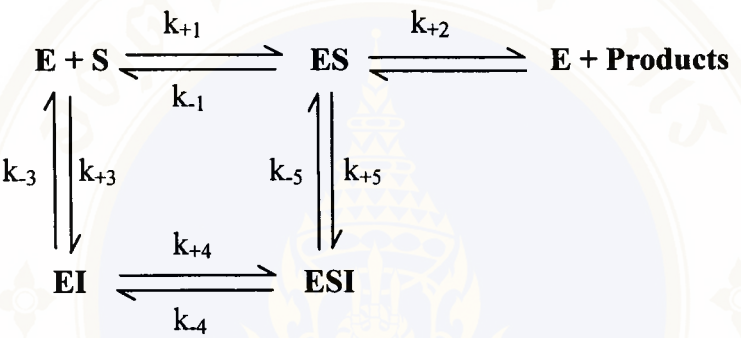


Figure 4.3 Inhibition kinetics of adGST1-2 with permethrin. (A) The 1/V vs 1/S plot shows parallel straight lines with a slope = 0.004 ± 0.0002. (B) The S/V vs S plot shows an intercept on the vertical axis at 0.004 ± 0.0002.

For adGST1-3, both the 1/V versus 1/S plot and the S/V versus S plot show the intercept on the horizontal axis as shown in **Figure 4.4**. The plots indicate that the permethrin is a non-competitive inhibitor for adGST1-3. This type of inhibition can be described by the following reaction scheme:



The permethrin (I) can combine with both the adGST1-3 /CDNB (ES) complex and the free adGST1-3 (E) hence it affects only V_{max} of the adGST1-3 reaction.

According to the 1/V versus 1/S plot, the inhibitor constant (K_i) is calculated from the following equation:

$$\text{y-intercept} = \frac{(1+i/K_i)}{V}$$

Hence, K_i for adGST1-3 = 0.034 ± 0.016 mM

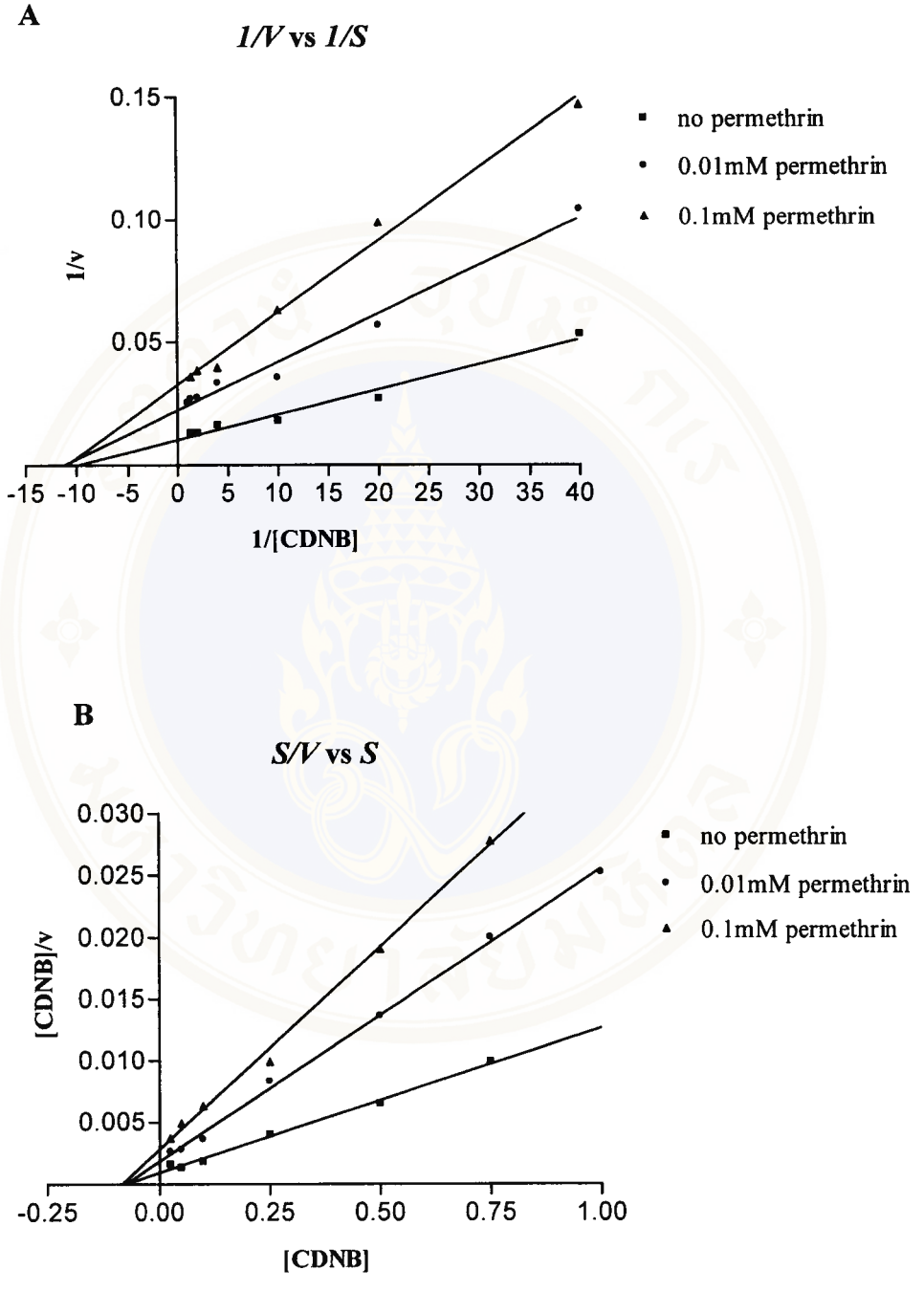
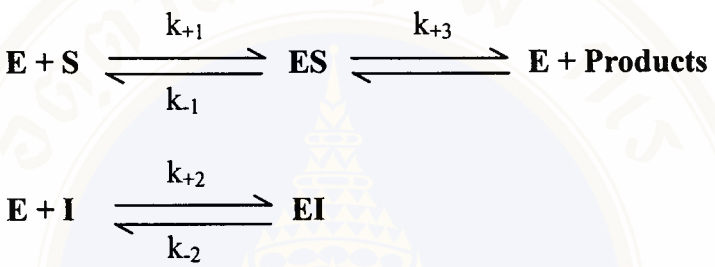


Figure 4.4 Inhibition kinetics of adGST1-3 with permethrin. (A) The 1/V vs 1/S plot shows an intercept on the horizontal axis at 10.73 ± 0.42. (B) The S/V vs S plot shows an intercept on the horizontal axis at 0.081 ± 0.003.

For adGST1-4, the 1/V versus 1/S plot shows an intercept on the vertical axis while the S/V versus S plot shows parallel straight lines as shown in **Figure 4.5**. The plots indicate that the permethrin is a competitive inhibitor for adGST1-4. This type of inhibition can be described by the following reaction scheme:



In this scheme, the amount of inhibition will depend on the relative amount of the permethrin (I) and the CDNB present. The permethrin will compete to bind with the free adGST1-4 (E) and increase the apparent of K_m of the adGST1-4 reaction.

According to the 1/V versus 1/S plot, the inhibitor constant (K_i) is calculated from the followed equation:

$$\text{x-intercept} = \frac{1}{K_m(1+i/K_i)}.$$

Hence, K_i for adGST1-4 = 0.013 ± 0.005 mM

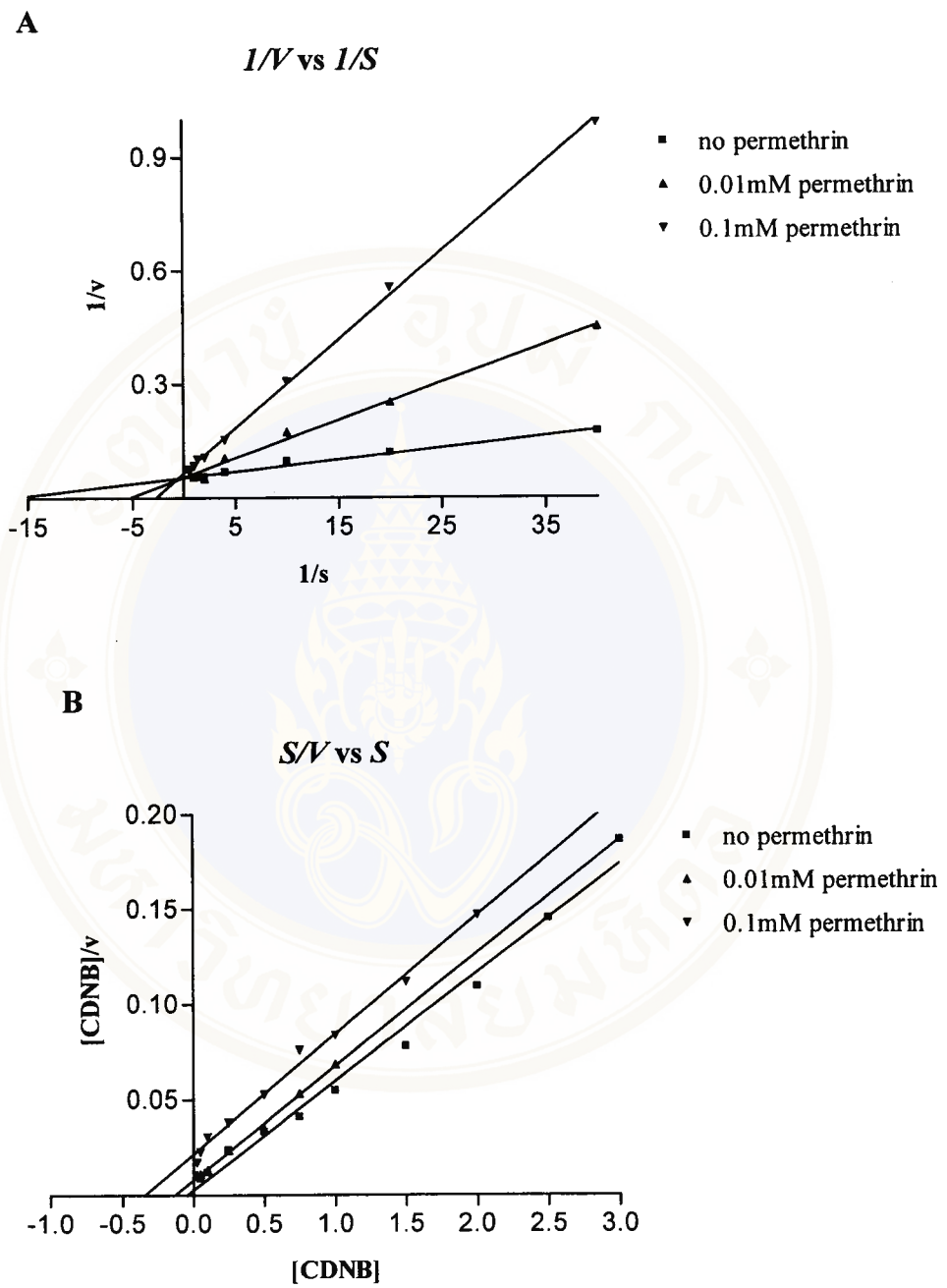


Figure 4.5 Inhibition kinetics of adGST1-4 with permethrin. (A) The 1/V vs 1/S plot shows an intercept on the vertical axis at 0.054 ± 0.006. (B) The S/V vs S plot shows paralell straight lines with a slope = 0.060 ± 0.002.

Among the three adGSTs, adGST1-2, adGST1-3 and adGST1-4, the inhibition kinetics of permethrin insecticide with these three enzymes are very different as shown in **Table 4.5.**

Table 4.5 Inhibition kinetics of permethrin to *Anopheles dirus* GSTs

Parameters	AdGST1-2	AdGST1-3	AdGST1-4
% Inhibition	81.4	70.2	39.7
K_i (mM)	0.019 ± 0.005	0.034 ± 0.016	0.013 ± 0.005
Inhibition type	Uncompetitive	Non-competitive	Competitive

The %inhibition of permethrin was performed in duplicate using the standard assay of CDNB conjugating activity. Final concentrations of GSH, CDNB and permethrin are 10.0, 1.0 and 0.1 mM respectively.
The K_i are the mean ± standard error of three separate experiments; adGST1-2 were tested at 0, 0.01 and 0.05 mM of permethrin, adGST1-3 were tested at 0, 0.01 and 0.1 mM of permethrin and adGST1-4 were tested at 0, 0.01 and 0.1 mM of permethrin.

The permethrin insecticide have been reported showing the inhibition of CDNB conjugating activity with 85.4% to adGST1-1 (106), 58.3% to agGST1-5, 9.8% to agGST1-6 (108), and 45.6% to adGST1-1 without T_7 tag protein (117). In this thesis project, permethrin can also inhibit the CDNB conjugating activity of three alternative splicing products of *adgst1*AS1 gene, it showed the greatest inhibition, 81.4% to that of adGST1-2 whereas the lowest, 39.7% to that of adGST1-4, although

adGST1-4 can bind at higher affinity to permethrin, which indicated by K_i = 0.013 ± 0.005 mM. Furthermore, permethrin inhibitions of all three adGSTs have shown distinct kinetics types, although they were spliced from the same gene. These results also support that the differences of C-termini in adGSTs affect their properties.

CHAPTER V

DISCUSSION

After one complete GST gene, *adgst1AS1* (112), was identified from a genomic library of *Anopneles dirus*, it was found to be the orthologous gene of the *aggst1α* gene identified from *An. gambiae* (107). This gene contains 6 exons, which was predicted to produce four alternatively spliced products. One splicing product of this gene, *adgst1-1*, has been isolated and characterized previously by Prapanthadara *et al* (106).

In this thesis, three novel cDNA species, *adgst1-2*, *adgst1-3* and *adgst1-4*, which encode the full-length glutathione S-transferases have been isolated from a susceptible strain of *Anopheles dirus* species B, using the primers designed from the *adgst1AS1* gene. Their nucleotide sequences are different from the genomic sequence at some positions, changed from C to T or G to A but the translated amino acid sequences are not changed. These changes occurred because the total RNA used in the RT-PCR amplification was extracted from the 4th instar larvae, which may contain multiple alleles of RNA species. The splicing products of *aggst1α* were reported simultaneously expressed in all developmental stages and in both males and females mosquitoes (107). It seems likely that the choice of 3' splice site is not developmentally regulated. However, the factors determining the choice of 3'splice site of *adgst1AS1* gene in *A. dirus*, either for the developmental regulation or tissue dependent regulation, remain to be elucidated.

To determine the role of these GSTs in insecticide resistance, the enzymes were expressed in *E. coli* under the strong T_7 promotor system without a 13-residues leader sequence tag attached to the N-termini, which may affect the activity of the enzymes. They could not be compared to the previous data of other GSTs because those enzymes contain that leader tag sequence. From the comparison of the specific activities of the adGST1-2, adGST1-3 and adGST1-4 with the various GST substrates, all of them have high activity with the general substrate, CDNB, 43.3, 59.7 and 29.1 μ mole/min/mg respectively and detectable activity to another substrate, DCNB. AdGST1-3 has highest activities with these two general substrates. Several studies have suggested that GST activities as measured with the substrates, CDNB and DCNB are correlated positively with resistance (see for example 24, 93, 96, 88, 102, 106). Therefore the activity with CDNB and DCNB exhibited by adGST1-2, adGST1-3 and adGST1-4 indicates that these enzymes may be involved in insecticide resistance.

Comparison of these three translated amino acids among all classes of GSTs showed that they are classified as class I or delta class GSTs because they share 50-92% identity with the previously characterized Dipteran insects class I GSTs, except 39-40% to *An. gambiae* GST1-7. They also share 32-45% identity with that of Lepidopteran insects. While they are only distantly related to the insect class II or sigma class GSTs and the other classes; theta, pi, alpha, mu, kappa and zeta GSTs because they share <23% identity. Furthermore, these three adGSTs also have very low activities against ethacrynic acid and *p*-nitrophenyl bromide, which are known substrates of mammalian pi class GSTs (119-120) and mammalian theta class GSTs (38). Hence the difference in the amino acid sequences and the substrate specificities

can confirm that insect class I or delta class GSTs are distinct to all other classes GSTs.

The insect class I or delta class GSTs of *M. domestica*, *Drosophila spp.* and *A. gambiae* are highly variable and encoded by complex gene families. Many isoenzymes from these species are reported in the GenBank database; 9 isoforms from *M. domestica*, >20 isoforms from *Drosophila spp.* and 7 isoforms from *An. gambiae*, but only 1 isoforms from *An. dirus* were isolated to date. It suggests that other isoforms will exist in this species.

The exon arrangement of the *adgst1AS1* gene was found to be similar with the *An. gambiae aggst1*α gene and showed >78.5% nucleotide sequence identity (112). Their alternatively spliced products also shared very high amino acid sequence identity; 85% between adGST1-2 and agGST1-4, 92% between adGST1-3 and agGST1-5 and 85% between adGST1-4 and agGST1-3, which confirms that the orthologous GSTs have occurred across the anopheline species, although, they are malaria vectors in different geographically epidemic areas. The high conservation of the splicing products of *adgst1AS1* gene and *aggst1*α gene at the amino acid sequence level and the splicing pattern suggest that these two GST genes should have important physiological roles and they must be maintained in these anopheline species.

Among the alternatively spliced products of the *adgst1AS1* gene, their steady state kinetic properties were compared. The adGST1-3 has highest affinity for both GSH, K_m = 0.404 ± 0.054 mM and CDNB, K_m = 0.100 ± 0.012 mM. The adGST1-3 is also the most reactive enzyme in catalyzing CDNB conjugation, k_{cat} = 26.9 s^{-1}. All adGST enzymes have higher affinity for CDNB than GSH, which suggests that these

enzymes normally operate close to a state of saturation with respect to electrophilic compounds.

Several data suggests that GSTs have an important role in the acquisition of resistance to insecticide (89). Many resistant insects have been shown to contain elevated levels of GST activity. For example, the overexpression of a GST is thought to be responsible for resistance to organophosphates in a strain of *Musca domestica* (88) and elevated levels of GST activity have also been found in DDT-resistant strains of *Aedes aegypti* (94). DDT-dehydrochlorinase activity catalyzed by GST is the major mechanism responsible for DDT resistance (121) and glutathione conjugation catalyzed by GST is also a secondary mechanism after oxidation in some organophosphate resistance (122). For the cloned insect GSTs, the DDT dehydrochlorinase activity has been reported in GSTD1 and GSTD21 of *Drosophila* (123) and adGST1-1 of *Anopheles dirus* (106). To determine whether the cloned GSTs are able to recognize any insecticides, the ability of the insecticides to block the CDNB conjugation activity of the recombinant GSTs will be determined. The simple inhibition (% inhibition) is the method that will be used in the insecticide resistance investigation. This thesis is the first demonstration of the permethrin or pyrethoid insecticide inhibition kinetics. Previously, the inhibition of CDNB conjugating activity by this insecticide was reported as the simple inhibition in adGST1-1 (106), agGST1-5 and agGST1-6 (108). The inhibition kinetic studies of three splicing product adGSTs in the presence of various CDNB concentrations have been observed. Surprisingly, the permethrin inhibition on the CDNB conjugation displayed markedly different inhibition kinetics among the three splicing products of *adgst1AS1* gene, uncompetitive for adGST1-2, noncompetitive for adGST1-3 and competitive for

adGST1-4. All three enzymes could bind permethrin at high affinity due to their inhibition constants are approximately 0.02-0.03 mM. Furthermore, this insecticide also blocked the CDNB activity of all four alternatively spliced products of *adgst1AS1* gene at high values, 85.5% in adGST1-1 (107), 70% in adGST1-2, 80% in adGST1-3 and 40% in adGST1-4. The inhibition kinetics of the recombinant proteins suggests that these four enzymes could play a role in resistance to permethrin.

Although all adGSTs to date are the splicing products of the same gene which share N-terminus sequence, but their substrate specificities, steady state kinetics with respect to both GSH and CDNB, and kinetic pattern to the permethrin insecticide are very different. Hence, it is suggested that C-terminal changes in these GSTs could produce enzymes with very different potentials for conferring resistance.

Crystal structures of mammalian and insect GSTs show that the majority of active site residues involved in the binding and activation of GSH are found within the N terminus and hence this region of protein is highly conserved between GSTs. The divergence in the C terminus confers the variation in substrate specificities of different GST isoenzymes (40, 118). Based on the crystal structure of a *Lucilia cuprina* delta class GST, a serine residue near the N terminus of theta class GSTs and insect class I GSTs, which plays a vital role in the catalytic mechanism of the GSTs, is highly conserved in N-termini (40). Three splicing products of *adgst1AS1* gene and the previous reported sequence of adGST1-1 (106) are identical at the N termini (residues 1-45) and therefore possess this N-terminal serine at position 9 (**Figure 3.9**). In contrast, the rest of sequence showed variation in some parts of the amino acid sequence which are proposed to form the C-terminal domain of the enzyme subunit, especially in all loops and in alpha helices, α5 and α7 (**Figure 3.8**). However, only a

one or two amino acid change in the adGST1-1 allelic form was shown to affect the activity and kinetic properties of the enzyme (117). The adGST1-2, adGST1-3 and adGST1-4 are different in the translated amino acid lengths, being 217, 209 and 219 residues respectively and the difference occurs in the C terminal part, which is the hydrophobic binding site. While the GST fold itself is well conserved, variations in the sequences, due to substitutions, deletions, or insertions in the β-strands and α-helices, have led to some overall readjustments of amino acid residues such as the rotation of structural domains (29). Therefore, the sequence difference of these three enzymes may lead to overall conformational changes and subsequently changes in both substrate and inhibitor binding.

Alternative splicing is known to be a mechanism, which generates functionally diverse isoforms for virtually every type of protein involved in metazoan development, cell function and physiology. Alternative splicing influences cell function by generating structural diversity among products of C-terminal difference (124). Therefore the *adgst1AS1* gene possesses an alternative splicing mechanism to produce four different mature products in order to increase the diversity of GSTs and expand their substrate range with a minimal length of gene.

Knowing the three-dimensional structure of a native enzyme, preferably when complexed with a substrate, transition-state analogue, or specific inhibitor, is essential for unraveling and understanding structure-function relationships. Structure analysis by X-ray crystallography is distinctive in its ability to determine atomic positions in a crystal lattice and hence yield a detailed and precise description of a protein's architecture (29). Therefore, further investigation of the structure-function relationship

of the *An. dirus* GST isoenzymes are essential in understanding the physiological importance of GSTs in xenobiotic detoxication.

CHAPTER VII

CONCLUSIONS

1. In this thesis, three cDNA sequences of glutathione S-transferases (GSTs); *adgst1-2*, *adgst1-3* and *adgst1-4*, which are the alternatively spliced products of the *adgst1AS1* gene have been isolated from the 4th instar larvae of *Anopheles dirus* B mosquito using the primers designed from the *adgst1AS1* gene. Their nucleotide sequences share >67% identity and the translated amino acid sequences share >61% identity.

2. They have been classified as insect class I GSTs on the basis of their high sequence identity of >39% to members of the insect class I family but low identity of <23% to all other classes GSTs.

3. The translated amino acid sequence identities between adGST1-2 versus agGST1-4, adGST1-3 versus agGST1-5, and adGST1-4 versus agGST1-3 show >85% identity which confirms that orthologous forms of isoenzymes occur across the anopheline species.

4. These three recombinant GSTs may be localized in peak 4b or 4c of the GST sequential chromatography pattern from larvae on the basis of binding to S-hexylglutathione agarose column and lack of homology to peak 4a sequence.

5. The purified enzymes were shown to be homogeneous preparations on SDS-PAGE. The subunit sizes of adGST1-2, adGST1-3 and adGST1-4 are 24.3, 23.9 and 25.1 kDa respectively on the basis of the amino acid compositions.

6. The differences of the C-termini of these three enzymes lead to distinct substrate specificities, kinetic properties and insecticide inhibition properties.

REFERENCES

1. Floore, T. Mosquito Information. [Online]. The American Mosquito Control Association, Mulrennan Research Center, Florida, U.S.A. Available from: http://www.mosquito.org/mosquito.html [Accessed April 2000].
2. Nomina Insecta Neacrtica. [Online]. Nearctica.com, Inc. Available from: http://www.nearctica.com/Nomina/nomina.htm [Accessed April 2000].
3. FAQ's on mosquitoes. [Online]. Rutgers Entomology NJMCA New Jersy Mosquito Homepage. Available from: http://www-rci.rutgers.edu/~insects/mosfaq.htm [Accessed April 2000].
4. Rietveld, A. Frequently-asked-questions about malaria. [Online]. Available from http://www.outbreak.org/cgi-unreg/dynaserve.exe/Malaria/faq.html [Accessed October 1999].
5. Malaria. [Online]. Department of Microbiology and Immunology, University of Leicester, U.K. Updated: 4 January 2000. Available from: http://www-micro.msb.le.ac.uk/224/Malaria.html [Accessed April 2000].
6. Malaria vector species. An AnoDB Database. [Online]. The Institute of Molecular Bilogy and Biotechnology, Greece. Available from: http://konops.imbb.forth.gr/AnoDB/Species/malariaspecies.html [Accessed April 2000].
7. Matsumura, F. Classification of insecticides. In: Matsumura, F., editor. Toxicology of insecticide. New York: Plenum Press; 1975. p. 47-103.

8. Roush, R.T. and Tabashnik, B.E. Pesticide Resistance in Arthropods. New York: Chapman and Hall; 1990.

9. Wood, R.J. Insecticide resistance. Genes and mechanisms. In: Bishop, J.A. and Cook, L.M., eds. Genetic consequences of man made change. New York: Academic Press; 1981. p. 53-96.

10. Noppun, V., Saito, T. and Miyata, T. Cuticular penetration in fenvalerate-resistant and susceptible strains of the diamondback moth, *Plutella xylostella* (L). Pestic Biochem Physiol 1989; 33: 83-87.

11. Golenda, C.F. and Forgash, A.J. The distribution and metabolism of fenvalerate in pyrethoid-resistant and susceptible strains houseflies. Pestic Biochem Physiol 1989; 33: 37-48.

12. Terriere, L.C. Induction of detoxication enzymes in insects. Ann Rev Entomol 1984; 29: 71-88.

13. Jakoby, W.B.and Ziegler, D.M. The enzymes of detoxication. J Biol Chem 1990; 265(34): 20715-20718.

14. Daniel, V. Glutathione S-transferases: Gene structure and regulation of expression. Critical Reviews in Biochemistry and Molecular Biology 1993; 28(3): 173-207.

15. Salinas, A.E. and Wong, M.G. Glutathione S-transferase – A review, Current Medical Chemistry. Bentham Science Publishers B. V. 1999; 6: 279-309.

16. Ishikawa, T. The ATP-dependent glutathione S-conjugate export pump. Trends Biochem Sci 1992; 17(11): 463-468.

17. Hayes, J.D., Wolf, C.R. Molecular mechanisms of drug resistance. Biochem J 1990; 272(2): 281-295.

18. Habig, W.H., Pabst, M.J. and Jakoby, W.B. Glutathione S-transferase: the first enzymatic step in mercapuric acid formation. J Biol Chem 1974; 249 (22): 7130-7139.

19. Mannervik, B. and Danielson, U.H. Glutathione transferases- structure and catalytic activity. Crit Rev Biochem Mol Biol 1988; 23: 283-337.

20. Kuzmich, S., Vanderveer, L.A., Walsh, E.S., LaCreta, F.P. and Tew, K.D. Increased levels of glutathione S-transferase π transcript as a mechanism of resistance to ethacrynic acid. Biochem J 1992; 281: 219-224.

21. Board, P., Coggan, M., Johnston, P., Ross, V., Suzuki, T. and Webb, G. Genetic heterogeneity of the human glutathione transferases: a complex of gene families. Pharmacol Ther 1990; 48(3): 357-369.

22. Clark, A.G. The comparative enzymology of the glutathione S-transferases from non-vertebrate organisms. Comp Biochem Physiol 1989; 92B(3): 419-446.

23. Clark, A.G., Dick, G.L. and Smith, J.N. Kinetic studies on a glutathione S-transferase from the larvae of *Costelytra zealandica.* Biochem J 1984; 217: 51-58.

24. Clark, A.G., Shamaan, N.A., Sinclair, M.D. and Deuterman, W.C. Insecticide metabolism by multiple glutathione S-transferases in two strains of the house fly, *Musca domestica* (L). Pestic Biochem Physiol 1986; 25: 169-175.

25. Dierickx, P.J. Anionic and neutral glutathione S-transferase isoenzymes in the freshwater worm *Tubifex tubifex* (O.F.M.). Arch Int Physiol Biochim 1985; 93(3): 193-198.

26. Hayes, J.D. Rapid purification of glutathione S-transferases by gradient affinity elution of the glutathione-agarose and the S-hexylglutathione-agarose chromatography matrices. In: Hayes, J.D., Pickett, C.B. and Mantle, T.J., eds. Glutathione S-transferases and Drug resistance. London: Taylor and Francis; 1990. p. 17-33.

27. Simon, P.C. and Vanre Jagt, D.L. Purification of glutathione S-transferases from human liver by glutathione-affinity chromatography. Anal Biochem 1977; 82: 334-341.

28. Guthenberg, C., Akerfeldt, K. and Mannervik, B. Purification of glutathione S-transferase from human placenta. Acta Chem Scan 1979; B33: 595-596.

29. Dirr, H., Reinemer, P. and Huber, R. X-ray crystal structures of cytosolic glutathione S-transferases. Implication for protein architecture, substrate recognition and catalytic function. Eur J Biochem 1994; 220: 645-661.

30. Danielson, U.H. and Mannervik, B. Kinetic independence of the subunits of cytosolic glutathione transferase from the rat. Biochem J 1985; 231: 263-268.

31. Epp, O., Ladenstein, R. and Wendel, A. The refined structure of the selenoenzyme glutathione peroxide at 0.2-nm resolution. Eur. J. Biochem. 1983; 133: 51-69.

32. Holmgren, A., Soderberg, B.O., Eklund, H. and Branden, C. Three-dimensional structure of *Escherichia coli* thiredoxin-S2 to 2.8 Å resolution. Proc Natl Acad Sci USA 1975; 72: 2305-2309.

33. Eklund, H., Ingelman, M., Soderberg, B.O., Uhlin, T., Nordlund, P., Nikkola M., Sonnerstam, U. and Joelson, T. Structure of oxidized bacteriophage T_4

glutaredoxin (thioredoxin). Refinement of native and mutant proteins. J Mol Biol 1992; 228: 596-618.

34. Karplus, P.A. and Schulz, G.E. Refined structure of glutathione reductase at 1.54 A resolution. J Mol Biol 1987; 195: 701-729.

35. Sinning, I., Kleywegt, G.J., Cowan, S.W., Reinemer, P., Dirr, H.W., Huber, R., Gilliland, G.L., Armstrong, R.N., Ji, X., Board, P.G., Olin, B., Mannervik, B. and Jones, T.A. Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with the mu and pi class enzyme. J Mol Biol 1993; 232: 192-212.

36. Rushmore, T.H. and Pickett, C.B. Glutathione S-transferase: structure, regulation and therapeutic implications. J Biochem Chem 1993; 268(16): 11475-11478.

37. Hiratsuka, A., Sebata, N. Kawashima, K., Okuda, H., Ogura, K., Watabe, T., Satoh, K., Hatayama, I., Tsuchida, S., Ishikawa, T.and Sato, K. New class of rat glutathione S-transferase Yrs-Yrs inactivating reactive sulfate esters as metabolites of carcinogenic arymethanols. 1990; 265: 11973-11981.

38. Meyer, D.J., Coles, B., Pemple, S.E., Gilmore, K.S. and Ketterer B. Theta, a new class of glutathione transferases purified from rat and man. Biochem J 1991; 274: 409-414.

39. Buetler, T.M. and Eaton, D.L. Complementary DNA cloning, messenger RNA expression, and induction of alpha-class glutathione S-transferases in mouse tissues. Cancer Res 1992; 52(2): 314-318.

40. Wilce, M.C.J., Board, P.G., Feil, S.C. and Parker, M.W. Crystal structure of a theta class glutathione S-transferase. EMBO J 1995; 14(100: 2133-2143.

41. Wilce, M.C.J., Parker, M.W. Structure and function of glutathion S-transferases. Biochem & Biophys Acta 1994; 1205: 1-18.

42. Armstrong, R.N. Glutathione S-transferases: structure and mechanism of an archetypical detoxication enzyme. Adv Enzymol Relat Areas Mol Biol 1994; 69: 1-44.

43. Graminski, G.F., Kubo, Y. and Armstrong, R.N. Spectroscopic and kinetic evidence for the thiolate anion of glutathione at the active site of glutathione S-transferases. Biochemistry 1989; 28: 3562-3568.

44. Graminski, G.F., Zhang, P., Sesay, M.A., Ammon, H.L. and Armstrong, R.N. Formation of the 1-(S-glutathionyl)-2,4,6-trinitrocyclohexadienate anion at the active site of glutathione S-transferase. Biochemistry 1989; 28: 6252-6258.

45. Huskey, S.E.W., Huskey, W.P. and Lu, A.Y.H. Contributions of thiolate "desolvation" to catalysis by glutathione S-transferase isoenzymes 1-1 and 2-2: evidence from kinetic solvent isotope effects. J Am Chem Soc 1991; 113: 2283-2290.

46. Jakoby, W.B. The glutathione S-transferases: a group of multifunctional detoxication proteins. Adv Enzymol 1978; 46: 383-414.

47. Prohaska, J.R. The glutathione peroxidase activity of glutathione S-transferases. Biochim Biophys Acta 1980; 611(1): 87-98.

48. Mannervik, B. The isoenzymes of glutathione transferases. Adv Enzymol Rel Areas Mol Biol 1985; 57: 357-417.

49. Hayes, J.D., Judah, D.J., McLellan, L.I., Kerr, L.A., Peacock, S.D. and Neal, G.E. Ethoxyquin-induced resistance to aflatoxin B_1 in the rat is associated with the expression of a novel Alpha-class geltathione S-transferase subunit, Yc_2 , which possesses high catalytic activity for aflatoxin B_1-8,9-epoxide. Biochem J 1991; 279: 385-398.

50. Clark, A.G., Smith, J.N. and Speir, T.W. Cross specificity in some vertebrate and insect glutathione S-transferases with methyl parathion (dimethyl p-nitrophenyl phosphorothionate), 1-chloro-2,4-dinitrobenzene and *S*-crotonyl-N-cysteamine as substrates. Biochem J 1973; 135: 385-392.

51. Habig, W.H. and Jakoby, W.B. Assays for differentiation of glutathione S-transferases. Methods Enzymol 1981; 77: 398-405.

52. Meyer, D.J.,Christodoulides, L.G., Tan, K.H. and Ketterer, B. Isolation, properties and tissue distribution of rat glutathione transferase. E FEBS Lett 1984; 173: 327.

53. Grove, G., Zarlengo, R.P., Timmerman, K.P., Li, N.Q., Tam, M.F. and Tu, C.-P.D. Characterization and heterospecific expression of cDNA clones of genes in the maize GSH S-transferase multigene family. Nucleic Acids Res 1988; 16: 425-438.

54. Hoesch, R.M and Boyer, T.D. Localization of a portion of the active site of two rat liver glutathione S-transferases using a photoaffinity label. J Biol Chem 1989; 264(30): 17712-17717.

55. Board, P.G. and Mannervik, B. The contribution of the C-terminal sequence to the catalytic activity of GST2, a human alpha-class glutathione transferase. Biochem J 1991; 275: 171-174.

56. Askelof, P., Guthenberg, C, Jakobson, I. and Mannervik, B. Purification and characterization of two glutathione S-aryltransferase activities from rat liver. Biochem J 1975; 147(3): 513-522.

57. Ivanetich, K. M. and Goold, R.D. A rapid equilibrium random sequential bi-bi mechanism for human placental glutathione S-transferase. Biochim Biophys Acta 1989; 998(1): 7-13.

58. Chen, W.-J., Graminski, G.F. and Armstrong, R.N. Dissection of the catalytic mechanism of isozyme 4-4 of glutathione S-transferase with alternative substrates. Biochemistry 1988; 27: 647-654.

59. D'Silva, C. Inhibition and recignition studies on the glutathione-binding site of equine liver glutathione S-transferase. Biochem J 1990; 271: 161-165.

60. Young, P.R. and Briedis, A.V. Purification and kinetic mechanism of the major glutathione S-transferase from bovine brain. Biochem J 1989; 257: 541-548.

61. Usui, K., Fukami, J.I. and Shishido, T. Insect glutathione S-transferases from the bodies of American cockroaches active on organophosphorus triesters. Pestic Biochem Physiol 1977; 7: 249-260.

62. Chang, C.K., Clark, A.G., Fieldes, A. and Pounds, S. Some properties of glutathione S-transferase from the larvae of the *Gallena mellonella*. Insect Biochem 1981; 11: 179-186.

63. Wood, E., Casabe, N., Melgar, F. and Zerba, E. Distribution and properties of glutathione S-transferase from *Triatoma infestans*. Comp Biochem Physiol 1986; 84B: 607-618.

64. Hayes, J.D. and Mantle, T.J. Anomalous electrophoretic behavior of the glutathione S-transferases Ya and Tk subunits isolated from man and rodents. Biochem J 1986; 237: 731-740.

65. Walholm, M., Gythenberg, C., Mannervik, B., von Bahr, C. and Glaumann, H. Identification of a new glutathione S-transferase in human liver. Acta Chem Scand 1980; B34: 607-610.

66. Hales, B.F., Jaeger, V. and Neims, A.H. Isoelectric focusing of glutathione S-transferases from rat liver and kidney. Biochem J 1978; 175: 937-943.

67. Aceto, A., Di Ilio, C., Angelucci, S., Felaco, M. and Federici, G. Glutathione S-transferase isoenzymes from human testis. Biochem Pharmacol 1989; 38: 3653-3660.

68. Sherman, M., Campbell, J.A.H., Titmus, S.A., Kew, M.C. and Kirsch, R.E. Glutathione S-transferase in hepato-cellular carcinoma. Hepatology 1983; 3: 170-176.

69. Ostlund Farrants, A.K., Meyer, D.J., Coles, B., Southan, C., Aitken, A., Johnson, P. J. and Ketterer, B. The separation of glutathione transferase subunits by using reversed-phase high-pressure liquid chromatography. Biochem J 1987; 245: 423-428.

70. Yeh, H.I., Hsieh, C.H., Wang, L.Y., Tsai, S.P., Hsu, H.Y. and Tam, M.F. Mass spectrometric analysis of rat liver cytosolic glutathione S-transferases. Modifications are limited to N-terminal processing. Biochem J 1995; 308: 69-75.

71. Hayes, J.D. and Mantle, T.J. Use of immuno-blot techniques to discriminate between the glutathione S-transferase Yf, Yk, Ya, Yn/Yb, and Yc

subunits and to study their distribution in extra-hepatic tissues. Evidence for three immunochemically distrinct groups of transferases in the rat. Biochem J 1986; 233: 779-788.

72. Hayes, J.D., Gilligan, D., Chapman, B.J. and Beckett, G.J. Purification of human hepatic glutathione S-transferases and the development of a radioimmunoassay for their measurement in plasma. Clin Chim Acta 1983; 134: 107-121.

73. McCusker, F.M.G., Boyces, S.J. and Mantle, T.J. The development of glutathione S-transferase subunits in rat liver. Sensitive detection of the major subunit forms of rat glutathione S-transferase by using an ELISA method. Biochem J 1989; 262: 463-467.

74. Peter, W.H.M., Kock, L., Nagengast, F.M. and Roelofs, H.M.J. Immuno-detection with a monoclonal antibody of glutathione S-transferase mu in patients with and without carcinomas. Biochem Pharmacol 1990; 39: 591-597.

75. Peter, W.H.M., Boon, C.E.W., Roelofs, H.M.J., Wobbes, T., Nagengast, F.M. and Kremers, P.G. Expression of drug-metabolising enzymes and P-170 glycoprotein in colerectal carcinoma and normal mucosa. Gastroenterology 1992; 103: 448-455.

76. Beckett, G.J. and Hayes, J.D. Glutathione S-transferases: biomedical applications. Adv Clin Chem 1993; 30: 279-378.

77. Jakoby, W.B. The glutathione S-transferase: a group of multifunctional detoxication proteins. Adv Enzymol 1978; 46: 383-414.

78. Mannervik, B., Ålin, P., Guthenberg, C., Jensson, H., Tahir, M.K., Warholm, M. and Jörnvall, H. Identification of three classes of cytosolic glutathione transferase common to several mammalian species: correlation between structural data and enzymatic properties. Proc Natl Acad Sci USA 1985; 82: 7202-7206.

79. Pemble, S.E. and Taylor, J.B. An evolutionary perspective on glutathione transferases inferred from class-theta glutathione transferase cDNA sequences. Biochem J 1992; 287: 957-963.

80. Pemble, S.E., Wardle, A.F. and Taylor, J.B. Glutathione S-transferase class kappa: characterization by the cloning of rat mitochondria GST and identification of human homologue. Biochem J 1996; 319: 749-754.

81. Board, P., Baker, R.T., Chelvanayagam, G. and Jermiin, L.S. Zeta, a novel class of glutathione transferases in a range of species from plants to humans. Biochem J 1997; 328: 929-935.

82. Ahmad, H., Singhal, S.S., Saxena, M. and Awasthi,Y.C. Characterization of two novel subunits of the alpha-class glutathione S-transferases of human liver. Biochim Biophys Acta 1993; 1161(2/3): 333-336.

83. Mannervik, B., Awashi, Y.C., Board, P.G., Hayes, J.D., Di Ilio, C., Ketterer, B., Listowsky, I., Morgenstern, R., Muramatsu, M., Pearson, W.R., Pickett, C.B., Sato, K., Widersten, M.and Wolf, C.R. Nomenclature for human glutathione transferases. Biochem J 1992; 282: 305-306.

84. Motoyama, N. and Dauterman, W.C. Molecular weight, subunits, and multiple forms of GST from the house fly. Insect Biochem1978; 8: 337-348.

85. Clark, A.G. and Dauterman, W.C. The characterization by affinity chromatography of glutathione S-transferase from different strains of house fly. Pestic Biochem Physiol 1982; 17: 307-314.

86. Clark, A.G. and Shamaan, N.A. Evidence that DDT-dehydrochlorinase from the house fly is a glutathione S-transferase. Pestic Biochem Physiol 1984; 22: 249-261.

87. Clark, A.G., Shamaan, N.A., Dauterman, W.C. and Hayaoko, T. Characterization of multiple glutathione S-transferases from the house fly, *Musca domestica*(L). Pestic Biochem Physiol 1984; 22: 51-59.

88. Wang, J.Y., McCommas, S. and Syvanen, M. Molecular cloning of a glutathione S-transferase overproduced in an insecticide resistant strain of the housefly, *Musca Domestica.* Molec Gen Genet 1991; 227: 260-266.

89. Fournier, D., Bride, J.M., Poirie, M., Berge, J.B. and Plapp, F.W. Insect glutathione S-transferases: Biochemical characterization of the major forms from houseflies susceptible and resistant to insecticides. J Biol Chem1992; 267: 260-266.

90. Clark, A.G., Dick, G.L., Martindale, S.M. and Smith, J.N. Glutathione S-transferases from the New Zealand grass grub, *Costelytra zealandica.* Their isolation, characterization and the effect on their activity of endogeneous factors. Insect Biochem 1985; 15: 35-44.

91. Cochrane, B.J., Morrissen, J.J. and LeBlanc, G.A. The genetics of xenobiotic metabolism in *Drosophila*-IV. Purification and characterization of the major glutathione S-transferase. Insect Biochem 1987; 17: 731-738.

92. Toung, Y.P.S., Hsieh, T.S. and Tu, C.P.D. The glutathione S-transferase D genes. A divergently organized, intronless gene family in *Drosophila melanogaster.* J Biol Chem 1990; 268: 9737-9746.

93. Grant, D.F. and Matsumura, F. Glutathione S-transferase 1 and 2 in susceptible and insecticides resistant *Aedes aegypti*: purification, characterization and isozyme-specific regulation. Insect Biochem1989; 21: 421-433.

94. Grant, D.F., Dietze, E.C. and Hammock, B.D. Glutathione S-transferase isozymes in *Aedes aegypti* purification, characterization and isozyme-specific regulation. Insect Biochem 1991; 18: 615-622.

95. Kotze, A.C. and Rose, H.A. Purification and properties of glutathione S-transferase from the larvae of the Australian sheep blowfly, *Lucilia cuprina.* Insect Biochem 1989; 19: 703-713.

96. Prapanthadara, L., Hemingway, J. and Ketterman, A.J. Partial purification and characterization of glutathione S-transferase involved in DDT-resistance from mosquito, *Anopheles gambiae*. Pestic Biochem Physiol 1993; 47: 119-133.

97. Tu, C.P.D., Toung, Y.P., Hsieh, T., Simkovich, N.M., and Tu, Y.S.L. Glutathione-transferase and Drug resistance. London: Taylor and Francis; 1990. p. 379-386.

98. Motoyama, N., Rock, G.C. and Dauterman, W.C. Studies on the mechanism of azinphosmethyl resistance in the predaceous mite, *Neoseiulus (T.) fallacis*. Pestic Biochem Physiol 1971; 1: 205-215.

99. Oppenoorth, F.J., Rupes, V., Elbashir, S., Houxn, W.H. and Voerman, S. Glutathione-dependent degradation of parathion and its significance for resistance in the housefly. Pestic Biochem Physiol 1972; 2: 262-269.

100. Motoyama, N. and Dauterman, W.C. Purification and properties of housefly GST. Insect Biochem 1977; 7: 361-364.

101. Beall, C., Fyrberg, C., Song, S. and Fyrberg, E. Isolation of a *Drosophila* gene encoding glutathione S-transferase. Biochem Genet 1992; 30: 515-527.

102. Fournier, D., Bride, J.M., Poire, M., Berge, J.B. and Plapp, F.W. Insect glutathione S-transferases. Biochemical characteristics of the major forms from houseflies susceptible and resistant to insecticides. J Biol Chem 1992; 267: 1840-1845.

103. Reiss, R.A. and James, A.A. A glutathione S-transferase gene of the vector mosquito, *Anopheles gambiae*. Insect Mol Biol 1993; 211: 421-425.

104. Toung, Y.P.S., Hsieh, T. and Tu, C.P.D. The glutathione S-transferase D genes. A divergently organized, intronless gene family in *Drosophila melanogaster.* J Biol Chem 1993; 268: 9737-9746.

105. Zhou, Z.H. and Syvanen, M. A complex glutathione transferase gene family in the housfly *Musca domestica.* Mol Gen Gent 1997; 256: 187-194.

106. Prapanthadara, L., Ranson, H., Somboon, P. and Hemingway, J. Cloning, expression and characterization of an insect class I glutathione S-transferase from *Anopheles dirus* species B. Insect Biochemistry and Molecular Biology 1998; 28: 321-329.

107. Ranson, H., Collins, F. and Hemingway J. The role of alternative mRNA splicing in generating heterogeneity within the *Anopheles gambiae* class I

glutathione S-transferase family. Proc Natl Acad USA 1998; 95: 14284-14289.

108. Ranson, H., Prapanthadara, L. and Hemingway, J. Cloning and characterization of two glutathione S-transferases from a DDT-resistant strain of *Anopheles gambiae*. Biochem J 1997; 324: 97-102.

109. Prapanthadara, L., Hemingway, J. and Ketterman, A.J. DDT-resistance in *Anopheles gambiae* (Diptera: *Culicidae*) from Zanzibar, Tanzania, based on increased DDT-dehydrochlorinase activity of glutathione S-transferases. Bull Entomol Res 1995; 85: 267-274.

110. Prapanthadara, L., Koottathep, L., Promtet, N., Hemingway, J. and Ketterman, A.J. Purification and characterization of a major glutathione S-transferase from the mosquito *Anopheles dirus* (Species B). Insect Biochem Mol Biol 1996; 26: 277-285.

111. Ranson, H., Cornel, A.J., Fournier, D., Vaughan, A., Collins, F.H. and Hemingway, J., Cloning and localization of a glutathione S-transferases class I gene from *Anopheles gambiae*. J Biol Chem 1997; 272: 5464-5468.

112. Pongjaroenkit, S., Jirajaroenrat, K., Boonchauy, C., Chanama, U., Leetachewa, S., Prapanthadara, L. and Ketterman, A.J. Genomic organization and putative promotors of highly conserved glutathione S-transferases originating by alternative splicing in *Anopheles dirus*. 2000; (in press).

113. Baimai, V. In The third conference on malaria researching, Thailand. 18-20 October: 1989: 146-152.

114. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976; 72: 248-254.

115. Dixon, M., Webb, E.C., Thorne, C.J.R., and Tipton, K.F. Enzyme inhibition and activation. In: Enzymes. 3rd ed. London: Longman Group Limited; 1979. p. 332-467.

116. Snustard, D.P., Simmons, M.J. and Jenkins, J.B. Chapter 13: Mutation, DNA repair, and Recombination. In Harris, D., editor. Principle of genetics. USA: John Wiley & Sons, Inc.; 1997. p. 325-326.

117. Ketterman, A.J., Prommeenate, P., Boonchauy, C., Chanama, U., Leetachewa, S., Promtet, N. and Prapanthadara, L. Single amino acid changes outside the active site significantly affect activity of glutathione S-transferases. 2000; (in press).

118. Sternberg, G., Board, P.G., Carlberg, I. and Mannervik, B Effects of directed mutagenesis on conserved arginine residues in a human class alpha glutathione S-transferase. 1991; 274: 549-555.

119. Phillips, M.F. and Mantle, T.J. The initial-rate kinetics of mouse glutathione S-transferase YfYf. Evidence for an allosteric site for ethacrynic acid. Biochem J 1991; 275: 703-709.

120. Phillips, M.F. and Mantle, T.J. Inactivation of mouse liver glutathione S-transferase YfYf (Pi class) by ethacrynic acid and 5,5'-dithiobis-(2-nitrobenzoic acid). Biochem J 1993; 294: 57-62.

121. Brown, A.W.A. Insecticide resistance in mosquitoes: a pragmatic review. J Am Mosq Control Assoc 1986; 2: 123-140.

122. Hemingway, J., Miyamoto, J. and Herath, P.R.J. A possible novel link between organophosphorus and DDT insecticide resistance genes in Anopheles: Supporting evidence from fenitrothion metabolism studies. Pestic Biochem Physiol 1991; 39: 49-56.

123. Tang, A.H and Tu, C.-P.D. Biochemical characterization of *Drosophila* glutathione S-transferase D1 and D21. J Biol Chem 1994; 269: 27876-27884.

124. Lopez, A.J. Alternative splicing of pre-mRNA: Developmental consequences and mechanisms of regulation. Annu Rev Gent 1998; 32: 279-305.

APPENDIX I

A fragment of the 8A.2 clone which contains the nucleotide sequence of *adgst1AS1* gene (accession no. AF251478). A box indicates each exon.

CGCATCTCGTTTGATCGCGTGCGCACAGGGACACACCAATACACCAAAAACTCCGCCAACCGGCAGCGCGTCGTCGACCGTTTATAT

Exon 1

CGCGGCCGGCACCCGAGCGGATCGCTTACTATTGTCTCGCTGGTTGAAGATCGTAGTCGTCGTCGCAGCAGCCCGTGTGCGTGTGTC

CAACCGTTTTGTCAGTGCTTGCTTAGTGTGTGGTTTTACGTGTGTGTATTACGGTGAGTGTTGGGAAGCGAGAATCGAAATTCGCCA

GAGAAACTGCAGTTGGTGGTTGAACCGTGCGTTTTAATTAACCGACGTTAATTGTTATCTTGTTCGCTTCCACTCTCCATCCTCGCA

Exon 2

GAATCATGGATTTCTACTACCTTCCCGGATCTGCGCCGTGCCGTGCCGTTCAGATGACGGCGGCCGCCGTTGGCGTCGAGCTGAACC

TGAAGCTCACCAACCTGATGGCGGGCGAGCACATGAAGCCGGAATTCCTGAAGGTAAGTGCGGCGACAGTTTCTGCCATTTTCTGTC

CTCCTCCGGGAGGGCTTGCATCGGCCATCACATAACCTTTGTGCCTGTGTGTGTCTGCTCGCGTGTGTTTGCCGCTGTGTGCCCTGT

GCTGTGTTCGTCACGATGGGAAATTTTCCATCCAGCATGCCCTCCAACCGTGTGTCTGTGTGTGCGTCCATTTTCCCCACGGCCAAA

GTGTGCCAAGAAAAATCAATTTCACTCCCATCTCCCATTTCCTTCGTATGTTTGTGTCGCGCTTTTCGAAACGGACGTCAAAATGGG

GAACACGAAGAACACGAAATTCGTGGGTCGTAAAAGAGAATGGCCGTTAGAGTGACGCCGCGGTCGAAAGACGTTTTTGTAAGTGGC

AGTTGTTGCTATGCCAACCGGGGGCCCCCACGCCACAATGTAACGCCATCTTGGATCGGGATAAATGTGCCTCACCTATCTCACCTC

CATCGTCGCACCTTGCCATCCGCCATTGCCCGTTTAAGGAATGTACCTGTGTGCAAAAAAAAAAAGAATTTAAATGCAAATGCATGCA

TCAACCAAGCCGGCAGGGATTGCCACTGAGAGATTGTCGCTGCTTGTTTGAGGACGTTGATTGCGTTGTTGTACGTTTTGTCGTCAA

CCGTCAATCGGCAAATGAAAGCGGGTCGCTCACTCATAAAATCATGCTCGTTTGTGCTGTTTAACGAACACTTGCATACGCCTTAAA

Exon 3A

CAAACCACCCCCCCAACTTTAAAGCTAAACCCGCAACACTGCATCCCGACGCTGGTCGACGAGGACGGGTTTGTGCTGTGGGAGTCG

CGCGCCATCCAGATCTACCTGGTCGAGAAGTACGGCGCCCACGATGCGGACCTGGCCGAGCGGCTTTACCCAAGCGATCCGCGTCGC

CGTGCCGTCGTCCACCAGCGGCTGTTCTTCGATGTGGCCGTCCTGTACCAGCGCTTCGCCGAGTACTACTATCCGCAGATCTTCGGC

CAGAAGGTGCCGGTGGGTGATCCGGGCCGGTTGCGCTCGATGGAGCAGGCACTCGAGTTCCTCAACACGTTCCTCGAGGGCGAGCAG

TACGTTGCCGGTGGTGATGATCCGACGATCGCCGATCTGAGCATACTGGCGACGATCGCCACGTACGAGGTGGCCGGCTACGATCTG

CGCCGGTACGAGAACGTTCAGCGGTGGTACGAGCGGACGAGCGCGATCGTCCCCGGGGCTGATAAGAACGTGGAGGGAGCTAAGGTT

TTCGGGCGCTACTTCACACAAAAATGATACGCCCGTCGTCTTATCGCGCTTATGTTTGTTGTCTATGTTAAGTGCAATACCCAGCTT

CCGCTCAGTCTGGCTCAGGCGAATGTACTCCGTCTCTTTTGATGTTTCTGTAGCGTAGCGTCGTACATGTTTGTCTGTTTTTCGTTG

TTTGATGTGTAATACAAAGTTTGTGGCTTTCTAAATTTCTACAATCATTAACACAAAACTCCTATTGCCGGCTTGTGCTTTCTTCTC

TGTACCTCCAACGCCTGCGGCTCGTACGTTCCGCTGCTAACCCGTTGTTGTTCCCCGTTCTTCCGAAACCCCAACCCCAACCGCCCC

Exon 3B

CTCCCTCAGATTAACCCCCAACACTGTATTCCGACGCTGGTCGACAATGGCTTTGCGCTGTGGGAGTCGCGCGCCATCTGCACGTAC

TTGGCGGAGAAGTACGGCAAGGACGACAAGCTGTACCCGAAGGACCCGCAGAAGCGCGCCGTCGTCAACCAGCGACTGTACTTCGAC

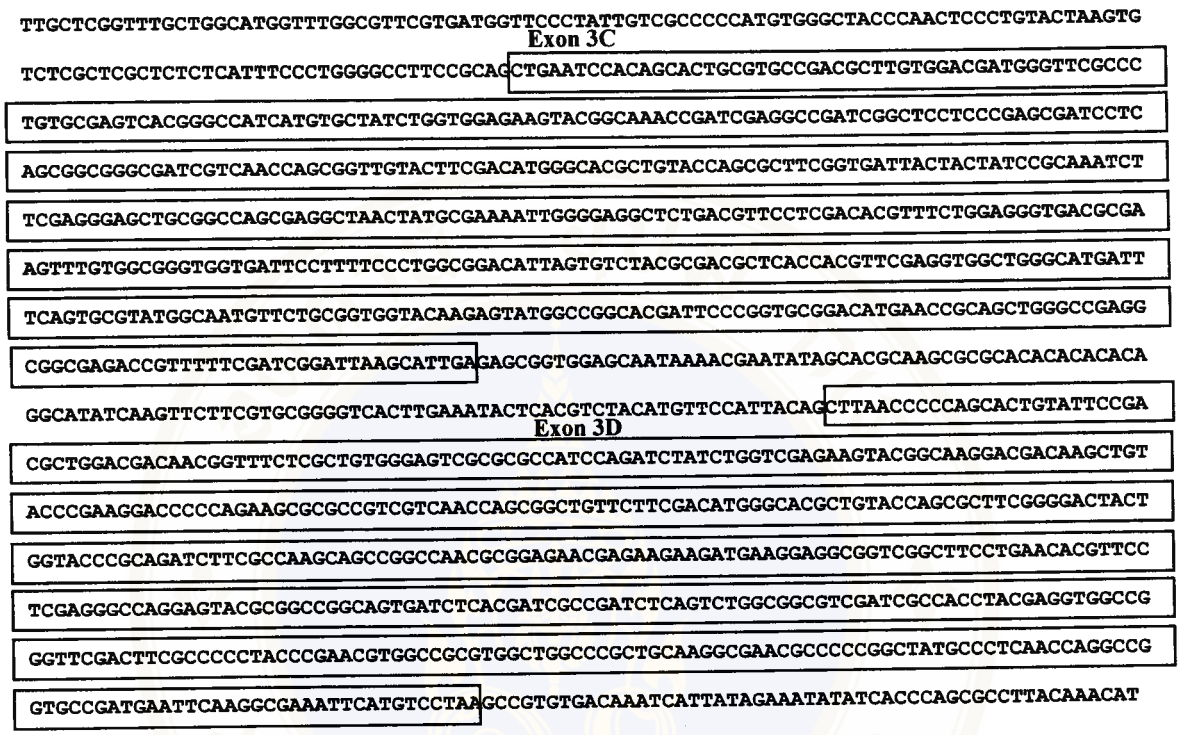
ATGGGCACGCTGTACCAGCGCTTTGCCGACTACTACTACCCGCAGATCTTCGCCAAGCAGCCGGCCAACGCGGAGAACGAGAAGAAG

ATGAAGGATGCGGTCGACTTCCTGAACACCTTCCTGGACGGGCACAAGTACGTGGCGGGCGACAGTCTGACGATCGCGGATCTGACC

GTGCTGGCCACGGTTTCGACGTACGACGTGGCGGGCTTCGAGCTGGCCAAGTACCCGCACGTGGCGGCGTGGTACGAGCGCACCCGC

AAGGAAGCGCCCGGTGCCGCCATCAACGAGGCCGGCATCGAGGAGTTCAGGAAGTACTTCGAGAAGTGAGCCGGGCGGAGTCAGCTG

TGATGTGGCAATAAAAGCTCGGGCAGTTTTGTTGATGTTTGTTCCTGTTTAACCGCCTCCCAGCGAGAGTGTTCCCGTGTGTTTTAC



APPENDIX II

Amino acid sequence comparison between *An. dirus* GSTs and *An. gambiae* GSTs

```
AdGST1-1 x AgGST1-6

   1 MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTNLMAGEHMKPEFLKLNPQHCIPTLDDNGFSLWESRAIQIYLVEKYGKDDKLYPKDPQKRAVVNQRLFFD 100
     ||||||||||||||||||||||||||||||||.|| |||||||||||||||||||| ||||.|||||||||| |||||||||||||||||||||||:||
   1 MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTDLMKGEHMKPEFLKLNPQHCIPTLVDNGFALWESRAIQIYLAEKYGKDDKLYPKDPQKRAVVNQRLYFD 100

 101 MGTLYQRFGDYWYPQIFAKQPANAENEKKMKEAVGFLNTFLEGQEYAAGSDLTIADLSLAASIATYEVAGFDFAPYPNVAAWLARCKANAPGYALNQAGA 200
     |||||||| || |||||||||| |||||||:||||||||||||||||.||||||||||.|||||||||||||||||| ||||||||||||||||
 101 MGTLYQRFADYHYPQIFAKQPANPENEKKMKDAVGFLNTFLEGQEYAAGNDLTIADLSLAATIATYEVAGFDFAPYPNVAAWFARCKANAPGYALNQAGA 200

 201 DEFKAKFMS 209
     |||||||:|
 201 DEFKAKFLS 209



AdGST1-2 x AgGST1-4

   1 MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTNLMAGEHMKPEFLKLNPQHCVPTLVDDGFALCESRAIMCYLVEKYGKPIEADRLLPSDPQRRAIVNQRL 100
     ||||||||||||||||||||||||||||||.|| |||||||||||||||||||| |||| ||||||||||||||||   | | |.|||:||||||||
   1 MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTDLMKGEHMKPEFLKLNPQHCVPTLVDSGFALWESRAIMCYLVEKYGKPCNNDSLYPTDPQKRAIVNQRL 100

 101 YFDMGTLYQRFGDYYYPQIFEGAAASEANYAKIGEALTFLDTFLEGDAKFVAGGDSFSLADISVYATLTTFEVAGHDFSAYGNVLRWYKSMAGTIPGADM 200
     |||||||||||||||||||||| |.|||:||||||| |||||||| :|||||. :||||||.||||||||||:||||| |||||||||   || .|
 101 YFDMGTLYQRFGDYYYPQIFEGAPANEANFAKIGEALAFLDTFLEGE.RFVAGGNGYSLADISLYATLTTFEVAGYDFSAYVNVLRWYKSMPELIPASDT 199

 201 NRSWAEAARPFFDRIKH 217
     |||||||||||||::||
 200 NRSWAEAARPFFDKVKH 216
```

AdGST1-3 x AgGST1-5

```
   1 MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTNLMAGEHMKPEFLKINPQHCIPTLVDNGFALWESRAICTYLAEKYGKDDKLYPKDPQKRAVVNQRLYFD 100
     ||||||||||||||||||||||||||||||.|| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
   1 MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTDLMKGEHMKPEFLKINPQHCIPTLVDNGFALWESRAICTYLAEKYGKDDKLYPKDPQKRAVVNQRMYFD 100

 101 MGTLYQRFADYYYPQIFAKQPANAENEKKMKDAVDFLNTFLDGHKYVAGDSLTIADLTVLATVSTYDVAGFELAKYPHVAAWYERTRKEAPGAAINEAGI 200
     |||||||||||||||||||||| |||.|||||| |||.||||||||||||||||||.:|||:||||||||:|||| ||||||| ||||||||||:|||
 101 MGTLYQRFADYYYPQIFAKQPANPENEQKMKDAVGFLNSFLDGHKYVAGDSLTIADLSILATISTYDVAGFDLAKYQHVAAWYENIRKEAPGAAINQAGI 200

 201 EEFRKYFEK 209
     |||:|||||
 201 EEFKKYFEK 209
```

AdGST1-4 x AgGST1-3

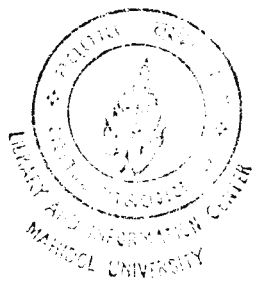
```
   1 MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTNLMAGEHMKPEFLKLNPQHCIPTLVDEDGFVLWESRAIQIYLVEKYGAHDADLAERLYPSDPRRRAVVH 100
     ||||||||||||||||||||||||||||||.|| |||||||||||||||||||||||||||||||||||||||||| |||  ||||||| |||||||||
   1 MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTDLMKGEHMKPEFLKLNPQHCIPTLVDEDGFVLWESRAIQIYLVEKYCAHDPALAERLYPGDPRRRAVVH 100

 101 QRLFFDVAVLYQRFAEYYYPQIFGQKVPVGDPGRLRSMEQALEFLNTFLEGEQYVAGGDDPTIADLSILATIATYEVAGYDLRRYENVQRWYERTSAIVP 200
     ||||||||:|||||||||||||||.||  ||| ||||||||||||||||||.:|||||||||| |||||.|||:: |||||||||: ||||.|  | |
 101 QRLFFDVAILYQRFAEYYYPQIFGKKV.AGDPDRLRSMEQALEFLNTFLEGERFVAGGDDPTIADFSILASIATFDAAGYDLRRYENIHRWYEQTGNIAP 199

 201 GADKNVEGAKVFGRYFTQK 219
      ||||. |||:|| || ||
 200 AADKNLAGAKIFGLYFRQK 218
```

APPENDIX III

The complete nucleotide sequence of *Anopheles dirus* GST 1-1 clone 10 (accession no. AF273041). Underline indicates the restriction endonuclease site. * indicates stop codon.

```
1     ATG GAT TTT TAT TAC CTA CCC GGA TCT GCG CCG TGC CGT GCC GTT   45
1      M   D   F   Y   Y   L   P   G   S   A   P   C   R   A   V   15

46    CAG ATG ACG GCG GCC GCC GTT GGC GTC GAG CTG AAC CTG AAG CTC   90
16     Q   M   T   A   A   A   V   G   V   E   L   N   L   K   L   30
                                                 EcoRI
91    ACC AAC CTG ATG GCG GGC GAG CAC ATG AAG CCG GAA TTC CTG AAG   135
31     T   N   L   M   A   G   E   H   M   K   P   E   F   L   K   45

136   CTT AAC CCC CAG CAC TGC ATT CCG ACG CTG GAC GAC AAC GGT TTC   180
46     L   N   P   Q   H   C   I   P   T   L   D   D   N   G   F   60

181   TCG CTG TGG GAG TCG CGC GCC ATC CAG ATA TAT CTG GTC GAG AAG   225
61     S   L   W   E   S   R   A   I   Q   I   Y   L   V   E   K   75

226   TAC GGC AAG GAC GAC AAG CTG TAC CCG AAG GAC CCC CAG AAG CGC   270
76     Y   G   K   D   D   K   L   Y   P   K   D   P   Q   K   R   90

271   GCC GTC GTC AAC CAG CGG CTG TTC TTC GAC ATG GGC ACG CTG TAC   315
91     A   V   V   N   Q   R   L   F   F   D   M   G   T   L   Y   105

316   CAG CGC TTC GGG GAC TAC TGG TAC CCG CAG ATC TTC GCC AAG CAG   360
106    Q   R   F   G   D   Y   W   Y   P   Q   I   F   A   K   Q   120

361   CCG GCC AAC GCG GAG AAC GAG AAG AAG ATG AAG GAG GCG GTC GGC   405
121    P   A   N   A   E   N   E   K   K   M   K   E   A   V   G   135

406   TTC CTG AAC ACG TTC CTC GAG GGC CAG GAG TAC GCG GCC GGC AGT   450
136    F   L   N   T   F   L   E   G   Q   E   Y   A   A   G   S   150

451   GAT CTC ACG ATC GCC GAT CTC AGT CTG GCG GCG TCG ATC GCC ACC   495
151    D   L   T   I   A   D   L   S   L   A   A   S   I   A   T   165

496   TAC GAG GTG GCC GGG TTC GAC TTC GCC CCC TAC CCG AAC GTG GCC   540
166    Y   E   V   A   G   F   D   F   A   P   Y   P   N   V   A   180

541   GCG TGG CTG GCC CGC TGC AAG GCG AAC GCC CCC GGC TAT GCC CTC   585
181    A   W   L   A   R   C   K   A   N   A   P   G   Y   A   L   195

586   AAC CAG GCC GGT GCC GAT GAA TTC AAG GCG AAA TTC ATG TCC TAA   630
196    N   Q   A   G   A   D   E   F   K   A   K   F   M   S   *
```

BIOGRAPHY



NAME	Miss Kanya Jirajaroenrat
DATE OF BIRTH	21 April 1972
PLACE OF BIRTH	Chonburi, Thailand
INSTITUTIONS ATTENDED	Mahidol University, 1990-1993: Bachelor of Science (Medical Technology)
	Mahidol University, 1998-2000: Master of Science (Molecular Genetics and Genetic Engineering)
FELLOWSHIP/ RESEARCH GRANT	Graduate Fellowship Program from National Science and Technology Development Agency (NSTDA), 1998-2000.