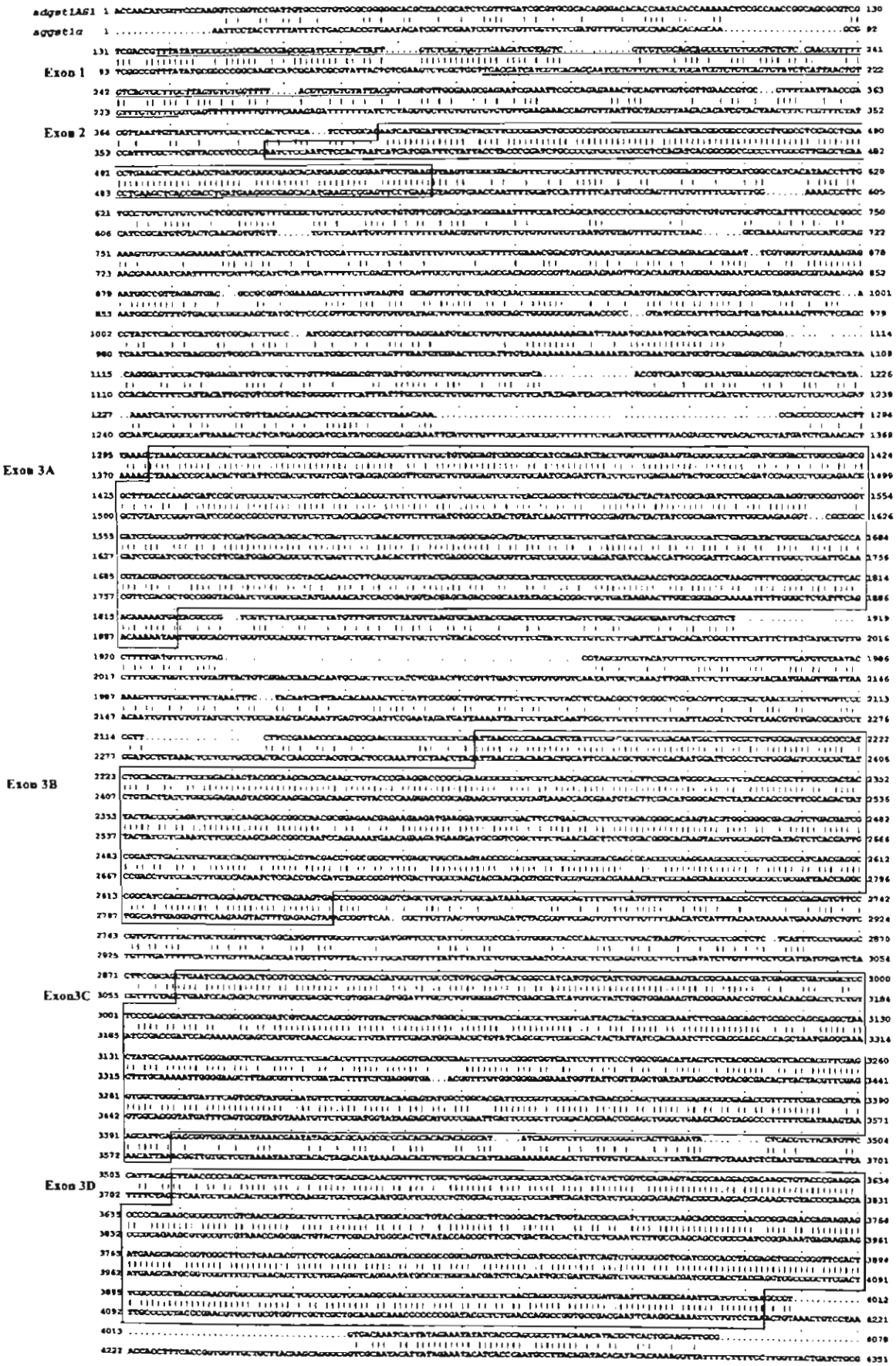
adgstLAS1
aggstla
Exon 1
Exon 2
Exon 3A
Exon 3B
Exon3C
Exon 3D

Fig. 3. A comparison of deduced amino acid sequences of *An. dirus* with their orthologous *An. gambiae* GSTs by using Gap in the GCG program package. The identities of each comparison are 93%, 85%, 92% and 86% for AdGST1-1, AdGST1-2, AdGST1-3 and AdGST1-4, respectively.

AdGST1-4. The AdGST1-1 and AdGST1-3 have the same number of amino acids as their orthologous gene products (AgGST1-6 and AgGST1-5, respectively), which are 209 amino acids, whereas AdGST1-2 and AdGST1-4 (217 and 219 amino acids) are one amino acid longer than AgGST1-4 and AgGST1-3.

3.3. Promoter prediction

In the present report a combination of two programs were used, TSSW (Smith et al., 1996) (human RNA polymerase II recognition using the TRANSFAC database) and the MatInspector program (Quandt et al., 1995) (transcription factor binding sites identified using TRANSFAC database with user defined parameters). Two putative promoters were identified within 2353 bp upstream of *adgst1AS1* by TSSW and have been named Ad214 (−1868 to −2074) and Ad2112 (+32 to −268). The potential transcription factor binding site of these two regions are shown in Fig. 4.

MatInspector results show the Ad214 region contains two hunchback binding sites, two dorsal morphogen (dl) binding sites, two Broad-Complex 4 (BR-CZ4) binding sites and three fork head domain protein (Croc, HFH3 and FREAC7) binding sites (Fig. 4). There is also one binding site for C/EBP. There are two of Homeotic selector (HOM) gene product binding sites that are *Deformed* (*Dfd*) and GATA. The Ad2112 region contains two *Dfd* recognition sites, three of tissue-specific transcription factor binding sites, two GATA binding sites and one NF1. There are two steroid inducible elements (BR-CZ4 and C-Ets). There also are three regulatory factors that could regulate xenobiotic metabolism, one AP-1 and two Arnt (Ah receptor nuclear translocator) recognition sites and three of transcriptional activator binding sites (C-Myb, V-Myb and HB). Nested 5′RACE sequence showed the transcription start site and the presence of a 122-nucleotide intron in the 5′untranslated region (Fig. 2). The transcription start site, which is located near Ad2112 suggests that Ad2112 regulates *adgst1AS1* at least in 4th instar larvae. However, this does not preclude the function of this promoter in other developmental stages.

Fig. 2. A comparison of *adgstAS1* nucleotide sequence with the nucleotide sequence of *aggst1α* (AF 071160) by using Gap in the GCG program Package. The values in the margins indicate nucleotide base length. Exon 2 and Exon 3A–3D of both GST genes are boxed. Exon 1, the 5′UTR, is shown in italics and underlined in both genes. Exon 2 codes for the N-terminus and exon 3A–3D code for different C-termini.

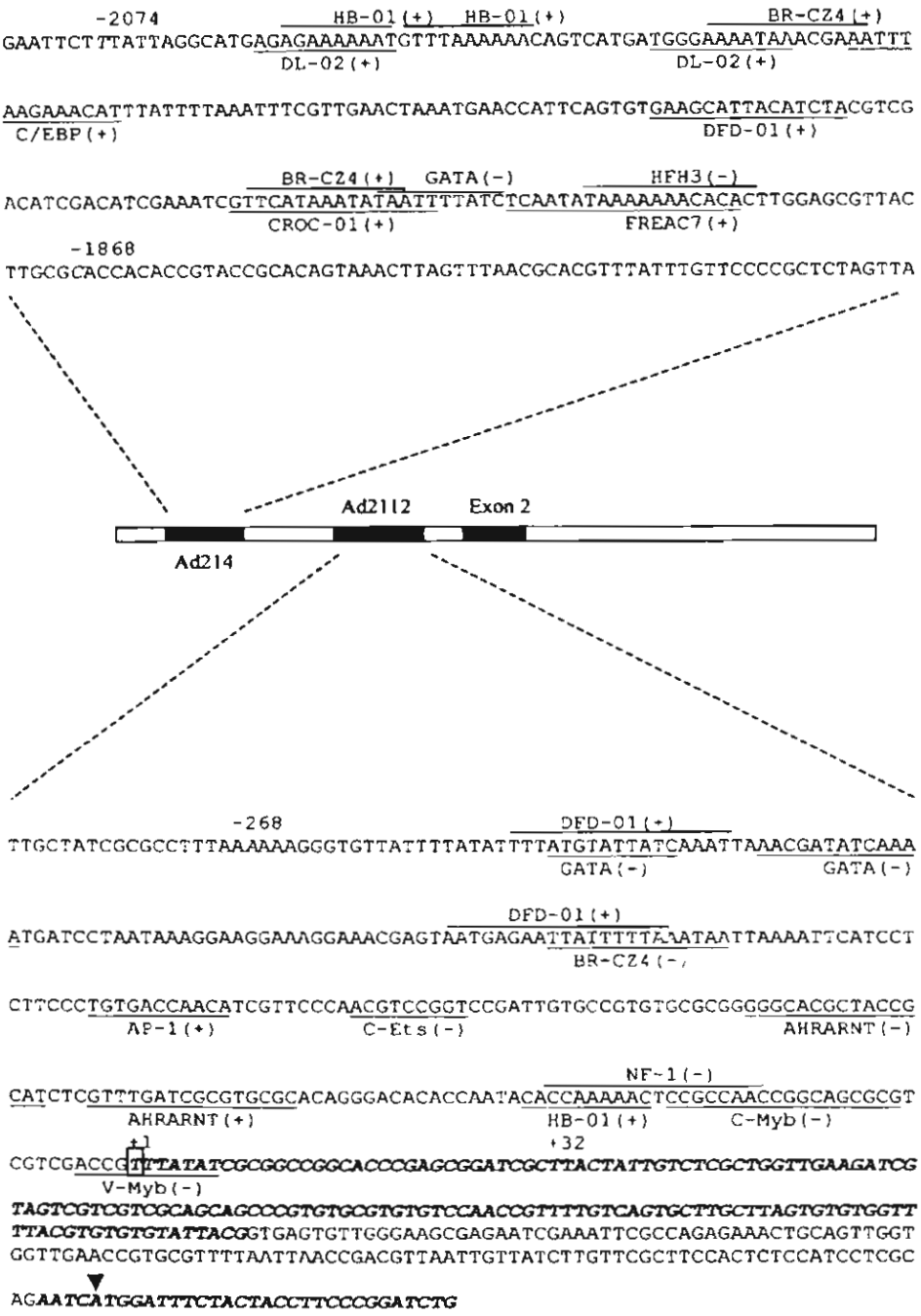

Fig. 4. Sequence of 5'-flanking region of *adgst1AS1*. Each putative promoter sequence is shown relative to exon 2 in the schematic. The two putative promoters (Ad214 and Ad2112) were searched for potential transcription factor binding sites by using MatInspector program at 0.9 for matrix similarity and 1 for core similarity. Some abbreviations of binding sites are: HB-01, hunchback; DL-02, dorsal morphogen; BR-CZ4, broad-complex 4; HFH3, hepatocyte nuclear factor 3; CROC-01, crocodile; FREAC7, fork related activator 7; DLD-01, deformed; AP-1, activator protein 1; AHRARNT, Ah receptor nuclear translocator. The potential transcription factor binding sites are underlined or overlined. The (+) indicates the binding site on the sense strand and (−) indicates the binding site on antisense strand. The *italic* and bold letters represent the transcript sequences and the transcription start site is boxed. The arrowhead indicates translation start site of *adgst1AS1*.

3.4. Identification of adgst1-5 an aggst1β orthologous gene

Two other positive fragments (5.6 and 6 kb fragment from *SalI* digestion of 5A.1 and 8A.2, respectively, as shown in Fig. 1) were detected with low stringency hybridization of DIG-labeled probe (200 bp). A restriction map and southern hybridization of 5A.1 and 8A.2 recombinant phage clones revealed the location of these positive fragments to be downstream of *adgst1AS1* (Fig. 1). These fragments were subcloned in pBluescript II SK$^+$ and digested with several restriction enzymes. A

positive signal was detected from *PvuII* digestion of both *SalI* fragments in pBluescript II SK$^+$ with the fragments about 900 bp in size. These *PvuII* fragments were sequenced and the amino acid translation of the *PvuII* fragment is shown in Fig. 5. Comparison of the *An. dirus* sequence with the reported *An. gambiae* nucleotide sequence showed a high identity to the *aggst1-7* sequence, 177 from 200 bp or 88.5%. Amino acid comparison of these sequences showed 89% identity and 94% similarity. Therefore, this GST coding region was identified as the 5′sequence of *adgst1-5*. The predicted splice site in the genomic sequence is shown in Table 1 and the difference in the predicted splice sites of *adgst1-5* and *aggst1β* are shown in Fig. 5.

4. Discussion

4.1. Genomic organization of the GSTs

One complete class I GST gene was identified from an *An. dirus* genomic library. A comparison with GST sequences from other species shows the highest identity (>78.5%) to *aggst1α* (class I GST from *An. gambiae*) and <30% similarity to insect GST class II. Therefore this gene is a class I GST. The same arrangement and

```
CTGTTGTTATCGTTCTCGGTCATTAATTTAATCCTCGCATGGTTCGGCTGTCTCGAACTC
CGCTGCCCCTGCCACCTGACTCACGAATCGGGCCCCATGCCCCCCATGAGTTTGTCTTAT
GCCACCGTGGCCAGCACTGGAAACGGGAGCCCACTCCGGGGACACGGACTGTAATTAGTT
CAACCGTGCTAATGATTGTTGTGATTCCGCTTTGCAGGCGCGGTCCGATGACGACGGTGC
  T  V  L  M  I  V  V  I  P  L  C  R  R  G  P  M  T  T  V  L
TGTACTATCTGCCGGCATCGCCACCGTGCCGATCGGTGCTGCTGCTGGCCAAGATGATCG
  Y  Y  L  P  A  S  P  P  C  R  S  V  L  L  L  A  K  M  I  G
GCGTGGAGCTGGATCTAAAGGTCCTGAACATTATGGAGGGCGAACAGCTAAAGCCGGACT
  V  E  L  D  L  K  V  L  N  I  M  E  G  E  Q  L  K  P  D  F
TCGTCGAGCTGAATCCGCAGCACTGCATACCAACGATGGACGATCACGGGCTGGTGCTGT
  V  E  L  N  P  Q  H  C  I  P  T  M  D  D  H  G  L  V  L  W
GGGAGAGGTAAGCAATGCCACCTATCGATTTTCCCATCCGTCGTTGCCCCTTCCGTGTAC
  E  R  *  A  M  P  P
TGTTGTGTGCGAGGGCGCCTTCAAAAGGGCGAGCTTAAACTTAAACCTTGGCCCCGTCCA
AGGGCTGTGCGTGATGTGTTTTTCTTATCATCATTTCGATTGCCGTACCGTGCGAAGAAG
GTTACGAGTGTGCGTGTGCGTTTCTTTTTCCTTTCATTTCGGGGTTTCGGATCGTTTCAT
TTTCCAGGCCGTGCGTTTATTGGATTTTCACCTACCCACACGTCTTGTGGTCCGACCGGA
TCGGACCGGTGATCTATTTTGGGGAGGGGTCGCGCACGTGCTACGAGGACGTAAGCGGAT
GTTGCGAGCGTGGGAAAGTGGCACACCGGGCACGAAATCAATGACACAATTTGGGAGACG
CGGCGCTGGTTCGGGGAGCACATTTGAGCTCGCGTCGAAAGCGCCAG
```

Fig. 5. The nucleotide and translated sequence of the 888 bp *PvuII* fragment, which is the *aggst1β* orthologous gene. The underlined sequences are the amino acid sequences with homology to *aggst1-7* (deduced amino acid sequence of *aggst1β*). The seven bold letters show the amino acids different from *aggst1-7*. The arrow is at the predicted splice site in *aggst1-7* and the arrowhead is at the predicted splice site in the 888 bp *PvuII* fragment.

the high identity of each exon and mRNA product suggest that the *adgst1AS1* is the orthologous gene of *aggst1α* (Figs. 2 and 3) (Ranson et al., 1998). The partial sequence of *adgst1-5* (the orthologous gene of *aggst1β*) also was identified (Fig. 5) therefore *adgst1-5* is still an *adgst1AS1* neighbor as *aggst1β* and *aggst1α*. Southern hybridization and sequence analysis (~4 kb upstream of *adgst1AS1*) does not show an *An. dirus* gene orthologous to *aggst1-2*, an intronless *An. gambiae* GST gene. The intronless GST gene appears to be absent from the region around the *adgst1AS1* gene. In addition the intron sizes of *adgst1AS1* (122, 724, 330, 231, and 114 bp for intron 1 to intron 5, respectively) and *aggst1α* (146, 837, 442, 230, and 130 bp for intron 1 to intron 5, respectively) differ as well as possessing a lower identity than the coding sequence. A comparison of the two genes nucleotide sequence for identity shows 27% for intron 1, 36% for intron 2, 24% for intron 3, 38% for intron 4, and 30% for intron 5. A reported phylogenic study was performed in *Anopheles* subgenus *Cellia* including *Anopheles (Cellia) leucosphyrus* Dönitz and *An. gambiae* complex using a cladistic approach based on morphological characteristics (Anthony et al., 1999). The *An. dirus* complex belongs to the *Anopheles (Cellia) leucosph. us* Dönitz group which is an Oriental species and the *An. gambiae* complex is an Afrotropical species. The data analysis reveals the Oriental species was established as a monophyletic group whereas the Afrotropical species as a paraphyletic group. These phylogenic results suggest the Oriental and Afrotropical groups have been diverged for a long time, possibly for more than several million years. However the high conservation of *adgst1AS1* and *aggst1α* in all coding sequences, splice sites and the gene organization suggest that these two GST genes in the two distantly related anopheline species have important physiological roles that have been maintained.

The splicing products of *adgst1AS1* use the same 5′exon (exon 2) which is spliced to one of four alternative 3′ exons (exon 3A–3D) to produce the mature transcripts (*adgst1-1, adgst1-2, adgst1-3* and *adgst1-4*) similar to *aggst1α*. The *adgst1-1* product (the splicing product of exon 1, 2 and exon 3D) has been characterized previously as a recombinant enzyme (Prapanthadara et al., 1998). The remaining splicing forms from *adgst1AS1* will be cloned and the recombinant proteins enzymatically characterized. The different C-termini of these GSTs may be involved in the determination of substrate specificity and these enzymes need to be investigated to elucidate their roles in insecticide resistance.

Genomic organizations of insect class I GST have been reported previously such as the eight intronless genes in *D. melanogaster* (Toung et al., 1993) or the fusion forms in *M. domestica* (Zhou and Syvanen, 1997). Currently, alternative splicing of pre-mRNA has been

shown in *An. gambiae* (Ranson et al., 1998) and also in this work for *An. dirus*. Alternative splicing generates functionally diverse isoforms for virtually every type of protein involved in metazoan development, cell function and physiology. Modular organization in the genome is a powerful and versatile regulatory mechanism, which can effect quantitative control of gene expression and functional diversification of proteins (Lopez, 1998). This concept also supports the idea that this GST gene must have an important physiological function.

4.2. Promoter prediction

The various genome projects have generated large amounts of sequence data, which is continually increasing. This immense amount of data has led to the development of computer programs to analyze and characterize these data. A major outcome is the appearance of eukaryotic polymerase II promoter prediction programs that identify transcription factor binding sites.

The Ad214 promoter contains several binding sites for fork head protein (Croc, HFH3 and FREAC7), HOM, dl and one of C/EBP (Fig. 4). The fork head proteins were suggested to have a central role in embryonic development (Häcker et al., 1995). *Dfd* is part of the development regulatory system of multicellular organisms through transcriptional modulation of specific sets of downstream genes (McGinnis and Krumlau, 1992; Rorth, 1994). The dl is the nuclear protein that determines the size and the fate of regions along the dorsal–ventral axis of the Drosophila embryo by setting the expression limits of key zygotic regulatory genes (Thisse et al., 1992). The C/EBP have been isolated from Drosophila and shown to play a role during development (late embrogenesis) (Rorth, 1994). Therefore putative Ad214 would appear to be active in the developmental stage (Häcker et al., 1995; Thisse et al., 1992; Rorth, 1994). As GSTs are multifunctional proteins then one question that is raised is: is the *adgst1AS1* regulated by the developmental transcription factor binding sites of the Ad214 promoter for an oxidative stress response activity? Since in the developmental stage of Drosophila, many oxidative stress response proteins are expressed such as Jun, Fos, JNK, JAK (www.Virtual Library — Developmental Biology). Recently a Pi class GST has been shown to interact with Jun N-terminal kinase and provide additional regulation of this important kinase (Adler et al., 1999). This observation suggests that there may be more physiological roles that GSTs perform remaining to be elucidated.

The *aggst1α* alternative splicing was proposed to be regulated in a tissue-dependent manner (Ranson et al., 1998). Therefore its orthologous gene (*adgst1AS1*) would be regulated in the same manner. There are several well-known tissue-specific elements in both of the putative promoters, Ad214 and Ad2112, including three GATA binding sites, one C/EBP binding site and one NF-1 binding site. It has been reported that NF-1 can activate the human elastin promoter in the Drosophila Schneider cell via its binding site and the elastin transcriptional level in different cell types may be regulated by cell specific expression of different NF-1 family members (Degterev and Foster, 1999). These data suggest that *adgst1AS1* may be regulated in a tissue-specific/cell-specific manner via these putative elements.

However, the expression of GSTs has been reported to be influenced by both intracellular and extracellular compounds. Endogenous compounds such as steroid hormones can regulate GST gene expression in rat, hamster and marine fish (Fan et al., 1992; Pulford and Hayes, 1996; Leaver et al., 1997). Several steroid inducible elements (BR-CZ4 and Ets) are identified in the putative promoter Ad2112. BR-C family and Ets related transcription factor (E74 in Drosophila) directly mediates a temporal and tissue-specific response to ecdysone as larvae become committed to metamorphosis (Bayer et al., 1996; Kalm et al., 1994). The vertebrate *myb* protooncogene product was reported to play an important role in cell cycle as well as activating transcription (Nishina et al., 1989). The *myb* gene identified from Drosophila also was shown to have a role in the cell division cycle like in vertebrates (Katzen et al., 1998). Therefore the *myb* product may act as a transcription activator similar to the vertebrate *myb*. The hunchback protein was shown to be a concentration-dependent activator of transcription in the cultured Drosophila cell (Zuo et al., 1991). Then the induction of *adgst1AS1* may be by C-Myb, V-Myb and HB. Moreover, the expression of many GST genes are induced with several xenobiotic compounds. For example, β-naphthoflavone induces GST Ya via ARE (2 AP-1 binding sites) and XRE (Rushmore et al., 1990) or a GST from marine fish via ARE (Leaver et al., 1997). The aflatoxin B_1-8,9-epoxide induced rat GST Yc_2 expression via ARE and barbie box (Pulford and Hayes, 1996). The AP-1 and the Ah receptor nuclear translocator (Arnt) have been reported capable of inducing the GST Ya expression (Rushmore et al., 1990) and these two elements also exist in the Ad2112 promoter. All the above data suggests that the regulation of *adgst1AS1* by Ad2112 occurs for intra- or extracellular compounds such as steroids through BR-C and Ets or xenobiotic compounds through AP-1 and Arnt. The 5′RACE shows that the Ad2112 probably controls *adgst1AS1* expression at least in the 4th instar larvae stage. However, Ad214 may be a distal promoter that will function in other stages such as the developmental stage, or be an enhancer since it is located 1599 bp upstream of Ad2112. One may speculate that the putative elements and the gene organization suggests that *adgst1AS1* is involved in oxidative stress response and expressed as a housekeeping gene. However, during development the increased oxidative stress in the cell requires an

increased GST expression. The increased GST expression may be regulated by the putative elements in promoter Ad214 that would be active during development.

In conclusion, a highly conserved alternatively spliced gene now found in two anopheles species was identified from *An. dirus*, a Thai malaria vector. The high conservation and type of gene organization suggests this gene is very important and has not changed during the course of evolution.

Acknowledgements

This work was supported by the Royal Golden Jubilee to S.P. and the Thailand Research Fund. We thank Dr Witoon Tirasophon and Mr Jongrak Kittiworakarn for their helpful comments.

References

Adler, V., Yin, Z., Fuchs, S.Y., Benezra, M., Rosario, L., Tew, K.D., Pincus, M.R., Sardana, M., Henderson, C.J., Wolf, C.R., Davis, R.J., Ronai, Z., 1999. Regulation of JNK signaling by GSTp. EMBO J. 18, 1321–1334.

Anthony, T.G., Harbach, R.E., Kitching, I.J., 1999. Phylogeny of the Pyretophorus Series of *Anopheles* subgenus *Cellia* (Diptera: Culicidae). Systematic Entomology 24, 193–205.

Baimai, V. 1989. Speciation and species complexes of the anopheles malaria vectors in Thailand. In: Proceedings of the Third Conference on Malaria Research, Thailand, pp. 146–152.

Bayer, C.A., Holly, B., Fristrom, J.W., 1996. A swith in broad-complex zinc-finger isoform expression is regulated posttranscriptionally during the metamorphosis of Drosophila imaginal discs. Dev. Biol. 177, 1–14.

Board, P.G., 1981. Biochemical genetics of glutathione S-transferase in man. Am. J. Hum. Genet. 33, 36–43.

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

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

Clark, A.G., Shamaan, N.A., Dauterman, W.C., Hayaoka, T., 1984. Characterization of multiple glutathione transferases from the house fly, *Musca domestica* (L). Pestic. Biochem. Physiol. 22, 51–59.

Degterev, A., Foster, J.A., 1999. The role of NF-1 factor in regulation of elastin gene transcription. Matrix Biol. 18, 295–307.

Fan, W., Trifiletti, R., Cooper, T., Norris, J.S., 1992. Cloning of a μ-class glutathione S-transferase gene and identification of the glucocorticoid regulatory domains in its 5′ flanking sequence. Proc. Natl. Acad. Sci. USA 89, 6104–6108.

Fournier, D., Bride, J.-M., Poirié, M., Bergé, J., Plapp, F.W., 1992. Insect glutathione S-transferases. Biochemical characteristics of the major forms from houseflies susceptible and resistant to insecticides. J. Biol. Chem. 267, 1840–1845.

Grant, D.F., Matsumura, F., 1989. Glutathione S-transferase 1 and 2 in susceptible and insecticide resistant *Aedes aegypti*. Pestic. Biochem. Physiol. 33, 132–143.

Häcker, U., Kaufmann, E., Hartmann, C., Jürgens, G., Knöchel, W., Jäckle, H., 1995. The drosophila fork head domain protein crocodile is required for the establishment of the head structures. EMBO J. 14, 5306–5317.

Hayes, J.D., Pulford, D.J., 1995. The glutathione S-transferase supergene family: Regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. CRC Crit. Rev. Biochem. Molec. Biol. 30, 445–600.

Kalm, L.V., Crossgrove, K., Seggern, D.V., Guild, G.M., Beckendorf, S.K., 1994. The broad-complex directly controls a tissue-specific response to the steroid hormone ecdysone at the onset of Drosophila metamorphosis. EMBO J. 13, 3505–3516.

Katzen, A.L., Jackson, J., Harmon, B.P., Fung, S.M., Ramsay, G., Bishop, J.M., 1998. Drosophila myb is required for the G2/M transition and maintenance of diploidy. Genes Dev. 12, 831–843.

Leaver, M.J., Wright, J., George, S.G., 1997. Structure and expression of a cluster of glutathione S-transferase genes from a marine fish, the plaice (*Pleuronectes platessa*). Biochem. J. 321, 405–412.

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

Lougarre, A., Bride, J.M., Fournier, D., 1999. Is the insect glutathione S-transferase I gene family intronless? Insect Molecular Biology 8, 141–143.

Mannervik, B., Danielson, U.H., 1988. Glutathione transferases — structure and catalytic activity. CRC Crit. Rev. Biochem. 23, 283–337.

McGinnis, W., Krumlau, R., 1992. Homeobox genes and axial patterning. Cell 68, 283–302.

Motoyama, N., Dauterman, W.C., 1978. Molecular weight, subunits, and multiple forms of glutathione S-transferase from the house fly. Insect Biochem. 8, 337–348.

Nishina, Y., Nakagoshi, H., Imamoto, F., Gonda, M.J., Ishii, S., 1989. Transactivation by the c-myb proto-oncogene. Nucleic Acids Res. 17, 107–117.

Pemble, S., Schroeder, K.R., Spencer, S.R., Meyer, D.J., Hallier, E., Bolt, H.M., Ketterer, B., Taylor, J.B., 1994. Human glutathione S-transferase Theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. Biochem. J. 300, 271–276.

Pickett, C.B., Lu, A.Y.H., 1989. Glutathione S-transferases: Gene structure, regulation, and biological function. Ann. Rev. Biochem. 58, 743–764.

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

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

Prapanthadara, L., Ranson, H., Somboon, P., Hemingway, J., 1998. Cloning, expression and characterization of an insect class I glutathione S-transferase form *Anopheles dirus* species B. Insect Biochem. Molec. Biol. 28, 321–329.

Pulford, D.J., Hayes, J.D., 1996. Characterization of the rat glutathione S-transferase Yc_2 subunit gene, GSTA5: identification of a putative antioxidant-responsive element in the 5′-flanking region of rat GSTA5 that may mediate chemoprotection against aflatoxin B_1. Biochem. J. 318, 75–84.

Quandt, K., Frech, K., Karas, H., Wingender, E., Werner, T., 1995. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 23, 4878–4881.

Ranson, H., Collins, F., Hemingway, J., 1998. The role of alternative mRNA splicing in generating heterogeneity within the *Anopheles gambiae* class I glutathione S-transferase family. Proc. Natl. Acad. Sci. USA 95, 14284–14289.

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

Rorth, P., 1994. Specifisation of C/EBP function during Drosophila development by the bZIP basic region. Science 266, 1878–1881.

Rushmore, T.H., King, R.G., Paulson, K.E., Pickett, C.B., 1990. Regulation of glutathione S-transferase Ya subunit gene expression: Identification of a unique xenobiotic-responsive element controlling inducible expression by planar aromatic compounds. Proc. Natl. Acad. Sci. USA 87, 3826–3830.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. In: Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press.

Smith, R., Wiese, B., Wojzynski, M., Davison, D., Worley, K., 1996. BCM Search Launcher — an integrated interface to molecular biology data base search and analysis services available on the World Wide Web (http://dot.imgen.bcm.tmc.edu:9331/seq-search/gene-search.html). Genome Res. 6, 454–462.

Thisse, C., Perrin-Schmitt, F., Stoetzel, C., Thisse, B., 1992. Sequence-specific transactivation of the Drosophila *twist* gene by the dorsal gene product. Cell 65, 1191–1201.

Toung, Y.-P.S., Hsieh, T.-S., Tu, C.-P.D., 1990. *Drosophila* glutathione S-transferase 1-1 shares a region of sequence homology with the maize glutathione S-transferase III. Proc. Natl. Acad. Sci. USA 87, 31–35.

Toung, Y.-P.S., Hsieh, T.-S., Tu, C.-P.D., 1991. The Drosophila glutathione S-transferase 1-1 is encoded by an intronless gene at 87B. Biochem. Biophys. Res. Comm. 178, 1205–1211.

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

Tsuchida, S., Sato, K., 1992. Glutathione transferases and cancer. CRC Crit. Rev. Biochem. Molec. Biol. 27, 337–384.

Vaughan, A., Rodriguez, M., Hemingway, J., 1995. The independent gene amplification of electrophoretically indistinguishable B esterases from the insecticide-resistant mosquito *Culex quinquefasciatus*. Biochem. J. 305, 651–658.

Widersten, M., Pearson, W.R., Engström, Å., Mannervik, B., 1991. Heterologous expression of the allelic variant Mu-class glutathione transferases μ and ψ. Biochem. J. 276, 519–524.

www.Virtual Library — Developmental Biology. http://sdb.bio.purdue.edu/Other/VL_DB.html

Zhou, Z., Syvanen, M., 1997. A complex glutathione transferase gene family in the housfly *Musca domestica*. Mol. Gen. Genet. 256, 187–194.

Zuo, P., Stanojevic, D., Colgan, J., Han, K., Levine, M., Manley, J.L., 1991. Activation and repression of transcription by the gap proteins hunchback and Kruppel in cultured Drosophila cells. Genes Dev. 5, 254–264.

Pergamon

Insect Biochemistry and Molecular Biology 31 (2001) 867–875

Insect Biochemistry and Molecular Biology

www.elsevier.com/locate/ibmb

Heterologous expression and characterization of alternatively spliced glutathione *S*-transferases from a single *Anopheles* gene

Kanya Jirajaroenrat [a], Saengtong Pongjaroenkit [a], Chartchai Krittanai [a], La-aied Prapanthadara [b], Albert J. Ketterman [a,*]

[a] *Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakorn Pathom 73170, Thailand*
[b] *Research Institute for Health Sciences, Chiang Mai, Thailand*

Received 18 October 2000; received in revised form 11 December 2000; accepted 15 December 2000

Abstract

Three cDNA sequences of glutathione *S*-transferase (GST), *adgst1–2*, *adgst1–3* and *adgst1–4*, which are alternatively spliced products of the *adgst1AS1* gene, were obtained from fourth instar larvae of *Anopheles dirus* mosquito by reverse transcriptase PCR reactions. The nucleotide sequences of these three cDNAs share >67% identity and the translated amino acid sequences share 61–64% identity. A comparison of the *An. dirus* to the *An. gambiae* enzymes shows that 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 orthologous isoenzymes occur across anopheline species. These three proteins were expressed at high levels, approximately 15–20 mg from 200 ml of *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 1-chloro-2,4-dinitrobenzene (CDNB), detectable activity with 1,2-dichloro-4-nitrobenzene but markedly low activity with ethacrynic acid and *p*-nitrophenethyl bromide. adGST1–3 was shown to be the most active enzyme from the kinetic studies. Permethrin inhibition of CDNB activity, at varying concentrations of CDNB, was significantly different, being uncompetitive for adGST1–2, noncompetitive for adGST1–3 and competitive for adGST1–4. In contrast, permethrin inhibition with varying glutathione concentrations was noncompetitive for all three GSTs. Despite the enzymes being splicing products of the same gene and sharing identical sequence in the N-terminal 45 amino acids, these GSTs show distinct substrate specificities, kinetic properties and inhibition properties modulated by the differences in the C-terminus.

Keywords: Glutathione *S*-transferase; *Anopheles dirus*; Mosquito; Isoenzymes; Alternatively spliced proteins

1. Introduction

Glutathione *S*-transferases (GSTs; EC 2.5.1.18) are a multiple gene family of multifunctional dimeric enzymes, which catalyze a broad range of substrates and play an important role in detoxication of xenobiotic compounds (Hayes and Pulford, 1995). GSTs catalyze the nucleophilic attack of glutathione on the electrophilic center of a range of endogenous hydrophobic molecules. On the basis of N-terminal amino acid sequence comparisons, insect GSTs were classified as belonging to the class Theta family (Hiratsuka et al., 1990; Pemble and Taylor, 1992). However, the multiple forms of GST in *Drosophila* can be classified into two groups belonging to distinct gene clusters. Sequence data for insect GSTs are limited, but there is evidence for at least two classes of insect GSTs; Class I and Class II (Beall et al., 1992; Fournier et al., 1992).

The sequences of two insect Class II or Sigma class GST genes have been published, *aggst2–1* from *Anopheles gambiae* (Reiss and James, 1993) and *DmGST2* from *Drosophila melanogaster* (Beall et al., 1992). Both *aggst2–1* and *DmGST2* are single copy genes with no closely related sequences present in the genome of either species. The other insect Class II GSTs are *MdGST2* from *Musca domestica* (Franciosa and Berge, 1995)and *MsGST2* from *Manduca sexta* (Snyder et al., 1995).

The insect Class I or Delta class GSTs, in contrast,

* Corresponding author. Fax: +66-2-441-9906.
E-mail address: frakt@mucc.mahidol.ac.th (A.J. Ketterman).

0965-1748/01/$ - see front matter
PII: S0965-1748(01)00032-7

are encoded by members of a large gene family. In *D. melanogaster*, eight divergent intronless genes are found within a 14-kb DNA segment (Toung et al., 1993), at least five different Class I GST genes are present in *M. domestica* (Zhou and Syvanen, 1997). Furthermore, 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* (Ranson et al., 1998). The *aggst1–3*, *aggst1–4*, *aggst1–5* and *aggst1–6* genes are alternatively spliced products from a single gene and share the same N-terminus.

Recently, 7.5 kb of the *adgst1AS1* gene (Fig. 1) was identified from an *An. dirus* genomic library (Pongjaroenkit et al., 2001). This gene had >78.5% nucleotide sequence identity to *aggst1α* from *An. gambiae* (Ranson et al., 1998). The arrangement and the nucleotide sequence identity of each exon between these two genes are similar. The alternatively spliced products of the *adgst1AS1* gene were predicted to produce four mature transcripts as in the *aggst1α* gene. One of the splicing products, *adgst1–1*, which is the splicing product of exons 1, 2 and 3D, has been characterized previously (Prapanthadara et al., 1998). We now report the cloning, heterologous expression and characterization of the other three splicing products of the *adgst1AS1* gene from *An. dirus*. The relationship of the isoenzymes to Class I insect GSTs and other GSTs is also discussed.

2. Materials and methods

2.1. Mosquito strain

The *Anopheles dirus* B colony established at the Department of Parasitology, Faculty of Medicine, Chiang Mai University, was used in this study. The colony was identified on the basis of its morphological and chromosomal characteristics. This strain was reviewed by Baimai (1989).

2.2. RNA extraction and PCR amplification

The total RNA was extracted from fourth instar *An. dirus* larvae using TRIzol™ Reagent and cleaned up using RNeasy® Mini Kits (QIAGEN) as described in the manufacturer's instructions. The primers used in these experiments were designed based on the *adgst1AS1* gene. Two sets of primers were used for cloning into different vectors, pUC19 was used as the cloning vector and pET3a (Novagen) was used as the expression vector. The first-strand cDNA was synthesized using SUPERSCRIPT™ II RNase H⁻Reverse Transcriptase (GIBCO BRL) in the presence of oligo(dT)$_{15}$ primer and then cleaned using a QIAquick PCR Purification Kit (QIAGEN) as described in the manufacturer's instructions.

For cloning into the pUC19 vector, the PCR reactions were performed using 5′-CCGGCGGGATCC-ATGGATTTTTATTACCTACCC-3′ as the 5′ primer which contains a *Bam*HI site and 5′-TCAATGCTT-AATCCGATCGAAAAACG-3′, 5′-CGCCGTCGACA-TATGTTACTTCTCAAAGTACTT-3′ and 5′-TC-ATTTTTGTGTGAAGCGCCCGA-3′ as the 3′ primers of *adgst1–2*, *adgst1–3* and *adgst1–4*, respectively. For each reaction, 50 μl consisted of 500 ng of 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% Triton X-100), 200 μM each of dATP, dGTP, dCTP and dTTP, and 0.5 units of Vent® DNA Polymerase (New England BioLabs). The PCR amplification was performed using a P-E Thermal Cycler 2400. After the initial denaturing step at 94°C for 5 min, 35 cycles were performed. Each cycle consisted of a denaturing step at 94°C for 30 s, an annealing step specific for *adgst1–2*, *adgst1–3* and *adgst1–4* at 60°C, 64°C or 62°C, respectively, for 30 s and an extension step at 72°C for 1 min. A final extension step was performed at 72°C for 7 min. PCR products were purified from an agarose gel with a GENECLEAN II® Kit (Bio 101). The products were digested with *Bam*HI (New England BioLabs) and ligated into pUC19 vectors that had been double digested with *Bam*HI and *Hin*cII (New England BioLabs). The inserts were sequenced in both directions at least twice using an ABI Prism 377 DNA Sequencer (Perkin Elmer). A multiple alignment of the nucleotide sequences and the translated amino acid sequences to other insect GSTs was analyzed using the ClustalX program.

2.3. Preparation of expression constructs

For subcloning, PCR was performed using the 5′ primer containing the *Nde*I restriction site 5′-CCGA-

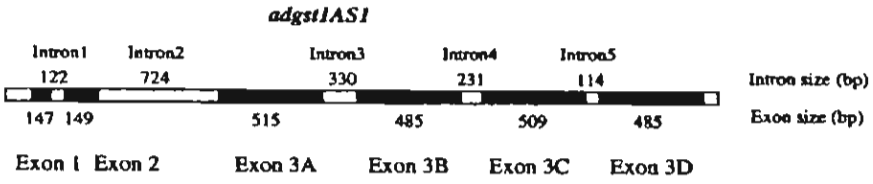

Fig. 1. A schematic of the *adgst1AS1* gene isolated from an *An. dirus* genomic library. The genomic sequence Genbank accession number is AF251478.

GAGCATATGGATTTCTACTACCTTCCC-3′ in conjunction with the previous 3′ primers specific to *adgst1–2*, *adgst1–3* and *adgst1–4*. The PCR conditions were performed as described above. The PCR products were digested and ligated into the pET3a vectors at the *Nde*I site. The inserts were again sequenced twice in both directions to confirm that no changes had been introduced by the PCR. Then *E. coli* BL21(DE3)pLysS was used as an expression host.

2.4. Recombinant protein expression and purification method

A colony of *E. coli* BL21(DE3)pLysS which contained a recombinant plasmid was grown at 37°C until the OD_{600} was approximately 0.6. After induction with 0.1 mM IPTG (final concentration) for 3 h, the cells were placed on ice for 20 min, and collected by centrifugation at 5000 rpm, 4°C for 10 min. The cell pellets from 50 ml of culture were suspended by mixing with 4.8 ml of 50 mM Tris–HCl (pH 7.4), 1 mM EDTA, 200 μl of 100 mg/ml lysozyme and 3.6 μl of 1.4 M β-mercaptoethanol. The suspension was incubated on ice for 20 min, then 50 μl of 1 M DTT was added and the suspension lysed at 900 psi in a French Pressure cell. The lysate was then centrifuged at 10,000g at 4°C for 20 min. The supernatant containing the soluble form of the recombinant protein was separated from the pellet.

The soluble target protein was purified at 4°C by using *S*-hexylglutathione immobilized on agarose (Sigma). The soluble target protein was mixed with the *S*-hexylglutathione agarose gel and incubated on ice for 5–10 min. The gel was then washed 6× with 50 mM Tris–HCl (pH 7.4) containing 0.2 M NaCl and 1 mM EDTA. The GST recombinant protein was eluted from the *S*-hexylglutathione agarose gel with 4 gel volumes of 50 mM Tris–HCl (pH 7.4), 0.2 M NaCl, 1 mM EDTA, containing 5 mM *S*-hexylglutathione and 10 mM DTT. After a concentration step using Centriprep-10 (Amicon), *S*-hexylglutathione that was bound to the recombinant GST was eliminated by gel filtration using PD-10 columns (Pharmacia) equilibrated with 50 mM phosphate buffer, pH 6.5. The concentrated supernatant containing the target protein was applied to the column and the purified protein eluted in the same buffer supplemented with 10 mM DTT. This *S*-hexylglutathione elimination step was performed twice. All the steps were carried out at 4°C. Glycerol was then added to a final concentration of 40% (v/v) and the purified concentrated GSTs stored at −20°C.

2.5. Characterization of the expressed enzymes

The purity and subunit size of the enzyme preparations were determined by SDS–PAGE using Bio-Rad broad range standards as molecular weight markers. Various substrate specificities and steady-state kinetic studies were performed as described previously (Prapanthadara et al., 1996; Ranson et al., 1997). GST activity with 1,2-dichloro-4-nitrobenzene (DCNB) was measured at 345 nm with 1 mM DCNB and 10 mM glutathione in 100 mM potassium phosphate buffer, pH 7.5. DDT-dehydrochlorinase activity was determined by the following method. A 1-ml reaction mixture contained an appropriate amount of enzyme in 0.1 M phosphate buffer pH 6.5, with a final concentration of 0.05 mM *p,p′*-DDT. Glutathione, to give a final concentration of 10 mM, was added to initiate the enzymatic reaction. The reaction mixtures were incubated at 28°C for 2 h, then 5 nmol dicofol were added before the samples were extracted three times each with 1.5 ml of chloroform. The dicofol was used as an internal control to determine the efficiency of extraction. The chloroform extracts (4.5 ml) were pooled and left to air-dry overnight, then stored at −20°C until used. The extracts were taken up in 100 μl of isopropanol and 100 μl of mobile phase before analysis. The HPLC was performed with a Waters Model 510 HPLC pump, Waters 490 Programmable Multi Wavelength detector and Waters Data Module 745 integrator. The reverse phase chromatographic analysis of DDT metabolites was performed by injecting 20 μl of the extracts into Waters Radial-PAK cartridge column type 8NVC18 (4 μm) with a Guard-PAK pre-column using a Rheodyne model 7215 manual injector with a 20-μl loop. The isocratic mobile phase contained methanol:acetonitrile:water (72.5:12.5:15, v/v) and the flow rate was 0.8 ml/min. The detector sensitivity was set at 0.5 absorbance units full scale (AUFS) at a wavelength of 236 nm. Peaks were integrated on the Data Module 745 integrator using the internal standard control. The amount of DDE produced (nmol/mg protein) was calculated by comparison with integrated values for reference standards of known concentrations of *p,p′*-DDT injected onto the HPLC column. The recovery of dicofol was also calculated from the peak area of a known concentration injected onto the HPLC column. The recovery factor generated from dicofol was then used to correct the DDE concentration. The lower limit of sensitivity for DDE was 0.025 nmol per injection.

K_m and V_{max} for 1-chloro-2,4-dinitrobenzene (CDNB) and glutathione (GSH) were determined by non-linear regression analysis using GraphPad Prism 2.01 Software. Protein concentration was determined using the Bio-Rad protein reagent with bovine serum albumin as the standard protein (Bradford, 1976).

Insecticide inhibition studies were performed using permethrin (Chem Service), a pyrethroid insecticide. The GST assay was performed at 0.025–2.0 mM CDNB concentrations with GSH held constant at a saturating concentration and at 0.25–10 mM GSH concentrations with CDNB held constant at a saturating concentration. The inhibition assay was determined in the absence and pres-

ence of various concentrations of permethrin from 0.01 to 0.1 mM. The initial rate of reaction was used to construct a double reciprocal plot, $1/V$ versus $1/S$, and the inhibition constant (K_i) was determined from the values of the appropriate intercepts on the axes (Dixon and Webb, 1979).

Mass spectrometry was used to determine if GST conjugation or metabolism of permethrin occurred. Electrospray ionization was performed using an API-365 LC/MS/MS triple quadrupole mass spectrometer equipped with an atmospheric pressure ionization source (PE SCIEX, USA). The liquid samples were loaded in a 100-μl glass syringe and transferred directly to the ion spray needle by a motor driven plunger at flow rate of 5 μl/min. The ionization potential was set at 4800–5200 V. Q1 of the mass spectrometer was scanned from 150 to 2000 of an *m/z* unit. The orifice potential was varied from 20 to 120 eV and the electron multiplier detector was set at 2200 eV.

3. Results

3.1. Cloning of three alternatively spliced products of the adgst1AS1 gene

Using primers based on the *adgst1AS1* gene, three cDNA sequences were amplified by PCR and named *adgst1-1*, *adgst1-3* and *adgst1-4*. Nucleotide sequences of these three cDNAs share >67% identity. The cDNAs encode enzymes of 217, 209 and 219 amino acids, respectively. The translated amino acid sequences share >61% identity and residues 1–45 in the N-termini are coded for by exon 2 and are identical (Fig. 2). adGST1-2, adGST1-3 and adGST1-4 show >50% identity to members of the Delta class GST family which has been identified previously as insect Class I GSTs, and <12% identity to the Sigma class GSTs sometimes called the insect Class II GST family. The *An. dirus* GSTs shows the highest identity to the *An. gambiae* GSTs, being 85% for adGST1-2 and agGST1-4, 92% for adGST1-3 and agGST1-5, and 85% for adGST1-4 and agGST1-3 (Table 1).

3.2. Expression of recombinant GSTs in E. coli

The expression of adGST1-2, adGST1-3 and adGST1-4 after 3 h of IPTG induction yielded approximately 15–20 mg from 200 ml of *E. coli* cultures. After the final step of purification on *S*-hexylglutathione agarose, the enzyme yields were approximately 40–70% of the total activity observed in the *E. coli* lysate. The purified enzymes were shown to be homogeneous preparations that appeared as single protein bands on SDS-PAGE. The expected bands are about 23–25 kDa, which corresponds to the calculated molecular weight of the GST subunits.

3.3. Substrate specificities

Measurements of DDT-dehydrochlorination performed by the alternatively spliced GSTs showed similar activities, that is 1.17, 2.80 and 3.72 nmol DDE formation/mg protein for adGST1-2, adGST1-3 and adGST1-4, respectively. The specific activities measured for the recombinant *An. dirus* GSTs with various substrates are shown in Table 2. adGST1-2, adGST1-3 and adGST1-4 have high specific activity with CDNB and detectable specific activity with DCNB. However, all three enzymes have very low specific activity with the mammalian Pi class substrate, ethacrynic acid, and the mammalian Theta class substrate, *p*-nitrophenethyl bromide. The use of these four substrates demonstrates the differences in specificity of these three enzymes.

3.4. Kinetic properties

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 and compared among the four alternatively spliced products of the *adgst1AS1* gene (Table 3). The K_m values with respect to CDNB and GSH for adGST1-3 indicate that this enzyme has the greatest affinities for these substrates. To determine the catalytic properties, the turnover number for CDNB, k_{cat}, and the catalytic efficiency, k_{cat}/K_m, with respect to CDNB and GSH were calculated. Among the alternatively spliced products of the *adgst1AS1* gene, adGST1-3 is the most reactive in catalyzing CDNB conjugation.

3.5. Insecticide inhibition

The percent inhibition by 0.1 mM permethrin of adGST1-2, adGST1-3 and adGST1-4 on CDNB conjugating activity ranged from 40 to 80% (Table 4). In addition, insecticide inhibition kinetics were determined with various permethrin concentrations in the presence of various CDNB or GSH concentrations. Double reciprocal plots of permethrin inhibition with varying CDNB concentrations showed uncompetitive inhibition of adGST1-2, non-competitive inhibition of adGST1-3 and competitive inhibition of adGST1-4 (Fig. 3). When varying GSH, permethrin behaved as a non-competitive inhibitor of all three adGSTs. The K_i values of the three enzymes for permethrin with respect to either CDNB or GSH are similar which indicates that these three enzymes have similar affinities to permethrin, although the different inhibition kinetics demonstrate a different mechanism of interaction.

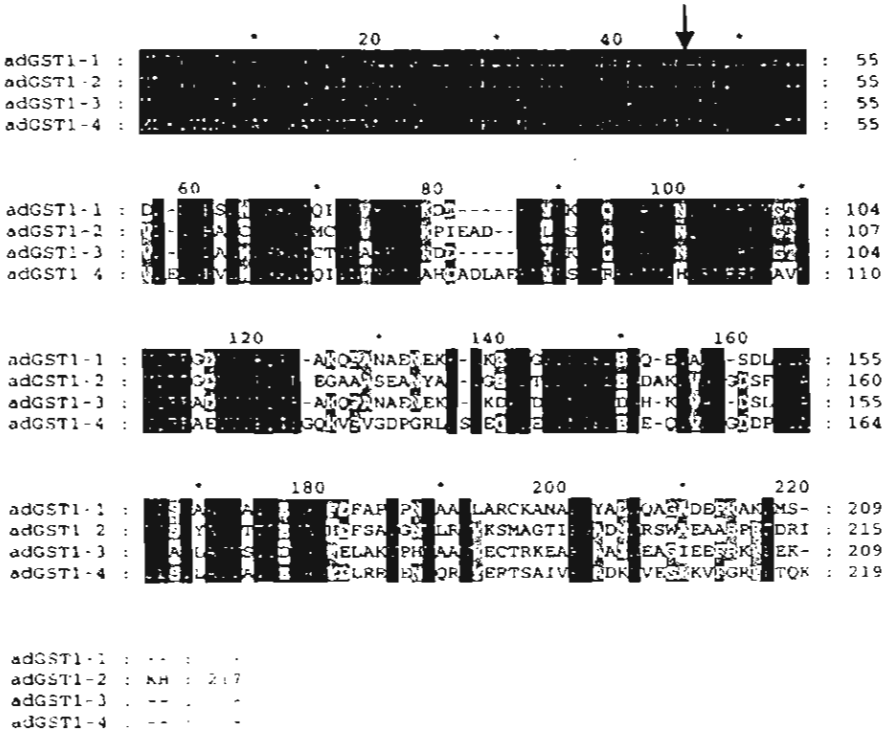

Fig. 2. Amino acid alignment of the *adgst1AS1* splicing products. Gaps introduced to maximize sequence similarity are shown by a horizontal dash. Black shadings represent 100% sequence homology and gray shadings represent 80% sequence homology. A vertical arrow indicates the junction between exon 2 and exon 3 in *adgst1-1*, *adgst1-2*, *adgst1-3* and *adgst1-4* (Genbank accession numbers AF273041, AF273038, AF273039 and AF273040, respectively).

Table 1
Amino acid identity of *Anopheles* class I GSTs

	Ad1-2	Ad1-3	Ad1-4	Ad1-1	Ag1-1	Ag1-3	Ag1-4	Ag1-5	Ag1-6	Ag1-7
Ad1-2	100	64	61	61	47	57	**85**	62	62	41
Ad1-3		100	63	77	50	60	63	**92**	80	44
Ad1-4			100	63	52	**85**	59	64	65	39
Ad1-1				100	48	58	63	76	**91**	45
Ag1-1					100	48	47	50	50	36
Ag1-3						100	59	62	61	37
Ag1-4							100	64	66	41
Ag1-5								100	82	41
Ag1-6									100	44
Ag1-7										100

The amino acid sequences were obtained from the GenBank database. The abbreviations are Ad for *Anopheles dirus* GST; Ad1-1 (AF273041), Ad1-2 (AF273038), Ad1-3 (AF273039), Ad1-4 (AF273040), Ag for *Anopheles gambiae* GST; Ag1-1 (2842739), Ag1-3 (3549276), Ag1-4 (3549274), Ag1-5 (2842717), Ag1-6 (2842718) and Ag1-7 (3549271).

The mass spectrometry data indicated that these GSTs bound but did not further interact with permethrin. Multiple *m/z* peaks were detected for GST corresponding to the deduced mass. GSH and GSSG peaks were identified at *m/z*=308 and 615 with Na and K adducts (330, 346 and 638, 653). A single peak at *m/z*=413 was detected for Permethrin+Na containing neighbor isotropic peaks of intensity ratio 9:6:1 for two chlorine atoms.

4. Discussion

The GST gene *adgst1AS1* was identified from a genomic library of *An. dirus*. The exon arrangement of the *adgst1AS1* gene was found to be similar with the *An. gambiae aggst1α* gene and showed >78.5% nucleotide sequence identity (Pongjaroenkit et al., 2001). This gene contains six exons, which was predicted to produce four

Table 2
Substrate specificities of the recombinant *Anopheles dirus* GSTs

Substrate	Specific activity (μmol/min/mg of protein)		
	adGST1-2	adGST1-3	adGST1-4
CDNB	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 1 mM; DCNB 1 mM; ethacrynic acid 0.2 mM; and *p*-nitrophenyl bromide 0.1 mM.

Table 3
Kinetic parameters of *Anopheles dirus* GSTs

Kinetic parameters	adGST1-1[a]	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	121	269	32
$k_{cat}/K_{m\ GSH}$	6	20	67	20

The units are: V_{max}, μmol/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.
[a]Ketterman et al., 2001.

Table 4
Permethrin inhibition kinetics of *Anopheles dirus* GSTs

Parameter	AdGST1-2	AdGST1-3	AdGST1-4
% inhibition	81	70	40
K_i CDNB (μM)	8.9±5.3	52.5±32.1	33.8±14.9
K_i GSH (μM)	8.7±1.1	11.0±4.6	34.1±9.7
Inhibition to CDNB	Uncompetitive	Non-competitive	Competitive
Inhibition to GSH	Non-competitive	Non-competitive	Non-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.

alternatively spliced products. One splicing product of this gene, *adgst1-1*, has been isolated and characterized previously (Prapanthadara et al., 1998).

Three novel cDNA species, *adgst1-2*, *adgst1-3* and *adgst1-4*, which encode full-length glutathione *S*-transferases, have been isolated from the mosquito *An. dirus*, using primers designed from the *adgst1AS1* gene. Their nucleotide sequences are different from the genomic sequence at several residues but the translated amino acid sequences were not changed. The nucleotide sequences from position 1–135 of *adgst1-1*, *adgst1-2*, *adgst1-3* and *adgst1-4* are identical to exon 2 of the *adgst1AS1* gene while the remaining sequences are nearly identical to each of exons 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 yield mature transcripts produced from the *adgst1AS1* gene. The alternatively spliced products also shared a very high amino acid sequence identity between products from an *An. gambiae* alternatively spliced gene (Ranson et al., 1998). The amino acid sequence identity between the *An. dirus* and the *An. gambiae* proteins were 85% between adGST1-2 and agGST1-4, 92% between adGST1-3 and agGST1-5 and 85% between adGST1-4 and agGST1-3. This high identity confirms that the orthologous GSTs have occurred across the anopheline species, although the species are malaria vectors in geographically distant epidemic areas, i.e. Southeast Asia and Africa. The high conservation of the amino acid sequence of the spliced

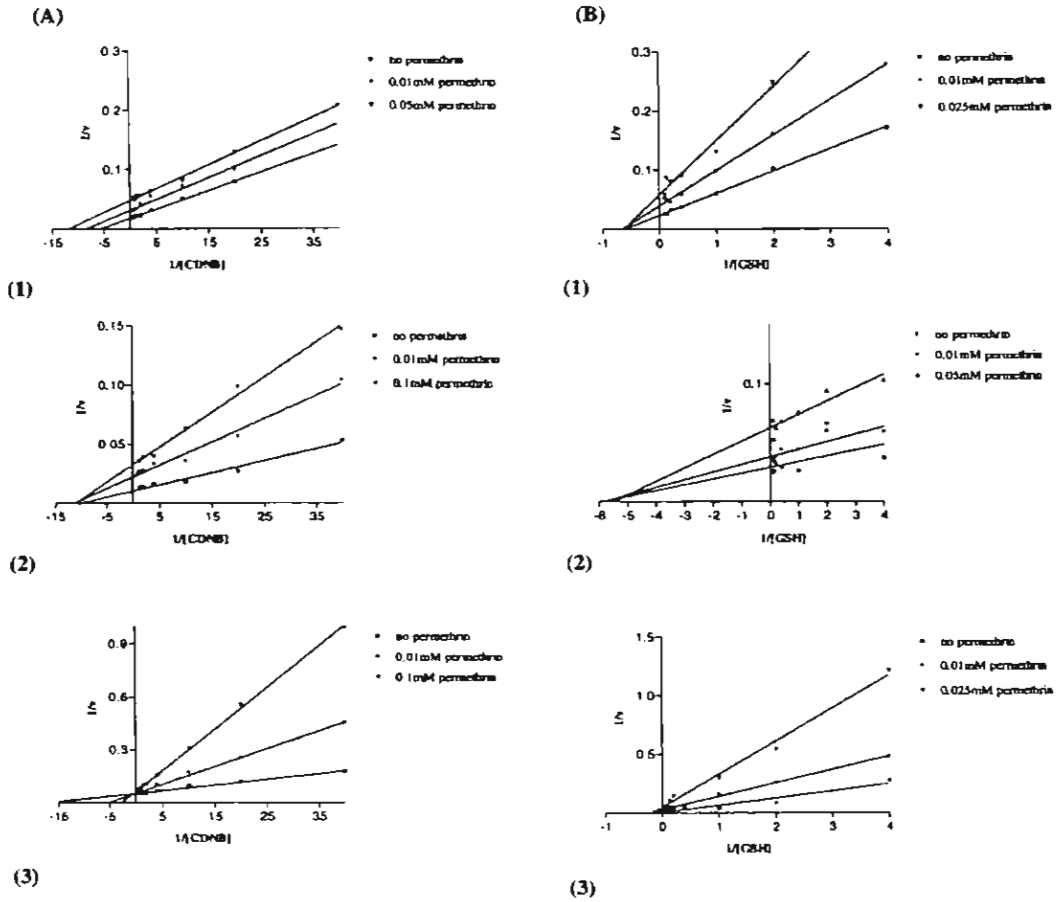

Fig. 3. Double reciprocal plots of permethrin inhibition kinetics on (1) adGST1–2, (2) adGST1–3 and (3) adGST1–4. (A) The initial rate of the enzymatic reaction was measured at 0.025–2.0 mM CDNB concentrations with GSH held constant at a saturating concentration. (B) The initial rate of the enzymatic reaction was measured at 0.25–10 mM GSH concentrations with CDNB held constant at a saturating concentration.

products of the *adgst1AS1* gene and the *aggst1α* gene, as well as conservation of the genomic splicing pattern, suggests that these two GST genes have important physiological roles that must have been maintained across millions of years of evolution.

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-residue leader sequence tag attached to the N-termini. 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 adGST1–2, adGST1–3 and adGST1–4 with the various GST substrates, all of them have high activity with the general substrate CDNB, and detectable activity with other substrates. adGST1–3 has the highest activities with the two general GST substrates CDNB and DCNB. Several studies have suggested that elevated DCNB activity correlates positively with resistance (Clark et al., 1986; Grant and Matsumura, 1989 Hemingway et al., 1985; Motoyama and Dauterman, 1975). The activity with CDNB and DCNB exhibited by adGST1–2, adGST1–3 and adGST1–4 suggests that these enzymes may be involved in insecticide resistance.

A comparison of these three translated amino acids among all classes of GSTs showed that they are members of the Class I or Delta class GSTs because they share 50–92% identity with the previously characterized insect Class I GSTs. They are only distantly related to the insect Class II or Sigma class GSTs and the mammalian classes (Theta, Pi, Alpha, Mu, Kappa and Zeta GSTs) because they share <23% identity. Furthermore, these three adGSTs have very low activities against ethacrynic acid and *p*-nitrophenethyl bromide, which are known substrates of mammalian Pi class GSTs (Phillips and Mantle 1991, 1993) and mammalian Theta class GSTs (Meyer et al., 1991). Hence the difference in the amino acid sequences and the substrate specificities confirms that insect Class I or Delta class GSTs are distinct from all other classes of GSTs.

Among the alternatively spliced products of the *adgst1AS1* gene, their steady-state kinetic properties were compared. adGST1–3 has the highest affinity for both GSH (K_m=0.404±0.054 mM) and CDNB (K_m=0.100±0.012 mM). 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

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

Several data suggest that GSTs have an important role in the acquisition of resistance to insecticide (Brown, 1986). 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* (Wang et al., 1991) and elevated levels of GST activity have also been found in DDT-resistant strains of *Aedes aegypti* (Grant et al., 1991). DDT-dehydrochlorinase activity catalyzed by GST is the major mechanism responsible for DDT resistance (Brown, 1986) and glutathione conjugation catalyzed by GST is also a secondary mechanism in some organophosphate resistance (Hemingway et al., 1991). For the recombinant insect GSTs, DDT-dehydrochlorinase activity has been reported in GSTD1 and GSTD21 of *Drosophila* (Tang and Tu, 1994) and adGST1–1 of *Anopheles dirus* (Prapanthadara et al., 1998). The three alternatively spliced GSTs in the present report also possess DDT-dehydrochlorinase activity and approximately to the same extent. A similar affinity was also shown by the three enzymes for permethrin interaction, with K_Is between 9 and 53 μM. In contrast to the similar affinities of permethrin interaction, the enzymes displayed quite different types of inhibition of CDNB activity (Fig. 3). These data suggest that the enzymes bind the permethrin in different ways and/or in different sites, with the binding differences being mediated by the approximately 36–39% differences in the amino acid sequences in the enzymes. Although these enzymes bound permethrin, no metabolism could be detected by mass spectrometry. However, binding and sequestration of insecticides is a well-known mechanism of resistance (Brown, 1986).

The translation products of these adGST transcripts are identical at the N-termini (residues 1–45) but highly variable at the C-termini. From crystal structures of mammalian and insect GSTs, it was found that the N-terminal domain, which has been identified as the glutathione-binding site, is very conserved within a GST class, while the C-terminal domain, which has been identified as the hydrophobic binding site, is variable even for enzymes within the same class (Dirr et al., 1994b; Wilce et al., 1995). Alternative splicing is known to be a mechanism which generates functionally diverse isoforms for almost every type of protein involved in metazoan development, cell function and physiology (Lopez, 1998). Our results indicate that the *adgst1AS1* gene possesses an alternative splicing mechanism that can produce four different mature products. This increases the diversity of GSTs and expands their substrate range with a minimal length of gene.

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 this region of protein is highly conserved within a class of GSTs. The divergence in the C-terminus confers the variation in substrate specificities of different GST isoenzymes (Wilce et al., 1995; Stenberg et al., 1991). We have shown that a one or two amino acid change in the adGST1–1 allelic forms can affect the activity and kinetic properties of the enzyme (Ketterman et al., 2001). adGST1–2, adGST1–3 and adGST1–4 have only 61–64% amino acid identity as well as having differences in translated amino acid lengths of 217, 209 and 219 residues, respectively. These differences occur in the C-terminus of the enzymes, which contains the hydrophobic binding site residues. While folding of the GST polypeptide is well conserved, variations in the sequences due to substitutions, deletions, or insertions have led to some overall readjustments of amino acid residues such as the rotation of structural domains (Dirr et al., 1994a). The sequence differences of these alternatively spliced enzymes may lead to overall conformational changes and subsequently to changes in both substrate and inhibitor interactions.

In conclusion, although these adGSTs are the splicing products of the same gene and share N-terminus sequence, their substrate specificities, steady-state kinetics with respect to both GSH and CDNB, and inhibition kinetics to the pyrethroid insecticide permethrin are very different. Hence, we postulate that the C-terminal changes resulting from the differential mRNA splicing in these GSTs produce enzymes with very different potentials for enzyme specificities.

Acknowledgements

This work was funded by the Thailand Research Fund (TRF). S.P. was funded by a Royal Golden Jubilee scholarship.

References

Baimai, V., 1989. Speciation and species complexes of the anopheles malaria vectors in Thailand. In: The Third Conference on Malaria Research, Thailand, pp. 146–152.

Beall, C., Fyrberg, C., Song, S., Fyrberg, E., 1992. Isolation of a *Drosophila* gene encoding glutathione *S*-transferase. Biochem. Genet. 30, 517–527.

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

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

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

Dirr, H., Reinemer, P., Huber, R., 1994a. X-ray crystal structures of cytosolic glutathione *S*-transferases. Implications for protein archi-

tecture, substrate recognition and catalytic function. Eur. J. Biochem. 220, 645–661.

Dirr, H., Reinemer, P., Huber, R., 1994b. Refined crystal structure of porcine class pi glutathione *S*-transferase (pGSTP1-1) at 2.1 Å resolution. J. Mol. Biol. 243, 72–92.

Dixon, M., Webb, E.C., 1979. Enzyme inhibition and activation. In: Enzymes. Academic Press Inc, New York, pp. 332–467.

Fournier, D., Bride, J.-M., Poirié, M., Bergé, J.-B., Plapp, F.W., 1992. Insect glutathione *S*-transferases. Biochemical characteristics of the major forms from houseflies susceptible and resistant to insecticides. J. Biol. Chem. 267, 1840–1845.

Franciosa, H., Berge, J.B., 1995. Glutathione *S*-transferases in housefly (*Musca domestica*): location of GST-1 and GST-2 families. Insect Biochem. Mol. Biol. 25, 311–317.

Grant, D.F., Dietze, E.C., Hammock, B.D., 1991. Glutathione *S*-transferase isozymes in *Aedes aegypti*: purification, characterization, and isozyme-specific regulation. Insect Biochem. 21, 421–433.

Grant, D.F., Matsumura, F., 1989. Glutathione *S*-transferase 1 and 2 in susceptible and insecticide resistant *Aedes aegypti*. Pestic. Biochem. Physiol. 33, 132–143.

Hayes, J.D., Pulford, D.J., 1995. The glutathione *S*-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. CRC Crit. Rev. Biochem. Mol. Biol. 30, 445–600.

Hemingway, J., Malcolm, C.A., Kissoon, K.E., Boddington, R.G., Curtis, C.F., Hill, N., 1985. The biochemistry of insecticide resistance in *Anopheles sacharovi*: comparative studies with a range of insecticide susceptible and resistant *Anopheles* and *Culex* species. Pestic. Biochem. Physiol. 24, 68–76.

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

Hiratsuka, A., Sebata, N., Kawashima, K., Okuda, H., Ogura, K., Watabe, T., Satah, K., Hatayama, I., Tsuchida, S., Ishikawa, T., Sato, K., 1990. A new class of rat glutathione *S*-transferase Yrs–Yrs inactivating reactive sulfate esters as metabolites of carcinogenic arylmethanols. J. Biol. Chem. 265, 11973–11981.

Ketterman, A.J., Prommeenate, P., Boonchauy, C., Chanama, U., Leetachewa, S., Promtet, N., Prapanthadara, L., 2001. Single amino acid changes outside the active site significantly affect activity of glutathione *S*-transferases. Insect Biochem. Mol. Biol. 31, 65–74.

Lopez, A.J., 1998. Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. Annu. Rev. Genet. 32, 279–305.

Meyer, D.J., Coles, B., Pemble, S.E., Gilmore, K.S., Fraser, G.M., Ketterer, B., 1991. Theta, a new class of glutathione transferases purified from rat and man. Biochem. J. 274, 409–414.

Motoyama, N., Dauterman, W.C., 1975. Interstrain comparison of glutathione-dependent reactions in susceptible and resistant houseflies. Pestic. Biochem. Physiol. 5, 489–495.

Pemble, S.E., Taylor, J.B., 1992. An evolutionary perspective on glutathione transferases inferred from class-Theta glutathione transferase cDNA sequences. Biochem. J. 287, 957–963.

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

Phillips, M.F., Mantle, T.J., 1993. Inactivation of mouse liver glutathione *S*-transferase YfYf (Pi class) by ethacrynic acid and 5,5′-dithiobis-(2-nitrobenzoic acid). Biochem. J. 294, 57–62.

Pongjaroenkit, S., Jirajaroenrat, K., Boonchauy, C., Chanama, U., Leetachewa, S., Prapanthadara, L., Ketterman, A.J., 2001. Genomic organization and putative promotors of highly conserved glutathione *S*-transferases originating by alternative splicing in *Anopheles dirus*. Insect Biochem. Mol. Biol. 31, 75–85.

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

Prapanthadara, L., Ranson, H., Somboon, P., Hemingway, J., 1998. Cloning, expression and characterization of an insect class I glutathione *S*-transferase from *Anopheles dirus* species B. Insect Biochem. Mol. Biol. 28, 321–329.

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

Ranson, H., Collins, F., Hemingway, J., 1998. The role of alternative mRNA splicing in generating heterogeneity within the *Anopheles gambiae* class I glutathione *S*-transferase family. Proc. Natl. Acad. Sci. USA 95, 14284–14289.

Reiss, R.A., James, A.A., 1993. A glutathione *S*-transferase gene of the vector mosquito, *Anopheles gambiae*. Insect Mol. Biol. 2, 25–32.

Snyder, M.J., Walding, J.K., Feyereisen, R., 1995. Glutathione *S*-transferases from larval *Manduca sexta* midgut: sequence of two cDNAs and enzyme induction. Insect Biochem. Mol. Biol. 25, 455–465.

Stenberg, G., Board, P.G., Carlberg, I., Mannervik, B., 1991. Effects of directed mutagenesis on conserved arginine residues in a human class alpha glutathione transferase. Biochem. J. 274, 549–555.

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

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

Wang, J., McCommas, S., Syvanen, M., 1991. Molecular cloning of a glutathione *S*-transferase overproduced in an insecticide-resistant strain of the housefly (*Musca domestica*). Mol. Gen. Genet. 227, 260–266.

Wilce, M.C.J., Board, P.G., Feil, S.C., Parker, M.W., 1995. Crystal structure of a theta-class glutathione transferase. EMBO J. 14, 2133–2143.

Zhou, Z., Syvanen, M., 1997. A complex glutathione transferase gene family in the housefly *Musca domestica*. Mol. Gen. Genet. 256, 187–194.

Acta Crystallographica Section D
Biological
Crystallography
ISSN 0907-4449

Crystallization of two glutathione S-transferases from an unusual gene family

Aaron J. Oakley,[a] Kanya Jirajaroenrat,[b] Thasaneeya Harnnoi,[b] Albert J. Ketterman[b] and Matthew C. J. Wilce[a]*

[a]Department of Pharmacology/Crystallography Centre, University of Western Australia and the Western Australian Institute for Medical Research, Nedlands WA 6907, Australia, and [b]Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakom Pathom 73170, Thailand

Correspondence e-mail: mwilce@receptor.pharm.uwa.edu.au

Received 18 December 2000
Accepted 19 March 2001

Two glutathione S-transferase isozymes from the mosquito *Anopheles dirus* (AdGST1-3 and AdGST1-4) from an alternately spliced gene family have been expressed, purified and crystallized. The isozymes share an N-terminal domain derived from a single exon and C-terminal domains from unique exons. Despite the high level of sequence identity (64% overall), the two isozymes crystallize in different space groups, the 1-3 isozyme in $P3_121$ or $P3_221$ (unit-cell parameters $a = 49.9$, $c = 271.8$ Å at 100 K) and the 1-4 isozyme in $P4_1$ or $P4_3$ (unit-cell parameters $a = 87.8$, $c = 166.1$ at 100 K). Determination of these structures will advance our understanding of how these enzymes inactivate pesticides and the structural consequences of alternate splicing.

1. Introduction

Glutathione S-transferases (GSTs; E.C. 2.5.1.18) are widely distributed in nature. Found in many bacteria and eukarya (Wilce & Parker, 1994), they play an important role in cellular detoxification. The most important reaction catalysed by GSTs is the nucleophilic addition of the thiol group of glutathione (γ-Glu-Cys-Gly; GSH) to compounds with electrophilic centres (Mannervik & Danielson, 1988). Glutathione conjugates are more water-soluble and act as a molecular marker, being selectively removed from the cells by specific pumps in the cell membrane (Hayes & McLellan, 1999).

Traditionally, GSTs have been divided into pi, alpha, mu and theta classes (Mannervik *et al.*, 1992). However, it has become apparent that this enzyme family is far more structurally and functionally diverse than previously thought. Recently, a number of extra classes has been proposed, including beta (Rossjohn *et al.*, 1998), delta (Toung *et al.*, 1993; Ketterman *et al.*, 2001), zeta (Board *et al.*, 1997), phi (*e.g.* Jepson *et al.*, 1994), sigma (Ji *et al.*, 1995), kappa (Pemble *et al.*, 1996) and Omega (Board *et al.*, 2000).

While mammalian GSTs have been well characterized structurally and biochemically, insect GSTs have not been comprehensively studied. Insect GSTs have been identified and reported in multiple forms from the house fly (Clark & Dauterman, 1982; Clark *et al.*, 1984; Motoyama & Dauterman, 1978), *Drosophila melanogaster* (Toung *et al.*, 1990, 1993), grass grub (Clark *et al.*, 1985) and mosquito (Grant & Matsumura, 1989; Prapanthadara *et al.*, 1993). Elevated GST levels have been detected in pesticide-resistant strains of insects and insect GSTs have been shown to break down pesticides (Prapanthadara *et al.*, 2000). Thus, the understanding of the structure and function of these enzymes is of considerable importance. Only one insect GST structure has been determined to date (Wilce *et al.*, 1995) in the absence of any pesticide or xenobiotic. AdGST1-3 and AdGST1-4 have 71 and 60% sequence identity with *Lucilia* GST, repectively (Fig. 1). The recombinant enzymes have high activities with 1-chloro-2,4-dinitrobenzene (CDNB) and detectable activity with 1,2-dichloro-4-nitrobenzene, but markedly low activity with ethacrynic acid and *p*-nitrophenethyl bromide (Jirajaroenrat *et al.*, 2001). DDT-dehydrochlorination performed by the alternatively spliced GSTs showed 2.80 and 3.72 nmol DDE formation per milligram of protein for AdGST1-3 and AdGST1-4, respectively (Jirajaroenrat *et al.*, 2001). We hope to utilize biochemical and structural information to understand how these GST attack such pesticides. Our interest also lies in understanding the effects of the unusual gene arrangement on structure and the mechanism of pesticide detoxification by these GSTs.

2. Materials and methods

2.1. Recombinant protein expression and purification

The recombinant proteins of AdGST1-3 and AdGST1-4 were cloned, overexpressed and purified as previously described (Jirajaroenrat *et al.*, 2001).

2.2. Crystallization

Using the hanging-drop vapour-diffusion method, initial screens of crystallization

Table 1
X-ray data-collection statistics for GST1-3.

Resolution shell	No. of reflections	Completeness (%)	R_{sym}	$I/\sigma(I)$
15.00–3.75	10024	79.0	0.042	26.4
3.75–2.99	8330	66.1	0.062	20.1
2.99–2.61	7447	59.1	0.068	16.2
2.61–2.37	8078	64.5	0.078	10.3
2.37–2.20	9928	79.0	0.120	5.1
2.20–2.07	10757	85.8	0.134	4.4
2.07–1.97	10652	84.7	0.155	4.3
1.97–1.88	10397	82.9	0.192	3.8
1.88–1.81	10555	83.9	0.208	3.3
1.81–1.75	10613	85.1	0.303	2.2
Overall	96781	77.0	0.073	9.4

conditions were carried out employing a sparse-matrix kit (Hampton Research, CA, USA).

Crystals of GST1-3 were observed in about 50% of the Crystal Screen I drops. The largest crystals grew in condition 10 (0.2 *M* ammonium acetate, 0.1 *M* sodium acetate pH 4.6, 30% PEG 4000 at room temperature). The optimized conditions consisted of 0.2 *M* ammonium acetate, 0.1 *M* sodium acetate pH 4.4, 25–40% PEG 4000. The final drops were made up of 2 µl GST1-3 at 12 mg ml^{-1} (preincubated for 2 h on ice with an equal volume of 10 m*M* GSH and 10 m*M* DDT) and 2 µl well solution. All crystals were grown at room temperature. The largest measured 0.5 × 0.1 × 0.1 mm, appearing as a thick rod. These crystals could be flash-frozen without the need for an additional cryoprotectant.

Crystals of GST1-4 were obtained using condition 28 of Crystal Screen I (0.2 *M* sodium acetate, 30% PEG 8000, 0.1 *M* sodium cacodylate pH 6.5 at room temperature). Optimum conditions for crystal growth were subsequently found with a solution containing 10–20% PEG 8000, 0.1 *M* sodium cacodylate pH 6.2 and 1% glycerol. The crystals were also improved when reduced glutathione was added to the drop. The final drops were made up of 2 µl GST1-4 at 14 mg ml^{-1}, 2 µl 10 m*M* GSH and 2 µl well solution; crystals were grown at room temperature. The presence of PEG 8000 in the crystallization buffer allowed the crystals to be cryofrozen without the need for any additional reagents. Crystals appeared after 4 d. The largest was 0.2 × 0.05 × 0.05 mm and appeared as fine rods.

2.3. Collection of crystallographic diffraction data

Diffraction data from crystals of GST1-3 and GST1-4 were collected at 100 K using a MAR Research image-plate detector and Cu $K\alpha$ radiation from a Rigaku rotating-anode generator operating at 40 kV and 100 mA equipped with focusing mirrors. The crystals were cooled using an Oxford cryocooler. 0.25° oscillations were used.

The X-ray data were processed with the *HKL* package (Otwinowski & Minor, 1997). Space-group identification utilized the programs *XDISPLAYF* and *SCALEPACK*. Data were reduced and merged using *DENZO* and *SCALEPACK*. The quality of the data are shown in Tables 1 and 2.

3. Results and discussion

Two closely related GSTs from *A. dirus* have been successfully crystallized under different conditions utilizing PEG as the main precipitant (PEG 4000 and 8000). The different space groups indicate that the crystal packing is different for the two enzymes.

Table 2
X-ray data-collection statistics for GST1-4.

Resolution shell	No. of reflections	Completeness (%)	R_{sym}	$I/\sigma(I)$
20.0–5.25	1640	97.3	0.034	29.5
5.25–4.18	1542	98.9	0.049	26.4
4.18–3.65	1522	99.0	0.063	21.3
3.65–3.32	1466	98.2	0.085	14.9
3.32–3.08	1448	97.2	0.119	10.3
3.08–2.90	1415	95.0	0.178	6.6
2.90–2.76	1366	92.3	0.230	4.5
2.76–2.64	1362	90.4	0.337	3.2
2.64–2.54	1261	85.7	0.410	2.4
2.45–2.45	1215	83.7	0.444	2.2
Overall	14237	93.9	0.077	13.1

For AdGST1-3, diffraction spots were observed to 2.4 Å. Processing in *DENZO* revealed the Bravis lattice to be hexagonal, with unit-cell parameters $a = b = 49.9$ Å, $c = 271.8$ Å. Attempts to reduce the data in *SCALEPACK* revealed the overall Patterson symmetry to be $P\bar{3}m1$. sing observations of systematic absences, the range of possible space groups was reduced to $P3_121$ or its enantiomorph $P3_221$. The resolution dependence and overall strength and completeness of the data are shown in Table 1.

For AdGST1-4, diffraction spots were observed to 1.6 Å. Processing with *DENZO* showed that the Bravis lattice was tetragonal, with unit-cell parameters $a = 87.8$, $c = 166.1$ Å at 100 K. *SCALEPACK* was used to demonstrate that the Patterson symmetry was $P4/m$. Analysis of systematic absences revealed the space group to be $P4_1$ or its enantiomorph $P4_3$. The resolution dependence and overall strength and completeness of the data are shown in Table 2.

It is anticipated that these GSTs will be solvable using the *Lucilia* GST structure (Wilce *et al.*, 1995) as a molecular-replacement model. These crystals will be used in soaking experiments with DDT and other pesticides. Together with enzymatic data, it is hoped that an understanding of how these important enzymes confer resistance to pesticides upon insects such as the mosquito *A. dirus*.

The structures of the GST isozymes will reveal information on how alternate gene splicing affects structure. The first 45 amino-acid residues, which comprise most of the glutathione-binding domain, are identical between the two isozymes, with all of the differences occurring at the C-terminus (Fig. 1). The structures of these enzymes will

```
1-3        MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTELMAGEHMKPEFLKINPQHCIPTLVDN-G
1-4        MDFYYLPGSAPCRAVQMTAAAVGVELNLKLTELMAGEHMKPEFLKLNPQHCIPTLVDEDG
Lucilia    MDFYYLPGSAPCRSVLMTAKALGIELNKKLLELQAGEHLKPEFLKINPQHTIPTLVDG-D
           1.......10........20........30........40........50........60

1-3        FALWESRAICTYLAEKYGKDDK-----LYPKDPQKRAVVNQRLYFDMGTLYQRFADYYYP
1-4        FVLWESRAIQIYLAEKYGAHDADLAERLYPSDPRRRAVVHQRLFFDVAVLYQRFAEYYYP
Lucilia    FALWESRAIMVYLAEKYGKNDS-----LFPKCPKKRAVINQRLYFDMGTLYKSFADYYYP
           ........70........80........90.......100.......110.......120

1-3        QIFAKQP--ANAENEKKMKDAVDFLNTFLDGHKYVAG-DSLTIADLTVLATVSTYDVAGF
1-4        QIFGQKVPVGDPGRLRSMEQALEFLNTFLEGEQYVAGGDDPTIADLSILATIATYEVAGY
Lucilia    QIFAKAP--ADPELYKKMEAAFDFLNTFLEGHQYVAG-DSLTVADLALLASVSTFEVAGF
           .......130.......140.......150.......160.......170.......180

1-3        ELAKYPHVAAWYERTRKEAPGAAINEAGIEEFRKYFEK-
1-4        DLRRYENVQRWYERTSAIVPGADKNVEGAKVFGRYFTQK
Lucilia    DFSKYANVAKWYANAKTVAPGFDENVEGCLEFKKFFN--
           .......190.......200.......210.........
```

Figure 1
BLAST comparison of GST1-3 and GST1-4 from *A. dirus* with GST from *L. cuprina*. Regions with identical residues are shaded.

also aid in the identification of catalytic residues involved in pesticide degradation and the explanation of differences in catalytic activities of the isozymes.

This research is supported by a grant from the Thailand Research Fund (AJK) and a UWA/ARC Small Research Grant (AJO). AJO is a recipient of a UWA postdoctoral research fellowship. MCJW is supported by a Raine Medical Research Foundation Fellowship.

References

Board, P. G., Baker, R. T., Chelvanayagam, G. & Jermiin, L. S. (1997). *Biochem J.* **15**, 929–935.

Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F. & Pandit, J. (2000). *J. Biol. Chem.* **275**, 24798–24806.

Clark, A. G. & Dauterman, W. C. (1982). *Pestic. Biochem. Physiol.* **17**, 307–314.

Clark, A. G., Dick, G. L., Martindale, S. M. & Smith, J. N. (1985). *Insect Biochem.* **15**, 35–44.

Clark, A. G., Shamaan, N. A., Dauterman, W. C. & Hayaoka, T. (1984). *Pestic. Biochem. Physiol.* **22**, 51–59.

Grant, D. F. & Matsumura, F. (1989). *Pestic. Biochem. Physiol.* **33**, 132–143.

Hayes, J. D. & McLellan, L. I. (1999). *Free Radic. Res.* **31**, 273–300.

Jepson, I., Lay, V. J., Holt, D. C., Bright, S. W. & Greenland, A. J. (1994). *Plant Mol. Biol.* **26**(6), 1855–1866.

Ji, X., von Rosenvinge, E. C., Johnson, W. W., Tomarev, S. I., Piatigorsky, J., Armstrong, R. N. & Gilliland, G. L. (1995). *Biochemistry*, **34**(16), 5317–5328.

Jirajaroenrat, K., Pongjaroenkit, S., Krittanai, C., Prapanthadara, L. & Ketterman, A. J. (2001). In the press.

Ketterman, A. J., Prommeenate, P., Boonchauy, C., Chanama, U., Leetachewa, S., Promtet, N. & Prapanthadara, L. (2001). *Insect Biochem. Mol. Biol.* **31**, 65–74.

Mannervik, B., Awasthi, 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 & Wolf, C. R. (1992). *Biochem. J.* **282**, 305–306.

Mannervik, B. & Danielson, U. H. (1988). *Crit. Rev. Biochem. Mol. Biol.* **23**, 283–337.

Motoyama, N. & Dauterman, W. C. (1978). *Insect Biochem.* **8**, 337–348.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* 307–326.

Pemble, S. E., Wardle, A. F. & Taylor, J. B. (1996). *Biochem. J.* **319**, 749–754.

Prapanthadara, L., Hemingway, J. & Ketterman, A. J. (1993). *Pestic. Biochem. Physiol.* **47**, 119–133.

Prapanthadara, L., Promtet, N., Koottathep, S., Somboon, P. & Ketterman, A. J. (2000). *Insect Biochem. Mol. Biol.* **30**, 395–403.

Rossjohn, J., Polekhina, G., Feil, S. C., Allocati, N., Masulli, M., Di Ilio, C. & Parker, M. W. (1998). *Structure*, **6**, 721–734.

Toung, Y.-P. S., Hsieh, T.-S. & Tu, C.-P. D. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 31–35.

Toung, Y. P., Hsieh, T. S. & Tu, C. P. (1993). *J. Biol. Chem.* **268**, 9737–46.

Wilce, M. C., Board, P. G., Feil, S. C. & Parker, M. W. (1995). *EMBO J.* **14**, 2133–2143.

Wilce, M. C. J. & Parker, M. W. (1994). *Biochim. Biophys. Acta*, **1205**, 1–18.

Pergamon

Insect Biochemistry and Molecular Biology •• (2001) ••–••

Insect Biochemistry and Molecular Biology

www.elsevier.com/locate/ibmb

Expression and characterization of a novel class of glutathione S-transferase from *Anopheles dirus*

R. Udomsinprasert, A.J. Ketterman *

Institute of Molecular Biology and Genetics, Mahidol University, Salaya campus, Nakhon Pathom 73170, Thailand

Received 28 December 2000; received in revised form 26 June 2001; accepted 29 June 2001

Abstract

A new *Anopheles dirus* glutathione S-transferase (GST) has been obtained and named adGST4-1. Both genomic DNA and cDNA for heterologous expression were acquired. The genomic sequence was 3188 bp and consisted of the GST gene as well as flanking sequence. The flanking sequence was analyzed for possible regulatory elements that would control gene expression. In *Drosophila* several of these elements have been shown to be involved in development and cell differentiation. The deduced amino acid sequence has low identity compared with the four alternatively spliced enzymes, adGST1-1 to 1-4, from another *An. dirus* GST gene *adgst1AS1*. The percent identities are 30–40% and 11–12% comparing adGST4-1 to insect GSTs from Delta and Sigma classes, respectively. Enzyme characterization of adGST4-1 shows it to be distinct from the other *An. dirus* GSTs because of low enzyme activity for customary GST substrates including 1-chloro-2, 4-dinitrobenzene (CDNB). However, this enzyme has a greater affinity of interaction with pyrethroids compared to the other *An. dirus* GSTs. © 2001 Published by Elsevier Science Ltd.

Keywords: Glutathione transferase; Anopheles dirus; Gene regulation; Promoter

1. Introduction

The glutathione S-transferases (GST) are a super-family of dimeric enzymes which currently has at least 25 possible families of GST-like proteins (Snyder and Maddison, 1997). Based on their sequences, the mammalian GSTs can be divided into seven distinct classes termed Alpha, Mu, Pi, Sigma, Theta, Zeta and Omega (Hayes and Pulford, 1995; Board et al., 1997, 2000). The insect GSTs can be grouped into two distinct classes termed Class I or Delta class and Class II or Sigma class (Toung et al., 1990 Fournier et al., 1992). Generally amino acid sequence identity within a class is 50% or greater, while inter-class identity is less than 30% (Mannervik et al., 1992). Therefore, the GST classes span multiple species with enzymes from the same class but from different species being more similar to each other than enzymes from different classes from within a single species. Currently Delta class has only been reported in insects although Sigma class GSTs have also been reported from cephalopods (Ji et al., 1995; Board et al., 1997).

Previously we have reported *adgst1AS1* which is an *Anopheles dirus* alternatively spliced GST gene (Pongjaroenkit et al., 2001). This gene encoded four Delta class GST enzymes, adGST1-1, 1-2, 1-3, and 1-4, that we had heterologously expressed and characterized (Jirajaroenrat et al., 2001 Ketterman et al., 2001). Although splice products from the same gene, the four enzymes possessed distinct enzyme kinetic properties for substrates and inhibitors including insecticides. Several allelic variants for one of the spliced products, adGST1-1, each had single amino acid changes outside the active site that significantly affected kinetic properties of the enzymes (Ketterman et al., 2001). Molecular modeling showed that the single residue change appeared to modulate the conformations attainable by the different variants.

In this report, we describe a novel *An. dirus* GST gene including putative regulatory elements. The gene encodes a protein from a new class of insect GST that we heterologously expressed and characterized. This enzyme appears to possess little activity for customary GST substrates and may be regulated by several

* Corresponding author. Fax: +66-2-441-9906.
E-mail address: frakt@mahidol.ac.th (A.J. Ketterman).

0965-1748/01/$ - see front matter © 2001 Published by Elsevier Science Ltd.
PII: S0965-1748(01)00119-9

elements that have been shown in *Drosophila* to be involved in development and cell differentiation.

2. Materials and methods

2.1. An. dirus genomic DNA sequencing

The recombinant bacteriophage, derived from an *An. dirus* genomic library (Pongjaroenkit et al., 2001) was partially digested with *Sal*I and *Xho*I. The DNA hybridization was performed by using the Digoxigenin (DIG)-labeled 5′ part of *adgst1-1*, the first 200 bp of the coding sequence, as the probe. The probe preparation and hybridization procedure were described previously (Boehringer Mannheim). Positive signal DNA were subcloned into pBluescript II SK(+). The contiguous sequence of 3188 bp was obtained by automated sequencer (ABI PRISM™ 377, Perkin–Elmer) and assembled using BioEdit software.

2.2. RT-PCR and cDNA sequencing

Total RNA of fourth instar larvae of *An. dirus* was extracted by TRIzol™ Reagent. This RNA was used as the template to synthesize first strand cDNA. PCR was performed by using primers 5′CCGAGAGCATATGGATTACTACTACAGCCTC3′ and 5′CCGAGAGCATATGTCACTTTC-GGCTCGCGAC3′ and 300–500 ng of cDNA as template. Optimal PCR conditions (40 cycles of 94°C for 30 s, 62°C for 30 s, 72°C for 1 min) were carried out in a Perkin–Elmer thermocycler to generate the coding sequence of adGST4-1. A single product of 630 bp was obtained and subcloned into pET3a (Novagen) by the *Nde*I restriction site contained in the primers, underlined above. The DNA sequencing was performed in both directions several times. The expression of the recombinant proteins was performed in *E. coli* BL21(DE3)pLysS.

2.3. Preparation of recombinant protein

The protein expression was performed as previously described (Ketterman et al., 2001). The soluble target protein was purified by HiTrap affinity columns (glutathione ligand coupled via a 10-carbon linker arm) as described in the user's instructions (Amersham Pharmacia Biotech). The bound proteins were eluted with 10 mM reduced-glutathione. The fractions containing active enzymes were concentrated using centriprep-10 ultrafiltration units (Amicon) by centrifugation at 2500*g* for 3 h, at 4°C and passed through Hitrap desalting columns (Amersham Pharmacia Biotech) to remove the glutathione. The enzymes were stored in 50 mM phosphate buffer (pH 6.5), 10 mM DTT, 40% (v/v) glycerol at −20°C. Protein was assayed by the method of Bradford using the Bio-Rad protein reagent with BSA as the standard protein (Bradford, 1976). The purity and subunit size of the enzyme preparations were confirmed by SDS-PAGE with Bio-Rad broad-range standards as molecular mass markers.

2.4. Characterization of expressed enzyme

The method for determination of GST activity with 1-chloro-2, 4-dinitrobenzene (CDNB) was described previously (Habig et al., 1974). The activity with 3 mM CDNB and 16 mM glutathione was monitored at 340 nm using a SpectraMax 250 (Molecular Devices) in 0.1 M phosphate buffer pH 6.5. This is the standard assay used for the adGST4-1 enzyme.

The kinetic studies were performed by varying the concentration of CDNB from 0.1 to 3.2 mM and glutathione from 0.25 to 20 mM. The V_{max} and K_m were determined by non-linear regression analysis using GraphPad Prism 2.01 software.

The inhibition studies were performed with the standard assay in the presence of various compounds as inhibitors. Determinations of IC_{50} were performed with ethacrynic acid and S-hexylglutathione by varying the inhibitor concentrations as previously described (Prapanthadara et al., 1996).

3. Results

3.1. Isolation and cloning of adgst 4-1

Several positive signal fragments were detected after the recombinant bacteriophage was double digested with restriction enzymes, *Sal*I and *Xho*I. These fragments were subcloned into pBluescriptII SK$^+$ and sequenced. The program BLAST was used to analyze the 3188 bp contiguous sequence. The full-length gene contained two coding exons and one 59 bp intron (Fig. 1). The 794 bp down stream sequence was also analyzed for other GST coding sequences. The BLAST program was used to determine whether there were more coding exons that might yield multiple alternatively spliced products as previously observed (Pongjaroenkit et al., 2001). To determine the GST classification, the translated amino acid sequence was compared to other insect GSTs in the Genbank database by using BLAST search programs (Fig. 2). The amino acid sequence alignment was also performed to generate percent identities and similarities (Table 1). The percent identities are 33–43% and 11–12% for adGST4-1 and insect GSTs Delta class (Class I) and Sigma class (Class II), respectively. However, the percent similarities with GST Delta class were about 50–60%. Sequence alignment with other *An. dirus*, adGST1-1, 1-2, 1-3, and 1-4, showed adGST4-1 to have sequence

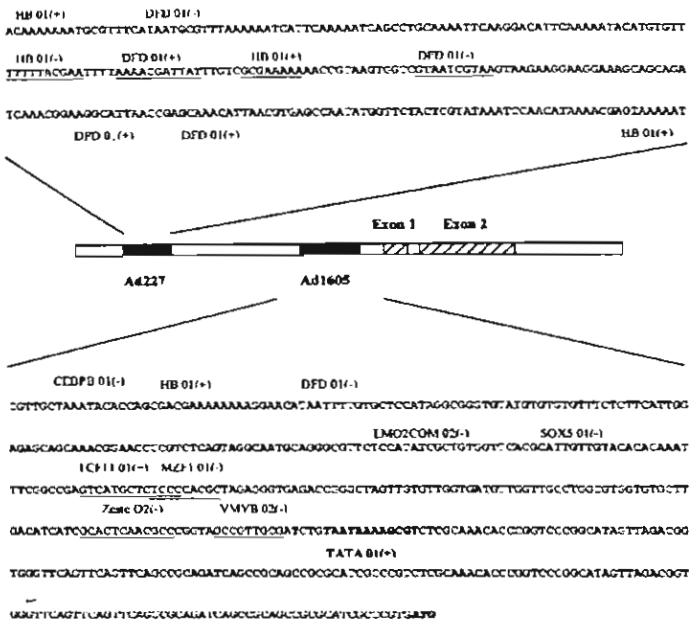

Fig. 1. A schematic of the *adgst4-1* gene isolated from an *An. dirus* genomic library. The genomic sequence for *adgst4-1* of 3188 bp is available in Genbank accession number AY014405. The two putative promoters, distal and proximal, are shown as black boxes with the protein coding exons shown as hatched boxes. Identified regulatory elements are underlined and labeled. The TATA box is shown in bold. The start site for translation is shown in bold italics.

variation even in the N-terminus which is highly conserved within a class (Fig. 2).

3.2. Promoter prediction

The 1702 bp upstream sequence of *adgst4-1* was analyzed and characterized by using a combination of two programs, TSSW (human RNA polymerase II recognition using the TRANSFAC database; http://dot.imgen.bmc.tmc.edu:9331/gene-finder/gf.html) and the MatInspector program (www.gsf.de2cgi-bin/matsearch.pl). The promoter regions were predicted by the TSSW program and the putative elements were predicted by the MatInspector program. Two putative promoters located at positions 227 and 1605 were analyzed and named Ad227 and Ad1605, respectively (Fig. 1). A TATA binding site was determined at position 1605 upstream of the coding gene. The proximal promoter Ad1605 was identified as the *adgst4-1* promoter. This putative promoter contained nine different regulatory protein binding sites (Fig. 1). These binding sites or regulatory elements may control expression of the GST gene in a tissue or stage specific manner. Additionally, the distal Ad227 promoter contained multiple Hunchback and *Dfd* recognition sites.

3.3. Expression and purification of adGST4-1

A 630 bp product from mRNA was obtained by PCR using two primers specific to the 5′ and 3′ end of the genomic sequence. The PCR product contained only coding sequence. This cDNA was subcloned into pET3a expression vector to produce a GST with a subunit size of 24,237 Da. The yield obtained after purification by glutathione ligand affinity chromatography was about 35% of the total *E. coli* lysate activity or approximately 20 mg l^{-1} of culture. During the purification of the recombinant enzyme it was observed that this GST possessed a lower affinity of interaction with the affinity column compared to other *An. dirus* GSTs in our laboratory. However, a high degree of purification, >99%, could be obtained as shown by SDS-PAGE (Fig. 3).

3.4. Characterization of adGST4-1 recombinant protein

The kinetic parameters of adGST4-1 for GSH and CDNB were determined (Table 2). Comparison of adGST4-1 with other *An. dirus* GSTs (adGST1-1, 1-2, 1-3, and 1-4) showed the adGST4-1 K_m for glutathione and CDNB was relatively high indicating low binding affinities for these substrates (Table 2). The V_m and K_m data for adGST4-1 could only be estimated because of the limitation of CDNB solubility. This also makes the k_{cat} an estimate and it is shown only for purposes of comparison. Although these kinetic parameters are estimates due to physical limitations, the values obtained were reproducible as shown in Table 2. The plots of V versus S also show the data obtained were approaching saturation thereby contributing to the reproducibility (Fig. 4). The parameters obtained, k_{cat}, k_{cat}/K_{mGSH} and k_{cat}/K_{mCDNB}, are relatively low compared to the other *An. dirus* GSTs indicating a very slow rate of turnover for CDNB and GSH. Several other substrates were used to determine activity including 1,2-dichloro-4-nitrobenzene, *p*-nitrophenethyl bromide and ethacrynic acid. However, adGST4-1 had no detectable activity for these substrates.

3.5. Inhibition study of adGST4-1

The inhibition of several compounds on CDNB conjugating activity of adGST4-1 is shown in Table 3. All compounds inhibited CDNB conjugating activity of the different *An. dirus* GSTs although to different extents. The extent of each insecticide inhibition appears to be very similar for each enzyme except the inhibition by permethrin and λ-cyhalothrin of adGST4-1 shows a significant difference from the other three adGSTs enzymes. For adGST4-1, ethacrynic acid was a good inhibitor of CDNB conjugating activity.

The IC_{50} for ethacrynic acid and S-hexylglutathione,

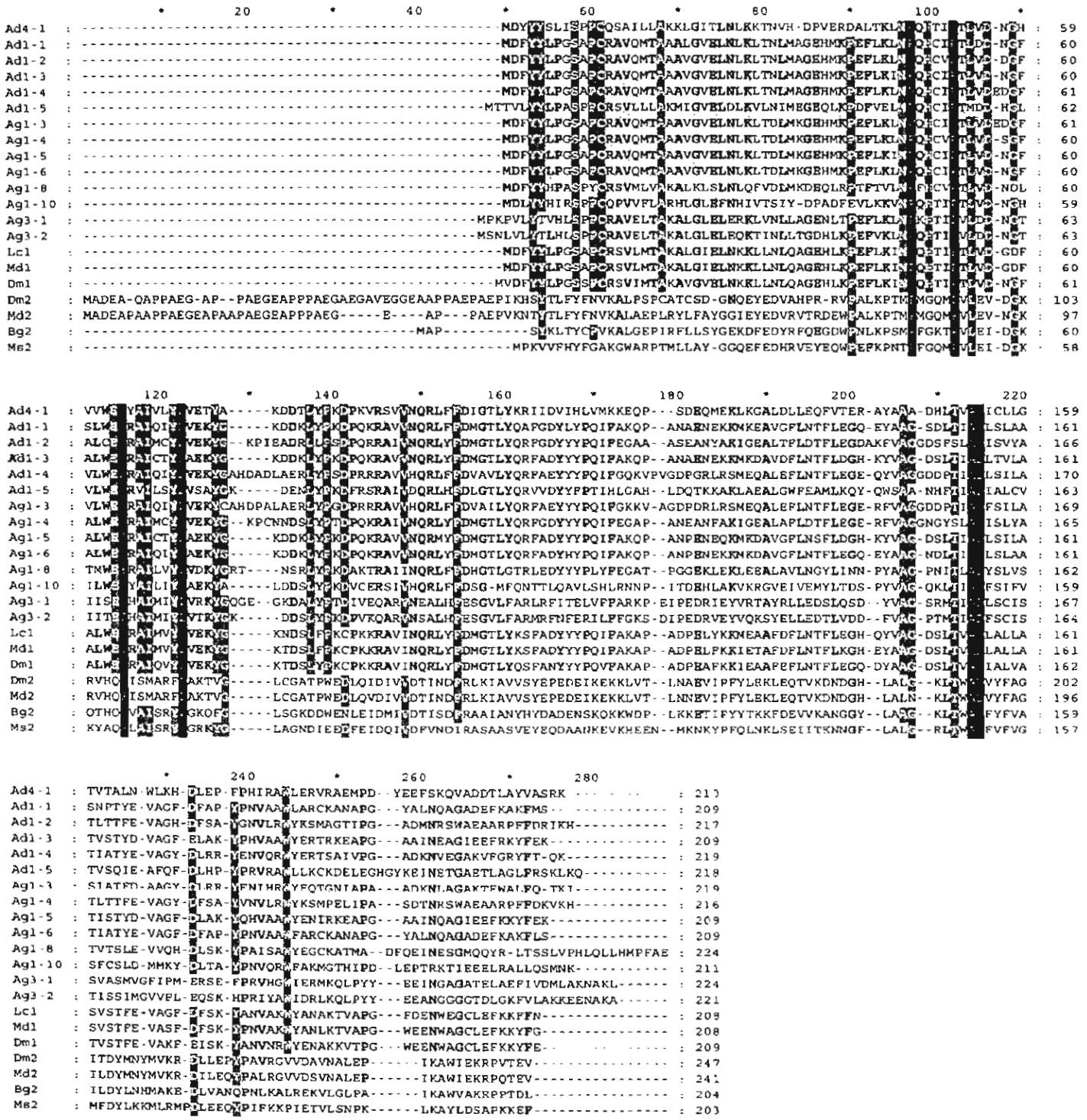

Fig. 2. The deduced amino acid sequence alignment of multiple adGSTs gene. Shading indicates degree of conservation of residue where black is 100% conserved and dark grey is >80% conserved and light grey is 60–80% conserved. The figure was produced by GeneDoc version 2.5.

23.2±3.9 μM and 126±30 μM, respectively, were determined by interpolation from the plots of fractional activity (y) versus log inhibition concentration (log [I]). This value is 100- and 13-fold greater than what was observed for adGST1-1 (Prapanthadara et al., 1996).

4. Discussion

4.1. Promotor prediction

Previously we have reported an *An. dirus* GST gene that was alternatively spliced to generate four different

Table 1
Comparison of the amino acid sequence of adGST4-1 with other insect GSTs

	Ad1-1	Ad1-2	Ad1-3	Ad1-4	Ad1-5	Ag1-3	Ag1-4	Ag1-5	Ag1-6	Ag1-8	Ag1-10	Ag3-1	Ag3-2	Lc1	Md1	Dm1	Dm2	Md2	Bg2	Ms2
	43%	33%	42%	35%	37%	33%	34%	40%	42%	37%	41%	29%	30%	38%	40%	39%	14%	15%	17%	13%
Ad4-1	60%	55%	60%	55%	58%	53%	57%	59%	61%	56%	66%	49%	52%	58%	58%	59%	27%	29%	30%	26%
		61%	77%	63%	44%	60%	62%	76%	91%	41%	38%	38%	36%	67%	68%	65%	13%	12%	16%	15%
Ad1-1		74%	87%	73%	63%	72%	74%	86%	95%	59%	54%	53%	54%	78%	78%	77%	26%	26%	29%	27%
			63%	62%	40%	60%	85%	61%	61%	40%	31%	32%	33%	58%	56%	56%	10%	9%	14%	14%
Ad1-2			78%	74%	59%	72%	91%	79%	75%	57%	53%	50%	50%	73%	72%	72%	22%	23%	29%	28%
				64%	42%	60%	62%	92%	80%	40%	37%	36%	36%	70%	70%	70%	13%	12%	16%	14%
Ad1-3				77%	62%	76%	78%	97%	90%	62%	55%	55%	56%	82%	81%	81%	25%	26%	30%	26%
					39%	84%	60%	63%	65%	37%	33%	34%	34%	57%	55%	57%	12%	10%	15%	14%
Ad1-4					56%	90%	74%	76%	75%	55%	50%	52%	54%	73%	73%	74%	22%	23%	29%	25%
						38%	40%	41%	43%	41%	32%	32%	34%	40%	41%	41%	9%	9%	12%	12%
Ad1-5						57%	62%	63%	65%	58%	50%	52%	56%	59%	59%	59%	20%	23%	25%	24%
							61%	62%	62%	36%	35%	33%	33%	54%	52%	54%	11%	10%	16%	14%
Ag1-3							74%	76%	75%	55%	51%	51%	53%	72%	71%	72%	21%	22%	29%	24%
								63%	65%	42%	34%	33%	34%	59%	57%	58%	11%	10%	14%	13%
Ag1-4								80%	76%	58%	56%	52%	51%	74%	73%	74%	22%	23%	28%	26%
									82%	40%	37%	34%	35%	68%	68%	68%	13%	12%	16%	14%
Ag1-5									91%	62%	53%	54%	56%	82%	81%	82%	24%	26%	29%	26%
										41%	38%	36%	36%	67%	68%	67%	13%	12%	16%	15%
Ag1-6										61%	55%	53%	53%	80%	81%	80%	25%	25%	29%	27%
											30%	31%	32%	41%	42%	41%	10%	10%	14%	14%
Ag1-8											52%	48%	52%	60%	60%	60%	23%	24%	28%	25%
												30%	30%	32%	32%	32%	12%	12%	18%	14%
Ag1-10												49%	50%	55%	55%	54%	25%	28%	32%	27%
													65%	35%	36%	35%	10%	10%	14%	14%
Ag3-1													83%	54%	55%	56%	22%	24%	30%	27%
														34%	35%	36%	9%	9%	13%	12%
Ag3-2														56%	56%	56%	23%	24%	30%	27%
															92%	82%	9%	8%	14%	13%
Lc1															96%	92%	23%	24%	29%	26%
																84%	10%	9%	15%	14%
Md1																92%	23%	25%	30%	28%
																	10%	10%	15%	13%
Dm1																	24%	26%	30%	28%
																		81%	38%	30%
Dm2																		84%	53%	44%
																			40%	32%
Md2																			58%	49%
																				40%
Bg2																				59%

The top number in each cell represents the percent amino acid sequence identities and the bottom number represents the percent similarity. The amino acid sequences were obtained from the GenBank database. The sequences are Ad4-1 (AY014406), Ad1-1 (AF273041), Ad1-2 (AF273038), Ad1-3 (AF273039), Ad1-4 (AF273040), Ad1-5 (AF251478), Ag1-3 (AAC79999.1), Ag1-4 (AAC79998), Ag1-5 (Q93112), Ag1-6 (Q93113), Ag1-8 (AF316637), Ag1-10 (AF316638), Ag3-1 (AF316635), Ag3-2 (AF316636), Lc1 (P42860), Md1 (P28338), Dm1 (P20432), Md2 (P46437), Dm2 (P41043), Bg2 (O18598), Ms2 (P46429). The abbreviations are Ad for *Anopheles dirus*, Ag for *Anopheles gambiae*, Lc for *Lucilia cuprina*, Md for *Musca domestica*, Dm for *Drosophila melanogaster*, Bg for *Blattella germanica*, Ms for *Manduca sexta*

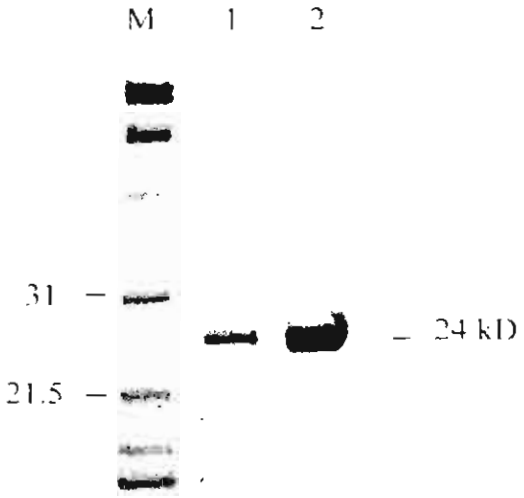

Fig. 3. Purification of recombinant protein adGST4-1. Lane M is molecular weight markers as shown in dD. Lane 1 is 3 μg and Lane 2 is 6 μg of the purified recombinant adGST 4-1.

protein products (Pongjaroenkit et al., 2001). We have now obtained a novel GST gene that codes for a single expressed protein product which we name adGST4-1. The 3188 bp genomic sequence contained two coding exons and 1702 bp 5′ flanking sequence that was analyzed for promoters and regulatory elements which may control expression of this gene. Response elements are short conserved sequences that regulate expression of a gene. Several of these elements were identified in two possible promoters (Fig. 1). Most putative elements contained in the Ad227 and Ad1605 promoters have been shown in other species to be functionally involved in developmental stage regulation as well as responding to xenobiotic modulation.

Although, GSTs are involved in protecting an organism from toxic and mutagenic xenobiotics, it has been reported that the over-expression of the Pi class GST has been associated with tumor development and carcinogenesis (Batist et al., 1986). Therefore, understanding the transcriptional regulatory mechanism of these genes is of interest. Several regulating elements of GSTs have been identified. In mammals, the GSTP1 promoter included a putative AP-1 response element as well as a negative regulatory element in a multidrug resistant derivative of a human mammary carcinoma cell line (Moffat et al., 1994). Sp1 binding sites, the GC box motif, have also been shown to play a role in regulating basal levels of GSTP1 transcription (Moffat et al., 1996). The rat GST-P gene is regulated by two enhancers and a silencer (Sakai et al., 1988). The protein bound to the silencer sequence belongs to the CCAATT/enhancer-binding protein (C/EBP) family (Osada et al., 1995). A rat Alpha class GST has been shown to be negatively regulated by C/EBP protein interaction with an antioxidant/electrophile response element (ARE/EpRE) in vascular smooth muscle cells to function as oxidative stress protection for blood vessels (Chen and Ramos, 2000).

TCF11 is a widely expressed transcription factor that binds to a subclass of AP-1 sites. The complexes of TCF11/LCR-F1/Nrf1 form heterodimers with a small Maf protein to increase stringency of specific binding to the NF-E2 site, the antioxidant response element and the heme-responsive element, and contribute to negative regulation of this specific target site (Johnsen et al., 1998). In human Pi-class GST, disruption of a putative AP-1 response element (Xia et al., 1991) within the *GSTP1* promoter abrogated *GSTP1* transcription while increased levels of GSTP1 transcription can play a major role in regulating overexpression of GSTP1-1 in multidrug-resistant cell line (Moffat et al., 1994).

Many of the putative elements contained in both *adgst4-1* promoters are DNA-binding regions for transcription factors expressed during a developmental or cell differentiation stage (Martinez-Arias and Lawrence, 1985; Siegfried and Perkins, 1990; Blair, 1994; Stauber et al., 2000). The data suggest that adGST4-1, a phase II detoxication enzyme, is expressed and regulated by these putative elements. The distal Ad227 promoter located upstream of the Ad1605 promoter may act as an enhancer/repressor to regulate the expression of adGST4-1 as has been reported for a Pi class rat GST (Sakai et al., 1988).

Table 2
Kinetic parameters of *Anopheles dirus* GSTs (the data are the mean±standard error of at least three separate experiments)

Kinetic parameters	Ad4-1	Ad1-1	Ad1-2	Ad1-3	Ad1-4
V_m (μmol^{-1} min^{-1} mg)	2.2±0.3	12.9±0.6	63.9±3.50	67.5±1.97	40.3±1.89
K_{mGSH} (mM)	1.8±0.4	0.86±0.2	1.30±0.15	0.40±0.05	0.83±0.08
K_{mCDNB} (mM)	5.3±0.8	0.10±0.03	0.21±0.03	0.10±0.01	0.52±0.67
k_{cat} (s^{-1})	0.9	5.03	25.9	26.9	16.9
k_{cat}/K_{mGSH} (mM^{-1} s^{-1})	0.5	5.86	20	67	20
k_{cat}/K_{mCDNB} (mM^{-1} s^{-1})	0.2	48.4	121	269	32

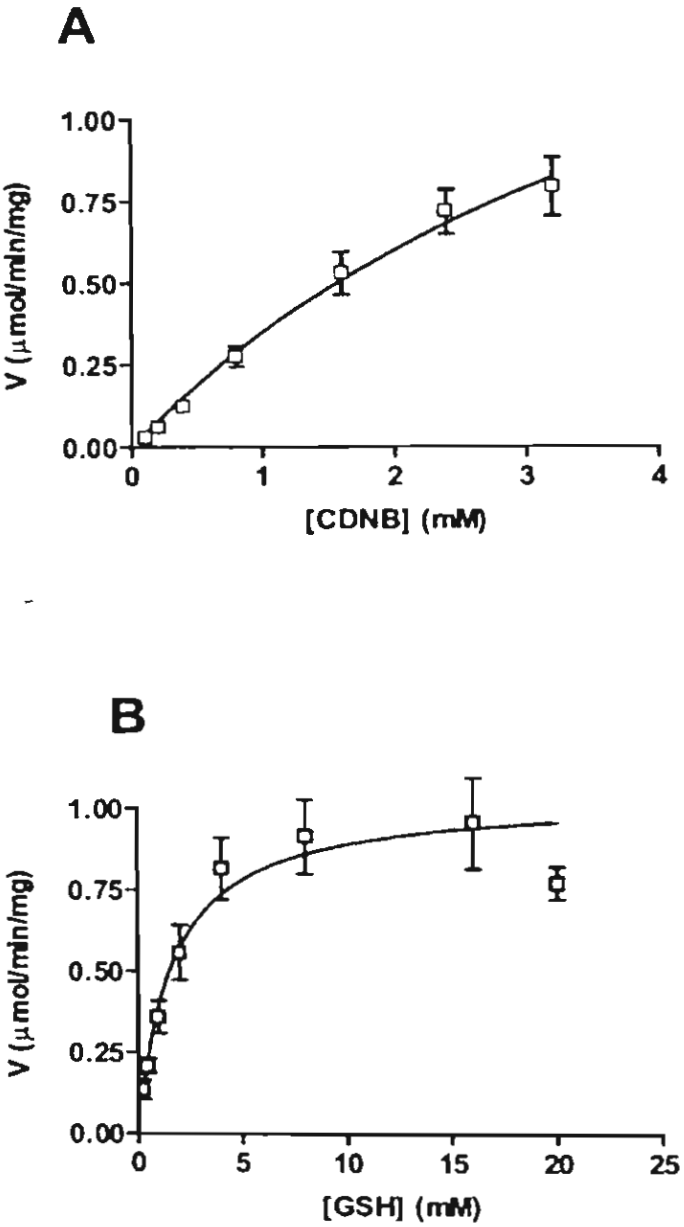

Fig. 4. Plot of velocity versus substrate concentration for kinetic constant determination of adGST4-1. (A) The glutathione concentration was fixed at 16 mM and the CDNB concentration was varied from 0.1 to 3.2 mM. (B) The CDNB concentration was fixed at 3 mM and the glutathione concentration was varied from 0.25 to 20 mM. The kinetic constants were calculated by non-linear regression. The data shown are means±SEM error bars from four independent experiments.

4.2. Protein characterization

As suggested by amino acid sequence identities and similarities compared with other insect GSTs Class I and Class II (Table 1), adGST4-1 is very distinctive. The percent similarity of 50–60% shows that adGST4-1 is more similar to insect GST Class I or Delta class. Comparison of amino acid sequence with other *An. dirus* GSTs, 1-1 to 1-4 (Fig. 2), shows adGST4-1 has a high variation in the N-terminal, contributing to the low affinity binding with the GSH ligand during purification.

Studies of the enzymatic properties indicate that adGST4-1 is very different from the other known adGSTs (Table 2). The model substrate, CDNB, could be turned over only at a slow rate as described by the very low k_{cat} and k_{cat}/K_{mCDNB}. This enzyme has behavior similar to several mammalian GSTs which lack activity toward CDNB such as GST Theta class (Meyer et al., 1991), GST Zeta class (Board et al., 1997) and the new GST Omega class (Board et al., 2000). While the other *An. dirus* GSTs are more similar to mammalian GSTs in class Alpha (Schramm et al., 1984; Stenberg et al., 1991), Mu (Schramm et al., 1984; Vorachek et al., 1991) and Pi (Widersten et al., 1992). The inhibition studies (Table 3) shows that the CDNB activity of adGST4-1 can be inhibited by several compounds. Although there was no detectable enzyme activity the inhibition indicates that there is some interaction between adGST4-1 and these compounds especially for ethacrynic acid which has been shown to bind to an effector site in GST Pi class as well (Phillips and Mantle, 1993). The inhibition of CDNB conjugation by pyrethroid compounds, permethrin and λ-cyhalothrin, was also obviously different for adGST4-1 compared to other adGSTs indicating a greater affinity of interaction. The greater IC_{50} value for ethacrynic acid and S-hexylglutathione for adGST4-1 compared to the previous data for adGST1-1 (Prapanthadara et al., 1995) indicate a lower interaction with these compounds and show the inhibition characteristics of a homodimer or single affinity site for interaction (Tahir and Mannervik, 1986). Despite low CDNB conjugating activity, adGST4-1 may possess some activity for a compound that is significant in a metabolism pathway as shown by other GST classes. For example, recombinant human omega class GST (GSTO1-1) exhibited a glutathione-dependent thiol transferase activity and catalyzed glutathione-dependent reduction of dehydroascorbate (Board et al., 2000) or the novel function of human Theta GSTT2-2 with 1-menaphthylsulphate demonstrating it to be a glutathione-dependent sulphatase (Tan et al., 1996). In conclusion, the recent reports on diverse roles of GSTs in regulation of Jun N-terminal kinase (Adler et al., 1999) or in tyrosine catabolism (Dixon et al., 2000) indicate that there may be GST proteins with little traditional GST activity but having other physiological functions still to be elucidated.

Acknowledgements

This work was supported by the Thailand Research Fund.

Table 3
Inhibition of *Anopheles dirus* recombinant GSTs CDNB activity by various compounds (the GSH and CDNB concentrations were 16 and 3 mM, respectively. The data are the mean±standard deviation of at least three separate experiments, each of which was performed in duplicate)

Compounds	Concentration	% Inhibition			
	(mM)	Ad4-1	Ad1-2	Ad1-3	Ad1-4
DCNB	0.1	**29.8±0.2**	21.1±12.0	9.3±2.5	28.3±6.6
p-Nitrophenyl bromide	0.1	**22.5±2.2**	7.1±12.3	25.5±4.2	31.0±6.4
Cumene hydroperoxide	2.5	**43.6±5.7**	51.9±9.1	17.8±7.3	30.1±5.1
p-Nitrobenzyl chloride	1.0	**16.5±3.7**	29.1±10.5	46.1±2.5	49.8±6.7
Ethacrynic acid	0.001	**8.6±2.1**	74.7±3.0	30.9±4.5	34.5±5.3
	0.01	**39.5±5.0**	97.9±0.4	79.2±2.3	77.6±3.0
	0.1	**88.5±1.5**	100±0.0	100±0.0	100±0.0
Permethrin	0.01	**11.8±3.4**	47.4±8.5	17.4±4.7	66.3±14.4
	0.1	**28.0±3.6**	89.0±0.8	100±0.0	100±0.0
λ-cyhalothrin	0.01	**15.2±5.1**	40.1±11.3	25.0±11.8	91.5±7.6
	0.1	**25.6±1.0**	86.3±3.9	100±0.0	100±0.0

References

Adler, V., Yin, Z., Fuchs, S.Y., Benezra, M., Rosario, L., Tew, K.D., Pincus, M.R., Sardana, M., Henderson, C.J., Wolf, C.R., Davis, R.J., Ronai, Z., 1999. Regulation of JNK signaling by GSTp. EMBO J. 18, 1321–1334.

Batist, G., Tulpule, A., Sinha, B.K., Katki, A.G., Myers, C.E., Cowan, K.H., 1986. Overexpression of a novel aniotic glutathione transferase in multidrug-resistant. J. Biol. Chem. 261, 15544–15549.

Blair, S.S., 1994. A role for the segment polarity gene shaggy-zeste white 3 in the specification of regional identity in the developing wing of Drosophila. Dev. Biol. 162, 229–244.

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

Board, P.G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L.S., Schulte, G.K., Danley, D.E., Hoth, L.R., Griffor, M.C., Kamath, A.V., Rosner, M.H., Chrunyk, B.A., Perregaux, D.E., Gabel, C.A., Geoghegan, K.F., Pandit, J., 2000. Identification, characterization, and crystal structure of the omega class glutathione transferases. J. Biol. Chem. 275, 24798–24806.

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

Chen, Y.-H., Ramos, K.S., 2000. A CCAAT/Enhancer-binding protein site within antioxidant/electrophile response element along with CREB-binding protein participate in the negative regulation of rat *GST-Ya* gene in vascular smooth muscle cells. J. Biol. Chem. 275, 27366–27376.

Dixon, D.P., Cole, D.J., Edwards, R., 2000. Characterisation of a zeta class glutathione transferase from *Arabidopsis thaliana* with a putative role in tyrosine catabolism. Arch. Biochem. Biophys. 384, 407–412.

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

Habig, W.H., Pabst, M.J., Jakoby, W.B., 1974. Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. J. Biol. Chem. 249, 7130–7139.

Hayes, J.D., Pulford, D.J., 1995. The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. CRC Crit. Rev. Biochem. Mol. Biol. 30, 445–600.

Ji, X., Von Rosenvinge, E.C., Johnson, W.W., Tomarev, S.I., Paitigorsky, J., Armstrong, R.N., Gilliland, G.L., 1995. Three-dimensional structure, catalytic properties, and evolution of a sigma class glutathione transferase from squid, a progenitor of the lens S-crystallins of cephalopods. Biochemistry 34, 5317–5328.

Jirajaroenrat, K., Pongjaroenkit, S., Krittanai, C., Prapanthadara, L., Ketterman, A.J., 2001. Heterologous expression and characterization of alternatively spliced glutathione S-transferases from a single *Anopheles* gene. Insect Biochem. Mol. Biol. (in press).

Johnsen, O., Murphy, P., Prydz, H., Kolsto, A.-B., 1998. Interaction of the CNC-bZIP factor TCF11/LCR-F1/Nrf1 with MafG: binding-site selection and regulation of transcription. Nucleic Acids Res. 26, 512–520.

Ketterman, A.J., Prommeenate, P., Boonchauy, C., Chanama, U., Leetachewa, S., Promtet, N., Prapanthadara, L., 2001. Single amino acid changes outside the active site significantly affect activity of glutathione S-transferases. Insect Biochem. Mol. Biol. 31, 65–74.

Mannervik, B., Awasthi, 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., Wolf, C.R., 1992. Nomenclature for human glutathione transferases. Biochem. J. 282, 305–306.

Martinez-Arias, A., Lawrence, P.A., 1985. Parasegments and compartments in *Drosophila* embryo. Nature 313, 639–642.

Meyer, D.J., Coles, B., Pemble, S.E., Gilmore, K.S., Fraser, G.M., Ketterer, B., 1991. Theta, a new class of glutathione transferases purified from rat and man. Biochem. J. 274, 409–414.

Moffat, G.J., McLaren, A.W., Wolf, C.R., 1994. Involvement of Jun and Fos proteins in regulating transcriptional activation of the human pi class glutathione S-transferase gene in multidrug-resistant MCF7 breast cancer cells. J. Biol. Chem. 269, 16397–16402.

Moffat, G.J., McLaren, A.W., Wolf, C.R., 1996. Sp1-mediated transcriptional activation of the human pi class glutathione S-transferase promoter. J. Biol. Chem. 271, 1054–1060.

Osada, S., Takahashi, K., Nishihira, T., Suzuki, T., Muramatsu, M., Imagawa, M., 1995. CCAAT/enhancer-binding proteins α and β interact with the silencer element in the promoter of glutathione S-transferase P gene during hepatocarcinogenesis. J. Biol. Chem. 270, 31288–31293.

Phillips, M.F., Mantle, T.J., 1993. Inactivation of mouse liver glutathione S-transferase YfYf (Pi class) by ethacrynic acid and 5,5′-dithiobis-(2-nitrobenzoic acid). Biochem. J. 294, 57–62.

Pongjaroenkit, S., Jirajaroenrat, K., Boonchauy, C., Chanama, U., Leetachewa, S., Prapanthadara, L., Ketterman, A.J., 2001. Genomic

organization and putative promotors of highly conserved glutathione S-transferases originating by alternative splicing in *Anopheles dirus*. Insect Biochem. Mol. Biol. 31, 75–85.

Prapanthadara, L., Ketterman, A.J., Hemingway, J., 1995. DDT-resistance in *Anopheles gambiae* Giles from Zanzibar Tanzania based on increased DDT-dehydrochlorinase activity of glutathione S-transferases. Bull. Entomol. Res. 85, 267–274.

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

Sakai, M., Okuda, A., Muramatsu, M., 1988. Multiple regulatory elements and phorbol 12-O-tetradecanoate 13-acetate responsiveness of the rat placental glutathione transferase gene. Proc. Natl. Acad. Sci. U.S.A. 85, 9456–9460.

Schramm, V.L., McCluskey, R., Emig, F.A., Litwack, G., 1984. Kinetic studies and active site-binding properties of glutathione S-transferase using spin-labeled glutathione, a product analogue. J. Biol. Chem. 259, 714–722.

Siegfried, E., Perkins, L.A., 1990. Putative protein kinase product of the *Drosophila* segment-polarity gene *zeste-white3*. Nature 345, 825–829.

Snyder, M.J., Maddison, D.R., 1997. Molecular phylogeny of glutathione-S-transferases. DNA Cell. Biol. 16, 1373–1384.

Stauber, M., Taubert, H., Schmidt-Ott, U., 2000. Function of bicoid and hunchback homologs in the basal cyclorrhaphan fly *Megaselia* (Phoridae). Proc. Natl. Acad. Sci. U.S.A. 97, 10844–10849.

Stenberg, G., Board, P.G., Carlberg, I., Mannervik, B., 1991. Effects of directed mutagenesis on conserved arginine residues in a human class alpha glutathione transferase. Biochem. J. 274, 549–555.

Tahir, M.K., Mannervik, B., 1986. Simple inhibition studies for distinction between homodimeric and heterodimeric isoenzymes of glutathione transferase. J. Biol. Chem. 261, 1048–1051.

Tan, K.-L., Chelvanayagam, G., Parker, M.W., Board, P.G., 1996. Mutagenesis of the active site of the human theta-class glutathione transferase GSTT2-2: catalysis with different substrates involves different residues. Biochem. J. 319, 315–321.

Toung, Y.P.S., Hsieh, T.-S., Tu, C.P.D., 1990. *Drosophila* glutathione S-transferase 1-1 shares a region of sequence homology with the maize glutathione S-transferase III. Proc. Natl. Acad. Sci. U.S.A. 87, 31–35.

Vorachek, W.R., Pearson, W.R., Rule, G.S., 1991. Cloning, expression, and characterization of a class-mu glutathione transferase from human muscle, the product of the GST4 locus. Proc. Natl. Acad. Sci. U.S.A. 88, 4443–4447.

Widersten, M., Kolm, R.H., Björnestedt, R., Mannervik, B., 1992. Contribution of five amino acid residues in the glutathione binding site to the function of human glutathione transferase P1-1. Biochem. J. 285, 377–381.

Xia, C.L., Cowell, I.G., Dixon, K.H., Pemble, S.E., Ketterer, B., Taylor, J.B., 1991. Glutathione transferase π: its minimal promoter and downstream cis-acting element. Biochem. Biophys. Res. Commun. 176, 233–240.

06-08-01 15:17.04 REV 16.02x IB$$$$096P