Homology Modeling and Prediction of Catalytic Amino Acid in the Neurotoxin from Indian Cobra (*Naja naja*)

Pallavi Somvanshi* and Vijai Singh

**Bioinformatics centre, Biotech Park, Sector-G Jankipuram, Lucknow-226021, Uttar Pradesh, India**

**Abstract:** The neurotoxin secreted by Indian cobra (*Naja naja*) binds to acetylcholine receptor of nerve cells, which leads to a lump in the nerve impulse, ceases breathing, thereby causing the death of a person due to suffocation. These neurotoxins are small peptide with approximately 7 kDa. Homology modeling was performed to generate the 3D structure of neurotoxins (A, B, C and D) of *N. naja*, using the known protein template crystal structure (PDB: 2CTX). The validation of 3D structure was done using PROCHECK. Furthermore, the prediction of catalytic amino acid residues in the active site domain of the 3-D structure of neurotoxin was identified. The 3-D structures of neurotoxin and catalytic amino acid residue may be used to target and design the antivenom drugs against the Indian cobra.

**Keywords:** *Naja naja*, Neurotoxin, Acetylcholine receptor, Antivenom.

**INTRODUCTION**

*Naja naja* is a most famous species of venomous snakes native to the Indian subcontinent. The venom of Indian cobra contains a powerful post-synaptic neurotoxin. The venom acts on the synaptic gaps of the nerves, thereby paralyzing the muscles and possibly leads to respiratory failure or cardiac arrest. The venom components include enzymes such as hyaluronidase that cause lysis and increase the spread of the venom [1]. A symptom of cobra envenomation begins from 15 minutes to two hours after the bite and can be fatal in less than an hour [2]. Indian cobra is one of the most dangerous venomous snakes and a polyvalent serum is available for treating snakebites by these snakes.

A neurotoxin particularly acts in the acetylcholine receptor of nerve cells by interacting with the membrane proteins such as ion channels. Many of the venoms and other toxins that organisms use in defense against vertebrates are neurotoxins. The venom of bees, scorpions, pufferfish, spiders and snakes contains many different toxins. A new drug was found for the treatment against Alzheimer's, using the similarity search technique. These molecules had good alignment with GTS-21, and docking studies with alpha7 nicotinic acetylcholine receptor was done [3]. The homology modeling also, known as comparative modeling, is a class of methods used for constructing an atomic-resolution model of a protein from its amino acid sequence (target). Homology modeling techniques rely on the identification of one or more known protein structures (templates) likely to resemble the structure of the query sequence, and on the production of an alignment that maps residues in the query sequence to residues in the template sequence [4]. The crystal structure of acetylcholine-binding protein, used for the construction of 3-D structures of extracellular domain of the monomer, homodimer, and homopentamer of the alpha7 nicotinic acetylcholine receptor, was determined [5].

The crystal structure of the neurotoxin alpha cobratoxin was refined by using X-ray data to 2.4-Å resolution. The polypeptide chain forms three loops I, II, III, knotted together by four disulfide bridges. Loop I is stabilized by one beta -turn, and two beta sheet hydrogen bonds. Loop II by eight beta sheet hydrogen bonds with the tip folded into two distorted right handed helical turns, stabilized by two alphahelical and two beta turn hydrogen bonds. Loop III forms an antiparallel triple-pleated beta sheet and tight anchoring of the Asn3 side chain fixes the tail segment. The alphacobratoxin molecules, dimerize by beta sheet formation between strands 63 and 57 of symmetry-related molecules, could be of importance for interaction with acetylcholine receptor [6]. GFAT and its interaction with UDP-GlcNAc may become a novel therapeutic target for the treatment of type 2 diabetes. The 3-D structure of human GFAT1 monomer and dimer was done [7].

The alpha toxin from *N. nigriceps* has been determined with NMR spectroscopy and molecular modeling. The overall folding of alpha toxin consists of three major loops, which are stabilized by three disulfide bridges and one short C terminal loop stabilized by a fourth disulfide bridge. All the disulfides are grouped in the same region of the molecule, forming a highly constrained structure from which the loops protrude. Minor differences have been observed between two structures, which are partly related to the deletion of one residue from the chain of alpha toxin [8]. The structure of alpha-cobratoxin, a neurotoxin, is purified from the venom of the snake *Naja naja siamensis*. The NMR data showed the presence of a triple-stranded beta-sheet (residues 19-25, 36-41, and 52-57), a short helix, and turns. These were applied as distance constraints in a molecular modeling comprised of distance geometry and dynamical simulated annealing calculations [9].

The distribution of phi and psi angles represented in the Ramachandran plot is probably the most frequently used
measure for the reliability of the experimentally derived model of a protein molecule. A model under test is assessed in PROCHECK by comparing the percentage of observations in the most highly favored regions of phi and psi space for the model to a standard plot of percentage-in-CORE versus resolution, derived from a dataset of previously solved protein structures. The major advancement was the use of G-factors, quantifying the plausibility of certain geometrical parameters such as the mainchain and sidechain torsion angles [10].

There is a scarcity of effective medicines, drugs and antivenom used for the treatment of cobra bite in humans. The homology modeling of neurotoxin provides a new approach to target and design the suitable drugs and antivenom. The present study enables an identification of a new lead molecule for targeting neurotoxin of snakes. The 3-D structure of four known snake neurotoxins of *N. naja* is so far, unknown. The present aim of the study is to generate and validate the 3-D structure of neurotoxin of *N. naja* (Indian Cobra) and to predict the catalytic amino acid residues in the active site domain of 3-D structure of neurotoxin.

1. MATERIALS AND METHODS

1.1. Retrieval of Sequence Data

The complete protein sequences of neurotoxins of *Naja naja* were retrieved from the biological database, National Centre for Biotechnology Information (NCBI), cited at http://www.ncbi.nlm.nih.gov.

1.2. Prediction and Identification of Neurotoxin:

The protein sequences of *N. naja* were confirmed through online server (www.imtech.res.in/raghava/ntxpred/) for the prediction of nature [11].

1.3. Blast

The relatedness of sequences deposited in databases was evaluated by BLAST (Basic Local Alignment Search Tool) [12], implemented via the NCBI website [www.ncbi.nlm.nih.gov/blast/] against the complete training dataset which is extracted from Genbank database. The BlastP (protein query – protein database comparison) was done with the protein data base (PDB). The alignment was performed with the target protein sequences with Protein Data Bank (PDB code: 2CTX) template using CLUSTALX 1.8.

1.4. Homology Modeling

The X-crystal structural coordinates of neurotoxin of *Naja naja siamensis* were available at 2.4 Å resolution (PDB: 2CTX) and were used as template structure to generate the three-dimensional model of neurotoxin. The Modeler9v2 homology modeling software was used to generate the 3-D structure of neurotoxin.

1.5. Evaluation and Validation of the 3D Structure

The method implied for the evaluation of the 3D model used two parameters like minimum model score and dope score of model and template. The 3-D structure was validated with the PROCHECK. This program generates Ramachandran plot, accuracy, and overall G-factor.

1.6. Identification of Catalytic Amino Acids

The 3D structure of neurotoxin was used to predict the catalytic amino acid residues in the active site domain. The European Molecular Biology Laboratory (EMBL) online catalytic site atlas program was used.

1.7. Protein Model Database

The 3-D structure of the Long neurotoxin A, C and D of Indian cobra have been assigned under PM0075157, PM0075161, and PM0075162, respectively.

RESULTS AND DISCUSSION

The size of neurotoxin (A, B, C and D) of Indian cobra was 71 amino acids in length. The physicochemical properties like molecular weight, pl value, and percentage identity with the template protein were given (Table 1). The nature of neurotoxin was confirmed by NTXPRED. All proteins showed the neurotoxin nature and may bind with the acetylcholine receptor of nerve cells. The molecular weight varied from 7.821 to 7.889 kDa, due to some mutation in the peptide chain. The pl value of all neurotoxins was ranged between 7.45 to 7.78; it provides the stability of peptide. The alignment result of neurotoxin including the template protein was given (Fig. 1). The alignment result showed approx 98% similarity of amino acids. However, mutation was also obtained in this alignment. This mutation may cause changes in the confirmation of protein 3-D structure. The 3-D structure of neurotoxins A, B, C and D of Indian cobra was given (Fig. 2), and was evaluated with scores of target protein and template. In this study, we observed significant similarity of score and dope score (minimum energy was considered), target and template protein. The score of 3-D structure of four neurotoxins was given (Table 2). The 3-D structures of human alpha-13 subunit of guanine nucleotide-binding protein (G alpha 13) and human thromboxane A2 (TXA2) receptor were generated [13]. Agaritine has both antiviral and anticancer properties. The most favorable free energy was used to distinguish compound agaritine among 9,000 com-

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Designation of Toxin</th>
<th>NCBI Acc. No.</th>
<th>MW (Da)</th>
<th>pl</th>
<th>Identity with PDB Template</th>
<th>Nature Predicted Through NTXPRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Neurotoxin A</td>
<td>P25668</td>
<td>7847.6</td>
<td>7.79</td>
<td>94 %</td>
<td>Neurotoxin</td>
</tr>
<tr>
<td>2</td>
<td>Neurotoxin B</td>
<td>P25669</td>
<td>7821.5</td>
<td>7.79</td>
<td>92 %</td>
<td>Neurotoxin</td>
</tr>
<tr>
<td>3</td>
<td>Neurotoxin C</td>
<td>P25671</td>
<td>7833.6</td>
<td>7.78</td>
<td>97 %</td>
<td>Neurotoxin</td>
</tr>
<tr>
<td>4</td>
<td>Neurotoxin D</td>
<td>P25672</td>
<td>7889.7</td>
<td>7.45</td>
<td>92 %</td>
<td>Neurotoxin</td>
</tr>
</tbody>
</table>
Fig. (1). The multiple sequences alignment of Indian cobra.

![Multiple sequences alignment of Indian cobra](image)

Table 2. Validated 3-D Structure of Indian Cobra (N. naja) Neurotoxin

<table>
<thead>
<tr>
<th>S. No</th>
<th>Designation of Toxin</th>
<th>Dope Score (kcal/mol)</th>
<th>Amino Acid Residue in Favor Region</th>
<th>Amino Acid Residue in Disallowed Region</th>
<th>Overall G- Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Neurotoxin A</td>
<td>-5226.004395</td>
<td>91.5 %</td>
<td>0.0 %</td>
<td>0.9</td>
</tr>
<tr>
<td>2</td>
<td>Neurotoxin B</td>
<td>-4935.359863</td>
<td>93.2 %</td>
<td>0.0 %</td>
<td>0.9</td>
</tr>
<tr>
<td>3</td>
<td>Neurotoxin C</td>
<td>-5187.339355</td>
<td>91.7 %</td>
<td>0.0 %</td>
<td>0.9</td>
</tr>
<tr>
<td>4</td>
<td>Neurotoxin D</td>
<td>-5326.658691</td>
<td>89.8 %</td>
<td>0.0 %</td>
<td>0.9</td>
</tr>
</tbody>
</table>

A novel short neurotoxin, cobrotoxin C (CBT C), was isolated from the venom of monocellate cobra (Naja kaouthia), using a combination of ion-exchange chromatography and FPLC. Its primary structure was determined by Edman degradation. CBT C is composed of 61 amino acid

pounds in binding against HIV protease [14]. The atomic coordinates of SARS coronavirus main proteinase were used for the docking studies with drugs KZ7088 and AVLQSGFR octapeptide. It was concluded that both drugs interact on the active sites of SARS enzyme by six hydrogen bonds [15].
residues. It differs from cobrotoxin B (CBT B) by only two amino acid substitutions, Thr/Ala11 and Arg/Thr56, which are not located on the functionally important regions by sequence similarity. The structure-function relationship analysis suggests the existence of a functionally significant domain on the outer region of Loop III of CBT C. The functionally important basic residues on the outside of Loop III might have a pair wise interaction with alpha subunit, instead of gamma or delta subunits of the nicotinic acetylcholine receptor [16].

The crystal structure of neurotoxin, a cobratoxin, was refined to an R-factor of 9.5%, using 3271 X-ray data to 2.4Å resolution. The polypeptide chain forms three loops I, II, II1, knotted together by four disulfide bridges, with the most prominent, loop II, containing another disulfide close to its lower tip. Loop I is stabilized by one beta-turn, and two beta sheet hydrogen bonds. Loop II by eight beta sheet hydrogen bonds with the tip folded into two distorted right handed helical turns, stabilized by two alpha-helical and two beta turn hydrogen bonds. Loop III forms an antiparallel triple-pleated beta sheet, and tight anchoring of the Asn3 side chain fixes the tail segment. The alpha-cobratoxin molecules dimerize by beta sheet formation between strands 63 and 57 of symmetry-related molecules could be of importance for interaction with acetylcholine receptor [6]. The 3-D-QSAR model of (S)-3-Aryl-5-substituted oxazolidinones was developed and provided new insight in designing potential antibacterial agents [17]. CYP2C19 is a member of cytochrome P-450 enzyme superfamily and played an important role in drug metabolism. The ligands of CEC, Fluvoxamine, Lescol and Ticlopidine were prepared through the structure-activity relationship strategies [18].

The influence of long-range interactions on local structures is an important issue in understanding the folding process and structure stability of protein. The conformational properties of short chain of snake venom neurotoxin studied were by Monte Carlo simulated annealing method. The surrounding protein structure was found to be crucial in stabilizing the loop conformation. Although eight peptides prefer type V beta turn in solution, three of them (KPGI, KPGV, KSGI) turned to type II beta turn and the other five (KKGI, KKV, KNGI, KQGI, and KRGV) were confined to more rigid type V beta turn conformation in the protein structure. Using flexible tetra-glycine-peptide to screen the backbone conformational space in the protein environment, also validates the results. This study shows that long-range interactions do contribute to the stability and types of conformation for a surface loop in protein, whereas short-range interactions may only provide candidate conformations, which then have to be filtered by the long-range interactions [19].

GABA is an inhibitory neurotransmitter present in the mammalian CNS which targets clinically important drugs. The 3-D models of GABA may provide a logical structural frame for designing suitable drugs. The given models might be useful to understand the mechanism of operation of GABA-A receptors, stimulating novel strategies for developing more specific drugs and better treatments [20]. SARS is a respiratory disease caused by coronavirus. The cleavage mechanism of SARS coronavirus main protease on the octapeptide was studied through molecular mechanics and quantum mechanics simulations. It has been observed that the catalytic dyad (His-41/Cys-145) site between domains [21].

The alpha toxin from *Naja nigricollis* has been determined by NMR spectroscopy and molecular modeling. The overall folding of alpha toxin consists of three major loops which are stabilized by three disulfide bridges and one short C terminal loop, stabilized by a fourth disulfide bridge. The disulfides are grouped in the same region of the molecule, forming a highly constrained structure from which the loops protrude. Minor differences have been observed between the two structures are partly related to the deletion of one residue from the chain of alpha toxin [8].

In this analysis, Ramachandran plot for neurotoxins of Indian cobra fulfilled the test ranged from 89.8 to 93.2% in the most favored region. There was no amino acid residue present in the disallowed region of the Ramachandran plot. The overall G – factor of neurotoxins was 0.9 for all 3-D structures (Table 2). The result deduced that 3-D structure of neurotoxin was favored for a good satisfactory model. Ramachandran plot for neurotoxin of Indian cobra was given (Figs. 3 and 4). The solution structure of alpha-cobratoxin, a neurotoxin, is purified from the venom of the snake *Naja*.

Fig. (3). Favoured amino acid residues in red area in Ramachandran plot for 3-D structure of neurotoxin.
Homology Modeling and Prediction of Catalytic

**Table 3. Prediction of Catalytic Amino Acid Residues Present in the Active Site Domain**

<table>
<thead>
<tr>
<th>S. no.</th>
<th>Designation of Toxin</th>
<th>Conserved Catalytic Amino Acid Residue Present in a Active Site Domain</th>
<th>Variable Catalytic Amino Acid Residue Present in a Active Site Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Neurotoxin A</td>
<td>TRP&lt;sub&gt;25&lt;/sub&gt;, CYS&lt;sub&gt;26&lt;/sub&gt;, GLY&lt;sub&gt;34&lt;/sub&gt;, ARG&lt;sub&gt;35&lt;/sub&gt;, ARG&lt;sub&gt;36&lt;/sub&gt;, ASP&lt;sub&gt;37&lt;/sub&gt;</td>
<td>CYS&lt;sub&gt;20&lt;/sub&gt;, HIS&lt;sub&gt;18&lt;/sub&gt;, CYS&lt;sub&gt;3&lt;/sub&gt;, CYS&lt;sub&gt;62&lt;/sub&gt;</td>
</tr>
<tr>
<td>2</td>
<td>Neurotoxin B</td>
<td>TRP&lt;sub&gt;25&lt;/sub&gt;, CYS&lt;sub&gt;26&lt;/sub&gt;, GLY&lt;sub&gt;34&lt;/sub&gt;, ARG&lt;sub&gt;35&lt;/sub&gt;, ARG&lt;sub&gt;36&lt;/sub&gt;, ASP&lt;sub&gt;37&lt;/sub&gt;</td>
<td>CYS&lt;sub&gt;20&lt;/sub&gt;, HIS&lt;sub&gt;18&lt;/sub&gt;, CYS&lt;sub&gt;3&lt;/sub&gt;, CYS&lt;sub&gt;62&lt;/sub&gt;</td>
</tr>
<tr>
<td>3</td>
<td>Neurotoxin C</td>
<td>TRP&lt;sub&gt;25&lt;/sub&gt;, CYS&lt;sub&gt;26&lt;/sub&gt;, GLY&lt;sub&gt;34&lt;/sub&gt;, ARG&lt;sub&gt;35&lt;/sub&gt;, ARG&lt;sub&gt;36&lt;/sub&gt;, ASP&lt;sub&gt;37&lt;/sub&gt;</td>
<td>Absent</td>
</tr>
<tr>
<td>4</td>
<td>Neurotoxin D</td>
<td>TRP&lt;sub&gt;25&lt;/sub&gt;, CYS&lt;sub&gt;26&lt;/sub&gt;, GLY&lt;sub&gt;34&lt;/sub&gt;, ARG&lt;sub&gt;35&lt;/sub&gt;, ARG&lt;sub&gt;36&lt;/sub&gt;, ASP&lt;sub&gt;37&lt;/sub&gt;</td>
<td>CYS&lt;sub&gt;20&lt;/sub&gt;, HIS&lt;sub&gt;18&lt;/sub&gt;, CYS&lt;sub&gt;3&lt;/sub&gt;, CYS&lt;sub&gt;62&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

**naja siamensis.** A sequence-specific assignment of the NMR resonances was attained by a combination of generalized main-chain-directed strategy and of the sequential method. The NMR data showed the presence of a triple-stranded beta-sheet (residues 19-25, 36-41, and 52-57), a short helix, and turns. These were applied as distance constraints in a molecular modeling, which includes distance geometry and dynamical simulated annealing calculations. A single family of structures is observed, which is folded in such a way that three major loops emerge from a globular head. The solution and crystal structures of alpha-cobratoxin are very similar [9].

Two novel alpha-type neurotoxins isolated from Taiwan cobra (*Naja naja atra*), possess distinct primary sequences and different postsynaptic neurotoxicity, were taken as prototype for short and long chain neurotoxins and compared with the major lethal short-chain neurotoxin in cobrotoxin, based on the derived 3-D structure of this toxin in solution by NMR spectroscopy and computer homology-modeling [22].

Furthermore, the conserved catalytic amino acid residues (TRP<sub>25</sub>, CYS<sub>26</sub>, CYS<sub>30</sub>, GLY<sub>34</sub>, ARG<sub>35</sub>, ARG<sub>36</sub>, ASP<sub>37</sub>) present in the 3-D structure of studied neurotoxins as well as the template 3D structure (2CTX), were predicted. This catalytic amino acid can be used for the drug target against Indian cobra, whereas additional catalytic amino acid residues (CYS<sub>20</sub>, HIS<sub>18</sub>, CYS<sub>3</sub>, CYS<sub>62</sub>) were also identified in the 3-D structure of neurotoxin A, B and D of Indian cobra. In the neurotoxin C, there is no any variable catalytic amino acid residue present. This catalytic residue may help to target and design the anti venom drugs against the neurotoxins (Table 3).

Torpedo acetylcholine receptor alpha 1 subunit was changed through mutagenesis in a recombinant fusion protein. Mutation of His-186, Tyr-189, Tyr-190, Cys-192, Cys-193, Pro-194, and Asp-195 greatly reduced or abolished alpha-bungarotoxin binding, while mutation of Tyr-198 reduced binding, indicating participation of these residues in binding, either through functional interaction with neurotoxin residues. Molecular modeling of acetylcholine receptor and knowledge of neurotoxin and receptor residues necessary for binding, allow analysis of possible structure-function relationships [23].

In conclusion, the investigation was done to generate and validate the 3-D structure of four neurotoxins of Indian cobra. This 3-D structure may provide a new insight to better understand the structure and function. The prediction of catalytic amino acid residues in the 3-D structure of neurotoxin may be helpful to design new antivenom drugs.

**REFERENCES**


