Molecular Cloning, Expression Pattern, and Phylogenetic Analysis of the Lysyl-tRNA Synthetase Gene from the Chinese Oak Silkworm Antheraea pernyi

Yan-Qun Liu and Li Qin
Department of Sericulture, College of Bioscience and Biotechnology, Shenyang Agricultural University, China

1. Introduction

Aminoacyl-tRNA synthetases (AARS) are a class of enzymes that charge tRNAs with their cognate amino acids. There are two classes of tRNA synthetases, classes I and classes II, that are distinguished by the architectures of their active-site catalytic cores (Guo et al., 2008). Lysyl-tRNA synthetase (LysRS) is an AARS, a group of ancient proteins known for their critical role in translation. LysRS is a homodimer localized to the cytoplasm which belongs to the class II family of tRNA synthetases. Its assignment to class II AARS is based upon its structure and the presence of three characteristic sequence motifs in the core domain. The catalytic core domain is primarily responsible for the ATP-dependent formation of the enzyme bound aminoacyl-adenylate.

The LysRS gene has been cloned and characterized from various kinds of organisms. In Escherichia coli, there are two distinct LysRSs encoded by two widely separated genes, LysRS and LysRU (VanBogelen et al., 1983; Emmerich & Hirshfield, 1987), while in Campylobacter jejuni only one LysRS is present (Chan & Bingham, 1992). In E. coli, the LysRS gene is expressed constitutively while LysRU can be induced by growth at high temperature (Hirshfield et al., 1981). In human, only one LysRS belonging to the class II family of AARSs is present (Guo et al., 2008). The crystal structure of tetrameric form of human LysRS has been recently determined (Guo et al., 2008). Human LysRS has been shown to be secreted and to trigger a proinflammatory response as a target of autoantibodies in the human autoimmune diseases, polymyositis or dermatomyositis (Park et al., 2005). And, LysRS is required for the translocation of calreticulin to the cell surface in immunogenic death (Kepp et al., 2010). However, none of insect LysRS gene has been characterized to date.

The Chinese oak silkworm, Antheraea pernyi (Lepidoptera: Saturniidae), is one of the most well-known economic insect species used for silk production and insect food source. This insect is known to be successfully domesticated in China around the 16th century (Liu et al. 2010a), and it is commercially cultivated in China, India, and Korea. To isolate the functional genes of A. pernyi, we have constructed a full-length cDNA library (Li et al. 2009). By EST
sequencing, several *A. pernyi* genes encoding important enzymes have been characterized, such as two enolase genes and a lysophospholipase gene (Liu et al. 2010b, 2010c).

In this chapter, we describe the cloning and characterization of the *A. pernyi* LysRS gene from the full-length pupal cDNA library by random EST sequencing. The expression patterns at various developmental stages and different tissues were investigated. Finally, the deduced protein sequence of the LysRS gene from *A. pernyi* and other organisms were used to examine the relationship among these species, and to test the potential use of LysRS protein in phylogenetic study.

2. Materials and methods

2.1 Silkworms and tissues

The *A. pernyi* strain *Shenhuang No. 1* was used in this study. Larvae were reared routinely on oak trees (*Quercus liaotungensis*) in the field. Blood, fat body, midgut, silk glands, body wall, Malpighian tubules, spermaries, ovaries, brain and muscle were taken from silkworm larvae at day 10 of fifth instar. Eggs at day 5, larvae at day 10 of fifth instar, pupae and moths were also sampled. All the samples were immediately frozen in liquid nitrogen and stored at –80°C for later use.

2.2 Cloning of the *A. pernyi* LysRS gene

A full-length pupal cDNA library of *A. pernyi* has been constructed in our lab (Li et al. 2009). An EST encoding LysRS homolog (GenBank accession no. GH335029) was isolated by random EST sequencing. So, the cDNA clone was used to complete the full-length cDNA sequence of the *A. pernyi* LysRS gene.

2.3 Sequence analysis

DNASTAR software (DNASTAR Inc., Madison, WI) was used to identify open reading frame (ORF), deduce amino acid sequence, and predict the isoelectric point and molecular weight of the deduced amino acid sequence. Blast search was performed at http://www.ncbi.nlm.nih.gov/blast/. Conserved Domains was predicted at http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi/. The *in silico* gene expression analysis based on the available EST resources was employed at http://www.ncbi.nlm.nih.gov/Unigene/ESTprofileViewer/.

2.4 Total RNA extraction and first strand cDNA synthesis

Total RNA was extracted by using RNAsimple Total RNA Extraction Kit (TIANGEN Biotech Co. Ltd., Beijing) according to the manufacturer’s instruction. DNase I was used to remove contaminating genomic DNA. The purity and quantity of the extracted RNA was quantified by the ratio of OD260/OD280 by ultraviolet spectrometer. First strand cDNA was generated by using 2 μg of total RNA per sample with TIANScript cDNA Synthesize Kit (TIANGEN Biotech Co. Ltd., Beijing).

2.5 RT-PCR analyses

The cDNA samples were amplified by semi-quantitative PCR method using the gene-specific primer pair LYQ146 (5’-TCCGA GTGGG GAAGA AGTTG-3’) and LYQ147 (5’-TTCAG TCAGT CCTGG TATGT-3’) for the *A. pernyi* LysRS gene, which generated a 322 bp
fragment. An actin gene (GU073316) was used as an internal control, and a 468 bp fragment was amplified using the primer pair LYQ85 (5’-CCAAA GCACCA ACAGA GAGAA GA 3’) and LYQ86 (5’CAAGA ATGAG GGCTG GAAGA GA 3’) (Wu et al., 2010). PCRs were performed with the following cycles: initial denaturation at 95°C for 5 min; followed by 30 cycles of 1 min at 95°C, 30 s annealing at 55°C, 30 s extension at 72°C; and a final extension at 72°C for 10 min. To avoid sample DNA contamination, the negative RT-PCRs control reactions were performed with every total RNA as templates. The amplification products were analyzed on 1.0% agarose gels, purified from the gel, and directly sequenced to confirm the specificity.

2.6 Phylogenetic analysis
The amino acid sequences of LysRS homologs from different organisms were retrieved from GenBank database, and SilkDB database (Duan et al., 2010). Multiple sequence alignments were performed using Clustal X software (Thompson et al. 1997). A phylogenetic tree was constructed by MEGA version 4 (Tamura et al. 2007) using Neighbour-Joining (NJ) method (Saitou and Nei 1987) with bootstrap test of 1000 replications.

3. Results and discussion

3.1 Sequence analysis of the A. pernyi LysRS gene
We identified the A. pernyi LysRS gene from a pupal cDNA library constructed in our lab by random EST sequencing (Li et al., 2009). Based on the cDNA clone Appu0107, we isolated and sequenced a full-length cDNA of the A. pernyi LysRS gene. The cDNA sequence and deduced amino acid sequence of the A. pernyi LysRS gene were shown in Figure 1. The obtained 2136 bp cDNA sequence contains a 5’-untranslated region (UTR) of 70 bp, a 3’ UTR of 292 bp with a polyadenylation signal sequence AATAAA at position 2082 and a poly (A) tail, and an ORF of 1740 bp encoding a polypeptide of 579 amino acids. The LysRS protein has a predicted molecular weight of 65.62 kDa and isoelectric point of 6.1. Blast search revealed that the deduced amino acid sequence of the A. pernyi LysRS gene had 66% identities and 79% positives with that of Homo sapiens LysRS (NP_005539), which belongs to the class II family of AARSs (Guo et al., 2008). Conserved Domains prediction showed that the A. pernyi LysRS protein contained the LysRS class II core domain (Wolf et al., 1999), including the conserved active sites and three characteristic sequence motifs (Desogus et al., 2000). Moreover, the characteristic signature HIGH sequence of LysRS class I was not present. We therefore referred to the protein as LysRS of A. pernyi. This cDNA sequence has been deposited in GenBank under accession no. JF773568.

3.2 Homologous alignment
By searching in database, the A. pernyi LysRS protein homologues were found in various kinds of life organisms, including bacteria, fungi, plants, invertebrates and vertebrates. To assess the relatedness of A. pernyi LysRS to LysRS proteins from other organisms, identities were calculated based on a Clustal alignment including 43 representative LysRS protein sequences (Figures 2 and 3). The other 42 LysRS protein or homologue sequences used in this study were downloaded from GenBank database, with one exception of B. mori LsyRS from SilkDB database (Duan et al., 2010). These protein sequences were from 3 bacteria, 1 fungi, 6 plants, 22 invertebrates and 11 vertebrates. By sequence alignment, the A. pernyi
Fig. 1. The complete nucleotide and deduced amino acid sequence of the *A. pernyi* LysRS gene. (A) cDNA sequence and deduced amino acid sequence of the LysRS gene. The amino acid residues are represented by one-letter symbols. The initiation codon ATG is bolded and the termination codon TAG is bolded and marked with an asterisk. The polyadenylation signals AATAAA are double-underlined. The gene specific primer sequences used in the semi-quantitative RT-PCR experiment are underlined. The cDNA sequence was deposited in GenBank under accession no. JF773568. (B) Conserved domains of *A. pernyi* LysRS determined by http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi/.
Fig. 2. Sequence alignment of LysRS proteins from *A. pernyi* and other organisms. These LysRS proteins were included from *A. pernyi*, *Bombyx mori* (BGIBMGA002984-PA in SilkDB), *Drosophila melanogaster* (NP_572573), *Homo sapiens* (NP_005539), *Zea mays* (NP_001146902), *Saccharomyces cerevisiae* (AAA6691), *Escherichia coli* (NP_417366). The number sign (#) show the residues which form the active site. Identical amino acids are highlighted in black, and positive amino acids are highlighted in gray.
LysRS revealed 84% sequence identity to LysRS of *Bombyx mori* (BGIBMGA002984-PA in SilkDB), 68-74% to LysRSs of other insects. The *A. pernyi* LysRS also revealed 57-72% identity to other invertebrates, 64-71% identity to vertebrates, 56-59% identity to plants, and 57% identity to fungi (*Saccharomyces cerevisiae*). Note that the *A. pernyi* LysRS protein showed 37-42% sequence identity to LysRSs of bacteria. In Figure 2, the sequence is shown aligned with six LysRS proteins from *A. pernyi, B. mori* (BGIBMGA002984-PA in SilkDB), *Drosophila melanogaster* (NP_572573), *H. sapiens* (NP_005539), *Zea mays* (NP_001146902), *S. cerevisiae* (AAA6691), *E. coli* (NP_417366). High level of conservation of the amino acid sequence among these LysRS proteins indicates that the LysRS proteins are highly conserved during the evolution of life organisms.

Moreover, by sequence alignment, we found that, the active site residues of LysRS (Desogus et al., 2000), responsible for the ATP-dependent formation of the enzyme bound aminoacyl-adenylate, are identical among various kinds of organisms including *A. pernyi* collected in this study (Figure 2 and data not shown). This finding suggested that the *A. pernyi* LysRS is sufficient to have the catalytic activity. The high level of conservation and identical active sites are suggestive of a critical function that these proteins must play in all the organisms where they are found.

### 3.3 Phylogenetic analysis

A total of 43 representative LysRS protein sequences from various organisms including *A. pernyi* were used to reconstruct the phylogenetic relationship. The final alignment resulted in 678 amino acid sites including gaps. Of these sites, 101 were conserved, 562 were variable, and 477 were informative for parsimony. A neighbor-joining tree was constructed using amino acid sequences and a poisson-corrected distance with bootstrap test of 1000 replications (Figure 3). The *A. pernyi* LysRS protein sequence was found to be closely related to that of *B. mori* with 100% confidence support. The used LysRS sequences were well divided into five groups corresponding to the known bacteria, fungi, plants, invertebrates and vertebrates. Within the plant group, three subgroups were well defined corresponding to algae (*Chlamydomonas reinhardtii*), magnoliopsida (*Vitis vinifera*) and liliopsida (*Z. mays*).

Within the vertebrate group, four subgroups were well clustered corresponding to mammalia (*H. sapiens*), archosauiria (*Gallus gallus*), amphibia (*Xenopus laevis*) and actinopterygii (*Danio rerio*). Within the invertebrate group, four subgroups were also well defined corresponding to insects and hard tick (*Branchiostoma floridae*), cnidaria (*Hydra magnipapillata*), nematoda (*Caenorhabditis elegans*). Moreover, within the insect subgroup, all the lepidopterans, dipterans and hymenopterans species were further clearly separated. These results agreed with the topology tree on the classical systematics and other molecular data, such as the *will die slowly* gene (Li et al., 2011), suggesting the potential value of LysRS protein in phylogenetic inference of life organisms.

### 3.4 Expression patterns

We performed semi-quantitative RT-PCR to detect the quantify the *A. pernyi* LysRS gene expression levels, by using an *actin* gene as an internal control that was a constitutively expressed gene (Wu et al, 2010). The results showed that the *A. pernyi* LysRS gene was expressed during four developmental stages (egg, larva, pupa and adult) (Figure 4), suggesting that the product of the *LysRS* gene plays an essential role throughout the entire life cycle of *A. pernyi*. 

www.intechopen.com
Fig. 3. Phylogenetic tree based on the amino acid sequence comparisons of LysRS proteins from various organisms including *A. pernyi*. Numbers at nodes represent bootstrap P-values (>50%). Public database accession numbers of LysRS proteins are shown following the names of organisms. Identity (%) in parentheses following accession number is obtained by pairwise alignment of amino acid sequence of *A. pernyi* LysRS with indicated LysRSs from other organisms.
Fig. 4. Expression patterns of the *A. pernyi* LysRS mRNA in different developmental stages and different tissues of fifth instar larvae preformed by semi-quantitative RT-PCR. RT-PCR was amplified after 30 cycles with specific primer pair for the *A. pernyi* LysRS gene. The *actin* gene was used as an internal standard to normalize the templates. Relative expression profiles of *A. pernyi* LysRS were normalized with *actin* level. Lanes: 1, eggs at day 5; 2, larvae of fifth instar; 3, pupae; 4, moths; 5, blood; 6, fat body; 7, midgut; 8, silk glands; 9, body wall; 10, Malpighian tubules; 11, spermaries; 12, ovaries; 13, brain; 14, muscle. (A) The electrophoretic results. (B) The relatively intensity.

Tissue distributions in fifth instar larvae of the *A. pernyi* LysRS gene were also analyzed. The results showed that the *A. pernyi* LysRS gene mRNA was present in all tissues tested including blood, fat body, midgut, silk glands, body wall, Malpighian tubules, spermaries, ovaries, brain and muscle (Figure 4). Large-scale EST resource for *B. mori*, a lepidopteran model insect, are available at GenBank database. The *in silico* gene expression analysis based on the available EST resources showed the *B. mori* LysRS gene was also expressed in imaginal disks and pheromone gland not analyzed in this study. The results also showed that the mRNA levels of the *A. pernyi* LysRS gene were most abundant in Malpighian tubules. Malpighian tubules system is a type of excretory and osmoregulatory system in insect, including *A. pernyi*. The system consists of branching tubules extending from the alimentary canal that absorbs solutes, water, and wastes from the surrounding hemolymph. The wastes then are released from the organism in the form of solid nitrogenous compounds. Therefore, the high expression of the gene in the Malpighian tubules corresponds to its role in the development of *A. pernyi*.

4. Conclusion

In the present study, the full length cDNA of the *A. pernyi* LysRS gene was isolated and characterized from a pupal cDNA library by random EST sequencing. The obtained cDNA sequence consists of 2136 bp nucleotides encoding a polypeptide of 579 amino acids which contains the LysRS class II core domain including the conserved active sites and three characteristic sequence motifs. RT-PCR analysis showed that the *A. pernyi* LysRS gene was transcribed during four developmental stages (egg, larva, pupa, and moth) and in all the tissues tested (blood, midgut, silk glands, Malpighian tubules, spermaries, ovaries, brain, muscle, fat body and body wall), with most abundance in Malpighian tubules. By searching in database, the *A. pernyi* LysRS protein homologues were found in various kinds of life organisms, including bacteria, fungi, plants, invertebrates and vertebrates, with 37–84% amino acid sequence identity, suggesting that they are highly conserved during the
evolution of life organisms. Phylogenetic analysis based on the LysRS protein homologue sequences clearly separated the known bacteria, fungi, plants, invertebrates and vertebrates, consistent with the topology tree on the classical systematics, suggesting the potential value of LysRS protein in phylogenetic inference of life organisms.

5. Acknowledgment

This work was supported by grants from the National Natural Science Foundation of China (No. 31072082), the National Modern Agriculture Industry Technology System Construction Project (Silkworm and Mulberry), the Support Project for Tip-top Young and Middle-aged Talent of Shenyang Agricultural University, the Scientific Research Project for High School of the Educational Department of Liaoning Province (No. 2008643).

6. References


www.intechopen.com
The development of molecular cloning technology in the early 1970s created a revolution in the biological and biomedical sciences that extends to this day. The contributions in this book provide the reader with a perspective on how pervasive the applications of molecular cloning have become. The contributions are organized in sections based on application, and range from cancer biology and immunology to plant and evolutionary biology. The chapters also cover a wide range of technical approaches, such as positional cloning and cutting edge tools for recombinant protein expression. This book should appeal to many researchers, who should find its information useful for advancing their fields.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
