Innovare
Academic Sciences

International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 7, Issue 4, 2015

Original Article

NEUROTOX: A UNIQUE DATABASE FOR ANIMAL NEUROTOXINS

R. BARANI KUMAR*, M. XAVIER SURESH

Department of Bioinformatics, Faculty of Bio and Chemical Engineering, Sathyabama University, Chennai 600119, India
Email: baranikumar22@gmail.com

Received: 22 Oct 2014 Revised and Accepted: 20 Nov 2014

ABSTRACT

Objective: A wide range of toxic effects against other species are affected through potential protein molecules called neurotoxins. Among various kinds of toxins, neurotoxins play a critical role in life of other organisms (especially humans). So it is relatively important to know more about neurotoxins and their action on biological systems. All around the world several research works are progressing in the area of neurotoxins.

Methods: Here we present NEUROTOX, a database which contains more than 1400 toxin entries originating from more than ~248 different species. The data was enriched with the related data derived from various other protein sequence as well as structural databases and several literature resources. Each entry of toxin contains 18 fields providing detailed information on neurotoxins and their source, fasta formatted protein sequences, molecular weight, toxicity information along with their literature references.

Results: In addition, 1086 computationally resolved 3D structures of toxin proteins with validation results were made available for download.

Conclusion: The easy to use and freely accessible NEUROTOX database is available online at http: //www. neuro-tox. org and it provides the large range data on toxins.

Keywords: Neurotoxins, Biological databases, Protein modeling, RDBMS, Toxin databases.

INTRODUCTION

Among a large number of toxins available in nature, neurotoxins play a very important role in affecting the nervous system either directly or indirectly of the other species especially the main target and threats is a human [1]. Toxins from animal origin produce large numbers of venomous proteins and most of them are species specific. The tremendous amount of research work is progressing across the world against toxins and toxin bites. Since several toxins are having high economic and medicinal values to treat varieties of diseases from chronic pains to anticancer therapies and most of the toxins are tested and used as a drug (for eg., ω-conotoxins and its variant are used to treat severe chronic pain and neuropathic pains) [2, 3]. Varieties of toxins from snakes, spiders, sea anemones and cone snails are acting as a modulator for voltage gated Na^+ion channels [4]. Toxins have been used as a tool to explore the structural and functional role of ion channels.

More number of biological databases available on World Wide Web and each one concern about a purpose and expertise. In the similar way toxin databases are also present on the web for a particular theme or areas of thrust. For example, animal toxin database (ATDB) and Tox-Prot, provides the information on toxins, which are functionally annotated [5, 6]. Toxin target database (T3DB) is a database providing the basic information on toxin from both natural (from venomous species) and synthetic origin (pesticides, pollutants, drugs and food toxins) [7]. In addition to these databases, some of the databases were developed specific to the species. Toxin-antitoxin database (TADB) is an integrated database containing type 2 toxin and antitoxin gene loci and its related genomic information of bacteria and archaea [8]. It includes Arachnoserver, manually curated database for spider toxins. This database contains sequence and structural information on arachinidae family toxins (spider toxins) [9]. Conoserver is a database which contains information on cone snails and their conopeptides. It covers most of the marine cone snails data, sequence and structures of conopeptides [10]. SCORPION is a database having the information of toxin sequences, structures, and functions of each toxin belong to scorpion species [11]. DBETH is another kind of toxin database providing information on bacterial exotoxins and their sequence, structure, interaction network and analytical information [12]. MvirDB database contains information on microbial infection and virulence nature. Genes responsible for antibiotic resistance, etc., were given in a detailed way with their primary data resources [13].

Though there are several databases on toxins, the neurotoxin data are scattered across public databases which provide sequence descriptions, but very limited structural and functional annotation. The exponential growth of newly identified toxin data has created a need for better data management and currently there is no specific resource available for this relatively important class of proteins. Hence we report a database which contains a comprehensive collection of all the neurotoxin data and integrated with the primary resources like Uniprot [14], Protein Data Bank (PDB) [15], PubMed, etc. This database provides the enriched information of neurotoxins sequence and structures.

MATERIALS AND METHODS

Data collection and integration

All neurotoxin data were collected from various resources, redundant data are removed and further cross checked thoroughly to ensure the quality of each data that is included in the database. The major repositories used for data collection are Uniprot, Pfam, Protein data bank (PDB). The primary data such as sequence and structure are retrieved from Swissprot and PDB. Other data such as literature are collected from PubMed resource and other publicly available open access literature resources. Prior to database development all important attributes are set for trouble-free way of sorting the collected data. Appropriate care has been given to have an error free data reserve. Each entry has been manually inspected before incorporating into the data base. The schematic overview and the data flow of the database are represented in fig.1.

Database description

This database is populated with large sets of data, including 1400 toxin entries from 248 different species which includes 80 species of snakes, 56 species of scorpions, 46 species of spider, 36 species of venomous snails, 22 species of sea anemones, 5 species of ferocious and venomous honeybees, 2 species of lizard and 1 species of wasp. The distribution of neurotoxins across various source organisms that are covered in this database is shown in fig. 2. The entire data are arranged as relational database management system with MySQL. The sequential, functional description, source organism, Fasta formatted protein sequence, structures (if found), structural homologs obtained from similarity search for every toxin, molecular weight, protein family classification and supporting literature data obtained from PubMed database (http: //www. ncbi. nlm. nih. gov/pubmed/)are also included in this database. In addition to these fields, users can directly access the primary databases like NCBI, Uniprot, Pfam, PDB through the links provided in the NEUROTOX database. The sequence length distribution is shown in fig. 3.

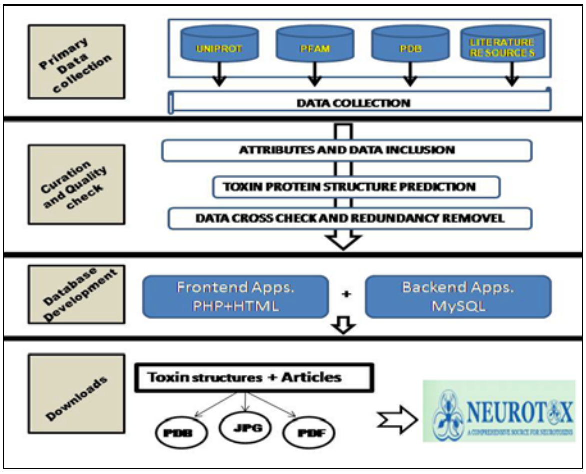

Fig. 1: Data flow in NEUROTOX database

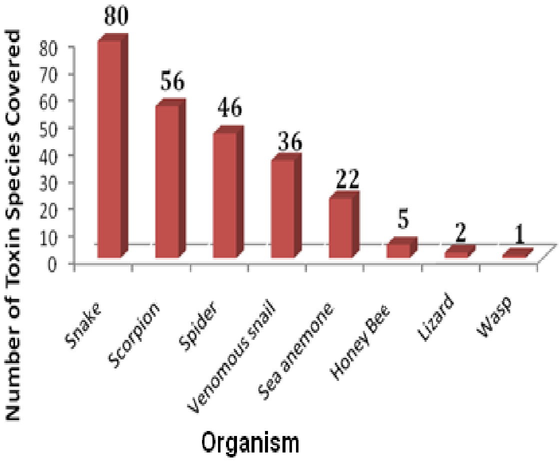

Fig. 2: Distribution of toxins and their source organisms curated in NEUROTOX database

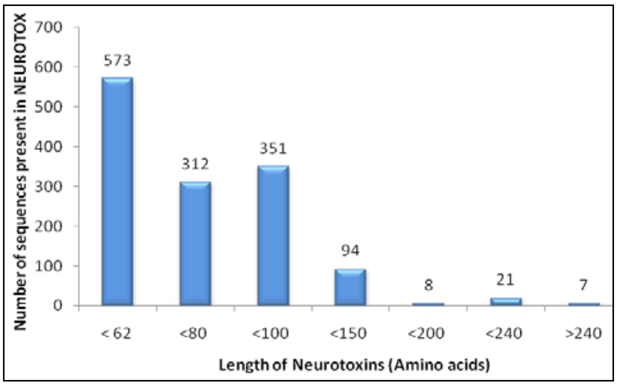

Fig. 3: Sequence length data distribution in NEUROTOX database

The search, structure, references, menus integrated into the database portal will provide the maximal information about the particular toxins. The search menu, allows the user to search a particular toxin entry in three different ways, by using (i) toxin name, (ii) Uniprot identifier and (iii) NTx database identifier. Once the search is executed, the user can read the hit page of toxin. Each toxin hit contains the information arrayed in 19 fields of data, including the primary data extracted from Uniprot database [14].

NEUROTOX database also contains computationally predicted structures (homology models) of K^+channel blocker toxin families, short chain neurotoxins (SCNs) and almost all toxins entries of NEUROTOX were modeled using Modeller 9v7 [16,17]. In structure menu, users can access and download all the modeled toxin structures (initially 1086 structures are predicted) in the form of. JPG and. PDB (Brookheaven) formats. It provides the details about the toxin protein name, amino acid length and reported function. In addition to this, all modeled structures were validated and the resulting Ramachandran plot obtained from PROCHECK program can also be downloadable as PDF format [18]. The reference menu provides the downloadable free research articles about neurotoxins research.

RESULTS AND DISCUSSION

Database implementation

NEUROTOX is a Java based web application, the front end was designed using PHP, and it interfaced to a MYSQL as the backend application. PHP was connected to the web using HTML and it provides a user friendly application environment for both design and wide access of users. This database is available online with user-friendly platform and it provides advanced search methods to obtain a specific taxin entity in much easier way. In addition to this, it supports all web browsers to access this database worldwide.

Features of NEUROTOX database

NEUROTOX provides information of neurotoxins which are updated on a regular basis, such as structural information of toxins. Each toxin record show, nineteen toxin attributes along with literature references for each toxin. NeuroTox database allows the user to fetch the prompt information in three different ways like, by using (i) toxin name (ii) Uniprot identifier (iii) NTxdb accession number (Neurotox database own identifier number). In addition to this, user can directly access PDB, NCBI, Uniprot and Pfam databases.

For instance, the modeled structures of toxins (K^+channel blocker toxins and short chain neurotoxins) and their 3D coordinates are available in NEUROTOX. The structures were modeled through computational homology modeling using Modeller version 9v7. In addition to these all modeled proteins are easily downloadable in. PDB (RCSB) format and all the structures have been validated using PROCHECK and resulting Ramachandran plots are also available in downloadable form. Along with structures, some of the useful open access research articles are collected and included in this database which will be useful to the neurotoxin research community. NEUROTOX provides information of neurotoxin along with predicted structures of neurotoxins (1086 computationally resolved neurotoxin structures). An overview of the total numbers of structures, both experimentally determined and computationally derived, which are available in the database, is given in fig. 4. In addition it also provides full text research articles relevant to neurotoxin research which will provide a comprehensive data of neurotoxins.

NEUROTOX Web interface

NEUROTOX consists of a collection of primary information along with protein sequences, reported functions, predicted structures of the whole toxin protein family, especially K+channel blockers toxin and short chain neurotoxin families along with literature collections are in downloadable forms. The important feature of NEUROTOX is the integrated animal neurotoxin database. Protein structures menu provides the information about the toxin protein name, length of protein and reported functions. Three dimensional structure of a particular toxin protein image is displayed along with downloadable form of protein structure in a protein data bank format and Ramachandran plot in PDF and protein structure images in JPG formats. The sample screen shots of NEUROTOX database are given in fig. 5. The search menu (fig. 5) allows the user to search a particular toxin entry in three different ways, by using (i) toxin name, (ii) Uniprot identifier and (iii) NTx database identifier (NEUROTOX own identifier). Once the parameters are inserted in the search boxes, search is executed; user can be directed to the hit page where the results of the search are listed. From this list user can select a result. Each toxin entry/ result page contains the information arrayed in all attributes of data.

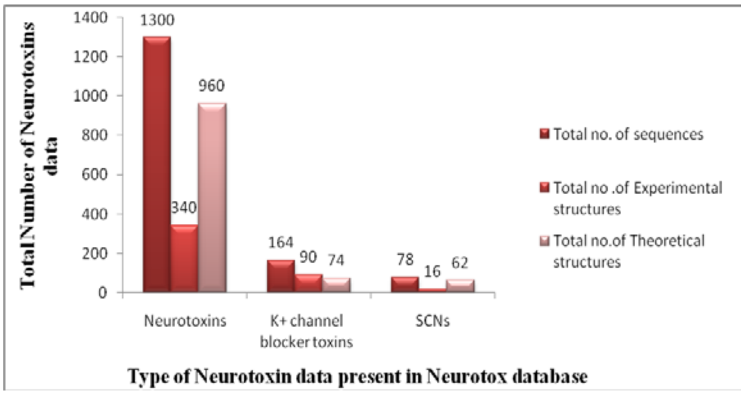

Fig. 4: Experimental vs. theoretical structural data in NEUROTOX database

Fig. 5: Snap shot of NEUROTOX web interface

CONCLUSION

NEUROTOX is an exclusive database for animal neurotoxins and it provides lots of information along with three dimensional structures. Due to several research activities on toxins like sequencing, characterization and other biological assays lead to sequence deposition rate in public repositories. Information about neurotoxins is covered in the huge data warehouses; however, there is a need to segregate the information present in the database based according to the user requirements. NEUROTOX database is a unique web server provides the data on animal neurotoxins and it provides enormous amount of information along with three dimensional structures. Data has been annotated at the structural level through computational structure prediction methodologies. It is speculated that this will provide a great insight into the structural biology studies and will assist to predict the function of each toxins.

In addition to the current status, we plan to add new entries of neurotoxin to the database and predict and validate the toxins structure for the new entries of NEUROTOX and add more information on receptor-toxin interaction and information about existing drugs for each toxin and exploring the ligands for each toxins and toxin-ligand interaction details will be added in this database in future.

ACKNOWLEDGEMENT

The authors thank the Management of Sathyabama University, Chennai for setting up the computational facilities in the Cluster Computing Laboratory, Department of Bioinformatics. The authors also thank database providers whose links have been used and the anonymous reviewers for their valuable comments and suggestions.

CONFLICT OF INTERESTS

Declared None

REFERENCES

1. Schiavo G, Matteoli M, Montecucco C. Neurotoxins Affecting Neuroexocytosis. Physiol Rev 2000;80:718-66.
2. Shen GS, Layer RT, McCabe RT. Conopeptides: from deadly venoms to novel therapeutics. Drug Discovery Today 2000;5:98–06.
3. Dutton JL, Craik DJ. Alpha-Conotoxins: nicotinic acetylcholine receptor antagonists as pharmacological tools and potential drug leads. Curr Med Chem 2001;8:327-44.
4. Rogers DF. Scorpion venoms: taking the sting out of lung disease. Thorax 1996;51:546–8.
5. He QY, He QZ, Deng XC, Yao L, Meng E, Liu ZH, *et al*. ATDB: a uni-database platform for animal toxins. Nucleic Acids Res 2008;36:D293–7.
6. Jungo F, Bairoch A. Tox-Prot, the toxin protein annotation program of the Swiss-Prot protein knowledgebase. Toxicon 2005;45:293–01.
7. Lim E, Pon A, Djoumbou Y, Knox C, Shrivastava S, Guo AC, *et al*. T3DB: a comprehensively annotated database of common toxins and their targets. Nucleic Acids Res 2010;38:D781–6.
8. Shao Y, Harrison EM, Bi D, Tai C, He X, Ou HY, *et al*. TADB: a web-based resource for Type 2 toxin–antitoxin loci in bacteria and archaea. Nucleic Acids Res 2010;39:D606–11.
9. Herzig V, Wood DLA, Newell F, Chaumeil PA, Kaas Q, Binford GJ, *et al*. Arachno Server 2.0, an updated online resource for spider toxin sequences and structures Nucleic Acids Res 2011;39:D653–7.
10. Kaas Q, Westermann JC, Halai R, Wang CK, Craik DJ. Cono Server, a database for conopeptide sequences and structures. Bioinformatics 2008;24:445–6.
11. Tan PT, Veeramani A, Srinivasan KN, Ranganathan S, Brusic V. Scorpion2: a database for structure-function analysis of scorpion toxins. Toxicon 2006;47:356–63.
12. Chakraborty A, Ghosh S, Chowdhary G, Maulik U, Chakrabarti S. Dbeth: A database of bacterial exotoxins for human. Nucleic Acids Res 2012;40:D615–20.
13. Zhou CE, Smith J, Lam M, Zemla A, Dyer MD, Slezak T. MvirDB-a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. Nucleic Acids Res 2007;35:D391–94.
14. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, *et al*. The Universal Protein Resource (Uni Prot). Nucleic Acids Res 2008;33:D154–9.
15. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, *et al*. The protein data bank. Nucleic Acids Res 2000;28:235–42.
16. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779-15.
17. Barani Kumar R, Shanmuga Priya B, Xavier Suresh M, Narendrakumar G. Homology modeling of short chain neurotoxins: an initiation towards understanding their functional inference. Int J Pharm Pharm Sci 2014;6:164-8.
18. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. Procheck: a program to check the stereochemical quality of protein structures. J Appl Cryst 1993;26:283-91.