



รายงานวิจัยฉบับสมบูรณ์

โครงการอนุรักษ์ชีววิทยาในจุลชีพ

โดย

ศาสตราจารย์ สกล พันธ์ภูมิ และคณะ

31 สิงหาคม 2546

รายงานวิจัยฉบับสมบูรณ์

โครงการอนุรักษ์วิทยาในจุลชีพ

คณะผู้วิจัย

สังกัด

- | | |
|--------------------------------|---|
| 1. นายสกล พันธุ์ยิ้ม | ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล |
| 2. นายบุรชัย สมนานนท์ | สถาบันอนุรักษ์วิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล |
| 3. นางสาววิภา จิงจุพรชัย | สถาบันอนุรักษ์วิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล |
| 4. นางสาวสุณี เกิดบัณฑิต | สถาบันอนุรักษ์วิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล |
| 5. นายชนันท์ อังสุรณสมบัติ | สถาบันอนุรักษ์วิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล |
| 6. นายอภิรักษ์ อุดมกิจ | สถาบันอนุรักษ์วิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล |
| 7. นายชาติชาย กฤตชัย | สถาบันอนุรักษ์วิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล |
| 8. นายวิเชษฐ ลีถาณิตย์ | ภาควิชาชีวเคมี คณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล |
| 9. นางสาวสุมาลี ตั้งประดับกุล | ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล |
| 10. นางวิไล หนูนุกดี | ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล |
| 11. นายศราวุฒ จิตรภักดี | ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล |
| 12. นางอัญชลี ทัศนากจร | ภาควิชาชีวเคมี คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย |
| 13. นายชัยณรงค์ วงศ์ธีระทรัพย์ | ภาควิชาชีวเคมี คณะแพทยศาสตร์ มหาวิทยาลัยมหิดล |

ศรีนครินทร์วิโรฒ ประสานมิตร

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย

ชุดโครงการเมธีวิจัยอาวุโส สกว. สกล พันธุ์ยิ้ม

สารบัญ

	หน้า
บทคัดย่อ	1
เนื้อหาทางานวิจัย	
- FOLDING AND STABILITY CHARACTERIZATION OF THE MOSQUITO-LARVICIDAL CRY4B PROTEIN	3
- MOLECULAR BASIS OF MEMBRANE PORE-FORMATION BY THE <i>BACILLUS THURINGIENSIS</i> Cry4 LARVICIDAL PROTEINS	11
- CLONING AND EXPRESSION OF KAZAL-TYPE SERINE PROTEINASE INHIBITORS OF THE BLACK TIGER SHRIMP, <i>PENAEUS MONODON</i>	20
- BLACK TIGER SHRIMP HEMOCYTE CDNAS : FROM EXPRESSED SEQUENCE TAGS TO PROSPECTIVE MARKERS OF MOLECULAR EVENT UPON YELLOW HEAD VIRUS INFECTION	31
- STRUCTURE AND ORGANIZATION OF CRUSTACEAN HYPERGLYCEMIC HORMONE-LIKE GENES OF <i>PENAEUS MONODON</i>	44
- PCR-BASED METHOD FOR EPIDEMIOLOGICAL IDENTIFICATION OF THE PATHOGENIC <i>BURKHOLDERIA PSEUDOMALLEI</i>	59
- CLONING AND CHARACTERIZATION OF A LIGHT-INTENSITY-RESPONSIVE GENE OF <i>SYNECHOCOCCUS</i> PCC7942	75
- RESISTANCE OF TRANSGENIC PAPAYA PLANTS TO PAPAYA RINGSPOT VIRUS THAI ISOLATE	89
- ISOLATION OF NISIN-PRODUCING <i>LACTOCOCCUS LACTIS</i> STRAINS FROM TRADITIONAL FERMENTED MEAT AND FISH PRODUCTS	111
- <i>IN VIVO</i> VARIABILITY AND KINETIC PROPERTIES OF THE SUBTYPE E HIV-1 PROTEASE FROM INFECTED-THAI PATIENTS	125
- MOLECULAR CLONING OF GP65 STRUCTURAL PROTEIN OF THE YELLOW HEAD VIRUS	130
- ANALYSIS OF GENETIC DIVERSITY IN RIBONUCLEOTIDE REDUCTASE GENE OF WHITE SPOT SYNDROME VIRUS	153
- SUPPRESSION OF VIRAL REPLICATION BY GENE SILENCING MECHANISM IN BLACK TIGER PRAWN (<i>Penaeus monodon</i>)	160

Research Ouput (จากทุนเมธีวิจัยอาวุโส สกว.)

- ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ	169
- การเสนอผลงานในประชุมวิชาการนานาชาติ	172
- การเสนอผลงานในประชุมวิชาการในประเทศ	174
- การประชุมวิชาการอณูชีววิทยาประจำปี	176
รายชื่อที่ปรึกษาโครงการและคณะผู้ร่วมวิจัย	177

บทคัดย่อ

โครงการงานวิจัยอนุชีววิทยาในจุลชีพ ประกอบด้วยคณะผู้ร่วมวิจัย จำนวน 27 คน ได้ดำเนินการเป็นระยะเวลา 3 ปี มีนักวิจัยผู้มีความเชี่ยวชาญ จำนวน 13 คน มีการผลิตผลงานวิจัยเป็นที่ยอมรับในระดับสากลโดยการตีพิมพ์ผลงานวิจัยในวารสารนานาชาติ จำนวน 34 เรื่อง เสนอผลงานวิจัยในที่ประชุมนานาชาติ จำนวน 19 เรื่อง และในที่ประชุมในประเทศ จำนวน 19 เรื่อง และได้ผลิตนักวิจัยระดับปริญญาเอก จำนวน 3 คน ระดับปริญญาโท จำนวน 2 คน และกำลังศึกษาปริญญาเอก จำนวน 3 คน ระดับปริญญาโท จำนวน 4 คน มีการจัดประชุมวิชาการอนุชีววิทยาในจุลชีพประจำปี จำนวน 3 ครั้ง มีผู้เข้าร่วมประชุม จำนวนทั้งสิ้น 227 คน ผลงานวิจัยโดยสังเขป ได้องค์ความรู้ใหม่ด้านกลไกการออกฤทธิ์ที่ระดับอนุของโปรตีน Cry4 จาก *Bacillus thuringiensis* โดยเฉพาะความรู้ด้านการมันพับของ Cry4 ทำให้เกิดโครงสร้าง 3 มิติที่มีผลต่อการออกฤทธิ์ฆ่าลูกน้ำยุงและการสอดแทรกของ Cry4 เข้าในเนื้อเยื่อเซลล์ ทำให้เกิดรูรั่วอันส่งผลให้เซลล์เสียสมดุลย์และตาย องค์ความรู้ในแบคทีเรียสังเคราะห์แสง *Synchococcus* PCC7942 ได้แยกโปรโมเตอร์ที่ถูกกระตุ้นโดยความเข้มของแสง ศึกษาผลความเข้มแสงต่อการแสดงออกของยีน ORF76 ซึ่งถูกควบคุมโดยโปรโมเตอร์ดังกล่าว องค์ความรู้ในแบคทีเรีย *Burkholderia pseudomallei* ซึ่งก่อโรค melioidosis ได้ทราบลำดับเบสในยีน flagellin ซึ่งจำเพาะต่อชนิดก่อโรคและได้พัฒนาวิธีตรวจหาแบคทีเรียก่อโรคในเลือดคนไข้และในดินที่มีความไวและความแม่นยำสูง ด้านแบคทีเรียแลคติด (LAB) ซึ่งแยกจากอาหารหมักดองไทย ได้แยกยีน Nisin ซึ่งออกฤทธิ์ฆ่าแบคทีเรียแกรมบวกในอาหาร Nisin เป็นโปรตีนเล็กมีกรดอะมิโน 56 ตัว ได้มีการศึกษาไวรัส HIV โดยเน้นการกลายพันธุ์ของโปรตีเอสพบมีการกลายพันธุ์สูงขึ้นในคนไข้ที่ได้รับยา ได้ศึกษาไวรัส PRSV ก่อโรคใบด่างจุดวงแหวนในมะละกอไทยและใช้ลำดับเบสของยีน CP นำไปสร้างมะละกอแปลงพันธุ์มียีน CP แล้วคัดเลือกพันธุ์ที่ต้านทานการติดเชื้อ PRSV ในระดับแปลงทดลองได้ 5 สายพันธุ์ การต้านทาน PRSV ในสายพันธุ์ G2 สามารถถ่ายทอดสู่รุ่นที่ R1, R2 และ R3 ได้ ได้ศึกษาโปรตีน GP116 และ GP64 จาก ORF3 ของ YHV ก่อโรคหัวเหลืองในกุ้งกุลาดำ พบส่วน Transmembrane ทางด้านปลายคาร์บอกซิล ซึ่งทำให้การสร้างโปรตีนดังกล่าวใน *E.coli* หรือเซลล์ Sf9 ลดลง ได้ศึกษาพบความแตกต่างในจำนวน repetitive DNA ของ WSSV ซึ่งแยกจากแหล่งต่างๆ ในประเทศ ได้ศึกษา cDNA จาก Hemocyte กุ้งพบยีน selenoprotein W เพิ่มขึ้น เมื่อกุ้งติดเชื้อ YHV หรือมี stress ได้ clone ยีน serine protease inhibitors จาก hemocyte กุ้ง และ express ใน *E.coli* ซึ่งแสดงผลยับยั้ง trypsin และ chymotrypsin ได้ศึกษาโครงสร้างยีน CHH จากก้านตากุ้ง พบ 3 ชนิดอยู่ใน cluster ขนาด 10 Kb โดยแต่ละยีน (CHH1, CHH2, CHH3) แสดง regulatory elements ที่แตกต่างกัน

โครงการนำร่องในการศึกษา suppression viral replication by gene silencing mechanism in black tiger prawn เพื่อจะได้ข้อมูลเบื้องต้นของกลไกการเกิด gene silencing ในกุ้งกุลาดำ ได้ออกแบบ siRNA ซึ่งสามารถกดการแสดงออกของยีน GAPDH ได้ประมาณ 70% และพบ RNA เส้นคู่ขนาด 500-800 bp กดการแสดงออกของยีนและการเพิ่มจำนวนไวรัส YHV ในเซลล์ กุ้งได้ แสดงว่ากุ้งมีกลไก กดการแสดงออกยีน (gene silencing) เหมือนกรณีที่พบในหนอนตัวกลม แมลงหวี่ และสัตว์บางชนิด

เนื้อหางานวิจัย

FOLDING AND STABILITY CHARACTERIZATION OF THE MOSQUITO-LARVICIDAL CRY4B PROTEIN

Chartchai Krittanai, Apichai Bourchookarn, Weerachon Taepanunt, Anchane Sangcharoen, Wanwarang Pathaichindachote and Sakol Panyim

Laboratory of Molecular Biophysics, Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakhonpathom 73170, Thailand

Abstract

The 65-kDa active cry4B protein from *Bacillus thuringiensis* was characterized for an unfolding pathway. The purified proteins were incubated in a series of guanidinium hydrochloride concentration in which their conformational states were monitored by circular dichroism and intrinsic fluorescence spectroscopy. The combined data from steady state and stopped flow kinetic measurements established an unfolding free energy of 17.87 kcal/mol and the activation energy of 25.42 kcal/mol for wild type. The mutant R203Q with an internal tryptic site removed showed an increased stability for the folded state by 5.23 kcal/mol. Several sets of mutant were constructed to assess for the stabilizing role of residues on the five conserved blocks. The data revealed that the substitution of the hydrophobic residues, L175 in block I has a significant destabilizing effect on the packing of domain I interior. Analysis based on the energy profile of mutants showed an increase of conformational energy for the folded state and no effect on the activation energy. A number of mutants with residue substitutions in conserved block I to block V were obtained either as insoluble protein or product that is highly sensitive to proteolytic degradation. Based on an analysis of the recently elucidated X-ray structure, it is suggested that the mutation introduced into these mutants has abolished the critical interaction that stabilizes the folding of tertiary structure. The other mutants, which demonstrate the native-like property were also analyzed and compared. However, the results showed that the removal of secondary structure stabilization did not affect folding and stability of the Cry4B.

Keywords: *Bacillus thuringiensis*, protein folding, molecular stabilization

Introduction

Delta-endotoxins produced from *Bacillus thuringiensis*, are found selectively toxic to a wide range of insect larvae [1]. The genes encoded for Cry4 toxin was cloned, expressed and characterized in *E. coli* [2]. A mechanism of action for Cry toxins has been under investigation by several research groups using site-directed mutagenesis [3, 4, 5, 6].

X-ray structures of Cry3A [7] and Cry1Aa [8] toxins were elucidated showing a general fold comprising of three domains. The seven-helix bundle of domain I has been extensively studied for its role on membrane insertion and channel activity [5, 6, 9, 10, 11]. Domain II was proposed to be responsible for docking and specificity [12,

13]. Studies for folding and stability of Cry toxins has begun in the early 1990s to determine the number of structural domain and conformation changes during proteolytic activation of the native protoxin [14, 15, 16]. The folding investigation of an induced transition of Cry toxin has been reported recently for Cry1A and Cry3A [17, 18] with proposed nucleation core on domain I [19].

Based on site-directed mutagenesis studies of Cry toxins, a numbers of mutant toxins with unusual physical properties were found. The toxins were found insoluble in native buffers, highly sensitive to protease digestions or complete lost of activity [6, 20, 21]. These mutant toxins were proposed to be a result from the incorrect folding or molecular instability due to changes of amino acid residue at specific locations.

This work aims to characterize for a structural unfolding of the 65-kDa active Cry4B toxin. Site-direct mutagenesis was applied to assess for the contribution of residues in the five conserved blocks on the molecular folding and stability. Each directed mutation was assayed by the unfolding experiment yielding an energy profile of unfolding. Energy profile was then applied to describe the role of mutated residues on folding and stability.

Materials and Methods

Sited-directed mutagenesis

Mutant plasmid were constructed by a method based on QuickChange site-directed mutagenesis kit (Stratagene, USA.) using high fidelity Pfu DNA polymerase. Synthetic oligonucleotide primers were from GeneSet. Mutant plasmids were constructed from PCR and digested with *Dpn* I for template removal. Transformants were screened by restriction enzyme digestion and DNA sequences were confirmed by automated DNA sequencing.

Protein expression and purification

Protein expression was induced with 0.1 mM IPTG at 37 °C for 4 hrs. The packed cell was then applied to French Press Cell at 1200 psi. Cell lysate was collected and washed several times before resuspended in 50 mM carbonate buffer pH 10.0. The obtained protoxin was activated by incubating with TPCK-treated trypsin (1:20 of enzyme: toxin). Activated toxin was injected to an AKTA Purifier FPLC system (Amersham Pharmacia Biotech) equipped with Superdex200-10/30, size-exclusion columns. Mobile phase was a carbonated buffer pH 10.0, flowing at 0.5 ml/min. Chromatograms were recorded according to absorption at 280 and 214 nm. Fractions were collected every 1 ml on a Frac-900 fraction collector. Purity was verified with a single peak for each sample before uses.

Determination of protein concentration

Protein assay kit based on Bradford's dye binding method [22] was used for routine determination of protein concentration. For spectroscopic studies, accurate concentration of the purified protein was obtained from measurement of absorption in the far UV at 215 and 225 nm using a double-beam Cary Bio300 spectrophotometer (Varian, Australia). The calculation was based on an equation, $\text{Conc. (mg/ml)} = 0.144 \times (A_{215} - A_{225})$

Unfolding experiments

Chemical unfolding was performed by a series of guanidinium hydrochloride titration from 0 to 6 M. Accurate concentration of guanidinium hydrochloride stock solutions were determined from a measurement of refractive index [23]. Protein concentrations used in the experiment were 0.1-0.2 mg/ml

Circular dichroism and fluorescence spectroscopy

Circular dichroism analysis was performed on a Jasco J-715 spectropolarimeter. The instrument was calibrated with CSA and continuously purged with nitrogen gas during the measurements. Samples were loaded on a water-jacketed cylindrical cell with 0.02 cm pathlength. The temperature was controlled by a Naslab circulation bath. Scanning rate was controlled at 20 nm/min with 2-millisecond response time, 50-milledegree sensitivity, and 3-4 accumulations. Intrinsic fluorescence spectra were measured on a Perkin Elmer LS-50 spectrofluorometer. Samples were excited at 280 nm and the emission spectra were recorded from 300-450 nm. Scanning rate was set at 50 nm/min using 5-nm bandwidth.

Results and Discussions

Thermodynamic analysis for unfolding free energy

Steady state analysis of conformational states of protein under various GuHCl concentrations gave a series of CD and fluorescence spectra. These spectra represent various states of protein during the transition from the native to unfolded conformation. (**Figure 1**). Unfolding curves derived from intensity changes of CD and fluorescence spectra revealed a similar sigmoidal curve (**Figure 2**). The obtained unfolding curves demonstrated two-state transition with a transitional midpoint at 3.54 M GuHCl. Unfolding free energy of this unfolding in an absence of denaturant ($\Delta G^{\circ}_{H_2O}$) calculated by two-state model fitting was 17.87 kcal/mol.

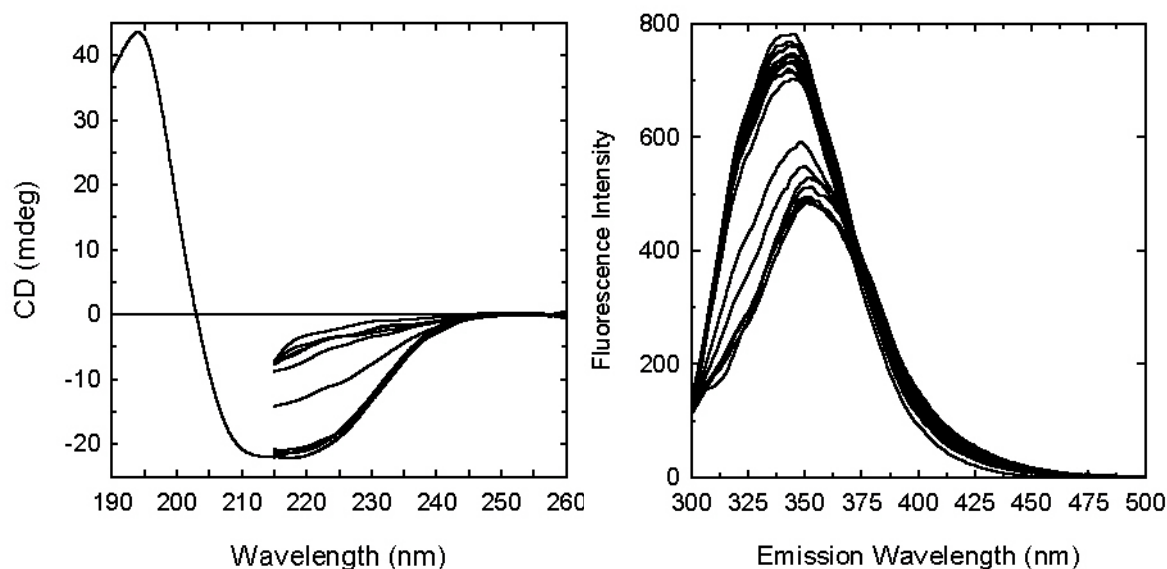


Figure 1: Typical changes of circular dichroism (A) and fluorescence spectra (B) upon conformational unfolding of Cry4B toxin.

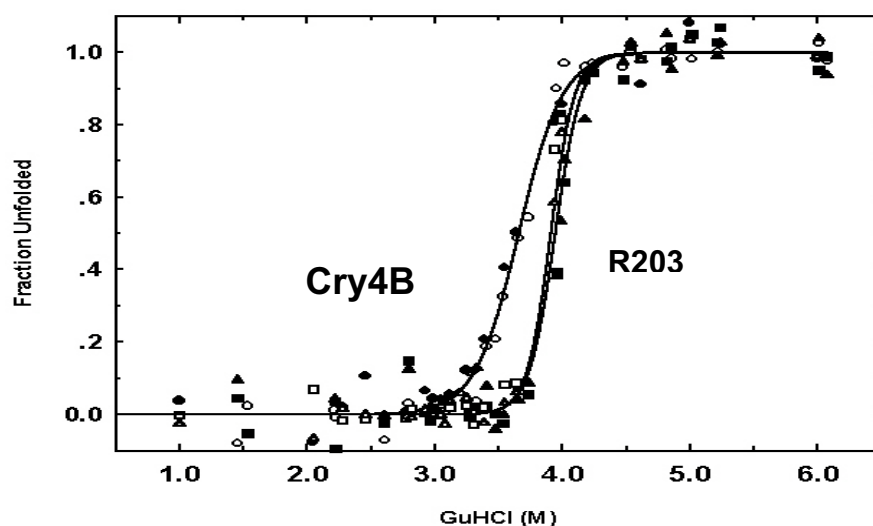


Figure 2: Unfolding curves of wild type and mutant toxins obtained from a plot of CD and fluorescence intensity

Kinetics analysis for activation energy of unfolding

To investigate for the activation energy of unfolding transition, conformational changes was analysis in the kinetic mode. The rapid mixing of protein with various concentration of GuHCl analysis yielded an exponential decay of the spectral intensity at the selected wavelength (**Figure 3**). Curve fitting using the first order reaction revealed the activation energy of unfolding in an absence of denaturant at 25 kcal/mol.

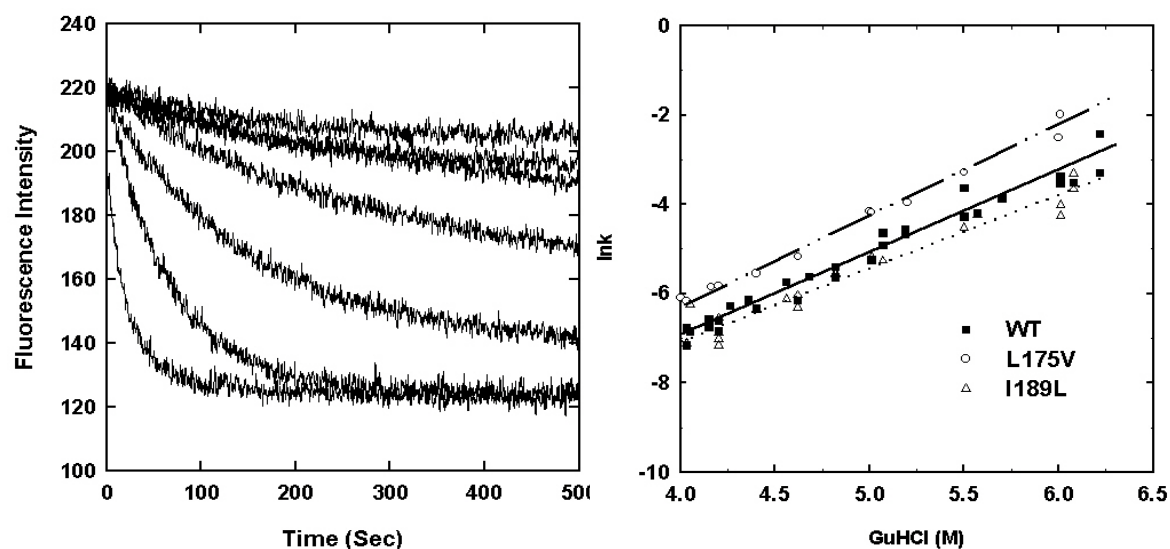


Figure 3: Exponential changes of fluorescence intensity upon mixing with GuHCl (A) and a plot of rate constant as a function of denaturant (B)

Construction of energy profile for the unfolding

A combined data of free energy and activation energy was applied for a construction of an energy profile of the wild type Cry4B. The profile simply demonstrated a pathway from the native state at lower energy to the unfolded state at higher energy. The energetic gap between the two states was corresponding to the free energy and the transitional barrier was according to the activation energy (**Figure 4**).

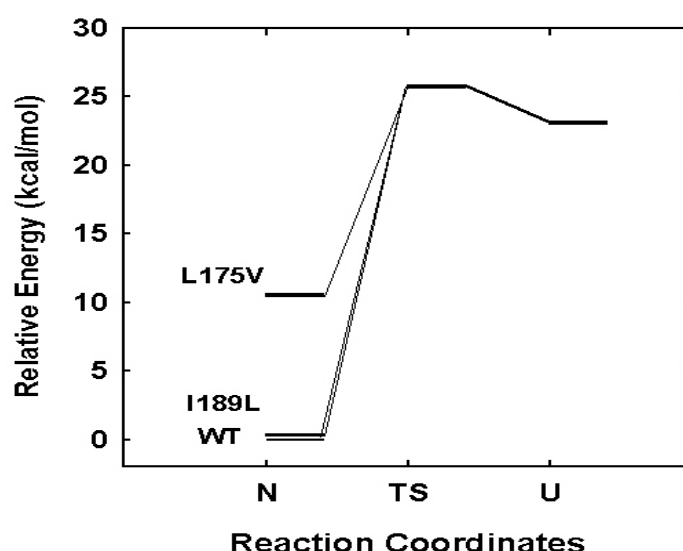


Figure 4: Energy profiles for the unfolding of mutant toxins compared to their template protein

Enhancing of molecular stability by a removal of tryptic site

The active Cry4B protein obtained from a proteolytic processing is generally processed further at the residue R203 to yield two associated polypeptide fragments. The R203Q mutant was constructed with the tryptic site removed to assess for the stability compared to the wild type. After the unfolding analysis in both steady state and kinetic mode, the energy profile was found with a decreased energy gap between folded and unfolded states. This is implied that the folded state of R203Q mutant containing a single chain molecule is more stable than the original wild type with two associated fragments. This mutant was then used further as a template protein for other site-directed mutagenesis experiments.

Destabilization of molecular structure by perturbation on the domain interior

Hydrophobic residues on the conserved block I (L175 and I189) were assessed by amino acid substitution. The generated mutant L175V was characterized by

unfolding experiment and found to have an increased energy in the folded state (**Figure 4**). However the energy profile of I189L was found to be identical to the template protein. The destabilization effect in the L175V mutant was described from the crystal structure of Cry4B as a direct perturbation on the side-chain packing within the hydrophobic interior of domain I.

Critical role of interaction among the conserved block

Stabilization role of residues on the five conserved blocks was tested by site-direct mutagenesis. A group of mutants was found to behave like their template protein. The solubility, proteolytic processing and energy profile of unfolding was identical to their template. However another group of mutants could not be solubilized in the native buffer or they were subject to proteolytic degradation during the processing. It is suggesting that these proteins are unable to fold to the native conformation as a result from amino acid changes at specific positions (**Figure 5**). Examination on the crystal structure has identified that the stabilizing interactions in these mutants were abolished. These interactions are either hydrogen bonding or electrostatic interaction forming between the conserved blocks that help stabilize the tertiary structure of protein. This type of interaction is highly critical for the Cry4B protein to adopt the correct native conformation. The group of mutants with native-like property was also examined and compared. The data showed that the removal of stabilizing interaction on secondary structure did not affect the ability to folded to the correct structure and the molecular stability upon unfolding.

Mutation	Block	Interaction	Distance (Angstrom)	Stabilization	Effect on folding
R190A	1	R190 --- N279	2.89	H-bond	Yes
D191A	1	D191 --- Y126	2.79	H-bond	Yes
R251A	2	R251 --- D248	2.89	Electrostatic	No
D259N	2	D259 --- R289	2.78	Electrostatic	Yes
H466Y	3	H466 --- I 468	3.00	H-bond	No
D470N	3	D470 --- N472	2.98	H-bond	No
K487A	3	K487 --- N473	2.91	H-bond	Yes
R540A	4	R540 --- E629	2.82	Electrostatic	Yes
R542A	4	R542 --- R627	2.94	H-bond	Yes
R627A	5	R627 --- Y272	3.12	H-bond	Yes
E629Q	5	E629 --- Y272	2.52	H-bond	Yes

E629 --- R627

2.73

H-bond

Acknowledgements

This work was supported by the Thailand Research Fund under RTA43 to Sakol Panyim and RSA44 to Chartchai Krittanai.

References

1. Hofte, H. and Whiteley, H. R. (1989) Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiol. Rev.* **53**: 242-255.
2. Angsuthanasombat, C., Chungjatupornchai, W., Kertbundit, S., Luxananil, P., Settasatian, C., Wilairat, P. and Panyim, S. (1987) Cloning and expression of 130-kd mosquito-larvicidal δ -endotoxin gene of *Bacillus thuringiensis* var. *israelensis* in *Escherichia coli*. *Mol. Gen. Genet.* **208**: 384-389.
3. Rajamohan, F., Hussain, S. R. A., Cottrill, J. A., Gould, F., and Dean, D. H. (1996) Mutations at domain II loop 3, of *Bacillus thuringiensis* CryIAa and CryIAc δ -endotoxin suggest loop 3 is involved in initial binding to lepidopteran midguts. *J. Biol. Chem.* **271**: 25220-25226.
4. Schwartz, J. L., Juteau, M., Grochuski, O., Cyler, M., Prefontaine, G., Brouss, R. and Masson, L. (1997). Restriction of intramolecular movements within the CryIAa toxin of *Bacillus thuringiensis* through disulfide bond engineering. *FEBS. Lett.* **410**: 397-402.
5. Uawithya, P., Tuntitipawan, T., Katzenmier, G., Panyim, S. and Angsuthanasombat, C. (1998) Effects on larvicidal activity of single proline substitutions in $\alpha 3$ and $\alpha 4$ of the *Bacillus thuringiensis* Cry4B toxin. *Biochem. Mol. Biol. Int* **44**: 825-832.
6. Sramala, I., Uawithya, P., Chanama, U., Leetachewa, S., Krittanai, C., Katzenmeier, G., Panyim, S. & Angsuthanasombat, C. (2000) Single proline substitutions of selected helices of the *Bacillus thuringiensis* Cry4B toxin affect inclusion solubility and larvicidal activity. *J. Biochem. Mol. Biol. & Biophys.* **3**: 1-7.
7. Li, J., Carroll, L. and Ellar, D. J. (1991) Crystal structure of insecticidal δ -endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution. *Nature* **353**: 815-821.
8. Grochuski, P., Masson, L., Borisova, S., Pusztai-Carey, M., Schwartz, J. L., Brousseau, R. and Cyler, M. (1995) *Bacillus thuringiensis* CryIAa insecticidal toxin: crystal structure and channel formation. *J. Mol. Biol.* **254**: 447-464.
9. Gazit, E. and Shai, Y. (1993) Structural and functional characterization of the $\alpha 5$ segment of *Bacillus thuringiensis* δ -endotoxin. *Biochemistry* **32**: 3429-3436.
10. Gazit, E., Bach, D., Kerr, I. D., Sansom, M. S., Chejanovsky, N., and Shai, Y. (1994) The $\alpha 5$ segment of *Bacillus thuringiensis* δ -endotoxin: *in vitro* activity, ion channel formation and molecular modelling. *Biochem. J.* **304**: 895-902.
11. Gazit, E. and Shai, Y. (1995) The assembly and organization of the $\alpha 5$ and $\alpha 7$ helices from the pore-forming domain of *Bacillus thuringiensis* δ -endotoxin. *J. Biol. Chem.* **270**: 2571-2578.
12. Lee, M. K., Milne, R. E., Ge, A. Z. and Dean, D. H. (1992) Location of a *Bombyx mori* receptor binding region on a *Bacillus thuringiensis* δ -endotoxin. *J. Biol. Chem.* **267**: 3115-3121.
13. Wu, D., and Aronson, A. I. (1992) Localized mutagenesis defines regions of the *Bacillus thuringiensis* δ -endotoxin involved in toxicity and specificity. *J. Biol. Chem.* **267**: 2311-2317.
14. Choma, C. T., Surewicz, W. K., Carey, P. R., Pozsgay, M., Raynor, T. and Kaplan, H. (1990) Unusual proteolysis of the protoxin and toxin from *Bacillus thuringiensis* structural implications. *Eur. J. Biochem.* **189**: 523-527.
15. Choma, C. T., Surewicz, W. K. and Kaplan, H. (1991) The toxin moiety of the *Bacillus thuringiensis* protoxin undergoes a conformational change upon activation. *Biochem. Biophys. Res. Comm.* **179**: 933-938.
16. Convents, D., Houssier, C., Lasters, I. and Lauwereys, M. (1990) The *Bacillus thuringiensis* δ -endotoxin: Evidence for a two domain structure of the minimal toxin fragment. *J. Biol. Chem.* **265**: 1369-1375.

17. Venugopal, W. G., Wolfersberger, M. G. and Wallace, B. A. (1992) Effects of pH on conformational properties related to the toxicity of *Bacillus thuringiensis* δ -endotoxin. *Biochem. Biophys. Acta.* **1159**: 185-192.
18. Feng, Q. and Becktel, W. J. (1994) pH-induced conformational transitions of CryIAa, CryIAC and CryIIIA δ -endotoxins in *Bacillus thuringiensis*. *Biochemistry* **33**: 8521-8526.
19. Potekhin, S. A., Loseva, O. I., Tiktopulo, E. I. and Dobritsa, A. P. (1999) Transition state of the rate-limiting step of heat denaturation of Cry3A δ -endotoxin. *Biochemistry* **38**: 4121-4127.
20. Chen, X. J., Lee, M. K. and Dean, D. H. (1993) Site-directed mutagenesis in a highly conserved region of *Bacillus thuringiensis* δ -endotoxin affect inhibition of short circuit current across *Bombyx mori* midguts. *Proc. Natl. Acad. Sci. USA.* **90**: 9041-9045.
21. Bosch, D., Visser, B. and Stiekema, W. J. (1994) Analysis of non-active engineered *Bacillus thuringiensis* crystal proteins. *FEMS Microbiol. Lett.* **118**: 129-133.
22. Bradford, M. M. (1976) *Anal. Biochem.*, **72**, 248.
23. Nozaki, Y. (1972) in *Methods in zymology*, vol. **26**, Academic Press, Orlando, FL, pp. 145-306.

MOLECULAR BASIS OF MEMBRANE PORE-FORMATION BY THE *BACILLUS THURINGIENSIS* Cry4 LARVICIDAL PROTEINS

Chanan Angsuthanasombat, Yodsoi Kaninthronkul, Walirat Pornwiroon, Issara Sramala, Sompob Leetachewa, Somsri Sakdee, Chaweewan Shimwai, Panadda Boonserm, Gerd Katzenmeier and Sakol Panyim

Laboratory of Molecular Biophysics, Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakornpathom 73170, Thailand

Summary

We have previously demonstrated that $\alpha 4$ and $\alpha 5$ of the 130-kDa *Bacillus thuringiensis* Cry4B mosquito-larvicidal protein are important determinants of toxicity against *Aedes aegypti* larvae, likely being involved in membrane-pore formation. Further analysis revealed a crucial role in toxicity for the positively charged side-chain of Arg-158 in helix 4, conceivably involved in the passage of ions through the pore. In this report, directed mutations within the loop linking $\alpha 4$ and $\alpha 5$ of Cry4B were performed and revealed that Asn-166 and Tyr-170 are critically involved in larvicidal activity, especially polarity at position 166 and aromaticity at position 170. A crucial role in toxin activity was also revealed for the conserved aromatic residue at position 202 within the $\alpha 4$ - $\alpha 5$ loop of the Cry4A toxin. These results support the notion that an aromatic structure of the highly conserved tyrosine residue within the $\alpha 4$ - $\alpha 5$ loop is an essential prerequisite for toxic action of the Cry δ -endotoxins.

Keywords: Aromaticity, *Bacillus thuringiensis*, δ -endotoxins, ion channel, larvicidal activity

Introduction

The general mechanism of gut epithelial cell disruption by different *Bacillus thuringiensis* (Bt) δ -endotoxins is evidenced to be the formation of lytic pores in the susceptible insect membrane [see Schnepf *et al.*, 1998 for reviews]. When ingested by susceptible larvae, toxin inclusions are solubilised by the alkaline pH of the larval midgut and the protoxins are activated by gut proteases. It is believed that the activated toxins then bind to midgut epithelial cells *via* specific receptors, and insert into the microvillar membrane to form ion channels or leakage pores that cause cell swelling and eventually death by colloid-osmotic lysis [Knowles & Ellar, 1987]. However, an entire characterisation at the molecular level of the pore-forming process mediated by these insecticidal proteins has not yet been obtained.

The X-ray crystal structure of five different Cry toxins, Cry1Aa [Grochulski *et al.*, 1995], Cry2Aa [Morse *et al.*, 2001], Cry3A [Li *et al.*, 1991], Cry3Bb [Galitsky *et al.*, 2001] and Cry4B [Boonserm *et al.*, 2003], reveals Cry proteins consisting of three distinct domains, and it is believed that each domain has a defined function. Structurally, it is

immediately apparent that domain I is likely to be the transmembrane pore-forming apparatus. This domain contains five amphipathic helices ($\alpha 3$, $\alpha 4$, $\alpha 5$, $\alpha 6$ and $\alpha 7$) that are theoretically long enough to span the bilayer lipid membrane and form a lytic pore [Li *et al.*, 1991]. The possibility that this α -helical bundle in domain I is essential for pore formation is supported by the feature that it is highly conserved in all Cry toxins [Höfte & Whiteley, 1989].

The molecular mechanism of membrane insertion and pore formation of the Cry toxins is now described in an ‘umbrella’ model [Knowles, 1994]. In this model, $\alpha 4$ and $\alpha 5$ form a helical hairpin to initiate membrane penetration upon specific receptor-binding in which structural rearrangement of the toxin occurs. After insertion of this hairpin, the other helices spread over the membrane surface followed by oligomerisation of the toxin [Gazit *et al.*, 1998; Guereca & Bravo, 1999], resulting in formation of an initial tetrameric pore [Schwartz *et al.*, 1997]. Currently, this model is supported by a number of experiments demonstrating a crucial role of $\alpha 4$ and $\alpha 5$ in membrane penetration and pore formation of different Cry toxins [Schwartz *et al.*, 1997; Gazit *et al.*, 1998; Kumar & Aronson, 1999; Masson *et al.*, 1999; Nunes-Valdes *et al.*, 2001]. Recent studies have clearly demonstrated that the helix 4-loop-helix 5 hairpin is more active in membrane penetration than each of the isolated helices or their mixtures, consistent with its function as the membrane-inserted portion of the Cry toxins [Gerber & Shai, 2000].

Experimental Procedures

Construction of Mutant Toxin Plasmids

Recombinant plasmids, pMEx-B4A encoding the 130-kDa Cry4A toxin, which has been reconstructed in the pMEx8 expression vector [Buttcher *et al.*, 1990] under control of the *tac* promoter together with the *cry4B* promoter [Angsuthanasombat *et al.*, 1987] and pMU388 encoding the 130-kDa Cry4B toxin [Angsuthanasombat *et al.*, 1987], were used as a template for site-directed mutagenesis. Mutant plasmids were generated by polymerase chain reaction (PCR) using a pair of mutagenic primers purchased from Prologo Inc. (Singapore) and *Pfu* DNA polymerase, following the procedure of the QuickChangeTM Mutagenesis Kit (Stratagene). All mutant clones with the required mutation were first identified by restriction endonuclease digestion of the plasmids, and verified by DNA sequencing, using a BigDyeTM Terminator Cycle Sequencing Kit (Perkin-Elmer).

Toxin Expression and Characterisation

The wild type and mutant toxin genes were expressed in *E. coli* strain JM109 under control of the *lacZ* promoter. Cells were grown in LB medium plus 100 μ g/ml of ampicillin until OD₆₀₀ reached 0.4-0.5 and incubation was continued for another 4 hrs after addition of IPTG to a final concentration of 0.1 mM. *E. coli* cultures expressing each mutant as inclusion bodies were harvested by centrifugation, resuspended in 1 ml of distilled water and then disrupted in a French Pressure Cell at 16,000 psi. The crude lysates were centrifuged at 8,000 *g* for 5 min and pellets obtained were washed 3 times in distilled water. Protoxin inclusions (1 mg/ml) were solubilised in 50 mM Na₂CO₃, pH 9.0 and incubated at 37°C for 60 min as described

previously [Angsuthanasombat *et al.*, 1991]. After centrifugation for 10 min, the supernatants were analysed by SDS-15% (w/v) PAGE in comparison with the inclusion suspension. The solubilised protoxins were assessed for their proteolytic stability by digestion with TPCK treated) at a protoxin:trypsin ratio of 20:1 (w/w) for 16 hrs [Angsuthanasombat *et al.*, 1991].

Mosquito-Larvicidal Assays

Larvicidal activity assays were performed as previously described [Angsuthanasombat *et al.*, 1992] using 2-day old *A. aegypti* larvae reared from eggs supplied by the mosquito-rearing facility of the Institute of Molecular Biology and Genetics, Mahidol University. About 500 larvae were reared in a container (22×30×10 cm deep) with approximately 3 l of distilled water supplemented with 0.2-0.3 g of rat diet pellets. In the assays, 1 ml of *E. coli* suspension (ca. 10⁸ cells) was added to a 48-well microtitre plate (11.3 mm well diameter), with 10 larvae per well and a total of 100 larvae for each type of *E. coli* sample. *E. coli* cells containing pMU388 and pUC12 vector were used as positive and negative controls, respectively. Mortality was recorded after incubation for 24 hrs.

Results and Discussion

Directed Mutations within the α 4- α 5 Loop of Cry4B

As predicted from the homology-based 3D model of Cry4B, the α 4- α 5 loop comprises eight amino acids of which one is charged *i.e.* Glu-171 and two are polar *i.e.* Asn-166 and Tyr-170. With the exception of Asn-166, both Tyr-170 and Glu-171 are structurally conserved among the known Cry toxins (see **Fig. 1a & b**) that these loop residues could play a crucial role in Cry4B toxicity. In this study, we therefore initially generated three Cry4B loop mutants in which Asn-166, Tyr-170 and Glu-171 were substituted with alanine. When each mutant toxin was expressed in *E. coli* upon IPTG induction, they were all predominantly produced as sedimentable inclusion bodies and the protein expression level was comparable to the wild type Cry4B toxin. *E. coli* cells expressing each type of the mutant toxin were tested for their relative biological activity against *A. aegypti* larvae. Alanine substitutions of Asn-166 and Tyr-170 almost completely abolished the Cry4B bioactivity, although mutation at Glu-171 showed only a small decrease in larvicidal activity (**Fig. 2**). These results suggested that Asn-166 and Tyr-170 play an important role in larvicidal activity of the Cry4B toxin.

When Asn-166 was further substituted with Asp, Gln, Arg, Cys or Ile, it revealed that substitutions of Asn-166 with polar amino acids could retain over 80% of the wild type toxicity while substitution with a non-polar residue *i.e.* isoleucine almost totally abolished the toxicity (see **Fig. 2**). These results suggested that the polarity at position 166 located in the α 4- α 5 loop is important for larvicidal activity of the Cry4B toxin. From molecular modelling of a putative toxin-induced pore (see **Fig. 3a**), Asn-166 points toward the pore lumen and could form hydrogen bonds with water molecules. A crucial role in toxin mechanism at this critical position is conceivably to be involved in the formation of hydrogen bonds with water to stabilise

the loop structure or it may be involved in the passage of ions through the channel. Whether these possibilities can be generalised remains to be elucidated.

For Tyr-170, it was also converted to Asp, Arg, Leu, Trp or Phe. Interestingly, substitutions of this critical tyrosine residue with only the aromatic residues (Trp or Phe) were shown to remain Cry4B toxicity against mosquito larvae (**Fig. 2**). These results together with the fact that the tyrosine residue at this position is a highly conserved amino acid among the Cry toxins strongly imply a general requirement for an aromatic structure at this crucial position. A function for Tyr-170 within the $\alpha 4$ - $\alpha 5$ loop of Cry4B may conceivably be an interaction with the phospholipid head groups for stabilising the oligomeric pore structure (**Fig. 3b**).

Directed Mutations within the $\alpha 4$ - $\alpha 5$ Loop of Cry4A

Based on multiple sequence alignments of the known *Bt* Cry toxin structures and the homology-based Cry4 models, the interhelical loop connecting $\alpha 4$ and $\alpha 5$ of the Cry4A toxin is composed of sixteen amino acids with the majority being polar and charged residues (**Fig. 1a**). Previously, we have shown that polarity and aromaticity for Asn-166 and Tyr-170, respectively, in the $\alpha 4$ - $\alpha 5$ loop are critically involved in larvicidal activity of the Cry4B toxin (see above). Here, we have also constructed several mutants in the $\alpha 4$ - $\alpha 5$ loop region of Cry4A in order to determine a residue responsible for the toxin activity. Two negatively charged (Asp-198 and Asp-200) and four polar (Asn-190, Asn-195, Tyr-201 and Tyr-202) residues were selected for initially substitution with alanine *via* PCR-based directed mutagenesis. The loop mutant toxins were expressed in *E. coli* under inducible control of the *tac* promoter. Upon addition of IPTG to the mid-exponential phase cultures, all mutant Cry4A protoxins were predominantly produced in the form of sedimentable inclusion bodies. When *E. coli* lysates were analysed by SDS-PAGE, the levels of protein expression for all mutants were found to be comparable to that of the wild-type (data not shown).

To assess the solubility of the mutant protoxin inclusions in comparison with that of the wild-type, experiments were carried out using carbonate buffer, pH 10.0. The amounts of the 130-kDa Cry4A soluble proteins in the supernatant were compared with those of the proteins initially used so as to determine the percentage of toxin solubilisation. The toxin inclusions of the N190A, D198A, D200A and Y201A mutants were soluble in this buffer, giving approximately 60-70% solubility, which is comparable to the solubility of the wild-type inclusions under similar conditions. On the other hand, a nearly complete loss of the inclusion solubility was observed for the two remaining mutants, N195A and Y202A (data not shown). However, toxin inclusions of the two closely related loop-Cry4B mutants, N166A and Y170A as mentioned earlier, were found to be relatively soluble in this buffer (see above). At this stage, the reason for this difference in solubility between the two loop-mutants of Cry4A and Cry4B is unclear. It does however lead to the interesting possibility that single-alanine substitutions at Asn-195 and Tyr-202 of the Cry4A toxin could disturb the structural characteristics that consequently affect toxin-inclusion formation as shown by a drastic decrease in solubility.

To determine the effect of the $\alpha 4$ - $\alpha 5$ loop mutations on the Cry4A bioactivity, *E. coli* cells expressing each mutant toxin were tested for their biological

activity towards *A. aegypti* mosquito-larvae. Replacement at only Tyr-202 with alanine almost completely abolished larvicidal activity, whereas alanine-substitutions at the other positions (Asn-190, Asn-195, Asp-198, Asp-200 and Tyr-201) did not affect the Cry4A toxicity. Further analysis *via* specific mutations revealed that conversion of this critical tyrosine residue to cysteine resulted in a drastic loss of toxicity, whilst replacement with the aromatic residue *i.e.* phenylalanine, still retained the high level of larvicidal activity (see Fig. 4). The level of protein expression of both Y202C and Y202F mutant toxins was approximately the same as that of the wild-type. These results, together with the highly structural conserved level of the tyrosine residue at this position among the Cry toxins (see Fig. 1b), suggest an essential feature of an aromatic structure at this critical position for the toxin activity. The data further support our previous findings that Tyr-170 in the α 4- α 5 loop plays an important role in larvicidal activity of the 130-kDa Cry4B toxin, since substitutions with only the aromatic residues, *i.e.* Phe or Trp, were shown to restore the bioactivity towards mosquito-larvae (see above).

For *in vitro* solubility, like the Y202A mutation, the substitution of Tyr-202 with Cys reduced the solubility *in vitro* of toxin inclusions, while a conversion to Phe still exhibited the same solubility characteristics as that of the wild-type (data not shown). Although insolubility of toxin inclusions and the loss of toxicity are seemingly correlated for both Y202A and Y202C mutants, the inclusion solubility *in vitro* may not necessarily reflect toxin activity *in vivo* as observed for the N195A mutant which was insoluble in the carbonate buffer, but still bioactive (see Fig. 4). It has been demonstrated that single-proline substitution in α 6 of Cry4B dramatically perturbed the inclusion dissolvability, but did not affect its larvicidal activity [Sramala *et al.*, 2000]. Also, it has been shown that the difference detected in solubilisation *in vitro* for the cloned Cry4A toxin inclusions, which were purified from two different *Bt* recipient strains, is not a factor for toxicity *in vivo* [Angsuthanasombat *et al.*, 1992]. Presumably, the larval gut proteases *in vivo* might facilitate the dissolution of the ingested toxin inclusions that would negate the differences between the observed larvicidal activities of the bioactive N195A and non-active Y202A or Y200C mutants.

Studies with several membrane proteins have indicated that the aromatic residues are predominantly found at or near the lipid-water interface [Ulmschneider & Sansom, 2001]. These aromatic residues have been proposed to function in anchoring the proteins into the membrane through interactions of their aromatic rings with phospholipid head groups [Yau *et al.*, 1998; Killian & Von Heijne, 2000], maintaining rigidity in the periphery of the transmembrane segments [Tsang & Saier, 1996], allowing vertical mobility of the transmembrane helical region with respect to the membranes [Schiffer & Deber, 1990], facilitating translocation of the periplasmic portion of proteins through the membrane, thereby acting as determinants of protein orientation [Schiffer *et al.*, 1992]. Taken together, a function for Tyr-202 in the α 4- α 5 loop of the Cry4A toxin may conceivably be an interaction with the phospholipid head groups for stabilising the oligomeric pore structure.

In conclusion, this study additionally demonstrates that the aromaticity of Tyr-202 in the α 4- α 5 loop plays a crucial role in the Cry4A toxicity that further supports the notion that the aromatic structure of the highly conserved tyrosine residue within the loop connecting the two transmembrane helices, α 4 and α 5, is essential for toxic

action of the *Bt* Cry δ -endotoxins. However, further studies are still required to elucidate the exact role of this critical residue in toxin function.

Acknowledgements

This work was supported in part by the Thailand Research Fund to SP.

References

1. Angsuthanasombat, C., Chungjatupornchai, W., Kertbundit, S., Luxananil, P., Settasatian, C., Wilairat, P. and Panyim, S. (1987) Cloning and expression of 130-kDa mosquito-larvicidal δ -endotoxin gene of *Bacillus thuringiensis* var *israelensis* in *Escherichia coli*. *Mol. Gen. Genet.* **208**, 384-389.
2. Angsuthanasombat, C., Crickmore, N. and Ellar D. J. (1991) Cytotoxicity of a cloned *Bacillus thuringiensis* subsp. *israelensis* CryIVB toxin to an *Aedes aegypti* cell line. *FEMS Microbiol. Lett.* **83**, 273-276.
3. Angsuthanasombat, C., Crickmore, N. and Ellar, D. J. (1992) Comparison of *Bacillus thuringiensis* subsp. *israelensis* CryIVA and CryIVB cloned toxins reveals synergism *in vivo*. *FEMS Microbiol. Lett.* **94**, 63-68.
4. Boonserm, P., Ellar, D.J. and Li, J. (2003) Crystallisation and preliminary X-ray diffraction studies of a mosquito-larvicidal toxin from *Bacillus thuringiensis* subsp. *israelensis*. *Acta Cryst.* **D59**, 591-594.
5. Buttcher, V., Ruhlmann, A. and Cramer, F. (1990) Improved single-stranded DNA producing expression vectors for protein manipulation in *Escherichia coli*. *Nucleic Acids Res.* **18**, 1075.
6. Galitsky, N., Cody, V., Wojtczak, A., Ghosh, D., Luft, J. R., Pangborn, W. and English, L. (2001) Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1101-1109.
7. Gerber, D. and Shai, Y. (2000) Insertion and organization within membranes of the delta-endotoxin pore-forming domain, helix 4-loop-helix 5, and inhibition of its activity by a mutant helix 4 peptide. *J. Biol. Chem.* **275**, 23602-23607.
8. Grochulski, P., Masson, L., Borisova, S., Pusztai-Carey, M., Schwartz, J.-L., Brousseau, R. and Cygler, M. (1995) *Bacillus thuringiensis* CryIA(a) insecticidal toxin: crystal structure and channel formation. *J. Mol. Biol.* **254**, 447-464.
9. Guereca, L. and Bravo, A. (1999) The oligomeric state of *Bacillus thuringiensis* Cry toxins in solution. *Biochem. Biophys. Acta.* **1429**, 342-350.
10. Höfte, H. and Whiteley, H. R. (1989) Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiol. Rev.* **53**, 242-255.
11. Killian, J. A. and Von Heijne, G. (2000) How proteins adapt to membrane-water interface. *Trends Biochem. Sci.* **25**, 429-434.
12. Knowles, B. H. (1994) Mechanism of action of *Bacillus thuringiensis* insecticidal δ -endotoxins. *Adv. Insect. Physiol.* **24**, 275-308.
13. Knowles, B. H. and Ellar, D. J. (1987) Colloid-osmotic lysis is a general feature of the mechanism of action of *Bacillus thuringiensis* delta endotoxin with difference insect specificity. *Biochim. Biophys. Acta* **924**, 509-518.
14. Kumar, A. S. and Aronson, A. I. (1999) Analysis of mutations in the pore-forming region essential for insecticidal activity of a *Bacillus thuringiensis* delta-endotoxin. *J. Bacteriol.* **181**, 6103-6107.
15. Li, J., Carroll, J. and Ellar, D. J. (1991) Crystal structure of insecticidal delta-endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution. *Nature* **353**, 815-821.

16. Masson, L., Tabashnik, B. E., Liu, Y. B., Brousseau, R. and Schwartz, J.-L. (1999) Helix 4 of the *Bacillus thuringiensis* Cry1Aa toxin lines the lumen of the ion channel. *J. Biol. Chem.* **274**, 31996-32000.
17. Morse, R. J., Yamamoto, T. and Stroud, R. M. (2001) Structure of Cry2Aa suggests an unexpected receptor binding epitope. *Structure (Camb.)* **9**, 409-417.
18. Nunez-Valdez, M.-E., Sanchez, J., Lina, L., Guereca, L. and Bravo, A. (2001) Structural and functional studies of α -helix 5 region from *Bacillus thuringiensis* Cry1Ab δ -endotoxins. *Biochim. Biophys. Acta* **1546**, 122-133.
19. Pawagi, A. B. and Deber, C. M. (1990) Ligand-dependent quenching of tryptophan fluorescence in human erythrocyte hexose transport protein. *Biochemistry* **29**, 950-955.
20. Schiffer, M., Chang, C. H. and Stevens, F. J. (1992) The functions of tryptophan residues in membrane proteins. *Protein Eng.* **5**, 213-214.
21. Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. R. and Dean, D. H. (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* **62**, 775-806.
22. Schwartz, J.-L., Juteau, M., Grochulski, P., Cygler, M., Prefontaine, G., Brousseau, R. and Masson, L. (1997) Restriction of intramolecular movements within the Cry1Aa toxin molecule of *Bacillus thuringiensis* through disulfide bond engineering. *FEBS Lett.* **410**, 397-402.
23. Sramala, I., Leetachewa, S., Krittanai, C., Katzenmeier, G., Panyim, S. and Angsuthanasombat, C. (2001) Charged residues screening in helix 4 of the *Bacillus thuringiensis* Cry4B toxin reveals one critical residue for larvicidal activity. *J. Biochem. Mol. Biol. Biophys.* **5**, 219-225.
24. Sramala, I., Uawithya, P., Chanama, U., Leetachewa, S., Krittanai, C., Katzenmeier, G., Panyim, S. and Angsuthanasombat, C. (2000) Single proline substitutions of selected helices of the *Bacillus thuringiensis* Cry4B toxin affect inclusion solubility and larvicidal activity. *J. Biochem. Mol. Biol. Biophys.* **4**, 187-193.
25. Tsang, S. and Saier, M. H., Jr. (1996) A simple flexible program for the computational analysis of amino acyl residue distribution in proteins: application to the distribution of aromatic versus aliphatic hydrophobic amino acids in transmembrane alpha-helical spanners of integral membrane transport proteins. *J. Comput. Biol.* **3**, 185-190.
26. Ulmschneider, M. B. and Sansom, M. S. P. (2001) Amino acid distributions in integral membrane protein structures. *Biochim. Biophys. Acta* **1512**, 1-14.
27. Yau, W. M., Wimley, W. C., Gawrisch, K. and White, S. H. (1998) The preference of tryptophan for membrane interfaces. *Biochemistry* **37**, 14713-14718.

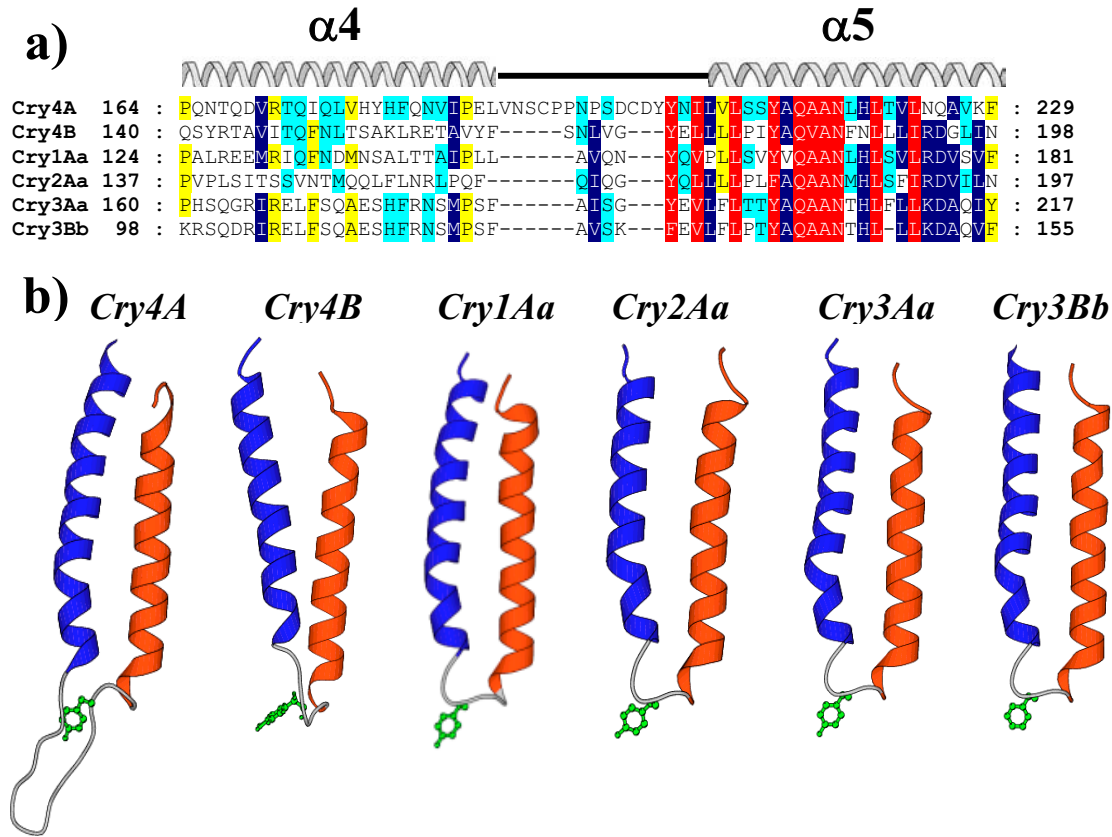


Figure 1 (a) Multiple sequence alignment of the transmembrane $\alpha 4$ -loop- $\alpha 5$ fragment of Cry4A with those of the crystal structures of Cry1Aa, Cry2Aa, Cry3Aa and Cry3Bb toxins and the homology-based Cry4B model. The sequences were aligned using the program Clustal X. The corresponding $\alpha 4$ -loop- $\alpha 5$ is shown above the sequences. (b) Side views of the $\alpha 4$ -loop- $\alpha 5$ helical hairpins of the homology-based Cry4A and Cry4B models, and the known structures of Cry1Aa, Cry2Aa, Cry3Aa and Cry3Bb. Gray ribbons represent $\alpha 4$ and $\alpha 5$. The conserved loop-residue, Tyr, is shown in ball-and-stick in all six helical hairpins.

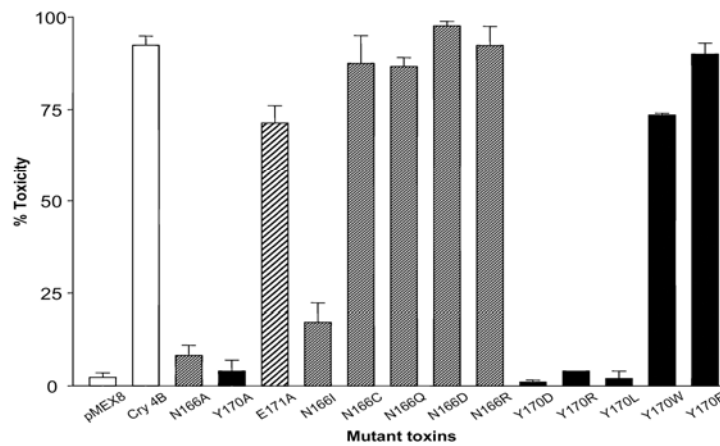


Figure 2 Mosquito-larvicidal activities of *E. coli* cells expressing the wild type Cry4B toxin or mutant toxins against *A. aegypti* larvae. Error bars indicate standard error of the means from three independent experiments.

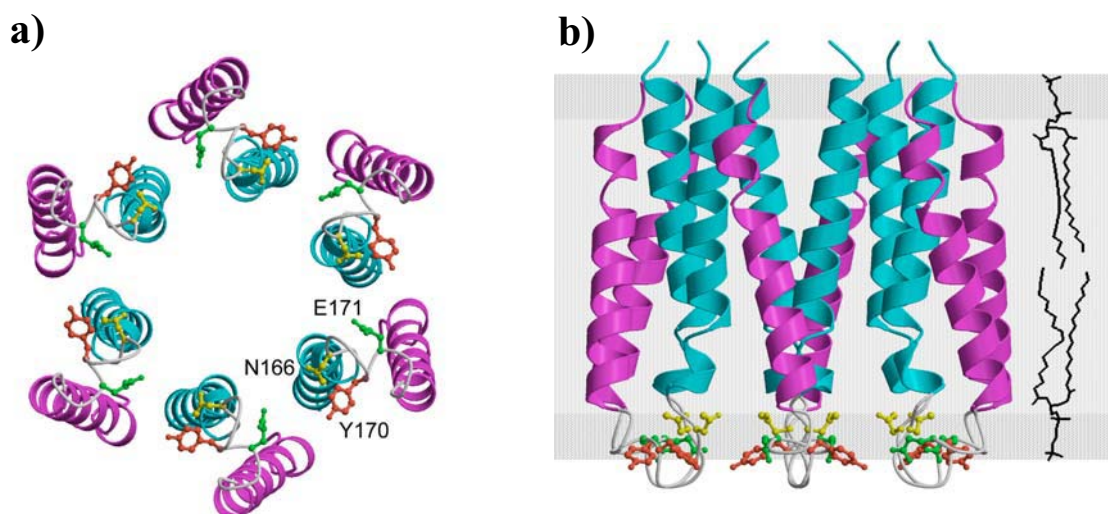


Figure 3 (a) Bottom view of the model of a putative oligomeric pore consisting of six copies of the α 4- α 5 hairpin of the Cry4B toxin. Asn-166, Tyr-170 and Glu-171 are shown. (b) The Tyr-170 residues are found at the membrane-water interface after oligomerisation and pore formation. Rectangular boxes represent the planar lipid membrane bilayers. Helices 4 and 5 are shown as gray strands.

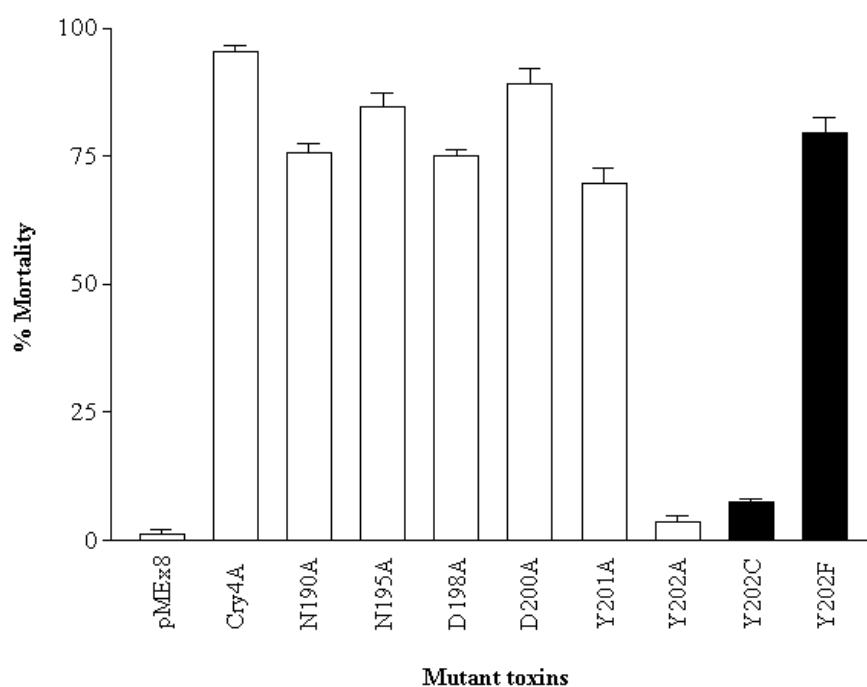


Figure 4 Mosquito-larvicidal activities of *E. coli* cells expressing the Cry4A wild-type or its mutant toxins (N190A, N195A, D198A, D200A, Y201A, Y202A, Y202C and Y202F) against *A. aegypti* larvae. Error bars indicate standard errors of the mean from three independent experiments.

CLONING AND EXPRESSION OF KAZAL-TYPE SERINE PROTEINASE INHIBITORS OF THE BLACK TIGER SHRIMP, *PENAEUS MONODON*

Boonyarin Jarasrassamee,¹ Anchalee Tassanakajon,¹ Sarawut Klinbunga,² and Sakol Panyim³

¹Shrimp Molecular Biology and Genomics Research Laboratory, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330

²Marine Biotechnology Research Unit, National Center for Genetic Engineering and Biotechnology, National Science and Development Agency, Bangkok 10400

³Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakhon Pathom 73170

Abstract

We have isolated cDNA clones of Kazal proteinase inhibitors from haemocyte cDNA libraries of the black tiger shrimp, *Penaeus monodon*. The deduced amino acid sequences of the cDNA clones showed that they consist of 2-5 Kazal domains with variation in the inhibitory reactive site residue, P1 amino acid, indicating variable in inhibitory specificity. The five-domain Kazal proteinase inhibitor from *P. monodon* was expressed in *Escherichia coli*. The recombinant protein inhibits trypsin, chymotrypsin but not subtilisin and elastase.

The mRNA expression of the Kazal inhibitors in shrimp haemocytes examined by Northern analysis showed two transcripts with sizes of approximately 1.7 and 1.8 kb. Tissue specific expression study in several tissue types of shrimps by RT-PCR suggests that the Kazal proteinase inhibitors are exclusively expressed in haemocytes. The expression level of the inhibitors was further analyzed in unchallenged and *Vibrio harveyi* challenged shrimps to examine the influence of bacterial stimulation on transcription of the inhibitors. No significant change in the expression level was found during 48 hr post injection.

Introduction

Serine proteinase inhibitors are widely distributed among living organisms. The important role of these inhibitors is protection against the proteolytic enzymes of many pathogens and insects [1,2]. In plants, they also regulate endogenous proteinases during seed dormancy and reserve protein mobilization [3]. Arthropod serine proteinase inhibitors are believed to have important roles in immune system including inhibition against proteinases of microorganisms or being regulators of host-defense reactions involving blood coagulation, prophenoloxidase activation, or cytokine activation. Four families of serine proteinase inhibitor have been reported in arthropod haemolymph. They are Kazal, Kunitz, α -macroglobulin and serpin families [4].

Among these serine proteinase inhibitors, a four-domain Kazal proteinase inhibitor and its cDNA clone were isolated from blood cells of the crayfish *Pacifastacus leniusculus*. The inhibitor was found to have an inhibitory activity toward chymotrypsin or subtilisin but not trypsin, elastase or thrombin [5].

Nonetheless, the biological function of the Kazal inhibitor in the blood cells of crayfish is not known.

From an expressed sequence tag (EST) analysis of haemocytes of the black tiger shrimp *P. monodon* [6], we isolated cDNAs that showed high homology to the crayfish Kazal proteinase inhibitor. In this study, we expressed the shrimp Kazal-type proteinase inhibitor in *E. coli* and assayed the inhibitory activity against some proteinases. mRNA expression of the inhibitor was examined in several tissue types and in response to bacterial stimulation.

Materials and Methods

Samples

Sub-adult shrimps (weighing 20-25 g) were purchased from a farm and maintained in 80 L aquarium tanks with 20 parts per thousand salinity sea water. Haemolymph was taken and haemocytes were collected by centrifuged at 800 g for 10 minutes. After haemolymph collection, animals were dissected for tissue preparation.

Immune-challenged shrimps were prepared by injection of 100 µl of *vibrio harveyi* innoculum (10^7 CFU/ml). The luminescent bacterium, *V. harveyi* 1526, used in this study was kindly provided by Shrimp Culture Research Center, Charoenpokaphand Group of Company.

RNA preparation

Total RNA was isolated from haemocyte, heart, gill, lymphoid organ, intestine and hepatopancrease of freshly killed animals by using Trizol (Gibco BRL). For time course experiment, total RNA was isolated from haemocytes of 10 animals at 3, 6, 12, 24 and 48 hr after bacterial challenge. The quantity of RNA was determined spectrophotometrically and the quality was checked by electrophoresis on a formaldehyde-agarose gel.

Northern Hybridization

Total RNAs were electrophoretically separated on 1.5% agarose gels containing formaldehyde and transferred to a Hybond N+ nylon membrane (Amersham) by capillary blotting. The RNA blot was hybridized with the [α - 32 P] dCTP-labeled insert of the cDNA clone (sh415) of the shrimp proteinase inhibitor. Hybridization and washing were carried out at 60°C according to the manufacturer's instructions and the membranes were subjected to autoradiography.

RT-PCR

First-strand cDNA was synthesised from total RNA using the AMV Reverse Transcriptase First-strand cDNA synthesis kit (Life Sciences) and 50 ng were used as templates for amplification by the polymerase chain reaction. The PCR was conducted in a volume of 25 µl with a final concentration of 1xPCR buffer (Gibco BRL), 1.5 mM MgCl₂, 0.5 µM of forward (5'TGGCGTGAGTGTCACCTTTCCA3') and reverse primers (5'AAGTCTTGCCATCACTGCCAC3'), 0.2 mM of each dNTP (Gibco BRL) and 1 µg of the cDNA template. Reactions were amplified on a thermal cycler using a 95°C, 2-minute initial denaturation followed by a 95°C, 30-sec denaturation, a 54 °C, 30-sec annealing temperature, a 72°C, 1 minute extension for

30 cycles, and a 72°C, 5-minute final extension. The PCR products were resolved on a 1.5% gel, visualized with ethidium bromide staining and UV illumination.

Semi-quantitative RT-PCR

The PCR conditions were optimized to amplify the proteinase inhibitor and β -actin that was used as an internal control in the same tube. The primers for amplification of β -actin were 5'GCTTGCTGATCCACATCTGCT3' and 5'AGTCGAACATGCAGGCCTATCC3'. Each primer set was used at a concentration of 0.3 μ M. The PCR conditions were tested to select suitable conditions for semiquantitative RT-PCR with both primers (7). The conditions are slightly modified from the standard RT-PCR protocol by increasing the concentration of $MgCl_2$ to 2.5 mM and decreasing the number of PCR cycle to 25 cycles so that the amplification is in the exponential range and has not reach a plateau yet. The PCR products were loaded onto 1.5% agarose gel and stained with ethidium bromide. Images of the PCR products were acquired with a CCD camera and quantified using Genetools analysis software (Syngene)

The expression level of the inhibitor at a particular time was normalized with the internal control (β -actin). Significantly different expression levels were treated using One Way Analysis of Variance (ANOVA) following by a post hoc test (Duncan's new multiple range test),

PCR amplification of the Kazal inhibitor cDNA cloning into pGEM-T easy vector

The plasmid (pSH415) isolated from the cDNA clone, sh415, was used as a PCR template. The following primers (BIOBASIC Inc.) were used for PCR amplifications. Restriction sites are underlined.

5'-CCATGGGATCCGGGCTACGGAAAAGGGGGGAAAATCC-3'

Deleted signal sequence forward primer with *Bam* HI site

5'-ATGGTCGACTAGGTACAGTCTGCGACCACAGATTCC-3'

Reverse primer with *Sal* I site

PCR amplification of the Kazal inhibitor gene and pSH415 was used as PCR template. PCR amplifications were run for 30 cycles of 60s at 94°C, 120s at 56°C and 60s at 72°C. All reacting were performed using standard conditions: in a final volume of 15 μ l, 0.45 μ M of each primer was added in the presence of 1.5 mM $MgCl_2$, 0.2 mM of each dNTP and 0.45 U *Tag* DNA polymerase.

Amplified products were digested with *Bam* HI and *Sal* I and cloned into the pGEM-T easy vector at the same sites. Recombinant pGEM-T easy vector was electroporated into *E.coli* strain JM 109 cells and selected transformants using Ampicillin agar plates. Ampicillin resistance clones were inoculated in 1.5 ml LB medium. pGEM-T/Sh 415 was extracted, cut and run in 1.2% agarose gel electrophoresis for size determination. The band in size of approximately 750 bp was cut and purified using Nucleospin gel extraction kit (MACHEREY-NAGEL).

Subcloning of the Kazal inhibitor gene into pTrcHis 2C

The purified digested serine proteinase inhibitor gene was cloned into pTrcHis 2C vector. Recombinant pTrcHis 2C construct was electroporated into *E.coli* strain JM 109 cels. Transformant selections were performed as described above. DNA

sequencing was used to confirmed the correction of vector and inserted DNA junction.

Expression of the inhibitor

Single colony of *E.coli* containing recombinant pTrcHis 2C was inoculated in 1.5 ml of LB broth including 50 µl/ml of ampicillin and grown overnight in 37°C and 250 rpm controlled incubator as starter. Two hundred microlitres of the culture were added to 10 ml the LB broth and grown the same as starter until OD₆₀₀ reaching 0.6. To induce expression, 100 mM IPTG was then added to the culture to make final concentration being 1 mM. The culture was incubated at 37°C with vigorous shaking. A 1 ml aliquot of cultured cells was collected at this moment for the zero time point sample. The collected cells were centrifuged at 12,000 g for 30 seconds at room temperature. The supernatant was aspirated and the cells were stored at -20°C until used. A 1 ml-expression culture was collected every hour for 5 hours. Each collected aliquot was centrifuged as above. Protein expression was analyzed by SDS-PAGE. *E. coli* strain JM 109 containing uninserted pTrcHis 2C was utilized as a control. It was treated at the same condition as the experimental group.

Proteinase inhibitory gelatin-SDS/PAGE activity

Proteinase inhibitors were detected by activity gel. Proteins were electrophoresed through a gelatin-12% SDS-PAGE. After electrophoresis, the gel was incubated in 2% Triton X-100 overnight and washed with distilled water and incubated in 0.1 M sodium phosphate, pH 7.8 containing 40 µg/ml proteinase for 4 hours, 37°C to degrade gelatin. Bands of undegraded blue-stained gelatin indicate the presence of proteinase inhibitors. (modified from [6]).

Assay of proteinase inhibitor

The inhibitory activity of recombinant inhibitor toward serine proteinases was tested as follows (modified from [7]). A 10 µl portion of the *E.coli* lysate was preincubated with appropriate amount of serine proteinase in 0.1 M Tris/HCl, pH 8.0 at 30 °C for 10 minutes. Residual serine proteinase activity was determined by addition of chromogenic substrate and incubated at 30 °C for 15 minutes. The reaction was terminated by adding equal volume of 50%(v/v) acetic acid. The absorbance was measured at 405 nm. In this experiment, the activities of trypsin and subtilisin (substrate N-benzoyl-PHE-VAL-ARG-*p*-nitroanilide), chymotrypsin (substrate N-succinyl-ALA-ALA-PRO-PHE-*p*-nitroanilide) and elastase (substrate N-succinyl-ALA-ALA-ALA-*p*-nitroanilide) were assayed.

Protein concentration

The protein concentration was assayed according to Bradford [8] with BSA as a standard.

Results

cDNAs of shrimp serine proteinase inhibitors

As a result of EST analysis of haemocytes of the black tiger shrimp *Penaeus monodon* [9], six clones were identified as putative serine proteinase inhibitors by the BLASTX program. They showed 50 to 58 % homology to a four-domain Kazal

serine proteinase inhibitor from crayfish. The open reading frame (ORF) of each clone was identified and 4 full-length cDNA clones were obtained. From the deduced amino acid sequence of the inhibitors, they differ in the number of the Kazal domain. Each domain contains 6 cysteine residues. The deduced amino acid of the two clones, sh415 and sh610, contain complete five and four Kazal domains, respectively, whereas that of the sh1069 clone has two complete and two incomplete Kazal domains. The clone, sh1064, has only two Kazal domains with half of the cysteine residues spacing in different manner (Fig. 1).

It has been suggested that each Kazal domain of the proteinase inhibitor can function independently. Proteinase inhibition is due to the formation of a very stable complex between conformationally constrained reactive site of the Kazal domain and the active site of the proteinase. The reactive site residue, P1 amino acid, generally corresponds to the specificity of the enzyme, e.g. inhibitors with P1 Lys and Arg tend to inhibit trypsin and trypsin-like enzymes, those with P1 Tyr, Phe, Trp, Leu and Met inhibit chymotrypsin and chymotrypsin-like enzymes, and those with P1 Ala and Ser inhibit elastase-like enzymes [10]. Chicken ovoidinhibitor has seven domains, the first four domains are active against trypsin, the fifth domain against chymotrypsin, the sixth and seventh domains against chymotrypsin and elastase [11]. The shrimp inhibitors have different amino acid residues at the putative P1 sites of the Kazal domains suggesting different enzyme specificity (Fig. 1).

```

CRAYFISH
AAR
*
CPSI--CPLNYKPVCGSDLKTYGNSCQLNAAICRNPSLKLYDGP----CIDKP
CPSI--CPLDYNPVCGTDGKTYSNLCALRIEACNNPHLNLRVDYQGE---CRP
CRNG--CTLQYDPVCGTDGKTYSNLCDELVAAACNNPQLNLKVAYKGE---CKQ
CPTI--CTQYDPVCGTDGKTYGNSCELGVAAACNNPQLNNKIAYKGA---CNF
PQQQT

SH415
MANKVALLTLLAVAVAVSGYGKGGKIRL
*
CAKH--CTTIS--PVCGSDGKTYDSRCHLENAA--CGGVSVTFHHAGPCPPPKR
CPGI--CPAVYAPVCGTNGKTYSNLCQLENDRTCNGAFVSKKH-DGRCG
CNPVACPEIYAPVCGSDGKTYDNDCYFQAAV--CKNPDLKKVRDGNCD
CTPLIGCPKNYRPVCGSDGVTYNNDCEFFKVAQ--CKNPALVKVSDTRCE
CNHV--CTEYYPVCGSNGVTYSNICLLNNA--CLDSSIYKVSDBGICG
RRLYLZ

SH1069
MANKVALLTLLAVAVAVSGYGKGGKIRL
*
CAKH--CTTIS--PVCGSDGKTYDSRCPGL----CPAVYA
PVCGTNGKTYSNLCQLENDRTCNGAFVSKKH-DGRCG
CNPVACPEIYAPVCGSDGKTYDNDCYFQAAV--CKNPDLKKVRDGNCD
CTPLIGCPKNYRPVCGSDGVTYNNDCEFFKVAQ--CKNPALV-KVSDTRCE
CLLNNAACLDSSIYKVSDBGICGRKMYL

SH1064
MANKVALLTLLAVAVAVSGYGKGGKIRL
*
CAKH--CTTIS--PVCGSDGK---TYDSR-----CHLENAACGGVSVTFHHAGPCPPPKR
CPGL--CPVYA--PVCGTTGKLTRTYANLRMTPEATVLSFPPSTMDVVGATQC-----CVPDL-----CSRVL

SH610
MLLCKITLIHLLQGFVFNDA NSDHD
*
CIGY--CPEVYDPVCGASNGWTYNDCELQAMIKCQGWNITKTHDQACE
CLKA--CPTTFAPVCGSDNKTYNECVFEVAS--CWDHSLDKASEGACGWGIH
CLQY--CPEVYDPVCGSNGQTYNEELQAAIQCRGLQIAKRHDQACE
CHAT--CPLIHDPVCGTDDRTTYNECFEFTKAS--CWDRSILKKKNGPCD
RKWKYLLEI

```

Figure 1. Alignment of the Kazal domains of the crayfish and *P.monodon* serine proteinase inhibitors. Conserved amino acids are in bold. * indicates putative reactive site P₁ amino acid.

Expression of a five-domain Kazal proteinase inhibitor in E.coli

We cloned a five-domain Kazal proteinase inhibitor sequences of the shrimp into *E.coli*. The full-length clone contains an open reading frame of 801 bp encoding 266 amino acids (Fig. 2). The serine proteinase inhibitor of the crayfish has been reported to consist of a signal sequence of 19 amino acids and a mature protein of 209 amino acids [5]. In *P. monodon*, a signal sequence of 18 amino acids was predicted using SignalP VI.1 software. Therefore, the recombinant clones containing N-terminal deleted sequences were constructed. The cDNA insert was PCR amplified and cloned into pGEM-T easy vector and then subcloned into the expression vector pTrHis2C. The nucleotide sequences of the transformants were confirmed to ensure the correct reading frame.

Expression of the serine proteinase inhibitor in the transformants was induced with 1mM IPTG for 0 to 5 h. After the induction step, the cells were lysed and the lysate was subjected to 12% SDS/PAGE. A protein band with size of approximately 35 kDa was found with increasing banding intensity at 1 to 3 h after induction comparing to the transformant with the parental plasmid (Fig. 3).

SH415

```

1      ATGGCCAAACAAAGTGGCACTCTTGACCCTTCTTGCAGTGGCCGTTGCAGTCTCTGGCTAC
      M  A N K V A L L T L L A V A V A V S G Y
61     GGAAAAGGGGGGAAAATCCGCCTCTGCGCCAAACACTGTACGACCATCTCCCCGTGTGTG
      G K G G K I R L C A K H C T T I S P V C
121    GGCTCTGATGGAAAACTTATGACAGCCGATGCCACCTGGAGAATGCTGCCTGTGGTGGC
      G S D G K T Y D S R C H L E N A A C G G
181    GTGAGTGTCACTTTCCACCATGCCGGACCCTGCCCTCCCCCAAAGAGATGTCCAGGAATA
      V S V T F H H A G P C P P P K R C P G I
241    TGCCCCGCGGTATATGCCCTGTGTGCGGGACCAACGGGAAAACCTACTCGAAGTTATGC
      C P A V Y A P V C G T N G K T Y S N L C
301    CAACTTGAGAAATGACAGAACCTGCAACGGTGCTTTCGTTTCCAAGAAGCACGATGGACGT
      Q L E N D R T C N G A F V S K K H D G R
361    TGTGGTTGCAACCCCATTTGTCGCGTGCCCTGAGATCTATGCTCCCGTGTGTGGCAGTGAT
      C G C N P I V A C P E I Y A P V C G S D
421    GGCAAGACTTATGATAACGACTGCTATTTCCAGGCAGCTGTTTGCAAGAATCCAGATCTT
      G K T Y D N D C Y F Q A A V C K N P D L
481    AAGAAGGTTTCGAGACGGTAACTGCGACTGCACTCCTCTCATCGGCTGTCCCAAGAACTAC
      K K V R D G N C D C T P L I G C P K N Y
541    AGGCCTGTGTGTGGCAGCGACGGTGTAACCTACAACAACGACTGCTTCTTCAAGGTTGCT
      R P V C G S D G V T Y N N D C F F K V A
601    CAGTGAAGAACCCCGCGCTCGTCAAAGTCTCTGATACTCGCTGTGAATGCAACCACGTC
      Q C K N P A L V K V S D T R C E C N H V
661    TGTACTGAAGAATATTACCCCGTGTGCGGAAGCAATGGTGTACGATTTTCAACATTTGT
      C T E E Y Y P V C G S N G V T Y S N I C
721    CTGTTGAATAATGCAGCGTGTTTAGATTCTCCATTTACAAGGTTTCGGACGGAATCTGT
      L L N N A A C L D S S I Y K V S D G I C
781    GGTGCGCACTGTACCTATAA
      G R R L Y L *

```

Figure 2. Nucleotide and amino acid sequences of the cDNA clone, sh415. Amino acids are shown as single letter abbreviation. The putative signal peptide is underlined.

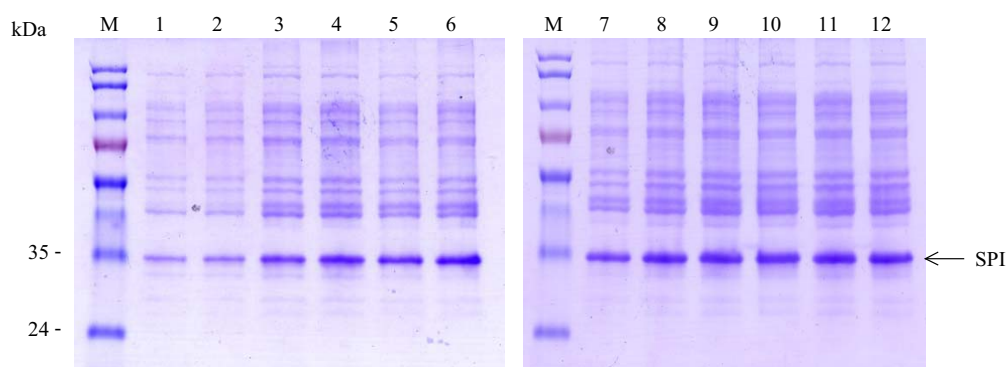


Figure 3. SDS-PAGE analysis of the recombinant serine proteinase inhibitor in *E.coli*. Lanes 1, 3, 5, 7, 9 and 11; the lysate of *E.coli* containing the parental pTrcHis2C after induction with 1 mM IPTG for 0, 1, 2, 3, 4 and 5 hours, respectively. Lanes 2, 4, 6, 8 and 10; the lysate of *E.coli* containing the shrimp proteinase inhibitor (SPI)-pTrcHis2C after induction for 0, 1, 2, 3, 4 and 5 hours, respectively. Lane M; prestained protein ladder

Assay of inhibitory activity

To ascertain that the 35-kDa band observed by SDS/PAGE is the cloned Kazal inhibitor, the inhibitory activity of the crude *E.coli* lysate was assayed by activity staining in a gelatin-SDS/PAGE. The 35-kDa protein was shown to have proteinase inhibitory activity against trypsin and chymotrypsin but not subtilisin (Fig. 4). Studies on proteinase inhibition of the crude proteins showed that the inhibitor blocked the activity of trypsin and chymotrypsin whereas no inhibition of subtilisin and elastase was detected (Table 1). As predicted on the basis of the P1 residue at the reactive site of each Kazal domain, the five-domain Kazal inhibitor of *P.monodon* contains Ala at the second domain and Lys at the forth domain suggesting the inhibitory activity against elastase and trypsin, respectively [10]. The inability of the protein to inhibit elastase suggested that the second domain may be inactive. The role of Thr at the P1 position of the first domain and Glu at the third and fifth domains has not yet been specified. The biological function of Kazal-type proteinase inhibitors in haemocytes of *P. monodon* is not known. In several organisms, Kazal-type proteinase inhibitors have been found and their activities against the proteinase substrates have been study extensively *in vitro* [10, 12]. However, little work has been published on the function of the inhibitors *in vivo*. In crayfish, it was found that the Kazal inhibitor blocks subtilisin, thus the biological function as the inhibitor of microbial proteinases was proposed [5] whereas another serine proteinase inhibitor from crayfish plasma inhibits prophenol oxidase activation [7]. Target proteinases of the Kazal inhibitors in *P.monodon* have yet to be identified to reveal the biological function of the inhibitors in the shrimp haemocytes.

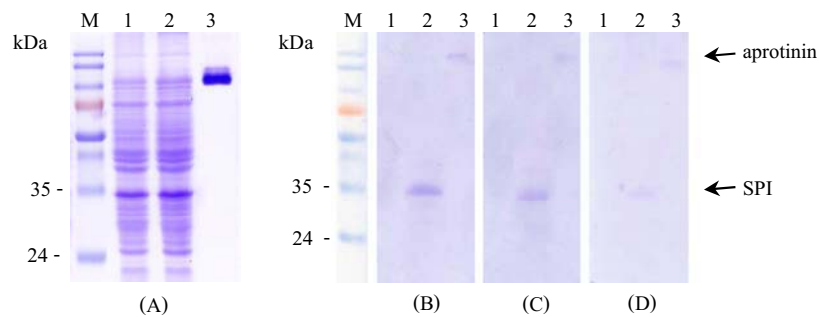


Fig. 4 PAGE analysis of the recombinant Kazal inhibitory activity. The *E.coli* lysate was electrophoresed on 12% SDS/PAGE showing the protein pattern (A). For activity gel, crude *E.coli* lysate was run on a gelatin-12%SDS/PAGE under reducing conditions. After electrophoresis, the gel was incubated in 2% TritonX-100 for protein refolding, then incubated with 40 µg/ml of trypsin (B), chymotrypsin (C), subtilisin (D). Lanes 1; the *E.coli* lysate containing the parent pTrcHis2C, Lanes 2; the *E.coli* lysate containing the recombinant plasmid pTrcHis2C, Lanes 3; 2 µg of aprotinin, Lane M; prestained protein ladder .

Table1. Inhibition of some serine proteinases by the inhibitor

The enzymes were assayed with chromogenic substrates as described above. The activity was monitored as the release of *p*-nitroaniline at 405 nm. The *E.coli* lysate (31.6 µg) was used in the experiment. The experiments were performed in duplicate.

Proteinase	A_{405} without the inhibitor	A_{405} with the inhibitor	Inhibition (%)
Trypsin (0.08 µg/assay)	0.82	0.11	89
Chymotrypsin (0.05 µg/assay)	0.40	0.12	70
Subtilisin (0.3 µg/assay)	1.28	1.28	0
Elastase (0.6 µg/assay)	0.62	0.72	0

Tissue Expression of the Kazal proteinase inhibitors

mRNA expression of the shrimp Kazal proteinase inhibitor was examined in several tissues including haemocyte, heart, intestine, gill, lymphoid organ and hepatopancrease. RT-PCR was performed using specific primers designed from the nucleotide sequence of the inhibitor cDNA . The results showed that the proteinase inhibitor is exclusively expressed in haemocytes whereas the control reactions using primers specific to 18S rRNA showed that the 18S rRNA gene was equally expressed in all tissues examined (Fig. 5).

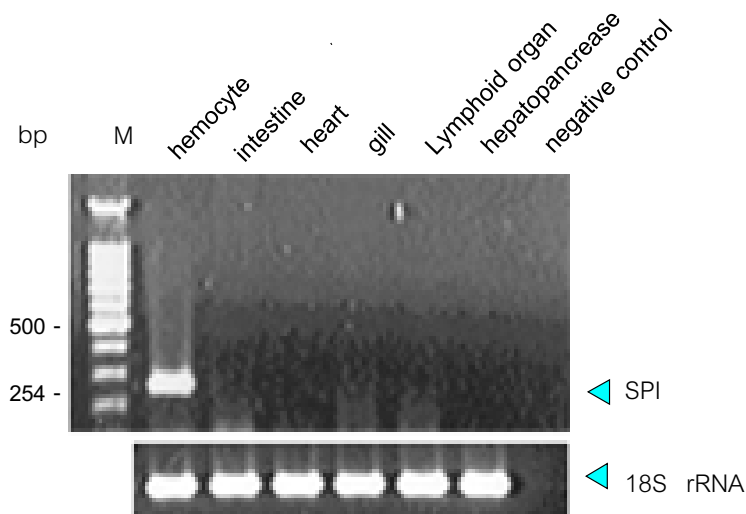


Figure 5. Analysis of the expression of the Kazal proteinase inhibitor in *P. monodon* tissues. RT-PCR was performed with cDNA templates from hemocytes, intestine, heart, gill, lymphoid organ and hepatopancrease. Amplification of 18S rRNA transcript was performed as control. M, 100 bp DNA ladder.

mRNA expression level of the inhibitor in bacterial infected shrimp

The influence of bacterial stimulation on transcription of the proteinase inhibitor in haemocyte of *P.monodon* was determined. Shrimps were experimentally challenged by injection with *Vibrio harveyi* and the haemolymph was collected at 3, 6, 12, 24 and 48 hr post injection. Detection of *Vibrio harveyi* in shrimps was performed by spreading suspensions of hepatopancrease onto thiosulfate citrate bile-salt sucrose agar (TCBS-agar). The presence of the luminescent bacteria was observed in hepatopancrease at 42-48 hr post injection. Therefore, the mRNA level was first compared between unchallenged and 48 hr bacterial challenged shrimps by Northern analysis (Fig. 6). Two transcripts with size of approximately 1.7 and 1.8 kb were detected in haemocytes. The major transcript is probably the transcript of the four-domain Kazal inhibitor whereas the minor transcript is that of the five-domain inhibitor. In any case, no difference in the mRNA level of both transcripts was observed after 48 hr of injection.

To determine mRNA level of the Kazal inhibitor at different time interval, a semi-quantitative RT-PCR was performed by addition of 2 sets of primers into the same reaction tube. One set amplified the cDNA of the inhibitor and another set amplified a constitutive β -actin gene that was used as an internal control to normalize for sample to sample variations in total RNA amounts. A time course of mRNA expression was performed on total RNA from haemocyte of unchallenged and bacterial challenged shrimps at 3, 6, 12, 24 hr post injection. Three sets of experimental shrimps (5 individuals each) were tested at different time point. No significant difference in the transcript level was found ($p>0.05$) (Fig. 7).

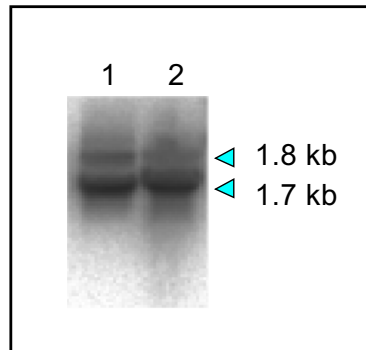


Figure 6. Northern analysis of mRNA expression of the proteinase inhibitor in haemocyte of unchallenged (lane 1) and *Vibrio harveyi* challenged (lane 2) *P. monodon*. Twenty μ g of total RNA from haemocytes were separated on a 1% agarose-formaldehyde gel and hybridized with the 32 P-labeled cDNA of the inhibitor.

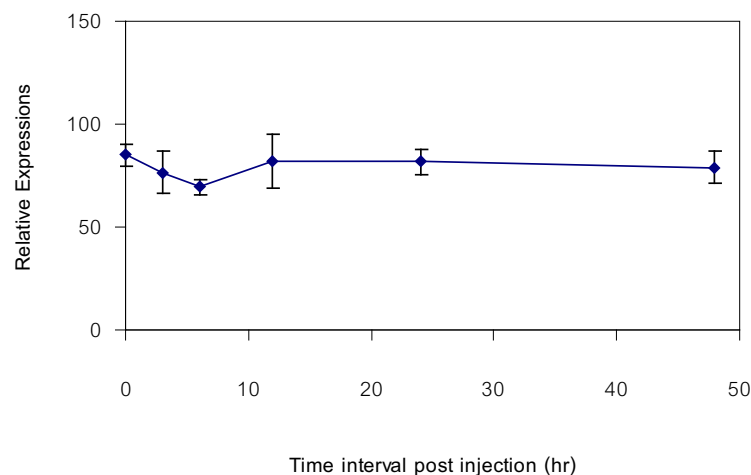


Figure 7. Relative expression levels of the Kazal inhibitor at different time intervals after injected with *V. harveyi*. Three sets of experimental shrimps were tested at each time point. Relative expression of the inhibitor was normalized with the β -actin expression (internal control).

Acknowledgement

This work is supported in part by The Thailand Research Fund's Senior Research Scholar to Prof. Sakol Panyim and the research grant (BT-B-06-2B-09-309) from National Center for Genetic Engineering and Biotechnology (BIOTEC) to Dr. Anchalee Tassanakajon.

References

1. Menegatti, E., Tedeschi, G., Ronchi, S., Bortolotti, F., Ascenzi, P., Thomas, R.M., Bolognesi, M. and Palmieri, S. (1992) *FEBS Letters* **301**(1), 10-14.
2. Ceciliani, F., Bortolotti, F., Menegatti, E., Ronchi, S., Ascenzi, P. and Palmieri, S. (1994) *FEBS Letters* **342**, 221-224.

3. Ascenzi, P., Ruoppolo, M., Amoresano, A., Pucci, P., Consonni, R., Zetta, L., Pascarella, S., Bortolotti, F. and Menegatte, E. (1999) *Eur J Biochem* **261**, 275-284.
4. Kanost, M.R. (1999) *Dev Comp Immunol* **23**, 291-301.
5. Johansson, M.W., Keyser, P. and Söderhäll, K. (1994) *Eur J Biochem* **223**, 389-394.
6. Volpicella M., Schipper A., Jongsma M.A., Spoto N., Gallerani R., Ceci L.R. (2000) *FEBS Letters* **468**, 137-141.
7. Hergenhahn H.-G., Aspan A., Söderhäll K. (1987) *Biochem. J.* **248**, 223-228.
8. Bradford M. (1976) *Anal. Biochem.* **72**, 248-254.
9. Supangul, P., Klinbungna, S., Pichyangkura, R., Jitrapakdee, Hirono, I., Aoki, T. and Tassanakajon, A. *Mar. Biotechnol.* **4**: 487-494.
10. Laskowski Jr, M. and Kato, I. (1980) *Annu Rev Biochem* **49**, 593-626.
11. Scott, M.J., Huckaby, C.S., Kato, I., Kohr, W. and Laskowski Jr., M., Tsai, M. J. and O'Malley, B.W. (1987) *J Biol Chem* **262**, 5899-5907.
12. Van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kroger, B., Hoffken, W., and Bode, W. (1995) *EMBO J.* **14**, 5149-5157.

**BLACK TIGER SHRIMP HEMOCYTE CDNAS : FROM
EXPRESSED SEQUENCE TAGS TO PROSPECTIVE
MARKERS OF MOLECULAR EVENT UPON
YELLOW HEAD VIRUS INFECTION**

Burachai Sonthayanon¹, Apinunt Udomkit, and Sakol Panyim

*Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Phutthamonthon District,
Nakhon Pathom 73170, Thailand*

¹ To whom correspondence should be addressed :- bsonthayanon@hotmail.com Fax. + 662 4419906

Abstract

Yellow Head Virus is a shrimp Nidovirus which frequently causes significant economic loss to shrimp aquaculture industry in Thailand and Asia Pacific region. To gain insight information on molecular event of the infection in black tiger shrimp (*Penaeus monodon*), we characterized random cDNA clones from the key defense cell type, hemocytes, and study some of their expression profile using membrane based macroarray hybridization. Over 600 cDNA clones were randomly picked and determined their partial sequences from the 5' end. Ninety six cDNA clones, most of which were of known putative identities based on our EST sequence were amplified for each insert fragment by PCR and spotted as a gridded array onto a nylon membrane. Seventy five clones (78%) gave detectable hybridization signal, when probed with three preparations of ³²P-labelled probes, prepared from total RNA of zero time non-infected shrimps, 24 hr post-injection buffer control shrimps, 24 hr post-YHV infection shrimps. Selected 38 clones from the first arrayed hybridization were then subjected to a confirmatory experiment using a smaller membrane and new RNA preparations. Our results revealed some altered transcript levels from the YHV-infected compared to buffer injected hemocytes at 24 hr. The changed transcription profile could be due to change in subpopulation of the hemocytes upon infection. One of the responsive transcripts were identified by a cDNA clone, PMC038, whose putative identity was selenoprotein W. Interestingly, there were two full-length cDNA clones from our cDNA library which we have sequenced, predicted to code for an identical polypeptide chain, and were differed mainly in their 3' UTR sequence and the length, one of which did not seem to respond to the YHV challenge (PMCP0207). The two cDNA might have come from two distinct genes. Limited known role of selenoprotein W, based on information from mammalian systems, was in a cellular control of oxidative damage. The function of shrimp gene for PMC0238 should further studied to elucidate its exact role in the molecular responses of the phagocytic hemocytes either due to defense response or viral pathogenesis of this economically important marine shrimp and other arthropods.

Key words: RNA level, pathogenesis, defense, phagocytes, hemolymph, decapoda, prawn, aquaculture

Introduction

Yellow Head Virus (YHV) is a lethal shrimp virus which was first observed in Thailand in 1990 (Chantanachookin *et al* 1993, Boonyaratpalin *et al* 1993, Nadala *et al* 1997, Flegel *et al* 1997) and then has been discovered elsewhere in Asia (Wang *et al* , 1996, Wang *et al* 2000, Walker *et al* 2001). YHV could kill most of the farm Penaeid shrimp species in a matter of few days. For *Penaeus monodon* (black tiger shrimp) the top economic species for aquaculture industry in Thailand, YHV infected shrimp would display a yellowish color of cephalothorax. (Lu *et al* 1995) However, the yellow color was not seen in *Litopenaeus* (*Penaeus*) *vannamei* and *Penaeus stylirostris*. (Lu *et al* 1994) Numerous studies have been conducted on detection of YHV by PCR or monoclonal antibodies (Tang and Lightner 1999 , Sithigorngul *et al* 2000) as well as their genome sequences (Sittidilokratna *et al* 2002, Jitrapakdee *et al* 2003). Nowadays, however, there were some reports on late or little symptom development and emergence of strains of YHV.

The YHV virus is an enveloped, spike-studded, virus particle, of the size around 150-186 nm (in length) and 38-50 nm (in width). Nuclear core is around 20-30 nm in diameter. The genome is positive single-stranded RNA (Wongteerasupaya *et al* 1995). Four major proteins were found from the purified virion, with the sizes of 170 kDa, 135 kDa (glycosylated), 67 kDa, 22 kDa. Recently availability of partial nucleotide sequence of the viral genome suggested YHV is a Nidovirus (Order Nidovirales) encompassing the Coronaviridae and Arteriviridae (Cavanagh 1997, Enjuanes *et al* 2000) and is a distinct species from another shrimp virus found in Australia, the Gill Associated Virus (GAV) (Cowley *et al* , 1999, Cowley *et al* 2000a, 2000b, Walker *et al* , 2001). Study in GAV have shown that it can be vertically transmitted into subsequent generation. (Cowley *et al* 2002) although it has not been demonstrated if YHV would do as such. Several shrimp viruses could infect the shrimp concurrently and mixed infections are commonly found in hatcheries and farms (for example, Manivannan *et al* 2002)

Studies of YHV infected tissues in selected shrimp species indicated virus particles were found in connective tissues (including of muscles, midgut caecum, nerve tracts, hematopoietic organ, antennal gland, heart, hepatopancreas), gills, cuticular epithelium, lymphoid organ (Lu *et al* 1995, Kasornchandra and Boonyaratpalin 1995). The virus could cause systemic infection of the black tiger shrimp as well as some other *Penaeus* shrimp species, such as *P.stylirostris*, *P.vannamei* (Lu *et al* 1994), and *P.japonicus* . Infected cells were then die due to apoptosis (Khanobdee *et al* 2002) . A wide area epizootic could be economically catastrophic to shrimp farms, due to a high mortality rate close to 100 % although there are circulating news of non-symptomatic infection. Despite much progress in the study of YHV infection and detection which led to wide-spread use of PCR in survey of post-larvae shrimps in aquaculture industry, information at molecular level on a virus infection, including those of YHV, has still been missing (Wongteerasupaya *et al* 1997, Tang and Lightner 1999, Sithigorngul *et al* 2000). Most of the research on host gene response to virus infection has been primarily on other lethal shrimp viruses, White Spot Virus (Astrofsky *et al* 2002, Bangrak *et al* 2002, Rojtinakorn *et al* 2002, van de Braak *et al* 2002), Taura syndrome virus (Song *et al* 2003), but so far not YHV. A number of studies were also directed toward viral genes and their protein products (van Hulten *et al* 2001).

Hemocytes is a major cell type involved in cell mediated defense mechanism of the crustaceans. They can be classified into subtypes based on microscopic morphology and the presence of some surface antigens. They are natural choice for us to work on. So far, few invertebrate genes from hemocytes are known to involve in immune or defense function, such as beta-1,3-glucan binding protein, peroxinectin (Sritunyalucksana et al 2001, 2003), superoxide dismutase (Campa-Cordova et al 2002), serine proteases (Gorman & Paskewitz 2001), serine proteinase inhibitor or Serpin (De Gregorio et al 2002), penaeidins (Cuthbertson et al 2002, Munoz et al 2003), antifungal peptide from hemocyanin (Destoumieux-Garzon et al 2001). We have earlier prepared and characterized a number of EST from hemocyte and muscle tissues, a number of which have been deposited in GenBank database. A number of groups have also reported a limited characterization of EST from various organs (for example, Gross et al 2001). Compare to other eukaryotic species, much less is known about DNA sequences from shrimp, esp. black tiger shrimp, (Boonchuoy et al 1999, Burks et al 1991, Udomkit et al 2000, Thepparit et al 2002).

Arrays of cDNA have been used to compare gene expression profiles of cells under different conditions (Cirelli and Tononi 1999, Cox *et al* 1999, Mochii *et al* 1999, Eickhoff et al 1999, Tanaka et al 2000). Using microarray techniques yet required a large number of clones (typically over 10,000 clones), plus sophisticated equipment and expensive reagents. A number of studies in mammalian system have enjoyed the availability of several thousands of available cDNA clones and commercially available microarrays. Array have been used to study gene expression in cancerous cells, in developmental stages of embryos, in viral infection, etc. In shrimp system, however, we have a limited number of available cDNA. Due to a low degree of DNA sequence similarity among genes of crustaceans, even among arthropods, which can be extended to mammals, arrays from heterologous organisms could not be used. This work took advantage on our limited number of existing cDNA clones. Using macroarrays we have identified a number of prospective cDNA of certain putative identities that showed altered expression level upon infection with YHV. This result provided the leads in which further studies could be conducted to characterize the molecular event in the shrimp cells upon the infection and pathogenesis of this economically devastating epizootic

Materials and Methods

DNA clones were those isolated from black tiger prawn hemocyte cDNA libraries in our laboratory and partially sequenced earlier from the 5' ends by J. Wongsantichon and C. Boonchuoy (unpublished). Some of the sequences determined as EST have been deposited into the dbEST division of the GenBank database. Few other clones, those in the NF series, were cloned by RT-PCR by C. Thepparit in our laboratory. Ninety six cDNA clones with known tentative or putative identities based on either BLASTN or BLASTX searches to existing online database at the time were chosen for this study. All the clones were amplified to get the insert part by mainly using M13 universal forward and reverse primers, except for those in NF series where M13 reverse primer and T7 primer were used. For the first experiment, a grid of 10x12 array was designed as a template and a small GeneScreenPlus nylon membrane, approximate size 5.5 cm x 6.5 cm, was used as substrate for manual DNA spotting. Spotting quantity of most DNA samples were 100 ng and the volume

for each sample was 1-2 μ l. Actin cDNA clone from hemocytes which also shared a high sequence identity with other shrimp actin types of clones were spotted into 3 different corners as a positive control at 10 ng per spot. Most of other clones were spotted at 100 ng per spot. Plasmid DNA, pBluescript SK (-) was used as negative control and 10 ng was spotted once onto the membrane. For the second experiment, a 9x5 arrayed grid was used to spot a 38 clone subset for hybridization in a similar manner.

Shrimp specimens

Farm grown shrimps were used in this study. Live healthy shrimps of about 4 months old, each weight about 30 g, were purchased from local farms in eastern Thailand. They were visually examined for good healthy signs and fed once with standard shrimp feeds prior to the experiment. They were also tested later for the absence of the two known lethal viruses of the *Penaeus* shrimps, Yellow Head Virus (YHV) and White Spot Virus (WSV) by PCR method. Specific primers for those were kindly provided by Prof. Vichai Boonsaeng's laboratory at the Department of Biochemistry, Faculty of Science, Mahidol University. Two different lots of shrimp were used for this work. For each lot, the shrimps were divided equally into 3 groups and put into 3 aquariums in the laboratory. The first group of the shrimps was collected for their hemolymph by cardiac puncture in the same day of the sample delivery. The second group was manually and individually injected with sterile NTE buffer, 0.1 ml each, into the tail section of each shrimp. The third group was also injected with 0.1 ml of (1/50) diluted YHV preparation, also kindly provided by Prof. V. Boonsaeng and kept as stock in the freezer at -20 C. The two groups of injected shrimps were kept 24 hr in each aquarium without any feeding. Samples from all the shrimp in each group were pooled. Hemolymph from the two groups were collected at 24 hr after respective injection and the hemocytes of both pools were separately pelleted and total RNA were extracted. It is noteworthy that only the group of the shrimp which was infected with YHV showed clear clinical symptoms of the yellow head disease and were tested positive to the YHV by PCR (data not shown) while the buffer-injected control appeared normal and tested negative for YHV. None of the shrimp PCR-tested were positive for WSV.

RNA preparation.

Hemocytes were obtained from hemolymph by pelleting on a bench top centrifuge for 10 min at 800 rpm. Procedure for RNA isolation was according to Chomzynski and Sacchi, using TRI Reagent (Molecular Research Center, Inc.). Equal amount of total RNA, around 5 μ g from each preparation, was separately labeled by reverse transcription with SuperscriptTM II reverse transcriptase (GibcoBRL, USA) (200 U/ reaction), using oligo(dT)₁₂₋₁₈ as the primer in 50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM MgCl₂, in the presence of [α -³²P]-dCTP 3000 Ci/mmol (Amersham Pharmacia Biotech, UK).

Hybridization was conducted in small plastic test tubes for 24 hr at 65 C. Washing was according to established protocol.

Autoradiography was conducted with Kodak XAR X-ray film for 15 min to few hours. Films were developed by automatic film processor. Autoradiograms

obtained from each type of reversed transcribed probe were selected for image analysis.

Imaging and quantification of radioactivity present in each spot was by the use of Bioimage Analysis software (BioRad). Intensity of spots from each film was normalized for determining the expression level by using one spot of actin DNA at the same corresponding position on the array as reference.

Results

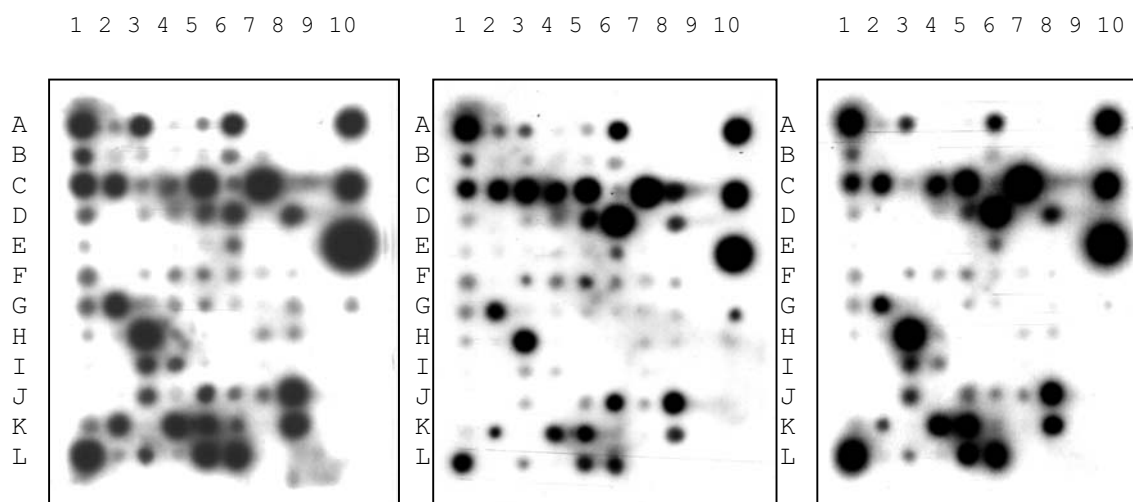


Figure 1 First array hybridization

A single nylon membrane was hybridized with probes from 3 new preparations of total RNA. The total RNAs were extracted from shrimp haemocytes of control, 24 hr. post NTE injection and 24 hr. post YHV injection.

Figure 1 showed a result of the first array hybridization (using a single membrane) to 32 -P labeled probes prepared from three preparations of hemocytes-derived total RNA, one from uninfected zero time total RNA, one from 24 hr buffer-injected total RNA, and one from 24 hr YHV-infected total RNA. Most of the cDNAs (94 clones) were originally isolated from our black tiger prawn hemocyte cDNA library and only two of the cDNAs were cloned by RT-PCR in a separate work. It can be seen that the majority of the spotted hemocyte cDNA clones (75 clones, or about 78%) could hybridize to the 3 reversed transcribed probes, also prepared from black tiger prawn hemocytes' RNAs. Only 21 out of 96 clones (22%) did not hybridize, as verified by their densitometric values to about the same as negative control background. A likely explanation was that the nonhybridizing clones were of low abundant type and their mRNA level were lower than the sensitivity of this reverse hybridization technique. It is noteworthy that when compared between the zero time probe (left pane) and the 24 hr post buffer injection probe (central pane), the majority of spots (in the central pane) showed reduced intensities. This occurred despite our attempted optimization of autoradiographic exposure time to get similar level of spot intensity of the positive control spots (three hemocyte actin spots, two on the top left and right corners, and one on the low left corner, all at 1/10 amount compare to most of other spots.) The reduced spots' intensity in central pane of the

buffer-injected probe was likely due to either stress of the shrimp from being injected with the buffer or deprivation of feeds at 24 hr post-injection, or both.

Although some spots appeared to hybridize differently from probes from the hemocyte RNA of the two conditions, it was difficult to judge only by naked eyes. To help categorize our result, we chose densitometric analysis to quantify spot intensities from selected X-ray films. After scanning, the output of the software was then further normalized against a base line using one spot of the positive control (top right corner) as reference so that spot intensities of the 3 membranes could be compared.

We confirmed the result of the first array hybridization by conducting a second array hybridization. To improve on hybridization efficiency, we selected only a subset of 38 clones, to spot onto a smaller membrane. New preparation of total RNA from a new batch of farm shrimps was also used to prepared probes. The result, shown in figure 2, revealed that most of the clones, except three, did not significantly change their expression levels. The discrepancy of our results between the first and the second array hybridization seemed to suggest that the different lots of farm shrimps used in our work has significant influence in the array screening.

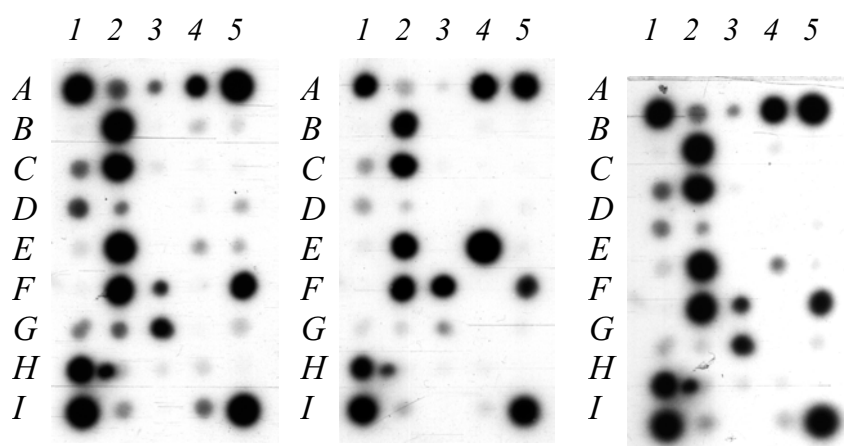


Figure 2 Second array hybridization.

A new nylon membrane with new arrayed cDNA clones was hybridized by radioactive cDNA labelled using RT-PCR from 3 new preparation of total RNA.

The only transcript of known putative identity which was found to change the expression level in both arrays were selenoprotein W as probed by clone PMCP0238 (spot G10 in fig.1, and spot E4 in fig.2). Densitometric scanning in fig. 2 confirmed the visualized change of spot density at 0.03 compare to zero time control (right pane versus the middle pane).

Interestingly, one of another cDNA clone for selenoprotein W (PMCP207) did not quite behave the same (spot F6 in fig.1, spot H4 in fig.2). In order to clarify their differences, their nucleotide sequences were verified and confirmed.

Black Tiger Shrimp Hemocyte cDNAs : From Expressed Sequence Tags to Prospective Markers of Molecular Event upon Yellow Head Virus Infection

```

>PMC0238                                     734 nt vs.
>PMCP207                                     606 nt
80.6% identity;                             Global alignment score: 2005

      10      20      30      40      50      60
AAGGAAAGACCATCGTGGTAGGACTTTCTTGCAGGAGCGATAAGGAGCCGGCAAGCTGTTT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- A-GGAAAGGCCATCGTGGTAGGACTTTCTTGCAGGAGCGATAAGGAGCCGGCAAGCTGTTT
      10      20      30      40      50
GCGGTCTTCAGGCGGGCTTTTTCATCAGCCATGACTAAGCTTCAGATTAAAGTGGTGTAT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- GCGGTCTTCAGGCGGGCTTTTTCATCAGCCATGACTAAGCTTCAGATTAAAGTGGTGTAT
60      70      80      90      100     110
TGCGGTGCCCTGAGGGTACGCCCCAAGTTCCGTAAACTTAAACGGAAGTGAAGCCGAA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- TGCGGTGCCCTGAGGGTACGCCCCAAGTTCCGTAAACTTAAACGGAAGTGAAGCTGAA
120     130     140     150     160     170
TTTCCAGATAAACTGGATATTCATGGCGAGGCTACTCCTTCAGCTACAGGATATTTTGAA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- TTTCCAGATAAACTGGATATTCATGGCGAGGCTACTCCTTCAGCTACAGGATATTTTGAA
180     190     200     210     220     230
GTAAGTGTGCGGGGATAAACTGGTCCATAGCAAAAAGATGGCGATGGCTATGTTGATTTCG
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- GTAAGTGTGCGGGGATAAACTGGTCCATAGCAAAAAGATGGCGATGGCTATGTTGATTCA
240     250     260     270     280     290
GAGAAGAAAAGTCAACAAGATAATTGATGCAGTCCATGATGCACTTGGTGAAGTAGAATTTA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- GAGAAGAAAAGTCAACAAGATAATTGATGCAGTCCATGATGCACTTGGTGAAGTAGAATTTA
300     310     320     330     340     350
CCTCTTATGGACGGCCTGACTTGTGTGCTTCGTAGGCCCTCTTTGTGGCAGGGCTACTG
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- CCTCTTATGGACGGCCTGACTTGTGTGCTTCGTAGGCCCTCTTTGTGGCAGGGCTACTG
360     370     380     390     400     410
AGTATGAAACATTGTGCTGAAAACCCCTCTGTTTGGATGGCACTTTGTGGAGGAGTAGAGA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- AGTATGAAACATTGTGCTGAAAACCCCTCTGTTTGGATGGCACTTTGTGGAGGAA-AAAAA
420     430     440     450     460     470
AGCCCTTTCAAGATTGCCATAAAGAGGCCCTGCTAGAGCACATCAAGGGCAGGTTGTCCTC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- AGCCCTTTCAAGATTGCCATAAAGAGGCCCTGCTAGAACACATCAAGGGCAGGTTGTCCTC
480     490     500     510     520     530
CTTTACCAGT-ACCTTATTCAGGGATCTTAACCTTTGGCTTTTGGCCCATC-TTGCATTAAT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- CTTTACCAGTTACCTTATTCAGGGATCTTA-CTTGTCTTT-GGCCCATCCTTGTCTTATT
540     550     560     570     580     590
AAATATATACTAAAGATGTGCCTGTCTTTCCCATAGCTTTGGCAAGTAATTAAGGCCTT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
- AAATATATACT-----
600
TTTCATGCCAAGGATTTTTTTTTTGGACATGCATGGCTTTTATATCCCTATTATTATTAGA
-----
720     730
ATAAAGGTTTCATTTC
-----

```

Figure 3 Full-length DNA sequence alignment of two different cDNA clones for Selenoprotein W from the shrimp hemocytes. The two highly similar sequences (PMC0238, top, and PMCP207, below) encode identical amino acid sequence despite the presence of two nucleotide substitution inside the open reading frame. Poly A tails were found at the end of both sequences at different positions (not shown). In PMC0238, translation start were found at position 91 and termination codon at 352 TGA at 130 was believed to encode for Selenocysteine. SECIS element which could form secondary structure in mRNA was found from nucleotide position 594-675. Two polyadenylation signals were found at nucleotide positions 596 and 718 of PMC0238 (boxed nucleotides).

As can be seen in fig. 3, there were some differences between the two sequences. the two cDNA shared 99% (602/608 bp) sequence identity, and the presence of 125 bp longer 3' UTR in clone PMC0238. The sequence discrepancies suggested that these were different transcripts from two different Selenoprotein W gene. Despite some differences in the ORF, however, the deduced amino acid sequences from both clones were identical. The amino acid sequence was highly similar to mammalian and *Drosophila's* Selenoprotein W (Bairoch *et al* 1997, Whanger 2000)

Discussion

Although the first array hybridization experiment seemed to show that a number of the transcripts have altered expression, many could not be confirmed by the second hybridization. Only one transcript was of known putative identity, selenoprotein W (as probed by cDNA clone PMC0238).

Studies on functional role of selenoprotein in higher eukaryotes have been limited even in mammalian systems. Selenoproteins are known to exist in human granulocytes (Liu *et al* 1999). Some have suggested on its regulatory function of oxidative state inside the cell, since the hemocytes would need to use significant oxidative power to destroy phagocytosed pathogens. Our data showed for the first time that the level of selenoprotein W dropped in YHV infected hemocytes with respect to buffer-injected control group thus perhaps indicating the loss of control on the intracellular oxidation-reduction status, and consequently, its capability to destroy invading pathogens. This control is apparently important. Human lymphocytes infected with HIV were found to have the level of selenoproteins decreased (Gladyshev *et al* 1999). Alsina *et al* (1999) reported that high levels of *Drosophila's* ptuf (Selenoprotein D) mRNA in dividing cells and low or undetectable levels in non-dividing cells. Most selenoproteins identified so far are antioxidants, the role of ptuf in cell proliferation through the control of the cellular redox balance has been proposed. Their biological function of the protein could be related to eliminating H₂O₂ generated in the respiratory burst reaction of granulocytes, thus protecting these cells from oxidative damage during phagocytosis. Downstream effects of Selenoproteins are still not thoroughly known. Morey *et al* 2001 reported that Ras/MAPK signalling pathway could be modified by selenoproteins.

Regarding the nature of two transcripts of selenoprotein W which we have found, study in *Drosophila* genome by Castellano *et al* have found 4 Selenoprotein genes. Thus it is possible that there could be few copies of selenoprotein W genes in the shrimp genome as well.

Currently the role of selenoprotein PMC0238 and other genes of unknown identity in response to viral infection is not clear, whether or not their change in expression level is due to the defense response of the cell or they signify the diseased status of the phagocytes. It is known that, upon infected with WSSV or YHV, the titers of the hemocytes in infected shrimps dropped significantly and likely the change in subpopulation of the hemocytes as a result of the virus infection. Our result was the first and is in line with those observations and should well serve as a useful basis for further study on the molecular responses of the hemocytes, whether they were either due to defense response or viral pathogenesis.

Acknowledgements

We thank Dr. Duncan Smith for kindly provided valuable suggestions on this work. This work is supported by Thailand Research Fund's Senior Research Scholar to Prof. Sakol Panyim. We thank Mr. Lerdchai Chintapitaksakul, Miss Jantana Wongsantichon, who did much of the work described in this paper. Mr. Lerdchai received a graduate scholarship as part of the same fellowship. We thank Miss Chanikarn Boonchuoy of the Central Equipment Laboratory Institute of Science and Technology for Research and Development Mahidol University, for a part of her DNA sequencing work which also contributed to this work ,and Miss Chutima Theparit, for providing two of her cDNA clones derived from RT-PCR cloning. We also are grateful to Prof. Dr. Vichai Boonsaeng of the Department of Biochemistry, Faculty of Science, Mahidol University, for kindly providing specific oligonucleotide primers for PCR detection of YHV and White spot viruses, and to Prof. Dr. Boonsirm Witthayachumnarnkul for numerous useful advises. We also appreciate generous helps from several members of the B315 lab at the Department of Biochemistry.

References

1. Alsina B, Corominas M, Berry MJ, Baguna J, Serras F. 1999 Disruption of selenoprotein biosynthesis affects cell proliferation in the imaginal discs and brain of *Drosophila melanogaster*. J Cell Sci **112** (Pt 17):2875-84
2. Astrofsky KM, Roux MM, Klimpel KR, Fox JG, Dhar AK. 2002 Isolation of differentially expressed genes from white spot virus (WSV) infected Pacific blue shrimp (*Penaeus stylirostris*). Arch Virol **147**(9):1799-812
3. Bairoch,A., Bucher,P., and Hofmann,K. 1997 The PROSITE database, its status in 1997. Nucleic Acids Res.**25**, 217-221
4. Bangrak P, Graidist P, Chotigeat W, Supamattaya K, Phongdara A. 2002 A syntenin-like protein with postsynaptic density protein (PDZ) domains produced by black tiger shrimp *Penaeus monodon* in response to white spot syndrome virus infection. Dis Aquat Organ **49**(1):19-25
5. Boonchuoy, C., Boonyawan, B., Panyim, S., and B. Sonthayanon. 1999 A cDNA sequence of phosphopyruvate hydratase (enolase) from black tiger prawn, *Penaeus monodon*. Asia Pacific J Mol Biology and Biotechnology. **7**,89-94
6. Boonyaratpalin, S., Supamattaya, K., Kasornchandra, J., Direkbusaracom, S., Aekpanithanpong, U., and Chantanachooklin, C. 1993 Non-occluded baculo-like virus, the causative agent of yellow head disease in the black tiger shrimp (*Penaeus monodon*) Gyobyo Kenkyu, **28**(3) 103-109
7. Burks, C., Cassidy, M., Cinkosky,M.J., Cumella, K.E., Gilna, P., Hayden, J.E.D., Keen, G.M., Kelley,T.A., Kelly, M., Kristofferson,D., Ryals,J. 1991 GenBank. Nucleic Acids Res. **19**, 2221-2225
8. Campa-Cordova AI, Hernandez-Saavedra NY, Ascencio F. 2002 Superoxide dismutase as modulator of immune function in American white shrimp (*Litopenaeus vannamei*).Comp Biochem Physiol C Toxicol Pharmacol (4):557-65
9. Campa-Cordova AI, Hernandez-Saavedra NY, De Philippis R, Ascencio F. 2002 Generation of superoxide anion and SOD activity in haemocytes and muscle of American white shrimp (*Litopenaeus vannamei*) as a response to beta-glucan and sulphated polysaccharide. Fish Shellfish Immunol **12**(4):353-66
10. Castellano S, Morozova N, Morey M, Berry MJ, Serras F, Corominas M, Guigo R. 2001 In silico identification of novel selenoproteins in the *Drosophila melanogaster* genome. EMBO Rep **2**(8):697-702

11. Cavanagh, D. 1997 Nidovirales : a new order comprising Coronaviridae and Arteriviridae. Arch. Virol **142/3** 629-633
12. Chantanachookin, C., Boonyaratpalin, S., Kasornchandra, J., Direkbusaracom, S., Ekpanithanpong, U., Supmataya, K., Sriurairatana, S., Flegel, T.W., 1993 Histology and ultrastructure reveal a granulosis like virus in *Penaeus monodon* affected by yellow-head disease. Diseases of Aquatic Organisms. **17**,145-157
13. Chintapitaksakul, L. Wongsantichon, J., Smith, D.R., Udomkit, A., Panyim, S. and B. Sonthayanon. 2002 Altered expression level of black tiger shrimp haemocyte transcripts infected with yellow head virus (YHV) as determined by arrayed hybridization, Fourth HUGO Pacific Concerence and the Fifth International Conference on Human Genetics. Pattaya, Thailand, October 27-30, 2002
14. Chintapitaksakul, L. Wongsantichon, J., Udomkit, A., Panyim, S. and B. Sonthayanon. 2002 Transcriptional profile of yellow head virus (YHV) infected haemocytes of black tiger prawn (*Penaeus monodon*). Proc. STT 28 Conference, Bangkok, Thailand, October 24-26, 2002
15. Cirelli C, Tononi G. 1999 Differences in brain gene expression between sleep and waking as revealed by mRNA differential display and cDNA microarray technology. J Sleep Res Jun;8 Suppl **1**:44-52
16. Cowley JA, Dimmock CM, Spann KM, Walker PJ. 2001 Gill-associated virus of *Penaeus monodon* prawns. Molecular evidence for the first invertebrate nidovirus. Adv Exp Med Biol **494**:43-8
17. Cowley JA, Hall MR, Cadogan LC, Spann KM, Walker PJ. 2002 Vertical transmission of gill-associated virus (GAV) in the black tiger prawn *Penaeus monodon*. Dis Aquat Organ **50**(2):95-104
18. Cowley, J.A. Dimmock, C.M., Spann,K.M., Walker, P.J. 2000 Gill-associated virus of *Penaeus monodon* prawns : an invertebrate virus with ORF1a and ORF1b genes related to arteri- and coronaviruses. Journal of General Virology. **81**, 1473-1484
19. Cowley, J.A. Dommock, C.M., Wongteerasupaya, C., Boonsaeng, V., Panyim,S., Walker, P.J. 1999 Yellow head virus from Thailand and gill-associated virus from Australlia are closely related but distinct prawn viruses. Diseases of Aquatic Organisms **36**, 153-157
20. Cowley, J.A., Dimmock, C.M., Spann, K.M., Walker P.J. 2000 Gill-associated virus of *Penaeus monodon* prawns : molecular evidence of the first invertebrate nidovirus. In Lavi, E.(ed.) The nidoviruses. Kluwer Academic / Plenum Publications, New York.
21. Cowley, J.A., Dimmock, C.M., Walker, P.J. 2002 Gill-associated nidovirus of *Penaeus monodon* prawns transcribes 3' coterminal subgenomic mRNAs that do not possess 5'-leader sequences. J. Gen Virol. **83**, 927-935
22. Cox KH, Pinchak AB, Cooper TG. 1999 Genome-wide transcriptional analysis in *S. cerevisiae* by mini-array membrane hybridization. Yeast Jun **15**;15(8):703-13
23. Cuthbertson BJ, Shepard EF, Chapman RW, Gross PS. 2002 Diversity of the penaeidin antimicrobial peptides in two shrimp species. Immunogenetics **54**(6):442-5
24. De Gregorio E, Han SJ, Lee WJ, Baek MJ, Osaki T, Kawabata S, Lee BL, Iwanaga S, Lemaitre B, Brey PT. 2002 An immune-responsive Serpin regulates the melanization cascade in *Drosophila*. Dev Cell **3**(4):581-92
25. Destoumieux-Garzon D, Saulnier D, Garnier J, Jouffrey C, Bulet P, Bachere E. 2001 Crustacean immunity. Antifungal peptides are generated from the C terminus of shrimp hemocyanin in response to microbial challenge. J Biol Chem. **276**(50):47070-7.
26. Eickhoff B, Korn B, Schick M, Poustka A, van der Bosch J. 1999 Normalization of array hybridization experiments in differential gene expression analysis. Nucleic Acids Res Nov **15**;27(22):e33
27. Enjuanes, L., Spann, W., Snijder, E., Cavanagh D. 2000 Nidovirales. In Regenmortel M.H.V., Fauquet C.M., Bishop, D.H.L., Carstens, E.B., and 7 others (eds.) Virus taxonomy. Academic Press, New York, pp.27-34

28. Flegel, T.W., Boonyaratpalin, S, Withyachumnarnkul, B. 1997 Progress in research on yellow-head virus and white-spot virus in Thailand. In Flegel TW., MacRae, IH (eds.) Diseases in Aquaculture III, Asian Fisheries Society, Manila, pp. 285-302
29. Gladyshev VN, Stadtman TC, Hatfield DL, Jeang KT. 1999 Levels of major selenoproteins in T cells decrease during HIV infection and low molecular mass selenium compounds increase. Proc Natl Acad Sci U S A 1999 Feb 2; **96**(3):835-9
30. Gorman MJ, Paskewitz SM. 2001 Serine proteases as mediators of mosquito immune responses. Insect Biochem Mol Biol. **31**(3):257-62.
31. Gross, P.S., Bartlett, T.C., Browdy, C.L., Chapman, R.W., Warr, G.W. 2001 Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific White Shrimp, *Litopenaeus vannamei*, and the Atlantic White Shrimp, *L. setiferus*. Dev. Compare. Imm. **25**,565-577
32. Jitrapakdee S, Unajak S, Sittidilokratna N, Hodgson RA, Cowley JA, Walker PJ, Panyim S, Boonsaeng V. 2003 Identification and analysis of gp116 and gp64 structural glycoproteins of yellow head nidovirus of *Penaeus monodon* shrimp. J Gen Virol **84**(Pt 4): 863-73
33. Kasornchandra, J., Boonyaratpalin S. 1995 Electron microscopic observation on the replication of yellow head baculovirus in the lymphoid organ of *Penaeus monodon*. In "Diseases in Asian Aquaculture II" (Shariff, M. Arthur, J.R., Subasinghe, R.P. editors) Asian Fishery Society. Manila, pp. 99-105
34. Khanobdee K, Soowannayan C, Flegel TW, Ubol S, Withyachumnarnkul B. 2002 Evidence for apoptosis correlated with mortality in the giant black tiger shrimp *Penaeus monodon* infected with yellow head virus. Dis Aquat Organ **48**(2):79-90
35. Khodarev NN, Advani SJ, Gupta N, Roizman B, Weichselbaum RR. 1999 Accumulation of specific RNAs encoding transcriptional factors and stress response proteins against a background of severe depletion of cellular RNAs in cells infected with herpes simplex virus 1. Proc Natl Acad Sci U S A Oct **12**; **96**(21):12062-7
36. Lee MH, Shiau SY. 2002 Dietary vitamin C and its derivatives affect immune responses in grass shrimp, *Penaeus monodon*. Fish Shellfish Immunol **12**(2):119-29
37. Lee YK, Soh BS, Wu JH. 2001 Quantitative assessment of phagocytic activity of hemocytes in the prawn, *Penaeus merguensis*, by flow cytometric analysis. Cytometry. **43**(1):82-5.
38. Lei XG. 2001 Glutathione peroxidase-1 gene knockout on body antioxidant defense in mice. Biofactors; **14**(1-4):93-9
39. Liu Q, Lauridsen E, Clausen J. 1999 The major selenium-containing protein in human peripheral granulocytes. Biol Trace Elem Res **68**(3):193-207
40. Loh, P.C., Tapay, L.M., Lu, Y., and Nadala, E.C.B., Jr. 1997 Viral pathogens of the penaeid shrimp. Advances in Virus Research, **48**, 263- 312
41. Lu, Y, Loh P.C. 1994 Infectivity studies of rhabdovirus in penaeid blue shrimp. Aquaculture International. **2**,123-127
42. Lu, Y. Tapay, L.M. , Loh, P.C., Brock, J.A. 1994 Infection of the yellow head baculo-like virus in two species of Penaeid shrimp, *P. stylirostris* and *P. vannamei*. Journal of Fishery Diseases **17**, 649-656
43. Lu, Y. Tapay, L.M., Loh, P.C., Brock, J.A., Gose, R.B. 1995 Distribution of yellow-head virus in selected tissues and organs of penaeid shrimp *Penaeus vannamei*. Diseases of Aquatic Organisms **23**, 67-70
44. Manivannan S, Otta SK, Karunasagar I, Karunasagar I. 2002 Multiple viral infection in *Penaeus monodon* shrimp postlarvae in an Indian hatchery. Dis Aquat Organ **5**; **48**(3):233-6
45. Morey M, Serras F, Baguna J, Hafen E, Corominas M. 2001 Modulation of the Ras/MAPK signalling pathway by the redox function of selenoproteins in *Drosophila melanogaster*. Dev Biol **238**(1):145-56

46. Munoz, M., Vandenbulcke, F., Saulnier D., Bachere, E. 2002 Expression and distribution of penaeidin antimicrobial peptides are regulated by haemocyte reactions in microbial challenged shrimp. *Eur. J Biochem.* **269**, 2678-2689
47. Nadala, E.C.B. Jr., Tapay, L.M., and Loh, P.C., 1997 Yellow-head virus : a rhabdovirus-like pathogen of penaeid shrimp. *Diseases of Aquatic Organisms* **31**,141-146
48. Rojtinakorn J, Hirono I, Itami T, Takahashi Y, Aoki T. 2002 Gene expression in haemocytes of kuruma prawn, *Penaeus japonicus*, in response to infection with WSSV by EST approach. *Fish Shellfish Immunol* **13**(1):69-83
49. Roux MM, Pain A, Klimpel KR, Dhar AK. 2002 The lipopolysaccharide and beta-1,3-glucan binding protein gene is upregulated in white spot virus-infected shrimp (*Penaeus stylirostris*). *J Virol* Jul;76(14):7140-9 Erratum in: *J Virol* 2002 Sep;76(17):8978
50. Serras F, Morey M, Alsina B, Baguna J, Corominas M. 2001 The *Drosophila* selenophosphate synthetase (selD) gene is required for development and cell proliferation. *Biofactors* **14**(1-4):143-9
51. Sithigorngul, P., Chauchuwong, P., Sithigorngul, W, Longyant, S., Chaivisuthangkura, P, Menasveata, P. 2000 Development of a monoclonal antibody specific to yellow head virus (YHV) from *Penaeus monodon*. *Diseases of Aquatic Organisms* **42**, 27-34.
52. Sittidilokratna, N., Hodgson, R.A.J., Cowley, J.A., Jitrapakdee, S., Boonsaeng, V., Panyim, S., Walker, P.J. 2002. Complete ORF1b-gene sequence indicates yellow head virus is an invertebrate nidovirus. *Diseases of Aquatic Organisms.* **50**(2):87-93
53. Song YL, Yu CI, Lien TW, Huang CC, Lin MN. 2003 Haemolymph parameters of Pacific white shrimp (*Litopenaeus vannamei*) infected with Taura syndrome virus. *Fish Shellfish Immunol* **14**(4):317-31
54. Spann, K.M., Cowley, J.A., Walker, P.J., Lester, R.J.G., 1997 A yellow head-like virus from *Penaeus monodon* cultured in Australia. *Diseases of Aquatic Organisms* **31**, 169-179
55. Sritunyalucksana K, Lee SY, Soderhall K. A beta-1,3-glucan binding protein from the black tiger shrimp, *Penaeus monodon*. 2002 *Dev Comp Immunol* **26**(3):237-45
56. Sritunyalucksana K, Wongsuebsantati K, Johansson MW, Soderhall K. 2001 Peroxinectin, a cell adhesive protein associated with the proPO system from the black tiger shrimp, *Penaeus monodon*. *Dev Comp Immunol* **25**(5-6):353-63
57. Tanaka TS, Jaradat SA, Lim MK, Kargul GJ, Wang X, Grahovac MJ, Pantano S, Sano Y, Piao Y, Nagaraja R, Doi H, Wood WH 3rd, Becker KG, Ko MS. 2000 Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc Natl Acad Sci U S A* Aug **1**;97(16):9127-32
58. Tang, K.F.J. Lightner, D.V. 1999 A yellow head virus gene probe : nucleotide sequence and application for in situ hybridization. *Diseases of Aquatic Organisms* **35**, 165-173
59. Thailand's National Statistic Office, Statistical Yearbook Thailand 1995, Vol. **42**, Office of the Prime Minister, Bangkok, Thailand, 1995
60. Thepparit, C., Sonthayanon, B., Udomkit, A., Panyim, S., and Smith, D.R. 2002 Molecular characterization of cDNAs obtained from black tiger shrimp (*Penaeus monodon*) hemocytes through the use of degenerate primers. *APJMBB* **10**(1), 31-39
61. Udomkit, A., Chooluck, S., Sonthayanon, B., and Panyim, S., 2000 Molecular cloning of a cDNA encoding a member of CHH/MIH/GIH family from *Penaeus monodon* and analysis of its gene structure. *J Exp. Marine Biol. and Ecology.* **244**, 145-156
62. van de Braak CB, Botterblom MH, Huisman EA, Rombout JH, van der Knaap WP. 2002 Preliminary study on haemocyte response to white spot syndrome virus infection in black tiger shrimp *Penaeus monodon*. *Dis Aquat Organ* **51**(2):149-55
63. van Hulten MC, Witteveldt J, Snippe M, Vlak JM. 2001 White spot syndrome virus envelope protein VP28 is involved in the systemic infection of shrimp. *Virology* **285** (2):228-33
64. Walker P.J., Cowley, J.A., Spann, K.M., Hodgson, R.A.J. Hall M.A., Withyachumnarnkul. B. 2001 Yellow head complex viruses : transmission cycles and topographical distribution

- in the Asia-Pacific region. In Browdy C.L., Jory, D.E. (eds.) The new wave : Proceedings of the Special Session of Sustainable Shrimp Culture, Aquaculture 2001. The world Aquaculture Society. Baton Rouge, LA. p 227-237.
65. Wang YT, Liu W, Seah JN, Lam CS, Xiang JH, Korzh V, Kwang J. 2002 White spot syndrome virus (WSSV) infects specific hemocytes of the shrimp *Penaeus merguensis*. Dis Aquat Organ **52**(3):249-59
66. Wang, C.S. Tang, K.F.J., Kou, G.H., Chen, S. N. 1996 Yellow head disease like virus infection in the kuruma shrimp *Penaeus japonicus* cultured in Taiwan. Fish Pathology. **31**, 177-182
67. Wang, Y., -C., and Chang, P.-S. 2000 Yellow head virus infection in the giant tiger prawn *Penaeus monodon* cultured in Taiwan. Fish Pathol. **35**(1), 1-10
68. Whanger PD. 2000 Selenoprotein W: a review. Cell Mol Life Sci **57**(13-14):1846-52
69. Whitney LW, Becker KG, Tresser NJ, Caballero-Ramos CI, Munson PJ, Prabhu VV, Trent JM, McFarland HF, Biddison WE. 1999 Analysis of gene expression in multiple sclerosis lesions using cDNA microarrays. Ann Neurol Sep;**46**(3):425-8
70. Wisconsin Package Version 9.1, Genetics Computer Group (GCG), Madison, Wisconsin, 1997
71. Wongteerasupaya, C., Sriurairatana, S., Vickers, J.E., Anutara, A., Boonsaeng, V, Panyim, S. Tassanakajon, A., Withayachumnarnkul, B. Flegel, T.W. 1995 Yellow-head virus of *Penaeus monodon* is an RNA virus. Diseases of Aquatic Organisms **22**, 45-50
72. Wongteerasupaya, C., Tongchuea, W., Boonsaeng, V, Panyim, S. Tassanakajon, A., Withayachumnarnkul, B. Flegel, T.W. 1997 Detection of Yellow-head virus (YHV) of *Penaeus monodon* by RT-PCR amplification. Diseases of Aquatic Organisms **31**, 181-186

STRUCTURE AND ORGANIZATION OF CRUSTACEAN HYPERGLYCEMIC HORMONE-LIKE GENES OF *PENAEUS MONODON*

Apinunt Udomkit, Amporn Wiwegweaw, Burachai Sonthayanon and Sakol Panyim
Institute of Molecular Biology and Genetics, Mahidol University, Nakhon Pathom 73170

Abstract

Crustacean hyperglycemic hormone (CHH), a major eyestalk peptide in decapod crustacean, is involved in several physiological processes including growth and reproduction. A gene encoding hyperglycemic hormone of *Penaeus monodon* (Pem-CHH1) has been characterized previously. In this study, two additional genes for *P. monodon*'s hyperglycemic hormones, *Pem-CHH2* and *Pem-CHH3* were cloned and characterized. The presence of three exons in both *Pem-CHH2* and *Pem-CHH3* genes makes them structurally different from the gene encoding *Pem-CHH1* that contains only two exons. The genomic sequences flanking the three *Pem-CHH* genes were determined by means of PCR-based genome walking. Nucleotide sequence analysis showed that both the upstream and the downstream sequences are highly conserved between *Pem-CHH2* and *Pem-CHH3* whereas the same regions of *Pem-CHH1* are less conserved. The TATA box-like sequence was found 28-30 nucleotides upstream of the transcription start site of each gene. Putative binding sites for several transcription factors e.g. CREB and Pit-1 were identified by computer programs. Finally, preliminary data on the organization of *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3* in *P. monodon*'s genome was obtained by Southern blot hybridization. At least one copy of each *Pem-CHH* gene is located in the same 10 kb *Hind* III-digested genomic fragment suggesting a gene cluster.

Keywords: black tiger shrimp, CHH, gene organization, promoter

Introduction

The X-organ sinus gland complex located in the eyestalk of decapod crustacean produces a variety of peptide hormones that plays important role in several physiological processes. Among all the eyestalk peptides, the hormones in the CHH/MIH/GIH family have been extensively studied. This peptide family, whose functions are essential for growth and reproduction, consists of crustacean hyperglycemic hormone (CHH), molt-inhibiting hormone (MIH), gonad inhibiting hormone (GIH) (Chang, 1997) and mandibular organ-inhibiting hormone (MO-IH) (Wainwright et al., 1996). A distinct characteristic of this hormone family is the presence of six cysteine residues that are aligned at identical positions in each member (Keller, 1992).

CHH is the most abundant eyestalk hormone. The major role of this hormone is to regulate glucose level in the haemolymph (Cooke and Sullivan, 1982). In

addition to its main function, CHH is also involved in the regulation of other processes such as lipid metabolism, ecdysteroid synthesis, ion transport and reproduction (Santos et al., 1997, Khayat, et al., 1998, de Kleijn et al., 1998, Spanings-Pierrot et al., 2000). CHH has been reported to exist as multiple isoforms in several species such as *Homarus americanus*, *Orconectes limosus*, *Metapenaeus ensis* and *Penaeus japonicus* (Tensen et al., 1991, De Kleijn et al., 1994, De Kleijn et al., 1995 Yang et al., 1997, Gu and Chan, 2000). Recently, the genes encoding different isoforms of CHH have been cloned and characterized in *M. ensis* (Gu and Chan, 1998, Gu et al., 2000). The MeCHH-A and MeCHH-B of *M. ensis* are encoded from at least eighteen genes, sixteen coded for MeCHH-A and the other two encodes MeCHH-B. These genes are arranged into two different clusters, each cluster contains at least six MeCHH-A genes and one copy of MeCHH-B gene (Gu et al., 2000). Moreover, different isoforms of CHH can be the result of alternative splicing. A gene encoding the precursor of such alternatively spliced CHHs has been identified in the shore crab *Carcinus maenas* (Dirksen et al., 2001). Similar type of CHH isoforms also occurs in *Macrobrachium rosenbergii* as the nucleotide sequence of the gene encoded for two alternatively spliced products, eyestalk form and gill form of CHH, was recently submitted to the GenBank by Chen et al. (AF372657).

In the black tiger shrimp, *P. monodon*, a cDNA encoding crustacean hyperglycemic hormone, so called Pem-CHH1, and its gene have been cloned and characterized (Udomkit et al., 2000). Recently, cDNAs encoding another two CHHs (Pem-CHH2 and Pem-CHH3) have been identified (unpublished data). The finding of gene cluster and alternatively spliced products of CHH in other species raises the possibility that the same might be true for *Pem-CHHs*. In this work, the organization of *Pem-CHH1-3* genes in the genome of *P. monodon* will be studied by means of PCR-based genome walking and Southern blot analysis. The putative promoter and regulatory elements of the three genes will also be analyzed.

Materials and Methods

Amplification of Pem-CHH2 and Pem-CHH3 genes

Genomic DNA was prepared from abdominal muscle tissues of *P. monodon* by using the QIAGEN Genomic-Tip and the Genomic DNA buffer set (QIAGEN). The PCR reaction contained 150 ng of genomic DNA template, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 2 mM MgCl₂, 200 nM each primer and 200 μM each dATP, dCTP, dGTP, dTTP in a total volume of 50 μl. The reaction was heated to 95°C for 5 minutes then, 2.5 units of *rTth* DNA polymerase (Biotools) were added. Amplification was achieved by 35 successive cycles of denaturation at 94°C for 1 minute, annealing at 50°C for 1 minute and extension at 72°C for 2 minutes, followed by a 10 minutes final extension at 72°C. The primers used to amplify the genes were designed from the 5' and 3' ends of the corresponding cDNA. The nucleotide sequences of these primers were CHH-F: 5' CGGAATTCTCAGTGCAGAGGGAGA GCC 3' and CHH2-R: 5' GCGGATCCCTGCTTTATGAAGACACTG 3' for *Pem-CHH2* gene and CHH-F and CHH3-R: 5' ATGCTTTATGAAGACATTAC 3' for *Pem-CHH3* gene.

Cloning and identification of the flanking nucleotide sequences of Pem-CHH genes

Universal Genome Walker kit (CLONTECH) was used to find the unknown 5' and 3' flanking region of *Pem-CHH* genes. *P. monodon*'s genomic DNA was separately digested to completion with four restriction enzymes, *Dra* I, *Hpa* I, *Sna* BI and *Xmn* I, that generate blunt ends after digestion. The digested genomic DNA fragments were ligated to the Genome walker Adapter. The pool of uncloned adaptor-ligated DNA fragments, referred to as Genome Walker libraries, was then amplified by polymerase chain reaction (PCR) with the outer adapter primer and gene specific primer for each *Pem-CHH* gene. The primary PCR products were diluted and used as a template for nested PCR with the nested adapter primer and nested gene specific primer. The sequences and relative position of gene specific primers for each walk from the three CHH-like genes were shown in Fig. 1. The PCR products were directly cloned into pGEMT[®] Easy (Promega). The nucleotide sequences of the cloned fragments were determined by automated DNA sequencing.

DNA sequence analysis

Double-stranded DNA sequencing was performed by the ABI PRISM[™] BigDye Terminator Cycle Sequencing Kit (Perkin Elmer). Transcription start site was predicted using the Berkeley Drosophila Genome Project WWW server (http://www.fruitfly.org/seq_tool/promoter.html). Search for transcription factor binding site was performed using Tfsitescan software on the MIRAGE WWW server (<http://www.ifti.org/cgi-bin/ifti/Tfsitescan.pl>) and Signal Scan software on the BIMAS WWW server (<http://bimas.dcrn.nih.gov/molbio/signal>).

Southern Blot Analysis

About 10 µg of *P. monodon*'s genomic DNA were digested with selected restriction enzymes. The genomic fragments were separated on 0.7% agarose gel at 30 V for 21 h and were subsequently transferred to GeneScreen Plus nylon membrane (Dupont) with 0.4 N NaOH using the vacuum blotter (Model 785, Bio-RAD).

The DNA probe was synthesized by incorporation of Fluorescein-11-dUTP (Fl-dUTP) in a random prime labelling reaction using the Gene Images random prime labelling module (Amersham Pharmacia Biotech). Three DNA probes specific to *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3* genes were synthesized from the intronic region of each gene. The fourth probe was synthesized from the coding region of the mature *Pem-CHH1* that is nearly identical to the same region of the other two *Pem-CHH* genes. The primers used for amplification of each probe were as follows: forward CHH1-IF (5' AACCCAACCCAGTTGTGTTGT 3') and reverse CHH1-IR (5' CCACTGGTAGGTCGTACAGGA 3') for CHH1-specific probe, forward CHH2-IF (5' TTCGTATAGCGGATCCGA 3') and reverse CHH2-IR (5' TCAGTTTTGTT CGGAGGG 3') for CHH2-specific probe, forward CHH-F (5' TCAGTGCAGAGGG AGAGCC 3') and reverse CHH2-IR followed by *Acc* I digestion for CHH3-specific probe and forward CHH-mat (5' GGAATTCCATATGAGCCTATCCTTCAG 3') and reverse CD-R (5' CGGGATCC CTACTTGCCGAGCCTCTG 3') for the mature CHH1 probe.

Hybridization was performed in 5xSSC, 0.1% (w/v) SDS, 5% (w/v) dextran sulphate and 20-fold dilution of liquid block (Amersham Pharmacia Biotech) at 65°C overnight. After hybridization, the membrane was washed at 65°C in 1xSSC,

0.1%SDS for 15 min followed by another 15 min- wash in 0.2xSSC, 0.1%SDS. The detection step was conducted according to the protocol of Gene Images CDP-Star detection module (Amersham Pharmacia Biotech).

Results

Nucleotide sequences and structure of Pem-CHH2 and Pem-CHH3 genes

PCR amplification of *P. monodon*'s genomic DNA with specific primers for *Pem-CHH2* and *Pem-CHH3* gene resulted in specific products of similar sizes around 1.2-1.3 kb. Nucleotide sequence analysis showed that *Pem-CHH2* and *Pem-CHH3* genes are 1,274 bp and 1,263 bp long, respectively. Both genes contain three exons that are interrupted by two introns at the same positions. The first intron is located between the codons for ⁶Leu and ⁷Val in the signal peptide whereas the second intron interrupts the codon for ⁴⁰Arg of the mature peptide (Fig. 2). The exon-intron boundary conforms to the splice donor and acceptor consensus sequence (Mount, 1982). The schematic structure of *Pem-CHH2* and *Pem-CHH3* genes in comparison to that of the previously characterized *Pem-CHH1* is shown in Fig. 3.

Cloning and characterization of the flanking regions of Pem-CHH genes

The length of the sequences walked into both 5' and 3' regions of each *Pem-CHH* gene was summarized in Table1. DNA sequences analysis of the amplified fragment from the first 5' walk of *Pem-CHH1* gene revealed two types of the 5' upstream sequences. The differences between these two types of sequences were due to both nucleotide variations and deletions as shown in Fig. 4A, whereas a single type of sequences was detected in the 3' downstream region. A few types of dinucleotide repeats, AC, AT and GT, were found in the sequences about 300 nucleotides upstream of *Pem-CHH1* transcription initiation site. Longer repeats of 27 nucleotides were found in the 3' flanking region of this gene (Fig. 4B). However, these repeats were not found within the sequences upstream and downstream of *Pem-CHH2* and *Pem-CHH3* genes that were characterized in this study. Analyses of the upstream and downstream sequences of the three *Pem-CHH* genes by BLASTN and BLASTX program did not show significant match to any sequences in the databases (data not shown).

Comparison of the 5' flanking region among the three *Pem-CHH* genes revealed a marked sequence similarity within 200 nucleotides upstream from the transcription start site (+1) of *Pem-CHH2* and *Pem-CHH3* genes, whereas the same region of *Pem-CHH1* showed higher degree of variation. The sequences further upstream were more diverged among the three genes (Fig. 5A). Similarly, the 3' flanking regions of *Pem-CHH2* and *Pem-CHH3* contained nearly identical sequences, whereas the sequence of *Pem-CHH1* showed only scattered identical nucleotides to that of *Pem-CHH2* and *Pem-CHH3* in this region (Fig. 5B).

Analysis of the putative promoter and regulatory elements

Prediction for potential upstream regulatory elements within 500 nucleotides of the upstream sequences was carried out by using MIRAGE Tfsitescan software and the BIMAS Signal Scan software. A putative TATA box-like element (TATAA) was found located at 23 nucleotides upstream of the +1 position of each gene. In addition,

several potential recognition sites for transcription factors were identified (Fig. 6). The putative binding sites for cAMP responsive element binding protein (CREB), pituitary-specific transcription factor-1 (Pit-1) and activator protein-1 (AP-1) were found in the upstream region of all three *Pem-CHH* genes.

Copy numbers and organization of Pem-CHH genes in the genome of P. monodon

Southern blot analysis was performed in order to gain preliminary data on the copy numbers and organization of *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3* genes in *P. monodon*'s genome. The nucleotide sequences specific to each *Pem-CHH* gene were used as a probe to detect individual *Pem-CHH* gene in the genome. These probes were unable to hybridize to one another (Fig. 7 A). *P. monodon*'s genomic DNA was separately digested with three restriction enzymes that do not cut within any of the three *Pem-CHH* genes i.e. *Bgl* I, *Eco*R I and *Hind* III. The result in Fig. 7 B showed that there were at least 3 copies of *Pem-CHH1* gene, one copy of *Pem-CHH2* genes and 2 copies of *Pem-CHH3* gene in the genome according to the minimum number of the bands hybridized to each probe. Some common DNA fragments were detected by more than one probe. The fragment of about 7 kb in *Eco*R I-digested DNA was detected by both CHH2 and CHH3 probes and the 9 kb band in *Eco*R I-digested DNA was detected by both CHH1 and CHH3 probes. In addition, the band above 10 kb in *Hind* III-digested DNA was detected by all probes. All of the DNA fragments detected by each gene-specific probe together with several additional bands were hybridized to the mature CHH 1 probe (Fig. 8).

Discussion

The nucleotide sequence analysis of *Pem-CHH2* and *Pem-CHH3* genes indicated that both genes are encoded by three exons similar to the genes encoding CHH and MIH from other species that have been characterized so far (Gu and Chan, 1998, Lu et al., 2000, Gu et al., 2000). Although the nucleotide sequences and the length of introns in those genes are varied, their positions are relatively conserved. In CHH genes of *M. ensis* (Gu and Chan, 1998, Gu et al., 2000) as well as *Pem-CHH2* and *Pem-CHH3* genes, the first intron is located after the codon of ⁶Met and ⁶Val, respectively. The second intron is always located within the codon of Arg at position 40 or 41 in all of the genes mentioned above. Interestingly, the previously characterized *Pem-CHH1* gene is encoded by only two exons (Udomiit et al., 2000). However, the single intron of *Pem-CHH1* gene is also located at the same position as intron 2 of *Pem-CHH2* and *Pem-CHH3* genes (see Fig. 3). The presence of introns at conserved positions suggests that these genes are closely related.

In this study, about 0.7-1.2 kb of the upstream and downstream flanking regions of three *Pem-CHH* genes were characterized. The nucleotide sequences in both regions of each gene were not similar to any known sequences in the database. This, together with the absence of long open reading frame in these regions, suggests that all flanking regions characterized are composed of non-coding DNA. Although the coding regions are highly conserved among the three *Pem-CHH* genes, the introns of *Pem-CHH2* and *Pem-CHH3* are more similar to each other than to that of *Pem-CHH1*. Comparison of the upstream and downstream sequences of these three genes also showed that the nucleotide sequences in these regions of *Pem-CHH1* were less

conserved than that of *Pem-CHH2* and *Pem-CHH3*. Taken together with the fact that *Pem-CHH1* contains only a single intron, whereas both *Pem-CHH2* and *Pem-CHH3* contain two introns, *Pem-CHH1* seems to have gone through independent evolutionary path from its related copies, *Pem-CHH2* and *Pem-CHH3*. The presence of two types of the sequences in the 5' flanking region of *Pem-CHH1* may reflect variations in different alleles.

The presence of the putative site for CREB (cAMP responsive element binding protein), a ubiquitous transcription factor that induces gene transcription by activating protein kinase A (Silva *et al.*, 1998), in the upstream region of all *Pem-CHH* genes suggests that the regulation of expression of these genes is cAMP dependent. The presence of the sequence with high similarity to the binding site for the major pituitary transcription factor controlling expression of growth hormone and prolactin genes in vertebrates, Pit-1 site (Bradford *et al.*, 2000), suggests the involvement of a tissue specific transcription factor in the regulation of expression of *Pem-CHH* genes. Another significant match is the homologue of the binding site for AP1 transcription factor. In *Drosophila*, AP1 protein controls cell growth and differentiation (Zhang *et al.*, 1990). Binding sites for similar transcription factors were also found in the other characterized CHH, MIH and MO-IH genes (Gu and Chang, 1998, Gu *et al.*, 2000, Lu *et al.*, 2000). This demonstrates that the genes in the CHH/MIH/GIH family might be under control of similar mechanisms. Further investigation of these putative transcription factors in *P. monodon* and other crustaceans is thus promising in order to understanding more clearly about the transcriptional control of the hormones in this family.

Southern blot analysis showed that the genome of *P. monodon* contains at least three, one and two copies of *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3* genes, respectively. The DNA probe generated from the mature peptide-coding region of *Pem-CHH1* can detect several extra bands in addition to those detected by each *Pem-CHH*-specific probe. This probe is 60-80% homologous in its nucleotide sequence to those of the other five CHH-like genes (Pm-sgp-I to -V) of *P. monodon* reported by Davey *et al.*, 2000. It is therefore possible that the additional hybridizing bands are corresponding to the presence of Pm-sgp-I to -V and probably other genes in the CHH family.

In the shrimp *M. ensis*, there are at least eighteen genes encoding two major types of CHH peptides that share more than 98% identity in their amino acid sequences (Gu *et al.*, 2000). These genes were arranged into two different clusters. In addition to CHH, other hormones in this family, MIH and MO-IH, of *Cancer pagurus* were also encoded from genes that are clustered within a 6.5 kb region of the genome (Lu *et al.*, 2000). Our results also give preliminary information on the clustering of *Pem-CHH* genes in the genome of *P. monodon*. A single genomic band about 10 kb obtained from *Hind* III digestion hybridized to all *Pem-CHH* probes suggesting that at least one copy of each *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3* genes are clustered within this *Hind* III-digested fragment. Further studies on genomic organization of *Pem-CHH* genes in *P. monodon* would provide an implication for their evolution as well as the control of their expression.

Acknowledgements

This work was supported by The Thailand Research Fund (TRF). AW is supported by the Graduated Fellowship Program, National Science and Technology Development Agency (NSTDA). SP is a TRF Senior Research Scholar.

References

1. Bradford, A. P., Brodsky, K. S., Diamond, S. E., Kuhn, L. C., Liu, Y., Gutierrez-Hartmann, A., 2000. The Pit-1 homeodomain and beta-domain interact with Ets-1 and modulate synergistic activation of the rat prolactin promoter. *J. Biol. Chem.* **275**, 3100-3106.
2. Chang, E.S. 1997. Chemistry of crustacean hormones that regulate growth and reproduction. In *Recent advances in marine biotechnology* (ed. M. Fingerman, R. Nagabhushanam and M-F. Thompson) vol. 1, pp. 109-161. Science Publishers, U.S.A.
3. Cooke, J.M., Sullivan, R.E. 1982. Hormones and neurosecretion. In: Bliss, D.E. (Ed), *The Biology of Crustacea*, vol. 3., Academic Press, New York, USA, pp. 205-290.
4. Davey, M.L., Hall, M.R., Willis R.H., Oliver R.W., Thurn M.J., Wilson K.J., 2000. Five crustacean hyperglycemic family hormones of *Penaeus monodon*: complementary DNA Sequence and Identification in single sinus glands by electrospray ionization-fourier transform mass spectrometry. *Mar. Biotechnol.* **2**, 80-91.
5. De Kleijn, D.P.V., Janssen, K.P., Mortens, G.J., Van Herp, F., 1994. Cloning and expression of two crustacean hyperglycemic hormone mRNAs in the eyestalk of the crayfish *Orconectes limosus*. *Eur. J. Biochem.* **224**, 623-629.
6. De Kleijn, D. P., de Leeuw, E. P., van den Berg, M. C., Martens, G.J., van Herp, F., 1995. Cloning and expression of two mRNAs encoding structurally different crustacean hyperglycemic hormone precursors in the lobster *Homarus americanus*. *Biochim. Biophys. Acta.* **1260**, 62-66.
7. De Kleijn, D.P.V., Janssen, K.P., Waddy, S.m L., Hegeman, R., Lai, W. Y., Martens, G. J. M., Van Herp, F. 1998. Expression of the crustacean hyperglycemic hormones and the gonad-inhibiting hormone during the reproductive cycle of the female American lobster *Homarus americanus*. *J. Endocrinol.* **156**, 291-298.
8. Gu, P.L. and Chan, S.M., 1998. The shrimp hyperglycemic hormone-like neuropeptide is encoded by multiple copies of genes arranged in a cluster. *FEBS Lett.* **441**, 397-403.
9. Gu, P-L., Yu, K. L., Chan, S-M., 2000. Molecular characterization of an additional shrimp hyperglycemic hormone: cDNA cloning, gene organization, expression and biological assay of recombinant proteins. *FEBS Letts.* **472**, 122-128.
10. Keller, R., 1992. Crustacean neuropeptides: structures, functions and comparative aspects. *Experientia.* **48**, 439-448.
11. Khayat, M., Yang, W-J., Aida, K., Nagasawa, H., Tietz, A., Funkenstein, B., Lubzens, E., 1998. Hyperglycemic hormones inhibit protein and mRNA synthesis in *in vitro*-incubated ovarian fragments of the marine shrimp *Penaeus semisulcatus*. *Gen. Comp. Endocrinol.* **110**, 307-318.
12. Lu, W., Wainwright, G., Webster, S.G., rees, H.H., Turner, P.C., 2000. Clustering of mandibular organ-inhibiting hormone and moult-inhibiting hormone genes in the crab, *Cancer pagurus*, and implications for regulation of expression. *Gene.* **253**, 197-207.
13. Santos, E. A., Nery, L. E., Keller, R., Goncalves, A. A., 1997. Evidence for the involvement of the crustacean hyperglycemic hormone in the regulation of lipid metabolism. *Physiol. Zool.* **70**, 415-420.
14. Silva, A. J., Kogan, J. H. and Frankland, P. W., 1998. CREB and memory. *Annu. Rev. Neurosci.* **21**, 127-148.

15. Spanings-Pierrot, C., Soyeux, D., Van Herp, F., Gompel, M., Skaret, G., Grousset, E., Charmantier, G., 2000. Involvement of crustacean hyperglycemic hormone in the control of gill ion transport in the crab *Pachygrapsus marmoratus*. *Gen. Comp. Endocrinol.* **119**, 340-350.
16. Tensen, C.P., De Kleijn, D.P., Van Herp, F., 1991. Cloning and sequence analysis of cDNA encoding two crustacean hyperglycemic hormones from the lobster *Homarus americanus*. *Eur. J. Biochem.* **200**, 103-106.
17. Udomkit, A., Chooluck, S., Sonthayanon, B., Panyim, S., 2000. Molecular cloning of a cDNA encoding a member of CHH/MIH/GIH family from *Penaeus monodon* and analysis of its gene structure. *J. Exp. Mar. Biol. Ecol.* **244**, 145-156.
18. Wainwright, G., Webster, S. G., Wilkinson, M. C., Chung, J. S., Rees H. H., 1996. Structure and significance of mandibular organ-inhibiting hormone in the crab, *Cancer pagurus*. *J. Biol. Chem.* **271**, 12749-12754.
19. Yang, W. J., Aida, K., Nagasawa, H., 1997. Amino acid sequences and activities of multiple hyperglycemic hormones from the Kuruma prawn, *Penaeus japonicus*. *Peptides*. **18**, 479-485.
20. Zhang, K., Chaillet, J.R., Perkins, L.A., Halazonetis, T.D., Perrimon, N., 1990. Drosophila homolog of the mammalian *jun* oncogene is expressed during embryogenesis development and activates transcription in mammalian cells. *Proc. Natl. Acad. Sci. USA*. **87**, 6281-6285.

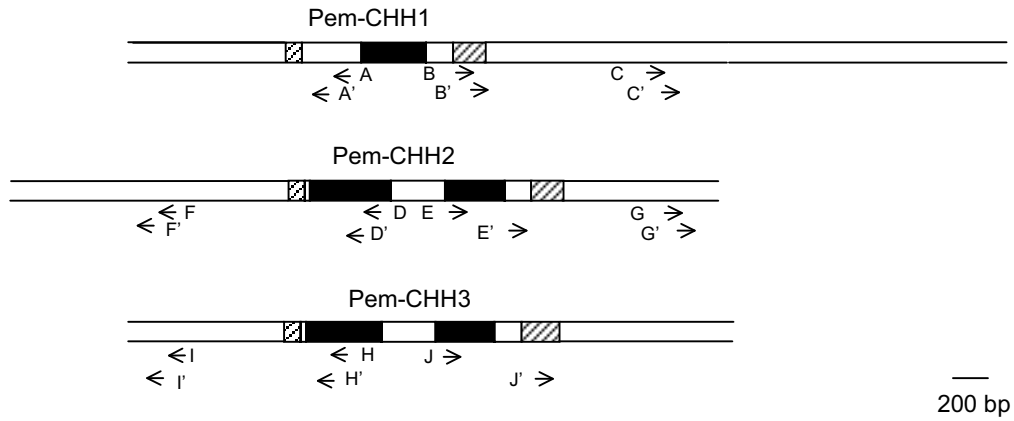


Figure 1. The diagram showing relative positions of primers used in genomic walking into the 5' and 3' flanking regions of *Pem-CHH1*, *Pem-CMG2* and *Pem-CMG3* genes. The hatched boxes represent 5' and 3' UTR, filled boxes represent introns and empty boxes represent exons. Double lines indicate the regions that have been characterized in this study. Broken lines indicate the region that have been cloned and remain to be characterized. The sequences of the primers are shown below.

A : 5' -CCTGAAGGATAGGCTGCGTTTGTCTGA-3'
A' : 5' -GTGAAGGGAAGTCACAGGAGAGGCTTCT-3'
B : 5' -AGACCTCGCCATGCGACTCCCAAGAC-3'
B' : 5' -GCAAGCTTCTGTACAACAAGCCTCCTGTCTAATA-3'
C : 5' -GATGCCTACACCGAGCCTCTGAGGTCC-3'
C' : 5' -CTATCAGACTGGCTGCACCATCACTGT-3'
D : 5' -TAAGGAAAAAATCGGACCTTGCTGAAA-3'
D' : 5' -CAAATGAATAATGCATACAGCGTCAGA-3'
E : 5' -GGGTAAATACTGTTATCATTTAATAGGTCG-3'
E' : 5' -GGTGTTCCTGTACTGCGTGGACTACATG-3'
F : 5' -TTACCGGTAGTTGAACTATGAACTGTGC-3'
F' : 5' -ATTCAAGATTTCGCATCCGCTGATGTTG-3'
G : 5' -TAGCCTAGACTTACTGCTGGCGTACA-3'
G' : 5' -CAGAACATTTCAGCATCTGTAGCCATAC-3'
H : 5' -AGGGGATATACAGGGCTTTACATTTCT-3'
H' : 5' -CAGGTTAAGCTCTGCAATTTACAGAGC-3'
I : 5' -GCAGTTTTTCCTATATATCGGTGAAGGGAAC-3'
I' : 5' -GATTGGAATATTTCGAGATTTCGCATTCG-3'
J : 5' -CAAGGGCAGATGTAATGGTATACTTTTCAG-3'
J' : 5' -GGTGTTCCTGTACTGCGTGGACTACATG-3'

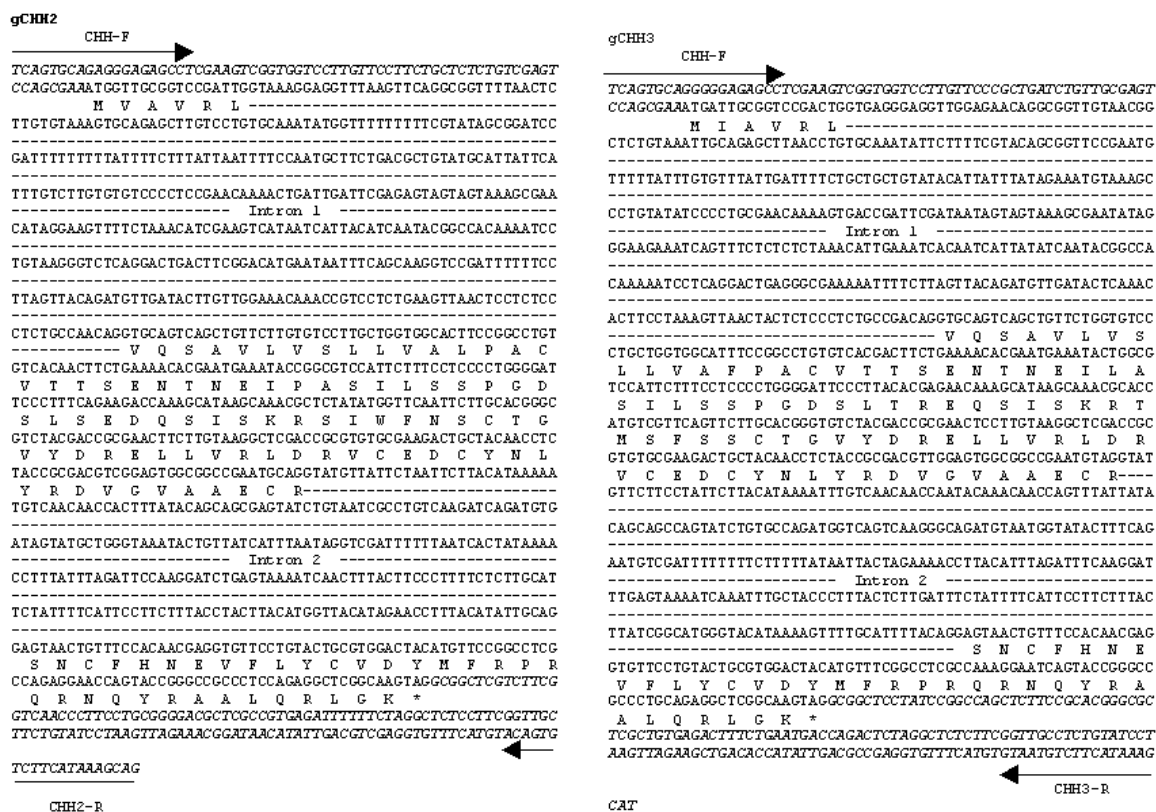


Fig. 2. Nucleotide and deduced amino acid sequences of *Pem-CHH2* (A) and *Pem-CHH3* (B) genes of *P. monodon*. Introns are indicated by dashed lines. Amino acids are shown in one-letter symbol and the asterisk marks the stop codon. Arrows indicate the nucleotide sequences corresponding to the primers used for amplification of each gene.

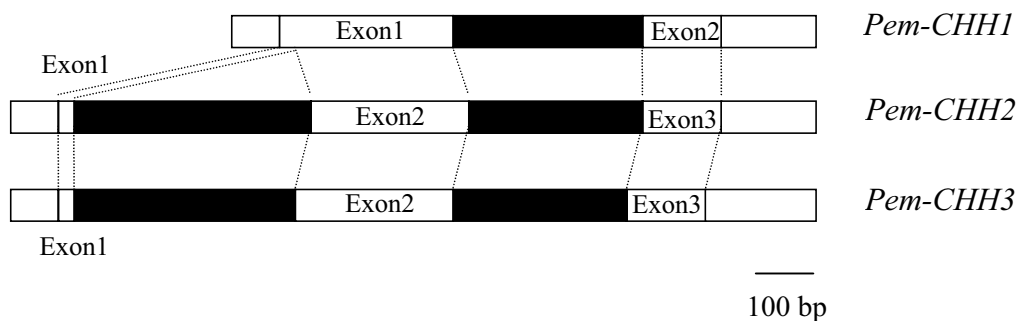


Fig. 3. Comparison of the gene structure of *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3*. The hatched boxes represent the 5' and 3' untranslated regions. Exons and introns are shown in white and black boxes, respectively.

Gene	Dircetion of walk	libraries	Sizes of the flanking fragment (bp)
<i>Pem-CHH1</i>	5'	<i>Dra</i> I	690
	3'	<i>Hpa</i> I (first walk)	1,124
		<i>Dra</i> I (second walk)	902
<i>Pem-CMG2</i>	5'	<i>SnaB</i> I (first walk)	782
		<i>Xmn</i> I (second walk)	371
	3'	<i>Hpa</i> I (first walk)	606
		<i>Dra</i> I (second walk)	53
<i>Pem-CMG3</i>	5'	<i>Dra</i> I (first walk)	685
		<i>Xmn</i> I (second walk)	144
	3'	<i>Hpa</i> I	721

Table 1. Summary of the 5' and 3' flanking sequences of *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3* genes obtained in this study.

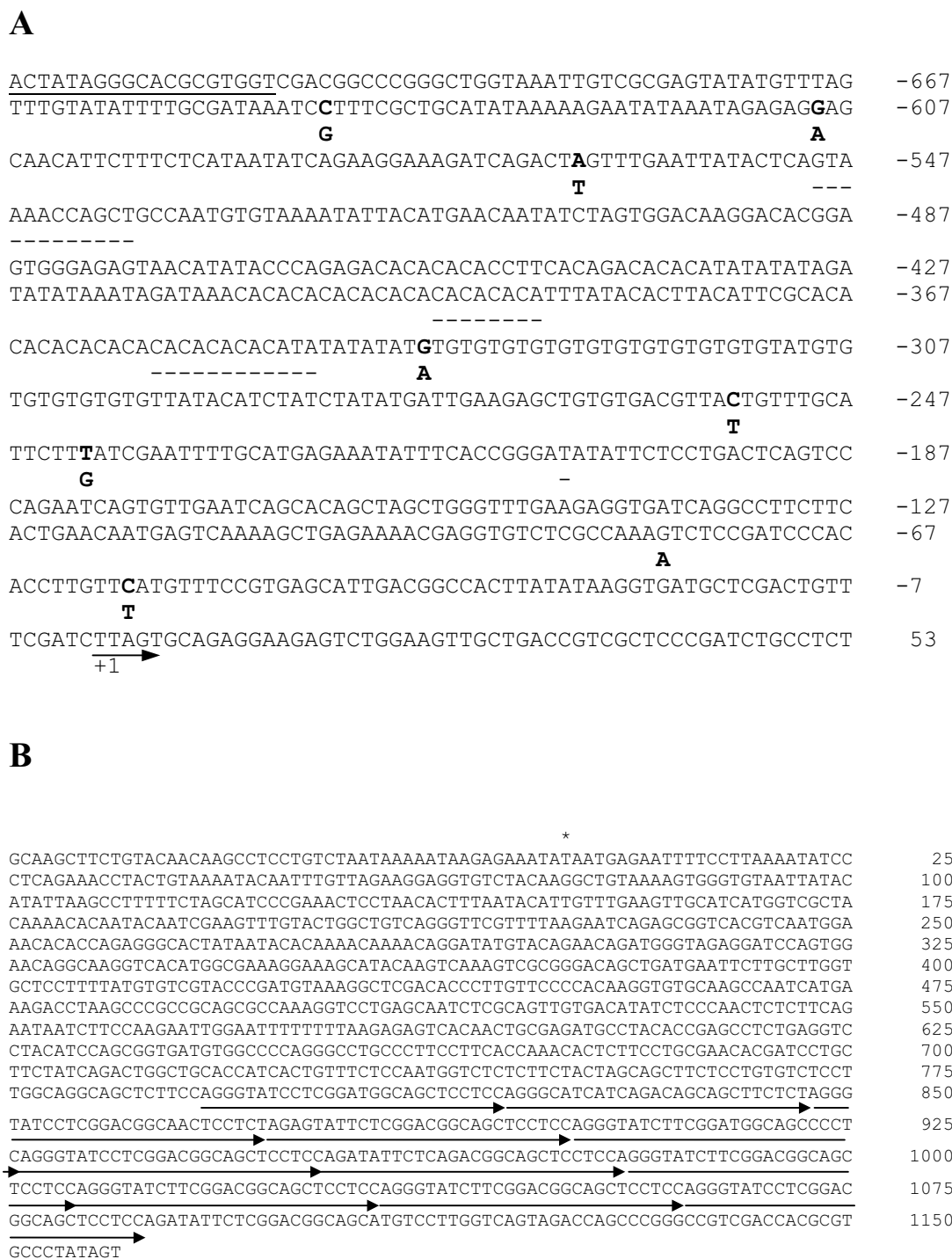


Fig. 4. Nucleotide sequences of the 5' and 3' flanking regions of *Pem-CHH1* gene.

A. The 5' flanking region of *Pem-CHH1*. The numbers on the right indicate the positions upstream of the transcription start site (+1). The sequence of AP2 primer is underlined. The variation in nucleotides and deletion regions are shown in bold letters and as dash line, respectively, below each line.

B. The 3' flanking region of *Pem-CHH1*. The asterisk represent the 3' end of the gene. The numbers on the right indicate the positions downstream of the gene. The 27 bp repeats were indicated by arrows.

5' CHH1 -----TACAAGGACACGGAGCTGGAGAGTAACATATATCCGACAGACACACACACCTTCACAGACACATATATATAGATATATATAAAT
5' CHH2 -----TCACAAAGAAATGGAGCTTATAATAATAGCCTGAATAATTCACCATTTGATTTGATTAATCATTTTATGTT
5' CHH3 TTTTITTTGGTTTCTATTTCTGTATTCTCCTAATGAACACATAGCTTAATAATC-TGCGATTATTGGCAATCACAAATAATGGACTCTTATTAAT

5' CHH1 GATAAACACACAACACACACACAACATTTATACAGTTAC--ATTTCGCACACACACACA-CACACACACACATATATATATGTGTGTGTGTGT
5' CHH2 TACTGAGCTATTATATTTATCTTATTTATTTTACTAAATCCCTGTAGGTGATGAGAAATCTAGGATTCAGATGAATGACCTGCTCTCTCC
5' CHH3 AATACCTGTTCCATACTCATATGCTCTTTTTCAAATTTTGGT--ATTCAGTTCCTTCACG-CATATATAGGAAATGCGCCATAGTCTTCATCT--

5' CHH1 TCTGTGTG-TCTGTATGCTGTGTGTGTGTTTATATCATCTATCTATATGATGAAGAGCTCTGTGTGACTTAATGTTTCATCTCTT--TTTCAAT
5' CHH2 GGTCAATGTCTAAAGTTTATGAGCAACCTAATATATCTCATTTATATAT--ATATTCTCTTCA-----CAGACATATATATTTACTTACAGCT
5' CHH3 -GTCAATATCAATTTATTTATTTATCTATCTCATTTTATTAATGTTATGAT--ATTTAAGTAAGGCTTCAGCTCGATAGG---AAGTCAAGCA

5' CHH1 TTGCATGAGAA-ATAATTTCACGGGAT--ATAATTCCTCTGACTCA-----GTCCC-GAATCAGTGTGTAATCCGCACAGCTAGGTGGGTTTGAA
5' CHH2 GGCACAGATCATACCTAACCTGAGATTTCTGTATGACACAGAGCAAGGGTCTCTTTACATCCGGGAATTCACACAGCTGACCAACTTTTGAA
5' CHH3 TTTCCACTGCT-ATATTTCGCACGAGAT--GAAGTTTTCCTGACA-CAGCGGCTCG---TGCAATCCGGGAATTAACACAGCTGACCAACTTTTGAA

5' CHH1 GAGGTGATCAGGCCCTC-----TTCACTGAACAATGAG-TCAAAAGCTTGAGAAACGAGGTGTCTCCGCAAACTTCCGATCCACACCTTGTTTAT
5' CHH2 GAGGTGATCAATCCTCACCATTCAACAGACCTTGAGATCAAAAGGCTTTAAACCAAGATGTTTCGCGAAGGCTCTGTAGCCACTACTCTTGGCTCAT
5' CHH3 GAGGTGATCAACCTTCAACATTTCACAAACATTTCCGGATCAAAAGGCTTCGAAACCAAGGCTTTTCGCGAAGGCTCTGTAGTTAATAGCTTGGTACAT

5' CHH1 GTTTCGGTGAG-CATTGACGGGCACCTTATATAAGCGATGCTCTGACTGTITC-GACC
5' CHH2 GTTTCATCATATACCAACGGGACGCTTATATAAGGGATGTTTCACTTGTTTCTGAG-
5' CHH3 GTTTCCTC-----GCTTATATAAGCGGGTTCCTACTGTTTCTGAG-

3' CHH1 -----AATGAGAAATTTTC-----TTAAATATCCCTCAGAAACCTACTGTAAATACATTTTGTTAGAGGAGCTGTCTACAGGCTGTAA
3' CHH2 AATATATAAGAGATTTTTTCCTGATTTTCAAT-TCCTTTTGATATATACT-AAAACATAATGTACGATCTGAGTGAGTCAAAATCTAAATA
3' CHH3 AAAACCAAAGGAAATATTT-ATCTTTAAATTCCTGTTGATACATACT-ATAAG--AGTGTACAAATCTGCAATGAGTCAAAATGGTGAATA

3' CHH1 -----GTTGGTGTAAATTATACATTTTA-AGCCTTTT-CTAGCAT-CCGGAANCTCTTAACATTTTAATACATTTGTTTGAGG-TCGA
3' CHH2 TACACATCAGTTTATGTA--AACTATAGACAGAAATATATGTAACAGACCAGATAAAGATATA--T-AGTCGCTTTTGTGCTT-TGTCA
3' CHH3 --CACACAAGTTTATTTTTATGCAATAGACAGAAACT--GTAATAAGACCAGATAAAGATAAAGTGTGTTAGTCATCTGTGTGCAATGATCGA

3' CHH1 CATGCTGCTTCAAAACACAATACATGGAAGTTTGTA--CTGCG-----TGTCGCGTTCGTTTAAAGATCAGACGCTACGCTCAATGG
3' CHH2 ACTGTACAGGTTTCAAACTCTTAAATGTAAA-----CA-----ATATTAAGAT-GAATTAATAAGATGAATTAAT-CTTAAATG
3' CHH3 ACTGTACAGGATTTAAATCTTTAAATACAAATTAATCGATCTGTCCTCAGATATCAAGAT-GAATTAATAAGATGAATGAG-CTTTAA-TG

3' CHH1 AACACACCAGAGGGCCTATAATTATACACAAAACAAAACAGGATATGACAGAAAGATGGGTAGAGCTCCAGTGGAACAGGCAGGCTCACATGGCC
3' CHH2 AAC-CACC-GATGAAGCCACAAGCAATAATATCTTTCTGCTTC-TGCGCTTAATTTCCGTACGGCTCC-----C-AGTTACGAAAG
3' CHH3 AAC-CACC-GATGAAGCCACAAGCAATAATGCTTTCTGCTTC-TGCGTAATATTTTGTATCTTATCC-----CAAAGTTACAAAAG

3' CHH1 AAAGGAAAGCATACAAGTCAAAGTGGCGGAGCGCTATGAATCTTGGCTGGTCTTCTTTATGTTGTCGTACCGCTGTAAAGGCTCGACACC
3' CHH2 CGAGTACAGCTTTCAAGTGTA--TCAAGTTAAAGCTGT--AAATATTTTTC-TGTTTCTATTTCTACT-----CCCTAACTTAGCGAGACTTA
3' CHH3 TGAGTACAGCTTTCAAGTGCA--TCAATTTAGACTT-TTTTTT-TTCTATTTCTACT-----CCCTGGTTTAGCGAGGACTTA

3' CHH1 TGTCTCCCAACAA-----GGTGTCTAGCCAAATATGAAAGACCTTAAGCC-----CGCCGACGCGCA-----
3' CHH2 TGTCTGCTACAAAAGAGCTTGAATTAATTTTGTAGCGGTTCCGCCATACATTTCAGATTACGTTACGTTTGTG
3' CHH3 TGTCTACCTACAAAAGAGCTTGGCAATTAATTTGTAGCGGTTCCCGTGGGCCCTTCAGATTACCTTACAA

Fig. 6. The putative promoter sequence and regulatory elements identified in the upstream region of *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3* genes. Putative TATA box and binding sites for transcription factors were highlighted. Arrows indicate the transcription start site (+1) of each gene (Udomkit *et al.*, 2000; unpublished data). Numbers on the right show the positions relative to the transcription start site.

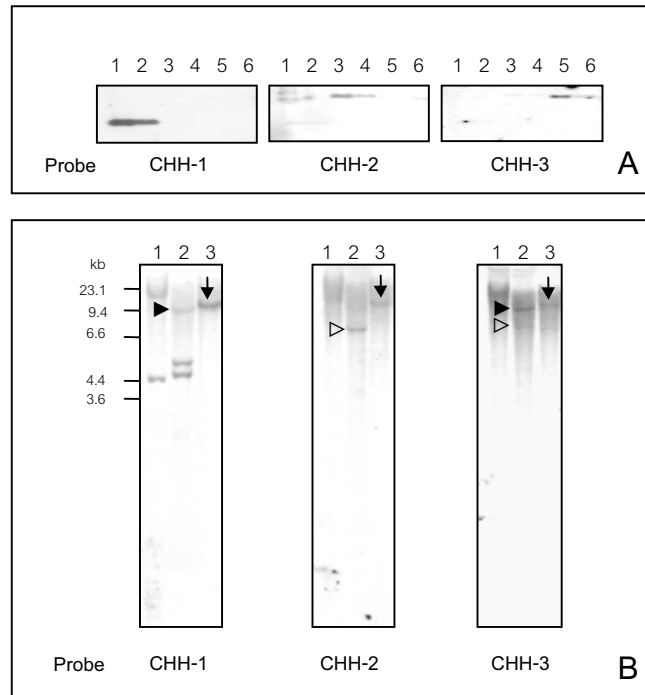


Fig. 7. Southern blot analysis of *Pem-CHH* genes of *P. monodon*. The same membrane was used in the hybridization with all DNA probes. (A) The DNA probes specific to each *Pem-CHH* genes were hybridized to 30 pg and 10 pg of *Pem-CHH1* genomic fragment (lanes 1 and 2) that are equivalent to three and one copies of *Pem-CHH1*, respectively. Lanes 3 and 4 represent three and one copy-equivalent of *Pem-CHH2* and lanes 5 and 6 represent three and one copy-equivalent of *Pem-CHH3* gene, respectively. (B) About 10 μ g of *P. monodon*'s genomic DNA digested with *Bgl* I, *Eco*R I and *Hind* III were loaded in lanes 1 to 3, respectively. The filled arrowheads indicate the bands of the same size that were detected by CHH1- and CHH3-specific probes. Opened arrowheads indicate the bands of the same size that were detected by CHH2- and CHH3-specific probes. The bands that were detected by all CHH1-, CHH2- and CHH3-specific probes were marked by arrows.

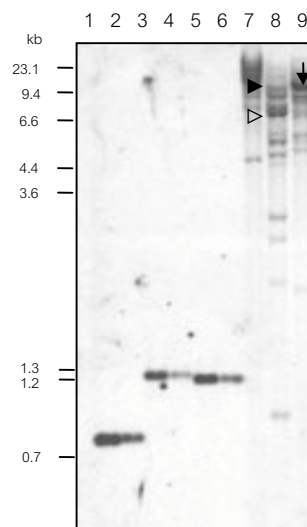


Fig. 8. The membrane in Fig. 7 was re-hybridized with the mature CHH1 probe. Lanes 1 to 6 represent three and one copy-equivalent of *Pem-CHH1*, *Pem-CHH2* and *Pem-CHH3*, respectively. Lanes 7 to 9 are genomic DNA of *P. monodon* digested with *Bgl* I, *Eco*R I and *Hind* III, respectively. The bands that were also detected by CHH1- CHH2- and CHH3-specific probes are indicated as described in Fig. 7B.

PCR-BASED METHOD FOR EPIDEMIOLOGICAL IDENTIFICATION OF THE PATHOGENIC *BURKHOLDERIA PSEUDOMALLEI*

Sumalee Tungpradabkul¹, Piengchan Sonthayanon², Piamnukul Krasao¹, Vannaporn Wuthiekanun³, and Sakol Panyim^{1,2}

¹Department of Biochemistry, Faculty of Science, ²Institute of Molecular Biology and Genetics, ³Wellcome Unit, Faculty of Tropical Medicine, Mahidol University, Thailand

Abstract

We have previously shown that *Burkholderia pseudomallei*, the causative pathogen of melioidosis, can be discriminated from the closely related non-pathogenic species *Burkholderia thailandensis* by the presence of a 15 base pair deletion in the flagellin gene of *B. thailandensis*. Using specific flagellin gene primers flanking the distinctive region, PCR products of 191 and 176 bp in size were detected for *B. pseudomallei* and *B. thailandensis*, respectively. The sensitivity of detection is 20-80 colonies forming units per reaction of *B. thailandensis* and *B. pseudomallei* cell suspension. To mimic the expected environmental situation, mixed populations of the two species were analyzed. The results showed that the PCR-based method could be used to distinguish the two species in a duplex reaction. However, the sensitivity of detection in *B. thailandensis* was higher than in *B. pseudomallei*. We have investigated and found a certain stem loop with a stable energy of -81.6 kcal/mole within a flagellin sequence of *B. pseudomallei* but not in the *B. thailandensis*. This finding suggested that the loop formation within the *B. pseudomallei* flagellin sequence is a factor involved in the lower sensitivity of the detection. We therefore improved the optimal condition by decreasing the primers concentration and that the sensitivity was increased to be the same sensitivity for detecting *B. thailandensis*. Finally, direct detection of the both bacteria from soils could be detected with the sensitivity of 1,000 cfu/reaction.

Keywords: Flagellin sequence, *Burkholderia pseudomallei*, *Burkholderia thailandensis*, polymerase chain reaction (PCR), stem loop formation

Introduction:

Burkholderia pseudomallei is the causative organism of melioidosis, a disease that is endemic to much of southeast Asia and northern Australia. The organism is a motile, Gram-negative rod and is a free-living saprophyte in the environment of endemic areas such as soil and stagnant water. *B. pseudomallei* had been differentiated into 2 distinct biotypes by the ability to assimilate the pentose sugar L-arabinose [1]. The ara⁺ biotype is capable of assimilating L-arabinose whereas ara⁻ biotype cannot. These two biotypes have also demonstrated differences in pathogenicity. The ara⁻ biotype is virulent while the ara⁺ biotype is nonvirulent. The differences in genotype, phenotype and pathogenicity of these two biotypes have been

reported, and it has been proposed that the nonvirulent ara⁺ biotype be named as a new species, *Burkholderia thailandensis*. In northeastern Thailand, 75% of soil isolates are ara⁻ biotypes and the other 25% are ara⁺ biotypes, although there are no reports documenting mixed population of *B. pseudomallei* and *B. thailandensis* in epidemiological surveys, possibly reflecting the inability of current assays to appropriately identify mixed population.

Recently we have developed a PCR-based method for differentiating these two species by using specific primers flanking a 15-bp deletion in the flagellin gene of the non-pathogenic *B. thailandensis* generating amplification products of 191 and 176 bp for *B. pseudomallei* and *B. thailandensis*, respectively [2]. We have also developed a simple method using a directly detected the bacteria culture, as well as analyzing mixed populations of the two species, mimicking the situation that may occur in the environment. The sensitivity of the detection was 80-20 colonies forming units per reaction of *B. pseudomallei* and *B. thailandensis* cell suspension, respectively. In addition, the sensitivity of a mimic mixed populations of the two species was shown a significant difference in which the amplified product intensity of *B. pseudomallei* was 4 times lower than *B. thailandensis* when using an equal number of template [3]. We therefore investigated a factor that involves in the sensitivity of the detection using Taq DNA polymerase with low fidelity and computer analysis of a stem loop formation program. We found a certain stem loop with a stable energy of -81.6 kcal/mole within a flagellin sequence of *B. pseudomallei* but not in *B. thailandensis*. Our finding clearly explain the unequally sensitivity for detecting *B. pseudomallei* and *B. thailandensis*.

In this study, we attempted to improve sensitivity and develop a simple method for detecting and differentiating *B. pseudomallei* and *B. thailandensis* for use in epidemiological studies. This simplified PCR protocol is expected to be considerable benefit in epidemiological status to understand the ecological relationships between *B. pseudomallei* and *B. thailandensis*.

Materials and Methods

PCR detection of *B. pseudomallei* and *B. thailandensis* from bacterial colony

A single colony from culture plate was selected and transferred to 50 ul of distilled water in 1.5 ml microtube. The suspension was boiled for 20 min and was diluted to 1:10, 1:100 and 1:1000. Approximately, 10 ul of each dilution was used as PCR template. The specific flagellin sequences of *B. pseudomallei* and *B. thailandensis* were amplified as follows. A total 50 ul reaction mixture containing 200 µM of each dNTPs, 1 µM or 0.5 µM of each primer, PMA-1 (5'-CTGTCGTCGACGGCCGT-3') and PMA-2 (5'-GGTTCGAGACCGTTTGCG-3'), 1x PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂ and 0.1% gelatin) 1 M Betaine, and 1 unit of Taq Polymerase enzyme. The 40 cycles-amplification were performed as follows. The mixture was preheated at 99 °C for 5 min, then add enzyme at 95°C followed by 40 cycles of 95°C 1 min, 65°C 30 sec. The final step was 72°C for 5 min. The amplification products were analysed on 3.5% agarose gel electrophoresis.

PCR detection of the mixed *B. pseudomallei* and *B. thailandensis* culture

1 ml overnight culture of *B. pseudomallei* and *B. thailandensis* were colony counting and were centrifuged at 3,000 rpm for 15 min. After colony counting, the cell pellet was suspended in 1 ml sterile distilled water. The suspension of *B. pseudomallei* and *B. thailandensis* were mixed at a ratio of 1:1, 1:2, 1:4, 1:8 (v/v) and 2:1, 4:1, 8:1(v/v), respectively. Approximately 10 ul of each 1:100 dilution mixture was used for PCR template as described above.

DNA sequencing and computer analysis

A short PCR product was obtained in *B. pseudomallei* template, after using low fidelity Taq DNA polymerase. Purification of a short PCR product was performed by GeneClean and subsequently sequenced by automated DNA sequencer. DNA sequence was aligned to the flagellin gene sequence of *B. pseudomallei*. Computer analysis of the flagellin sequence stem-loop formation was carried out using MFOLD from GCG program.

Detection of *B. pseudomallei* and *B. thailandensis* specific flagellin sequence from soil

Bacterial strains used in this experiment are *B. pseudomallei* NF10/38, NF47/38, E38(L58), and E271 and *B. thailandensis* E257, and E276. All bacterial strains were grown in LB media at 37° C for 20 hours with shaking at 250 rpm. The cultured cells were diluted 1:10 in sterile distilled water before 1 ml of the diluted cultured was inoculated into soil suspensions.

Three soil samples, two were collected from rice field in Hoi Kha Yung and Varinchamrab, Ubon Rachathani and one was collected from Hat Yai, Song Khla, were used in this experiment. They had pH (H₂O) of 5.0, 6.5, and 5.0 respectively as measured by pH paper and had different colors as shown in Fig. 3. Each soil was weighted into seven 1-gram portions and transferred to 7X15 ml conical tube for autoclave at 121° C for 20 min. A soil suspension was made by adding 1 ml of sterile distilled water to each portion of autoclaved soil and homogenized for 1 min using vortex mixer.

Inoculated soil was prepared by adding 1 ml of 1:10 diluted cells suspension into 1 portion of suspension soil and then mixed by vortex for 30 sec. The inoculated soils were left to sediment at room temperature overnight. Approximately 50 ul of each soil samples was boiled for 20 min. The soil suspension was diluted to 1:10, 1:100 and 1:1000. About 10 ul of each dilution was used for PCR template. A 10 mg/ml of BSA (Bovine serum albumin) were added in the reaction mixture to get 1 uM of a final concentration. The amplification method was performed as described above.

Results

PCR detection of *B.pseudomallei* and *B. thailandensis* from bacterial colony

The amplification product of specific flagellin gene sequence of *B.pseudomallei* is 191 bp whereas that of *B. thailandensis* is 176 bp. The product can be distinguished by 3.5 % agarose gel electrophoresis. The results show that as little as 1:1000 dilution of *B.pseudomallei* and *B. thailandensis* cell suspension, the expected product could be detected as shown in Figure 1. After colony counting, the

number of colony at 1:1000 dilution was approximately 20 to 80 cfu/ ml. Total 290 samples were amplified and analysed. The example of the results was shown in Figure 1 and Table 1. The PCR results of all 290 samples were corresponded with the arabinose assimilation test.

PCR detection of the mixed *B. pseudomallei* and *B. thailandensis* culture

A mixture of *B. pseudomallei* and *B. thailandensis* (7.2×10^7 cells/ml and 7.3×10^7 cells/ml respectively, from plate counts) at a ratio of 1:1 showed two DNA bands of PCR-amplified products of 191 bp (*B. pseudomallei*) and 176 bp (*B. thailandensis*). The intensity of DNA band of *B. pseudomallei* was lower than the intensity of the band of *B. thailandensis*. In mixtures which contained more number of *B. thailandensis* than *B. pseudomallei* at ratios of 2:1, 4:1 and 8:1 showed only one DNA band of PCR-amplified products of 176 bp. In contrast, mixtures containing more number of *B. pseudomallei* than *B. thailandensis* at ratios of 2:1, 4:1 and 8:1 showed two DNA bands of PCR-amplified products of 191 bp and 176 bp. The result was shown in Fig. 2.

PCR detection of *B. pseudomallei* and *B. thailandensis* using low fidelity Taq DNA polymerase

In order to reduce a cost for PCR detection of the two bacteria, Taq DNA polymerase was considered as a main step. Thus a homemade preparation of a recombinant *E. coli* harboring Taq DNA polymerase enzyme was used to express and purify the enzyme and that could reduce an expense in the PCR reaction. However, the enzyme has low fidelity property. To check this property of the assay, total 20 samples of environmental and clinical isolates were amplified and analyzed at a dilution of 1:1000. These isolates were analyzed in a blind test, without any prior bacterial data. The PCR amplification products were shown in Fig. 3. A 100% concordance was observed between PCR, size and arabinose utilization. However, sample number 6, 14 and 16 showed 2 bands of the PCR products. The upper band is equal to 191 bp or concordance with *B. pseudomallei* but the lower band is lower than 176 bp of *B. thailandensis*.

The short PCR products were then analyzed their sequence. The DNA sequence was aligned and compared to the flagellin gene sequence of *B. pseudomallei* as shown in Fig.4. The result demonstrated that a short PCR product was derived from the flagellin gene of *B. pseudomallei* by deleting 24 bp of the sequence generating a product of 167 bp in size. Using MFOLD from GCG program, a 191 bp sequence of *B. pseudomallei* was analyzed its stem loop formation as shown in Fig.5. A strongest loop formation was found within the region of 24 bp deletion of the short PCR product. This loop stabilizes with higher energy of -81 kcal/mol.

Optimization of PCR condition to improve sensitivity of PCR detection of *B. pseudomallei* and *B. thailandensis* culture

Since the loop formation within the flagellin gene in *B. pseudomallei* is a factor involving in the sensitivity of the detection, optimization of the PCR condition by decreasing primers concentration was performed. In Fig. 6, the result showed the optimal ratio between primers and template was 0.5 to 1 μ M of primers per 100-200

ng of DNA template and 0.5 μ M of primers per 0.1 ng of DNA template. Therefore the optimal PCR condition has changed to decrease the primers concentration into 0.5 μ M..

Detection from soil

PCR-amplified products of the variable regions of flagellin genes of *B. pseudomallei* NF10/38, NF10/38, NF47/38, E38(L58), and E271 and *B. thailandensis* E257 and E276 from three sources of soil suspensions (Fig 7.) could be observed in PCR reaction contained 10 μ M BSA (Fig.7.2) compared to a reaction without BSA (Fig. 7.1). The amplified products by PCR reaction including 10 μ M BSA showed a positive detection in all inoculated soils whereas a negative PCR-product were obtained in a reaction without BSA.

However, the new optimal condition with low concentration of primers was applied to detect the bacterial from soil, the reproducible result showed increasing in sensitivity of the detection with and without BSA (Fig. 8). Sensitivity of the detection in the optimal condition was performed as shown in Fig. 9. The result showed the same intensity detected from both species and could be detected to 1000 cfu per reaction. Figure 10 are the soil sample from difference sources.

Discussion

Detection of bacterial species using conventional microbiological techniques can be laborious and time-consuming. Polymerase chain reaction (PCR) is an alternative method for detection of microorganisms. The specificity, sensitivity, and rapidity of this technique are suitable for application and monitoring for the detection of small numbers of target microorganisms for which specific amplification primers are available [3]. A simple and rapid system to differentiate the two bacterial species by PCR-method was developed in this study. The DNA extract, by using a single colony suspended in sterile-distilled water and boiling for releasing DNA molecule, was used as template for PCR using the same primer pair specific to the conserved sequences of the variable regions of flagellin genes of both species. The detection by this technique was sensitive enough that could be detected approximately 20-80 cells / reaction.

The significant of arabinose utilization by *B. pseudomallei* and *B. thailandensis* is demonstrated to be the marker for pathogenesis. Our method has been compared to the arabinose utilization with blind test. The PCR results of 290 samples were corresponded to the method of arabinose utilization. The length of flagellin gene amplified product generated by PCR can be used to identify bacterial species [2]. The difference in product sizes between 191-bp of *B. pseudomallei* and 176-bp of *B. thailandensis* could be observed on 3.5% agarose gel. This method is rapid and simple. The difference in size of these amplified flagellin genes could be used as a DNA marker for detection of infectious species.

Although, there is no clearly evidence indicating the relationship of the two species, we have proposed that whether the both bacteria which is very similar in morphology, biochemistry and antigenicity should be survived and distributed in the same environment. In this experiment, we have designed the PCR amplification for differentiating the two species that has been mixed together by one reaction condition. Our study showed that the amplification product of mixed population of both bacteria

could be detected. Moreover, the mixed ratio of both bacteria gave significant difference in amplification product. When the template ratio of *B. pseudomallei* is twice times more than *B. thailandensis*, the amplification product gave same intensity. However, this PCR condition is preformed to mimic the environment situation. Thus, there is a possibility that the PCR method can be applied for investigation of the *B. pseudomallei* and *B. thailandensis* in natural environment. In addition, the sensitivity of a mimic mixed populations of the two species was shown a significant difference in which the amplified product intensity of *B. pseudomallei* was 4 times lower than *B. thailandensis* when using an equal number of template [3]. We therefore investigated a factor that involves in the sensitivity of the detection using Taq DNA polymerase with low fidelity and computer analysis of a stem loop formation program. We found a certain stem loop with a stable energy of -81.6 kcal/mole within a flagellin sequence of *B. pseudomallei* but not in *B. thailandensis*. Our finding clearly explain the unequally sensitivity for detecting *B. pseudomallei* and *B. thailandensis*.

The variable region of *B. pseudomallei* contained 70% GC-contents (GC=134 in 191 bp sequence) and *B. thailandensis* contained 71% GC-contents (GC=125 in 176 bp sequence) in their sequences [2]. To improve yield and specificity of target rich in GC content or ones that can form secondary structure in PCR amplification, enhancing agents were often included in the reaction [4, 5]. Addition of betaine in PCR reaction has been reported to improve the amplification of DNA containing high GC contents [4, 6]. Betaine is an isostabilizing agent, equalizes the contribution of GC- and AT-base pairing to the stability of the DNA duplex [4]. Betaine binds and stabilized AT based pairs whilst destabilises GC base pairing resulting in a net specific destabilisation of GC-rich region [7]. The use of betaine has demonstrated general benefits for a range of PCR amplification and should be considered to ensure the highest levels of quality and reproducibility [8]. In this study, we found that 1 M betaine was suitable to improve both levels of quality and reproducibility in the PCR reaction. In addition, decreasing the primers concentration to 0.5 μ M could also improved sensitivity of the detection to the same intensity of the both organisms.

The application of PCR-based method should be detect the so-called viable non-cultivable state of the bacteria to make the promising tool for environmental surveys. Thus, our attempt is to directly detect specific *B. pseudomallei* and *B. thailandensis* flagellin gene sequence from soil. Natural collected samples usually contain a number of different bacterial species and various unknown or known compounds. Depending on the source of sample, various known compounds can inhibit the PCR reaction, such as humic and fulvic acids [10, 11]. Humic substance, which can inhibit PCR reaction, is a mixture of complex polyphenolics produced during the decomposition of organic matter. They are ubiquitous in soil and water. Other factors involving a difficulty of DNA isolation are due to the presence of heavy metal ions or other DNA-damaging agent [10]. The PCR of soil samples could be improved to increase the reproducibility by addition of various proteins to the reaction. These proteins with a higher affinity for the inhibitors could relief inhibition of the PCR. Bovine serum albumin (BSA) is a widely used for reducing interference in the PCR. It had been reported that the inhibitors in soil were preventing amplification by binding to the polymerase or the target DNA. BSA may prevent binding of inhibitor to target DNA or Taq DNA polymerase [9, 10, 11]. Relief of

amplification inhibition in the PCR provided by BSA was tested in this study. Addition of 10 μ M BSA in the PCR reaction could overcome the inhibition and provide PCR-amplified products from soil suspension samples that could not be observed in PCR reaction without BSA in the high concentration of primers condition.

The limitation of PCR amplification of bacterial target DNA in soil suspension sample depends on several factors. The purity of the DNA from contaminants, the amount of template DNA added to the reaction mixture, the amount of background DNA present in the soil sample, and the condition of the extracted DNA can significantly affect the achievement. The variability in detection sensitivity is due to uncontrollable factors such as background DNA and inhibitory material that co-exist with target DNA. In this study, amplification of *B. pseudomallei* and *B. thailandensis* target DNA in soil suspension was found to decrease sensitivity compared with amplification of both species in the cultured broth. Sensitivity of amplification of *B. pseudomallei* and *B. thailandensis* target DNA in soil suspension were about 10^3 cells per reaction which were quite low. Low sensitivity of PCR method for detection of *Burkholderia* target gene might be caused by adsorption of bacterial cells in soil or by inhibitors in soil. In this study, we used a simple method, heating in boiling water bath, for released DNA without further extracted DNA before amplified by PCR so that all of inhibitors in soil might be high in the reaction and decrease sensitivity of detection. Moreover, we have also compared three different sources of soils resulting in differences in amplification-product. The results suggested that there are still many unknown factors may involved in the amplification detection. We have demonstrated here that using the specific flagellin primer, the 191 bp *B.pseudomallei* and 176 bp *B. thailandensis* amplification product, could be detected without enrichment and any extraction steps. Our study would be beneficial for epidemiological application in the real situation.

Acknowledgements

This work was supported by grants from the Thailand Research Fund and the National Center for Genetic Engineering and Biotechnology, the National Science and Technology Development Agency of Thailand.

References

1. Smith, M.D., Angus, B.J., Wuthiekanun, V., and White, N.J. (1997) Arabinose assimilation defines a nonvirulent biotype of *Burkholderia pseudomallei*. Infect. Immun. **65**: 4319-4321.
2. Wajanarogana, S., Sonthayanon, P., Wuthiekanun, V., Panyim, S., Simpson, A.J.H., and Tungpradabkul, S. (1999) Stable marker on flagellin gene sequences related to arabinose non-assimilating pathogenic *Burkholderia pseudomallei*. Microbiol. Immunol. **43**: 995-1001.
3. Alvarez AJ, Buttner MP, and Stetzenbach LD.(1995) PCR for bioaerosol monitoring: sensitivity and environmental interference. Appl Environ Microbiol. **61**: 3639-3644.
4. Varadaraj K, and Skinner DM. (1994) Denaturants or cosolvents improve the specificity of PCR amplification of a G+C-rich DNA using genetically engineered DNA polymerases. Gene **140**: 1-5.

5. Frakman S, Kobs G, Simpson D, and Storts D. (1994) Betaine and DMSO enhancing agents for PCR. *Promega notes* **65**: 27-29.
6. Henke W., Herdel K., Jung K., Schnorr D., and Loening SA. (1997) Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res.* **25**: 3957-3958.
7. Rees WA., Yager T.D., Krote J., Hippel PHV. (1993) Betaine can eliminate the base pair composition dependence of DNA melting. *Biochemistry* **32**: 137-141.
8. McDowell DG., Burns NA., Parks HC. (1998) Localised sequence regions possessing high melting temperatures prevent the amplification of a DNA mimic in competitive PCR. *Nucleic Acids Res.* **26**: 3340-3347.
9. McGregor DP., Forster S., and Steven J. (1996) Simultaneous detection of microorganisms in soil suspension based on PCR amplification of bacterial 16S rRNA fragments. *Biotechniques* **21**: 463-471.
10. Kreader CA. (1996) Relief of amplification in PCR with bovine serum albumin or T4 Gene32 protein. *Appl. Environ. Microbiol.* **62**: 1102-1106.
11. Brook MD., Currie B., and Desmarchelier PM. (1997) Isolation and identification of *Burkholderia pseudomallei* from soil using selective culture techniques and the polymerase chain reaction. *J Appl. Microbiol.* **82**: 589-596.

Table 1. Source of bacterial isolates tested for arabinose biochemical property comparing with PCR-based method

Source of bacterial isolates	Number of samples	Arabinose property			
		Biochemical		PCR	
		Ara-	Ara+	Ara-	Ara+

Clinical isolate from blood	63	63	-	63	-
sputum	4	4	-	4	-
urine	4	4	-	4	-
pleural fluid	2	2	-	2	-
pus	22	22	-	22	-
Chest wall	1	1	-	1	-
Synovial fluid	1	1	-	1	-
Ascitic fluid	1	1	-	1	-
Environment isolate (Thailand)	11	-	11	-	11
Central (non-endemic)					
Northeastern (endemic)	56	32	24	32	24
Laos	112	72	40	72	40
Vietnam	13	4	9	4	9

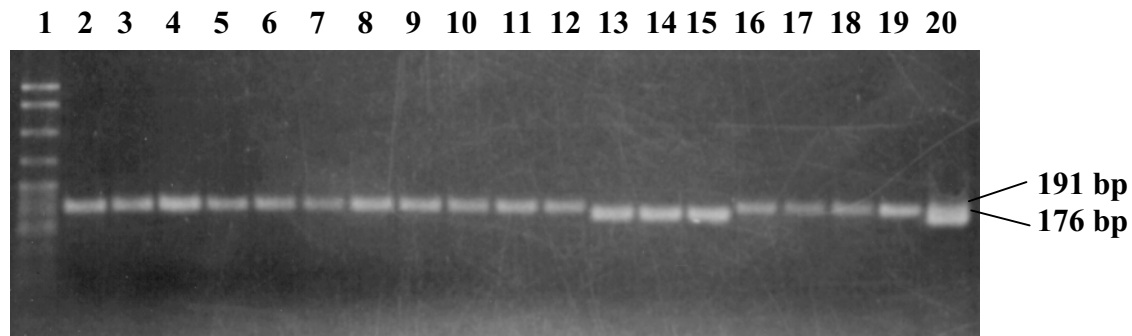


Fig. 1. The ethidium stained 3.5% agarose gel of amplification product of *B.pseudomallei* and *B. thailandensis* using specific flagellin gene primer. Lane 1 is a pBR 322 digested with MspI as standard marker. Lanes 2-20 are soil isolate samples

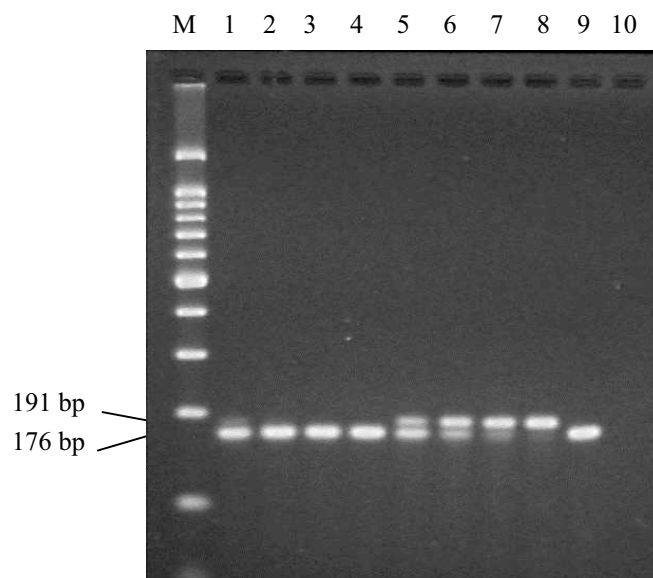


Fig. 2. PCR-amplified products of a mixture of *B. pseudomallei* and *B. thailandensis* (7.2×10^7 cells/ml and 7.3×10^7 cells/ml respectively) at various ratios. Lane M is a 100 bp ladder DNA marker. Lanes 1-4 are a mixture of *B. pseudomallei* and *B. thailandensis* at ratios of 1:1, 1:2, 1:4, and 1:8 respectively. Lanes 5-7 are a mixture of *B. pseudomallei* and *B. thailandensis* at ratios of 2:1, 4:1, and 8:1 respectively. Lanes 8 and 9 are positive control of PCR run, PCR reactions contain culture broth of E38(L58) and E276 at 1:100 dilution respectively. Lane 10 is a negative control of PCR run.

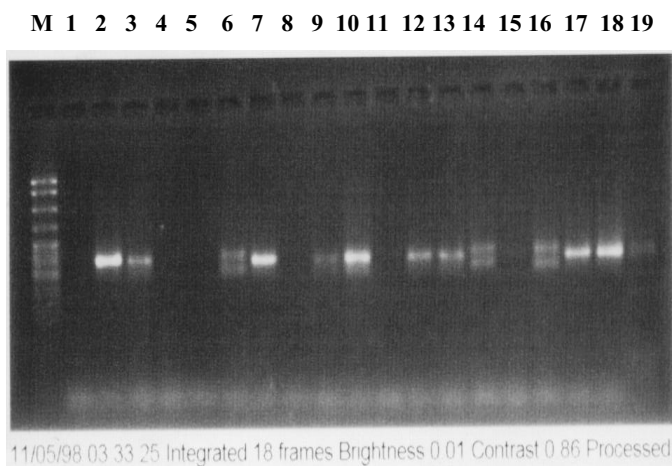


Fig. 3. Amplification products of *B. pseudomallei* and *B. thailandensis* from environmental and clinical isolates were analyzed at a dilution of 1:1000. Lanes 1, 4, 5, 8, 11 and 15 are *E. coli*. Lanes 2, 3, 7, 9, 10, 12, 13, 17 and 18 are *B. thailandensis*. Lanes 6, 14 and 16 are *B. pseudomallei* from environmental isolates and clinical isolate, respectively. Lane M is a 100 bp ladder DNA marker

PRIMER

191-bp CTG TCG TCG ACG GCC GTG ACC GCC GTG TTC GGC TCG TCG ACC GCC 45

167-bp CTG TCG TCG ACG GCC GTG ACC GCC GTG TTC GGC TCG TCG ACC GCC 45

15-bp deletion in *B. thailandensis*

191-bp GGC ACG GGC ACG GCG GCC TGG CCG TCG TTC CAG ACG CTG GCG CTG 90

167-bp GGC ACG GGC ACG GCG GCC TGG CCG TCG TTC CAG ACG CTG GCG CTG 90

191-bp TCG ACT TCG GCA ACC AGC GCG CTG TCC GCG ACC GAC CAG GCG AAC 135

167-bp TCG ACT --- --- --- --- --- --- --- --- ACC GAC CAG GCG AAC 111

24-bp deletion in a short PCR product

PRIMER

191-bp GCC ACG GCG ATG GTT GCG CAG ATC AAC GCG GTC AAC AAG CCG CAA 180

167-bp GCC ACG GCG ATG GTT GCG CAG ATC AAC GCG GTC AAC AAG CCG CAA 156

191-bp ACG GTC TCG AA 191

167-bp ACG GTC TCG AA 167

Fig. 4. Sequence analysis and sequence alignment of 167-bp a short PCR product of *B. pseudomallei* template generated by using low fidelity Taq DNA polymerase with 191-bp a fragment of flagellin gene sequence from *B. pseudomallei* template. The underlines are primer sequences, a box is a 15-bp deletion sequence found in *B. thailandensis*, and dashes are 24-bp deletion sequence found in a short PCR product.

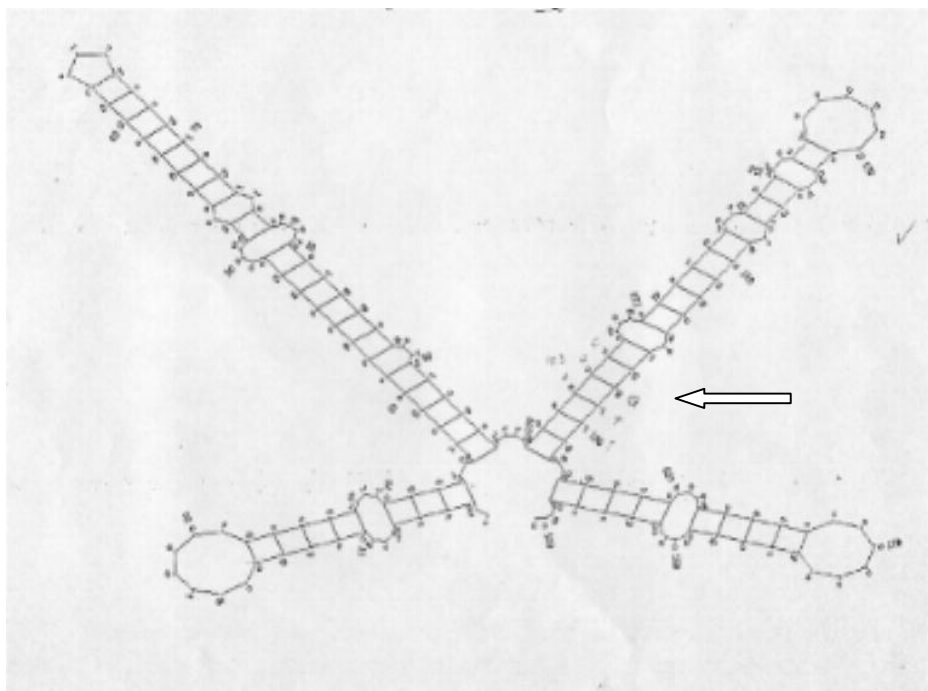


Fig. 5. A secondary structure of the 191 bp-flagellin gene sequence of *B. pseudomallei* having stabilization energy of -81.6 kcal/mole by MFOLD from GCG Squiggle plot program. An arrow is indicated a 24-bp deletion region.

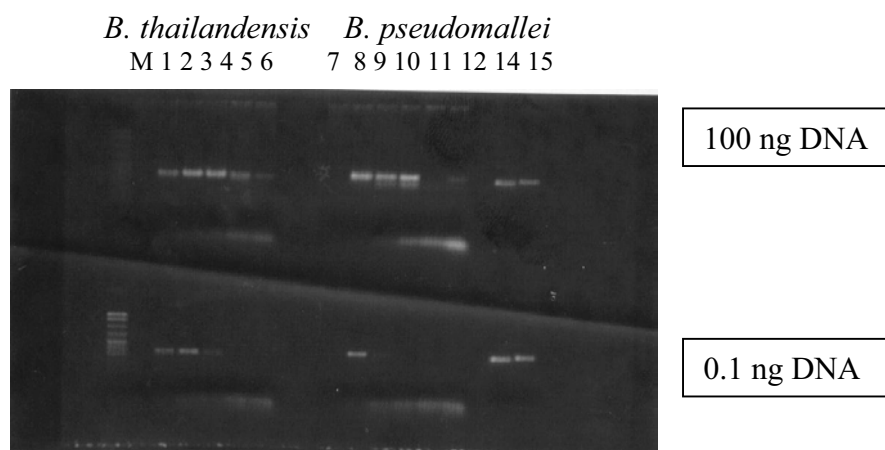


Fig. 6. Amplification products of *B. thailandensis* with DNA template 100 ng (upper) and 0.1 ng template (lower) in lanes 1, 2, 3, 4, 5 and 6 are various primers concentration of 0.1, 0.5, 1, 1.5, 2 and 3 μ M, respectively. Amplification product *B. pseudomallei* with DNA template 100 ng (upper) and 0.1 ng template (lower) in lanes 7, 8, 9, 10, 11 and 12, are various primers concentration as above. Lane M is a 100 bp ladder DNA marker, lanes 14 is Ara⁺ positive control and lane 15 is Ara⁻ positive control.