รายงานวิจัยฉบับสมบูรณ์

โครงการ

ผลิตงานวิจัยและส่งเสริมกลุ่มวิจัยด้านอณูชีววิทยา

โดย

ศาสตราจารย์ สกล พันธุ์ยิ้ม และคณะ

19 ธันวาคม 2542

สัญญาเลขที่ RTA / 01 / 2539

รายงานวิจัยฉบับสมบูรณ์

โครงการผลิตงานวิจัยและส่งเสริมกลุ่มวิจัยด้านอณูชีววิทยา

คณะผู้วิจัย	สังกัด
1. นายสกล พันธุ์ยิ้ม	ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล
2. นายชนันท์ อังศุธนสมบัติ	สถาบันอณูชีววิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล
3. นางสาววิภา จึงจตุพรชัย	สถาบันอณูชีววิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล
4. นางสาวสุณี เกิดบัณฑิต	สถาบันอณูชีววิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล
5. นายบุรชัย สนธยานนท์	สถาบันอณูชีววิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล
6. นายอภินันท์ อุดมกิจ	สถาบันอณูชีววิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล
7. นางสาวจรรยาภรณ์ ฟู่เจริญ	สถาบันอณูชีววิทยาและพันธุศาสตร์ มหาวิทยาลัยมหิดล
8. นายวิเชษฐ ลีลามานิตย์	ภาควิชาชีวเคมี คณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล
9. นางสาวมธุรส พงษ์ลิขิตมงคล	ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล
10. นางวิไล หนุนภักดี	ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล
11. นางสาวสุมาลี ตั้งประดับกุล	ภาควิชาชีวเคมี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย

ชุดโครงการเมธีวิจัยอาวุโส สกว. สกล พันธุ์ยิ้ม

สารบัญ

บทคัดย่อ

บทคัดย่อ

อณูชีววิทยาเป็นวิทยาการผสมผสานความรู้และหลักการทางเคมี ฟิสิกส์ และชีววิทยา เพื่อให้เกิดความรู้ความเข้าใจปรากฏการณ์ที่ระดับอณูของสิ่งมีชีวิต อันเนื่องมาจากโครงสร้างของสารพันธุกรรม องค์ความรู้ทางอณูชีววิทยาเป็นรากฐานสำคัญของเทคโนโลยีชีวภาพ ซึ่งส่งผลในแนวกว้างทางด้านการแพทย์ เกษตรกรรม อุตสาหกรรมและสิ่งแวดล้อม การวิจัยและพัฒนาวิชาการด้านอณูชีววิทยาในประเทศไทยยังไม่ก้าวหน้าเท่าที่ควรเพราะการขาดแคลน ผู้เชี่ยวชาญ และบุคลากร จึงจำเป็นต้องดำเนินการเร่งการพัฒนางานวิจัย สร้างกลุ่มวิจัย และผลิตบุคลากรวิจัยด้านอณูชีววิทยาขึ้นภายในประเทศ

โครงการผลิตงานวิจัยและส่งเสริมกลุ่มวิจัยด้านอณูชีววิทยา ประกอบด้วยคณะผู้ร่วมวิจัย 35 คน ได้ดำเนินการเป็นระยะเวลา 3 ปี โดยสนับสนุนและฝึกนักวิจัยด้านอณูชีววิทยาให้มีความเชี่ยวชาญ 11 คน ส่งผลให้มีการผลิตผลงานวิจัยเป็นที่ยอมรับในระดับสากลโดยการตีพิมพ์ในวารสารวิชาการระดับนานาชาติ จำนวน 16 เรื่อง และได้ผลิตนักศึกษาระดับปริญญาโท-เอก ซึ่งจะเป็นนักวิจัยในอนาคต จนจบการศึกษา จำนวน 12 คน และกำลังศึกษาอยู่ 6 คน ได้นำเสนอผลงานวิจัยในการประชุมวิชาการนานาชาติ 10 ครั้ง ในการประชุมวิชาการภายในประเทศ 12 ครั้ง และจัดประชุมวิชาการอณูชีววิทยาประจำปี 3 ครั้ง

ผลงานวิจัยโดยสังเขปได้องค์ความรู้ใหม่ด้านกลไกการฆ่าลูกน้ำยุงของโปรตีนCry4B ทราบโครงสร้าง αhelix ใน pore-forming domain มี 7 สาย โดยสาย helix 4 และ helix 5 เป็นสายที่ทะลุผ่านผนังเซลล์ทำให้ลูกน้ำยุงตาย และได้ทำการสร้างแบคทีเรียสายพันธุ์ที่พบในกระเพาะลูกน้ำยุงให้ฆ่าลูกน้ำยุงได้ดี ได้ศึกษาDNAควบคุมการแสดงออกของยีน(promoter)ในcyanobacteria พบpromoter หลายชุด และ 1 ชุดมีลำดับเบสเป็น t.RNA^{pro} ได้แยกพร้อมศึกษาการเรียงลำดับเบสของยีน flagellin ใน *B.pseudomallei* และสามารถพัฒนาการจำแนก *B.pseudomallei* และ *B.cepacia* ที่พบในคนไข้โดยวิธี PCR-RFLP ได้พัฒนาเทคนิคPCR-RFLPในการจำแนก *M.tuberculosis* complex ออกจาก mycobacteria และ pathoganic bacteria 24 ชนิด ได้แยกยีน catalase ของ lactobacillus และเปลี่ยน lactobacillus ที่พบในแหนมให้ผลิต catalase ซึ่งสลาย H_2O_2 ได้ดี ได้ศึกษาเชื้อไวรัสจุดวงแหวนมะละกอ (PRV) จนทราบลำดับเบสของ coat protein ซึ่งมีความแตกต่างกันใน PRV จากราชบุรี ชุมพร ขอนแก่น และจากฮาวาย ไต้หวัน ออสเตรเลีย และได้สร้างมะละกอข้ามสายพันธุ์ป้องกันการติดเชื้อPRV ได้ศึกษายีนที่แสดงออกในเซลล์กล้ามเนื้อและเม็ดเลือดกุ้งกุลาดำ และแยกยีนฮอร์โมนจากเซลล์ก้านตากุ้งซึ่งอาจควบคุมการลอกคราบหรือระดับน้ำตาลหรือการพัฒนาเซลล์สืบพันธุ์

เนื้อหางานวิจัย

เนื้อหางานวิจัย

SINGLE PROLINE SUBSTITUTIONS OF SELECTED HELICES OF THE *BACILLUS THURINGIENSIS* CRY4B TOXIN AFFECT INCLUSION SOLUBILITY AND LARVICIDAL ACTIVITY

ABSTRACT

PCR-based mutagenesis was employed to investigate the role in toxicity of putative transmembrane helices of the 130-kDa Cry4B mosquito-larvicidal delta-endotoxin produced by *Bacillus thuringiensis* subsp. *israelensis*. Mutant toxins with a single proline substitution in the central region of α5, α6 and α7 were constructed and expressed in *Escherichia coli* as cytoplasmic inclusion bodies with a yield similar to that of the wild-type toxin. Unlike inclusions of the wild-type and that of the previous mutants for proline replacements in α3 or α4, all proline-substituted inclusions were insoluble in carbonate buffer, pH 9.0, indicating that the bend introduced in these three helices possibly interferes with the protein inclusion packing as shown by the insolubility. Similar to the previous substitution in α4, an almost complete loss of toxicity against *Aedes aegypti* mosquito-larvae was demonstrated for *E. coli* cells expressing mutant toxins in which residues Val-181, Ala-182 or Leu-186 in α5 were changed to proline. In addition, a dramatic decrease in larvicidal activity was observed for the substitution at Thr-254 in α7 while the mutation Q215P in α6 did not affect the biological activity. These results suggest that the central helix (α5) and conceivably α7, but not α6, are important determinants of toxin function. Our data therefore further support the notion that the putative pore forming helical hairpin α4-α5 together with α7 plays a crucial role in Cry toxin activity.

INTRODUCTION

Bacillus thuringiensis (*Bt*) is a Gram-positive entomopathogenic bacterium that has been used successfully as an alternative insecticide for biological control of disease vectors and other pests. During sporulation, different *Bt* strains produce larvicidal proteins (classified as Cry and Cyt delta-endotoxins) in large quantities as cytoplasmic crystalline inclusions that are specifically toxic to a variety of dipteran, lepidopteran and coleopteran insect larvae [1,2]. When ingested by susceptible larvae, the 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 midget 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 (see [3] for reviews). However, the insecticidal mechanism at the molecular level of these *Bt* toxins is still not entirely established.

The X-ray crystal structure of two different Cry toxins, the coleopteran-specific Cry3A toxin [4] and the lepidopteran-specific Cry1Aa toxin [5], reveals Cry proteins consisting of three distinct domains, and it is believed that each domain has a defined function including pore formation and receptor recognition [4,5]. Domain I is a group of seven α-helices in which the central helix (α5) is relatively hydrophobic and encircled by six other amphipathic helices. Domain II is the most variable part of the Cry toxin family and is composed of three anti-parallel β-sheets, each terminating in a surface-exposed loop. Domain III is a tightly packed β-sandwich of two anti-parallel sheets. It has been proposed that other members of this family will have the same overall tertiary structure since the core of the molecule including all the domain interfaces is built up from five amino acid sequence segments that are highly conserved throughout the entire Cry toxin family [1,4].

Structurally, it is immediately apparent that domain I is likely to be the transmembrane pore-forming apparatus. This domain contains five amphipathic helices (α3, α4, α5, α6 and α7) that are theoretically long enough to span the bilayer lipid membrane and form a lytic pore [4,5]. 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 [1], and by analogy with the helical bundle pore-forming structures of two other well-characterized bacterial toxins, colicin A and diphtheria toxin, although they bear no sequence homology [6]. This notion is further supported by several studies with truncated proteins corresponding to domain I of Cry1Ac [7] and Cry3B2 [8] and with synthetic peptides of selected helices from Cry1Ac [9] or Cry3A [10] that demonstrated pore-forming activity either in phospholipid membrane vesicles or within planar lipid bilayers. A number of experiments *via* site-directed mutagenesis suggested that α4 and/or α5 of Cry1Aa [11,12] and of Cry1Ac [13] and the loop between the bottoms of α4 and α5 of Cry1Ab [14] are involved in pore formation. Recently, we have employed single proline substitutions and demonstrated that α4, but not α3, of the dipteran-specific Cry4B toxin plays a crucial role in larvicidal activity, possibly in the pore forming step rather than in receptor binding [15]. In this report, we have applied the same mutagenesis strategy to further investigate a possible involvement in toxicity of three other putative-transmembrane helices (α5, α6 and α7) of

this mosquito-active toxin, and found that these three helices are essentially involved in inclusion solubility. Our results also indicate that α5 and α7, but not α6, play a role in larvicidal activity of the Cry4B toxin.

MATERIALS AND METHODS

Construction and Expression of Mutant Toxins

Single proline substitutions *via in vitro* site-directed mutagenesis were performed using a Quickchange PCR-based mutagenesis kit (Stratagene) following the manufacturer's instructions. The plasmid pMU388, constructed by cloning the full-length *cry4B* toxin gene from *Bt* subsp. *israelensis* in a pUC12 vector [16], was used as a template. Mutagenic primers were purchased from Bio-synthesis Inc. (USA) as shown in Table I. All mutations were verified by DNA sequencing using an ABI prism 377 sequencer.

Partial Purification and Solubilisation of Protoxin Inclusions

The wild type and mutant Cry4B toxin genes were expressed in *E.coli* strain JM109 under control of the *lacZ* promoter. Cells were grown in LB medium plus 100 µg/ml of ampicillin until OD_{600} reached 0.4-0.5 and incubation was continued for another 4 hours after addition of isopropyl-β-D-thiogalacto pyranoside (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 minutes and pellets obtained were washed 3 times in distilled water. Protein concentrations were determined by using a protein microassay (Bio-Rad), with bovine serum albumin fraction V (Sigma) as a standard.

Protoxin inclusions (1 mg/ml) were solubilised in 50 mM Na_2CO_3, pH 9.0 and incubated at 37 °C for 60 minutes as described previously [15]. After centrifugation for 10 minutes, the supernatants were analysed by SDS-15% (w/v) polyacrylamide gel electrophoresis (PAGE) in comparison with the inclusion suspension.

Larvicidal Activity Assays

Bioassays were performed as described previously [15] with some modification using 2-day old *Aedes aegypti* mosquito-larvae reared in a container (22x30x10 cm deep) with approximately 3 litres of distilled water supplemented with a small amount of rat diet pellets. Both rearing and bioassays were done at room temperature (25 °C). The assays were carried out in 1 ml of *E. coli* suspension (10^8 cells suspended in distilled water) in 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* samples. Mortality was recorded after 24-hour incubation period.

RESULTS AND DISCUSSION

Recently, data on molecular determinants of membrane insertion and pore formation for *Bt* Cry toxins has increased substantially. Where studied, the helical domain of different Cry toxins has now been putatively demonstrated to be involved in membrane integration and toxin pore formation [7-15]. In the present study, we have made use of PCR-based mutagenesis to construct several more mutants in the putative pore-forming domain to determine which of the three other helices (α5, α6 or α7) would be responsible for toxicity of the mosquito-active Cry4B toxin. Proline substituted mutants were designed based on a 3D Cry4B model which was previously constructed by homology modelling using Cry3A as a template [17]. The designed mutants have single proline substitutions in the central region of the selected helices of Cry4B including α5 at residues Val-181, Ala-182 and Leu-186, α6 at Gln-215 and α7 at Thr-254 (Fig.1). Each mutant gene could be over-expressed in *E. coli* under inducible control of the *lacZ* promoter by addition of IPTG to mid-exponential phase cultures and the mutant toxins were predominantly produced as sedimentable inclusion bodies which were then partially purified. In addition, the level of protein expression of all mutant proteins was approximately the same as that of the wild type, and the expressed mutant derivatives still cross-reacted with Cry4B antibodies (data not shown).

Experiments were carried out to assess the solubility of mutant protein inclusions in carbonate buffer, pH 9.0 in comparison with that of the wild-type inclusion. The amounts of the 130-kDa soluble proteins in the supernatant were compared with those of the proteins initially used in order to determine the percentage of protein solubilisation. A complete loss of solubility at alkaline pH was observed for the inclusions of the mutants substituted in α5, α6 or α7, whilst the inclusions of the two previously constructed mutants, V118P and Q149P, exhibited the same solubility characteristics as that of the wild-type (Fig. 2). The reason for the difference in solubility between those two sets of point mutations is not clear. However, this may be explained by the fact that all the five mutated residues, *i.e.* Val-181, Ala-182 and Leu-186 in α5, Gln-215 in α6 and Thr-254 in α7 are buried in the protein core, unlike Val-118 in α3 and Gln-149 in α4 that are exposed at the protein surface (see Fig. 1C). Thus, the bend created by the proline substitution in these relatively conserved helices, *i.e.* α5, α6 or α7 possibly disturbs the structural characteristics either locally or globally that might consequently affect inclusion formation as demonstrated by a drastically

perturbed dissolvability. On the other hand, the proline mutations in α3 or α4 appear to affect only the individual helices, without significantly influencing the other helices or the overall structure, as earlier demonstrated for an α-lactalbumin folding intermediate [28].

To determine whether a single praline replacement in these three additional helices also affects the larvicidal activity of Cry4B, *E. coli* cells expressing each type of the mutant toxins were bioassayed using *Aedes aegypti* mosquito-larvae (Fig. 3). The mortality data recorded after a 24-hour incubation reveals that the α6 mutant (Q215P) still exhibited full larvicidal activity (98.7±0.7%) similar to the α3 mutant (V118P), whereas the T254P mutant produced merely 23.7±13.9%. However, mortality of all the α5 mutants, V181P, A182P and L186P, was almost totally abolished as approximately comparable to that observed with the previously described α4 mutant (Q149P). These results suggest that the integrity of α5 and α7 may indeed be important for toxin activity like that of α4.

Although toxin inclusion insolubility and larvicidal activity are seemingly correlated for all the three α5 mutants and the α7 mutant, the insolubility *in vitro* may not always necessarily reflect toxin activity *in vivo* as observed for the α6 mutant which was insoluble at pH 9.0 but still bioactive. It has been previously demonstrated that the difference detected in solubilisation *in vitro* for Cry4A inclusions which were purified from two different *Bt* recipient strains, is not a factor in larval toxicity [19]. Presumably, larval midgut proteases *in vivo* might facilitate the dissolution of the ingested toxin inclusions which would negate the differences between the observed toxicities of the α3 and α6 mutants. It was shown that incubation of the Cry2A toxin with gut proteases enhanced the solubility of the toxin inclusion, obviating the requirement for reducing agents at pH10.5 [20].

In conclusion, this report additionally demonstrates that the central helix (α5), α6 and α7, but not α3 or α4, are conceivably involved in protein packing for inclusion formation, thus destabilising these three helices individually could gives rise tc toxin insolubility *in vitro*. Moreover, our study also provides further support for a crucial role of the pore-forming helical hairpin α4-α5 as well as α7, which has been suggested to be a domain binding sensor [10], in larvicidal activity of the Cry4B toxin.

REFERENCES

1. Hofte H, Whiteley HR. Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiol. Rev. 1989; 53: 242-255.

2. Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean DH. Microbiol. *Bacillus thuringiensis* and its pesticidal crystal proteins. Mol. Biol. Rev. 1998; 62: 775-806.

3. Knowles BH. Mechanism of action of *Bacillus thuringiensis* insecticidal δ-endotoxins. Adv. Insect Physiol. 1994; 24: 275-308.

4. Li J, Carroll J, Ellar DJ. Crystal structure of insecticidal delta-endotoxin *from Bacillus thuringiensis* at 2.5 Å resolution. Nature 1991; 353: 815-821.

5. Grochulski P, Masson L, Borisova S, Pusztai-Carey M, Schwartz JL, Brousseau R, Cygler M. *Bacillus thuringiensis* CryIA(a) insecticidal toxin: crystal structure and channel formation. J. Mol. Biol. 1995; 254: 447-464.

6. Parker MW, Pattus F. Rendering a membrane protein soluble in water: a common packing motif in bacterial protein toxins. Trends Biochem. Sci. 1993; 18: 391-5.

7. Walters FS, Slatin SL, Kulesza CA, English LH. Ion channel activity of N-terminal fragments from CryIA(c) delta-endotoxin. Biochem. Biophys. Res. Commun. 1993; 196: 921-926.

8. Von Tersch MA, Slatin SL, Kulesza CA, English LH. Membrane-permeabilizing activities of *Bacillus thuringiensis* coleopteran-active toxin CryIIIB2 and CryIIIB2 domain I peptide. Appl. Environ. Microbiol. 1994; 60: 3711-3717.

9. Cummings CE, Ellar DJ. Structural and functional studies of a synthetic peptide mimicking a proposed membrane inserting region of a *Bacillus thuringiensis* delta-endotoxin. Microbiology 1994; 140: 2737-2747.

10. Gazit E, La Rocca P, Sansom MSP, Shai. Y. The structure and organization within the membrane of the helices composing the pore-forming domain of *Bacillus thuringiensis* delta-endotoxin are consistent with an "umbrella-like" structure of the pore. Proc. Natl. Acad. Sci. USA 1998; 95: 12289-12294.

11. Schwartz JL, Juteau M, Grochulski P, Cygler M, Prefontaine G, Brousseau R, Masson L. Restriction of intramolecular movements within the Cry1Aa toxin molecule of *Bacillus thuringiensis* through disulfide bond engineering. FEBS Lett. 1997; 410: 397-402.

12. Ahmad W, Ellar DJ. Directed mutagenesis of selected regions of a *Bacillus thuringiensis* entomocidal protein. FEMS Microbiol. Lett. 1990; 56: 97-104.

13. Kumar AS, Aronson AI. Analysis of mutations in the pore-forming region essential for insecticidal activity of a *Bacillus thuringiensis* delta-endotoxin. J. Bacteriol. 1999; 181: 6103-6107.

14. Chen XJ, Curtiss A, Alcantara E, Dean DH. Mutations in domain I of *Bacillus thuringiensis* delta-endotoxin CrylAb reduce the irreversible binding of toxin to *Manduca sexta* brush border membrane vesicles. J. Biol. Chem. 1995; 270: 6412-6419.

15. Uawithya P, Tuntitippawan T, Katzenmeier G, Panyim S, Angsuthanasombat C. Biochem. Effects on larvicidal activity of single proline substitutions in α3 or α4 of the *Bacillus thuringiensis* Cry4B toxin. Mol. Biol. Inter. 1998; 44: 825-832.

16. Angsuthanasombat C, Chungjatupornchai W, Kertbundit S, Luxananil P, Settasatian C, Wilairat P, Panyim S. Cloning and expression of 130-kd mosquito-larvicidal delta-endotoxin gene of *Bacillus thuringiensis* var. *israelensis* in *Escherichia coli*. Mol. Gen. Genet. 1987; 208: 384-389.

17. Uawithya P, Krittanai C, Katzenmeier G, Leetachewa S, Panyim S, Angsuthanasombat C. 3D models for *Bacillus thuringiensis* Cry4A and Cry4B insecticidal proteins based on homology modelling. Abst 32nd Meeting, Society for Invertebrate Pathology, Irvine, USA. 1999; 75.

18. Schulman BA, Kim PS. Proline scanning mutagenesis of a molten globule reveals non-cooperative formation of a protein's overall topology. Nat. Struct. Biol. 1996 Aug;3 (8):682-7. Nature Struct. Biol. 1996; 3: 682-687.

19. Angsuthanasombat C, Crickmore N, Ellar DJ. Comparison of *Bacillus thuringiensis* subsp. *israelensis* CrylVA and CrylVB cloned toxins reveals synergism *in vivo*. FEMS Microbiol. Lett. 1992; 94: 63-68.

20. Nicholls CN, Ahmad W, Ellar DJ. J. Bacteriol. Evidence for two different types of insecticidal P2 toxins with dual specificity in *Bacillus thuringiensis* subspecies. J. Bacteriol. 1989; 171: 5141-5147.

A)

α3 α4

Cry3A 123 DYAKNKALAELQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFA-ISGY

Cry4B 106 AYVRTDANAKMT**V**VKDYLDQYTTKFNTWKREPN---NQSYRTAVIT**Q**FNLTSAKLRETAVYFSNLVGY

118 149

α5 α6

Cry3A 190 EVLFLTTYAQAANTHLFLLKDAQIYGEEWGYEKEDIAEFYKRQLKLTQEYTDHCVKWYNVGLDKLRGS

Cry4B 171 ELLLLPIYAQ**VA**NFN**L**LLIRDGLINAQEWSLARSAGDQLYNTMV**Q**YTKEYIAHSITWYNKGLDVLRNK

181 186 215

182

α7

Cry3A 258 SYESWVNFNRYRREMTLTVLDLIALFPLYDVRL

Cry4B 239 SNGQWITFNDYKREM**T**IQVLDILALFASYDPRR

254

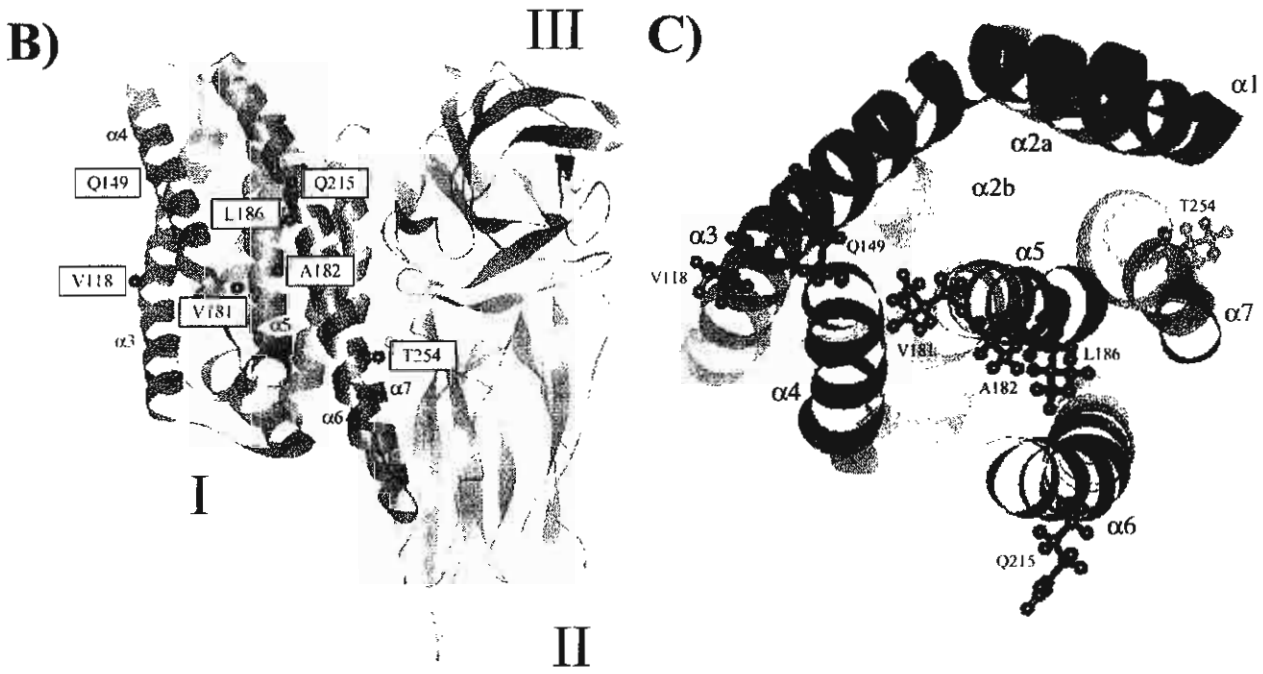

FIGURE 1 **A)** Amino acid sequence alignment of Cry4B with the known secondary structure elements of Cry3A. Single proline substitutions at each indicated position are underlined and emboldened. The corresponding helices are shown above the sequences. **B)** A ribbon representation (generated by Weblab viewer, Molecular Simulation Inc.) of the Cry4B model built by homology modelling [17], illustrating the three-domain organisation (I-III). Black beads indicate β-carbon atoms of the mutated residues. **C)** Top view of the helical bundle in domain I of Cry4B. The mutated residues are shown as ball and stick: α3 at Val-118, α4 at Gln-149, α5 at Val-181, Ala-182, and Leu-186, α6 at Gln-215 and α7 at Thr-254.

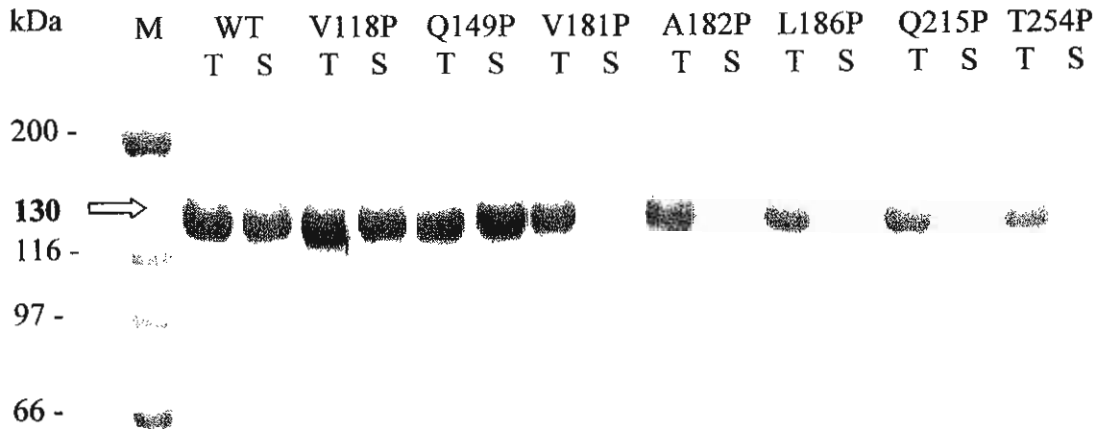

FIGURE 2 The figure shows a Coomasie brilliant blue-stained SDS-15% polyacrylamide gel of the partially purified 130-kDa protein inclusions extracted from *E. coli* expressing the wild-type (WT) and mutant Cry4B toxins and solubilised in carbonate buffer. (T) and (S) represent the total fractions and an equivalent volume of the supernatants after centrifugation, respectively. (M) represents the molecular mass standards.

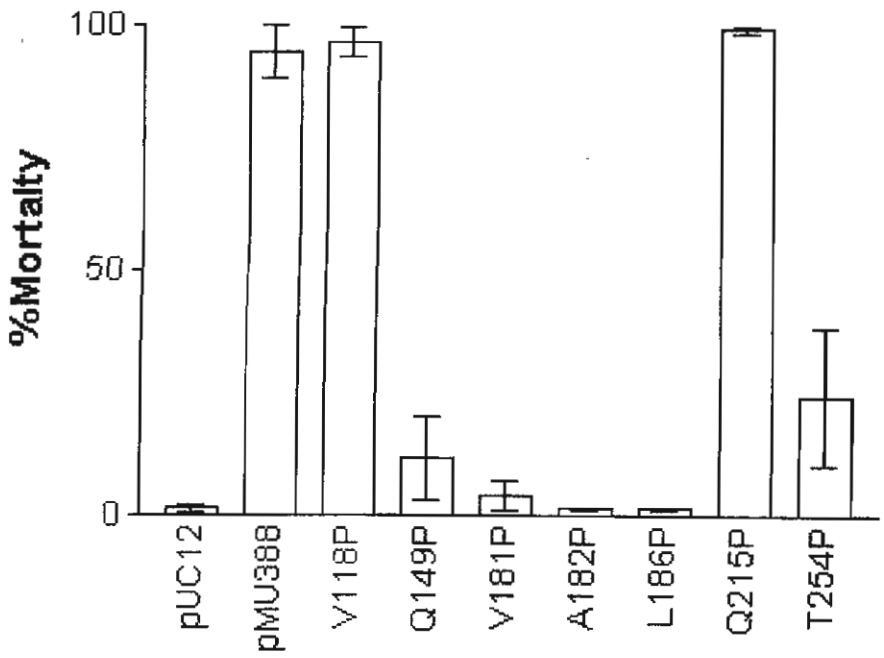

FIGURE 3 Lavicidal activities of *E. coli* cells expressing either the Cry4B wild-type toxin (pMU388) and its mutants (V118P, Q149P, V181P, A182P, L186P, Q215P and T254P) against *A. aegypti* larvae. Error bars indicate standard error of the mean from three independent experiments.

Table I Oligonucleotide primers designed to substitute a coded residue with proline

Primer	Location	Sequence[a]	Restriction Site
V181P-f	α5	P I Y A Q **P** A N F N L CCAATATACGCACAA**CCGG**CAAATTTCAATTTA	*Hpa*II
V181P-r		C GTAAATTGAAATTTG**CCGG**TTGTGCGTATATTG G	
A182P-f	α5	P I Y A Q V **P** N F N L CCAATATACGCACA**GGTCC**CAAATTTCAATTTA	*Sau*96I
A182P-r		C GTAAATTGAAATTTG**GGACC**TGTGCGTATATTG G	
L186P-f	α5	A N F N **P** L L I R D G L GCAAATTTCAAT**CC**ACTTTTAAT**CCGG**GATGGC	*Hpa*II
L186P-r		CTC GAGGCCATC**CCGG**ATTAAAAGT**GG**ATTGAAATT TGC	
Q215P-f	α6	T M V **P** Y T K E Y I CACTATGGTGCC**GTATAC**TAAAGAATATATTGC	*Acc*I
Q215P-r		GCAATATATTCTTTA**GTATAC**GGCACCATAGTG	
T254P-f	α7	D T K R E M **P** I Q V L GATTATAAAAGAGAGATGCC**GATTC**AAGTATTA	*Hinf*I
T254P-r		G CTAATACTT**GAATC**GGCATCTCTCTTTTATAATC	

[a]Introduced restriction enzyme recognition sites are underlined. Mutated nucleotide residues are shown as boldface. The deduced amino acid sequence is shown above each oligonucleotide sequence.

เนื้อหางานวิจัย

SELECTION OF STRONG PROMOTERS IN CYANOBACTERIUM *SYNECHOCOCCUS* PCC7942

ABSTRACT

Promoter-active fragments of *Synechococcus* PCC7942 were isolated by transcriptional gene fusion to the promoterless β-glucuronidase (GUS) gene of *E. coli*, which was used as a reporter gene. Several of the isolated promoter-active fragments expressed GUS activity in *Synechococcus* comparable with that of the λP_R promoter. Only 10% of the isolated promoter-active fragments also functioned in *E. coli*. The transcription initiation sites of the promoter-active fragment E3 were identified. The major transcription initiation site of E3 in *Synechococcus* was located within the nucleotides TTG which was identical to that corresponding in *E. coli*. Immediately upstream of the E3 transcription initiation sites was $tRNA^{pro}$ (GGG) gene, which contained two regions exhibiting strong homology to the major promoter elements in eukaryotic tRNA genes but did not contain the *E. coli* promoter element. Thus, the $tRNA^{pro}$ gene can act as a promoter. In addition, the promoter-active fragments E10 and E14, which contain the light-intensity responsive promoter and distant enhancer respectively, are under investigation.

INTRODUCTION

Cyanobacteria are the simplest organisms capable of performing oxygenic photosynthesis, with a thylakoid apparatus remarkably similar to that of higher plant chloroplasts [11]. They are model organisms for the cloning of genes involved in photosynthesis. Cyanobacteria have been used as hosts to express several heterologous genes. For example, attemps have been made to express the mosquitocidal protein genes of *Bacillus sphaericus* and *Bacillus thuringiensis* subsp. *israelensis* in order to provide an alternative biological insecticide for control of mosquito populations [2, 4, 23]. They have also been used in the expression of salmon growth hormone gene in order to produce a feed additive for fish [12], or the expression of ethylene-forming enzyme in order to exploit atmospheric CO_2 as a substrate [8]. However, the level of heterologous gene expression in cyanobacteria is very low when compared with that in *E. coli*. A possible way to improve the gene expression is to use an endogenous strong promoter. However, current knowledge of the structure and function of promoters recognized within cyanobacteria is still limited. Although cyanobacteria are classified as eubacteria [10], their RNA polymerase holoenzyme

is unique in that it contains a subunit γ, in addition to the $\alpha_2\beta\beta'\sigma$ structure common to the RNA polymerase of other eubacteria [19]. The genetics of cyanobacteria is highly heterogeneous, the GC content of their DNA ranges from 35% to 71% which is as varied as the whole kingdom of bacteria [17].

In order to isolate cyanobacterial promoters for structure and functional studies, Chungjatupornchai et al. [6] used the unicellular *Synechococcus* PCC7942 (previously referred to as *Anacystis nidulans* R2, Pasteur Culture Collection no. 7942) as a model system. Promoter-active fragments of *Synechococcus* PCC7942 were isolated by transcriptional gene fusion to the promoterless β-glucuronidase (GUS) gene of *E. coli*, which was used as a reporter gene. From a total of 640 transformants screened, 15 clones (2.3%) and 25 clones (3.9%) of the transformants expressed high and low GUS activity respectively. Out of these 40 clones, only 4 clones (10%) could express in both *Synechococcus* and *E. coli*. The majority of the isolated promoters could not function in *E. coli*, thus it indicated that their functional sequences were different from those of *E. coli*.

In this study, the three different isolated strong promoter-active fragments, E3, E10 and E14, were characterized.

MATERIALS AND METHODS

Organisms and culture conditions

Synechococcus PCC7942 strain R2-SPc, cured of the small plasmid pUH24 [13], was grown in liquid or on solid (1% Difco Bacto Agar) BG-11 medium as described previously [24]. *E. coli* strain MC1061 [3] was grown in LB broth or agar as previously described [18].

Construction of plasmids

A promoter-probe shuttle vector pKG, capable of replication in *Synechococcus* PCC7942 and in *E. coli*, was generated previously [6]. The pKG plasmid harbored the BamHI site upstream of the promoterless-GUS gene. Plasmid pKG-E3, pKG-E10 and pKG-E14 were the pKG plasmid harboring the chromosomal DNA of *Synechococcus* at the BamHI site.

β-glucuronidase and protein assay

For the GUS assay of *Synechococcus* and *E. coli*, the harvested cells cultures were suspended in buffer (50 mM $NaPO_4$ pH7.0, 10mM EDTA, 0.1% Triton X-100, 0.1% Sarkosyl, 10 mM β-mercaptoethanol). Cells were lysed using a hand-held motor homogenizer in the

presence of fine silica. The cell debris and fine silica were removed by centrifugation. GUS activity in the cell lysates was assayed using 4-methylumbelliferone-β-D-glucuronide as a substrate following the method of Scott *et al.* [20]. 4-methylumbelliferone fluorescence was measured with a TKO 100 Mini-Fluorometer (Hoefer Scientific Instruments, USA.). The values obtained were converted into specific activity using a calibration curve with known amounts of 4-methylumbelliferone product. Protein concentration of the cell lysates were determined using the Bio-Rad Detergent Compatible Protein Assay Kit (Bio-Rad Laboratories, USA).

Nucleotide sequence analysis

DNA sequences of the promoter-active fragments and recombinant plasmids from 5'RACE-PCR were determined by an automated sequence analyzer (Perkin Elmer, ABI, Model 377). Sequence comparisons and analysis were carried out using the Wisconsin Package Version 9.1 solfware, Genetics Computer Group (GCG, Madison, WI, USA) and the BLAST program [1] with the GenBank database and the genome database for *Synechocystis* PCC6803 [15].

5'Rapid Amplification of cDNA Ends (RACE) -PCR

Total RNA was extracted from *E. coli* or cyanobacterial cells using Trizol Reagent (GIBCO BRL, USA). The absence of DNA contamination in the total RNA sample was verified using PCR detection. 5'RACE-PCR method, used to identify the transcription initiation sites, were carried out following the manufacturer's manual (5'/3' RACE kit, Boehringer Mannheim, Germany) as described [6].

RESULTS AND DISCUSSION

Promoter strength of the isolated DNA fragments in *Synechococcus* PCC7942 and *E.coli*

The promoter strength of the isolated DNA fragments was determined by measuring GUS activity. In comparison with the *CaMV35S* and λP_R promoters, the GUS activity of the 7 isolated promoter-active fragments is shown in Table 1. Promoter-active fragments E3, E4, D13 and D21 were active in both *Synechococcus* and *E. coli*. E8, E10 and E14 were active only in *Synechococcus* but not active in *E. coli*. In *Synechococcus*, E3 was stronger than the λP_R promoter, whereas E10 and E14 were of similar strength. These 7 isolated promoters were not inducible by heat at 39°C. We also observed that light intensity affected expression of E10. GUS activity of E10 in *Synechococcus* grown at 1,500 Lux was 65 (±7) nmole/ min/ mg protein. This activity was approximately 4 times lower than that detected

when the *Synechococcus* were grown at 5,000 Lux, (see Table 1). GUS activity of the remaining 6 promoters was not significantly different when they were grown at 1,500 or 5,000 Lux. GUS activity of E3, E4, D13 and D21 in *Synechococcus* was higher than the corresponding activity in *E. coli*. The region immediately upstream of the initiation codon of the GUS gene did not contain a canonical Shine-Dalgarno sequence (GGAGG) similar to that found in *E. coli* (Fig. 1). This may be one of the factors that caused lower levels of GUS activity in *E. coli* than in *Synechococcus* (Table 1). Although it has been shown that in *E. coli*, the presence of canonical SD sequence increases expression of the reporter gene, SD-independent translation initiation system also exists [7]. No concensus sequence of a cyanobacterial counterpart to the SD sequence in *E. coli* has been previously reported.

None of the 7 isolated promoters were as strong as the λP_R promoter in *E. coli*. Several promoters from this study were stronger or as strong as the λP_R promoter which has been previously shown to be a strong promoter in *Synechococcus* [5, 14] and thus they are useful for high expression of heterologous genes in *Synechococcus*. Further characterization of the isolated promoters will reveal more detail of cyanobacterial promoter structures.

Nucleotide sequence analysis of promoter-active fragment E3

The insert fragment E3 exhibited the strongest promoter activity in *Synechococcus* (Table 1), and therefore further characterized. The nucleotide sequence at the 3'end of the 1.5 kb of E3 fragment, immediately upstream of the promoterless-GUS gene, was determined. Nucleotide sequence analysis revealed an ORF of 107 amino acids (Fig. 1) which was highly homologous to the hypothetical proteins slr1846 from *Synechocystis* PCC6803 and ORF107 from *Porphyra purpurea* with 76.6% and 53.3% amino acids identity respectively (Fig. 2). The 74 nucleotides following the ORF107 (Fig. 1) were 100% identical to the $tRNA^{pro}$ (anticodon GGG) gene of *Synechocystis* PCC6803 (accession no. D90900). Upstream of the 5'end $tRNA^{pro}$ gene, no conformed *E. coli* promoter element was detected (Fig. 1). At the 3'end of this gene there was an 15 bp inverted repeat (Fig. 1). Immediately downstream at the end of the $tRNA^{pro}$ gene of *Synechocystis* PCC6803, an 10 bp inverted repeat (TTCAGCTGAA) is observed (accession no. D90900). There is no Sau3A1 site within ORF107 and $tRNA^{pro}$ gene, thus the organizations of the gene could not be artifacts from cloning.

The tRNA^{pro} gene can act as promoter

Transcription initiation sites of E3-GUS were determined using 5'RACE-PCR. Sequencing analysis of 5'RACE-PCR independent clones revealed that transcription initiation sites in *E. coli* were located at TTG (6 out of 10 clones) and at G (4 clones) indicated as (a) and (b) respectively. Those in *Synechococcus* were located at TTG (11 out of 13 clones) and at A (2 clones) indicated as (c) and (d) respectively (Fig. 1). Thus, there are one major and one minor transcription initiation sites of E3-GUS. The major transcription initiation site in *E. coli* was located within nucleotides TTG identical to that in *Synechococcus*. Thus in both organisms, the GUS transcripts were probably controlled by the same functional promoter sequence. Immediately upstream of the transcription initiation sites was the tRNA^{pro} gene, and therefore the tRNA^{pro} gene can act as a strong promoter. However, it did not contain any -10 and -35 regions which conformed to those of *E. coli*σ^{70} promoter and may be one of the factors that caused low promoter activity in *E. coil* (Table 1). The tRNA^{pro} gene contained two regions, block A and B (Fig. 1), exhibiting strong homology respectively to block A (TGGCNNAGTGG) and B (GGTTCGANNCC) of the eukaryotic tRNA gene [9]. In the tRNA^{pro} gene, two additional nucleotides at nt.16 and 17 interrupt the block A (Fig. 1). It has been shown that blocks A and B are split promoter sequences of the tRNA^{Leu} gene of *Xenopus laevis* and recognized by RNA polymerase III [9]. Yeast tRNA gene containing the two conserved blocks can act as promoter in yeast [16]. Whether these two conserved blocks are the promoter elements that control the GUS transcripts in *E. coli* and in *Synechococcus* remain to be investigated. Further deletion analysis and site directed mutagenesis of the potential region in the tRNA^{pro} gene could elucidate the characteristics of a strong promoter in *Synechococcus*.

Nucleotide sequence analysis of promoter-active fragment E10

The nucleotide sequence at the 3'end of the 1.75 kb of E10 fragment, immediately upstream of the promoterless-GUS gene, was determined. Nucleotide sequence analysis revealed the 233 nucleotides (Fig. 3) whose deduced amino acid sequence was highly homologous to the *htpG* heat shock protein gene of *Synechocystis* PCC6803 (Fig. 4). However, only 3'end of the *htpG* gene was obtained. Since there was a Sau3A1 site upstream of the *htpG* gene, the missing of 5'end of the gene may be due to Sau3A1 partial digestion of the DNA used in the cloning procedure. No significant similarity to downstream sequence of the htpG gene of E10 was detected in the database.

Deletion analysis of promoter-active fragment E14

The effects of different length of E14 deletions on GUS expression are shown in Fig. 5. The GUS activity of *Synechococcus* harboring pKG-E14D2 was similar to that of cells harboring pKG-E14. Preliminary results of 5'RACE-PCR revealed that the transcription initiation sites of E14D2-GUS mRNA were located within 150 bp upstream of the PvuII restriction endonuclease site (data not shown). Thus, it was expected that the promoter region of E14D2 might locate within the HaeIII site (Fig. 5). However, the GUS activity of cells harboring further deleted plasmid pKG-E14D3 was reduced to background level. The GUS activity of cells harboring pKG-E14D4 was similar to that of cells harboring pKG-E14D2. These results indicated that the region between HincII and HaeIII sites was important for E14D2 promoter activity (Fig. 5). Whether the HincII/HaeIII region is distant enhancer remained to be investigated.

Multiple sigma factor genes have been identified in *Synechococcus* PCC7942. Their encoded products have conserved domains characteristic of principal sigma factors of the σ^{70} class [21, 22]. However, our results show that the majority of isolated *Synechococcus* promoters did not function in *E. coli*. The $tRNA^{pro}$ gene, which functions as a promoter and lacks the *E. coli* promoter element, was not recognized by sigma factors of the σ^{70} class. The use of these isolated promoters as DNA templates to probe for new RNA polymerase activity *in vitro* will clarify whether other uncharacterized sigma factors exist in *Synechococcus* PCC7942.

LITERATURE CITED

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410
2. Angsuthanasombat C, Panyim S (1989) Biosynthesis of 130-kilodalton mosquito larvicide in the cyanobacterium *Agmenellum quadruplicatum* PR-6. Appl Environ Microbiol 55:2428-2430
3. Casadaban MJ, Cohen SN (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J Mol Biol 138:179-207
4. Chungjatupornchai W (1990) Expression of the mosquitocidal-protein genes of *Bacillus thuringiensis* subsp. *israelensis* and the herbicide-resistance gene *bar* in *Synechocystis* PCC6803. Curr Microbiol 21:283-288
5. Chungjatupornchai W, Panyim S (1994) Advances in heterologous gene expression in cyanobacteria. In: Phang SM, Lee YK, Borowitzka MA, Whitton BA (eds) Algal biotechnology in the Asia-Pacific region. Malaysia: University of Malaya, pp186-189

6. Chungjatupornchai W, Senawong T and Panyim S (1999) Isolation and characterization of synechococcus PCC7942 promoters: tRNa pro gene functions as a promoter. Curr Microbiol 38:210-216
7. Fargo DC, Zhang M, Gillham NW (1998) Shine-Dalgarno-like sequences are not reguired or translation of chloroplast mRNA in *Chlamydomonas reinhardtii* chloroplasts or in *Escherichia coli.* Mol Gen Genet 257:271-282
8. Fukuda H, Sakai M, Nagahama K, Fujii K, Matsuoka M, Inoue Y, Ogawa T (1994) Heterologous expression of the gene for the ethylene-forming enzyme from *Psuedomonas syringae* in the cyanobacterium *Synechococcus.* Biotech Lett 16:1-6
9. Galli G, Hofstetter H, Birnstiel ML (1981) Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature 294:626-631
10. Giovannoni SJ, Delong EF, Olsen GJ, Pace NR (1988) Evolutionary relationships among cyanobacteria and green chloroplasts. J Bacteriol 170:720-726
11. Gray MW (1989) The evolutionary origins of organells. Trends Genet 5:294-299
12. Kawata Y, Yamano N, Kojima H, Itoh S (1991) Expression of salmon growth hormone in the cyanobacterium *Agmenellum quadruplicatum.* Biotech Lett 13:851-856
13. Kuhlemeier CJ, Thomas AAM, van der Ende A, van Leen RW, Borrias WE, van den Hondel CAMJJ, Arkel GA (1983) A host-vector system for gene cloning in the cyanobacterium *Anacystis nidulans* R2. Plasmid 10:156-163
14. Mermet-Bouvier P, Chauvat F (1994) A condition expression vector for the cyanobacteria *Synechocystis* sp. strains PCC6803 and PCC6714 or *Synechococcus* sp. strains PCC7942 and PCC6301. Curr Microbiol 28:145-148
15. Nakamura Y, Kaneko T, Hirosawa M, Miyajima N, Tabata S (1998). CyanoBase, a www database containing the complete nucleotide sequence of the genome of *Synechocystis* sp. strain PCC6803. Nucl Acids Res 26:63-67
16. Otter CA, Straby KB (1991) Transcription of eukartotic genes with impaired internal promoters: the use of a yeast tRNA gene as promoter. J Biotech 21:289-294
17. Rippka R (1988) Recognition and identification of cyanobacteria. Methods Enzymol 167:28-67
18. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press
19. Schneider GJ, Haselkorn R (1988) RNA polymerase subunit homology among cyanobacteria, other eubacteria, and archaebacteria. J Bacteriol 170: 4136-4140

20. Scott R, Draper J, Jefferson R, Dury G, Jacob L (1988) Analysis of gene organization and expression in plants. In: Draper J, Scott R, Armitage P, Walden R (eds) Plant genetic transformation and gene expression: a laboratory manual. London: Blackwell Scientific Publications, pp 265-337
21. Sugimoto Y, Tanaka K, Masuda S, Takahashi H (1997) The *rpoID1* gene of *Synechococcus* sp. strain PCC7942 encodes the principle sigma factor of RNA polymerase. J Gen Appl Microbiol 43:17-21
22. Tanaka K, Masuda S, Takahashi H (1992) Multiple rpoD-related genes of cyanobacteria. Biosci Biotech Biochem 56:1113-1117
23. Tandeau de Marsac N, de la Torre F, Szulmjster J (1987) Expression of the larvicidal gene of *Bacillus sphaericus* 1593M in the cyanobacterium *Anacystis nidulans* R2. Mol Gen Genet 209:396-398
24. William JGK (1988) Construction of specific mutations in photosystem II photosynthetic reaction center by genetic engineering methods in *Synechocystis* 6803. Methods Enzymol 167:766-778

Table 1. GUS activity of promoter-GUS fusion in *Synechococcus* PCC7942 and *E. coli* MC1061.

Plasmid	GUS activity (nmole/ min/ mg protein) [a]	
	Synechococcus [b]	*E. coli* [c]
Host (without pKG)	0.11 (± 0.01)	13 (± 2)
pKG (control)	0.13 (± 0.01)	13 (± 2)
pKG-35S	6 (± 1.7)	N
pKG-P_R (non induction) [d]	27 (± 10)	N
pKG-P_R (heat induction) [d]	276 (± 30)	505 (± 8)
pKG-E3	325 (± 20)	57 (± 8)
pKG-E4	36 (± 5)	21 (± 3)
pKG-E8	22 (± 9)	N
pKG-E10	298 (± 30)	N
pKG-E14	245 (± 22)	N
pKG-D13	145 (± 20)	70 (± 3)
pKG-D21	79 (± 15)	31 (± 8)

[a]: The specific activities are the means of three independent experiments (each in duplicate), with standard deviations indicated in parantheses. N: no expression (the GUS activity is not higher than the control).

[b]: *Synechococcus* was grown on BG-11 agar at 28°C with a light intensity of approximately 5,000 Lux, for 6 days.

[c]: *E.coli* was grown in LB broth at 37°C for 14 hr.

[d]: For non induction conditions, cells were grown at 28°C. For heat induction, a 5-day culture of *Synechococcus* and a log-phase culture of *E. coli* were further grown at 39°C, 24 hr and at 42°C, 18 hr respectively.

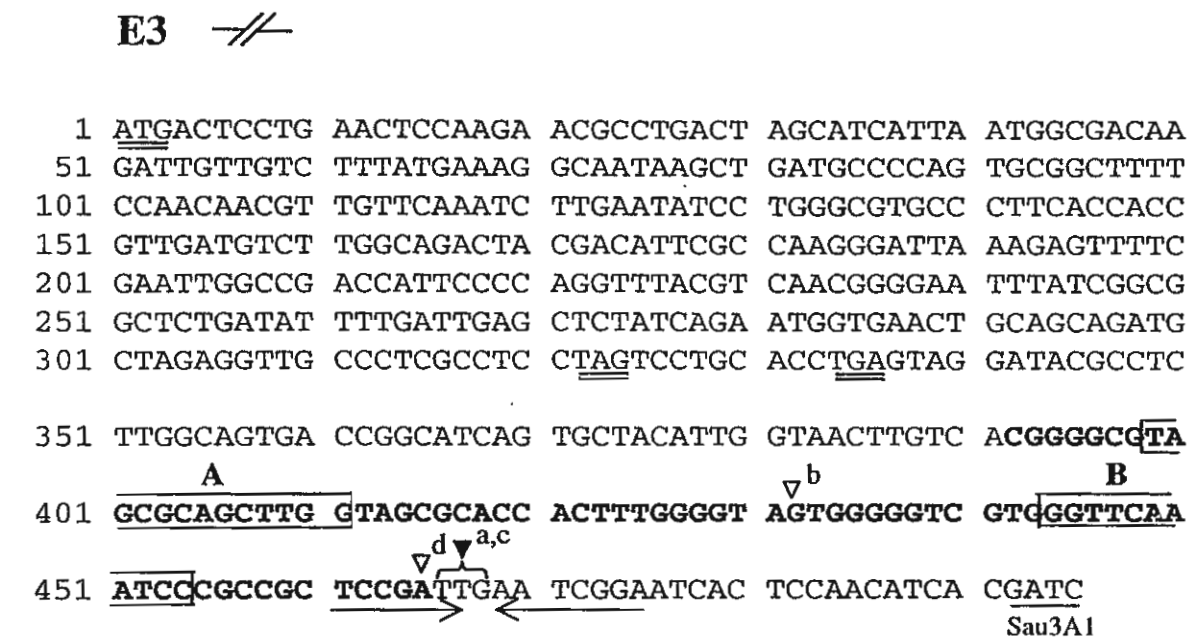

Fig. 1. Nucleotide sequence of the E3 promoter-active fragment.

The nucleotide sequence shown is the 3' end of the 1.5 kb E3 fragment immediately upstream of the promoterless-GUS gene (GenBank Database accession no. AF083392). The start and stop codons of an open reading frame of 107 amino acids (shown in Fig. 17) are marked by double lines. The 74 nucleotides (nt. 392-465) of $tRNA^{Pro}$ gene is in bold. The two boxed regions A and B conform to the two conserved sequence blocks within eukaryotic tRNA genes. The transcription initiation sites of GUS transcripts were determined by 5' RACE-PCR. In *E. coli*, the transcription initiation sites are indicated as (a) and (b). In *Synechococcus*, the transcription initiation sites are indicated as (c) and (d). Major and minor transcription initiation sites are marked by black and white triangles respectively. The inverted repeat sequence at the 5' end of the transcript is marked by arrows.

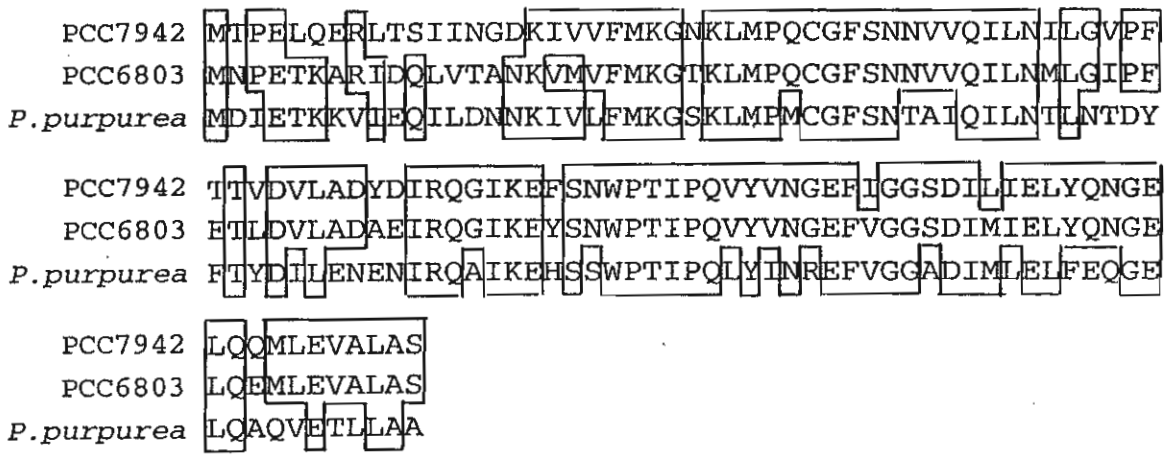

Fig. 2. Comparison of the ORF107 of E3 promoter-active fragment.

The ORF from *Synechococcus* PCC7942 (this study, accession no. AF083392), slr1846 from *Synechocystis* PCC6803 (accession no. P73056) and ORF107 from *Porphyra purpurea* (accession no. U38804). Boxed residues represent identical amino acids.

E10 //

```
  1  CCCAGTCAGT AGTCAGAGAG CAATACCGTT AAACCCTGCC CACGCGGGGT
 51  TTTTTTATGG GCGATCGCAG ACGCGTCCGT AAGATGAAGG GGTGATGTTG
                  Sau3A1
101  GCCATCGCCT CAATGCTGAT CCGGCTAATA TACCCGATGA CCATGTCCTG
                       Sau3A1
151  CTCGTGAATA CCGCGCATCC TCTTGTCCAG AACATCCTTA GTCTGCAGCA
201  GGGGGCTATC CTCAGCAGTG ATGGACACTC TCCAAGCCAG GTCCTGGCAG
251  AGCAGCTCTG TCGACACATC TATGACTTGG CCTTGATGAC CCAAAAAGGG
301  TTTGATGCTG AGGGAATGAA AGCCTTCATT GAGCGTTCTA ATGCGGTCTT
351  GACGGCGTTG ACGACTCGCC AGTGAAGGGT TTAGGGGAAC TCTAAACTTC
401  ACAGAACGTC ATCCTAGCTA TCTCGCCCTC AGACGTGAGC CTCGCTAAGC
451  TGAGGGCGAC GATCGCAGCG GAAAAATGTG ACCGGGGCGG CGCAGATC
                                                  Sau3A1
```

Fig. 3. Nucleotide sequence of the E10 promoter-active fragment.

The nucleotide sequence shown is the 3'end of the 1.75 kb E10 fragment immediately upstream of the promoter-less GUS gene. The 233 nucleotides of htpG heat shock protein gene are in bold. The stop codons of htpG are marked by double lines.

```
E10            PDDHVLXVNT AHXLVQNILS LQQGXILSSD GHSPSQVLAE
               P+ HVL  NT  H L++NILS L QG I++   G   SPSLA+
PCC6803  580 PEQHVLAINT NHPLIKNILS LSQGGIVTGS GESPSAELAK

E10          QLCRHIYDLA LMTQKGFDAE GMKAFIERSN AVLTALT
             LC+ H+YDLA LM QKGFDAE GMK FIERSN AVLT LT
PCC6803      SLCQHVYDLA LMAQKGFDAE GMKGFIERSN AVLTRLT 657
```

Fig. 4. Amino acid aligment of E10 and htpG protein of *Synechocystis* PCC6803.

Deduced amino acids of nt. 131-363 of E10 promoter-active fragment was highly homologous to the C-terminal of htpG heat shock protein of PCC6803. There are 68% (53/77) identities and 80% (62/77) positives.

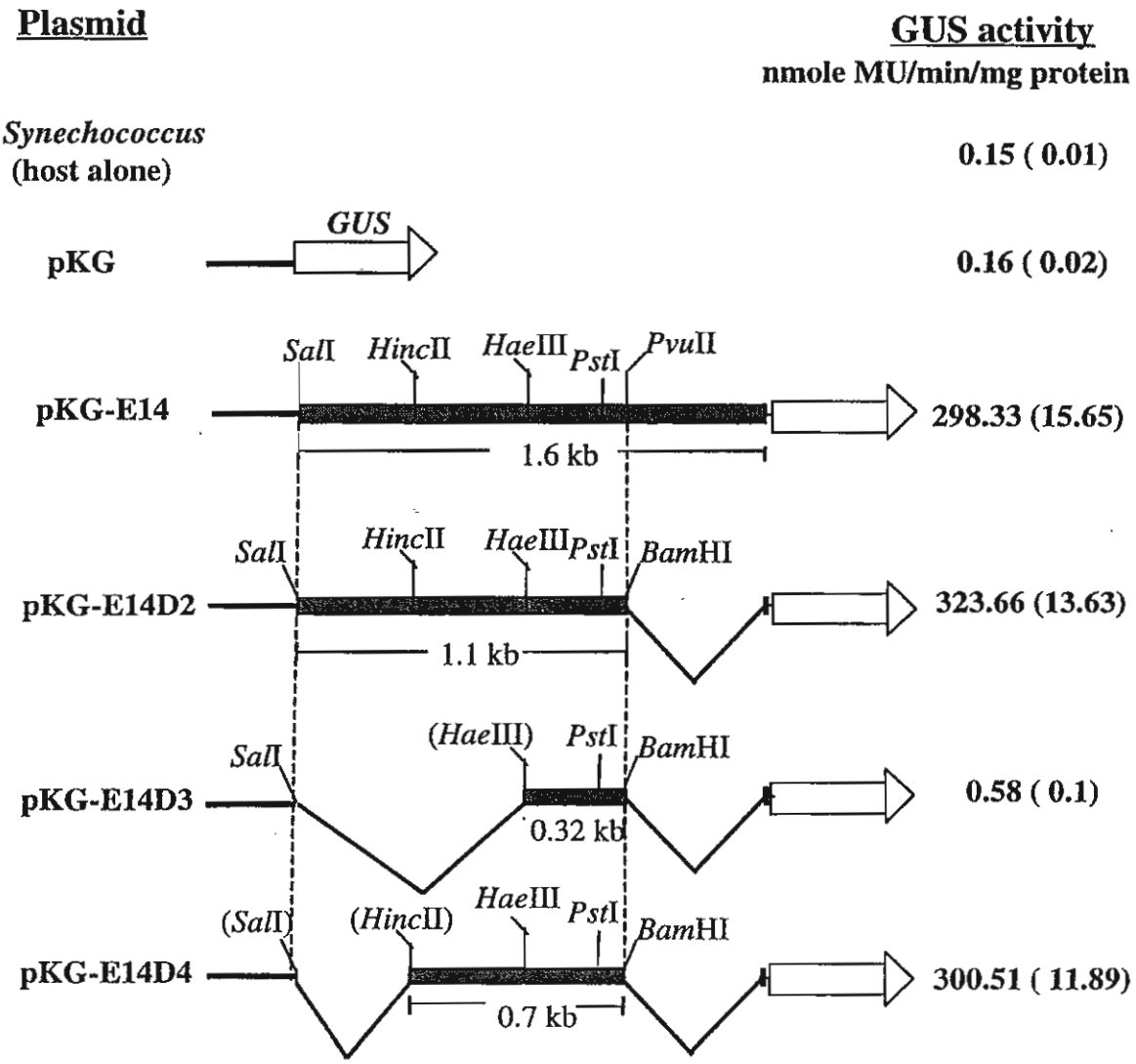

Figure 5. Effect of different lenght of E14 deletions on GUS expression

Synechococcus harbouring pKG, pKG-E14 and their derivatives were grown in BG-11 broth at 28 C with a light intensity of approximately 5,000 Lux, for 9 days (OD_{730} ~ 5). The GUS specific activities are means of three independent experiments (each in duplicate), with standard error mean indicated in parentheses. The figure is not drawn to scale. The restriction sites in parentheses are lost after ligation.

เนื้อหางานวิจัย

CHARACTERIZATION OF EXPRESSED SEQUENCE TAGS FROM BLACK TIGER SHRIMP 'S HEMOCYTES

ABSTRACT

Two lambda ZAPII-based cDNA libraries was prepared from poly(A)+ RNA of hemocytes from Black Tiger Shrimp (*Penaeus monodon* Fabricius). Random clones were isolated and their 5' end partial nucleotide sequence determined by automated procedure. A segment of each DNA sequence was used to search online nucleotide and protein sequence databases by two BLAST programs. From the first cDNA library with a lower phage titer, 53 DNA sequences were obtained. Among these, 12 sequences (22%) did not match existing DNA sequence in online databases by either BLASTN or BLASTX searches, while 12 sequences (23%) matched known DNA sequences for known protein identities and 29 sequences (55%) matched 16S ribosomal rRNA sequences. Among those matching protein coding sequences, 3 sequences (6%) matched those of unique genes, 5 sequences (9 %) matched those of ribosomal protein genes, 4 sequences (7%) matched those of mitochondrial genes. A second cDNA library constructed from higher amount of poly(A)+ RNA provided higher phage titer. Analysis of random cDNA sequences from the second cDNA library was in progress. The study of partial cDNA sequences should allow rapid characterization of the DNA sequences, some of which may be relevant to immune system of the shrimp.

INTRODUCTION

Shrimps in the subphylum Crustaceae are a group of highly diversed organisms of over 38,000 known species (1). Human consume a number of marine shrimps species, yet those in the Family Penaeidae constitute the bulk of the catches (2). With the adoption of certain fast growning *Penaeus* species for aquaculture around the World, its commercial impact is even higher. In South and Southeast Asia, *Penaeus mondon* is the choice species. For Thailand alone, the World 's number one exporter of frozen shrimps, a yearly export value for about 250,000 tonnes of frozen shrimps amounted to US$ 2 billions.

Despite its economic importance to the World, little attention has been paid to study the shrimps in Molecular Biology aspects. Online DNA sequence databases possessed only a hundred plus records, as of early 1999. In recently years' a number of viruses infecting

useful genetic marker for studying diversity and significance in pathogenesis[13] as well as variations within populations of closely related bacteria.[14]

The flagellin gene of *B. pseudomallei* has been studied and its sequence reported.[15] Ithas been shown to be a typical flagellin in that there are two conserved regions at the N- and C-terminal ends and a variable part at the middle region of the gene. Commonly, the conserved regions are most highly conserved among the same genus.[16] Thevariable region shows significant differences between species and even between strains in the same species.[17,18] In this study, the characteristics of the flagellin gene are used to isolate and compare the flagellin gene are used to isolate and compare the flagellin sequences between *B. pseudomallei* and *B. cepacia* from different clinical isolates. We also report here a simple method for identification of both organisms based on the flagellin sequence by PCR-restriction fragment length polymorphism (PCR-RFLP) analysis.

MATERIALS AND METHODS

Bacterial species, plasmid and growth conditions

Clinical isolated *B. pseudomallei* and *B. cepacia* used in this study were obtained from the National Institute of Health, Thailand and are listed in Table 1. All isolated were grown at 37°C on nutrient agar and broth. *Escherichia coli*K12 JM109,19 which was used as the recipient for recombinant plasmids was grown at 37°C on Luria-Bertani (LB) agar and broth without or with antibiotic (100 μg ml^{-1} ampicillin) for plasmid selection. PUC19[19] was used as a vector for cloning and sequencing analysis of the flagellin genes.

Genomic DNA purification and PCR amplification of flagellin gene

Genomic DNA purification was performed by QIAGEN Genomic-tip according to QIAGEN protocol. The purified genomic DNA was directly used as a template for PCR amplification reaction. The PCR reaction was carried out in a 50 μl mixture containing 1 μl of DNA template (about 200 ng), reaction buffer (10 mM Tris-HCl pH 8-8, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100), 200 μM of each dNTP, 0-4 μM of each primer (Forward primer; 5′ ATG CTC GGA ATC AAC AGE AAC AT 3′ and Reverse primer ; 5′ CAG GAG CTT CAG CAC TTG CTG 3′). All reactions were overlaid with two drops of light minerol oil (Sigma) and heated at 100°C for 10 min before adding 1 unit of *Vent* DNA polymerase (New England Biolabs). The synthesis was performed for 30 cycles in an automated DNA Thermal Cycler (480 Perkin-Clmer Cetus). The PCR condition of all isolates was carred out with the debaturation temperature at 94°C for 1 min, the annealing temperature at 65°C for 1 min, and followed by the synthesis temperature at 72°C for 1 min.

Cloning and DNA sequencing

The PCR-amplified products were excised from the agarose gel and purified with Gene Clean II kit (Bio 101, Inc.). Cloning in plasimid pUC19 vector was performed by blunt-end ligation at the *Sma*I cloing site following the standard protocol described proviously.[20] DNA sequencing fo both strands of pUC19-based clones and subclones was carried out by the dideoxy-chain termination method.[21] An automated DNA sequencer (Applied Biosystem, model 377) was utilized to determine the nucleotide sequences of both strands of the entire fragment. Computer analysis of the sequences was performed by using PC/Gene software package (Intelligenetics), the CLUSTAL W[22] for multiple sequence alignment, and the NCBI BLAST e-mail server.[23]

PCR-restriction endonuclease analysis

The PCR-amplified products (5 µl) were subsequently used for restriction endonuclease analysis either by *Pst*I or *Xho*I digestion. The reaction was carried out in a total volume of 20 µl using the condition srecommended by the sepplier (Ner England Biolabs).

RESULTS

PCR-amplification of *B. pseudomallei* and *B. cepacia* flagellin genes

Specific primers were designed based on the 5′-and 3′-end conserved sequences from the flagellin gene of *B. pseudomallei* 1026b, and the flagellin gene was seccussfully amplified from different clinical isolates of both *B. pseudomallei* and *B. cepacia*. As expected, a single PCR product of apporximately 1-1 kb in size was obtained from all *B. pseudomallei* isolated and a PCR product of the same size was also found from two clinical isolates of *B. cepacia*, as shown in Fig.1

Cloning and DNA sequencing

All of the PCR-amplified products were subsequently cloned into the pUC19 vector and were consuently analysed for their DNA sequences using primers for both orientations. The results can be accessed from the GenBank database under the accession numbers AF078151; AF078152; AF080259; and AF080260 for flagellin sequences of *B. pseudomallei* NF 10/38, NF 47/38, NF 105/37; *B. pseudomallei* NF 154/37; *B. cepacia* DMS 2555; and *B. cepacia* DMS 3027, respectively.

The deduced amino acid sequences were examined for similarity to sequences previously reported.[15] In *B. pseudomallei*, the sequences showed 99% homology among

themselves and with the sequences previously reported, whereas the sequences of *B. cepacia* showed 100% homology in our two isolates, and 97% homology with the *B. cepacia* flagellin type II sequences recently reported,[24] as shown in Fig. 2a and 2b, respectively. Comparison of the amino acid sequences between these two species in Fig.3 reveals the general features of the flagellin sequence in that there are the N- and C-terminal conserved regions and the central part is diverse, resulting in less homology (84%).

PCR-restriction endonuclease analysis

Figure 4 represents a restriction enzyme mapping of the flagellin sequences from both apecies. According to the maps, the restriction sites for endouncleases *Psi*I and *Xho*I are apparently distinct between the two species. Using either *Psi*I and *Xho*I digestion of the PCR products, the obtained restriction endonuclease patterns can be used to identify the two species, as shown in Fig. 5a and 5b, respectively, In the PCR-restricted *Psi*I patternm, three fragments of 676, 264 and 216 bp were obtained from *B. pseudomallei*. The PCR procuct form *B. cepacia* contains five restriction sites for *Psi*I and should yield fragments of 417, 264, 165, 164, 70 and 69 bp. Three bands can be visualized in the gel in Fig. 5a, the 417, 264 and the combined 165 and 164 bp fragments, In the PCR-restricted *Xho*I pattern, two fragments of about 584 and 567 bp were detected from *B. pseudomallei*, whereas the recognition site for *Xho*I in absent in *B. cepacia*.

DISCUSSION

PCR-based isolation of the flagellin gene from the bacterial species *Pseudomonas*[16-18] has been achieved, making use of the highly conserved 5′ and 3′ regions of the gene. Since, the specific primers designed from the first codon of the N-terminal and the third codon from the C-terminal sequences were not able the amplify flagellin gne of bacterial species in the genus *Burkholderia*, the conservation of these regions has been shown to be restrictive and more diverse across the genus.[16] In this study, we therefore designed specific primers based on the 5′ and 3′ conserved flagellin sequence from *B. pseudomallei* 1026b,15 and they wete secessfully employed to isolate the flagellin gene from four clinical isolates of *B. pseudomallei*. In an attempt to extend this method for isolation of the flagellin gene in other spceies in the same genus *Burkholderia* using the same pair of primers, two chinical isolates of *B. cepacia* were also successfully amplifiied. The PCR-amplified flagellin products from both species were of a similar size of 1-1 kb, which is the experted size of the flagellin gene of *B. pseudomallei*. In the case of *B. cepacia* there have been recently reported at least two types of flagellin genes.[24] Type I is 1-4 kb in size and type II is 1-1kb.

Thus, our isolate should contain the type II lfagellin gene. As the majority of *B. cepacia* clinical isolates have been found to have typeII flagellin,[24] it is similar to the size of flagellin gene found in *B. pseudomallei* clinical isolates. This finding also indicates a close relationship of the two organisms.

In order to be sure that the PCR-amplified products botained fromthese two species were not artifacts of contamination in the clinical samples or in the PCE reaction, DNA sequencing analysis was applied. The results demonstrated that the 1-1 kb PCR-amplified products from the two species have differents flagellin sequecnes. Moreover, the DNA sequences found in the four clinical isolates and a published sequence of *B. pseudomallei*[15] are highly conserved throughout the gene, especially at the amino acid level (more than 99% homology). This was also found in the two clinical isolates and a published sequence of a type II flagellin gene of *B. cepacia*24 (97% homology). In contrast, the amino acid sequence comparison between the two species in more divergent (84% homology). The variations are commonly located in the middle part of the gene which is the characteristic[25-28] of a flagellin gene. Interestionly, the variable region of the *B. cepacia* type II lfalgellin sequence has a four amino acid deletion a position 248-251 **(-PSFO-)** compared to the *B. pseudomallei* flagellin sequence. This finding suggests that a nested PCR could be developed for the bacteriological classification by designing specific primers in this region.

To simply identify these two species based on the PCR-amplified flagellin products, *Psl*I and *Xho*I restriction endonuslease reactions were performedI Both restriction endonucleases are also found and localize in the same regions of a type II *B. cepacia* flagellin sequence.[24] Either the *Psl*I and *Xho*I digestion is able to distinguish the two species, therefore the flagellin sequence is a useful marker for identification of the closely related species. Although PCR-RFLP analysis enabled flagellin type classification within population of *B. cepacia*,[24] it was not applicable for studying variation among species. Our inverstigation is mainly to study variation between species, *B. pseudomallei* and *B. cepacia*, parricularly from clinical isolates. Previously, 165 rRNA sequences have been applied for identification of the species.[8] the flagellin sequence is an interesting target because a correlation may be established with the serological and pathogenesis of the diseases, particularly from clinical isolates. A study of larfer numbers of clinical isolated from both species are required to berrer assess these connections as well as to study diversity of the organisms.

REFERENCES

1. Chaowagul, W., White, N. J., Dance, D. A., Wattanagoon, Y., Naigowi, P., Davis, T.M., Looareesuwan, S. and Pitakwatchara, N. (1989) Melioidosis : a major cause of community-acquired septicemia in northeastern thailand. Journal of Infectious Diseases 159, 890-9.
2. Govan, J. R. W., Hughes, J. E. and Van Damme, P. (1996) *Burkholderia cepacia* mediacl, taxonomic and ecological issues, Journal of Medical Microbiology 45, 395-407.
3. Dejsirilert, S., Butraporn, R., Chiewsilp, D., Kondo, E. and Knai, k. (1989) High activity of acid phosphatase of *Pseudomonas pseucomallei* as a possible attribute relating to its pathogenicity. Japanese Journal of Medical Science and Biology 42, 39-49.
4. Kondo, E., Naigowit, P., Phanichruttiwong, P., Thirawattanusuk, N. and Knai, K. (1991) Fatty acid profile and acid phosphatase activity of fresh isolates of *Burkholderia pseudomallei*. Japanese Journal of Medical Science and Biology 44, 195-211.
5. Ashdown, L. R. (1992) Rapid differentiation of *Pseudomonas pseudomallei* from *Pseudomonas cepacia*. Letter in Applied Microbiology 14, 203-5.
6. Desakorn, V., Smith, M. D., Wuthiekanun, V., Dance, D. A. B., Aucken, H., Suntharasamai, P., Rajchanuwong, A. and White, N. J. (1994) Detection of *Pseudomonas pseudomallei* antigen in urine for the diagnosis of melioidosis. American Journal of Tropical Medicine and Hygiene 51, 627-33.
7. Trakulsomboon, A., dance, D.A., Smith, M.D., White, N.J. and Pitt, T.L. (1997) Ribotype differences between clinical and environmental isolates of *Burkholderia pseudomallei*. Journal of Medical Microbiology 46, 565-570.
8. Dharakul, T., Songsivilai, S., Viriyachitra, S., Luangwedchakarn, V., Tassaneetritap, B. and Chaowagul, W. (1996) Detection of *Burkholderia pseudomallei* in patients with septicemic melioidosis. Journal of Clinical Microbiology 34, 609-614.
9. Rattanathongkorn, A., Sermswan, R.W. and Wongratanacheewin, S. (1997) Detection of *Burkholderia pseudomallei* in blood samples using polymerase chain reaction. Molecular and Cellular Probes 11, 25-31.
10. Picken, R.N. (1992) Polymerase chain reaction primers and probes derived from flagellin gene sequences for specific detection of the agents of *Lyme* disease and North America relapsing fever. Journal of Clinical Microbiology 30, 99-114.
11. Oyofo, B.A. and Rollins, D.M. (1993) Efficacy of filter types for detecting *Campylobacter jejuni* and *Campylobacter coli* in environmemtal water samples by polymerase chain recation. Applied Environmental Microbiology 59, 4090-4095.

12. Gray, D.I. and Kroll, R.G. (1995) Polymerase chain reaction amplification of the *fla*A gene for the rapid identification of *Listeria spp.* Letter in Applied Microbiology 20, 65-68.

13. Penn, C.W. and Luke, C.J. (1992) Bacterial flagellar diversity and significance in pathogenesis. FEMS Microbiology Letter 100, 331-336.

14. Winstanley, C. and Morgan, J.A.W. (1997) The bacterial flagellin gene as a biomarker for detection population genetics and epidemiological analysis. Microbiology 143, 3071-3084.

15. DeShazer, D., Brett, P.J., Carlyon, R. and Woods, D.E. (1997) Mutagenesis of *Burkholderia pseudomallei* with Tn5-OT185 : Isolation of motility mutants and molecular characterization of the flagellin structural gene. Journal of Bacteriology 179, 2116-2125.

16. Tanypradabkul, S., Senapin, S. and Panyim, S. (1998) PCR-based method for isolation of flagellin gene in *Pseudomonas species.* Journal of General and Applied Microbiology 44, 23-242.

17. Winstanley, C., Coulson, M.A., Wepner, B., Morgan, J.A.W. and Hart, C., (1996) Flagellin gene and protein variation amongst clinical isolates of *Pseudomonas aeruginosa.* Microbiology 142, 2145-2151.

18. Spangenberg, C., Heuer, T., Burger, C. and Tummler, B. (1996) Genetic diversity of flagellins of *Pseudomonas aeruginosa.* FEBS Letters 396, 213-217.

19. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains : nucleotide sequences of the M13mp 18 and pUC19 vectors. Gene 33, 103-119.

20. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning : a laboratory manual 2nd ed. Cold Spring Harber Laboratory Press. United States of America.

21. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. Proceedings of the National Academy Sciences of the United States of America 74, 5463-5467.

22. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sesityvity of progressive multiple sequence alignment through sequence weighting, position-speckfic gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673-4680.

23. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. Journal of Molecular Biology 215, 403-410.

24. Hales, B.A., Morgan, J.A.W., Hart, A. and Winstanley, C. (1998) Variation in flagellin genes and proteins of *Burkholderia cepacia.* Journal of Bacteriology 180, 1110-1118.

25. Joys, T.M. (1985) The covalent structure of the phase-1 flagellar filament protein of Salmonella typhimurium and its comparison with other flagellins. Journal of Biological Chemistry 260, 15758-15761.

26. Wei, L.N. and Joys, T.M. (1986) The nucleotide sequence of the H-1 r gene of *Salmonella rubislaw*. Nucleic Acids Research 14, 8227.

27. Kuwajima, G. W1988) Flagellin domain that affects H antigenicity of *Eschericahia* coli K-12. Journal of Bacterilogy 170, 485-488.

28. Nuijten, P.J., van Asten, F.J., Gaastra, W. and van der Zeijst, B.A. (1990) Structural an functional analysis of two *Campylobacter jejuni* flagellin genes. Journal of Biological Chemistry 265, 17798-17804.

Tabele 1. *B. pseudomallei* and *B. cepacia* from clinical isolates used in this study

Species and corresponding isolates number	Source of isolation
B. pseudomallei NF[a] 10/38	Blood
B. pseudomallei NF 47/38	Blood
B. pseudomallei NF 105/37	Pus
B. pseudomallei NF 154/37	Pus
B. cepacia DMS[b] 2555	Pus
B. cepacia DMS 3027	Blood

NF[a] is Nofermentation

DMS[b] is Department of Medical Sciences, Ministry of Public Health, Thailand

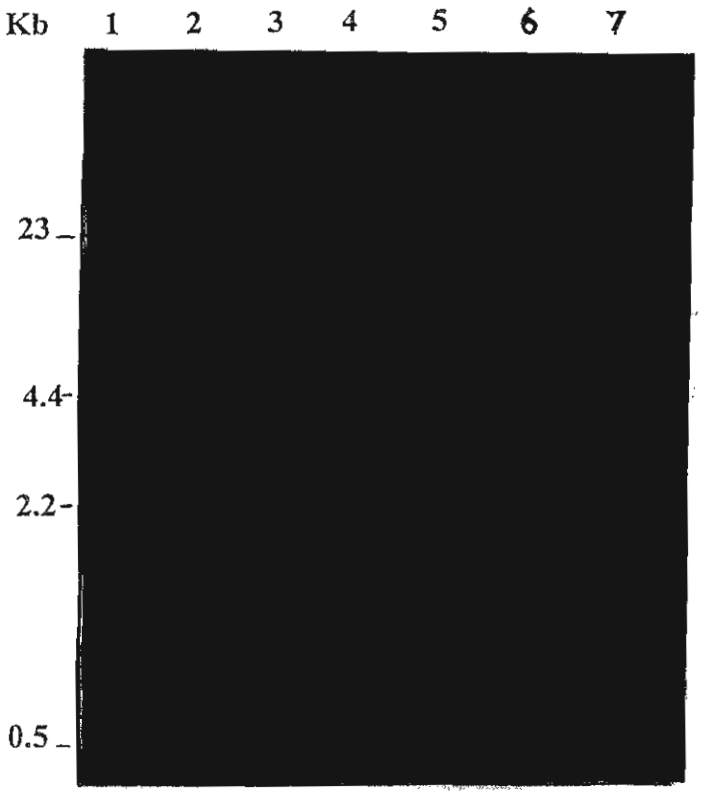

Figure1 Ethidium bromide staining of an 1% agarose gel electrophoresis of PCR-amplified flagellin gene from clinical isolates of B.pseudomallei NF 10/38 (lane 2), NF 47/38 (lane 3), NF 105/37 (lane 4), NF 154/37 (lane 5) and B.cepacia DMS 2555 (lane 6), DMS 3027 (lane 7), respcectively. Lane 1 is a standard size marker.

```
NF10/38 MLGINSNINSLVAQQNLNGSQGALSQAITRLSSGKRINSAADDAAGLAIATRMQTQINGL   60
NF47/38 MLGINSNINSLVAQQNLNGSQGALSQAITRLSSGKRINSAADDAAGLAIATRMQTQINGL
NF105/37MLGINSNINSLVAQQNLNGSQGALSQAITRLSSGKRINSAADDAAGLAIATRMQTQINGL
NF154/37MLGINSNINSLVAQQNLNGSQGALSQAITRLSSGKRINSAADDAAGLAIATRMQTQINGL
Bps      MLGINSNINSLVAQQNLNGSQGALSQAITRLSSGKRINSAADDAAGLAIATRMQTQINGL
         ************************************************************

NF10/38  NQGVSNANDGVSILQTASSGLTSLTNSLQRIRQLAVQASNGPLSASDASALQQEVAQQIS  120
NF47/38  NQGVSNANDGVSILQTASSGLTSLTNSLQRIRQLAVQASNGPLSASDASALQQEVAQQIS
NF105/37NQGVSNANDGVSILQTASSGLTSLTNSLQRIRQLAVQASNGPLSASDASALQQEVAQQIS
NF154/37NQGVSNANDGVSILQTASSGLTSLTNSLQRIRQLAVQASNGPLSASDASALQQEVAQQIS
Bps      NQGVSNANDGVSILQTASSGLTSLTNSLQRIRQLAVQASNGPLSASDASALQQEVAQQIS
         ************************************************************

NF10/38  EVNRIASQTNYNGKNILDGSAGTLSFQVGANVGQTVSVDLTQSMSAAKIGGGMVQTGQTL  180
NF47/38  EVNRIASQTNYNGKNILDGSAGTLSFQVGANVGQTVSVDLTQSMSAAKIGGGMVQTGQTL
NF105/37EVNRIASQTNYNGKNILDGSAGTLSFQVGANVGQTVSVDLTQSMSAAKIGGGMVQTGQTL
NF154/37EVNRIASQTNYNGKNILDGSAGTLSFQVGANVGQTVSVDLTQSMSAAKIGGGMVQTGQTL
Bps      EVNRIASQTNYNGKNILDGSAGTLSFQVGANVGQTVSVDLTQSMSAAKIGGGMVQTGQTL
         ************************************************************

NF10/38  GTIKVAIDSSGAAWSSGSTGQETTQINVVSDGKGGFTFTDQNNQALSSTAVTAVFGSSTA  240
NF47/38  GTIKVAIDSSGAAWSSGSTGQETTQINVVSDGKGGFTFTDQNNQALSSTAVTAVFGSSTA
NF105/37GTIKVAIDSSGAAWSSGSTGQETTQINVVSDGKGGFTFTDQNNQALSSTAVTAVFGSSTA
NF154/37GTIKVAIDSSGAAWSSGSTGQETTQINVVSDGKGGFTFTDQNNQALSSTAVTAVFGSSTA
Bps      GTIKVAIDSSGAAWSSGSTGQETTQINVVSDGKGGFTFTDQNNQALSSTAVTAVFGSSTA
         ************************************************************

NF10/38  GTGTAASPSFQTLALSTSATSALSATDQANATAMVAQINAVNKPQTVSNLDISTQTGAYQ  300
NF47/38  GTGTAASPSFQTLALSTSATSALSATDQANATAMVAQINAVNKPQTVSNLDISTQTGAYQ
NF105/37GTGTAASPSFQTLALSTSATSALSATDQANATAMVAQINAVNKPQTVSNLDISTQTGAYQ
NF154/37GTGTAASPSFQTLALSTSATSALSATDQANATAMVAQINAVNKPQTVSNLDISTQTGAYQ
Bps      GTGTAASPSFQTLALSTSATSALSATDQANATAMVAQINAVNKPQTVSNLDISTQTGAYQ
         ************************************************************

NF10/38  AMVSIDNALATVNNLQATLGAAQNRFTAIATTQQAGSNNLAQAQSQIQSADFAQETANLS  360
NF47/38  AMVSIDNALATVNNLQATLGAAQNRFTAIATTQQAGSNNLAQAQSQIQSADFAQETANLS
NF105/37AMVSIDNALATVNNLQATLGAAQNRFTAIATTQQAGSNNLAQAQSQIQSADFAQETANLS
NF154/37AMVSIDNALATVNNLQAMLGAAQNRFTAIATTQQAGSNNLAQAQSQIQSADFAQETANLS
Bps      AMVSIDNALATVNNLQATLGAAQNRFTAIATTQQAGSNNLAQAQSQIQSADFAQETANLS
         ********** ***************************************************

NF10/38  RAQVLQQAGISVLAQANSLPQQVLKLL-     388
NF47/38  RAQVLQQAGISVLAQANSLPQQVLKLL-
NF105/37RAQVLQQAGISVLAQANSLPQQVLKLL-
NF154/37RAQVLQQAGISVLAQANSLPQQVLKLL-
Bps      RAQVLQQAGISVLAQANSLPQQVLKLLQ
         ***************************
```

Figure 2(A). Alignmet of the amino acid sequences for flagellin gene from clinical isolates NF10/38, NF47/38, NF105/37, NF154/37 and a published sequence[15], Bps, of *B. pseudomallei*. Identical amino acids are indicated by an asterisk.

```
DMS3027  MLGINSNINSLVAQQNLNGSQNALSQAITRLSSGKRINSAADDAAGLAISTRMQTQINGL   60
DMS0704  MLGINSNINSLVAQQNLNGSQNALSQAITRLSSGKRINSAADDAAGLAISTRMQTQINGL
BcII     MLGINSNINSLVAQQNLNGSQNALSQAITRLSSGKRTNSAADDAAGLAISTRMQTQINGL
         ************************************************************

DMS3027  NQGVSNANDGVSMIQTASSALSSLTNSLQRIRQLAVQASTGTMSTTDQAALQQEVSQQIQ  120
DMS0704  NQGVSNANDGVSMIQTASSALSSLTNSLQRIRQLAVQASTGTMSTTDQAALQQEVSQQIQ
BcII     NQGVSNANDGVSMIQTASSALSSLTNSLQRIRQLAVQASTGTMSSTDQAALQQEVAQQIQ
         ******************************************.***********.****

DMS3027  EVNRIASQTTYNGTNILDGSAGIVSFQVGANVGQTISLDLSQSMSAAKIGGGLVQKGQTV  180
DMS0704  EVNRIASQTTYNGTNILDGSAGIVSFQVGANVGQTISLDLSQSMSAAKIGGGLVQKGQTV
BcII     EVNRIASQTTYNGTNILDGSAGIVSFQVGANVGQTMSLDLSQSMSAAKIGGGLVQKGQTV
         ***********************************.************************

DMS3027  GTVTGLSLDNAGAYTSSGAAITAINVLSDGKGGYTFTDQNGGAISQTVAQSVFGANATTG  240
DMS0704  GTVTGLSLDNAGAYTSSGAAITAINVLSDGKGGYTFTDQNGGAISQTVAQSVFGANATTG
BcII     GTVTGLSLDNAGAYVSSGATITAINVISDGQGGYTFTDQNGQSISSGAATAVFGSNATTG
         ************** ****.******.***.************.**    * .*********

DMS3027  TGTAVGNLTLQTGATGTGTSAAQQTAITNAIAQINAVNKPATVSNLDISTVSGANVAMVS  300
DMS0704  TGTAVGNLTLQTGATGTGTSAAQQTAITNAIAQINAVNKPATVSNLDISTVSGANVAMVS
BcII     SGTAVGALSLQPSATGANTTAAQLTAITNAIAQINAVNKPVTVSGLDISTVSGANVAMVS
         .***** *.**..***. *.*** ****************.*** ***************

DMS3027  IDNALQTVNNVQAALGAAQNRFTAIATSQQAESTDLSSAQSQITDANFAQETANMSKNQV  360
DMS0704  IDNALQTVNNVQAALGAAQNRFTAIATSQQAESTDLSSAQSQITDANFAQETANMSKNQV
BcII     IDNALQTVNNLQAALGAAQNRFTAIATAQQAESTDLSSAQSQITDANFAQETANMSKNQV
         **********.****************.**********************************

DMS3027  LQQAGISVLAQANSLPQQVLKLL-      384
DMS0704  LQQAGISVLAQANSLPQQVLKLL-
BcII     LQQAGISVLAQANSLPQQVLKLLQ
         ***********************
```

Figure 2(B). Alignment of the amino acid sequences for flagellin gene from clinical isolates DMS 2555, DMS 3027 and a published sequence24, BelI, of type II *B. cepacia.* Identical amino acide are indicated by an asterisk and similar amino acids are marked by a dot.

```
NF10/38  MLGINSNINSLVAQQNLNGSQGALSQAITRLSSGKRINSAADDAAGLAIATRMQTQINGL   60
DMS3027  MLGINSNINSLVAQQNLNGSQNALSQAITRLSSGKRINSAADDAAGLAISTRMQTQINGL   60
         ********************* ************************************

NF10/38  NQGVSNANDGVSILQTASSGLTSLTNSLQRIRQLAVQASNGPLSASDASALQQEVAQQIS  120
DMS3027  NQGVSNANDGVSMIQTASSALSSLTNSLQRIRQLAVQASTGTMSTTDQAALQQEVSQQIQ  120
         ************..***** *.***************** * .*..* .**********

NF10/38  EVNRIASQTNYNGKNILDGSAGTLSFQVGANVGQTVSVDLTQSMSAAKIGGGMVQTGQTL  180
DMS3027  EVNRIASQTTYNGTNILDGSAGIVSFQVGANVGQTISLDLSQSMSAAKIGGGLVQKGQTV  180
         *********.*** ******** .***********.*.**.***********.** ***.

NF10/38  GTIKVAIDSSGAAWSSGSTGQETTQINVVSDGKGGFTFTDQNNQALSSTAVTAVFGS-ST  240
DMS3027  GTVTGLSLDNAGAYTSS--GAAITAINVLSDGKGGYTFTDQNGGAISQTVAQSVFGANAT  240
         **.         * .*   *   * ***.******.******. **** *  .***. .*

NF10/38  AGTGTAASPSFQTLALSTSATSA-LSATDQANATAMVAQINAVNKPQTVSNLDISTQTGA  300
DMS3027  TGTGTAVG----NLTLQTGATGTGTSAAQQTAITNAIAQINAVNKPATVSNLDISTVSGA  300
         .*****       *.* * ** .  **. *.  *  .********* ********* .**

NF10/38  YQAMVSIDNALATVNNLQATLGAAQNRFTAIATTQQAGSNNLAQAQSQIQSADFAQETAN  360
DMS3027  NVAMVSIDNALQTVNNVQAALGAAQNRFTAIATSQQAESTDLSSAQSQITDANFAQETAN  360
           ********* ****.**.*************.*** *. *. *****  * *******

NF10/38  LSRAQVLQQAGISVLAQANSLPQQVLKLL-     388
DMS3027  MSKNQVLQQAGISVLAQANSLPQQVLKLL-     384
         .*. ************************
```

Figure 3. Alignment of the amino acid sequences of *B. pseydimallei* flagellin (NF10/38) and *B. cepacia* flagellin (DMS 3027). Identical amino acide are indicated by an asterisk and aimilar amino acids are marked by a dot.

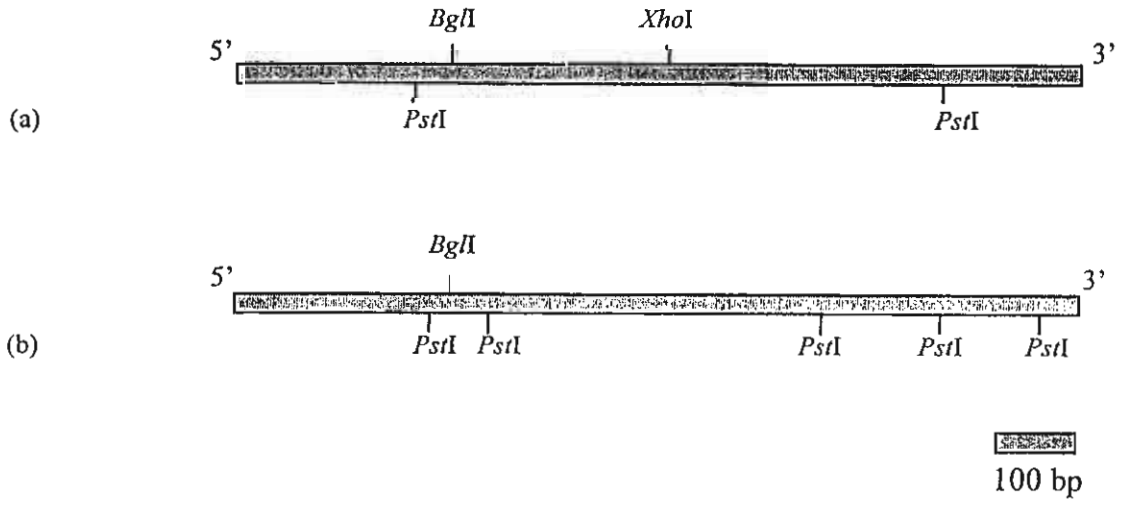

Figure 4. Partial restriction enzymes map of the flagellin genes of *B. pseydimallei* (a) and of *B. cepacia* (b).

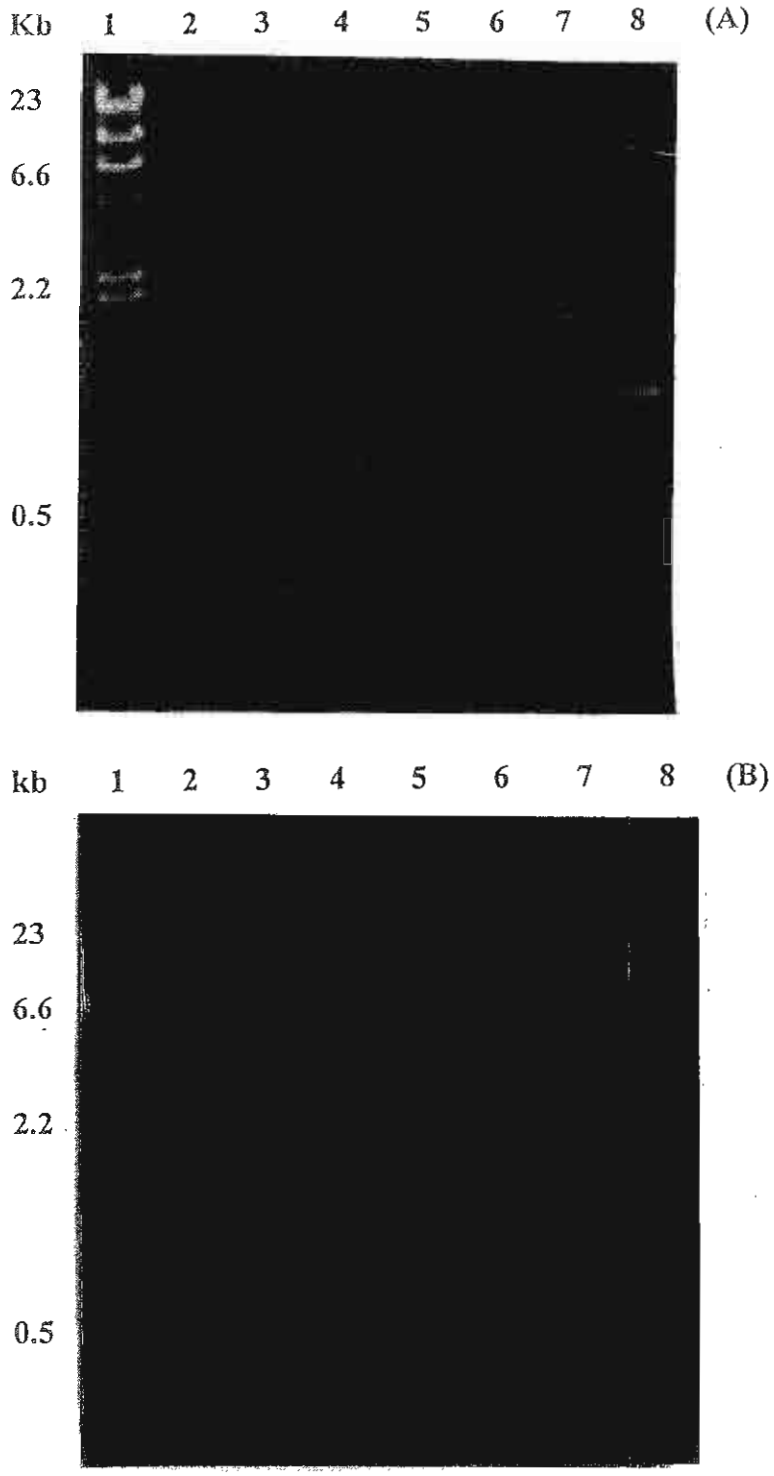

Figure 5. Ethidium bromide staining of an 1% agarose gel electrophoresis of the PCR-restricted PstI pattern (A) and PCR-restricted XhoI pattern (B) of *B. pseudomallei* NF 10/38 (lane 2), NF 47/38 (lane 3), NF 105/37 (lane 4), NF 154/37 (lane 5) and *B.cepacia* DMS 2555 (lane 6), DMS 3027 (lane 7) and undigested PCR-amplified product (lane 8). Lane 1 is a standard size marker.

เนื้อหางานวิจัย

DIFFERENTIAL ANALYSIS OF THE *RPO*B GENE OF MYCOBACTERIA AND PATHOGENIC BACTERIA BY PCR-RFLP TECHNIQUE

ABSTRACT

We have analysed the *rpo*B gene of *M.tuberculosis* and compared to other 13 mycobacteria and 11 pathogenic bacteria using PCR-RFLP technique. Results demonstrated that the 5'portions (region A, nucleotides 36-1,073 and region B, nucleotides 1,024-1,666) and the 3'portion (region F, nucleotides 2,856-3,645) of the *rpo*B gene were conserved among the *M.tuberculosis* complex (*M.tuberculosis, M.bovis* BCG and *M.africanum*). Due to the absence of these fragments in many bacteria, they could be used to distinquish the *M.tuberculosis* complex from other mycobacteria (*M.avium, M.chelonae, M.duvalli, M.gordonae, M.kansasii, M.phlei, M.scrofulaceum, M. smegmatis, M.vaccae and M.intracellulare)* and respiratory bacteria (*S.aureus, S.epidermidis, S.pyogenes, S.viridans, H.influenzae, Ps.aeruginosa, Ps.pseudomallei, K.pneumoniae, C.diptheriae, N.meningitidis and E.coli). M.microti*, one of the *M.tuberculosis* complex, differed from the group by giving PCR products only from regions B and F. PCR-RFLP patterns of region C (nucleotides 1,615-2611) using *Sal*I and region E (nucleotides 2,335-2962) using *Alu*I confirmed the specificity of the *rpo*B gene of *M.tuberculosis* complex. Two regions, D (nucleotides 2,335-2,492) and H (nucleotides 3,935-4,357) were highly conserverd in all mycobacteria and respiratory bacteria tested as they yielded similar sizes of PCR products. However, *Bsa*AI digestion pattern of the amplified region D could be used to differentiate the *M.tuberculosis* complex from *H.influenzae, E.coli* and other mycobacteria while *Hae*III digestion could be used to separate *M.tuberculosis* complex from *M.avium, M.gordonae, M.kansasii, M.scrofulaceum* and *M.intracellulare*. As a whole, we suggested that the PCR-RFLP pattern of regions D and E were useful for identification of the *rpo*B gene of the *M.tuberculosis* complex, hence assuring the detection of rifampicin resistance genotype in unpurified specimens.

INTRODUCTION

Tuberculosis (TB) is a serious health problem in Thailand. The rising incidence rate is believed to accompany the acquired immunodeficiency syndrome (AIDS) epidemic. Moreover, the high occurrence of drug resistant *M. tuberculosis*, the causal agent,

contributes considerably to an ineffective treatment of TB and therefore compromises the Tuberculosis Control Programme. Rifampicin is one of the first-line drugs used in TB chemotherapy. Its mechanism of action is the inhibition of bacterial ribonucleic acid (RNA) polymerase by forming a complex with the holoenzyme, thus blocking the transcription step[1-3]. Resistance to rifampicin is one major cause of ineffective treatment. Alterations in the hot spot region covering 25 amino acids in the *rpoB* gene (encoding for the β subunit of RNA polymerase) are responsible for the rifampicin resistant phenotype[3-8]. Several laboratories have developed rapid techniques for detection of points mutations lying in this region including heminested PCR[9], PCR-single stranded conformation polymorphism[6], dideoxy fingerprinting[10] specific probe hybridization (the line probe assay kit)[11] and DNA sequencing[12,13]. However, these techniques require a separate bacterial species identification procedure. Current identification of *Mycobacterium tuberculosis* complex utilizes at least two target genes for example; genes encoding 65kDa, 19kDa, MPB 64 and the MPB 70 proteins[14].

In this study, we have examined the specificity of each DNA fragment in the *rpoB* gene of *M. tuberculosis* comparing to other mycobacteria and respiratory bacteria using PCR-RFLP technique. The analysis was also extended to compare the identity of the *rpo*B gene sequence among 10 different isolates of *M. tuberculosis*. Specific RFLP pattern obtained from this study could be used to differentiate the *M. tubeculosis* complex from 13 other mycobacteria and 11 pathogenic bacteria therefore were useful for *rpo*B gene mutation analysis especially when the samples were unsure of the purity.

MATERIALS AND METHODS

Bacterial strains

The standard strain *M. tuberculosis* H37Rv and 13 other mycobacteria were generously obtained from Dr. Prasit Palittapongarnpim Department of Microbiology, Faculty of Science, and Dr. Angkana Chaiprasert Department of Microbiology, Faculty of Medicine, Mahidol University. Eleven other bacteria (Table 1) were obtained kindly from Dr. Roongnapa Prachaktam and Dr. Poonpilas Hongmanee, Department of Medicine, Ramathibodi Hospital.

Table 1 Mycobacteria and bacteria used in this work.

Mycobacteria species	Bacteria species
Mycobacterium tuberculosis H37Rv	*Corynebacterium diptheriae*
Mycobacterium africanum	*Escherichia coli*
Mycobacterium avium	*Haemophilus influenzae*
Mycobacterium bovis BCG	*Krebsiella pneumoniae*
Mycobacterium chelonae	*Neisseria meningitidis*
Mycobacterium duvalii	*Pseudomonas aeruginosa*
Mycobacterium gordonae	*Pseudomonas pseudomallei*
Mycobacterium intracellulare	*Staphylococcus aureus*
Mycobacterium kansasii	*Staphylococcus epidermidis*
Mycobacterium microti	*Streptococcus pyogenes*
Mycobacterium phlei	*Streptococcus viridans*
Mycobacterium scrofulaceum	
Mycobacterium smegmatis	
Mycobacterium vaccae	

DNA extraction

The chromosomal DNA was prepared by the enzymatic method. Briefly, a loopful of mycobacteria grown on Lowenstein-Jensen (LJ) medium was transferred to a 1.5 ml microcentrifuge tube containing 200 µl of 10 mM Tris HCl pH8.0, 1 mM EDTA (1XTE) buffer and heated at 80°C for 30 min. The solution was incubated at 37°C overnight with 200 µl of 10 mg/ml lysozyme. Then 75 µl 10% SDS/10 mg/ml proteinase K solution was added and the mixture was incubated at 65°C for 10 min. 50 µl of 5 M NaCl was then mixed and followed by 50 µl of cetylpyridinium chloride. The solution was further incubated at 65°C for 10 min. The DNA was extracted by chloroform/isoamyl alcohol and precipitated with 0.7 volume isopropanol. The DNA pellet was washed with 70% ethanol before air-drying and resuspending in 20 µl 1XTE buffer.

PCR-RFLP analysis

Nine regions covering the entire *rpoB* gene of *M. tuberculosis* were amplified by using primers designed from the *rpoB* gene sequence of *M. tuberculosis* H37Rv strain as shown in Table 2. Nucleotide sequences of the *rpo*B gene were based on Miller's report in 1994[8]. Regions D and E were amplified using different downstream primers (DD, for region

D and ED for region E) but the same upstream primer (DU). Generally, the amplification was carried out for 30 cycles in a total volume of 50 µl consisting of 50 ng genomic DNA, 10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.1% Triton X-100, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 10 pmoles of each primer and 1 unit of *Taq* polymerase. The cycles of reaction were denaturated at 94°C, 1 min, annealing at the appropriate temperature for 1 min, and extension at 72°C for 1 min (10 min in the last cycle). The amplified PCR products were analyzed by electrophoresis in 1.5% agarose gel.

Table 2. List of primers, oligonucleotide sequences and their melting temperatures.

Primer	Length (base)	Oligonucleotide sequence (5'→3')	Melting temperature (T_m, °c)	Nucleotides of amplified regions in the *rpo*B gene (region)
AU	19	CGATGAACTCAACGACTTG	56	36-1,073 (A)
AD	18	TTCATGCGACCAACGAGA	54	
BU	20	GTGGGCACCGCTCCTCTAAG	66	1,024-1,666 (B)
BD	20	GCTGTGCAGCGTCTTGTCGG	66	
CU	20	GGGTGTACTTCGACGAGACC	64	1,615-2,611 (C)
CD	19	TACGGCGTTTCGATGAACC	58	
DU	21	TGGTCGCCGCGATCAAGGAGT	68	2,335-2,492 (D)
DD	20	TGCACGTCGCGGACCTCCAG	68	
ED	20	CGCAGCTCCATCCCGGTGCC	72	2,335-2,962 (E)
FU	20	GAGCACGACGACGCCAACCG	68	2,856-3,645 (F)
FD	20	GCAACTCGTCCTCGTCCTGG	66	
GU	20	CCTGCTGGTGGCCATCATGC	66	3,216-4,028 (G)
GD	20	CAACAGGCCCTGCAGCTCGG	68	
HU	20	GCTGCCCGACGAACTGCTCG	66	3,935-4,357 (H)
HD	20	CGGCCGACGGTGTCATCGGA	68	
IU	20	GGTTCGGGGAGATGGAGTGC	68	4,261-4,652 (I)
ID	20	GACGTCGAGCACGTAACTCC	68	

Enzyme digestion

The amplified PCR products were digested with various kinds of restriction endonucleases (Table 3). The digestion reaction was performed in a total volume of 20 µl consisting of 2 µg DNA, 1X reaction buffer, and 10 units restriction endonuclease. The mixtures were incubated at the appropriate manufacturer-recommended temperature for 2 h before being electrophoresed on 1.5% agarose gel. The positive control for the *Bsa*AI

digestion was the amplified PCR product from a part of the *rpoB* gene containing the *Bsa*AI cutting site while for the *Alu* I digestion, pUC18 was used.

Table 3. Amplified regions and the restriction endonucleases used for species differentiation.

Region	Nucleotide regions in the *rpo*B gene	Size of PCR product (bp)	Restriction endonuclease
A	36 - 1,073	1,028	*Alu* I, *Eco*R V, *Mbo* I, *Msc* I, *Sac* II, *Sf* I, *Sph* I, *Stu* I, *Taq* I
B	1,024 - 1,666	644	*Alu* I, *Hae* III, *Msp* I, *Sal* I, *Taq* I
C	1,615 - 2,611	997	*Bcl* I, *Msc* I, *Sal* I
D	2,335 - 2,492	160	*Apa* I, *Bsa*A I
E	2,335 - 2,962	628	*Alu* I, *Cla* I, *Hae* III, *Msp* I, *Pvu* II
F	2,856 - 3,645	789	*Hae* III, *Msc* I, *Rsa* I
G	3,216 - 4,028	812	*Cla* I, *Sph* I, *Xho* I
H	3,935 - 4,357	423	*Bcl* I, *Bgl* II, *Msc* I, *Rsa* I, *Sal* I
I	4,261 - 4,652	392	*Alu* I, *Msp* I, *Pst* I

RESULTS AND DISCUSSION

We have used the PCR-RFLP technique to analyse the *M.tuberculosis rpo*B gene. The patterns and product sizes generated after restriction enzyme digestion of *rpo*B gene fragments could help distinquish different species of mycobacteria and bacteria, although many of them were similar.

Our data showed that region A (nucleotides 36-1,073), region B (nucleotides 1,024-1,666) and region F(nucleotides 2,856-3,645) were conserved to the *M. tuberculosis* complex due to the presence of expected DNA fragments in *M. tuberculosis, M. bovis* BCG *and M. africanum* (Tables 4 and 5). Regions A, B and F could not be amplified from other mycobacteria except for *M.microti* which gave PCR products only from regions B and F. To examine if the PCR product obtained from regions B was belonging to the *M.tuberculosis* complex, we further digested this fragment (644 bp) with *Sal*I and compared the pattern of digestion with the control *M.tuberculosis*. *Sal*I digestion of region B from *M.tuberculosis* complex and *M.microti* gave rise to specific products of 247 bp and 397 bp fragments. However, this enzyme could not be used to separate the *M.tuberculosis* complex from *M. microti*.

Region D of the *rpo*B gene covering the hot spot of mutations is always the target for analysis of rifampicin resistant genotype. This region exhibited high homology in DNA sequences among mycobacteria and other bacteria whereas the 5' and 3' regions of the *rpo*B gene seemed to be conserved only among mycobacteria, especially those belonging to *M. tuberculosis* complex. Unfortunately, PCR-RFLP pattern of region D investigated so far could not discriminate the *M.tuberculosis* complex from most pathogenic bacteria. Therefore, the combination of results from other regions of the *rpo*B gene is necessary. In this work, nested PCR of regions D and E was performed in order to eliminate 11 respiratory bacteria in the amplified DNA. Patterns of digestion of region D products excluded other mycobacteria, *H.influenzae* and *E.coli* from the *M. tuberculosis* complex. PCR-RFLP patterns of region E generated by *Hae*III should separate *M. tuberculosis* complex from *M.avium, M.gordonae, M.kansasii, M.scrofulaceum* and *M.intracellulare* (Figures 1and 2). To avoid false negative results, two additional control experiments were performed. In the first experiment, DNA extract of *M.tuberculosis* was added to the tested DNA and reamplified with the same set of primers. This experiment should confirm whether the PCR condition was optimal or not. The second experiment was conducted to test the quality of DNA templates by reamplifying with different set of primers. In both cases, specific PCR product were detected indicating the appropriate PCR condition (data not shown).

Regions D, E, H and I were more conserved among mycobacteria whereas regions D and H were conserved in both mycobacteria and bacteria. The results of PCR amplification were summarized in Tables 4 and 5.

Although we could not distinguish *M. tuberculosis* from the *M. tuberculosis* complex due to a very high homology in gene sequences among them, tuberculosis occured in human are mainly caused by *M. tuberculosis.* RFLP patterns tested by *Alu*I, *Mbo*I *Taq*I for region A, *Sal*I, *Hae*III *Taq*I for region B, *Cla*I, *Msp*I, *Pvu*II for region E, *Hae*III *Rsa*I for region F, and *Bcl*I, *Bgl*II, *Msc*I, *Rsa*I, *Sal*I for region H did not show any variations among 10 isolates of *M.tuberculosis.*

In conclusion, the PCR-RFLP patterns of the *rpoB* gene obtained in this study could rapidly differentiate the *M.tuberculosis* complex from 13 other mycobacteria and 11 pathogenic bacteria. This data should therefore be useful as a confirmation that DNA fragments amplified from the *rpo*B gene belongs to the *M. tuberculosis* complex prior to mutation analysis.

REFERENCES

1. McClure WR and Cech CL. On the mechanism of rifampin inhibition of RNA synthesis. J Biol Chem.1978;253(24):8949-56.
2. Ute IL and Guido RH. On the binding of rifampin to the DNA-directed RNApolymerase from *Escherichia coli*. Eur J Biochem.1973;38:336-45.
3. Williams DL, Waguespack C, Eisenach K, *et al.* Characterization of rifampicin resistance in pathogenic mycobacteria. Antimicrob Agents Chemother. 1994; 38 (10):2380-6.
4. Ovchinnikov YA, Monastyrskaya GS, Guriev SO, *et al.* RNA polymerase rifampicin resistance mutation in *Escherichia coli*: Sequence changes and dominance. Mol Gen Genet. 1993;190:344-8.
5. Jin DJ and Gross CA. Mapping and sequencing of mutation in the *Escherichia coli rpoB* gene that lead to rifampiçin resistance. J Mol Biol. 1988;202:45-58.
6. Telenti A, Imboden P, Marchesi F, *et al.* Detection of rifampicin-resistance in *Mycobacterium tuberculosis*. Lancet. 1993;34:647-50.
7. Kapur V, Li LL, Iordanescu S, Hamrick MR, *et al.* Characterization by automated DNA sequencing of mutations in the gene (*rpoB*) encoding the RNA polymerase β subunit in rifampicin-resistant *Mycobacterium tuberculosis* strains from New York City and Texas. J Clin Microbiol 1994;32(4):1095-8.
8. Miller LP, Crawford JT, Shinnick TM. The *rpoB* gene of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 1994;4:805-11.
9. Whelen AC, Felmlee TA, Hunt JM, *et al.* Direct genotypic detection of *Mycobacterium tuberculosis* rifampin resistance in clinical specimens by using single tube heminested PCR. J Clin Microbiol 1995; 33:556-561.
10. Felmlee TA, Liu Q, Whelen AC, *et al.* Genotypic detection of *Mycobacterium tuberculosis* rifampin resistance: comparison of single-strand conformation polymorphism and dideoxy fingerprinting. J Clin Microbiol 1995; 33:1617-1623.
11. Matso-Bernard P, Vrioni G and Marinis E Characterization of rpoB mutations in rifampicin resistant clinical *Mycobacterium tuberculosis* isolates from Greece. J Clin Microbiol 1998; 36:20-23.
12. Ohno H, Koga H, Kohno S, *et al* Relationship between rifampin MICs for and rpoB mutations of *Mycobacterium tuberculosis* strains isolated in Japan. Antimicrob Agents Chemother 1996; 40:1053-1056.

13. Kapur V, Li LL, Iordanescu S, *et al.* Characterization by automated DNA sequencing of mutations in the gene (*rpo*B) encoding the RNA polymerase β subunit in rifampicin-resistant *Mycobacterium tuberculosis* strains from New York City and Texas.J Clin Microbiol 1994; 32:1095-8.

14. Mustafa AS, Ahmed A, Abal AT *et al.* Establishment and evaluation of a multiplex polymerase chain reaction for detection of mycobacteria and specific identification of *Mycobacterium tuberculosis* complex. Tubercle and Lung Disease 1995; 76:336-343.

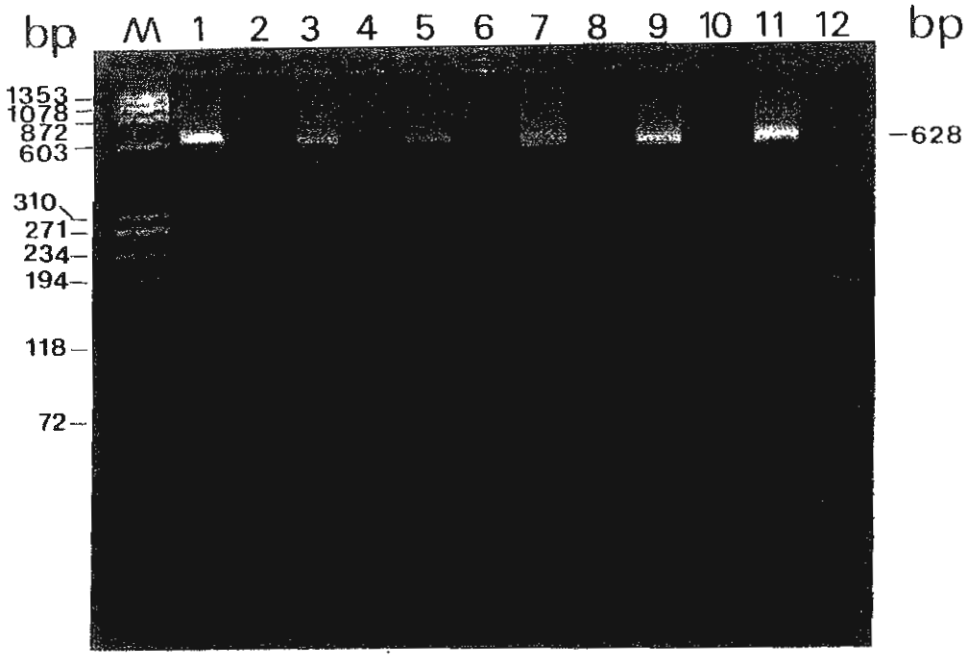

Figure 1. Digestion of region E of the *rpo*B gene with *Hae*III. Even and odd numbers indicate uncut and cut PCR products respectively. Amplified products of 628 bp from *M.tuberculosis* (Lanes1,2), *M.avium* (Lanes 3,4), *M.gordonae* (Lanes 5,6), *M. intracellulare* (Lanes 7,8), *M.kansasii* (Lanes 9,10) *and M.scrofulaceum* (Lanes11,12) generated distinct patterns of different sizes after cutting with *Hae*III.