CHAPTER IV

DISCUSSION

I: Cloning and characterization of Major Royal Jelly Protein 1 and 2 (MRJP1 and MRJP2 from *A. cerana*)

The full length of major royal jelly protein (MRJP) 1 and MRJP2 cDNAs of *Apis cerana* were first obtained using RT-PCR approach. The forward primer, designed from conserved nucleotide and amino acid sequences of AmMRJP1-5 of *A. mellifera*, and oligo (dT) reverse primer were used for synthesizing a fragment of AcMRJP1 and AcMRJP2 cDNAs from total cDNA prepared from toal RNA extracted from head of *A. cerana* nurse honeybee by reverse transcription. Two amplification fragments (1,500 bp and 1,600 bp) corresponding sizes of AmMRJP1 and AmMRJP2 cDNAs were isolated. Characterization of these fragments by digestion with several restriction enzyme revealed restriction pattern related to the restriction map from *A. mellifera* MRJPs cDNA. These results suggested that these fragments might be AcMRJP1 and AcMRJP2.

The amplified fragments were then cloned and sequenced. Sequence analysis of the 1,500 bp PCR fragments demonstrated an ORF of 1,302 nucleotides encoding 433 amino acid residues (Accession no. AF525776). Sequence analysis of the 1,600 bp PCR fragments demonstrated an ORF of 1,392 nucleotides encoding a polypeptide of 463 amino acid residues (Accession no. AF525777). The two sequences were blasted against data in GenBank. The result showed that these cDNAs of two protein were highly homologous to those of AmMRJP1 cDNA (93% nucleotide identity and 90% deduced amino acid similarity) and AmMRJP2 cDNA (92% nucleotide identity and 86% deduced amino acid similarity), respectively which suggested a relative close relationship between MRJP1 and MRJP2 from two different species.

The differences were found among the nucleotide sequences of AcMRJP1 cDNAs amplified by *Taq* and *Pfu* DNA polymerase. It is possible that these differences could be an error introduced during PCR amplification because the thermostable DNA polymerase *Taq* DNA polymerase enzyme used in PCR

amplification lacks 3'-to 5' –exonuclease or proofreading activity and thus, the nuclotide misincorporation might occur. Therefore the *Pfu* DNA polymeras was used in other experiments throughout this study.

The predicted MW without their signal peptide of AcMRJP1 and AcMRJP2 were 46.7 and 50.6 kDa, respectively whereas Srisuparbh *et al.*, (2003) reported the MW of native AcMRJP1 and AcMRJP2 proteins as 50 and 55 kDa, respectively. This result was correspond with Schimitzova *et al.*, (1998) which reported the predicted MW of AmMRJP1 and AmMRJP2 were 46.8 and 48.9 kDa and the native formed of AmMRJP1 and AmMRJP2 as 55 and 49 kDa, respectively. The MW of native proteins were expected to be larger than that of predicted MW because these proteins are glycoproteins. The attachment of the sugar chain to the protein can alter the appearance size of this proteins on SDS-PAGE (Schmitzova, *et al.*, 1998). The number of putative *N*-glycosylation site of AcMRJP1 and AcMRJP2 in *A.cerana* were 3 and 2 sites corresponded with those AmMRJP1 and AmMRJP2 in *A. mellifera*, respectively. The *O*-linked glycosylation site was not found in any AcMRJPs like in *A. mellifera* (Schmitzova, *et al.*, 1998).

All members of the MRJP family share an *N*-terminal hydrophobic sequence that functions as a cleavable signal peptide as well as putative *N*-linked glycosylation sites, suggesting that these proteins are secreted from the cell. The predicted signal peptidase clevage site of *N*-terminal amino acid sequence of AcMRJP1 and AcMRJP2 was SILRGESLNKS and AIIRQNSAKNLEN, respectively. Comparison of the *N*-terminal amino acid sequence of AcMRJP1 and AcMRJP2 obtained from this study with those previously reported in *A. mellifera* species, the predicted N-terminal amino acid sequence of AcMRJP1 was differ only one amino acid residue where compared with N-terminal amino acid sequence of *A. mellifera* MRJP1 (NILRGESLNKS). The predicted N-terminal amino acid sequence of AcMRJP2 were differ in five amino acid residues with those in *A. mellifera* that possed AIVRENSPRNLEK. It seems likely that this difference in N-terminal amino acid residue may originate from the species differences.

At the time when the work was in progress, Su *et al.*, (2004) reported the cloning and sequence analysis of MRJP1 cDNA from *Apis cerana cerana*. A cDNA library was constructed from eight-day-old worker heads of *Apis cerana cerana*. A

probe derived from part of *Apis cerana* genomic AcMRJP3 segment was used to screen this library. The cDNA sequence analysis indicated an ORF of 1,302 bp, which predicts a protein of 434 amino acid residues. The nucleotide homology was 93.78% between *Apis mellifera* MRJP1 and *Apis cerana* MRJP1 while the homology was 99.36% between *Apis cerana cerana* in China and *Apis cerana india* in Thailand. This result confirmed in molecular level that *Apis cerana cerana* and *Apis cerana india* had a common ancestor since the relationship between *Apis cerana* and *Apis mellifera* was more far.

Comparison of the AcMRJP1 cDNA sequence in this study with AcMRJP1 deduced from three clones from EST library (Srisuparph *et al.*, 2003) showed 14 mismatches. Ten of these did not cause amino acid replacement. Nevertheless, non–synonymous mutations were found from P_{106} to Q, A_{112} to S, G_{138} to V and L_{150} to M by substitutions of C_{317}, G_{334}, G_{413} and C_{448} to A, T, T and A, respectively. Internal peptide sequencing of purified AcMRJP1 supported the existence of Q_{106} and S_{112} for AcMRJP1 in this study. Nevertheless, N_{29} of AcMRJP1 found in both studies was D as revealed by internal peptide sequencing (Srisuparbh *et al.*, 2003). Moreover, comparison of the AcMRJP1 cDNA sequence in this study with MRJP1 from *A. cerana cerana* (Su *et al.*, 2004) showed 6 mismatches. Two of these cause amino acid replacement from N_{222} to Y and S_{252} to N by substitutions of A_{664} and G_{755} to T and A, respectively. This result correlated with the observed in *A. mellifera*, the single nucleotide polymorphism (SNP), were observed in AmMRJP3, AmMRJP4 and AmMRJP5 (Albert and Klaudiny, 2004) suggesting possible allelic variants of AcMRJP1.

The identical consensus polyadenylation signal sequence was observed almost at the same position in these cDNA sequences. The putative single (AATAAA) and multiple (AATAAATAAAATAAA) polyadenylation signals are found at 14 nucleotides upstream from the poly (A) tail of AcMRJP1 and AcMRJP2 cDNAs. The latter also contain a consensus AATAAA at 73 bp upstream from the multiple polyadenylation signal sequence. Sequence and the position of overlapping polyadenylation signal in AcMRJP2 were identical with those in AmMRJP3 (Klaudiny *et al.*, 1994) and AcMRJP3 (Srisuparph *et al.*, 2003) but different from AmMRJP2 (Klaudiny *et al.*, 1994).

The interesting character of the AcMRJP2 protein compared to AmMRJP2 is the direct repeat region located at C-terminal region (amino acid residue 423 to 457). The repeat unit consists of tandemly arranged NQKNN pentapeptide unit encoded by AATCAGAA(A/G)AATAAC, pentadecanucleotide. This motif appears in 6 copies making this part of the sequence nitrogen-rich domain. The nucleotide and deduced amino acid sequence of the repeat region was searched in the GenBank and the SwissProt databases. No region of significant identity was detected suggested this decapeptide motif is unique to AcMRJP2. The feature of the repeat regions is the appearance of polar uncharged, asparagine (N), glutamine (Q) and positively charge lysine (K) residue.

Bilikova *et al.,* (1999) illustrated that the single protein band of purified AmMRJP2 analyzed by SDS–PAGE was composed of at least 8 different isoelectricfocusing variants of pI 7.5 - 8.5. Schmitzova *et al.*, (1998) examined nucleotide differences of 4 and 2 isoforms of AmMRJP3 and AmMRJP5 and indicated that polymorphism of these proteins is related with length variability of repetitive regions among honey bee individuals in the colony. Therefore, intra– and inter–colonial variability of an AcMRJP2 gene in different populations of *A. cerana* in Thailand should be further examined.

In *A.cerana*, at least six types of the cDNA encoding MRJP protein in the MRJP family were reported, so far. Five of them (AcAMRJP1-5), were isolated by RT-PCR, and their nucleotide sequences were already submitted to the GanBank database (Imjongjirak *et al.,* 2005, Srisuparbp *et al.,* 2003, Cenphakdee, 2004). Comparison of the amino acid seqeunces deduced form the nucleotide sequences of these five AcMRJP clones revealed high degree of homology. Multiple alignments of deduced amino acid sequence of AcMRJP1-5 revealed some common structural feature: four conserved cysteines typically found across different families of MRJPs; the presence of several blocks of conserved amino acid sequence (PYPDWS, DCSGIVS, RLWVLDS, NLYYSP and LYYVNT); and a highly hydrophilic character mainly in their C-termini. This indicates similarity in tertiary structures and the possibility of similar biological function.

Comparison of the deduced AcMRJP1 and AcMRJP2 with those of the peptides in MRJP family from previously submitted to the Genbank database revealed

that AcMRJP1 and AcMRJP2 exhibited a significant degree of homology to the peptide in this family. The deduced amino acid of AcMRJP1 revealed all characteristics of MRJP peptide family including the four conserved cysteine residues which were also found conserved at identical positions in other MRJP family. This suggests that these cysteine residues might be involved in the structure and the conformation of MRJP peptide.

It could be presumed that AcMRJPs have nutritional function in honeybee larval food. The analysis of their amino acid composition showed that four of them (MRJP1, MRJP2, MRJP5 and MRJP6) contained a high amount of essential amino acids (45%-51.9%), comparable with other nutritional proteins [casein (49.1%), chicken ovalbumin (51.6%) and quail ovalbumin (48.8%) (Schmitzova *et al.*, 1998)], MRJP3 and MRJP4 possess a high amount of some essential amino acids; MRJP3 Arg (5.6%), Lys (6.6%) and MRJP4 Leu (8.6%), Val (7.7%).

The amino acids composition of AcMRJPs, and their dominant content in RJ, indicated that they together represent a balanced mixture of the amino acid essential for nourishing both honeybee larvae and the queen. Some of MRJPs (MRJP3 and MRJP4) that contain lower amounts of essential amino acids and exist in RJ with low content may play other roles in the honeybee physiology. Kucharski *et al.*, (1998) reported that MRJPs may had an other physiological function since AmMRJP1 was also found to expressed in honeybee brain.

Recently, Fontana *et al.*, (2004) reported four antimicrobial peptides purified from Royal Jelly of honeybees, by using reverse phase-HPLC and sequenced by using Q-Tof-MS/MS: PFKLSLHL-NH2 (Jelleine-I), TPFKLSLHL-NH2 (Jelleine-II), EPFKLSLHL-NH2 (Jelleine-III), and TPFKLSLH-NH2 (Jelleine-IV). The peptides were synthesized on-solid phase, purified and subjected to different biological assays: antimicrobial activity, mast cell degranulating activity and hemolysis. The Jelleines-I-III presented exclusively antimicrobial activities against yeast, Gram+ and Gram- bacteria; meanwhile, Jelleine-IV was not active in none of the assays performed. These peptides do not present any similarity with the other antimicrobial peptides from the honeybees; they are produced constitutively by the workers and secreted into Royal Jelly. Comparison to the MRJP1 of *A. mellifera* and *A. cerana*, it revealed that the last nine amino acid residues of MRJP1 peptide exhibited a significant degree of

homology to the Jelleine-II, antimicrobial peptides (TPFKISIHL-NH_2) from *Apis mellifera* (Fontana *et al.*, 2004). Therefore, the biological function of the peptides encoded by AcMRJP1 and AcMRJP2 cDNAs cloned in our laboratory needs to be determined by further analysis such as biological activity assay of the peptides expressed from AcMRJP1 recombinant cDNAs.

The origin of the MRJP family has been studied in *A. mellifera*. AmMRJP family showed similarity with yellow protein of *Drosophila melanogaster*. In structural features, four cysteines conserved in MRJP were also found in yellow protein (Albert, *et al.*, 1999a). It seems that MRJPs diverged from yellow protein to gain a novel nutritional function in honeybee. Recently, *A. mellifera* yellow protein was identified (Albert and Klaudiny, 2004).

The result of genetic distance and phylogenetic relationships of AcMRJP show that the average sequence divergence between the same MRJP family from different bee species (AcMRJP1-AmMRJP1 and AcMRJP2-AmMRJP2) was lower than that between divergence calculated from pairs of different protein family within the same species (AmMRJP1-AmMRJP2 and AcMRJP1-AcMRJP2). The sequence divergence of nucleotides and deduced amino acid were used to construct a bootstrapped NJ tree. The result revealed close relationships between AcMRJP1- AmMRJP1 and AcMRJP2-AmMRJP2 from different bee species. This result suggested gene duplication process (Mitsuo *et al.*, 2001). This result was consistent with Albert *et al.*, (1999a) who determined evolutionary relationships of AmMRJP families and reported that family variants of MRJP genes resulted from near–simultaneous gene duplication, with MRJP4 possibly being the earliest divergence within these gene families.

II: Cloning and characterization of MRJP1 and MRJP2 genes from *A. cerana*

In this study, polymerase chain reaction was exploited for the amplification of the AcMRJP1 and AcMRJP2 gene fragments of *A. cerana*. The specific and degenerated oligonucleotides primer were designed based on known cDNA sequences of AcMRJP1 and AcMRJP2 and were used to amplified the AcMRJP1 and AcMRJP2 genes on *A. cerana* genomic DNA. The genomic organizations of these genes were characterized and the putative promoter sequence and cis-acting regulatory elements

that may exist in the 5' upstream region of AcMRJP1 and AcMRJP2 genes were identified.

The AcMRJP1 gene sequences span over 3,663 bp in length. The intron-exon organization of AcMRJP1 gene was performed by comparison with the AcMRJP1 cDNA. Six exons and five introns were found in the gene. The sequence of AcMRJP1 gene was corresponded to the AcMRJP1 cDNA sequence. The exon-intron boundary was detected in the genomic fragment of AcMRJP1. The splice junction conformed to the splice donor and acceptor consensus sequence and followed the 'GT-AG rule' (Breathnach and Chambon, 1981). The presence of five introns in AcMRJP1 gene is consistent with the primary structure of the gene encoding AmMRJP1 in *A. mellifera* that have recently been characterized (Malecova *et al.*, 2003).

The AcMRJP2 gene sequences span over 3,963 bp in length. The intron-exon organization of AcMRJP2 gene was performed by comparison with the AcMRJP2 cDNA. Six exons and five introns were found in the gene. The nucleotide sequence of the AcMRJP2 gene and the AcMRJP2 cDNA was differed from each other at a single position. In AcMRJP2 cDNA sequence, ^{8}A was originating from primer sequence and should be corrected to ^{8}G from the AcMRJP2 genomic sequence. Therefore the encoded amino acids should be change from ^{3}Arg to ^{3}Lys. The exon-intron boundary was detected in the genomic fragment of AcMRJP2. The splice junction conformed to the splice donor and acceptor consensus sequence and followed the 'GT-AG rule'.

With respect to codons, there are three possible locations for an intron within a gene: between codons (phase 0), between the first and the second nucleotides of a codon (phase 1), and between the second and the third letters of a codon (phase 2). The distribution of intron phases is unequal, with, typically, a bias in favor of phase 0 introns (Fedorov *et al.*, 1992; Long *et al.*, 1995). More introns exist between codons (phase 0) than between the first and the second bases (phase 1) or between the second and the third base (phase 2) within the codon.

Introns 2, 3, and 5 of AcMRJP1 and AcMRJP2 lying between two codons (phase 0) whereas introns 1 interrupting a codon between the first and second base (phase 1); and intron 4 interrupting a codon between the second and third base (phase 2).

When AcMRJP1, AcMRJP2 and AmMRJP1 genes were compared, the similarities in the conservation of the position of the intron among these MRJP genes were observed. This result suggested that they had occurred through gene duplication events.

The putative transcription start site of the AcMRJP1 that was predicted by the computer software was identical to that in *A. mellifera* (Malecova *et al.*, 2003). However, the computer software predicted different transcription start site for AcMRJP2 gene. The precise transcription start sites of both AcMRJP1 and AcMRJP2 could be confirmed in future study by another method such as primer extension.

Regulatory regions of eukaryotic genes are mostly located in immediate proximity to the transcriptional start point. In this study, the length of approximately 0.6 and 0.8 kb of 5' upstream regions of AcMRJP1 and AcMRJP2 were obtained and characterized. The presence of common conserved putative binding sites for dead ringer (Dri) and ultraspiracle (USP) transcription factors were detected in the AcMRJP1 and AcMRJP2 genes.

Dri is a member of a family of DNA-binding proteins that share an A/T rich interaction domain (ARID). ARID containing proteins are responsible for both positive and negative transcriptional regulation. They are also known to play regulatory roles in cell proliferation, differentiation, and development (Kortschak *et al.*, 2000).

USP is a member of a family of nuclear hormone receptors, a family of ligand-modulated transcription factors that regulate cell differentiation and development as well as homeostasis and reproduction (Sergaves, 1991). It was reported, that two active juvenile hormone (JH) isoforms (JH acid and JH methylester), found in *D. melanogaster*, bind specifically to *Drosophila* USP transcription factor (Jones and Sharp, 1997). This binding induces conformational changes and homo-oligomerization of USP. This may confer different activities to USP influencing transcription. Based on these results it was proposed that JH is a physiological regulator of USP action (Jones and Sharp, 1997). There was found to be a correlation between JH titer in honeybee heamolymph, hypopharyngeal gland (HG) size and task performance of honeybee worker (Huang *et al.*, 1991, 1994). The USP is a candidate

to test for a role as the connecting element to the JH response of HG in honeybee. The negative result in functional AmMRJP1 promoter assay in S2 insect cell line derived from *Drosophila* supports the idea of a highly organ and physiological stage specific regulation development (Kucharski and Maleszka, 2002).

Both AcMRJP1 and AcMRJP2 contained a single USP binding site at the 5' UTR immediately following the TATA box. Malecova *et al.*, (2003) reported two ultraspiracle transcriptional factor (USP–TF, GGTCA) binding sites in AmMRJP1 and only one binding site in AmMRJP2 – 5 immediately downstream from the predicted TATA box.

In summary, this study showed that the AcMRJP1 and AcMRJP2 genes of *A. cerana* have similar structures. The putative binding sites for regulatory factors that were identified within the 5' upstream sequence of all the AcMRJP genes suggested that these genes encoding MRJP peptide family might be under the control of similar regulatory mechanisms. The information of this study would constitute a basis for understanding of molecular mechanism that regulates the expression of these genes.

III: Expression of AcMRJP1 and AcMRJP2 in *E. coli* expression system

Expression of AcMRJP1 and AcMRJP2 in the *E. coli* system was performed to investigate whether AcMRJP1 cDNA encoded the major royal jelly protein and to confirm the identity of these proteins. The AcMRJP1 and AcMRJP2 without signal peptide sequence was cloned into the pGEM®-3Zf(+), pTrcHis2c and pET17b expression vector and the expression was induced by IPTG.

The pGEM®-3Zf(+) (lac promoter) was initially attempted to express the AcMRJP1 and AcMRJP2 in *E. coli* system by fusion with lacZ and expressed in DH5α as an expression host cell. The result revealed the protein band of 43 and 55 kDa corresponded to the estimated molecular fusion protein of 402.2 and 54.4 kDa of AcMRJP1 and AcMRJP2, respectively. This expression construct was performed to confirm the translation terminating mutations in the AcMRJP1 amplified by *Taq* DNA polymerase, resulting in the expression of a truncated AcMRJP1 protein. The expressed protein band of AcMRJP1 and AcMRJP2 was not stable as it was decreased after 1 hour of induction.

In comparison with the control experiment, without IPTG induction, the expressed AcMRJP1 was also detected in the cell lysate suggested that the presence of basal level (leaky) expression might be occurred.

Use of an IPTG-inducible expression vector requires transformation of an *E. coli* strain that overexpresses the *lac* repressor. The *lac* repressor prevents transcription of foreign gene sequences in the absence of inducer. Since most IPTG-inducible expression plasmids are present in high copy number (30-600), excess repressor is needed to prevent titration of the protein and subsequent basal level (leaky) expression. The *lacI*^{q} allele, which overproduces the *lac* repressor by tenfold when present in single copy (Calos, 1978), is most commonly used to overexpression this protein. However, if the foreign protein is especially toxic to *E. coli* and very low basal expression levels cannot be tolerated, or if a very high-copy-number plasmid vector is used (e.g., pUC-based plasmids), then a lacI^{q} gene should be cloned into the expression plasmid to ensure tight regulation.

To facilitate purification of the recombinant AcMRJP1 and AcMRJP2, The protein was expressed as a recombinant protein fused to a C-terminal hexahistidine tag, allowing purification by nickel-agarose affinity chromatography. This could be achieved by cloning the gene to translate in frame with the initiation ATG and C-terminal hexahistidine tag in pTrcHis2c expression vector.

Studies on SDS-PAGE of time course analysis of small-scale expression revealed the protein at molecular mass of 50 and 55 kDa which consistent with the calculated molecular weight of 51.3 and 55.2 kDa of recombinant AcMRJP1 and AcMRJP2, respectively. The amount of induced protein may increase according to the time of induction. The production of the recombinant protein was observed at 1 hour after induction and showed highest expression level at 6 hours after induction with IPTG. In comparison with control, without IPTG induction, no the expressed protein was detected in the cell lysate suggested that the expression vectors pTrcHis2c (Trc promoter) used in our studies provide tight control of uninduced expression.

The expressed protein solubility was determined by bacterial lysate centrifugation. After centrifugation, the 50 and 55 kDa of AcMRJP1 and AcMRJP2 protein was present only in pellet fraction of induced sample suggested that the

recombinant protein was expressed predominantly in the pellet fraction as insoluble inclusion bodies which agree with many studies of *E. coli* expression system (Gu *et al.*, 2001; Liao *et al.*, 2003). The expression of foreign proteins at high levels in *E. coli* often results in the formation of cytoplasmic granules or inclusion bodies composed of insoluble aggregates of the expressed protein. Solubilization of the protein from the pellet generally requires exposure to strong chaotropic agents, such as urea (6-8 M) or guanidinium (5-8 M), to detergents, such as SDS, and sometimes nonphysiological pH (Marston, 1986; Marston and Hartley, 1990). Contaminating proteins and nucleic acids can be removed from inclusion bodies fraction by washing the pellet with solutions containing EDTA, detergents, and/or DNase. Usually, the aggregated protein isolated from washed inclusion bodies is ~50-75% pure and consists of denatured monomeric and oligomeric forms held together by disulfide bonds (Light 1985).

The AcMRJP1 and AcMRJP2 recombinant proteins tagged with a hexahistidine at its C-terminus could be isolated by affinity chromatography in a nickel affinity column, under denaturing condition with urea containing buffer. Urea is used to disrupt non-covalent interactions in the inclusion bodies resulting in a more open structure. Purification of hexahistidine tagged recombinant protines, Hi-Trap chelating HP when charged with Ni^{2+} ions, will selectively retain proteins if complex-forming amino acid residues, in particular histidine, are exposed on the surface of the protein, histidine-tagged protiens can be eluted from Hi-Trap chelating HP with buffers containing imidazole.

In present study, the 50 kDa of recombinant AcMRJP1 was elulted by 60 mM imidazole with minor protein band of approximately 38 kDa. The 55 kDa of recombinant AcMRJP2 protein was purified by elution from Ni^{2+} affinity column with about 60 mM imidazole. The histidine tag in this recombinant protein was used for further identification of this purified protein. The purified AcMRJP1 and AcMRJP2 were analyzed by Western blot analysis using anti-His antibody. For AcMRJP1, two positive band were observed at 50 kDa and 38 kDa. For AcMRJP2 one major protein band was observed with molecular weight of 55 kDa.

The expected purified AcMRJP1 band of 50-kDa and unknown band of 38 kDa was detected in SDS and Western blot analysis. This 38-kDa band may be

originated from protease digestion of AcMRJP1 or there are translation initiation ATG. The N-terminal peptide sequencing on these protein bands will give a direct evidence on these bands. However, this confirmatory experiment has not been done.

The third strategy, the expression vector pET17b was used in our studies. Since the low expression level of AcMRJP1 and AcMRJP2 was achieved from pGEM®-3Zf(+) and pTrcHis2c. Therefore, an alternative strategy was used for the expression of AcMRJP1 and AcMRJP2.

To facilitate purification of the recombinant AcMRJP1 and AcMRJP2, fusion of the AcMRJP1 and AcMRJP2 gene with N-terminal hexahistidine-tag coding sequence was performed. This could be achieved by cloning the gene with primer coding for six histidine residues into the original pET17b expression vector, and expressed under the control of T7 promoter as a His-tag fusion protein which can be purified by Ni^{2+} immobilized resin. The presence of T7 promoter and *lac* repressor (*lac* I) in the pET expression vectors permit a tight control of high expression of the target gene.

Time course analysis of IPTG induction revealed the molecular weight of induced protein at 50 and 55 kDa consistent with the calculated molecular weight of 47.9 and 51.7 kDa of recombinant AcMRJP1 and AcMRJP2, respectively. The amount of induced protein increased according to the time of induction.

Comparison of expression level in several *E. coli* host strains; BL21(DE3) pLysS, BL21(DE3) pLysS star and Rosetta (DE3) pLysS, revealed high expression level in Rosetta (DE3) pLysS host cell. This strain supply tRNAs for the codons AUA, AGG, AGA, CUA, CCC and GGA which rarely used in *E. coli* (Brinkmann *et al.*, 1989; Seidel *et al.*, 1992; Kane, 1995; Kurland and Gallant, 1996). It is corresponded to the sequence analysis of AcMRJP1 and AcMRJP2 sequence.

In the AcMRJP1 gene, out of 417 codons, there was ten AGA rare codons (at codon number 39, 40, 72, 101, 174, 274, 328, 361, 374 and 391), nine ATA rare codons (at 83,160, 187, 195, 358, 363, 368, 401 and 420), four CTA rare codons (at 45, 71, 175 and 178), and four CCC rare codons (at 98, 138, 250 and 357).

In the AcMRJP2 gene, out of 455 codons, there were seven AGA rare codons (at codon number 42, 43, 127, 138, 226, 329 and 103), six ATA rare codons (at 73, 86,163, 187, 195 and 379), four CTA rare codons (at 74, 97, 178 and 181), and three CCC rare codons (at 59, 225 and 250).

The production of both recombinant proteins was observed at 1 hour after induction and showed highest expression level at 4 hours after induction with 0.4 mM IPTG. In comparison with control, without IPTG induction, no the expressed protein was detected in the cell lysate suggested that the expression vectors pET17b (T7 promoter) used in our studies provide tight control of uninduced expression.

The expressed protein solubility showed that the 50 and 55 kDa of AcMRJP1 and AcMRJP2 was present only in pellet fraction of induced sample suggested that the recombinant proteins were expressed predominantly in the pellet fraction as insoluble inclusion bodies. It has been well documented that foreign proteins overproduced in *E. coli* are not always soluble, and often aggregate to form inclusion bodies. The mechanisms of inclusion bodies formation are not well understood, however, small changes in structure as well as an number of factors, i.e. growth temperature, culture media, expression vectors, and / or the gene sequence, can have profound effects on the tendency of the expressed protein to precipitate *in vivo*.

In present study, the 50 kDa protein was purified by successive elution from Ni^{2+} affinity column with about 250 mM imidazole. The purified rAcMRJP1 and rAcMRJP2 were analyzed by Western blot analysis. The N-terminal amino sequence of the purified AcMRJP1 and AcMRJP2 was determined by automated Edman degradation. The sequence of twenty residues at the N-terminus was ASHHHHHHSILRGESLNKSL and ASHHHHHHAIIRQN(S/N)(S/A)KNL which was identical to expected AcMRJP1 and AcMRJP2, respectively. The lack of an *N*–terminal methionine (M) which is often removed from expressed proteins in the *E. coli* expression system (Hirel *et al.*, 1989). Base on the N-terminal amino sequence, Western blot analysis and the expected mass of recombinant protein according to the amino acid sequence, it was concluded that the recombinant proteins in the present study were AcMRJP1 and AcMRJP2, respectively.

Judova *et al.*, (1998) cloned AmMRJP1 cDNAs into pQE32 vectors and expressed recombinant constructs in *E. coli* M15[pREP4]. The highest production of recombinant proteins was observed at 1 hour after 1 mM IPTG induction. The rAmMRJP1 protein was dominantly expressed in the insoluble form. Only 0.6 mg of purified rAmMRJP1 were obtained from 1-liter culture.

Bilikava *et al.*, (1999) cloned AmMRJP2 cDNAs into pQE–30 vectors and expressed recombinant constructs in *E. coli* M15[pREP4]. The highest production of recombinant proteins was observed at 5 hour after 1 mM IPTG induction. The rAmMRJP2 protein was dominantly expressed in the insoluble form. Low amount of rAmMRJP2 were obtained from 1 liter culture. Unlike rAcMRJP1, rAmMRJP2 was degraded by proteases of the host cells with prolonged culture period.

In summary, our results indicated successful isolation and expression of AcMRJP1 and AcMRJP2 in the *E. coli* expression system. Relatively high amount of recombinant proteins were obtained from a small scale culture. Larger quantity of rAcMRJP1 and rAcMRJP2 can be scaled up and used for further studies on antiallergic, antioxidative and/or antitumor activities of these recombinant proteins.

IV: Expression of AcMRJP1 in plant expression system

In this study, the potato and rice were used as the model plants for the experiments to test whether the AcMRJP1 can be transformed into and expressed in potato and rice. The rice callus and leaf disk of potato have already been shown to be compatible with *Agrobacterium*-mediated transformation (Hiei *et al.*, 1994 and Dietze *et al.*, 1995).

For potato transformation, *Solanum tuberosum* L. cultivar Atlantic was used. Atlantic is one of the main varieties culture in Thailand for processing (potato chips). Leaf disk transformation by *A. tumefaciens* was used as a general method for potato transformation. The transformation process was performed according to the method described by Dietze *et al.*, 1995 which works excellently with the variety *S. tuberosum* cv. Desiree. The transformed shoots were obtained in 8 weeks after co-cultivation process. However, the efficiency of potato cultivar Atlantic transformation was less than in cultivar Desiree (50%) suggested that these cultivars vary in their

physiological and agronomic characteristics and these differences appear to affect the efficiency of *Agrobacterium*-mediated transformation (Rockhold *et al.*, 2001).

The constitutive 35S promoter, tuber specific GBSS and patatin gene promoter was used for potato. The CaMV 35S is expressed in most cells of most plants, and in most tissues, although with different efficiency according to plant species. The promoters of the tuber-specific patatin genes and the promoter of GBSS gene allow expression of heterologous proteins in tubers. These promoter has been used successfully as an expression cassette in several potato plants. Many proteins have been successfully express in potato tuber by using these promoter, e.g., seed albumin gene from *Amaranthus* for increasing nutritive value (Chakraborty *et al.*, 2000). As expected in this study, AcMRJP1 was successfully expressed and accumulated in transgenic potato tuber under the control of these promoter.

If the constructed vector was transferred into potato explants, the regenerated plantlet should express the selectable marker, reporter gene and the gene of interest. The kanamycin resistance gene in the T-DNA was used to select the transformed plant and the gus gene was used as a reporter gene to verified the transformation by GUS histochemical test.

The use of *in vitro* grown microtubers has the potential to offer some considerable advantages over soil-grown tuber. First, it is notoriously difficult to assess the precise developmental status of a developing soil-grown tuber, particularly in the early stages (Ross *et al.*, 1994; Sowokinos, 1976). Soil-grown tubers of the same apparent size can show tremendous variability in their metabolic status (Merlo *et al.*, 1993; Sung *et al.*, 1989). By contrast, the speed synchronicity and reproducibility of the location of microtuber initiation in vitro allows the stage of harvesting to be precisely defined. Studies of tuber initiation are beset by similar problems to those described above, and here the use of in vitro grown microtubers has become widely accepted (e.g. Appeldoorn *et al.*, 1997; Hendriks *et al.*, 1991; Visser *et al.*, 1994). It has been shown that the activities of some enzymes important for the tuberization process (e.g. sucrose synthase, EC 2.4.1.13; invertase, EC3.2.1.26; fructokinase, EC 2.7.1.14, hexokinase, EC 2.7.1.1, UDP-glucose pyrophosphorylase, EC 2.7.7.9) show a similar pattern in in vitro microtubers to that found in developing soil-grown tubers (Appeldoorn *et al.*, 1997; Sweetlove *et al.*, 1996). Further, analysis

of the structure, pattern of cell division, protein and starch composition of *in vitro* microtubers has also showed that tissue culture microtubers are similar to soil-grown tuber (Debon *et al.*, 1998; Sweetlove *et al.*, 1996; Visser *et al.*, 1994).

For rice transformation, KDML105 is a photoperiod-sensitive variety, and must be grown under special conditions to obtain seed. The transformation efficiency varies with both plant genotype and with the transformation system used. In general, *japonica* cultivars of rice are more tissue culture-responsive and give higher frequency of transformation than *indica* cultivars. Like many *indica* varieties KDML105 was view as recalcitrant to both tissue culture and transformation. Transformation efficiency obtained in this work was 7.5% which has been as high as those reported for rice (Hiei *et al.*, 1994, Khanna and Raina, 1999, Zhang *et al.*, 1997 and Rashid *et al.*, 1996). Most importantly, all transformed rice plants were morphologically normal and fertile.

The double-strand caulimovirus family has provided the single most important promoter for transgene expression in plants: the cauliflower mosaic virus (CaMV) 35S promoter. The CaMV 35S is expressed in most cells of most plants, and in most tissues, although with different efficiency according to plant species. Many proteins have been successfully expressed in potato tuber by using this promoter. As expected in this study, AcMRJP1 was successfully expressed and accumulated in transgenic rice leaf under the control of this promoter.

If the constructed vector was transferred into rice explants, the regenerated plantlets should express the selectable marker, reporter gene and the gene of interest. The hygromycin resistance gene in the T-DNA was used to select the transformed plant and the gus gene was used as a reporter gene to verified the transformation by GUS histochemical test.

Although, histochemical staining of GUS activity in calli, roots of transformant showed uniform blue staining, the leaves segments showed blue staining at only cut surface. This result was similar to the studies of Dong *et al.*, (1996) which determined that this resulted from limited penetration of X-Gluc substrate.

The integration of AcMRJP1 was confirmed by PCR analysis and histochemical assay of GUS reporter gene. The PCR analysis showed expected band at 1.3 kb suggested that AcMRJP1 gene was present in the transformants. The absence of detectable PCR product amplification band of the AcMRJP1 gene and did not have GUS expression in some potato plants, despite kanamycin selection, could reflect development of antibiotic resistance or may equally reflect genetic instability after integration.

The advantage of assaying for the presence to the foreign gene in transformed tissue with PCR is that a positive results can signify that the full sequence is present (as well as being a quick test). On the other hand, information of the number and site of integration is obtainable only with the Southern blotting approach. Therefore, a further study may be concentrated on transformed status by estimation of copy number and site of integration by Southern blot analysis.

Recently, the other homologous MRJP1 from *A. mellifera* have been expressed in tobacco. Judova *et al.*, 2004 reported the expression of AmMRJP1 in plants, plasmid carrying the expression cassette composed of CaMV35S protmoter, cDNA encoding MRJP1 with its native signal peptide, and nos3' as transcription terminator in binary vector pBin19 was prepared. The plasmid was introduced into tobacco (*Nicotiana tabacum* L. cv. Wi38) plants by *Agrobacterium tumefaciens*-mediated transformation. Transgenic F1 and F2 generation was grown from the seeds of the primary obtained transgenic tobacco plants. The integration of AmMRJP in transformed tobacco was determined by PCR analysis.

The expression of AcMRJP1 was confirmed by RT-PCR analysis with AcMRJP1 specific primer. The RT-PCR amplification of RNA provide a mean to easily detect the presence of the AcMRJP1 transcripts (Tang *et al.*, 1999; Cary *et al.*, 2000). To confirm that there was no DNA contaminating the RNA samples all cDNAs from the reverse transcription of potato mRNA were amplified with GBSS primers. Due to the presence of introns in the GBSS coding regions, these primers were expected to generated a 173 bp product from mRNA, where as a product of 358 bp would be expected upon amplification of genomic DNA. All samples gave one product of 173 bp indicating that there was no DNA contamination. The result suggested that the PCR products observed with both primers were the result of

amplification of their respective gene transcripts. Different expression levels of AcMRJP1 were detected among the transformed plant. This phenomenon has been observed in several transformed plant species and could reflect the influence of many factors such as genetic background, position effects, copy number, transgenes rearrangements and the physiological and developmental status of the tissue.

Transformed potato plants were obtained which expressed the *A. cerana* gene encoding AcMRJP1 driven by the GBSS promoter, B33 patatin promoter and 35S promoter as shown by the presence of mRNA and protein for AcMRJP1 in tubers of transformed plant. AcMRJP1 is synthesized as a 48.3 kDa protein in both transformed potato and rice, corresponding to the expected size of the matured AcMRJP1. This accumulation was consistent with the construct. AcMRJP1 protein appeared to be synthesized as a nonglycosylated form and is not translocated into the ER. Targeting of the AcMRJP1 was cytosol.

The stabilizing effect of the short KDEL sequence was also demonstrated for other cytoplasmically expressed scFv fragments and may be based on C-terminal protection against proteolytic degradation (Schouten *et al.,* 1996; Schouten, 1997). Additional strategies for stabilizing cytosolically expressed antibodies have been reported. Schouten *et al.,* (1996, 1997) have demonstrated that the addition of a C-terminal KDEL peptide increased the cytosolic expression level of an scFv fragment. Maximum expression levels of up to 0.2% of the total soluble protein were measured, whereas the same scFv fragment without KDEL did not show any accumulation in the cytosol of transgenic plants. They suggested that short polypeptides such as the KDEL sequence may protect the scFv antibody from proteolytic degradation. Without signal peptide sequences, the recombinant antibodies are accumulated in cytosol which is the most difficult compartment to obtain high expression level although several successful examples in the cytosol has been reported (Owen *et al.,* 1992).

Immumunoblot analysis using a rabbit antiserum against the recombinant AmMRJP1 (Judova *et al.,* 1998), five immunopositive MRJP1 bands were detected on blot from plants expressing MRJP1 in higher level. Their molecular weight were: tM1-54 kDa, tM2-51 kDa, tM3-46.5 kDa, tM4-45 kDa, tM5-43 kDa. The band tM3 corresponded with its molecular weight 46.5 kDa to theoretical molecular weight of complete mature secreted honey bee MRJP1, which was calculated from its amino

acid sequence to be 46.8 kDa (Schmitzova *et al.*, 1998). The number and molecular mass profile of proteins in these bands suggest pot-translational processing of expressed plant AmMRJP1.

The evidence of honeybee protein expression in the experimental model of potato and rice plant in general suggests that it is possible to express the honeybee protein in plants. However, the utilization of this protein to influence the nutritional value of some crop plants would require increasing of its expression level in plant cells. Protein accumulation can be improved by post-translation targeting to subcellular compartments such as the ER, the vacuole or the chloroplast (Schouten *et al.*, 1996; Gruber *et al.*, 2001; Dierych, 1997). Limitations of nuclear genome transformation methods might be overcome by the site-specific introduction of genes into plastid (e.g. chloroplast) chromosomes (Gogorad, 2000; Daniell *et a.*, 2000; He *et al.*, 2001).

Data on physiological properties of RJ proteins, such as suppression of allergic reactions by RJ (Oka *et al.*, 2001), or antihypertensive activity of bioactive RJ peptides (Matsui *et al.*, 2002) broaden their potential application in pharmacy and indicate their natural function in honeybee evolution, where they could play role of inducer of defense mechanisms during larval development. This recent discovery, that RJ proteins may have important physiological functions as ingredients of functional foods showed potential of transgenic plants for their heterologous expression.

In summary, this technique will allow the genetic improvement of diverse varieties of rice, as well as many aspects of the molecular biology. Experiments are underway to utilize these procedures to study the expression of various genes driven by various promoters in rice, regulation of gene expression, functional analysis of genes, gene tagging and introduction of agronomically useful gene into plants.