Cloning and Expression of Crocodile (Crocodylus siamensis) Hemoglobin

Student: Mr. Thai Kabbua
Supervisor: Assoc. Prof. Sompong Thammasirirak
CSG member: Mr. Yospong Temsiripong

Summary

Hemoglobin (Hb) is a major component in red blood cell of vertebrates, plays an important role in oxygen transfer from lung to tissues. Moreover, hemoglobin can be a source of bioactive peptides with its antibacterial activity. The purpose of this study is to clone and express the Siamese crocodile (Crocodylus siamensis) hemoglobin from Thailand. The nucleotides of alpha and beta hemoglobin of C. siamensis were cloned from blood. The mRNA of two globin genes were amplified by using degenerate primers and perform complete cDNA sequences by using a gene-specific primer by the RACE method. The full-length gene of alpha hemoglobin cDNA consists of 558 nucleotides which contain an open reading frame for 141 amino acids, whereas the beta hemoglobin cDNA consists of 605 nucleotides which contain an open reading frame for 146 amino acids. Using Blastp tool, the deduce amino acid sequence of alpha hemoglobin from C. siamensis was identity to alpha hemoglobin of Crocodylus niloticus (99%), Alligator mississippiensis (88%) and Caiman crocodilus (85%). The deduced amino acid sequence of beta hemoglobin from C. siamensis was identity to beta hemoglobin of Crocodylus niloticus (95%), Alligator mississippiensis (80%) and Caiman crocodilus (75%). This study provided essential information for elucidating the possible roles of hemoglobin in oxygen binding and antimicrobial activity in C. siamensis. This data will be contributed to the understanding of the genetic background underlying the structures and functions of crocodile hemoglobin and help to understand the evolutionary relationships of the various species of crocodiles. Our studies in the near future will express the hemoglobin from C. siamensis in E. coli system and characterize of this protein for understanding their evolution and function.

Introduction

Hemoglobin (Hb), the key component of oxygen storage and regulation, plays an important role in oxygen transfer from lung to tissues. It is widely distributed in all living organism including animals, plants, bacteria, yeast, etc (Hardison, 1998). It is unique in its ability to adapt to a wide range of environmental conditions (Monica et al., 2003). The structure of hemoglobin varies across species. Among vertebrates, hemoglobin has been widely used as subject for evolution and molecular adaptation studies due to their highly conserved tridimensional structure and biological function, exhibiting a great variation regarding absolute affinities for oxygen and susceptibility to control by metabolic effectors (Petruzzelli et al., 1996). Moreover, hemoglobin can be a source of bioactive peptides with its antibacterial activity (Liepke et al., 2003). Alligator hemoglobin exerts antimicrobial activity (Hoffman et al., 2002). Crocodile hemoglobin has a specialized structure causing crocodiles to live in water for hours. This is due to the fact that the bicarbonate ions, which are the end product of respiration, drastically reduce the oxygen binding affinity of crocodile hemoglobin leading to the release of bound oxygen to the respiring tissue. Structural studies on crocodile (Crocodylus niloticus) and human hemoglobin are only 68% for their alpha chains and 51% for their beta chains are homologous. These results indicate that an entirely new function could evolve in a protein by a relatively small number of amino-acid substitutions in key
positions (Komiyama et al., 1995). However, there is no genetic information on properties of an alpha and beta chain of hemoglobin from Siamese crocodiles (*Crocodylus siamensis*), which is an important captive bred crocodilian in Thailand that have been bred widely in captivity. This research focused on the cloning and expression which was carried out on hemoglobin of *C. siamensis*.

**Methodology**

**Sample preparation**

*Crocodylus siamensis* (Siamese crocodile) were captured and housed at Sriracha Moda Farm, Chonburi Province, Thailand. Crocodile blood was harvested in heparin tube and immediately frozen in dry-ice and then transfer to laboratory and stored at -70°C for molecular analysis.

**Total RNA extraction and synthesis of cDNA**

Total RNA was extracted and purified from blood of *C. siamensis* using TRIZOL LS reagent (Invitrogen, Co., Ltd.) following the manufacturer’s protocol. After that, total RNA was digested with DNaseI to removed DNA contamination. Total RNA was used to the template in RT-PCR reaction. The first strand cDNA was synthesized using of total RNA by ImProm II reverse transcriptase (Promega, Madison, USA) with Oligo (dT). The amplification condition was carried out on MJ Mini™ Gradient Thermal Cycler (Bio-rad, USA) according to instructions.

**Primers design and amplification of *C. siamensis* hemoglobin fragment**

Degenerate primers were designed based on amino acid sequence of alpha hemoglobin (PIR ID: BOM2T2) and beta hemoglobin (PIR ID: P86919) from *C. siamensis* hemoglobin which was characterized in our laboratory by Srihongthong. The degenerated primer was analyses by using oligo analyzer program. Degenerate PCR reaction was conducted with MJ Mini™ Gradient Thermal Cycler (Bio-rad, USA). The gradient PCR condition was comprised of an initial denaturation step at 95 °C for 3 min and then 35 cycles were used as follows: at 95°C for 30 sec, at 50°C to 55°C for 30 sec, at 72°C for 45 sec and a final extension at 72°C for 5 min. The PCR products were separated in 1.5% agarose gel electrophoresis. The expected size of PCR product was extracted by the NucleoSpin® Extract II (Invitrogen, USA) following the manufacturer’s instructions.

**Cloning and sequencing of PCR product**

The purified PCR products were ligated into a pGEM®-T Easy vector (Promega, USA). The recombinant plasmids were transformed into *E. coli* DH5α and selected positive clone on LB agar containing 100 µg/mL ampicillin. The positive clones were checked by colony PCR method and restriction enzyme. Recombinant plasmids from selected clone were extracted and subjected to sequencing at Macrogen Inc. (Korea). The DNA fragment sequences were analyzed by using the multiple sequence alignment in the Basic Local Alignment Sequence algorithm program (BLAST).

**Synthesis and cloning of full length hemoglobin gene**

Rapid Amplification of cDNA Ends (RACE) was used to determine 3'-end and 5'-end fragment of *C. siamensis* hemoglobin gene. The gene-specific primers (GSP) derived from the degenerative partial cDNA sequence of alpha hemoglobin and beta hemoglobin of *C. siamensis*. The RACE synthesized cDNA were cloned into pGEM®-T Easy vector (Promega, USA). The nucleotide sequences of RACE fragment were extracted and subjected to
sequencing at Macrogen Inc. (Korea). The RACE DNA fragment sequences were analyzed by using the multiple sequence alignment in the Basic Local Alignment Sequence algorithm program (BLAST).

Sequence analysis
The nucleotide sequences were analyzed by using bioinformatics tool such as translate tool (http://web.expasy.org/translate), protein blast or Blastp (http://blast.ncbi.nlm.nih.gov /Blast) and clustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2).

Results

Synthesis and cloning of full length hemoglobin gene
The cDNA syntheses were achieved completely by RACE method. The full-length gene of alpha hemoglobin cDNA consists of 558 nucleotides which contain an open reading frame for 141 amino acids, whereas the beta hemoglobin cDNA consists of 605 nucleotides which contain an open reading frame for 146 amino acids.

Sequence analysis
Using translate tool, the nucleotide sequences of alpha hemoglobin and beta hemoglobin were deduced to 141 and 146 for alpha and beta hemoglobin, respectively.

To find the regions of similarity between DNA and protein sequences, we used the Blast search program. Using Blastn, the nucleotide sequence of alpha hemoglobin is identical to the top tree nucleotides databases which are Myotis adversus hemoglobin alpha chain (HBA) mRNA (GenBank: EU568365.1), Condylura cristata isolate MB1 hemoglobin subunit alpha 2 (HBA2) mRNA (GenBank: JN208869.1) and Pantholops hodgsonii hemoglobin alpha chain mRNA (GenBank: JF811751.1) for 75%. The nucleotide sequence of beta hemoglobin is identical to Ailuropoda melanoceua hemoglobin delta mRNA (GenBank: AY753986.1) Ailuropoda melanoceua hemoglobin beta mRNA (GenBank: AY753985.1) Mirounga angustirostris mRNA (GenBank: M73997.1) for 78%, 77% and 75%, respectively.

The deduced amino acid of C. siamensis hemoglobin was analyzed by using Blastp program to find the similar region with other organism. The deduced amino sequence of alpha hemoglobin is identical to alpha hemoglobin of Crocodylus niloticus (UniProtKB/Swiss-Prot: P01998.1), Alligator mississippiensis (UniProtKB/Swiss-Prot: P01999.2) and Caiman crocodilus (UniProtKB/Swiss-Prot: P02000.1) for 99%, 88% and 85%, respectively. The deduced amino sequence of beta hemoglobin is identical to beta hemoglobin of Crocodylus niloticus (UniProtKB/Swiss-Prot: P02129.1), Alligator mississippiensis (UniProtKB/Swiss-Prot: P02130.1) and Caiman crocodilus (UniProtKB/Swiss-Prot: P02131.1) for 95%, 80% and 75%, respectively.

To compare with other crocodile species and human hemoglobin, the deduced amino acid sequence of C. siamensis were align by using clustalW2 program. The results of alignment were shown in figure 1 and 2 for alpha and beta hemoglobin, respectively.
The result from the both figures show the conserve residues are the heme binding sites which are 59His (distal to heme) and 88His (proximal to heme) for alpha hemoglobin and 64His (distal to heme) and 93His (proximal to heme) for beta hemoglobin. Another region that conserved is the site that involve in hydroglen bonding with bicarbonate ion. The bicarbonate ions can reduce the oxygen binding affinity of crocodile hemoglobin. To study the predicted properties of primary and secondary structure of deduce amino acid from *C. siamensis*, we used secondary structure prediction program (GOR - Garnier et al., 1996) and Expasy program for analyse. The theoretical Mw and pl is 15.72 kD and 7.16, respectively for alpha hemoglobin and 16.92 kD and 7.83, respectively for beta hemoglobin. The predicted secondary structures of deduced amino acid of alpha and beta hemoglobin from *C. siamensis* were shown in figure 3 and 4, respectively.
Figure 3. The predicted secondary structure of alpha hemoglobin from *C. siamensis*.

```
<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MVLS...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3_10 helix (Gg)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pi helix (Ii)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Beta bridge (Bb)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Extended strand (Ee)</td>
<td>24</td>
<td>16.90%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta turn (Tt)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bend region (Ss)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Random coil (Cc)</td>
<td>65</td>
<td>45.77%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguous states (?)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other states</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
```

Sequence length : 142

GOR4 :

- Alpha helix (Hh) : 53 is 37.32%
- 3_10 helix (Gg) : 0 is 0.00%
- Pi helix (Ii) : 0 is 0.00%
- Beta bridge (Bb) : 0 is 0.00%
- Extended strand (Ee) : 24 is 16.90%
- Beta turn (Tt) : 0 is 0.00%
- Bend region (Ss) : 0 is 0.00%
- Random coil (Cc) : 65 is 45.77%
- Ambiguous states (?) : 0 is 0.00%
- Other states : 0 is 0.00%

Figure 4. The predicted secondary structure of beta hemoglobin from *C. siamensis*.

```
<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSA...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha helix (Hh)</td>
<td>79</td>
<td>53.74%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3_10 helix (Gg)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pi helix (Ii)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Beta bridge (Bb)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Extended strand (Ee)</td>
<td>21</td>
<td>14.29%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta turn (Tt)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bend region (Ss)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Random coil (Cc)</td>
<td>47</td>
<td>31.97%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguous states (?)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other states</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
```

Sequence length : 147

GOR4 :

- Alpha helix (Hh) : 79 is 53.74%
- 3_10 helix (Gg) : 0 is 0.00%
- Pi helix (Ii) : 0 is 0.00%
- Beta bridge (Bb) : 0 is 0.00%
- Extended strand (Ee) : 21 is 14.29%
- Beta turn (Tt) : 0 is 0.00%
- Bend region (Ss) : 0 is 0.00%
- Random coil (Cc) : 47 is 31.97%
- Ambiguous states (?) : 0 is 0.00%
- Other states : 0 is 0.00%

The secondary structure of deduce amino acid of both sequences are most of alpha helix and random coil form. Interestingly, the alpha helix is one of the secondary structures that found in the antimicrobial peptides or antimicrobial proteins. This result suggest that the antibacterial activity of hemoglobin from *C. siamensis* may result from the secondary structure of its
Conclusions and Discussion

The deduced amino acid sequence of alpha hemoglobin from *C. siamensis* was identity to alpha hemoglobin of *Crocodylus niloticus* (99%), *Alligator mississippiensis* (88%) and *Caiman crocodilus* (85%). The deduced amino acid sequence of beta hemoglobin from *C. siamensis* was identity to beta hemoglobin of *Crocodylus niloticus* (95%), *Alligator mississippiensis* (80%) and *Caiman crocodilus* (75%). for understanding their evolution and function. These results may help to better understand the evolution and relationships of crocodile species. The theoretical Mw and pI are 15.72 kD and 7.16, respectively for alpha hemoglobin and 16.92 kD and 7.83, respectively for beta hemoglobin. The secondary structure of deduced amino acid of *C. siamensis* hemoglobin sequences are most of alpha helix and random coil form. This study provided essential information for elucidating the possible roles of hemoglobin in oxygen binding and antimicrobial activity in *C. siamensis*. Our studies in the near future will express the hemoglobin from *C. siamensis* in *E. coli* system and characterize of this protein for better understanding their evolution and function.

References


