Systematic analysis of snake neurotoxins’ functional classification using a data warehousing approach

Joyce Phui Yee Siew¹, Asif M. Khan¹,², Paul T. J. Tan¹,², Judice L. Y. Koh², Seng Hong Seah², Chuay Yeng Koo¹, Siaw Ching Chai¹, Arunmozhiarasi Armugam¹, Vladimir Brusic² and Kandiah Jeyaseelan¹,*

¹Department of Biochemistry, Faculty of Medicine, National University of Singapore, 8 Medical Drive, Singapore 117597 and ²Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613

Received on June 4, 2004; revised on July 8, 2004; accepted on July 16, 2004
Advance Access publication July 22, 2004

ABSTRACT
Motivation: Sequence annotations, functional and structural data on snake venom neurotoxins (svNTXs) are scattered across multiple databases and literature sources. Sequence annotations and structural data are available in the public molecular databases, while functional data are almost exclusively available in the published articles. There is a need for a specialized svNTXs database that contains NTX entries, which are organized, well annotated and classified in a systematic manner.

Results: We have systematically analyzed svNTXs and classified them using structure–function groups based on their structural, functional and phylogenetic properties. Using conserved motifs in each phylogenetic group, we built an intelligent module for the prediction of structural and functional properties of unknown NTXs. We also developed an annotation tool to aid the functional prediction of newly identified NTXs as an additional resource for the venom research community.

Availability: We created a searchable online database of NTX proteins sequences (http://research.i2r.a-star.edu.sg/Templar/DB/snake_neurotoxin). This database can also be found under Swiss-Prot Toxin Annotation Project website (http://www.expasy.org/sprot/)
Contact: bchjeya@nus.edu.sg

1 INTRODUCTION
The number of venom components in venomous animals such as snakes, scorpions or cone snails ranges from 50 to 200 toxins (Tan et al., 2003). The natural library of toxins is thus estimated to contain millions of different toxins and variants. Toxin entries in the public DNA and protein databases represent only a small fraction of <1% of the estimated natural library (Tan et al., 2003). 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 functions of new toxins. The in silico approach can assist in designing experiments for functional characterization of newly identified snake venom neurotoxin (svNTX) sequences, particularly those identified as novel cDNAs.

Sequence annotations, functional and structural data on svNTXs are scattered across multiple databases and literature sources. Sequence annotations and structural data are available in the public molecular databases, while functional data are almost exclusively available in the published articles. There is a need for a specialized svNTX database that contains neurotoxin entries, which are organized, well annotated and classified in a systematic manner. Data warehousing has been used in molecular biology for creating subsets within existing database (Sorace and Canfield, 1998; Bassett et al., 1999). A specialized data warehouse (which contains a family of proteins that have similar structures and functions) stands in contrast to a general-purpose database (which contain all known proteins sequences; Schönbach et al., 2000). The integrated environment of a data warehouse allows consistency in naming conventions, measurements of variables and encoding structures (Schönbach et al., 2000). Thus, a specialized database for snake neurotoxins would assist in analytical tasks compared with other generalized sequence databases, such as GenBank and Swiss-Prot, which are primarily sequence repository and are not subject-oriented (Schönbach et al., 2000).

Dufton (1984) had systematically examined 139 homologous short and long svNTXs and cytotoxins from elapid snake venoms, and grouped the toxins according to similarity.
Fry et al. (2003) employed a phylogenetic analysis to study Elapids three-finger toxins and demonstrated that these three-finger proteins can be classified into several groups based on an evolutionary relationship. In spite of these efforts, no clear picture has yet emerged to explain how the various toxin types are functionally related to each other. A major problem is the current limited understanding of structure–function relationship of the NTX groups. Our approach in combining data warehousing and venom toxins study could provide a useful method in studying venom toxins from other species, such as bees and wasps, spiders and jellyfish. So far, three specialized venom databases have been built using a data warehousing approach from scorpions (Srinivasan et al., 2002), mollusks and snakes (accessible at http://tw.expasy.org/sprot/tox-prot/). To fill this gap, we have systematically analyzed svNTXs and classified them using structure–function groups based on their structural, functional and phylogenetic properties. Using conserved motifs in each phylogenetic group, an intelligent module for the prediction of structural and functional properties of unknown NTXs was built.

2 MATERIALS AND METHODS

2.1 Data collection and cleaning

Neurotoxin amino acid sequences were retrieved from public databases, such as GenBank, EMBL, DDBJ, PIR, Swiss-Prot and Trembl. The neurotoxin three-dimensional (3D) structures were extracted from the Protein Data Bank (PDB). The summary of entries extracted from various data sources is given in Table 1. Throughout the text we have used Swiss-Prot accession numbers to identify sequences unless otherwise mentioned.

Data processing and storage were further enhanced and facilitated by the BioWare system that is available at http://sdmc.i2r.a-star.edu.sg/Templar. This user-friendly application helps to extract data from multiple sources, annotate entries and organize data in a searchable database format. BioWare uses three main modules namely BioWare-Retrieve, BioWare-Preparation and Templar. BioWare-Retrieve extracts the required data from GenBank, Swiss-Prot and PDB data sources, whereas BioWare-Preparation allows manual annotation of the data records prior to creation of the data file. The final module creates a searchable database according to a given set of the users and the data file uploaded.

The initial retrieval of svNTX sequences revealed 900 known NTX entries from public databases using the keywords search such as ‘Serpentes AND neurotoxin’. Subsequently, we performed preliminary filtering to remove redundant amino acid fragments and duplicates. The BioWare-Preparation module that generates a sequence comparison report summary based on pairwise alignment of entries in the dataset. Entries having 100% amino acid sequence identity were considered as duplicates and merged into single entries. Partial sequences were cross-checked with the literature to determine their uniqueness and have been retained if found to be unique. The final, cleaned dataset is available online at http://research.i2r.a-star.edu.sg/Templar/DB/snake_neurotoxin.

2.2 Data annotation

Nearly 300 published journal articles and papers listed in the database were analyzed for data enrichment with functional and structural enrichment. Six new fields containing information normally absent in public database entries were created for each entry. These fields were Action_site, Physiological_function, Critical_residues, Toxin_activity, Binding_affinity and Miscellaneous. Action_site denotes the site of action of svNTX. Physiological_function and Toxin_activity refer to published svNTX function and LD50 values, respectively. Critical_residues reveal the critical amino acid residues involved in toxin–receptor interaction. Information from binding studies performed on svNTX-receptor is represented by the Binding_affinity field. Miscellaneous field refers to other additional information relevant for functional or structural description of an svNTX. The description of data annotation is further elaborated under discussion.

2.3 Data analysis

The svNTX data were divided into two main groups: presynaptic and postsynaptic NTXs. A classification
chart that incorporated molecular target and phylogenetic information was built to produce unique structure–function groupings. NTxs sharing the same structure–function information have been classified in an individual subgroup.

Several methods such as disulfide pairing bridge patterns, pharmacological function and evolutionary relationship were used in the analysis. Disulfide pairing pattern was generated using CysAlign tool (Lenffer et al., 2004) available in the BioWare software. CysAlign is a module in the BioWare system that searches through a flat file of multiple entries for information on disulfide pairing and returns a plot of disulfide pairing bridges. Entries with identical disulfide bridge patterns, as reported in the database entry annotations, are grouped together by this tool.

Signal and leader sequences were removed and only mature, full-length amino acid was used for phylogenetic studies. Phylogenetic studies were undertaken only for groups having more than four or more members for better statistical realiability. CLUSTAL-X (Thompson et al., 1997) was used for multiple alignment of amino acid sequences. The pairwise alignment parameters used were a gap opening penalty of 35 and a gap extension penalty of 0.30, whereas the multiple alignment parameters were a gap opening penalty of 15 and a gap extension penalty of 0.30.

Each phylogenetic subgroup was generated by using the SEQBOOT program of PHYLIP 3.6a2 (Felsenstein, 1985, 2001). Statistical reliability of the trees was assessed using 100 bootstrap replications. The bootstrap values of 90 and above are considered to be significant. The most parsimonious trees were calculated using the PROTPARS program using the maximum parsimony (MP) method (Saitou and Nei, 1987). MP method was chosen because this program takes into account the number of changes required at the nucleotide level to substitute one amino acid for another. In this scheme, the most accurate phylogenetic tree is one that is based on the fewest changes in the genetic code. The strict consensus trees were obtained by using the CONSENSE program (Felsenstein, 2001). The rooted tree diagrams were generated with the TREEVIEW program (Page, 1996). Bee venom PLA2 (Swiss-Prot A59055) and Lynx1 (Swiss-Prot AAF16899) were utilized as outgroups for the analysis of presynaptic and postsynaptic NTX subgroups, respectively. In this study, we did not attempt to differentiate between convergent and divergent processes.

2.4 Experimental validation: purification of a novel weak neurotoxin Bc-wntx4 from Bungarus candidus

Crude venom (4 mg) from B. candidus was purified on reverse phase–high-performance liquid chromatography (RP–HPLC) (Smart system, Pharmacia) using Jupiter C4 column. Buffer systems used were 0.1% trifluoroacetic acid (TFA) (Buffer A) and 80% acetonitrile in 0.1% TFA (Buffer B). Rate of flow was 1 ml/min. RP–HPLC purified proteins were subjected to N-terminal amino acid sequencing mass spectrometry (MS). One protein fraction represented Bc-wntx4 and this was confirmed by electron spray ionization MS (ESI–MS) analysis (data not shown). Amino acid sequence of Bc-ntx4 is LTCLICPEKYCQKVHTRDDEENLCVKRFYEGKRFGEKYPRGCAATLPeadPHEIVECCSTDCNK (Accession number AY611643).

2.5 Experimental validation: acetylcholine receptor binding assays

Preparation of muscle nAChR receptor from Torpedo californica was carried out according to the procedure described by Ishikawa et al. (1977). The preparation of PC12 membranes containing α-7 neuronal nAChR was carried out according to the method described by Meyer et al. (1998). Competitive binding assays were performed using nAChR. Increasing concentrations of unlabeled Bc-ntx4 was added to a single concentration of 125I-α-Bgt (5 nM) to study their inhibitory effects on high-affinity toxin binding. An aliquot of 2.5 µg of Torpedo membrane was incubated with 5 nM125I-α-bungarotoxin at room temperature for 1 h in the presence of increasing concentration of Bc-ntx4 in a total volume of 200 µl. Reaction was terminated on ice. The membranes were recovered by centrifugation, washed with 1 ml of buffer containing 0.1% BSA and dried before subjecting to radioactive monitoring in a Packard, COBRA Auto Gamma Counter (Packard Instruments Co. Inc.).

3 RESULTS AND DISCUSSIONS

3.1 Snake venom neurotoxin database (svNTXdB)

The svNTXdB provides a unique compilation of these toxins collected from public databases and literature sources. Each entry was analyzed for possible errors and inconsistencies, and annotated with functional information. This database served as a convenient tool for information storage and retrieval allowing analysis of svNTxs. The entries in svNTXdB have been classified and annotated to enhance new knowledge discovery. This database contained a set of tool including structure–function prediction module. As of March 31, 2004, the svNTXdB contained 272 unique snake NTX sequences.

Snake NTxs have five major search or extraction tools: BLAST search, Query svNTX, Structure viewer, Download FASTA and Annotate svNTX. BLAST search (Altschul et al., 1997) allows user to perform a sequence similarity search against snake NTX database. The user can specify output alignment result in either standard BLAST output or color-coded multiple sequence alignment generated using Mview program (Brown et al., 1998). The Query svNTX feature
allows users to search entries in the database by simple keywords, such as bungarotoxin, *Naja* or taipoxin. The search results are displayed in a tabular form as a list containing accession numbers, species and toxin name. The accession number of each match in the result output is hyperlinked to the full data record. *Structure viewer* feature (requires Internet Explorer 6.0 or Netscape 4.75) contains 3D structures of svNTXs that are extracted from the PDB. This simple feature allows users to view available 3D file using Chime viewer or download PDB files of user-specified records. *Annotate svNTX* is a prediction tool that uses sequence similarity, nearest neighbor analysis and rule-based system, which allows user to predict function of an unknown svNTX. Users can also choose to download FASTA formatted files of amino acids from the database under the option *Download FASTA*.

### 3.2 Description of the NEUROTOXIN database records

Each NTX entry in the database is stored as an individual record. Of the 272 records, 260 have complete mature protein (Table 2).

Each NTX entry in the database is given a unique accession number ‘DBACC’ in the form of [D][six digit number], where the six digit number represents the unique descriptor for each record. This is followed by ‘Date’ that identifies the date when the entry was made and ‘Name’ field that contains the name of the toxin used in a published journal. The ‘Accession’ field provides hyperlinks to corresponding entries of the relevant databases, such as GenBank or Swiss-Prot. The organism source of each toxin and its taxonomical classification can be found in the ‘Source’ and ‘Species’ fields, respectively.

In order to further enrich the database, we added functional information based on experimental studies from various literatures. These were represented by six fields known as ‘Action_site’, ‘Physiological_function’, ‘Critical_residues’, ‘Toxin_activity’, ‘Binding_affinity’ and ‘Miscellaneous’. The ‘Binding_affinity’ and ‘Critical_residues’ fields are unique in our database and not available in any public domain.

The ‘Action_site’ field indicates the site of action in which the NTX acts either presynaptically or postsynaptically. A brief description of the NTX function is denoted by the field ‘Physiological_function’. ‘Toxin_activity’ fields refer to the toxicity in terms of LD50 values. ‘Binding_affinity’ field refers to the toxin–receptor interaction in terms of KD or IC50 values. The critical amino acid residues involved in receptor binding and other miscellaneous comments are provided in the ‘Critical Residues’ and ‘Miscellaneous’ fields, respectively. These newly added data have corresponding references added to each record.

Structural features of the toxin, such as residues forming the disulfide bridges are described in the field ‘Features’. Putative structural information derived by similarity to known structures by amino acid alignment is indicated as ‘BY SIMILARITY’. An ‘svNTXdB’ sign followed by ‘BY SIMILARITY’ field implies data added by the authors which were not available in the public databases. ‘Conflict’ field list down the amino acid discrepancies between records of a particular NTX in different databases or published journals. The field ‘Translation’ provides the amino acid sequence of the toxin. ‘Structure’ field contains internal hyperlink to the PDB structure stored in the database, provided that the 3D structure is available.

The ‘Reference’ field contains a list of literature references, with author names and journal titles. Some reference entries have the sign ‘svNTXdB’, which denotes publications that were added by us that is not listed in any public databases.

The svNTX database was largely compiled from the public databases which contained some errors. Common errors that we encountered are summarized in Figure 1. For example, entries Q9PSN6, Q91138 and Q91139 in the public database showed the origin of these toxins from *Naja naja* (Indian cobra) when the correct species should be *Naja sputatrix* (Malayan cobra). Other common errors from these sources were different names for similar toxin and different sequence for similar toxins. Although Swiss-Prot typically provides lethal dose (LD50) values, we found one particular entry—long NTX OH-5 (P80965) did not have this information although its LD50 was available in the original paper. These errors were reduced by checking database entries and cross-checking with the original publications.

### 3.3 Classification of snake neurotoxins

svNTXdB was analyzed to form a classification system for the structure–function prediction. Phylogenetic analysis was used to classify the snake neurotoxins. However, phylogenetic groups alone did not correlate well to functional groups of svNTXs. Therefore, svNTXs were classified by combining pharmacological function and phylogenetic studies. Phylogeny was determined using MP method to test whether a direct structure–function prediction can be deduced.
Members in a clade having a bootstrap value of 90 are grouped together. Based on the available data, of 260 full-length svNTXs that had well-defined pharmacological functions, 53 (20.4%) proteins belonged to the presynaptic NTXs group, 155 (59.6%) proteins belonged to the postsynaptic NTXs, 3 (1.2%) had both pre- and postsynaptic effect and 49 (18.8%) had other functional properties. Those classified as ‘other’ may have pre- or postsynaptic effects, but this is not currently known. The classification trees of all svNTXs were shown in Figure 2A and B and the phylogenetic analysis is depicted in Figure 3A–H. This tree had six layers. Layer 1 denotes the NTX site of action; layers 2 and 3 represent functional and structural properties of NTXs, respectively; layer 4 denotes the molecular target of each svNTX and layer 5 represents the subclassification of NTX based on phylogenetic relationship that formed the basis of the group naming in the final layer.

Presynaptic NTXs exhibit at least eight different functions (Fig. 2A, layer (2), and can be classified into five major groups: (1) NTXs that inhibit acetylcholine release; (2) NTXs that facilitate acetylcholine release either by blocking voltage-gated potassium channel or some unknown mechanism; (3) NTXs that inhibit acetylcholinesterase; (4) NTXs that block calcium channel (5) NTXs that exhibit multiple functions, such as myotoxicity, hemorrhagic/hypotensive or bactericidal activity.

Group 1 NTXs had two distinct groups, those possessing inter disulfide (S–S) chains and those without them. The molecular target of these NTXs is the phospholipid membrane.

---

**Fig. 1.** Summary of number of entries having discrepancies represented by five different categories.

**Fig. 2.** (a and b) Classification of snake NTX based on the structure, function and phylogenetic information. Boxed number refers to the total number of full-length NTXs sharing the same physiological function. Layer 1 denotes the NTX site of action; layers 2 and 3 represent functional and structural properties of NTXs, respectively; layer 4 refers the molecular target of each svNTX; layer 5 represents the subclassification of NTX based on phylogenetic relationship; and layer 6 represents group labels.
The Group 1 of svNTXs formed six phylogenetic subgroups (Fig. 3A and B). Four of these svNTX groups originated from the Elapidae snake family and Groups 1B-1 and 1B-2 were from the Viperidae family.

Several interesting features can be observed from Figure 2A. PLA2 KPA2 (Q9DF52) from Bungarus caeruleus and PLA2 He (Q9PSN5) from Notechis scutatus scutatus in Group 1B-4 exhibit additional function such as anticoagulant or myotoxic activities. However, three other members from this group did not show anticoagulant activity. This could suggest that, based on evidence from phylogenetic analysis, these three svNTXs could possibly have anticoagulant/myotoxic activity that are yet to be determined. One svNTX from Group 1B-2 (F15; Toyama et al., 2003) and two svNTXs from Group 1B-3 (P00608 and P00610) also demonstrated additional functional properties besides neurotoxicity (Fig. 2a; Table 3). It may be possible that other NTX members in these groups possess multiple function(s) as well.

Groups 2A1-3 of presynaptic NTXs that originated from Dendroaspis facilitate acetylcholine release. These were dendrotoxins that block voltage-gated potassium channel
Group 2B was unique—it is the only three-finger NTX reported to date to act presynaptically to facilitate acetylcholine release (Kuhn et al., 2000). Group 3 consisted of presynaptic NTXs that inhibit acetylcholinesterase activity. Acetylcholinesterase-type NTXs had so far been isolated from African mambas (*Dendroaspis*) and were cross-linked by four disulfide bonds.

Group 4 consisted of one member, lethal peptide I (P24335) that contain only one disulfide bond and found to block calcium channels at the presynaptic membrane (Aiken et al., 1992). Group 5 svNTXs had multiple functions *in vivo*, such as myotoxicity, anticoagulant, hemorrhagic/hypotensive or bactericidal effects and has five distinct groups (Fig. 3C).

In contrast, postsynaptic svNTXs fall into two major groups based on their interaction with either nicotinic acetylcholine receptor (nAChR) or L-type calcium channels. The molecular targets of these NTXs are classified under different nAChR subunit or families (Fig. 2h, layer 4). Unknown groups of neuromuscular blockers refer to undetermined receptor subtypes.

The postsynaptic NTXs that interact with nAChR were subdivided into five subgroups such as short, long, weak, kappa and the unknown group of svNTXs (Groups 6–10). Groups 6–9 NTXs formed a similar folding pattern of three loops adopting the three-stranded antiparallel β-pleated sheet (Dufton and Hider, 1983) but Group 10 svNTXs had a non-three-finger fold with one disulfide bond (P24335). We employed a phylogenetic study on the short NTXs and found six subgroups. Surprisingly, all the short NTXs originated from the Elapidae family (Fig. 3D and E). From extensive literature search, 10 short NTXs had been reported to interact with muscle-type nAChR (Marchot et al., 1988; Weber and Changeux, 1974; Pillet et al., 1993; Gong et al., 1999, 2000).

One interesting svNTX was fasciatoxin from Group 7B-1 (Fig. 3G) (P14534). It is an unusual short NTX from *Bungarus fasciatus* (Liu et al., 1989). This protein is 63 amino acids long but had an unusual C-terminal tail of PSTPST not found in any other short NTXs. So far, short NTXs were found to interact specifically with muscle-type nAChRs (Fig. 2B, layer 4). Critical residues that were involved in the molecular recognition of...
Fig. 3. Continued.
Fig. 3. Continued.
Systematic analysis of snake neurotoxins' functional classification

Fig. 3. Continued.

Toxin–receptor were studied elsewhere by individual chemical modifications (review by Endo and Tamiya, 1991). Lys-27 and Lys-47 were critical for receptor binding (Hori and Tamiya, 1976; Ishikawa et al., 1977; Faure et al., 1983). Pillet et al. (1993) suggested that three ionogenic residues Lys-27, Asp-31 and Arg-33 in erabutoxin a from L. semifasciata were involved in direct interaction with AChR.

Long svNTXs from Elapidae family were subclassified into five subgroups (Fig. 2b, layer 5). Several members of long NTXs interact with alpha-7 neuronal nAChR (Group 6B) in contrast to short NTXs. Phylogenetic analysis could not be performed on this group because it had only three members at present. To date, five and three long NTXs were demonstrated experimentally to interact with muscle-type nAChR (Group 6A-1 and 6A-2; Fig. 3D) and alpha-7 neuronal nAChR (Group 6B-1), respectively (Ishikawa et al., 1977; Chang et al., 1993; Servent et al., 1997). The presence of the fifth disulfide bond in long NTX had been implicated in the binding specificity to neuronal nAChR (Servent et al., 1997). Arg-33 was especially important for this interaction (Antil et al., 1999). It would be of great interest if the important residues that interact with these two receptors can be delineated. Forty-five long NTXs were not experimentally determined whether they interact with either muscle-type or neuronal nAChR. Phylogenetic analysis indicated two main groups within this group: Groups 6C-1 and 6C-2 (Fig. 3E).

svNTXs in Groups 7A-1 and 7A-2 (Fig. 3F) did not have the consensus Lys-27, Asp-31 and Arg-33 residues that were implicated in receptor binding (Pillet et al., 1993). However, svNTXs in Groups 7A-3 and 7A-4 (except P01426) contain conserved residues of Lys-27, Asp-31 and Arg-33 for receptor interaction.

Known kappa svNTXs have 65–66 amino acid residues. Their disulfide bonding patterns are yet to be determined. They showed similarity to long-chain NTXs that contain five disulfide bonds (Fiordalisi et al., 1994). Unlike monomeric long SVNTXs, kappa svNTXs exist as dimers (Dewan et al., 1994). One distinguishing functional property of kappa svNTXs was their ability to interact with α3 neuronal nAChR (Fig. 2b, Group 8A). So far, kappa svNTXs were found only in Bungarus species and it was believed that they do not interact with muscle nAChR. The absence of Trp-32 in all kappa NTXs could be the reason for the lack of binding to muscle nAChR (Grant et al., 1988). In addition, the specific target of six kappa NTXs was unknown; and the phylogenetic analysis showed two distinct groupings (Fig. 3H).
Phylogenetic analysis on weak NTX was not performed because of the small number of NTXs (Groups 9A–C). This group of toxins was found to interact weakly with muscle- and α7-neuronal nAChR (Poh et al., 2002; Utkin et al., 2001). To date, our database contains 86, 50, 8 and 6 full-length short, long, kappa and weak NTXs, respectively.

Certain postsynaptic NTXs (Fig. 2b, Group 11) interact with L-type calcium channel (Strydom, 1977; Yasuda et al., 1994).
Fig. 3. Continued.

Crotoxin (P24027) and PLA₂ HTe (Q9PSN5) in Group 12 exerted their effect postsynaptically (Bon et al., 1979; Francis et al., 1995). However, their specific postsynaptic actions are yet to be determined.

Three NTXs in Groups 13 and 14 exerted neurotoxic effects both pre- and postsynaptically (Francis et al., 1995; Bon et al., 1979). Specific functions of Group 15 NTXs are still unknown. They are designated as NTX homologs, NTX-like proteins or NTX isoforms that are not yet well-studied. The individual records of NTX in each group in Figure 3 are listed in Table 3 according to their GenBank accession number.

3.4 Testing of ‘Annotate svNTX’ tool by experimental research

We tested Bcntx4 protein, a novel B.candidus neurotoxin using competitive binding studies with ¹²⁵I-α-Bgt in muscle (Torpedo) nAChR. The binding activity studies of Bcntx4 on AChR isolated from T.californica showed very low-binding affinity even in the micromolar range (Fig. 4). Bcntx4 was also found to bind weakly to neuronal nicotinic receptor (data not shown). Using the ‘Annotate svNTX’ function in the database, Bcntx4 was predicted to be a weak neurotoxin. This further validates the proof-of-concept of the in silico prediction tool.

4 CONCLUSION

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. In summary, this paper lists several contributions in the field of svNTXs. The svNTX dB was created with several special features that are uncommon in sequence databases: (i) 3D structure feature viewer, (ii) functional annotation of individual entries and (iii) functional properties prediction tool. This database aims to collate sequence, functional and structural information of svNTXs in a single repository; enable fast access to and retrieval of data; clean it of errors and redundancies; enable rapid data analysis and development of a prediction module. This classification system of svNTXs, based on the structure, function and phylogenetic analysis, represents a snapshot of current svNTX data. It can be used to predict function, structure and molecular targets of newly determined sequences. The field of toxinology is growing rapidly.
### Table 3. Summary of the individual entries based on GenBank Accession Number of each member in each group as depicted in Figure 2

<table>
<thead>
<tr>
<th>Group</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A-1</td>
<td>P00619, Q8AXW2, Q90251, P59018, Q9PTA6, Q9PTA5, Q9PU97, Q9PTA7, Q8QFW3, P00617, Q9PTA1, P17934, P00618, Q8QFW4</td>
</tr>
<tr>
<td>1A-2</td>
<td>PSKF3U, PSKF4U</td>
</tr>
<tr>
<td>1B-1</td>
<td>P14424, P11407, P00626</td>
</tr>
<tr>
<td>1B-2</td>
<td>Q8JFG0, Q10755, P14420, AAB36096, Q02471, P14421, 151381, P24027, F15</td>
</tr>
<tr>
<td>1B-3</td>
<td>P10116, Q8UHU8, Q8UHUKQ, 08U10, Q8UHU7, P23026, P00610, P00609, P00608, P20529, P20528</td>
</tr>
<tr>
<td>1B-4</td>
<td>Q9DEF52</td>
</tr>
<tr>
<td>2A-1</td>
<td>P00980, P00982, P00981</td>
</tr>
<tr>
<td>2A-2</td>
<td>P00979, P00980</td>
</tr>
<tr>
<td>2A-3</td>
<td>P00980</td>
</tr>
<tr>
<td>2B</td>
<td>P1782</td>
</tr>
<tr>
<td>3</td>
<td>P01403, P25681</td>
</tr>
<tr>
<td>4</td>
<td>P24335</td>
</tr>
<tr>
<td>5A-1</td>
<td>F15</td>
</tr>
<tr>
<td>5A-2</td>
<td>Q9DEF52</td>
</tr>
<tr>
<td>5A-3</td>
<td>Q9PSN5, P00610, P00608</td>
</tr>
<tr>
<td>6A-1</td>
<td>P01379</td>
</tr>
<tr>
<td>6A-2</td>
<td>P82662, P10382, P80156, P80965</td>
</tr>
<tr>
<td>6B</td>
<td>P01391, P01379, ACC83981</td>
</tr>
<tr>
<td>6C-1</td>
<td>P01363</td>
</tr>
<tr>
<td>6C-2</td>
<td>Q9W7J5, P13495, Q42257, P25671, P25672, P25673, 764177A, P25668, P01388, P01389, P01390, P25674</td>
</tr>
<tr>
<td>6C-3</td>
<td>P01366, P07526, P01387, P01393, P01396, P25667, P01397, P01395, P01394, P25670, BAA32922, 0901189A, 0901189B, Q8UW29, Q8UW28, P01381, P01380, P34473, P01384, P01385, ACC83990, ACC83998, ACC83992, ACC83985, ACC83987, CABS1841, ACC83995, ACC83997, ACC83996, ACC83986</td>
</tr>
<tr>
<td>7A-1</td>
<td>Q9W7J7, Q9W7J6, Q9W7K0, Q9W7K1</td>
</tr>
<tr>
<td>7A-2</td>
<td>Q9W7K2, QAD40971</td>
</tr>
<tr>
<td>7A-3</td>
<td>P01431</td>
</tr>
<tr>
<td>7A-4</td>
<td>P01426, P01435, N1LT1E</td>
</tr>
<tr>
<td>7B-1</td>
<td>P15434</td>
</tr>
<tr>
<td>7B-2</td>
<td>P10808, W-IV, W-III, P34076, Q9YHUJ, P01428, N1NJ2P, Q9YHV0, P59276, P01427, P59275, P80956, P80548, P01416, P01417, P01418, P25495, Q9YGC4, Q9YGC2, Q9YGW9, Q9YGW8, P25496, Q9PRJ3, Q9YGX1, P01460, P01458, Q9PRJ6, Q9YGC7, Q9YGX0, Q9PRJ7, Q9PRJ0, Q9PRJ5, Q9PW4, S14003, 090VW1, P01434, P25497, P19959, P25679, P19958, P19960, P01437, P25492, Q8UW26, Q8UW27, P25494, P25493, P01438, S24844, G71059A, P34075, P01429, P01424, P25675, P01422, P01423, P01420, P01421, P01425, P01433, P01432, P26849, O57327, Q91139, Q9YGJ5, Q57326, Q9YGJ6, Q91138, P14613, P01430, Q9PSN6, Q9DE57, CAA43097, P33309, Q9W7J9</td>
</tr>
<tr>
<td>8A</td>
<td>P01398, P15815</td>
</tr>
<tr>
<td>8B</td>
<td>P01398</td>
</tr>
<tr>
<td>8C-1</td>
<td>Q9W729, CAA72434, P15817</td>
</tr>
<tr>
<td>8C-2</td>
<td>P12961, P15816, P12962</td>
</tr>
<tr>
<td>9A</td>
<td>P82935, O142255, P01783</td>
</tr>
<tr>
<td>9B</td>
<td>P82935, O142255, P01783</td>
</tr>
<tr>
<td>9C</td>
<td>Q42256, Q9W713, Q9W714</td>
</tr>
<tr>
<td>10</td>
<td>P24335</td>
</tr>
<tr>
<td>11</td>
<td>P01414, AAB50805</td>
</tr>
<tr>
<td>12</td>
<td>Q9PSN5, P24027</td>
</tr>
<tr>
<td>13</td>
<td>Q9PSN5, P24027</td>
</tr>
<tr>
<td>14</td>
<td>P24335</td>
</tr>
<tr>
<td>15</td>
<td>Q9YG10, Q9PU7B, Q9YGH9, Q9YGJ0, Q9YGJ8, Q9YGJ6, Q9W796, 012963, Q9PRJ1, P58370, P15816, P43445, P81030, P82462, P13822, P82463, P82464, P80494, P80495, Q9W707, Q9YGJ12, Q9YGJ14, Q9YGJ11, P29179, P29180, P29181, P29182, Q9DE03, Q9W717, Q9W726, Q9W797, Q9YGJ17, CAA6886, AAM00185, P25676, P0145, AAB36087, P01399, CAD18648, S48648, P01401, S66418, P25680, P25679, P01412, P63346, P25683, P14612, P01400</td>
</tr>
</tbody>
</table>

Field in boldface and underlined indicate the presence of entry in more than one group. Fields in italic (F15, W-III and W-IV) represents entries found in the published journal but not in any public database.
especially in terms of information provided by cDNA and protein data in the public database. The availability of a specialized database allows easier systematic data analysis and also enables incremental data analysis upon updating with new data. This database and analysis system will serve as an example for bioinformatics application in the studies of snake NTXs.

ACKNOWLEDGEMENTS

The authors would like to thank the National Medical Research Council (Singapore) for the research grant, R183-000-068-213. We also appreciate all the assistance rendered to us by the members of Institute of Infocomm Research, Singapore.

REFERENCES


