Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

IN SILICO ANALYSIS AND HOMOLOGY MODELING OF ALPHA-CONOTOXIN FROM SELECTED VERMIVOROUS CONUS SPECIES

NOEL A. SAGUIL

Ph.D. is currently the Chief of the Center for Engineering and Technology Research (CETR), Under the Research Institute for Science and Technology (RIST) of the Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines. E-mail: noels70@yahoo.com

ALVIN N. CARIL

Currently a Faculty Researcher Under the Research Institute for Science and Technology (RIST) of the Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines. E-mail: carilus.alvinii@gmail.com

LLOYD VINCENT P. PAEZ

Graduate of Bachelor of Science in Biology, Under Department of Biology, College of Science of the Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines.

MA. ADRIENE JOY P. GONZALES

Graduate of Bachelor of Science in Biology, Under Department of Biology, College of Science of the Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines.

SOFIA HAZEL L. DE GUZMAN

Currently Pursuing a Master's Degree Program in Biology at the Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines.

PRINCESS CASEY BANTIGUE

Currently a Research Assistant Under the Research Institute for Science and Technology (RIST) and Currently Pursuing a Master's Degree Program in Biology at the Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines.

Abstract

Cone snails are members of the vast Conoidean group of venomous marine gastropods found in a variety of tropical and subtropical ocean ecosystems. They use their venom, like other venomous species, as a defensive mechanism as well as a tool for capturing their prey. As part of their neurochemical and biochemical strategy for catching prey, these species produce peptides that target specific ion channels and receptors, and they appear to have produced conotoxin, an unusual combination of pharmacologically active peptides. These conotoxins contain more than 100 components, resulting in a variety of active mono peptides used pharmacologically. However, studies on the structure of these peptides are still limited, leaving us with a gap in understanding their action, selectivity, and potency. Therefore, in silico and homology modeling studies on one (1) family of alpha-conotoxin from selected vermivorous *Conus* species were discussed in this study. The majority of the alpha-conotoxins were highly basic in nature, according to physicochemical characterization, with one alpha-conotoxin sequence (α-conotoxin *Lt28.1*) indicating acidic in nature. The index of instability ranges from stable α-conotoxin *RgIA* (14.62) to unstable α-conotoxin *ImII* (126.91). Additionally, the aliphatic index for the protein sequences studied ranges from 29.15 to 97.50, indicating that it is thermostable over a wide temperature range. Furthermore, GRAVY values for alpha-conotoxins range from -0.910 (most hydrophilic; α-conotoxin *SrIA/SrIB*) to 0.731 (most hydrophobic; alpha-conotoxin *Lp1.1*). Moreover, *toxin_8* was discovered to be the common domain among the species. A random coil dominates other secondary structures in terms of percentage score of amino acid distribution for secondary structure prediction, followed by α-helix, β-strand, and β-turns. Finally, the predicted models for α-conotoxin *Lp1.4*, α-conotoxin *ArIA*, α-conotoxin *VxXXC,* α-conotoxin *VxXXA*, and α-conotoxin *LvIA* are of high quality, with ≥ 90% coverage in their favored regions.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

Index Terms: *Conus*, Alpha-Conotoxin, in Silico Analysis, Homology Modeling, Ramachandran Plot, Cone Snails, Marine Gastropods

1. INTRODUCTION

Cone snails are members of the vast Conoidean group of venomous marine gastropods, including *Conidae, Turridae*, and *Terebridae*. They can be found in a wide range of tropical and subtropical ocean ecosystems, from shallow coral reefs to deep-sea environments [1]. Similar to other venomous species, cone snails utilize their venom as a defensive mechanism and a tool for capturing their prey [2], [3], [4], [5]. These species produce peptides that target specific ion channels and receptors as part of their neurochemical and biochemical strategy for catching prey [6].

The genus *Conus* has approximately 900 species of venomous marine cone snails, and as a result of the widespread occupation of their biological niche, they appear to have generated an unusual combination of pharmacologically active peptides known as conotoxin [7], [8], [9], [10]. These conotoxins contain more than 100 components, leading to various active monopeptides used pharmacologically wherein each *Conus* species uses a distinct combination of peptides to paralyze its prey rapidly. The disulfide bond frameworks responsible for its potency contain *alpha*-helices, *beta*-sheets, and *beta*-turns with exquisite selective of receptor subtype and resistance to proteases, making it a strong candidate for neurological tools and drug development [11].

Moreover, conotoxins, according to [12] are classified based on their precursor's signal sequence, defining the gene superfamily (A-Q, S, T, V, Y); cysteine structures, describing the cysteine pattern (I-XXVI); and pharmacology, determining its pharmacological target (α, γ, ε, δ, κ, μ, ρ, σ, τ, χ, ω). Among these, alpha-conotoxin is given more focus due to its ability to block receptors such as nicotinic acetylcholine receptors (nACHR), the well-known target of other animal toxins such as neurotoxins from snake venom. Alpha-conotoxin is the most prominent family of nACHR antagonists; it has rich peptides isolated from venomous marine snails composed of 12-19 amino acids [13], [14]. However, studies on the structure of these peptides are still limited, providing us with a gap in understanding their action, selectivity, and potency.

The goal of the drug discovery process is to find novel medicinal compounds that can bind to a recognized disease-causing target. While the most critical task in the traditional drug development approach is identifying a good therapeutic target [15]. According to [16], one of the major drawbacks of the experimental method used to characterize protein was its high cost and time consumption. As a result, these methods were incompatible with high throughput techniques. These issues, however, can be solved using computational (*in silico*) approaches, which are frequently used in the drug discovery process nowadays. Furthermore, researchers can use computational tools to learn about the physicochemical and structural properties of proteins, because the amino acid sequence contains the vast majority of the information required to determine and characterize the molecule's function, physical, and chemical properties [16]. In addition, the application of homology modeling has evolved into an improved method in structural

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

biology, significantly bridging the gap between known protein sequences and empirically determined structures [17].

Therefore, in this study, the *in silico* and homology modeling studies on one (1) family of alpha-conotoxin from selected vermivorous *Conus* species were reported. The research compares, analyzes, and structures this family of alpha-conotoxin compounds using several *in silico* approaches such as physical and chemical property prediction, domain scanning, protein secondary structure prediction, and peptide structural analysis. Since the three-dimensional (3D) structure of this protein family was not yet available, model structures for this alpha-conotoxin family were constructed to describe their structural features and to understand their molecular function. Through this research, a thorough understanding of the target molecule's composition and biological properties will serve as the foundation for developing and comprehending its design and effect. Furthermore, the evolution of its structure will benefit its overall structure as well as clinical development via computer-aided techniques.

2. METHODOLOGY

2.1 Protein Sequence Retrieval of the Target Protein

The protein database from the National Library of Medicine, National Center for Biotechnology Information (NCBI), which contains sequences from various sources, including annotated coding sections from GenBank, RefSeq, and TPA with the addition of records from SwissProt, PIR, PRF, and PDB, was used to analyze the structure and functional phenomena of the protein networks of alpha-conotoxin. The protein sequence of *C. imperialis, C. lividus, C. pulicaria, C. vexillum, C. planorbis, C. leopardus, C. regius, C. litteratus, C. arenatus,* and *C. spurius* were retrieved from the Universal Protein Resource (UniProt; http://www.uniprot.org) protein database and was saved in fast all (FASTA) file format for further analysis. Most of the sequences were from the translation of sequences, which is a great resource for understanding protein’s post-translational modifications [18].

2.2 Physicochemical Parameters Analysis

Determination of physicochemical parameters used Expasys’s ProtParam Tool (https://web.expasy.org/protparam/) to determine the nature of interactions between systems components. These parameters were used to examine the properties of chemicals identified as significant structural [19]. Briefly, Expasys’s ProtParam Tool was used to determine several physical and chemical parameters such as molecular weight, isoelectric pI (isoelectric point), instability index (II), aliphatic index (AI), and grand average of hydropathy (GRAVY; [20]). To calculate such parameters, the protein sequence was supplied using an accession number from UniProtKB/Swiss-Prot or UniProtKB/TrEMBL or as an amino acid sequence.

2.3 Domain Analysis

The Protein families (Pfam) database was used to analyze the molecular domain of alpha-conotoxin from selected vermivorous *Conus* species utilizing the protein sequence

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

obtained from the protein database. Pfam 34.0 is a comprehensive collection of multiple sequence alignments and hidden Markov modeling encompassing many common protein domains and families.

2.4 Secondary Structure Prediction

Self-optimized Prediction with Method and Alignment (SOPMA; https://www.npsa-ibcp.fr) was used to predict the secondary structure (*alpha*-helix, *beta*-sheet, and coil) of the alpha-conotoxins. SOPMA correctly predicts 69.5% of amino acids in a database [21] in which the researchers used this tool to visualize the results and the score curves for all predicted states, allowing the researcher to interpret the data accurately. The protein sequences saved in FASTA format from the protein database were submitted to the SOPMA server to yield secondary structure components. It shows the parameters, such as window width, number of states, etc., that are used for the prediction and provides a link to the prediction result file, which gives the result in a text format.

2.5 Homology Modeling of Alpha-Conotoxin and Structure Verification

Homology modeling was utilized in this study to bridge the gap between primary and three-dimensional (3D) structures, allowing the derivation of functional and beneficial properties in the same way as an experimental 3D structure does [22]. This approach allows access to more therapeutic targets as well as a variety of other applications, such as studying protein function, structural roles of proteins in the cell, and protein interactions [23], [24], [25], [26]. Briefly, the alpha-conotoxin's three-dimensional (3D) structure was constructed using the SWISS-Model Server, a bioinformatics tool widely used for protein structure prediction and changing the modeling parameters to produce the needed data according to the research of [27]. Wherein the protein ID would be entered into the SWISS-Model webserver to create a model with enough query sequence coverage and sequence identity. The most dependable 3D structure was chosen based on the Global Model Quality Estimation (GMQE) [28] and Qualitative Model Energy Analysis (QMEAN) values [29]. The GMQE values are normally between 0 and 1, and the greater the number, the more reliable the projected structure is, whereas a value below 4.0 indicates reliability for QMEAN [29]. The modeled alpha-conotoxin’s .pdb file format was uploaded to the European Bioinformatics Institute's PDBsum online server [30]. Lastly, for structure validation, acquiring the Ramachandran plot was used to analyze the quality of the modeled protein.

3. RESULTS AND DISCUSSIONS

3.1 Retrieved Target Protein Sequences

The protein sequences of alpha-conotoxins from vermivorous cone snails were retrieved and saved in FASTA format from the UniprotKB database and used for further analysis. Table 1 summarizes the retrieved protein sequences with their accession number, species name, length, and protein name. As shown in the results, the sequence from *C. imperialis*’ alpha-conotoxin *ImII* and *ImI* was the shortest, with only seventeen (17) amino acids, whereas α-conotoxins *VxXXB* from *C. vexillum*, was the longest, with 95 amino acids.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

Table 1: Retrieved Protein Sequences Using Uniprotkb

SPECIES	ACCESSION NUMBER	LENGTH	PROTEIN NAME
C. leopardus	A1X8B6	68	Alpha-conotoxin *Lp 1.4*
	Q6PTD5	68	Alpha-conotoxin *Lp1.1*
C. arenatus	P0C8R2	39	Alpha-conotoxin *ArIA*
C. spurius	P85886	69	Alpha-conotoxin *SrIA/SrIB*
C. vexilium	J7JU64	68	Alpha-conotoxin *VxXXIVA*
	P0C1W7	47	Alpha-conotoxin *VxXXC*
	P0C1W6	95	Alpha-conotoxin *VxXXB*
	P0C1W5	92	Alpha-conotoxin *VxXXA*
C. pulicarius	C6ZJQ2	69	Alpha-conotoxin *Pu14.1*
C. imperialis	Q8I6R5	17	Alpha-conotoxin *ImII*
	P50983	17	Alpha-conotoxin *ImI*
C. lividus	L8BU87	37	Alpha-conotoxin *LvIA*
	H9N3R7	31	Alpha-conotoxin *Li1.12*
C. planorbis	Q0N4U8	76	Alpha/Kappa conotoxin *Pl14a*
C. litteratus	A0A068B0Z6	41	Alpha-conotoxin *Lt1.3b*
	Q2I2R5	64	Alpha-conotoxin *Lt14.1*
	F6JWU7	85	Alpha-conotoxin *Lt28.1*
C. regius	P0C1D0	32	Alpha-conotoxin *RgIA*
	P85013	66	Alpha-conotoxin *RgIIA*

3.2 Physicochemical properties of the alpha conotoxins

The physicochemical properties of the retrieved alpha-conotoxins from the ten (10) selected vermivorous *Conus* species were determined and studied using Expasy's Protparam Tool and results are summarized in Table 2. The parameters include accession number, molecular weight, isoelectric pI, instability index (II), aliphatic index (AI), and grand average of hydropathicity (GRAVY).

The computed pI for the majority of the alpha-conotoxins was larger than seven (pI > 7), with one alpha-conotoxin sequence (α-conotoxin *Lt28.1*) having a pI less than seven (pI < 7), as shown in Table 2. The isoelectric pI and charge are important features for solubility, subcellular localization, and interaction of the proteins [31]. Results show that most alpha-conotoxins were found to be extremely basic in nature, with only one (1) alpha-conotoxin sequence (α-conotoxin *Lt28.1*) identified with less than seven (7) pI value, indicating that the protein is considered acidic [31].

The alpha conotoxins' instability index (II) reveals a wide range of proteins with stable and unstable values. The results show that nine (9) alpha-conotoxins out of the nineteen (19) identified proteins have a stable instability index (II) value, namely *α*-conotoxins *RgIA* (14.62), *VxXXA* (24.02), *Lp1.1* (26.33), *ImI* (28.68), *RgIIA* (33.24), *Pu14.1* (33.36), *Lp1.4* (36.97), *LvIA* (37.65), and *Lt14.1* (38.32). While ten (10) alpha-conotoxins have an unstable instability index (II) value, namely α-conotoxins *Lt1.3b* (41.86), *SrIA/SrIB* (46.65), *VxXXB* (48.67), *VxXXIVA* (51.71), *VxXXC* (53.59), *Li1.12* (66.50), *ArIA* (68.28), *Lt28.1* (73.31), and *ImII* (126.91), including alpha/kappa-conotoxin *Pl14a* (62.84). According to [32], the dipeptide composition-based Instability Index (II) is the protein's major structure-dependent approach for *in vivo* protein stability predictions. A protein with an instability index of less than 40 is considered stable; a value greater than 40 indicates that the protein may be unstable [33].

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

Moreover, the aliphatic index value of the nineteen (19) alpha-conotoxins was determined in order to assess the thermostability of each alpha-conotoxin. The aliphatic index for alpha-conotoxin *VvXXC* has the lowest value, 29.15, while alpha-conotoxin *VxXXIVA* has the highest, 97.50, as shown in Table 2. The aliphatic index is the relative volume filled by aliphatic side chains [alanine (Ala), valine (Val), isoleucine (Ile), and leucine (Leu)], and it could be viewed as a positive factor in globular proteins' increased thermostability [34]. This means that the higher the aliphatic index, the more thermally stable the protein.

Finally, the protein's grand average hydropathy, or GRAVY, is determined by dividing the total amount of hydropathy values of all amino acids by the number of residues in the sequence [34]. The GRAVY value of a protein determines whether it is hydrophobic (+) or hydrophilic (-), with most proteins ranging from -2 to +2. As shown in the results (Table 2), eleven (11) of the nineteen (19) identified proteins have positive values, suggesting that these alpha conotoxins are hydrophobic, whereas eight (8) of the nineteen (19) proteins have negative values, suggesting that these alpha conotoxins are hydrophilic.

Table 2: Determined Physicochemical Characteristics of the Nineteen (19) Alpha-Conotoxins from Ten (10) Selected Vermivorous *Conus* Species Using the Expasys' Protparam Tool

SPECIES	PROTEIN NAME	MOLECULAR WEIGHT	ISOELECTRIC PI	INSTABILITY INDEX	ALIPHATIC INDEX	GRAVY
C. leopardus	Alpha-conotoxin *Lp 1.4*	7291.66	9.30	36.97 (stable)	81.91	0.340
	Alpha-conotoxin *Lp1.1*	7154.61	9.37	26.33 (stable)	92.06	0.731
C. arenatus	Alpha-conotoxin *ArIA*	4437.04	10.59	68.28 (unstable)	52.56	-0.910
C. spurius	Alpha-conotoxin *SrIA/SrIB*	7587.92	8.93	46.65 (unstable)	73.09	0.201
C. vexilium	Alpha-conotoxin *VxXXIVA*	6697.00	9.96	51.71 (unstable)	97.50	0.078
	Alpha-conotoxin *VxXXC*	5286.11	8.81	53.59 (unstable)	29.15	-0.389
	Alpha-conotoxin *VxXXB*	10545.20	8.79	48.67 (unstable)	69.79	-0.277
	Alpha-conotoxin *VxXXA*	10112.86	8.59	24.02 (stable)	76.20	0.035
C. pulicarius	Alpha-conotoxin *Pu14.1*	7509.98	9.20	33.36 (stable)	83.33	0.281
C. imperialis	Alpha-conotoxin *ImII*	2096.50	10.41	126.91 (unstable)	45.88	-0.712
	Alpha-conotoxin *ImI*	1938.29	9.01	28.68 (stable)	45.88	-0.300
C. lividus	Alpha-conotoxin *LvIA*	3841.33	7.91	37.65 (stable)	60.81	-0.297
	Alpha-conotoxin *Li1.12*	3081.73	8.56	66.50 (unstable)	75.81	0.713
C. planorbis	Alpha/Kappa conotoxin *Pl14a*	8297.84	10.02	62.84 (unstable)	66.71	0.088
C. litteratus	Alpha-conotoxin *Lt1.3b*	4243.73	8.52	41.86 (unstable)	47.80	-0.407
	Alpha-conotoxin *Lt14.1*	6878.28	8.38	38.32 (stable)	79.38	0.341
	Alpha-conotoxin *Lt28.1*	9391.21	4.40	71.31 (unstable)	96.35	0.311
C. regius	Alpha-conotoxin *RgIA*	3725.38	9.93	14.62 (stable)	49.06	-0.831
	Alpha-conotoxin *RgIIA*	7038.18	9.00	33.24 (stable)	87.12	0.305

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

3.3 Domains and secondary structures of the alpha conotoxins

Table 3 summarizes the molecular domains of the nineteen (19) alpha conotoxins identified from the ten (10) vermivorous *Conus* species. According to the results, nine (9) of the nineteen alpha conotoxins contain a domain belonging to the protein family (Pfam) 07365 and known as *toxin_8*, whereas ten (10) have no domains in their sequence. The identified specific domain, *toxin_8*, is the sole member of the superfamily cl06417 (Pfam Database) [35]. This alpha-conotoxin domain-containing family also includes multiple alpha-conotoxin precursor proteins from various *Conus* species. Moreover, alpha-conotoxins are peptide neurotoxins found in the venom of various cone snail species that block nicotinic acetylcholine receptors (nAChRs). The discovery of domains in putative and conserved proteins raises the likelihood of proteins sharing the same domain function [36]. Conserved domain footprint identification on protein sequences is frequently the initial step in *in-silico* protein function analysis. Protein domains can be arranged into an evolutionary classification and considered units in the molecular evolution of proteins. The set of protein domains identified thus far appears to describe only a few thousand super families, with members of each superfamily descended from the same ancestor [37].

Table 3: Domain Analysis of the Nineteen (19) Alpha-Conotoxins from Ten (10) Selected Vermivorous *Conus* Species Using the Pfam Database

SPECIES	ACCESSION NUMBER	PROTEIN NAME	DOMAIN
C. leopardus	A1X8B6	Alpha-conotoxin *Lp 1.4*	*toxin_8*
	Q6PTD5	Alpha-conotoxin *Lp1.1*	*toxin_8*
C. arenatus	P0C8R2	Alpha-conotoxin *ArIA*	*toxin_8*
C. spurius	P85886	Alpha-conotoxin *SrIA/SrIB*	*toxin_8*
C. vexilium	J7JU64	Alpha-conotoxin *VxXXIVA*	N
	P0C1W7	Alpha-conotoxin *VxXXC*	N
	P0C1W6	Alpha-conotoxin *VxXXB*	N
	P0C1W5	Alpha-conotoxin *VxXXA*	N
C. pulicarius	C6ZJQ2	Alpha-conotoxin *Pu14.1*	*toxin_8*
C. imperialis	Q8I6R5	Alpha-conotoxin *ImII*	N
	P50983	Alpha-conotoxin *ImI*	N
C. lividus	L8BU87	Alpha-conotoxin *LvIA*	*toxin_8*
	H9N3R7	Alpha-conotoxin *Li1.12*	*toxin_8*
C. planorbis	Q0N4U8	Alpha/Kappa conotoxin *Pl14a*	N
C. litteratus	A0A068B0Z6	Alpha-conotoxin *Lt1.3b*	*toxin_8*
	Q2I2R5	Alpha-conotoxin *Lt14.1*	N
	F6JWU7	Alpha-conotoxin *Lt28.1*	N
C. regius	P0C1D0	Alpha-conotoxin *RgIA*	N
	P85013	Alpha-conotoxin *RgIIA*	*toxin_8*

The secondary structure elements alpha (α) helix, extended strand, and random coils of nineteen (19) alpha-conotoxins from the ten (10) selected vermivorous cone snails are shown in Table 4. *Conus pulicarius* α-conotoxin *Pu14.1*, *Conus leopardus* α-conotoxin *Lp1.1*, and *Conus spurius* α-conotoxin *SrIA* had the most α-helix present, with 38 (55.07%), 37 (54.41%), and 36 (52.94%) respectively. However, *Conus imperialis* α-conotoxin *ImI* lacks a α-helix in its sequence. *Conus imperialis* α-conotoxin *ImI* and α-conotoxin *VxXXA* of *Conus vexillum* have the largest number of β-strands, 6 (35.29%)

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

and 22 (23.91%), respectively, and the lowest number is α-conotoxin *RgIA* of *Conus regius*, which has no β-strand in the sequence. In contrast, α-conotoxin *Lt1.3b* of *Conus litteratus* and α-conotoxin *LvIA* of *Conus lividus* have the most random coils in their sequences, with 27 (65.85%) and 24 (64.86%), respectively, whereas α-conotoxin *Lp1.1* of *Conus leopardus* has the fewest, with just 12 (17.65%). Furthermore, in terms of the number of β-turns in each sequence, α-conotoxin *ImI* and α-conotoxin *ImII* of *Conus imperialis* have the same number of β-turns (4 (23.53%), however, both α-conotoxins of *Conus leopardus* do not have any indicated β-turns present in their sequences. The α-helical structure is made up of the amino acids methionine (M), alanine (A), leucine (L), glutamate (E), and lysine (K), while the β-helical structure is made up of tryptophan (W), tyrosine (Y), phenylalanine (F), valine (V), isoleucine (I), and threonine (T). Additionally, the amino acids glycine (G) and proline (P) help to make the required turns (Lee *et al*., 2004). These results suggest that the protein's secondary structure is exclusively influenced by the number of amino acids [38]. The amino acid distribution percentage score indicates that random coils dominate other secondary structures, followed by α-helix, β-strand, and β-turns. This result is in line with the results of [39], who discovered that the most common secondary structural elements are α-helices, β-sheets, and random coils. However, according to [40], alpha-helices are the most common secondary structure, accounting for 30% of the structure of a typical globular protein, which contradicts the stated results because random coils dominate most alpha-conotoxins. The properties of amino acids account for random coils' dominance over other elements. The molecular properties of the polypeptide chain, according to [41], determine the characteristics of a protein's random coil state.

Table 4: Predicted Secondary Structure of the Nineteen (19) Alpha-Conotoxins from Ten (10) Selected Vermivorous *Conus* Species Using Sopma Tool

SPECIES	PROTEIN NAME	ALPHA-HELIX	BETA-STRAND	RANDOM COIL	BETA TURN
C. leopardus	Alpha-conotoxin *Lp 1.4*	30 (44.12%)	13 (19.12%)	25 (36.76%)	
	Alpha-conotoxin *Lp1.1*	37 (54.41 %)	13 (19.12%)	12 (17.65%)	
C. arenatus	Alpha-conotoxin *ArIA*	13 (33.33%)	2 (5.13%)	20 (51.28%)	4 (10.26%)
C. spurius	Alpha-conotoxin *SrIA/SrIB*	36 (52.94%)	8 (11.76%)	22 (32.35%)	2 (2.94%)
C. vexilium	Alpha-conotoxin *VxXXIVA*	21 (35.00%)	8 (13.33%)	28 (46.67%)	3 (5.00%)
	Alpha-conotoxin *VxXXC*	5 (10.64%)	8 (17.02%)	25 (53.19%)	9 (19.15%)
	Alpha-conotoxin *VxXXB*	15 (15.79%)	15 (15.79%)	50 (52.63%)	15 (15.79%)
	Alpha-conotoxin *VxXXA*	15 (16.30%)	22 (23.91%)	40 (43.48%)	15 (16.30%)
C. pulicarius	Alpha-conotoxin *Pu14.1*	38 (55.07%)	8 (11.59%)	22 (31.88%)	1 (1.45%)
C. imperialis	Alpha-conotoxin *ImII*	4 (23.53%)	4 (23.53%)	5 (29.41%)	4 (23.53%)
	Alpha-conotoxin *ImI*	0	6 (35.29%)	7 (41.18%)	4 (23.53%)
C. lividus	Alpha-conotoxin *LvIA*	8 (21.62%)	3 (8.11%)	24 (64.86%)	2 (5.41%)
	Alpha-conotoxin *Li1.12*	4 (12.90%)	5 (16.13%)	18 (58.06%)	4 (12.90%)
C. planorbis	Alpha/Kappa conotoxin *Pl14a*	34 (44.74%)	13 (17.11%)	26 (34.21%)	3 (3.95%)
C. litteratus	Alpha-conotoxin *Lt1.3b*	9 (21.95%)	4 (9.76%)	27 (65.85%)	1 (2.44%)
	Alpha-conotoxin *Lt14.1*	29 (45.31%)	3 (4.69%)	28 (43.75%)	4 (6.25%)
	Alpha-conotoxin *Lt28.1*	27 (31.76%)	9 (10.59%)	42 (49.41%)	7 (8.24%)
C. regius	Alpha-conotoxin *RgIA*	13 (40.62%)	0	19 (59.38)	0
	Alpha-conotoxin *RgIIA*	32 (48.48%)	13 (19.70%)	17 (25.76%)	4 (6.06%)

Prediction of secondary structures is thought to be useful in conventional drug discovery. Identifying the secondary structure correctly can aid in sequence alignment [42], as the secondary structure is the polypeptide backbone of local conformation proteins. The α-

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

helix (H) and β-strand (E) secondary structure states are both regular, whereas the coil region is an irregular secondary structure type (C) [43]. Furthermore, in molecular recognition mechanisms, interactions between peptide substrates and receptors, and protein folding, it is assumed that irregular protein secondary structures represent key structural domains [44]. Lastly, according to [45], the secondary structural pattern of neurotoxins is utilized to assess whether an amino acid is in a helix, strand, or coil. Agreeing with their observations, random coil outperforms the other forms in secondary structure prediction of the snake venom neurotoxins studied, with C-turn being the least conformational shape.

3.4 Homology modeling of alpha-conotoxins and structure verification

Figure 1 (*Parts 1* and *2*) exhibit the predicted three-dimensional (3D) structure of the seventeen (17) alpha-conotoxin based on protein sequences acquired from the protein database. This was built using the SWISS-Model tool and is based on the Global Model Quality Estimation (GMQE) and Qualitative Model Energy Analysis (QMEAN) data. The GMQE values are generally between 0 and 1, and the higher the number, the more reliable the projected structure is, whereas a score less than 4.0 suggests reliability for QMEAN. SWISS-Model requires more than 30 amino acids to generate a model. Thus, *Conus imperialis'* alpha-conotoxins, namely α-conotoxin *ImII* and *ImI,* are the only proteins that cannot build a 3D model due to their short lengths of seventeen (17) amino acids.

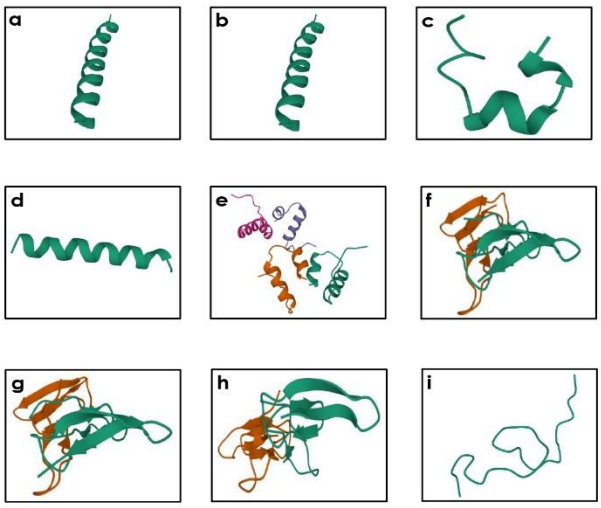

Fig 1: Homology models of the seventeen (17) alpha-conotoxin from ten (10) selected vermivorous *Conus* species using SWISS-Model Tool (*Part 1*). (a) Alpha-conotoxin *Lp1.4*, (b) Alpha-conotoxin *Lp1.1*, (c) Alpha-conotoxin *ArIA,* (d) Alpha-conotoxin *SrIA/SrIB*, (e) Alpha-conotoxin *VxXXIVA*, (f) Alpha-conotoxin *VxXXC*, (g) Alpha-conotoxin *VxXXB*, (h) Alpha-conotoxin *VxXXA,* and (i) Alpha-conotoxin *Pu14.1*

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

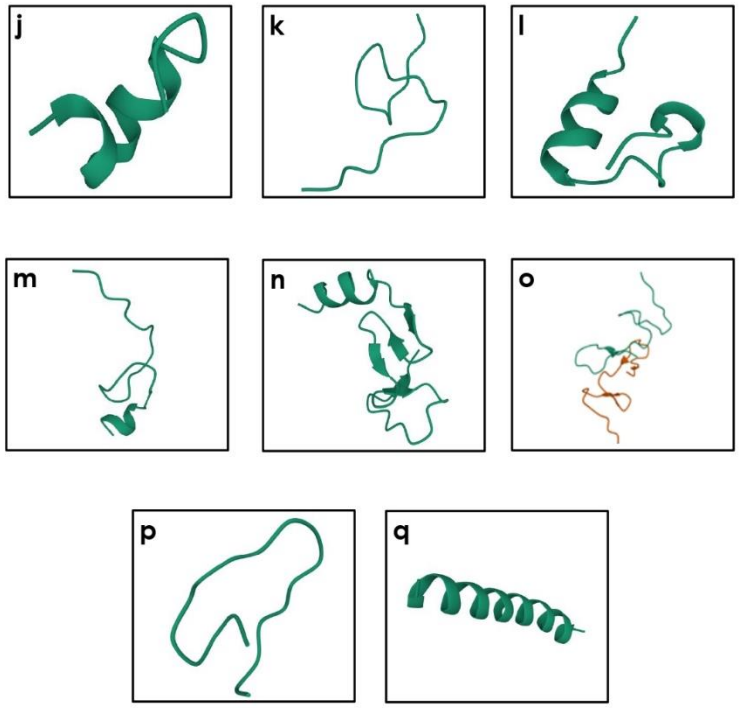

Fig 1: Homology models of the seventeen (17) alpha-conotoxin from ten (10) selected vermivorous *Conus* species using SWISS-Model Tool (*Part 2*). (j) Alpha-conotoxin *LvIA*, (k) Alpha-conotoxin *Li1.12,* (l) Alpha/Kappa conotoxin *Pl14a*, (m) Alpha-conotoxin *Lt1.3b*, (n) Alpha-conotoxin *Lt14.1*, (o) Alpha-conotoxin *Lt28.1*, (p) Alpha-conotoxin *RgIA*, and (q) Alpha-conotoxin *RgIIA*

A critical component of drug discovery is the structure-based design of small-molecule inhibitors of protein-ligand and protein-protein interactions. The structural similarity of the secondary structure parts can be used to determine the underlying protein interactions [46], and the application of homology modeling generates reliable structural models. The receptor binding assay is one example; when the receptor binds to the ligand, signal transduction events occur, which govern numerous biological processes essential for cell development and function. As a result, identifying chemicals that precisely bind to cellular receptors to inhibit (antagonists) or boost (agonists) a biological function is a key emphasis in drug discovery [47]. However, according to [48], while it was thought that the receptor binding method could be used as a primary screen to accelerate and simplify the identification of new drug candidates, experience has shown that ligand binding is most useful for drug discovery when combined with functional, phenotypic assays.

3.5 Validation of alpha-conotoxins models

According to [16], homology modeling was the most accurate computational method for generating reliable structural models, and it was used in many biological applications. The

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

models' reliability was estimated using model quality assessment tools. Following the refinement process, the stereochemical quality of the predicted model and the accuracy of the protein model were assessed using Ramachandran map calculations computed with the PROCHECK program. Ramachandran plot is a basic two-dimensional graphic representation of all conceivable protein structures. Each amino acid residue in a polypeptide can have a unique collection of angles; it can be represented as a point on a Ramachandran plot with associated angles [49].

Ramachandran plots were generated from the PDBsum web server and were shown in Figure 2 (*Parts 1 to 3*), where the red regions in the graph indicate the most favored regions. Glycine residues are represented by triangles and other residues are represented by squares. The results revealed, as shown in Table 5, that the modeled structure for alpha-conotoxins *Lp 1.4, ArIA, VxXXC, VxXXA, LvIA*, and *RgIIA* have a percentage of 90%, 95.2%, 100%, 100%, 91.7%, and 100%, respectively, under their favored regions. Such Ramachandran plot figures indicate that the predicted models are of high quality. Furthermore, the remaining alpha-conotoxins have a percentage of less than 90% in their favored regions, indicating that the predicted models are of low quality.

Table 5: Ramachandran Plot Calculation And Analysis Of The Models Of The Nineteen (19) Alpha-Conotoxins From Ten (10) Selected Vermivorous *Conus* Species From The Swiss Model And Computed With The Procheck Program

SPECIES	ACCESSION NUMBER	PROTEIN NAME	RAMACHANDRAN PLOT	
			FAVORED REGION	DISALLOWED REGION
Conus leopardus	A1X8B6	Alpha-conotoxin *Lp 1.4*	63 (90%)	0 (0%)
	Q6PTD5	Alpha-conotoxin *Lp1.1*	66 (85.7%)	1 (1.3%)
Conus arenatus	P0C8R2	Alpha-conotoxin *ArIA*	20 (95.2%)	0 (0%)
Conus spurius	P85886	Alpha-conotoxin *SrIA/SrIB*	103 (85.8%)	0 (0%)
Conus vexilium	J7JU64	Alpha-conotoxin *VxXXIVA*	11 (73.3%)	0 (0%)
	P0C1W7	Alpha-conotoxin *VxXXC*	23 (100%)	0 (0%)
	P0C1W6	Alpha-conotoxin *VxXXB*	63 (87.5%)	0 (0%)
	P0C1W5	Alpha-conotoxin *VxXXA*	23 (100%)	0 (0%)
Conus pulicarius	C6ZJQ2	Alpha-conotoxin *Pu14.1*	17 (73.9%)	0 (0%)
Conus lividus	L8BU87	Alpha-conotoxin *LvIA*	11 (91.7%)	0 (0%)
	H9N3R7	Alpha-conotoxin *Li1.12*	35 (83.3%)	1 (2.4%)
Conus planorbis	Q0N4U8	Alpha/Kappa conotoxin *Pl14a*	12 (57.1%)	1 (4.8%)
Conus litteratus	A0A068B0Z6	Alpha-conotoxin *Lt1.3b*	46 (82.1%)	1 (1.8%)
	Q2I2R5	Alpha-conotoxin *Lt14.1*	5 (35.7%)	0 (0%)
	F6JWU7	Alpha-conotoxin *Lt28.1*	12 (63.2%)	1 (5.3%)
Conus regius	P0C1D0	Alpha-conotoxin *RgIA*	14 (77.8%)	0 (0%)
	P85013	Alpha-conotoxin *RgIIA*	23 (100%)	0 (0%)

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

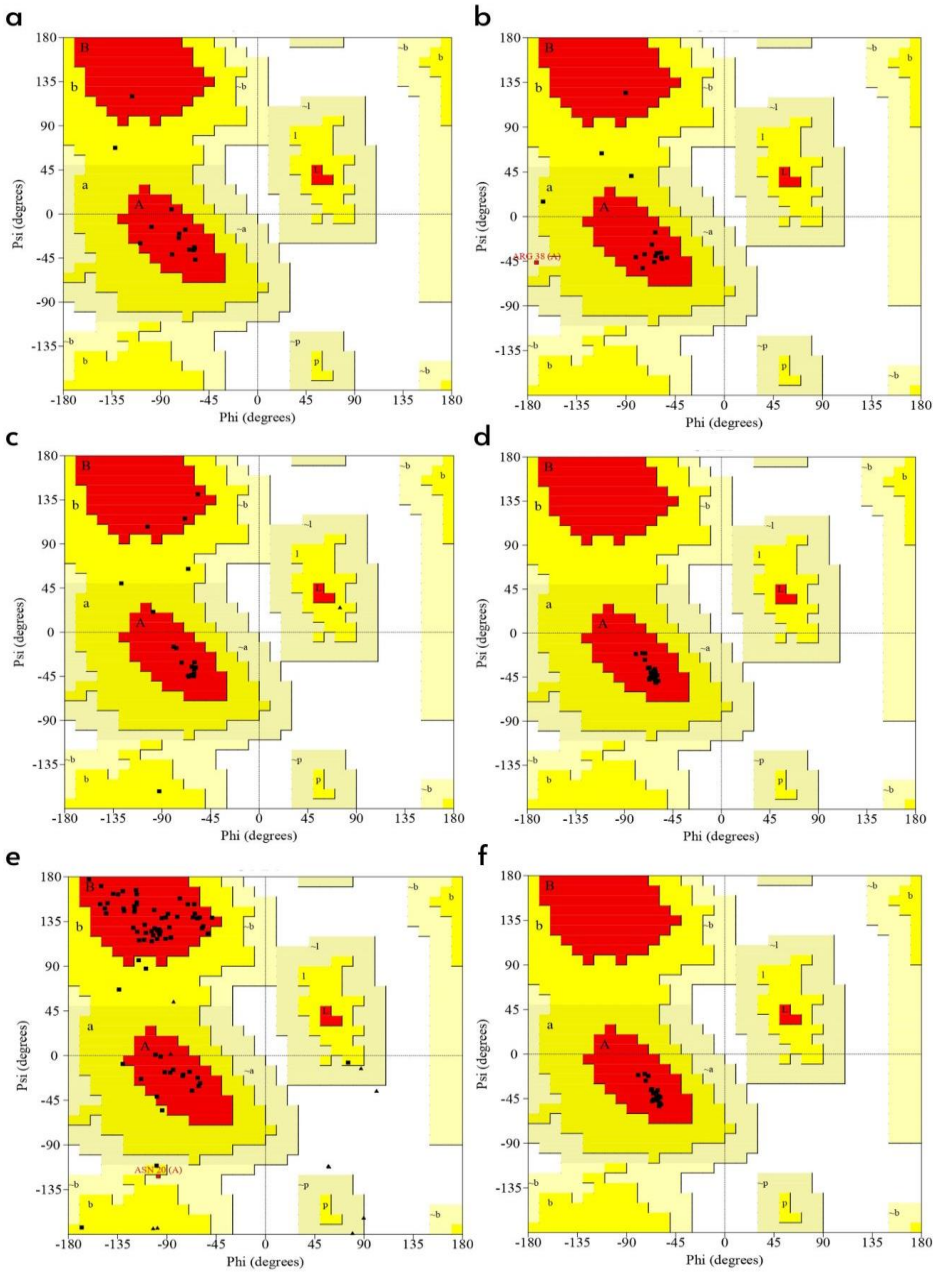

Fig 2: Ramachandran plots of the seventeen (17) alpha-conotoxin predicted models (*Part 1*). (a) Alpha-conotoxin *LvIA*, (b) Alpha-conotoxin *VxXXIVA*, (c) Alpha-conotoxin *RgIA*, (d) Alpha-conotoxin *RgIIA*, (e) Alpha-conotoxin *VxXXB,* and (f) Alpha-conotoxin *VxXXA*

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

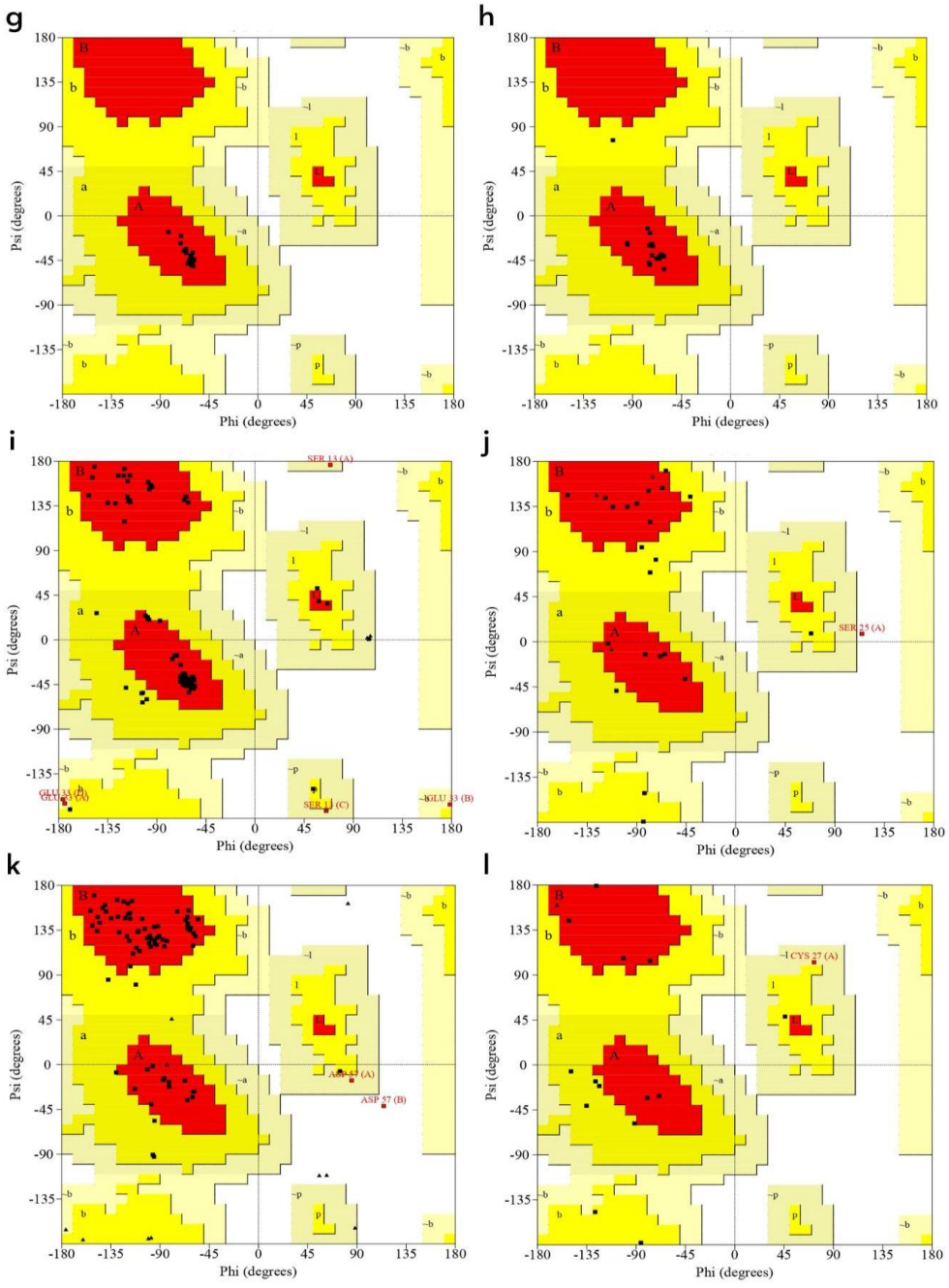

Fig 2: Ramachandran plots of the seventeen (17) alpha-conotoxin predicted models (*Part 2*). (g) Alpha-conotoxin *VxXXC*, (h) Alpha-conotoxin *ArlA*, (i) Alpha-conotoxin *SrlA/SrlB*, (j) Alpha/Kappa conotoxin *Pl14a*, (k) Alpha-conotoxin *Lp1.1*, and (l) Alpha-conotoxin *Lt14.1*

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

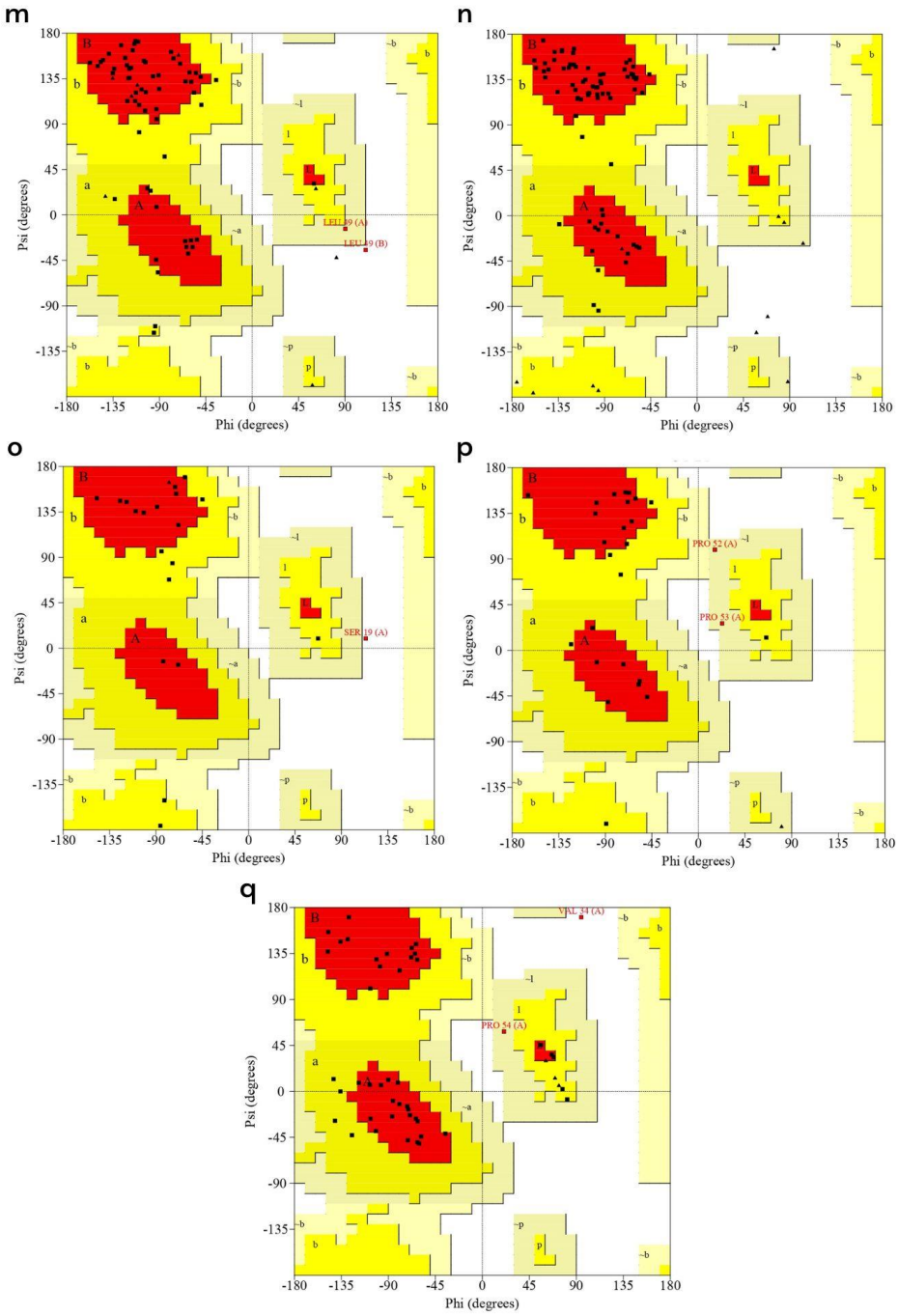

Fig 2: Ramachandran plots of the seventeen (17) alpha-conotoxin predicted models (*Part 3*). (m) Alpha-conotoxin *Lt1.3b*, (n) Alpha-conotoxin *Lp 1.4*, (o) Alpha-conotoxin *Lt28.1*, (p) Alpha-conotoxin *Pu14.1,* and (q) Alpha-conotoxin *Li1.12*

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

4. CONCLUSIONS AND FUTURE WORKS

The purpose of this study was to use in silico technologies to produce, evaluate, and compare the structure of alpha-conotoxin from selected vermivorous *Conus* species. Physicochemical qualities were utilized to calculate molecular weight, isoelectric pI, instability index, aliphatic index, and grand average of hydropathy, among other physical and chemical factors. Furthermore, property prediction, domain scanning, secondary structure prediction, homology modeling, and structural analysis were all included. The results revealed that the majority of the alpha-conotoxins were highly basic in nature, with one alpha-conotoxin sequence (Alpha-conotoxin *Lt28.1*) indicating acidic in nature. The instability index varies from stable Alpha-conotoxin *RgIA* (14.62) to unstable Alpha-conotoxin *ImII* (126.91). Additionally, the aliphatic index for the protein sequences included in this study ranges from 29.15 to 97.50, indicating that it is thermostable over a wide temperature range. Moreover, GRAVY values for alpha-conotoxins range from -0.910 (most hydrophilic; alpha-conotoxin *SrIA/SrIB*) to 0.731 (most hydrophobic; alpha-conotoxin *Lp1.1*). Furthermore, the common domain found among the species was *toxin_8*. The percentage score of amino acid distribution for secondary structure prediction shows a random coil dominates other secondary structures, followed by alpha-helix, beta-strand, and beta-turns. Lastly, the models for Alpha-conotoxin *Lp1.4*, Alpha-conotoxin *ArIA*, Alpha-conotoxin *VxXXC*, Alpha-conotoxin *VxXXA*, and Alpha-conotoxin *LvIA* are of high quality, as these proteins have a percentage of ≥ 90% in their favored regions.

This study proposes doing a more extensive investigation on the alpha conotoxins of vermivorous *Conus* species for their molecular properties, as studies on them are unusual. Furthermore, more research is needed to better understand the structure and method of blocking neuromuscular transmission by selective binding to muscle nicotinic acetylcholine receptors (nAChR). Finally, researchers who intend to continue this type of research might use the same methods to examine the alpha-conotoxin found in piscivorous and molluscivorous cone snails, in order to gain a deeper understanding of this family of *Conus* venom for the optimum application in drug discovery.

References

1) Teichert, R, E Jimenez, and B Olivera. *"Biology and Pharmacology of Conotoxins".* Botulinum Toxin: Therapeutic Clinical Practice and Science, 446-464. 2009. doi:10.1016/B978-1-4160-4928-9.00036-6.

2) Prentis, PJ, A Pavasovic, and RS Norton. *"Sea Anemones: Quiet Achievers in the Field of Peptide Toxins".* Toxins, 10(1), 36. 2018. Doi: 10.3390/toxins10010036

3) Pennington, MW, A Czerwinski, and RS Norton. *"Peptide therapeutics from venom: Current status and potential".* Bioorganic and Medicinal Chemistry, 26(10), 2738-2758. 2018. doi: 10.1016/j.bmc.2017.09.029

4) [Calderon-Celis, F, L Cid-Barrio, JR Encinar, a Sanz-Medel, and JJ Calvete. *"Absolute venomics: Absolute quantification of intact venom proteins through elemental mass spectrometry".* Journal of Proteomics, 164, 33-42. 2017.

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

5) Dutertre, S, I Vetter, B Hamilton, K Sunagar, V Lavergne . . . and D Venter. *"Evolution of separate predation- and defense-evoked venoms in carnivorous cone snails".* Nature Communications, 5, 3521. 2014.

6) Mari, F and J Tytgat. (2010). *"2.15 Natural Peptide Toxins".* Comprehensive Natural Products II: Chemistry and Biology, 2, 511-538. 2010.

7) Lewis, R, S Dutertre, I Vetter, and M Christie. *"Conus venom peptide pharmacology".* Pharmacological Reviews, 64, 259-298. 2012.

8) Olivera, B, D Hillyard, M Marsh, and D Yoshikami. *"Combinatorial peptide libraries in drug design: Lessons from venomous cone snails".* Trends in Biotechnology, 13(10), 422-426. 1995. Doi: 10.1016/S0167-7799(00)88996-9.

9) Bruce, G, R Ken, and K Zeinab. *"Drugs from the sea: Conopeptides as potential therapeutics".* Current Medicinal Chemistry, 11, 1715-1723. 2004.

10) Robinson, S and R Norton. *"Conotoxin gene superfamilies".* Marine Drugs, 12(12), 6058-6101. 2006. Doi: 10.3390/md12126058.

11) Jin, AH, M Muttenthaler, S Dutertre, S Himaya, Q Kaas, DJ Lewis, and PF Alewood. *"Conotoxins: Chemistry and Biology".* Chemical Reviews, 119(21), 11510-11549. 2019. doi: 10.1021/acs.chemrev.9b002

12) Akondi, K, M Muttenthaler, S Dutertre, Q Kaas, D Craik, R Lewis, and PF Alewood. *"Discovery, synthesis, and structure-activity relationships of conotoxins".* Chemical Reviews, 114, 5815–5847. 2012. doi: http://dx.doi.org/10.1021/

13) Azam, L and JM Mcintosh. *"Alpha-conotoxins as pharmacological probes of nicotinic acetylcholine receptors".* Acta Pharmacologica Sinica, 30(6), 771-783. 2009.

14) Duterte, S, A Nicke, and V Tsetlin, V. (n.d.). *"Nicotinic acetylcholine receptor inhibitors derived from snake and snail venoms".* Neuropharmacology, doi: 10.1016/j.neuropharm.2017.06.011.

15) Chikhale, HU and A Nerkar. *"Review on in-silico techniques: an approach to drug discovery".* Current Trends in Pharmacy and Pharmaceutical Chemistry. 2 (1), 24-32. 2020.

16) Sahay, A, A Piprodhe, and M Pise. *"In silico analysis and homology modeling of strictosidine synthase involved in alkaloid biosynthesis in Catharanthus roseus".* Journal of Genetic Engineering and Biotechnology. 18:44. 2020. doi: https://doi.org/10.1186/s43141-020-00049-3

17) Waterhouse, A, M Bertoni, S Bienert, G Studer, G Tauriello, R Gumienny, … and T Schwede. *"SWISS-MODEL: homology modeling of protein structures and complexes".* Nucleic Acids Research, 46(1), W296-W303. 2018. doi:10.1093/nar/gky427

18) Browne, P, R Eberhardt, and the UniProt Consortium. Capturing Protein Sequence Data. 2009.

19) Gad, SE and SC Gad. *"Toxicity Testing, Dermal".* Encyclopedia of Toxicology, 2, 252-262. 2005 doi:10.1016/B0-12-369400-0/00966-2

20) Dwivedi, V, S Arora, and A Pandey. *"Computational analysis of physicochemical properties and homology modeling of carbonic anhydrase from Cordyceps militaris".* Network Modeling and Analysis in Health Informatics and Bioinformatics, 2, 209-212. 2013. Doi: 10.1007/s13721-013-0036-8

21) Geourjon, C and G Deléage. *"SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments".* Computer applications in the biosciences: CABIOS, 11(6), 681-684. 1995. doi:10.1093/bioinformatics/11.6.681

22) Hamed, T, Y Haddad, V Adam, and Z Heger. *"Homology modeling in the time of collective and artificial intelligence".* Computational and Structural Biotechnology Journal, 18, 3494-3506. 2020. doi: 10.1016/j.csbj.2020.11.007

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

23) Muhammed, M and E Aki-Yalcin. *"Homology modeling in drug discovery: Overview, current applications, and future perspectives".* Chemical Biology and Drug Design, 93(1), 12-20. 2019.

24) Read, R and G Chavali. *"Assessment of CASP7 predictions in the high accuracy template-based modeling category".* Proteins 69(S8), 27-37. 2007.

25) Jalily, HH and K Barakat. *"Homology Modeling: An Overview of Fundamentals and Tools".* International Review on Modelling and Simulations, 10(2), 1-14. 2017.

26) Haddad, Y, V Adam, Z Heger. *"Ten quick tips for homology modeling of high-resolution protein 3D structures".* PLOS Computational Biology, 16(4), 1-19. 2020.

27) Oduselu, GO, OO Ajani, YU Ajamma, B Brors, and E Adebiyi. *"Homology Modelling and Molecular Docking Studies of Selected Substituted Benzo[d]imidazol-1-yl) methyl) benzimidamide Scaffolds on Plasmodium falciparum Adenylosuccinate Lyase Receptor".* Bioinformatics and Biology Insights 13, 1-10. 2019. Doi: 10.1177/117793221986553

28) Cardoso, J, L Fonseca, C Egas, and I Abrantes. *"Cysteine proteases secreted by the pinewood nematode, Bursaphelenchus xylophilus: in silico analysis".* Computational Biology and Chemistry, 77, 291-296. 2018.

29) Benkert, P, M Kunzli, and T Schwede. *"QMEAN server for protein model quality estimation".* Nucleic Acids Research, 37, W510-W514. 2009.

30) Laskowski, R. *"PDBsum new things".* Nucleic Acids Research, 37, D355-D359. 2008.

31) Moldoveanu, SC and V David. *"Chapter 5 - Properties of Analytes and Matrices Determining HPLC Selection".* Selection of the HPLC Method in Chemical Analysis, 189-230. 2017. doi:10.1016/B978-0-12-803684-6.00005-6.

32) Gamage, DG, A Gunaratne, GR Periyannan, and TG Russell. *"Applicability of Instability Index for In vitro Protein Stability Prediction".* Protein and Peptide Letters, 26(5), 339-347. 2019. Doi: 10.2174/0929866526666190228144219

33) Guruprasad, K, BVP Reddy, and MW Pandit. *"Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence".* Protein Eng 4:155–164. 1990. doi: https://doi.org/10.1093/protein/4.2.155

34) Gasteiger, E, C Hoogland, a Gattiker, S Duvaud, M Wilkins, R Appel, a Bairoch. In: Walker JM, editor. The Proteomics Protocols Handbook. Humana Press. 2005.

35) Finn, R, A Bateman, J Clements, P Coggill, R Eberhardt, S Eddy … and M Punta. *"Pfam: the protein families database".* Nucleic Acids Research, 42, D222-30. 2013. doi:10.1093/nar/gkt1223

36) Mohan, R and S Venugopal. *"Computational structural and functional analysis of hypothetical proteins of Staphylococcus aureus".* Bioinformation, 8(15), 722-728. 2012. Doi: 10.6026/97320630008722

37) Marchler-Bauer, A, JB Anderson, MK Derbyshire, C DeWeese-Scott, NR Gonzales, M Gwadz, M … C Liu. *"CDD: a conserved domain database for interactive domain family analysis".* Nucleic Acids Research, 35(suppl_1), D237-D240. 2007. doi:10.1093/nar/gkl951

38) Santhoshkumar, R and A Yusuf. *"In silico structural modeling and analysis of physicochemical properties of curcumin synthase (CURS1, CURS2, and CURS3) proteins of Curcuma longa".* J Genet Eng Biotechnol, 18(1), 24. 2020. Doi: 10.1186/s43141-020-00041-x

39) Gupte, TM, M Ritt, and S Sivaramakrishnan, S. *ER/K-link-Leveraging a native protein linker to probe dynamic cellular interactions.* Methods in Enzymology. 2020. doi: 10.1016/bs.mie.2020.10.002

40) Pace, CN and JM Scholtz. (1998). *"A helix propensity scale based on experimental studies of peptides and proteins".* Biophysical Journal. pp. 422-427. 1998. doi: 10.1016/s0006-3495(98)77529-0.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 06-2023
DOI 10.17605/OSF.IO/9MRVS

41) Smith, JR, M Nakanishi, RS Robetorye, SF Venable, and OM Pereira-Smith. *"Studies demonstrating the complexity of regulation and action of the growth inhibitory gene SDI1".* Exp. Gerontol. 31 (1-2):327-35. 1996. doi: 10.1016/0531-5565(95)00026-7. PMID: 8706802.

42) Shiragannavar, S. (2017, January). *Methods of prediction of secondary structures of proteins.* Biotech Articles. Retrieved from: http://www.biotecharticles.com/Bioinformatics-Article/Methods-of-Prediction-of-Secondary-Structures-of-Proteins-3759.html

43) Ma, Y, Y Liu, and J Cheng. *"Protein Secondary Structure Prediction Based on Data Partition and Semi-Random Subspace Method".* Scientific Reports, 8. 2018. Doi: 10.1038/s41598-018-28084-8

44) Pavone, V, G Gaeta, a Lombardi, F Nastri, O Maglio, C Isernia, M Saviano. *"Discovering protein secondary structures: classification and description of isolated alpha-turns".* Biopolymers, 38(6), 705-721. 1996. doi:10.1002/(SICI)1097- 0282(199606)38:6%3C705::AID-BIP3%3E3.0.CO;2-V

45) Hasan, M, N Belhaj, H Benachour, M Barberi-Heyob, CJF Kahn, E Jabbari, MLinder, and Arab-Tehrany. "*Liposome encapsulation of curmin: Physico-chemical characterizations and effects on MCF7 cancer cell proliferation*". International Journal of Pharmaceutics. 461(1-2), 519-528. 2014. doi: 10.1016/j.ijpharm.2013.12.007

46) Koch, AW, T Mathivet, B Larrivée, RK Tong, J Kowalski, L Pibouin-Franer, … and A Eichmann. *"Robo4 maintains vessel integrity and inhibits angiogenesis by interacting with UNC5B".* Developmental cell, 20(1), 33-46. 2011.

47) Mohs RC and NH Greig. *"Drug discovery and development: role of basic biological research".* Alzheimers Dement (NY). 3(4):651-657. 2017. doi: 10.1016/j.trci.2017.10.005.

48) Bylund, DB and SJ Enna. *"Receptor binding assays and drug discovery".* Advances in Pharmacology, 82, 21-34. 2018.

49) Kumar, P and an Arya. Ramachandran plot- A simplified approach. 2019.