ISSN 0975-6299 Vol 2/Issue 1/ Jan-Mar 2011

International Journal of Pharma and Bio Sciences

RESEARCH ARTICLE

BIOINFORMATICS

PHYLOGENETIC ANALYSIS AND HOMOLOGY BASED INHIBITOR DESIGN FOR SHORT NEUROTOXINS OF FOREST COBRA

Corresponding Author

DEEPA NAGARAJAN

Gensilico Biosolutions, Zamin Pallavaram, Chennai- 600043, India.

Co Authors

GOWRI KARTHIK[1], RAJASEKAR DURAIVELU[2] AND KUMARAN SANTHALINGAM[2]

[1]Department of Biotechnology, School of Life Sciences, Vels University, Chennai 600117, India.
[2]Gensilico Biosolutions, Zamin Pallavaram, Chennai- 600043, India.

ABSTRACT

Naja commonly referred as Cobra is a genus of venomous elapid snakes found in the Middle East, India, and southern eastern Asia posse's complex mixture of toxin proteins which are found to have pharmacological activities. These neurotoxins bind specifically to the nicotinic acetylcholine receptors, thereby blocking synaptic nerve transmission. Short neurotoxin protein sequences of Cobra were retrieved from Swiss prot database and multiple sequence alignment was performed on the retrieved sequences using ClustalX. Phylogenetic relationships were analyzed using Phylip software and sequences were modeled through Modeller 9v8. The quality of the model was validated by the Structural analysis verification server. Tropine derived from Ashwagandha was investigated for its skeletal muscle relaxant property. The interaction between the inhibitors and the predicted structures of short neurotoxin were analyzed Insilco by ArgusLab. From this study we conclude that short neurotoxin of *Naja oxiana* showed the higher affinity towards Tropine.

ISSN 0975-6299 Vol 2/Issue 1/ Jan-Mar 2011

KEYWORDS

Short neurotoxin, Tropine, Phylogenetics, Modeller 9v8, ArgusLab.

INTRODUCTION

Snake venoms are complex mixtures of proteins, nucleotides and inorganic ions. These combinations confer a formidable array of toxic properties on the venom, the peptides and polypeptides being responsible for a variety of toxic properties. The number of venom components in venomous animals like snake, scorpion or cone snail ranges from 50-200 toxins[1]. Snake venoms are important tools in toxinology, neuroscience, and pharmacology. The venom components are highly variable and functionally complex and they offer many research opportunities [2]. The main toxins from snake venom that affect the CNS are neurotoxins. Neurotoxins form one of the largest families of proteins with established primary structures. Snake neurotoxin is a toxic agent or substance that inhibits damages or destroys the tissues of the nervous system and neurons. Neurotoxic proteins isolated from various snake venoms have high affinity for a particular target site are used extensively as pharmacological tools to gain insights into the function of the nervous system. The vast majority of snake venom neurotoxic peptides competitively bind to the nicotinic acetylcholine receptor. The potency of these molecules lies in their affinities towards the biomolecules involved in the functioning of neuromuscular transmission. Nicotinic acetylcholine receptors are prototypes for the pharmaceutically important family of pentameric ligand-gated ion channels [3]. Among the best studied snake neurotoxins are the á-neurotoxins that bind to nicotinic acetylcholine receptors (nAChRs). They are capable of reversibly blocking nerve transmission by competitively binding to the nAChR located at the postsynaptic membranes of skeletal muscles and neurons, preventing neuromuscular transmission and thereby leading to death by asphyxiation[4].

Fossil records revealed that the human use of plants as traditional medicine date back to middle Paleolithic age, approximately 60,000 years ago[5] . At present, natural products (and their derivatives and analogs) represent over 50% of all drugs in clinical use, in which natural products derived from higher plants represent ca. 25% of the total. The World Health Organization estimated that over 80% of the people in developing countries rely on traditional remedies such as herbs for their daily needs and about 855 traditional medicines include used crude plant extracts. This means that about 3.5 to 4 billion of the global population rely on plants resources for drugs [6]. Medicinal plants have provided a good source of a wide variety of compounds, such as phenolic compounds, nitrogen compounds, vitamins, terpenoids and some other secondary metabolites, which are rich in valuable bioactivities[7]. Computer-aided techniques for the efficient identification and optimization of novel molecules with a desired biological activity have become a part of the drug discovery process. Such a approach could lead to a reduction in the cost of drug design and development by upto 50%[8]

In our present study we focused on the evolutionary relationship of all short neurotoxins of forest cobras. A medicinal plant compound called tropine derived from *Withania somnifera* was investigated for skeletal muscle relaxant activity against *Naja* short neurotoxins.

MATERIALS AND METHODS

Swiss Prot Database

Swiss-Prot is a manually curated biological database of protein sequences. Swiss-Prot strives to provide reliable protein sequences associated with a high level of annotation, a minimal level of redundancy and high level of integration with other databases.

Phylogenetic Analysis

All the sequences were aligned with ClustalX 2. The computed alignment was manually checked and corrected. Pair-wise evolutionary distances were computed protdist implemented in the PHYLIP program and a phylogenetic tree was constructed by neighbor-joining method. Bootstrapped values of 100 were sampled to determine a measure of support for each node on the consensus tree**.**

Templete Identification for the Target molecule*:*

The template may be a predefined layout to give an idea about the unknown structure of the query molecule[9]. The NCBI BLAST was used to identify the template for modeling the three dimensional structure of short neurotoxin. The sequence of the target molecule in FASTA format was submitted for blastp against pdb database which yields the suitable template.

Homology modeling

Among all current theoretical approaches, comparative modeling is the only method that can reliably generate a 3D model of a protein from its amino acid sequence. Modeling of protein structures usually requires extensive expertise in structural biology and the use of highly specialized computer programs for each of the individual steps of the modeling process[10]. The method of homology modeling is based on the observation that protein tertiary structure is better conserved than amino acid sequence[11]. The three dimensional structure of short neurotoxin has been predicted using MODELLER9v8 (http://www.salilab.org/modeller/).

Model refinements and evaluation

The model generated by MODELLER9v8 was subjected to energy minimization using the steepest descent technique to eliminate bad contacts between protein atoms using swiss pdb viewer. Validation of modeled structure was carried out using Structure Analysis and Validation Server. It performs structure validation calculations using PROCHECK, PROVE, Verify3D, ERRAT and WHAT_IF programs. The validated result of the modeled protein from the server is an important part of comparative modeling process.

Active site prediction

After obtaining the final model, the possible binding sites of short neurotoxin were searched using Computed Atlas of Surface Topography of Proteins (CASTp). These include pockets located on protein surfaces and voids buried in the interior of proteins. CASTp includes a graphical user interface, flexible interactive visualization, as well as on-the-fly calculation for user uploaded structures[12].

Docking the inhibitors against the active site of the Short Neurotoxin

Docking is a computational technique that samples conformations of small molecules in protein binding sites; scoring functions are used to assess which of these conformations best complements the protein binding site[13].The inhibitor and target protein was geomentrically optimized and docked using docking engine ArgusDock.

RESULTS AND DISCUSSION

Sequence retrieval of short Neurotoxin:

The protein sequences of short neurotoxin were retrieved from the Swiss prot database having the Swiss prot id as listed in Table1.

Table 1

Protein details of short neurotoxins retrieved from Swissprot database

S.no	Species	ID	Length	Protein sequence
1	*Naja pallida*	P01426	61	LECHNQQSSQPPTTKTCPGETNCYKKVWRDHR GTIIERGCGCPTVKPGIKLNCCTTDKCNN
2	*Naja oxiana*	P01427	61	LECHNQQSSQPPTTKTCSGETNCYKKWWSDH RGTIIERGCGCPKVKPGVNLNCCRTDRCNN
3	*Naja nivea*	P01423	61	MICHNQQSSQRPTIKTCPGETNCYKKRWRDHR GTIIERGCGCPSVKKGVGIYCCKTDKCNR
4	*Naja melanoleuca*	P01424	61	MECHNQQSSQPPTTKTCPGETNCYKKQWSDH RGTIIERGCGCPSVKKGVKINCCTTDRCNN
5	*Naja nivea*	P68419	61	LECHNQQSSQPPTTKTCPGETNCYKKRWRDH RGSITERGCGCPSVKKGIEINCCTTDKCNN
6	*Boulengerina annulata annulata*	P34075	61	KICYNQPSSQHPTTKACPGEKNCYRKQWSDHR GTIIERGCGCPTVKPGVKLHCCTTEKCNN
7	*Boulengerina christyi*	P34076	62	MECHNQQSSQPPTTTHCSGGETNCYEKRWHD HRGTIIERGCGCPTVKPGVKLNCCTTDKCNN
8	*Naja samarensis*	P60774	61	LECHNQQSSQAPTTKTCSGETNCYKKWWSDH RGTIIERGCGCPKVKPGVKLNCCTTDRCNN
9	*Naja mossambica*	P01431	62	LECHNQQSSEPPTTTRCSGGETNCYKKRWRD HRGYRTERGCGCPTVKKGIELNCCTTDRCNN
10	*Naja kaouthia*	P14613	62	LECHNQQSIQTPTTTGCSGGETNCYKKRWRDH RGYRTERGCGCPSVKNGIEINCCTTDRCNN
11	*Naja haje haje*	P68418	61	LECHNQQSSQPPTTKTCPGETNCYKKRWRDH RGSITERGCGCPSVKKGIEINCCTTDKCNN

Phylogenetic Analysis

Phylogenetic analysis provides the most accurate reconstruction of evolutionary relationships and distances between short neurotoxin sequences. Phylogenetic methods are widely used method to estimate the evolutionary rates of genes and genomes to detect footprints of natural selection, and the evolutionary information is used to interpret genomic data[14].Multiple sequence alignment was done using clustalx and found all the sequence has snake toxin signature and it is conserved which is shown in Fig.1.

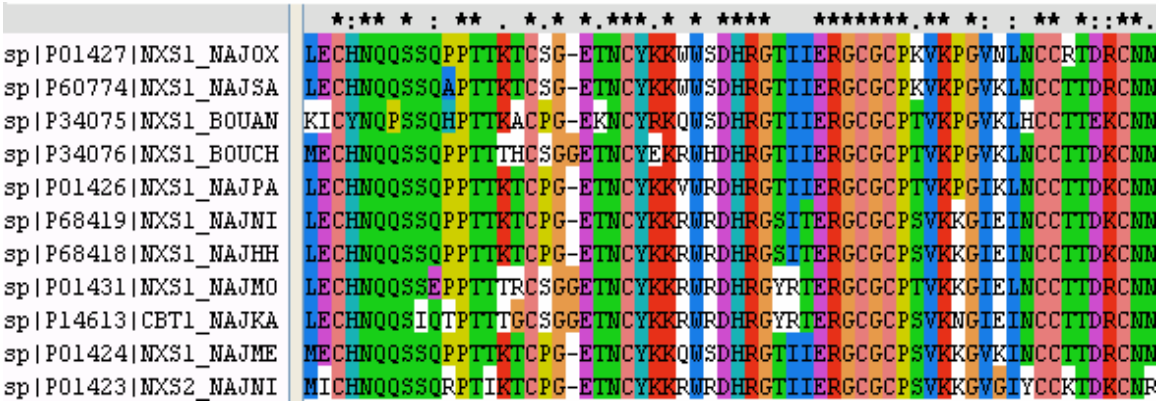


Fig. 1Multiple Sequence alignment of the target proteins which was carried out with clustalx. * represents the conserved region whereas the : represents the conserved substitution. . Represents the semi conserved substitution. From the above result, it was observable that nearly 90% of the sequences are similar.

Distance analysis of protein sequences was done to compute a distance matrix, under four different models of amino acid replacement. It can also compute a table of similarity between the amino acid sequences. The distance for each pair of species estimates the total branch length between the two species which is shown in figure 2.

Between	And	Length
1	sp|P60774|	0.01588
1	3	0.05682
3	4	0.01228
4	8	0.01783
8	sp|P01424|	0.04766
8	9	0.02030
9	sp|P01423|	0.16372
9	6	0.03965
6	7	0.07406
7	sp|P14613|	0.08065
7	sp|P01431|	0.04567
6	5	0.03460
5	sp|P68418|	0.00001
5	sp|P68419|	0.00000
4	sp|P01426|	0.03995
3	2	0.01420
2	sp|P34076|	0.07994
2	sp|P34075|	0.21311
1	sp|P01427|	0.03676

Figure 2: Pairwise distance of the short neurotoxin showing the branch Length between all the protein sequences.

Fitsch Analysis

```
  11 Populations

Fitch-Margoliash method version 3.69

                                                   2
                   __  __   (Obs - Exp)
Sum of squares =   \   \    -------------
                   /_  /_         2
                    i   j        Obs

Negative branch lengths not allowed

  +sp|P60774|
  !
  !       +--sp|P01424|
  !       !
  !     +-8 +----------sp|P01423|
  !     !  !  !
  !     ! +-9        +----sp|P14613|
  !     !    !    +---7
  !     !    !    !    +--sp|P01431|
  !   +-4    +---6
  !   !  !        !  +sp|P68418|
  !   !  !        +-5
  !   !  !           +sp|P68419|
  1--3  !
  !   !  +--sp|P01426|
  !   !
  !   !  +----sp|P34076|
  !   +-2
  !      +-------------sp|P34075|
  !
  +-sp|P01427|

Sum of squares =     1.54933

Average percent standard deviation =    11.97735
```

Fig.3: Shows the unrooted tree that is constructed using Fitsch analysis

Sum of squares (SSQ) is calculated as 1.54933. The APSD gives an indication of the average percentage error. The APSD was found to be 11.97735.

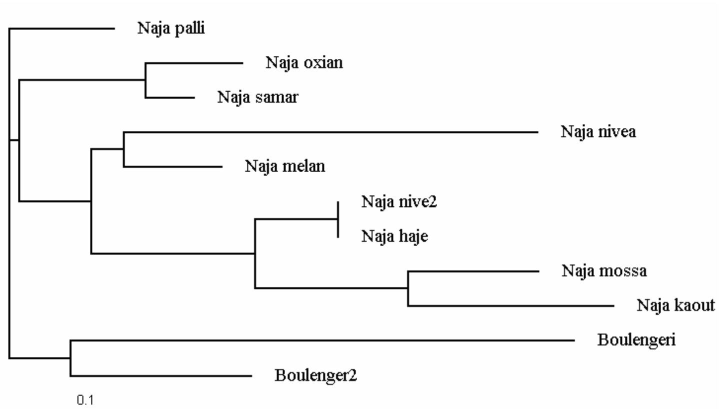


Fig.4: Showing the Phylogenetic tree constructed using NJ and UPGMA

The *B. annulata* and *B.cyteri, Naja kaout* and *Naja mossa, Naja melan* and *Naja nivea*, *Naja sammer* and *Naja oxia*n shows homologous respectively which arise from the common ancestor *Naja pallida*. *B.annu*lata was found to be outgroup for all the species (Fig.4).

HOMOLOGY MODELING OF TARGET STRUCTURES

The absence of the three dimensional structure for protein sequences of short neurotoxin in PDB prompted us to construct the 3D model of the protein. The three dimensional structure provides information into the function and also helps us to analyses of its interactions with the suitable inhibitors[15]. The various templates used for modeling were given in Table 2.

Table 2
Template structure for the target proteins

S.No	Species	Id	Templete	Id percentage
1	Naja melanoleuca	P01424	1IQ9	86
2	Naja nivea	P68419	1IQ9	88
3	Boulengerina annulata annulata	P34075	1IQ9	81
4	Boulengerina christyi	P34076	1IQ9	85
5	Naja samarensis	P60774	1ONJ	88
6	Naja mossambica	P01431	1V6P	90
7	Naja kaouthia	P14613	1V6P	99
8	Naja haje haje	P68418	1IQ9	88

The alignment of template and the target protein was performed using the script "align2d.py". Among the five models generated using the script “modelsingle. py” the thermodynamically stable model was chosen for further refinement and validation. The refined structure was the submitted in Structure validation and analysis server for validation. Ramachandran was used to visualize dihedral angles ø against ö of amino acid residues in protein structure[16] .The validation result were shown in table 3 which confirms that all the protein structure are thermodynamically stable.

Table 3
Energy values of the modeled Protein along with the RMS score

S.No	Id	Templete	Energy value	RMS Score
1	P01424	1IQ9	-3045.049	0.48
2	P68419	1IQ9	-3360.734	0.16
3	P34075	1IQ9	-3936.482	0.04
4	P34076	1IQ9	-3353.238	0.07
5	P60774	1ONJ	-2889.456	0.16
6	P01431	1V6P	-3396.821	0.05
7	P14613	1V6P	-3993.135	0.08
8	P68418	1IQ9	-3396.564	0.07

Docking between the target protein and the drug molecule*:*

Given the three dimensional structure of a target receptor molecule usually a protein, chemical compounds having potential affinity towards it are designed rationally, with the a id of computational methods[17,18]. Owing to the growing number of identified snake venom neurotoxin sequences, it is increasingly difficult to study them by experimentation alone. Detailed bioinformatics analysis offers a convenient methodology for efficient *in silico* preliminary analysis of possible function of new toxins[2]. The high specificity of neurotoxins for nAChRs has been utilized as a tool in understanding the structure and function of the nervous system[19]. The inhibitor Tropine derived from *Withania somnifera* is shown in figure 5.

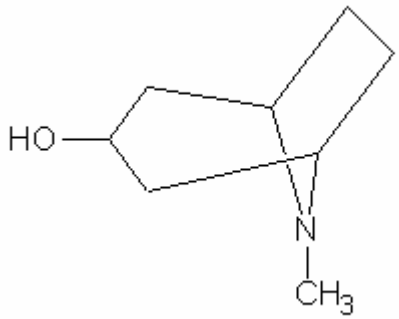


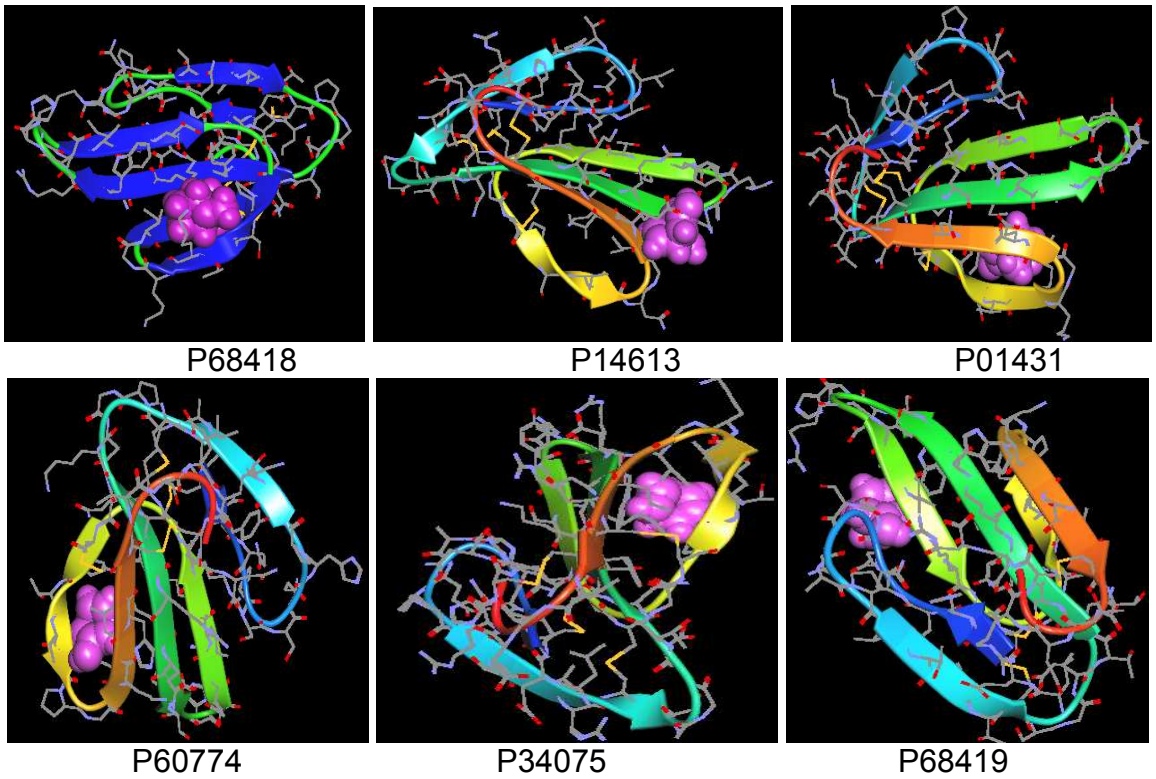
Fig: 5
Structure of Tropine

Tropine is a small sized molecule with a molecular weight of 141.2108(g/mol). It has one hydrogen bond donor and two hydrogen bond acceptors with no rotatable bonds. The compound Tropine has the LogP value of 0.477. Thereby it satisfies all the criteria of Lipinski's rule of five[20].

The scoring function takes a pose as input and returns the number indicating the likelihood that the pose represents a favorable

binding interaction. A low (negative) energy indicates a stable system and thus a likely binding interaction[21]. Fig. 6 shows the binding of ligand with the receptor molecule. The docking energy values were given in table 4.

S.No	Id	Docking Energy Level	No of Hydrogen Bonds	Hydrogen Bond Length	Bonded Residues
1	P01426	No result	nil	nil	Nil
2	P01427	-5.41696	3	1.9,2.2,2.2	Gln6, Gln10
3	P01423	-5.69051	1	2.8	Gln6
4	P01424	-6.55902	Nil	Nil	Nil
5	P68419	-5.99758	1	3.0	Gln7
6	P34075	-6.41765	1	3.0	Lys47
7	P34076	No result	Nil	Nil	Nil
8	P60774	-6.77322	1	2.1	Lys46
9	P01431	-6.42584	Nil	Nil	Nil
10	P14613	-5.33916	1	2.4	Arg30
11	P68418	-6.303	Nil	Nil	nil



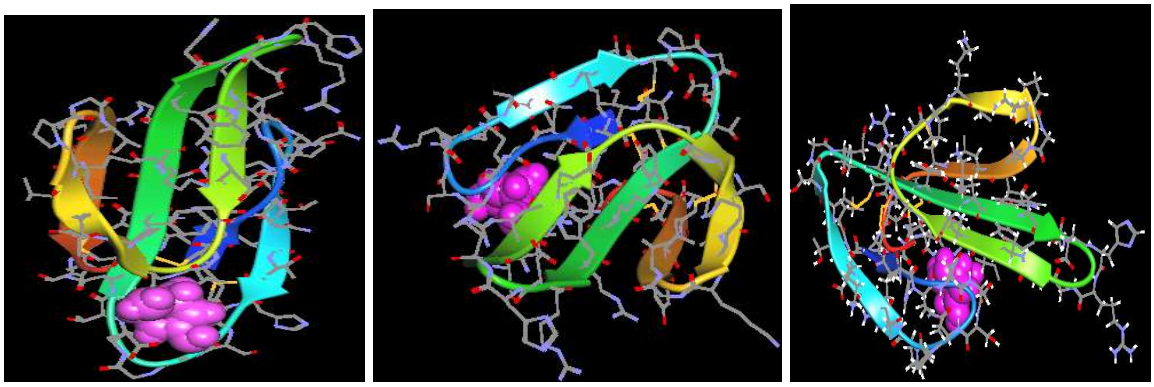


Fig 6: Docking of Short Neurotoxin with the Ligand Tropine

The short neurotoxin of Naja oxiana species showed the higher affinity towards the compound Tropine. It has the binding energy value of -5.41696(KJ/Mol) with three hydrogen bonds. The first hydrogen bond was formed between the Tropine and Gln6 with the distance of 1.9A. The second and third hydrogen bond is formed between Tropine and Gln10 with the distance of 2.2 and 2.2 A respectively which is shown in Fig.7.

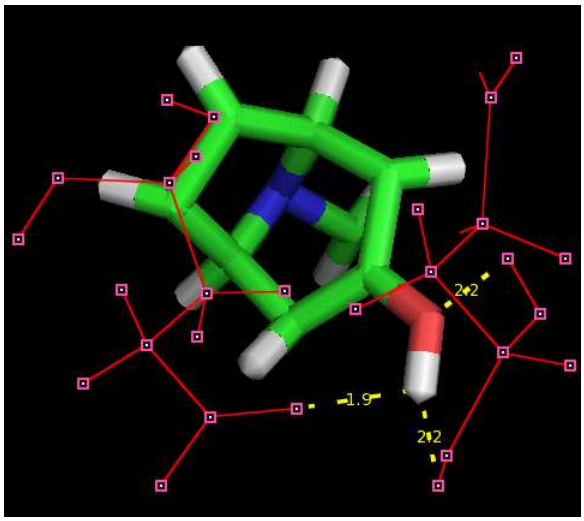


Fig 7: Shows the Hydrogen bond between the Tropine and the active site amino acids (Gln6 and Gln10)

The previous study showed that Anabesine has higher specificity and efficiency towards the short neurotoxin of Naja melanoleuca[22]. The present study showed that docking of Tropine has higher binding affinity towards the target protein short neurotoxins which causes

paralysis. The short neurotoxin of *Naja oxiana* showed the higher affinity towards Tropine. Hence Tropine could act as a better skeletal muscle relaxant.

CONCLUSION

Evolutionay studies of short neurotoxin sequence of the cobra conclude that *Naja pallida* found to be the common ancestor of all the species. The inhibitor Tropine derived from *Withania somnifera* satisfies all lipinskys's rule of five. The short neurotoxin of *Naja oxiana* showed the higher affinity towards Tropine. Hence Tropine could act as a better skeletal muscle relaxant.

REFERENCE

1. Tan PT, Khan AM and Brusic V, Bioinformatics for venom and toxin sciences. Brief Bioinform, 4(1):53-62, (2003).
2. Joyce P.Y.S, A.M. Khan, Paul T. J. Tan, J.L.Y. Koh, S.H Seah, C.Y Koo, S.C. Chai, A. Armugam, V. Brusic, and K. Jeyaseelan, Systematic analysis of snake neurotoxins functional classification using a data warehousing approach.Bioinformatics, 20: 3466 – 3480, (2004).
3. Celie P.H.N, S.E Van Rossum-Fikkert, W.J Van Dijk, K.Brejc, A.B.Smit and T.K Sixma, Nicotine Carbamylcholine Binding to Nicotinic Acetylcholine Receptors as Studied in Achbp Crystal Structures. Neuron, 41:907-914, (2004)
4. Tsetlin V.I. and F. Hucho, Snake and snail toxins acting on nicotinic acetylcholine receptors: fundamental aspects and sciences. Brief. Bioinform, 4: 53–62, (2004).
5. Solecki R and I.V Shanidar, A Neanderthal flower burial in northern, Iraq. Science 190:880–881, (1975)
6. Farnsworth, N.R, In: *Human Medicinal Agents from Plants.* ACS Symposium Series, 534: 2-12, (1998).
7. Maridass.M and A. John De Britto, Origins of Plant Derived Medicines Ethnobotanical Leaflets 12: 373-387,(2008).
8. DiMasi, J.A., Ronald W. Hansenb and Henry G. Grabowskic, The price of innovation: New estimates of drug development cost. J. Health Econ 22:151-185, (2003).
9. Brindha.V, A. Saravanan and R. Manimekalai, Drug designing for ring finger protein 110 involved in adenocarcinoma (human breast cancer) using casuarinin extracted from Terminalia arjuna. Indian Journal of Science and Technology 2:22-26,(2009).
10. Tramontano,A., Leplae,R. and Morea,V, Analysis and assessment of comparative modeling predictions in CASP4. Proteins, 45 (Suppl. 5),22–38, (2001).
11. Marti-Renom,M.A., Stuart,A.C., Fiser,A., Sanchez,R., Melo,F. and Sali,A, Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct., 29, 291–325,(2000).
12. Binkowski TA, S. Naghibzadeh and J. Liang, CASTp: Computed Atlas of Surface Topography of proteins. NucleicAcids Res. 31:3352-5, (2003).
13. Warren G.L, C.W Andrews, A.M Capelli, B. Clarke, J. LaLonde, M.H Lambert, M. Lindvall, N. Nevins, S.F Semus, S. Senger, G. Tedesco, I.D Wall, J.M Woolven, C.E Peishoff and M.S Head,A

critical assessment of docking programs and scoring functions. J Med Chem. 49(20): 5912-31, (2006).
14. Yang Z. The power of phylogenetic comparison in revealing protein function. Proc Natl Acad Sci USA. 102:3179–3180,(2005).
15. Natarajan Vidya, Bagavathkumar Vadivukkarasi*, Gopalakrishnan Manivannan and Krishnan Anbarasu, Molecular modeling and docking studies of glutamate racemase in *Vibrio vulnificus* CMCP6. Insilicobiology 8:0036,(2008*).*
16. Ramachandran G.N, C. Ramakrishnan and V. Sasisekharan, Stereochemistry of polypeptide chain configurations. J. Mol. Biol. 7: 95–9, (1963).
17. Jorgensen, W.L, The many roles of computation in drug discovery. Science.303:1813-1818, (2004).
18. Ooms, F, Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. Current Medicinal Chemistry, Vol. 7, No.2.pp.141-158,(2000).
19. Koh D.C.I., A. Armugam and K. Jeyaseelan, Snake venom components and their applications in biomedicine. Cell. Mol. Life Sci. 63,Pages 3030–3041, (2006).
20. Christopher A. Lipinski, Franco Lombardo, Beryl W. Dominy and Paul J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews. Volume 23, Issues 1-3, 15, Pages 3-25,(1997).
21. Kitchen DB, H. Decornez, J.R Furr and J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications. Nature reviews Drug discovery. 3(11), 935 – 949, (2007).
22. Gowri Karthik, Vidya Venkatasubramanian ,Gomathi Ravichandran, Karthikeyan Subbarayan, and Deepa Nagarajan, Homology Based Inhibitor Design For Naja melanoleuca Short Neurotoxin Using Plant Based Drugs. Journal of Herbal Science and Technology .6:16-20, (2010).