Glutathione S-transferase (GST) of American Cockroach, *Periplaneta americana*: Classes, Isoforms, and Allergenicity

Nitat Sookrung1,2, Onrapak Reamtong3, Rojana Poolphol4,5, Nitaya Indrawattana6, Watee Seesuay2, Nawannaporn Saelim2,5, Pongsakorn Tantilipikorn7, Chaweewan Bunnag7, Wanpen Chaicumpa2,5 & Anchalee Tungtrongchitr2,5

Insect glutathione S-transferases (GSTs) play important roles in insecticide/drug resistance and stress response. Medically, GSTs of house dust mites (*Dermatophagoides pteronyssinus* and *Blomia tropicalis*) and German cockroach (*Blattella germanica*) are human allergens. In this study, classes, isoforms and B-cell and allergenic epitopes of GST of American cockroach, *Periplaneta americana*, the predominant species in the tropics and subtropics were investigated for the first time. Enzymatically active native and recombinant *P. americana*-GSTs bound to IgE in sera of all *P. americana* allergic patients that were tested. By gel-based proteomics and multiple sequence alignments, the native GST comprises three isoforms of delta and sigma classes. All isoforms interacted with serum IgE of the cockroach allergic subjects. Molecularly, the protein contains six B-cell epitopes; two epitopes located at β1-α1 and β4-α3 regions bound to patients’ serum IgE, indicating that they are allergenic. *P. americana* are ubiquitous and their GST can sensitize humans to allergic diseases; thus, the protein should be included in the allergen array for component resolved diagnosis (CRD) of allergic patients, either by skin prick test or specific IgE determination. The GST is suitable also as a target of environmental allergen detection and quantification for intervention of cockroach sensitization and allergic morbidity.

Cockroaches (CR) are ubiquitous. They are pestiferous source of human pathogens and allergens1. Although worldwide prevalence of the CR allergy is relatively less than the house dust mite (HDM) allergy2,3 clinical manifestations caused by CR are usually more prolonged and severe and often require emergency-room visit, hospitalization and/or intensive care4. The CR allergens are prevailed in the environment especially in the infested areas5,6. In Thailand, the principal CR species causing human allergic sensitization and morbidity is *Periplaneta americana* (American CR)7. Currently, officially recognized *P. americana* allergens include Per a 1 (a protein that exists in multiple variants containing different numbers of repeated amino acid domains) which elicit 93% skin reactivity among CR allergic patients and bound to IgE in all patients’ sera8,9; Per a 2 or aspartic protease-like protein (42 kDa) bound to IgE in sera of 63% of allergic rhinitis or asthma and rhinitis patients10; Per a 3 (Cr-PI), an insect hemolymph or insect storage protein related to arylphorin which caused skin reaction in 93% of CR allergic patients11; Per a 6 or troponin C, a 17 kDa protein that bound to IgE in sera of 54% of atopic patients12,13; Per a 7 or tropomyosin (a 37 kDa protein with high homology to other invertebrate and vertebrate tropomyosins) which bound to IgE in sera of 41% of atopic and 57% of CR allergic patients14,15; Per a 9 or arginine kinase which is a 43 kDa pan-insect protein that reacted to IgE of all CR allergic patients tested16; native troponin T (47 kDa)

1Department of Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, 10700, Thailand. 2Center of Research Excellence on Therapeutic Proteins and Antibody Engineering, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, 10700, Thailand. 3Department of Molecular Tropical Medicine and Genetics, Faculty of Tropical Medicine, Mahidol University, Bangkok, 10400, Thailand. 4Graduate Program in Clinical Pathology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, 10700, Thailand. 5Department of Parasitology, Mahidol University, Bangkok, 10700, Thailand. 6Department of Microbiology and Immunology, Faculty of Tropical Medicine, Mahidol University, Bangkok, 10400, Thailand. 7Department of Oto-rhino-laryngology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, 10700, Thailand. Correspondence and requests for materials should be addressed to A.T. (email: anchalee.tun@mahidol.ac.th)
Glutathione S-transferases (GSTs) is a large family of intracellular enzymes of aerobic prokaryotes and eukaryotes that the host organisms use for protection against oxidative stress, degenerative conditions, aging, and cancers and also detoxify endogenous and xenobiotic electrophiles such as drugs, herbicides, insecticides\(^22,23\). GSTs are classified according to their cellular localizations into three major families, i.e., cytosolic, mitochondrial/peroxisomal, and microsomal GSTs\(^24\). All insect GSTs are cytosolic and further subdivided into at least six classes including delta, epsilon, omega, sigma, theta, and zeta, based on phylogenetic analysis\(^23\). Delta and epsilon GSTs are arthropod specific\(^25\). Each GST subunit is approximately 21–28 kDa and may exist as either homo- or hetero-dimeric form\(^26\).

The specific GST protein which is found in the cockroach Der. americana is a 28 kDa protein named Bla g 5\(^36\), which is the major allergen in this species. The allergenic properties of Bla g 5 have been extensively studied\(^34,35\). A homologous GST (Bla g 5) was also identified in another cockroach species Blattella germanica\(^37\). Similar GSTs have been identified in other species of arthropods, including Alternaria alternata\(^38\), Ascaris lumbricoides and Ascaris suis\(^39,40\), the B. germanica GSTs, i.e., sigma BgGSTS1 and delta BgGSTD1 have been reported as potent human allergens\(^35,38\). Data on allergenicity and several other attributions of P. americana GSTs are lacking. Therefore in this study, P. americana GST classes, isoforms, allergenicity, and B cell epitopes were investigated.

### Materials and Methods

**Serum samples.** This study was approved by Siriraj Ethical Committee (COA no. SI268/2008), Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok. All methods were performed in accordance with the
relevant guidelines and regulations by International Ethical Guidelines for Health-related Research Involving Humans. Informed consent was obtained from each subject. Serum samples were isolated from clotted blood aliquots collected from 15 patients who visited Allergy Clinic, Department of Oto-Rhino-Laryngology, Siriraj Hospital, Bangkok. All patients were multiply positive by skin prick test (SPT) to crude *P. americana* extract and other allergens, but positivity to the *P. americana* extract was more pronounced. Serum specific IgE levels to American CR extract were measured by using ImmunoCAP (UniCAP 250, Instrument Pharmacia Diagnostic AB, Uppsala, Sweden). Serum samples of five subjects who were negative by SPT, IgE ImmunoCAP and IgE-binding ELISA to the *P. americana* and other extracts served as non-allergic (normal) controls. A pool of sera of 10 *P. americana* allergic patients was prepared by mixing 0.5 ml of individual samples.

**Preparation of recombinant GST (rGST) of *P. Americana***. Adult cockroaches were caught from houses in Bangkok. They were identified entomologically and only the *P. americana* were kept frozen at −80 °C until use. Frozen CR was ground to fine pieces in liquid nitrogen and total RNA was isolated from the powder (100 mg) by using TRIzol reagent (Invitrogen, CA, USA). After checking RNA integrity by agarose gel electrophoresis, cDNA was synthesized from the RNA and used as a PCR template for amplification of full-length GST coding sequence (*gst*). The PCR primers were designed from GenBank database (accession number AY792949; UniProt ID Q1M0Y4); forward: 5′-CCG GAT ATC ATG ACC ATC GAC TTC TAC-3′; reverse: 5′-CGA AAG CTT TCA CTT CTT GGC GAG GTT-3′. The PCR reaction mixture was: 1 μl cDNA, 10 μM each of the forward and reverse primers, 2.5 μl of 10 × PCR buffer, 3 μl of 25 mM MgCl2, 2 μl of 2.5 mM dNTP, 0.2 μl of 5 units/μl DNA polymerase (Fermentas, Lithuania), and 40.3 μl ultrapure distilled water (UDW). The thermal cycles were: initial denaturation at 94 °C, 5 minutes; 30 cycles of 94 °C for 30 seconds, 50 °C for 30 seconds, and 72 °C for 40 seconds; and final extension at 72 °C for 7 minutes. The PCR product was verified by DNA sequencing before cloning into pKRX-T (Gentaur, Belgium), subcloned into pET20b+(Novagen, Merck, Germany) expression vector, and

---

**Figure 1.** Phylogram of glutathione-S-transferases of *P. americana* and other insects. Ad, *Anopheles dirus*; Ag, *Anopheles gambiae*; Bg, *Blattella germanica*; Bm, *Bombyx mori*; Dm, *Drosophila melanogaster*; Dp, *Dermatophagoides pteronyssinus*; Lc, *Lucilia cuprina*; Lm, *Locusta migratoria*; Ms, *Manduca sexta*; Ni, *Nilaparvata lugens*; Pa, *Periplaneta americana*.
put into BL21 (DE3) E. coli. A transformed E. coli colony was grown in isopropyl β-D-1-thiogalactopyranoside (IPTG) (Affimetrix, OH, USA) conditioned-Luria-Bertani (LB) broth (Himedia, India). The recombinant protein was purified from the bacterial lysate by using HISTrap FF affinity chromatography (GE Healthcare Lifesciences) and verified by SDS-PAGE and Coomassie Brilliant Blue G-250 (CBB) staining and LC-MS/MS. Nucleotide and deduced amino acid sequences of the P. americana rGST was subjected to phylogenetic analysis together with GSTs of other closely related insects to determine percent identity and GST class.

![Figure 2. SDS-PAGE-separated patterns of P. americana native and recombinant GSTs after staining with Coomassie Brilliant Blue G-250 dye. Lane 1, nGST (~23 kDa) eluted from glutathione affinity column; lane 2, rGST (~25 kDa) purified from gst-pET23b^-transformed E. coli lysate M, protein molecular weight (MW) marker. Numbers at the left are relative molecular masses of proteins.](image)

| Proteins | Orthologous protein | Accession no. | Number of matched peptides | Protein score | Matched peptide sequence (average peptide score; APS) |
|----------|---------------------|---------------|---------------------------|---------------|------------------------------------------------------|
| nGST     | Glutathione S-transferase class delta variant 1 (Periplaneta americana) | Gi 359326557 | 5 | 249 | LYFDIGTLYHR (102) FGEYYYPIYFAK (52) TIDFYLPGSAPCR (34) VTNLMAEHLTPEFLK (29) AILSYLADQYGDLDLYPK (34) |
| rGST     | Glutathione S-transferase class delta variant 1 (Periplaneta americana) | Gi 359326557 | 7 | 356 | AIGVDNLNKL (50) FKEMCDNLAK (45) AILSYLADQYGK (52) FGEYYYPIYFAK (64) VTNLMAEHLTPEFLK (48) AILSYLADQYGDLDLYPK (62) MNPQHTIPTLDNGFCLWESR (35) |

Table 2. LC-MS/MS Mascot results of peptides generated from in-gel tryptic digestion of P. americana nGST and rGST after searching against the Swiss-Prot database.
Preparation of native *P. americana* GST (nGST). Five ml of binding buffer (phosphate buffered saline, pH 7.3) were added to dissolve the CR powder (15 mg). The preparation was sonicated (Sartorius LABSONIC® P sonicator, Germany) in ice-bath at 20 kHz, 2 minutes pulse-on, 3 minutes pulse-off for a total of 15 minutes and then centrifuged at 10,000 g at 4 °C for 15 minutes. The clear supernatant was collected, filtered through a sterile 0.45 µm filter, and protein content was quantified by Bradford’s method (Bio-Rad, CA, USA) using bovine serum albumin (BSA) standard curve for calibration. The preparation was loaded onto a GSTrap FF affinity column (GE) that had been equilibrated with the binding buffer. The column was washed thoroughly with the binding buffer to eliminate unbound materials. The column-bound protein was eluted in 1 ml-fractions with a total of 10 ml elution buffer [50 mM Tris-HCl, 20 mM reduced glutathione (Affimextrix), pH 8.0, 5 mM DTT] at a flow rate of 1–2 ml/min. Fractions containing the eluted protein were verified by SDS-PAGE and CBB staining, pooled, dialyzed against distilled water, and lyophilized. The protein was verified by LC-MS/MS.

**SDS-PAGE and Western blot analysis (WB).** SDS-PAGE and WB were performed as described previously [17]. A 4% stacking and 12.5% separating polyacrylamide gels were used in the SDS-PAGE which was performed in a Mini-PROTEAN® 3 Cell (Bio-Rad). Separated proteins in the gels were either stained or electro-blotted onto a nitrocellulose membrane (NC) for WB. For WB, empty sites on the blotted NC were blocked with 3% BSA in PBS before placing the membrane in a solution of mouse anti-6× His tag (Abcam, UK).
### Table 3. Orthologous proteins of database that contained peptides matched with peptides of proteins in gel plugs 1–5 of 2DE-separated *P. americana*-native GST.

| Gel plug no. | Access. No. | Protein | Protein score | Mass (Da) | pI | No. of peptide | m/z | Ion score | Sequence | Possible phosphorylation sites | Possible O-glycosylation |
|-------------|-------------|---------|---------------|-----------|----|----------------|-----|-----------|----------|-------------------------------|--------------------------|
| 1 | Gi 359326557 | GST class delta (P. americana) | 368 | 24606 | 6.4 | 6 | 424.5294 | 31 | FKECDNLAK | 11, 42, 56, 72, 78, 83, 85, 100, 105, 114, 145, 166, 176, and 195 | 116 |
| | | | | | | | 466.5781 | 87 | LYFDIGTVYR | | | |
| | | | | | | | 780.8726 | 62 | FGEYVYPFVYAK | | | |
| | | | | | | | 830.3934 | 65 | TIDFYLYPGSAPCR | | | |
| | | | | | | | 720.6989 | 44 | AILSYLDQYKDDSLYPK | | | |
| | | | | | | | 849.3951 | 79 | MNPQHTPLNDNGCFLWESR | | | |
| 2 | Gi 359326557 | GST class delta (P. americana) | 217 | 24606 | 6.4 | 4 | 780.8727 | 48 | FGEYVYPFVYAK | 11, 42, 56, 72, 78, 83, 85, 100, 105, 114, 145, 166, 176, and 195 | 116 |
| | | | | | | | 830.3931 | 93 | TIDFYLYPGSAPCR | | | |
| | | | | | | | 720.697 | 55 | AILSYLDQYKDDSLYPK | | | |
| | | | | | | | 849.3917 | 20 | MNPQHTPLNDNGCFLWESR | | | |
| 3 | Gi 565341529 | GST class delta (Locusta migratoria) | 96 | 23459 | 6.19 | 2 | 434.2514 | 32 | LTYFPVK | 5, 8, 24, 42, 46, 65, 97, 105, 110, 114, 151, 163, 176, and 202 | N/A |
| | | | | | | | 579.8286 | 64 | YKLTYFPVK | | | |
| 4 | Gi 359326557 | GST class delta (P. americana) | 100 | 24606 | 6.4 | 3 | 780.8705 | 42 | FGEYVYPFVYAK | 11, 42, 56, 72, 78, 83, 85, 100, 105, 114, 145, 166, 176, and 195 | 116 |
| | | | | | | | 830.3964 | 37 | TIDFYLYPGSAPCR | | | |
| | | | | | | | 720.7007 | 22 | AILSYLDQYKDDSLYPK | | | |
| 5 | Gi 359326585 | GST class delta (B. germanica) | 80 | 23377 | 6.84 | 2 | 434.2532 | 31 | LTYFPVK | 5, 8, 24, 46, 52, 62, 65, 69, 71, 79, 95, 97, 126, 131, 144, and 151 | N/A |
| | | | | | | | 486.2729 | 48 | TPVLEIDGK | | | |

### Table 4. Sequences of *P. americana* synthetic peptides 1–6 which encompassed the 6 predicted *P. americana*-GST B-cell epitopes. These peptides were used for determining B-cell and allergic (IgE-binding) epitopes.

| *P. americana* GST peptide no. | Amino acid sequence and residue numbers | Location on the *P. americana*-GST molecule |
|-------------------------------|----------------------------------------|------------------------------------------|
| 1 | 1MTIDFYLYPGSAPCRVSLLA20 | β1-α1 |
| 2 | 47MNQHTPLNDNGCFLWESRA68 | α2-3β3-4 |
| 3 | 61GFCFLWESRASLYLDQYK80 | β4-α3 |
| 4 | 72SYLADQYKDLSYPKDAKRRALVD96 | α3-α4 |
| 5 | 116PYFAKQADPKMKKLEEAF137 | α5-α6 |
| 6 | 187KCKKIPVGYEELNHSCGLK206 | α8-α9 |

Anti-mouse immunoglobulin-alkaline phosphatase (AP) conjugate (Dako Cytomation, Denmark) and BCIP/NBT substrate (KPL, MD, USA) were used to reveal the 6× His-tagged-rGST band.

**Protein identification by LC-MS/MS.** Native and recombinant GSTs were verified by LC-MS/MS as described previously. The generated ion spectra of the peptides from tryptic-digested GSTs were interpreted using the Turbo SEQUEST algorithm in the BioWorks™ 3.1SR1 software package (Thermo Fisher Scientific) and the nr.fasta database. The protein search parameter was performed as described previously, which included mass tolerance of 1.25 amu, a fragment mass tolerance of ±0.4 amu, methionine (M) oxidation, and threonine (T) or serine (S) phosphorylation. The identified peptides were further evaluated using charge state versus cross-correlation numbers (Xcorr). The criteria for a positive identification of the peptides were Xcorr > 1.5 for singly charged ions, Xcorr > 2.0 for doubly charged ions, and Xcorr > 2.5 for triply charged ions. A delta correlation (ΔCn) of >0.08 was used as a cut-off for peptide acceptance.

**Determination of enzymatic activity of GST preparations.** Enzymatic activities of the nGST and the rGST were determined by using glutathione S-transferase assay kit (Sigma-Aldrich, MA, USA) which the 1-chloro-2,4-dinitrobenzene (CDNB) was used as the enzymatic substrate. GST standard (0.25 mg/ml) was provided with the test kit. Active GST catalyzes the conjugation of L-glutathione to CDNB via a thiol group of the glutathione. The reaction product, GS-DNB conjugate, could be detected at absorbance 340 nm (A340nm). The amount of the product is directly proportional to the GST activity in the sample. In a reaction mixture, 2 µl (0.5 µg) of native, recombinant, and standard GST were mixed individually with 1 ml of substrate solution containing 980 µl Dulbecco’s PBS, 10 µl of 200 mM reduced L-glutathione, and 10 µl of 100 mM CDNB in a 1-ml cuvette. Absorbance 340 nm of the reaction mixture was determined by spectrometer against blank (1 ml non-enzymatic conjugation substrate solution alone) using BioMate™3 series spectrophotometer (Thermo Fisher Scientific). A kinetic program was set for every 30 seconds over a period of 5 minutes after a lag time of 1 minute. Specific catalytic activities of both nGST and rGST (µmol of GST/ml/minute) were calculated.
Figure 5. Results of dot-ELISAs for determining binding activities of the synthetic peptides that contained predicted B-cell epitopes of the *P. americana*-GST. (A) All peptides were bound by antibodies in serum pool of *P. americana* sensitized subjects, indicating that the six peptides contain B cell epitopes. PBS was used as a negative antigen control. (B) Peptides 1 and 3 were bound by IgE in sera of all 15 *P. americana* allergic Thai patients (1–15), indicating that these two peptides contained allergenic epitopes. Sera of non-allergic controls (N1–N5) were used as negative IgE control in (B). Peptides 2, 4, 5, and 6 were not bound by the serum IgE (data not shown). Blank, peptide probed with PBS.

Figure 6. Homology model of the *P. americana*-GST three dimensional structure (grey) and locations of peptides 1 (yellow) and 3 (red) that were bound by IgE in sera of the 15 *P. americana* allergic patients, implying that these regions contain the GST allergenic epitopes. Peptide 1 is located on β1-α1 while peptide 3 is on β4-α3.
Allergenicities of nGST and rGST. Allergenicities (IgE binding frequencies) of the nGST and the rGST were determined by IgE-ELISA. The assay was performed as described previously. The nGST and rGST (5 μg/ml carbonate-bicarbonate buffer, pH 9.6) were added to separate wells (100 μl/well) of a microtiter plate (Costar, MA, USA) and kept at 37 °C until dried. All GST-coated wells were washed and blocked with 200 μl of a blocking solution (1% BSA in PBS) before adding with serial two-fold dilutions of individual sera and the plate was incubated for 3 hours. Wells added with only the serum diluent served as blank. All wells were washed and added with 100 μl of mouse anti-human IgE-biotin conjugate (Southern Biotech, AL, USA; diluted 1:1,000 in PBS-T). Streptavidin-horseradish peroxidase (HRP) conjugate (Dako Cytomation) and ABTS substrate solution (KPL) were used for color development. Absorbance at 405 nm (OD 405nm) of the content in each well was determined (ELISA reader, MultiscanEX, Labystem, Helsinki, Finland) against the blank. Cut-off OD 405nm, between positive and negative IgE-ELISA was arbitrarily set at ≥ mean OD 405nm of non-allergic sera + 2 standard deviations (SD).

Identification of nGST isoforms and their IgE reactivity. Isoforms of nGST were determined by a gel-based proteomics. Purified nGST was subjected to 2DE as described previously. For the first dimensional electrophoresis, 7 cm-IPG strips and 0.5% pH 3–10 IPG buffer (GE Healthcare) were used. The eletrophoresed-IPG strips were then subjected to 12% SDS-PAGE and proteins in the gel were stained by CBB. Gel pieces containing proteins of ~21 kDa were excised from the stained gel and subjected to in-gel tryptic digestion and LC-MS/MS, respectively. Protein orthologues were identified by comparing the peptide sequences of the P. americana-GST generated from the mass spectrometry with the Arthropoda/Insecta database sequences.

The nGST isoforms were checked for their reactivity to IgE in the pool of CR allergic patients’ sera by 2DE IgE-immunoblotting. The 2DE-separated nGST was electro-transblotted onto an NC, blocked with BSA, and the blot was allowed to react with the CR allergic patients’ serum pool. After keeping at 4 °C overnight, the NC was washed with TBS-T before placing in a solution of appropriately diluted mouse anti-human IgE-biotin conjugate (Southern Biotech) and kept at 25 °C on a rotating platform for 3 hours. Spots of the nGST isoforms bound by the specific serum IgE were revealed by using streptavidin-AP conjugate (Dako Cytomation) and BCIP/NBT substrate (KPL).

Identification of P. americana-GST B cell and allergenic epitopes. Linear B-cell epitopes of P. americana-GST were predicted by using BepiPred 1.0 server. The amino acid sequence of P. americana-GST was submitted to the server. Three methods available at the server, i.e., BCPred, AAP and FBPCpred were used for the epitope prediction. Specific threshold and the epitope length were set at 85% and 20–25 residues, respectively. All predicted peptides obtained from individual methods were aligned with the P. americana-GST sequence to obtain consensus sequences (predicted linear B-cell epitopes). Peptides containing the potential linear B-cell epitopes of P. americana-GST were synthesized. Moreover, in the case that two of the predicted epitope sequences were close to each other (peptides 2 and 4 of this study), the overlapped peptide sequences were also synthesized (peptide 3 of this study). Binding of the synthetic peptides to antibodies in a pool of P. americana sensitized human sera (for determining B-cell epitopes) and to IgE in individual allergic patients’ sera (for determining allergenic epitopes) were determined by dot-ELISAs.

Dot-ELISAs. For detecting binding of the synthetic peptides to antibodies in the pool of the CR allergic patients’ sera, individual synthetic peptides (1 μg) were dotted onto one cm-NC squares. PBS was used as negative control. The NC pieces were blocked with 1% BSA in PBS-T and then placed in the allergic patients’ serum pool. After keeping at 25 °C for 3 hours, the membranes were washed and allowed to react with goat-anti-human IgG-AP conjugate and BCIP/NBT substrate, respectively, with appropriate incubation and washing with PBS-T between the steps. The enzymatic reaction was stopped by rinsing the membranes with distilled water. Color appeared at the peptide-dotted spots indicated that the peptide contained B-cell epitope. For determining the IgE reactivity of the synthetic peptides, individual peptides were dotted separately onto NC squares (1 μg/dot) and let air-dried. They were blocked with 1% BSA in PBST and probed with individual serum samples (diluted 1:4 in PBST). Serum samples of normal subjects were included in the experiments. After incubating and washing, all NC squares were reacted sequentially with biotin-labeled-goat-anti-human IgE, streptavidin-AP conjugate, and BCIP/NBT substrate, respectively. The peptides that gave the typical colored spots indicated that they contained allergenic (IgE binding) epitopes.

Locations of the IgE-binding (allergenic) epitopes on the P. americana GST three dimensional (3D) structural model. Deduced amino acid sequence of the cloned P. americana GST was submitted to I-TASSER server. The templates used for GST 3D modeling (selected by the server) were PDB ID: 3WYW, 3VK9, 3AY8 and 4PNF. The allergenic peptides were mapped on the 3D modeled structure of the P. americana-GST.

Results

CR allergic patients. Dermographic data of the 15 P. americana allergic patients and 5 normal subjects as well as clinical diagnosis and results of the skin prick tests and the specific IgE levels among the patients are shown in Table 1.

P. americana nGST and rGST. Gene sequence coding for full-length P. americana-GST (648 bp) was cloned into pKRX-T cloning vector and subcloned into pET20b + protein expression vector. Supplementary Fig. 1 illustrates nucleotide and deduced amino acid sequences of the P. americana-GST of this study (accession number MG255130). Fig. 1 illustrates a phylogenetic tree of deduced amino acid sequence (216 residues) of the cloned P. americana-GST and GSTs of homologous and heterologous insects. Alignment (Clustal Omega Multiple
Sequence Alignment) and percent identity of amino acid sequence of _P. americana_-rGST of this study (accession number MG255130) with GSTs of other insects are shown in Supplementary Fig. 2. The _P. americana_-rGST was purified from the _gst-pET20b^-transformed _E. coli_ (Fig. 2, lane 2) and the protein was verified by the mass spectrometry (Table 2).

The nGST prepared from the frozen _P. americana_ whole body powder and purified by using GST trap column revealed a protein band at ∼23 kDa (Fig. 2, lane 1). LC-MS/MS verified that the protein band was the nGST (Table 2).

**Catalytic activities of nGST and rGST.** Enzymatic activities of the nGST and the rGST were 36.65 and 2.93 μmol/ml/min, respectively. The nGST had a much higher enzymatic activity than the rGST.

**Serum IgE reactivities of _P. americana_ nGST and rGST.** The mean + 2 SD of IgE-ELISA OD_{450\text{nm}} of the non-allergic control sera tested against the nGST and rGST were 0.460 and 0.649, respectively. These values were used arbitrarily as cut-off levels between positive and negative IgE-ELISA and it was found that all of the 15 allergic patients’ sera gave positive IgE-ELISA results to both nGST and rGST proteins (Fig. 3).

**_P. americana_ nGST isoforms.** The 2 DE-pattern of purified _P. americana_ nGST stained by CBB is shown in Fig. 4A. There are 5 protein spots at ∼21 kDa. The gel pieces containing the 5 protein spots were subjected to LC-MS/MS. The results (Table 3) revealed that tryptic peptides generated from spot nos. 1, 2, and 4 matched with peptides of _P. americana_ delta variant 1 (accession number Gi 359326557), which the molecular mass was 24606 Da and the pl was 6.44. This protein has possibility of 14 phosphorylation sites and a possible O-glycosylation at residue 116. Peptides of spot no. 3 matched with sigma GST of _Locusta migratoria_ (migratory locust) (accession number Gi 565341529) with molecular masses of 23459 Da and pl 6.19. This protein also has 14 possible phosphorylation sites; albeit most sites are different from the delta GST of spots 1, 2, and 4. Peptides of spot no. 5 matched with peptides of _B. germanica_ sigma GST (accession number Gi 359326585) which has a molecular mass of 23377 Da and pl 6.84 with possible 16 phosphorylation sites. Thus, the nGST of _P. americana_ of this study has three isoforms which belong to delta (one isoform) and sigma classes (two isoforms). By the 2 DE-IgE immunoblotting, all three isoforms of the nGST bound to IgE in a serum pool of the _P. americana_ allergic patients (Fig. 4B).

**Linear B-cell epitopes and allergenic epitopes of _P. americana_-GST.** Linear B-cell epitopes of _P. americana_-GST predicted by using BepiPred 1.0 server are shown in Supplementary Fig. 3. The BCPred, AAP, and FBCPred methods predicted 2, 3, and 4 epitopic sequences, respectively. All of the predicted peptides were aligned with the _P. americana_-GST sequence. They were found to match with five regions (marked in red in the Supplementary Fig. 3) of the GST including _α_1, _α_2-33-3 _α_4, _α_5-6, and _α_8-9. The consensus peptides of these locations were synthesized (Pep 1, 2, 4, 5, and 6; marked in blue in the Supplementary Fig. 3 and peptides 1, 2, 4, 5, and 6 in Table 4). Because the matched regions _α_2-33-3 _α_4 and _α_3-4 were close to each other, an overlapped peptide of the two regions was also synthesized (Pep 3 in the Supplementary Fig. 3 and peptide 3 in Table 4). All synthetic peptides gave positive binding to antibodies in a pool of 10 _P. americana_ allergic subjects (Fig. 5A), verifying the computerized results that the peptides contained B-cell epitopes. Sera of all _P. americana_ allergic patients gave positive IgE-dot-ELISA to peptide 1 (1MTIDFYYLPGSPACRSVLLA20) located between _α_1 and _α_1 and peptide 3 (61GFCLWESRAILYLDGYK80) located between _α_4 and _α_3 of the _P. americana_-GST molecule (Fig. 5B) and did not give positive IgE binding with peptides 2, 5, and 6 (data not shown).

Figure 6 illustrates locations of the two IgE-binding (allergenic) epitopes on the 3D modeled structure of the _P. americana_-GST.

**Discussion**

American cockroach, _P. americana_, is a predominant species and a major source of indoor allergen causing type 1 hypersensitivity in atopic inhabitants of subtropics (such as Taiwan) and tropics (such as Brazil, Malaysia, Singapore, and Thailand), especially in crowded and unhygienic urban cities where their infestation is enhanced. Several _P. americana_-derived proteins have been recognized as important (major) allergens as they sensitized >50% of cockroach allergic subjects. Glutathione-S-transferases (GSTs) are indispensable enzymes which insects use to protect themselves against oxidative damage and stress and insecticide toxicity. Most interest on the insect GSTs has been focused on their role in insecticide and drug resistance while other attributions of this protein family seem to be neglected. In this study, three _P. americana_-GST isoforms belonging to delta and sigma classes were identified. The roles of the protein in causing human allergy, i.e., allergenicity (specific serum IgE-binding frequency) as well as B-cell epitopes and allergenic epitopes were investigated. To our knowledge, this is the first report on the allergenic attribution of the American cockroach, _P. americana_-GST. The protein that was cloned from the GenBank database (accession number AT792949; UniProt ID Q1M0Y4) has been designated Per a 5 (Per a 5.0101) by the IUIS Allergen Nomenclature. The other isoforms reported in this study were based on the tryptic digested peptides generated from the 2DE protein spots that matched with the orthologous proteins of the database. Because the complete amino acid sequences of these isoforms were not available, they were not submitted to the IUIS Allergen Nomenclature.

Enzymatically active native and recombinant _P. americana_-GSTs were produced. On the equal weight of both proteins (0.5 μg), the native protein was about 12.5 times more active than the recombinant counterpart. The difference may be because the nGST contains several isoforms belonging to different GST classes whereas the recombinant one is produced from only one cDNA sequence. The multiple isoforms in the nGST might confer...
additive or synergistic enzymatic activity. Besides, the rGST produced from the transformed BL21 (DE3) E. coli contains additional 6× His tag (useful for subsequent protein detection and purification) but lacks putative post-translational modifications such as glycosylation and disulfide bridge formation which might impact on the protein folding compared to the native state and hence the less enzymatic activity. After subjecting the rGST and GST to SDS-PAGE and protein staining, the molecular masses of the native and recombinant proteins were ∼23 and ∼25 kDa, respectively. The larger size of the latter should be from the plasmid flanking regions and the 6× His tag. The rGST produced in this study (accession number MG255130) contained 216 amino acids which are similar in residue number to the cloned 24614 Da B. germanica GSTD145. The P. americana-rGST has 99.5 and 98.1% sequence identity to the previously reported P. americana-GSTs (UniProt ID G8XWU4 and Q1M0Y4) and 32.9–82.4% identity to the delta GSTs of other organisms and less so to the other GST classes. Thus, the recombinant P. americana-GST (accession number MG255130) should belong to the delta class.

Both native and recombinant P. americana GSTs reacted with IgE in sera of all cockroach allergic patients when tested by indirect ELISA, indicating that the GST is a novel and important (major) allergen of the P. americana. The ELISA using rGST as antigen gave higher background signal than when the nGST was used. This could be due to the contamination of residual proteins of the E. coli used as the rGST expression host in the rGST preparation which reacted to the anti-E. coli that exists naturally in the human sera.

Data on the allergenic repertoire of a protein are useful for understanding the patients’ allergenic response, cross-reacting allergenic determinants among allergens, as well as for properly designing of an engineered therapeutic allergen vaccine/diagnostic material, particularly for personalized immunotherapy and component resolved diagnosis (CRD)46. Several methods have been used for gaining information on B-cell and allergenic epitopes of an allergen. These include the use of overlapping synthetic peptides or fragments of recombinant allergens47-48; mimotope mapping49; peptide microarray immunoassay50,51; X-ray crystallography and nuclear magnetic resonance techniques52-55; computerized prediction56-59 and specific monoclonal antibody binding and IgE competition assay60-62. In this study, a combination of in silico methods and synthetic peptide based-immunoassays was used for predicting and determining B-cell and allergenic epitopes of the P. americana-GST. For the in silico prediction of B cell epitopes, three different methods of the BepiPred 1.0 server, i.e., BCPred, AAP, and FBCPred were used. All methods gave a conformed prediction of a peptide located at α2-33-34; the AAP and PBCPred methods predicted another peptide at α3-α4; the BCPred and the FBCPred methods predicted another peptide at α5-α6; and only the AAP method predicted two more peptides at β1-α1 and α8-α9. The different results so-obtained suggest that several in silico methods should be used for increasing the possibility of finding the potential B-cell epitopes of a particular protein. From the in silico prediction, consensus peptides encompassed the potential B-cell epitopes were synthesized and used in the dot-ELISAs for detecting peptides bound by antibodies (B-cell epitopes) and IgE (allergenic epitopes) in sera of P. americana sensitized subjects. By using the dot-ELISAs, all six synthetic peptides that contained predicted GST B-cell epitopes reacted with antibodies in a serum pool of P. americana exposed-subjects, implying that the six peptides were part or contained GST B-cell epitopes. Among them were two IgE-binding epitopes located at the β1-α1 and 34-α4 regions of the GST molecule.

In summary, this study provides an insight into characteristics and medically important role of the P. americana-GST beyond the previously established physiologic roles of the protein in the host defense against toxic substances and stress conditions. Both native and recombinant GSTs of the P. americana were bound by IgE in sera of cockroach allergic subjects and thus the GST is a novel and major P. americana allergen. Because cockroaches are ubiquitous, attention should be paid on reducing the insect derived-allergenic proteins from human environment for intervention of allergic sensitization of the naives and clinical aggravation of the sensitized subjects.

References
1. Arruda, L. K. et al. Cockroach allergens and asthma. J. Allergy Clin. Immunol. 107, 419–428 (2001).
2. Huss, K. et al. House dust mite and cockroach exposure are strong risk factors for positive allergy skin test responses in the Childhood Asthma Management Program. J. Allergy Clin. Immunol. 107, 48–54 (2001).
3. Bunjean, K. et al. Prevalence of allergic rhinitis and types of sensitized allergen in adult at Wat Intaram community, Hua Raeu, Phra Nakhon Si Ayutthaya District, Phra Nakhon Si Ayutthaya Province, Thailand. J. Med. Assoc. Thai 95, S63–68 (2012).
4. Yong, T. S. & Jeong, K. Y. Household arthropod allergens in Korea. Korean J. Parasitol. 47, S143–153 (2009).
5. Richman, P. G., Khan, H. A., Turkeltaub, P. C., Malveaux, F. J. & Baer, H. The important sources of German cockroach allergens as determined by RAST analyses. J. Allergy Clin. Immunol. 73, 590–595 (1984).
6. Menon, P., Menon, V., Hilmah, B., Stankus, R. & Lehrer, S. B. Skin test reactivity to whole body and fecal extracts of American (Periplaneta americana) and German (Blattella germanica) cockroaches in atomic asthmatics. Ann. Allergy 67, 573–577 (1991).
7. Tungtrongchitr, A. et al. The levels of cockroach allergen in relation to cockroach species and allergic diseases in Thai patients. Asian Pac. J. Allergy Immunol. 22, 115–121 (2004).
8. Schou, C., Lind, P., Fernandez-Caldas, E., Lockey, R. F. & Lewenstein, H. Identification and purification of an important cross-reactive allergen from American (Periplaneta americana) and German (Blattella germanica) cockroaches in atomic asthmatics. Ann. Allergy 86, 935–946 (1990).
9. Diraphat, P. et al. Recombinant American cockroach component, Per a 1, reactive to IgE of allergic Thai patients. Asian Pac. J. Allergy Immunol. 21, 11–20 (2003).
10. Wu, C. H., Lee, M. F. & Wang, N. M. Expression of the American cockroach Per a 1 allergen in mammalian cells. Allergy 55, 1179–1183 (2000).
11. Lee, M. F., Song, P. P., Hwang, G. Y., Lin, S. J. & Chen, Y. H. Sensitization to Per a 2 of the American cockroach correlates with more clinical severity among awayd allergic patients in Taiwan. Ann. Allergy Asthma Immunol. 108, 243–248 (2012).
12. Wu, C. H., Lee, M. F., Liao, S. C. & Luo, S. F. Sequencing analysis of cDNA clones encoding the American cockroach Cr-P1 allergens. Homology with insect hemolymph proteins. J. Biol. Chem. 271, 17937–17943 (1996).
13. Hindley, J. et al. Bla g 6: a troponin C allergen from Blattella germanica with IgE binding calcium dependence. J. Allergy Clin. Immunol. 117, 1389–1395 (2006).
14. Pomes, A. & Arruda, A. K. Investigating cockroach allergens: aiming to improve diagnosis and treatment of cockroach allergic patients. *Methods* 66, 75–85 (2014).
15. Witteman, A. M., Akkerdaas, J. H., van Leeuwen, J., van der Zee, J. S. & Aalberse, R. C. Identification of a cross-reactive allergen (presumably tropomyosin) in shrimp, mite and insects. *Int. Arch. Allergy Immunol.* 105, 56–61 (1994).
16. Asturias, J. A. et al. Molecular characterization of American cockroach tropomyosin (*Periplaneta americana*) allergen 7, a cross-reactive allergen. *J. Immunol.* 162, 4342–4348 (1999).
17. Sookrung, N. et al. Allergenicity of native/recombinant tropomyosin, per a 7, of American cockroach (CR), *Periplaneta americana*, among CR allergic Thais. *Asian Pac. J. Allergy Immunol.* 27, 9–17 (2009).
18. Sookrung, N. et al. *Periplaneta americana* arginine kinase as a major cockroach allergen among Thai patients with major cockroach allergies. *Environ. Health Perspect.* 114, 875–880 (2006).
19. Khantisitthiporn, O. et al. Native tropomin-T of the American cockroach (CR), *Periplaneta americana*, binds to IgE in sera of CR allergic Thais. *Asian Pac. J. Allergy Immunol.* 25, 189–197 (2007).
20. Sudha, V. T., Arora, N., Gaur, S. N., Pasha, S. & Singh, B. P. Identification of a serine protease as a major allergen (Per a 10) of *Periplaneta americana*. *Allergy* 63, 768–776 (2008).
21. Fang, Y. et al. Two new types of allergens from the cockroach. *Periplaneta americana*. *Allergy* 70, 1674–1678 (2015).
22. Hayes, J. D. & Pulford, D. J. The glutathione S-transferase supergene family: regulation of GST and the contribution of the enzymes to cancer chemoprotection and drug resistance. *Critt. Rev. Biochem. Mol. Biol.* 30, 445–600 (1995).
23. Enayati, A. A., Ransoh, H. & Hemingway. I. Insect glutathione transferase. *Insect Mol. Biol.* 14, 3–8 (2005).
24. Birben, E., Sahiner, U. M. & Sackesen, C. Erzurum. S. & Kalayci, O. Oxidative stress and antioxidant defense. *World Allergy Organ. J.* 5, 9–19 (2012).
25. Ding, Y., Orelli, E., Rossiter, L. C., Hemingway, J. & Ransoh, H. The *Aphanoele gambiae* glutathione transferase supergene family: annotation, phylogeny and expression profiles. *BMC Genomics* 4, 35, https://doi.org/10.1186/1471-2164-4-35 (2003).
26. Hayes, J. D., Flanagan, J. U. & Jowsey, I. R. Glutathione transferases. *Annu. Rev. Biochem.* 63, 369–376 (2006).
27. Chen, J., Liu, H., Yang, J. & Chou, K. C. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. *Protein Eng. Des. Sel.* 15, 422–430 (2016).
28. El-Manzalawy, Y., Dobbs, D. & Honavar, V. Predicting linear B-cell epitopes using string kernels. *J. Mol. Recognit.* 21, 243–255 (2008).
29. Chen, J., Liu, H., Yang, J. & Chou, K. C. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. *Amino Acids* 33, 423–428 (2007).
30. El-Manzalawy, Y., Dobbs, D. & Honavar, V. Predicting flexible length linear B-cell epitopes. *Compust. Syst. Bioinformatics Conf. 7*, 121–132 (2008).
31. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738 (2010).
32. Sookrung, N. & Chaiumpita, W. A revisit to cockroach allergens. *Asian Pac. J. Allergy Immunol.* 28, 95–106 (2010).
33. Ma, B. & Chang, F. N. Purification and cloning of a delta class glutathione S-transferase displaying high peroxidase activity isolated from the German cockroach *Blatella germanica*. *FEBS J.*, 274, 1793–1803 (2007).
34. Valenta, R. et al. The recombinant allergen-based concept of component-resolved diagnostics and immunotherapy (CRD and CRIT). *Clin. Exp. Allergy* 29, 896–904 (1999).
35. Pomes, A. Relevant B cell epitopes in allergic disease. *Int. Arch. Allergy Immunol.* 152, 1–11 (2009).
36. van Millingen, F. J., van ’t Hof, W., van den Berg, M. & Aalberse, R. C. IgE epitopes on the cat (*Felis domesticus*) major allergen Fel d 1: a study with overlapping synthetic peptides. *J. Allergy Clin. Immunol.* 93, 34–43 (1994).
37. Pocios, L. F. et al. Minitope mapping as a complementary strategy to define allergen IgE-epitopes: peach Pru p 3 allergen as a model. *Mol. Immunol.* 45, 2269–2276 (2008).
38. Vereda, A. et al. Identification of IgE sequential epitopes of lentil (*Lens culinaris*) by means of peptide microarray immunoassay. *J. Allergy Clin. Immunol.* 126, 596–601 (2010).
39. Otsu, K., Guo, R. & Dreski, S. C. Epitope analysis of Ara h 2 and Ara h 6: characteristic patterns of IgE-binding fingerprints among individuals with similar clinical histories. *Clin. Exp. Allergy* 45, 472–484 (2015).
40. Amit, A. G., Marriuza, R. A., Phillips, S. E. & Poljak, R. J. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science* 233, 747–753 (1986).
41. Sheriff, S. et al. Three-dimensional structure of an antibody-antigen complex. *Proc. Natl Acad. Sci. USA* 84, 8075–8079 (1987).
42. Padlan, E. A. et al. Structure of an antibody-antigen complex: crystal structure of the H-YHEL-10 Fab-lysozyme complex. *Proc. Natl Acad. Sci. USA* 86, 5938–5942 (1989).
43. Ichiikawa, S. et al. NMR study of the major mite allergen Der f 2: its refined tertiary structure, epitopes for monoclonal antibodies and characteristics shared by ML protein group members. *J. Biochem.* 137, 255–263 (2005).
44. Dall’Antonia, F., Giersas, A., Devanaboyina, S. C., Valenta, R. & Keller, W. Prediction of IgE-binding epitopes by means of allergen surface comparison and correlation to cross-reactivity. *J. Allergy Clin. Immunol.* 128, 872–879 (2011).
45. Dall’Antonia, F., Pavkov-Keller, T., Zanger, K. & Keller, W. Structure of allergens and structure based prediction methods. *Methods* 66, 3–21 (2014).
58. Tasaniyananda, N. et al. A novel IgE-binding epitope of cat major allergen, Fel d 1. Biochem. Biophys. Res. Commun. **470**, 593–598 (2016).
59. Olson, J. R. & Klapper, D. G. Two major human allergenic sites on ragweed pollen allergen antigen E identified by using monoclonal antibodies. *J. Immunol.* **136**, 2109–2115 (1986).
60. Chapman, M. D., Heymann, P. W. & Platts-Mills, T. A. E. Epitope mapping of two major inhalant allergens, Der p 1 and Der f 1, from mites of the genus *Dermatophagoides*. *J. Immunol.* **139**, 1479–1484 (1987).

**Acknowledgements**
The work was co-supported by the National Research University (NRU) Project of the Office of Commission on Higher Education, Ministry of Education, the Thailand Research Fund (RSA5780025), Faculty of Medicine Siriraj Hospital, Mahidol University, and the NSTDA Chair Professor grant (P-1450624) of the National Science and Technology Development Agency (NSTDA) funded by the Crown Property Bureau of Thailand.

**Author Contributions**
A.T.T., N.S.R. and W.C.C. contributed to the design of the experiments, result analysis, and manuscript writing. N.S.R., R.P.P., N.I.D., O.R.T. and N.S.L. performed the experiments. W.S.S. predicted the GST B cell epitopes and the locations of the IgE-binding epitopes on the 3D structure of the GST computerized model. P.T.T. and C.B.N. performed skin prick test, checked serum specific IgE levels and collected blood samples. N.S.R., W.S.S. and W.C.C. created Figures.

**Additional Information**
**Supplementary information** accompanies this paper at https://doi.org/10.1038/s41598-017-18759-z.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[Open Access] This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017