Int. J. Indust. Entomol. Vol. 6, No. 1, 2003, pp. 93~97

International Journal of *Industrial Entomology*

Molecular Cloning of a cDNA Encoding Putative Calreticulin from the Silkworm, *Bombyx mori*

Seong Ryul Kim, Kwang Sik Lee, Iksoo Kim[1], Seok Woo Kang[1], Si Kab Nho[2], Hung Dae Sohn and Byung Rae Jin*

College of Natural Resources and Life Science, Dong-A University, Busan 604-714, Korea.
[1]Department of Sericulture and Entomology, National Institute of Agricultural Science and Technology, Suwon 441-744, Korea.
[2]College of Agriculture and Life science, Kyungpook National University, Taegu 702-701, Korea.

(Received 31 December 2002; Accepted 11 March 2003)

We describe here the cloning of a cDNA encoding putative calreticulin (CRT) from the silkworm, *Bombyx mori*. The CRT cDNA comprised of 1,194 bp encoding 398 amino acid residues. *B. mori* CRT has a HDEL sequence at the end of the C-domain. The *B. mori* CRT showed 88% protein sequence identity to the *G. mellonella* CRT, 71% to *A. aegypti* CRT, and 63% to *H. sapiens* CRT. Phylogenetic analysis revealed that the deduced amino acid sequences of the *B. mori* CRT formed a highly inclusive subgroup with other insect CRTs. Northern blot analysis exhibited an expression of the *B. mori* CRT gene in the fat body, evidencing the fat body as a major site for CRT synthesis.

Key words: Insect, Silkworm, *Bombyx mori*, Calreticulin, cDNA sequences, Phylogeny

Introduction

Calreticulin (CRT) is a major intracellular calcium binding protein, found in a wide range of species and in all nucleated cell types. CRT has previously been known as high-affinity calcium binding protein, calregulin, Erp60, CRP55, CAB-63 and CaBP3 (Michalak *et al.*, 1992). CRT is a lectin and it is now established that this characteristic is important for its function as a molecular chaperone in the endoplasmic reticulum (ER) (Krause *et al.*, 1997; Helenius *et al.*, 1997). The human CRT has been well characterized (Coppolino and Dedhar, 1998).

*To whom correspondence should be addressed.
College of Natural Resources and Life Science, Dong-A University, Busan 604-714, Korea. Tel: +82-51-200-7553; Fax: +82-51-200-7594; E-mail: brjin@mail.donga.ac.kr

In insects, CRT has also identified in some insect species such as *Galleria mellonella* (Choi *et al.*, 2002), *Anopheles gambiae* (Francischetti *et al.*, 2001), *Drosophila melanogaster* (Smith, 1992) and *Aedes aegypti* (Valenzuela *et al.*, 2002). Recently a report from *G. mellonella* larvae suggested that CRT is involved in non-self recognition in invertebrate cellular defense reactions (Choi *et al.*, 2002). As indicated, molecular information on insect CRT is still limited to a few insect species. Thus, our objective in initiating this study is to illustrate the structure of the silkworm, *Bombyx mori*, CRT cDNA. In this study, we report the cDNA sequence for CRT homologue from *B. mori*.

Materials and Methods

cDNA library screening, nucleotide sequencing and data analysis

A cDNA library was constructed using poly(A)+ mRNA isolated from the whole body of *B. mori* larvae by Uni-ZAP XR vector and Gigapack III Gold Packing Extract (Stratagene). Sequencing of randomly selected clones harboring cDNA inserts was performed to generate the expressed sequence tags (ESTs). For DNA sequencing, plasmid DNA was extracted by Wizard mini-preparation kit (Promega). Sequence of each cDNA clone was determined using an automatic sequencer (model 310 Genetic Analyzer; Perkin-Elmer Applied Biosystems, Foster City, CA). The sequences were compared using the DNASIS and BLAST programs provided by the NCBI (www.Ncbi.nlm.nih.gov). GenBank, EMBL and SwissProt databases were searched for sequence homology using a BLAST algorithm program (www.Ncbi.nlm.nih.gov/BLAST).

MacVector (ver. 6.5, Oxford Molecular Ltd.) was used to align the amino acid sequences of CRTs. Including seventeen GenBank-registered CRTs amino acid sequences, phylogenetic analysis was performed using PAUP* (Phylogenetic Analysis using Parsimony) version 4.0 (Swofford, 2000). The tree was obtained by bootstrap analysis with the option of heuristic search (1,000 replications). The accession numbers of the sequences in the GenBank are: *Bombyx mori* (this study), *Galleria mellonella* (AB 063250), *Drosophila melanogaster* (X64461), *Anopheles gambiae* (AF457551), *Aedes aegypti* (AF466603), *Homo sapiens* (M84739), *Rattus norvegicus* (NM022399), *Mus musculus* (NM007591), *Oryctolagus cuniculus* (J05138), *Rana rugosa* (D78589), *Danio rerio* (AF195882), *Necator americanus* (AJ006790), *Meloidogyne incognita* (AF 402771), *Dirofilaria immitis* (AF052978), *Ricinus communis* (U74630), *Arabidopsis thaliana* (AY086745), *Zea mays* (2205314A), and *Hordeum vulgare* (L27349). Helical and non-helical regions of *B. mori* CRT were predicted by the Garnier-Robson method (Garnier *et al.*, 1978).

RNA isolation and Northern blot analysis

B. mori larvae were dissected under the Stereo-microscope (Zeiss, Jena, Germany), individual samples such as fat body, midgut, silk gland, ovary, and testis were harvested, and washed twice with PBS (140 mM NaCl, 27 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.4). Total RNA was isolated from fat body, midgut, silk gland, ovary, and testis of the *B. mori* using the Total RNA Extraction Kit (Promega, Madison, WI). Total RNA (10 μg/lane) from the *B. mori* was separated on glyoxalation gel (McMaster and Carmichael, 1977), transferred onto a nylon blotting membrane (Schleicher & Schuell, Dassel, Germany) and hybridized at 42°C with a probe in a buffer containing 2 × PIPES, 50% formamide, 1% sodium dodecyl sulphate (SDS) and blocking agent (Boehringer Mannheim, Mannheim, Germany). The probe used to detect the CRT gene was 1677 bp for *B. mori* CRT cloned in this study and labeled with [α-^{32}P] dCTP (Amersham, Arlington Heights, IL) using the Prime-It II Random Primer Labeling Kit (Stratagene, La Jolla, CA). After hybridization, the membrane filter was washed three times for 30 min each in 0.1% SDS and 0.2 × SSC (1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate) at 65°C, and finally exposed to X-ray film.

Results and Discussion

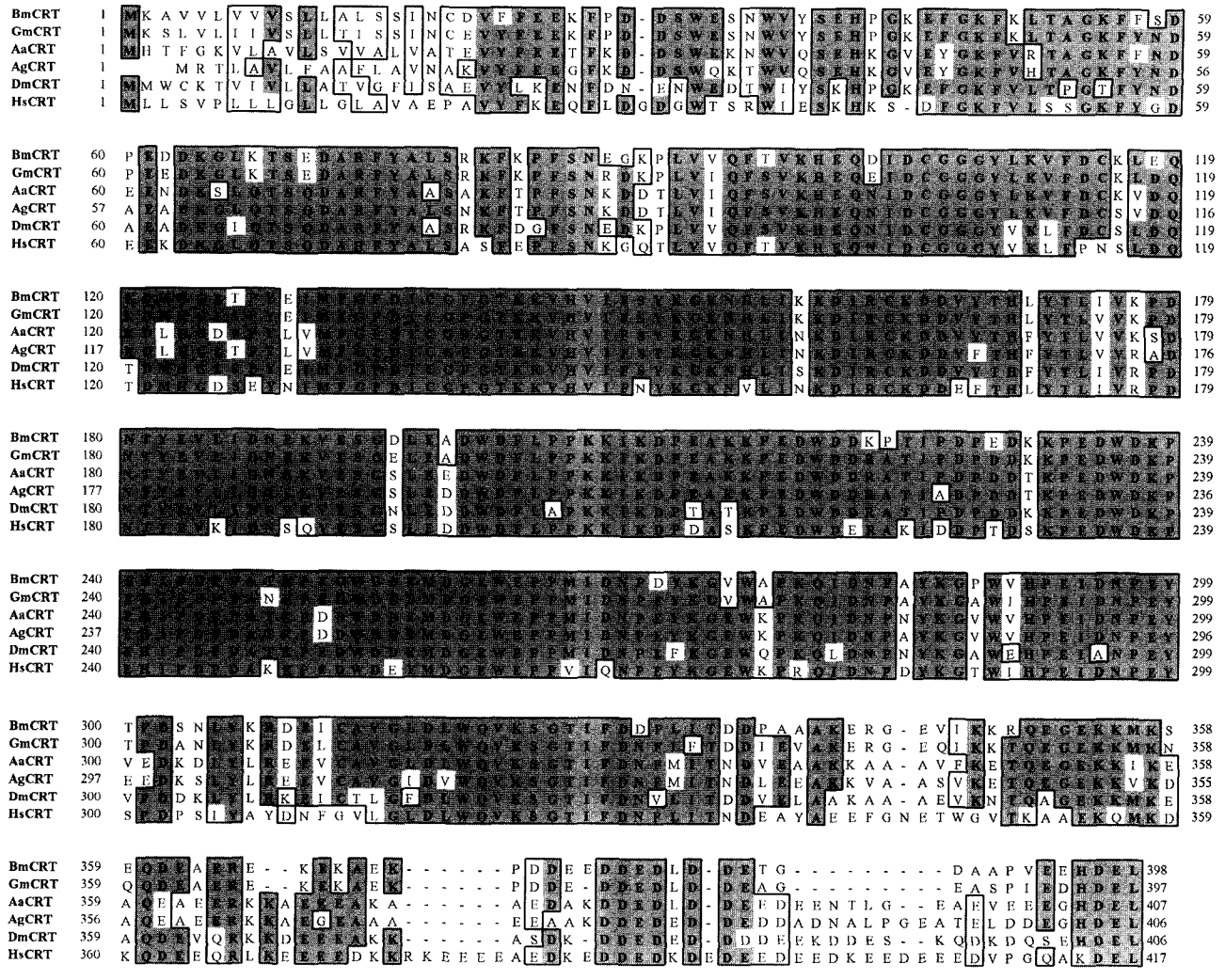
We screened the cDNA library constructed using whole bodies of *B. mori* larvae. Sequencing of randomly selected clones harboring cDNA inserts was performed to generate

```
-11                                                    GTGAATCGAAA
  1 ATGAAAGCTGTTGTACTGGTCGTTGTCAGCTTGCTGGCATTATCTTCGATAAATTGTGAT
  1 M  K  A  V  V  L  V  V  V  S  L  L  A  L  S  S  I  N  C  D

 61 GTATTCTTTGAGGAGAAGTTTCCTGATGATTCGTGGGAAAGCAATTGGGTGTACAGTGAG
 21 V  F  F  E  E  K  F  P  D  D  S  W  E  S  N  W  V  Y  S  E

121 CACCCAGGAAAAGAATTTGGAAAGTTTAAATTGACTGCTGGAAAGTTCTTCAGCGACCCA
 41 H  P  G  K  E  F  G  K  F  K  L  T  A  G  K  F  F  S  D  P

181 GAGGATGACAAAGGTTTGAAAACCTCTGAAGATGCGAGGTTCTATGCTCTCTCCCGTAAG
 61 E  D  D  K  G  L  K  T  S  E  D  A  R  F  Y  A  L  S  R  K

241 TTCAAACCGTTCAGTAATGAGGGTAAACCCTTGGTAGTCCAGTTCACTGTCAAACATGAA
 81 F  K  P  F  S  N  E  G  K  P  L  V  V  Q  F  T  V  K  H  E

301 CAAGACATTGACTGTGGAGGCGGATACCTCAAGGTCTTTGACTGCAAACTGGAGCAGAAG
101 Q  D  I  D  C  G  G  G  Y  L  K  V  F  D  C  K  L  E  Q  K

361 GACATGCACGGAGAGACTCCATATGAGATAATGTTCGGCCCTGACATTTGTGGTCCGGGA
121 D  M  H  G  E  T  P  Y  E  I  M  F  G  P  D  I  C  G  P  G

421 ACCAAGAAGGTACACGTAATCTTCAGTTACAAGGGTAAGAACCATCTCATCAAAAAGAT
141 T  K  K  V  H  V  I  F  S  Y  K  G  K  N  H  L  I  K  K  D

481 ATCCGCTGCAAAGATGATGTCTACACACATTTGTACACTCTGATTGTGAAACCTGACAAC
161 I  R  C  K  D  D  V  Y  T  H  L  Y  T  L  I  V  K  P  D  N

541 ACCTATGAAGTCCTCATTGACAATGAGAAAGTTGAATCTGGCGACCTAGAGGCAGACTGG
181 T  Y  E  V  L  I  D  N  E  K  V  E  S  G  D  L  E  A  D  W

601 GACTTCCTTCCGAATAAGAAAATCAAGGACCCTGAAGCCAAGAAACCAGAAGACTGGGAT
201 D  F  L  P  P  K  K  I  K  D  P  E  A  K  K  P  E  D  W  D

661 GACAAGCCCACTATTCCAGACCCCGAAGACAAGAAGCCTGAGGATTGGGACAAGCCTGAA
221 D  K  P  T  I  P  D  P  E  D  K  K  P  E  D  W  D  K  P  E

721 CACATTCCAGATCCAGATGCCACCAAACCTGAAGACTGGGATGATGAGATGGACGGAGAA
241 H  I  P  D  P  D  A  T  K  P  E  D  W  D  D  E  M  D  G  E

781 TGGGAGCCACCTATGATTGACAACCCAGACTACAAGGGCGTTTGGGCACCAAAACAGATT
261 W  E  P  P  M  I  D  N  P  D  Y  K  G  V  W  A  P  K  Q  I

841 GATAACCCAGCTTACAAAGGCCCATGGGTCCATCCCGAAATCGACAACCCAGAGTACACT
281 D  N  P  A  Y  K  G  P  W  V  H  P  E  I  D  N  P  E  Y  T

901 CCTGATTCAAACCTGTACAAGCGCGACGAAATCTGCGCTGTCGGTCTCGATTTGTGGCAA
301 P  D  S  N  L  Y  K  R  D  E  I  C  A  V  G  L  D  L  W  Q

961 GTCAAGTCCGGTACCATCTTTGACGACTTCCTTATTACTGACGACCCGGCCGCCGCCAAA
321 V  K  S  G  T  I  F  D  D  F  L  I  T  D  D  P  A  A  A  K

1021 GAGAGGGGAGAGGTCATCAAGAAAAGGCAGGAGGGAGAAAAGAAAATGAAAAGCGAACAA
341  E  R  G  E  V  I  K  K  R  Q  E  G  E  K  K  M  K  S  E  Q

1081 GACGAGGCCGAGAGGGAGAAGGAGAAGGCGGAAAAGCCTGACGATGAGGAAGATGATGAG
361  D  E  A  E  R  E  K  E  K  A  E  K  P  D  D  E  E  D  D  E

1141 GACCTCGACGATGAGACTGGAGACGCAGCGCCTGTGGAGGAACACGATGAGCTGTAGGCT
381  D  L  D  D  E  T  G  D  A  A  P  V  E  E  H  D  E  L  *

1201 GTTGTGATTAATTAGGTGAGTGCCAAGACATTAAGTACCTCGGGGTACGTAGTGTGCGTG
1261 CGTGGAGAATTCTTTTTTTAAAACCTCTCGTTGCGCAATCTTCTACAATTCCAATACTAT
1321 TGTCATTTGTTTTTAGCACTGTGCTGTGTGTTTGGCTCGAATTTTACCGATTCGTAGGGA
1381 CATGTTGTGTGTAAGTAAATATGTGTCCGTTGTTTCGTACTGATATCCGTGTACATGGGG
1441 ATTACGCTCAGAATTGATGTATGTAATTATTTAATGTTTTGTTATCGGTCTGGCACCGTG
1501 TGCTGGGACGTGAATGTAAAGTTTTACTGCACAATAACATTCGATATAGTAGCCCTAATG
1561 TTTTGTATGATTGTGATTGTTATTGTACAGTAAATCTGTCTTCTCATTTTAAGTTGTGAA
1621 ATGAGTCAATAAAATGGAATTAAATTGAAAAAAAAAAAAAAAAAAA
```

Fig. 1. The nucleotide and deduced amino acid sequences of *B. mori* CRT. The start codon of ATG is boxed and the termination codon is shown by asterisk. The polyadenylation signal is underlined.

B. mori ESTs. Of these clones, one had an insert of 1,666 bp, which contained the complete 3' end with a poly(A) tail, but it was not complete at the 5' end. Sequences of the clone exhibited similarity to previously reported CRTs. The nucleotide and deduced amino acid sequences of *B. mori* CRT are presented in Fig. 1. The *B. mori* CRT homologue cDNA sequences comprised of 1,194 bp encoding 398 amino acid residues. The *B. mori* CRT cDNA contained 1,666 nucleotides with a polyadenylation signal AATAAA occurred in the 3' untranslated region at positions 1628-1633.

A multiple sequence alignment of the deduced protein sequence of *B. mori* CRT cDNA with other CRT sequences is shown in Fig. 2. Alignment of the *B. mori* CRT sequences with those of CRTs from several other



(A)

Percent similarity

	1	2	3	4	5	6
1. BmCRT		95	78	78	76	72
2. GmCRT	88		79	78	78	72
3. AaCRT	71	73		90	80	74
4. AgCRT	69	71	84		79	72
5. DmCRT	69	72	71	71		73
6. HsCRT	63	63	64	62	64	

Percent identity

(B)

Fig. 2. Alignment (A) and identity (B) of deduced amino acid sequences of *B. mori* CRT with five CRTs. The identical residues are shown in solid boxes. BmCRT, *B. mori* CRT; GmCRT, *G. mellonella*; AaCRT, *A. aegypti* CRT; AgCRT, *A. gambiae*; DmCRT, *D. melanogaster*; and HsCRT, *H. sapiens* CRT.

species indicates the extent of the identity that exists. The *B. mori* CRT showed 88% protein sequence identity to the *G. mellonella* CRT, 71% to *A. aegypti* CRT, and 63% to *H. sapiens* (Fig. 2B).

The structure of CRT has three structural and function domains: N, P and C-domains (Coppolino and Dedhar, 1998; Corbett and Michalak, 2000). The first third of CRT is referred to as the N-domain (residues 1-180), which is the most conserved domain among CRTs. The middle third, or P-domain (residues 181 – 280), of CRT is Pro-rich segment. This P-domain contains a high-affinity, low capacity Ca^{2+}-binding site. The C-terminal third of CRT is highly negatively charged and contains the low affinity/ high capacity Ca^{2+}-binding site. Interestingly, all insect CRTs have a HDEL sequence at the end of the C-domain, but human CRT has KDEL sequence (Fig. 2A).

The hydropathy plot (Fig. 3A) of *B. mori* CRT by the Kyte and Doolittle method (1982) showed that *B. mori* CRT is hydrophilic except for N-terminal and is similar to insect CRTs such as *G. mellonella* (Choi *et al.*, 2002), *A. gambiae* (Francischetti *et al.*, 2001), *D. melanogaster* (Smith, 1992) and *A. aegypti* (Valenzuela *et al.*, 2002), suggesting that cleavage of the putative signal peptide occurred between Cys^{19} and Asp^{20} in *B. mori* CRT. The mature *B. mori* CRT is predicted to be 379 amino acid residues.

Helical and non-helical regions of *B. mori* CRT were predicted as shown in Fig. 3B. The *B. mori* CRT showed very similar structure when compared to other known CRTs. Furthermore, the *B. mori* CRT is very similar to the

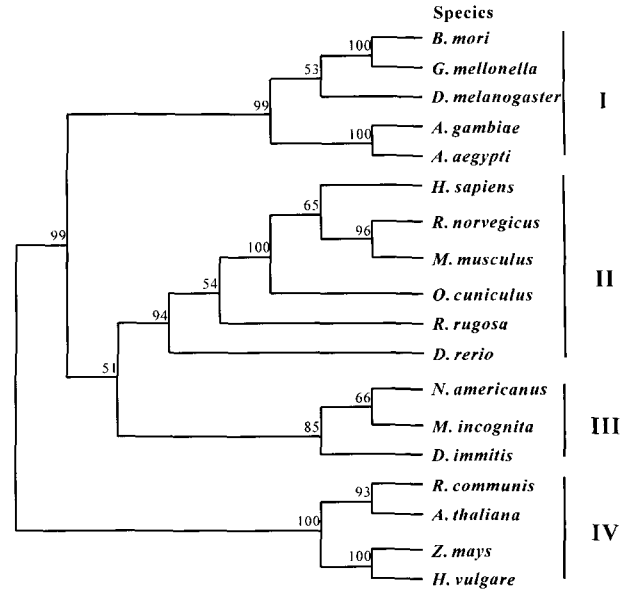


Fig. 4. Phylogenetic tree of aligned amino acid sequences of the *B. mori* CRT and the other known CRTs. The sequences were extracted from; *B. mori* (this study), *G. mellonella* (AB063250), *D. melanogaster* (X64461), *A. gambiae* (AF457551), *A. aegypti* (AF466603), *H. sapiens* (M84739), *R. norvegicus* (NM022399), *M. musculus* (NM007591), *O. cuniculus* (J05138), *R. rugosa* (D78589), *D. rerio* (AF195882), *N. americanus* (AJ006790), *M. incognita* (AF402771), *D. immitis* (AF052978), *R. communis* (U74630), *A. thaliana* (AY086745), *Z. mays* (2205314A), and *H. vulgare* (L27349). The tree was obtained by bootstrap analysis with the option of heuristic search and the numbers on the branches represent bootstrap values for 1,000 replicates.

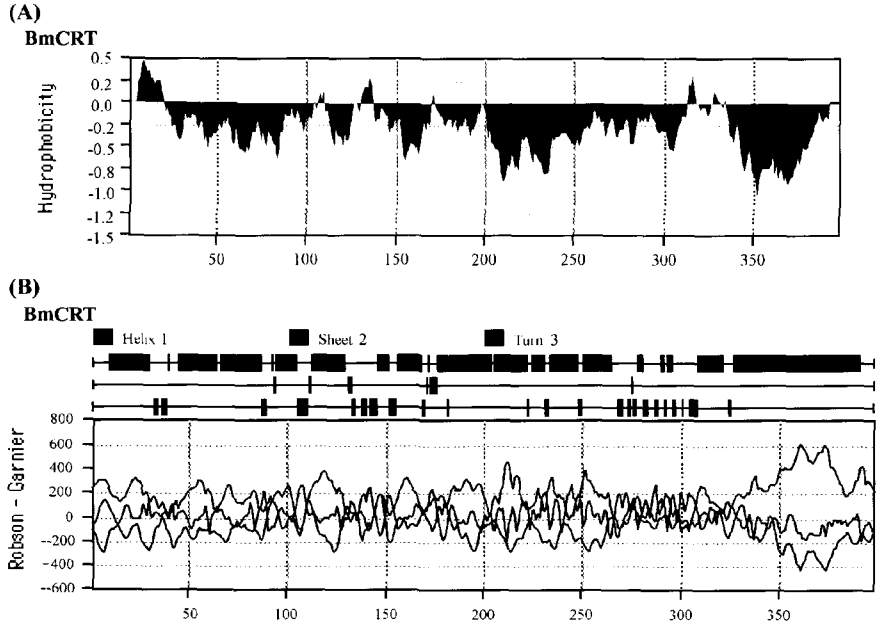


Fig. 3. The hydropathy profile (A) and predicted helical regions (B) of *B. mori* CRT.

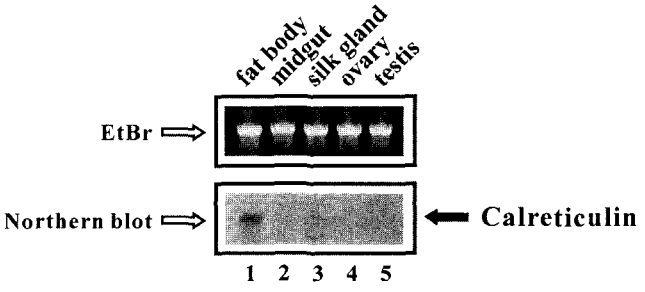


Fig. 5. Northern blot analysis of the *B. mori* CRT messages. Total RNA was isolated from fat body (lane 1), midgut (lane 2), silk gland (lane 3), ovary (lane 4) and testis (lane 5) of *B. mori* larvae. The RNA was separated by 1.0% formaldehyde agarose gel electrophoresis (upper panel), transferred onto a nylon membrane, and hybridized with the appropriate radiolabelled probe (lower panel). *B. mori* CRT messages are indicated by arrow on the right side of the panel.

H. sapiens CRT predicted to be globular in structure with eight anti-parallel âsheets and a helix-turn-helix motif at its N-terminus (Coppolino and Dedhar, 1998).

A phylogenetic analysis using deduced amino acid sequences of known CRT genes revealed four major groups. The deduced amino acid sequences of the *B. mori* CRT formed a highly inclusive subgroup with other insect CRTs (Fig. 4).

To confirm the expression of the *B. mori* CRT gene at transcriptional level, Northern blot analysis was performed using total RNA prepared from the fat body, midgut, silk gland, ovary and testis of *B. mori* (Fig. 5). A hybridization signal was detected as a single band in mRNA from the fat body. The Northern hybridization exhibited an expression of the *B. mori* CRT gene in the fat body, evidencing the fat body as a major site for CRT synthesis.

In conclusion, we hope that the molecular characterization of *B. mori* CRT in this study will expand the understanding of insect CRTs.

Acknowledgements

This work was supported by a grant from BioGreen21 Program, Rural Development Administration, Republic of Korea.

References

Choi, J. Y., M. M. A. Whitten, M. Y. Cho, K. Y. Lee, M. S. Kim, N. A. Ratcliffe and B. L. Lee (2002) Calreticulin enriched as an early-stage encapsulation protein in wax moth *Galleria mellonella* larvae. *Dev. Comp. Immunol.* **26**, 335-343.

Coppolino, M. G. and S. Dedhar (1998) Calreticulin. *Int. J. Biochem. Cell Biol.* **30**, 553-558.

Corbett, E. F. and M. Michalak (2000) Calcium, a signaling molecule in the endoplasmic reticulum? *Trends Biochem. Sci.* **25**, 307-311.

Francischetti, I. M., J. G. Valenzuela and J. M. Ribeiro (2001) Towards a catalog for genes and proteins from the salivary gland of the malaria vector, *Anopheles gambiae. GenBank accession number AF457551.*

Garnier, J., D. J. Osguthorpe and B. Robson (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**, 97-120.

Helenius, A., E. S. Trombetta, D. N. Hebert and J. F. Simons (1997) Calnexin, calreticulin and the folding of glycoproteins. *Trends Cell Biol.* **7**, 193-200.

Krause, K. H. and M. Michalak (1997) Calreticulin. *Cell* **88**, 439-443.

Kyte, J. and R. F. Doolittle (1982) A simple method for displaying hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.

McMaster, G. K. and G. G. Carmichael (1977) Analysis of single- and double-stranded nucleic acids on polyacryamide and agarose gels by using glyoxal and acridine orange. *Proc. Natl. Acad. Sci. USA* **74**, 4835-4838.

Michalak, M., R. E. Milner, K. Burns, M. Opas (1992) Calreticulin. *Biochem. J.* **285**, 681-692.

Smith, M. J. (1992) Nucleotide sequence of a *Drosophila melanogaster* gene encoding a calreticulin homologue. *DNA Seq.* **3**, 247-250.

Swofford, D. L. (1990) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4, Sinauer Sunderland, MA.

Valenzuela, J. G., V. M. Pham, M. K. Garfield, I. M. Francischetti and J. M. C. Ribeiro (2002) Toward the sialome of the adult female mosquito *Aedes aegypti. GenBank accession number AF466603.*