Cyclotides: a source of bioactive metabolites from plants

Chamari M Hettiarachchi

Department of Chemistry, University of Colombo

Introduction

In 1965, the professor in pharmacognosy at Uppsala University, Finn Sandberg reported after an expedition the traditional medicinal use of plants in the Central African Republic. In his report he mentioned the use of decoctions of the plant *Oldenlandia affinis* by African natives to facilitate childbirth.[1] A few years later in the 1970s, a Red Cross doctor, Lorens Gran, observed that tribes in Congo used the plant and he brought samples home to Norway for identification and investigation. He discovered that polypeptides in the plant had remarkably strong uterotonic activity. The main active component was almost completely sequenced and called kalata B1, after the native name of the plant ´Kalata-Kalata´.[2,3] The structure of kalata B1 was not fully elucidated until 1995, when its circular backbone and knotted arrangement of disulphide bonds were clarified by nuclear magnetic resonance (NMR) analysis.[4] By this time three different groups independently published studies of macrocyclic peptides with six cysteine residues from the plants *Viola arvensis*, *Chassalia parvifolia* and *Psychotria longipes*.[5-7] The macrocyclic peptides with a cystine knot grew over the following years and finally collective term for them was suggested as cyclotides after *cyclo-peptides*.[8]

Structure and sequence of cyclotides

Cyclotides are an exceptional family of gene-encoded plant proteins, that are cyclic and their N and C termini are joined by a peptide bond, forming a continuous circular backbone. The circular cyclotide chain consists of approximately 28-37 amino acids, including six cysteins residues that form three disulfide bonds arranged in a cyclic cystine knot (CCK) motif (Figure 1).[8,9] The unique cyclotide structure forces hydrophobic residues to be exposed on the surface of the protein making them amphipathic proteins. The remarkable CCK motif makes cyclotides extremely resistant to enzymatic, chemical and thermal degradation[10] and ideal for developing cyclotide based peptides with diverse medical and agricultural applications.[11,12]

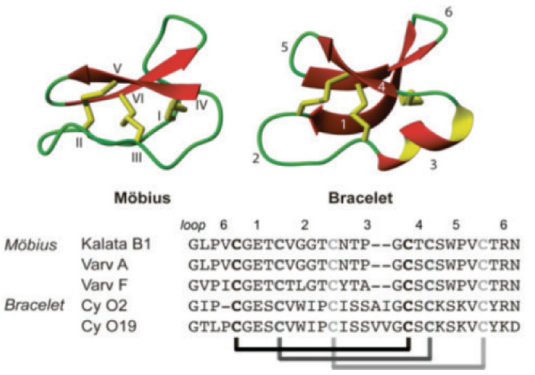

Figure 1. Schematic structures of Möbius and bracelet cyclotides, together with typical cyclotide sequences from both subfamilies. The abbreviated notation cyO2 and cyO19 stand for the cycloviolacins O2 and O19, respectively. The unique features of the CCK motif: a cyclic backbone with sequence loops (1-6) and three stabilizing disulphide bonds. These disulphides are arranged in a cystine knot, that is two of the disulphides form a ring structure together with the backbone connecting the four cysteines (I-IV; II-V), while the third disulphide is threaded through the ring (III-VI)

Cyclotides are divided into two main subfamilies, the möbius and the bracelet. They are characterized by the presence or absence of a *cis*-pro peptide linkage, respectivily.[8] The subfamilies also differ in size and amino acid content, the bracelets being the more structurally diverse of the two; to date according to cybase (the database of cyclic proteins)[13,14], 2/3 of the known cyclotides belong to the bracelet subfamily, and the rest to the möbius subfamily. Some residues are found in all/most cyclotides, the strictly conserved Cys residues with the intermediate residues defined as loops (Figure 1), a Glu residue in loop 1, and a Gly-Asn/Asp sequence in loop 6 (residues involved in the post-translational ring closure. The rest of the residues are

changeable, and although there are relatively few amino acids in a cyclotide sequence, variations are immense.

The connectivity of the six Cys residues have been debated and can theoretically form 15 possible variants. The real conformation has been deduced from NMR spectra[4,15,16], and verified by chemical proofs achieved by partial reduction of the disulfides with stepwise alkylation[17] and recently by X-Ray crystallography[18]. The consensus conclusion from these studies is that they connect as follows: CysI-CysIV, CysII-CysV and CysIII-CysVI forming the cystine knot.

Distribution and occurrence of cyclotides in plants

The plant families *Rubiaceae*[2], *Solanaceae*[19], *Fabaceae*[20], *Curcubitaceae*[21] and *Violaceae*[22] are known to contain cyclotides. More than 400 cyclotide sequences have been reported from these families.[23] The violaceae family includes about 23-31 genera and 1000-1100 species worldwide.[24,25] Three of these genera, *Viola*, *Rinorea* and *Hybanthus* accounts for 90% of the species in the family.[26] *Viola* is the largest genus in the family, having 580-620 species in total.[27] There are well established phylogenetic systems for the Violaceae family, and especially for the genus *Viola*. There are many cyclotides from Violaceae extracted worldwide and in Sri Lanka there are eight species reported however, species under *Rinorea* genus reported as extinct (EX) and other species are still under critical stage (CR), also exhibits only in wild forests.[24]

Biosynthesis of cyclotides

Cyclotides are one of few naturally derived classes of macrocyclic gene encoded peptides.[26] Analysis of cyclotide precursor sequences obtained from cDNA have shown that the genes encoding them consist of an endoplasmic reticulum (ER) signal domain, a pro-region and one to three mature cyclotide domains, each proceeded by an N-terminal repeat (NTR) sequence (Figure 2).[28,29] The cleavage points of a schematic cyclotide after a Lys/Gly/Asn residue in the NTR sequence and the Asn or Asp in the cyclotide domain. Details of the processing of the precursors, including the order of the events, are not fully understood, However, involve oxidative folding, excision of the mature cyclotide sequence and head-to-tail cyclization. An asparaginyl-endoproteinase has been suggested to be involved in cleavage of the C-terminal tail and simultaneous cyclization of the cyclotide domain, at least for the prototypic cyclotide kalata B1.[30] Additionally, a protein-disulfide isomerase seems to play a major role in the oxidative folding of cyclotides through re-shuffling (isomerization) of disulfide bonds.[31]

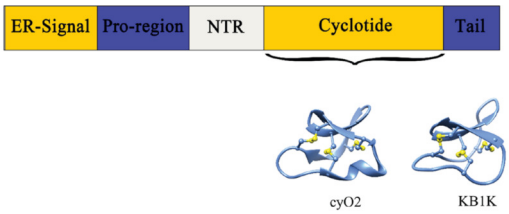

Figure 2. Biosynthesis and structure of cyclotides.

Biological activities of cyclotides

In 1970 the first report on activity of cyclotide for Kalata B1 which is responsible for uterotonic activity. In the early 1990's, a series of independent reports were published describing cyclotides discovered in bioassay-guided isolations, including the haemolytic violapeptide I from *viola arvensis*[5], the neurotensin-binding inhibitor cyclopsychotride A from *Psychotria longipes*[32], and the circulins A-F with anti-HIV properties from *Chassalia pravifolia*.[6] In subsequent assays cyclotides showed activities in the low micromolar range against a wide range of pests and other organisms; insecticidal effects against *Helicoverpa punctigera* and *H. armigera* larvae, toxicity towards golden apple snails (*Pomacea canaliculata*) and Nile tilapia fish (*Oreochromis niloticus*), gastrointestinal nematode parasites of sheep, and potent effects against fouling barnacles (*Balanus improvisus*) larvae from settling. Cyclotides, also showed activity against human pathogens including *Escherichia coli*, *Klebsiella oxytoca*, *Staphylococcus aureus*, *Candida kefyr* and *Candida tropicalis*.[33] These pesticide and anti-pathogen effects support the hypothesis that cyclotides are components of the plant defense systems.

In conclusion, Cyclotides are the largest class of cyclic peptides found within plants with a highly interesting motif (cyclic-cystine-knot) and bioactivities (anti-HIV, antibacterial, neurontensin-antagonist, insecticide)

References

1. Sandberg, F., *Etude sur les plantes medicinales et toxiques d'Afrique equatoriale*; Paris: Cahiers de la Maboké: Paris, 1965.
2. Lorents Gran., *Basic Clin. Pharmacol. Toxico*l. **1973**, *33* (5), 400–408.
3. Sletten, K., Gran, L., *Medd. Nor. Farm. Selsk*. **1973**, *35*, 69–82.
4. Saether, O., Craik, D. J., Campbell, I. D., Sletten, K., Juul, J., Norman, D. G., *Biochemistry*. **1995**, *34* (13), 4147–4158.
5. Schöpke, T., Hasan, A. M., Kraft, R., Otto, A., Hiller, K., *Sci. Pharm*. **1993**, *61*, 145–153.
6. Gustafson, K. R., Sowder, R. C., Henderson, L. E., Parsons, I. C., Kashman, Y., Cardellina, J. H., McMahon, J. B., Robert W, B. J., Pannell, L. K., Boyd, M. R., *J. Am. Chem. Soc*. **1994**, *116* (20), 9337–9338.
7. Witherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., Sardana, M., *J. Nat. Prod*. **1994**, *57* (12), 1619–1625.
8. Craik, D. J., Daly, N. L., Bond, T., Waine, C., *J. Mol. Biol.* **1999**, *294* (5), 1327–1336.
9. Yeshak, M. Y., Burman, R., Asres, K., Goransson, U., *J. Nat. Prod*. **2011**, *74* (4), 727–731.
10. Colgrave, M. L., Craik, D. J., *Biochemistry* **2004**, *43* (20), 5965–5975.
11. Craik, D. J., Clark, R. J., Daly, N. L., *Expert Opin. Investig. Drugs* **2007**, *16* (5), 595–604.
12. Kolmar, H., *Curr. Opin. Pharmacol.* **2009**, *9* (5), 608–614.
13. Mulvenna, J. P., *Nucleic Acids Res*. **2006**, *34* (90001), D192–D194.
14. Wang, C. K. L., Kaas, Q., Chiche, L., Craik, D. J., *Nucleic Acids Res*. **2008**, *36* (SUPPL. 1).
15. Rosengren, K. J., Daly, N. L., Plan, M. R., Waine, C., Craik, D. J., *J. Biol. Chem*. **2003**, *278* (10), 8606–8616.
16. Nair, S. S., Romanuka, J., Billeter, M., Skjeldal, L., Emmett, M. R., Nilsson, C. L., Marshall, A. G., *Biochim. Biophys. Acta - Proteins Proteomics*. **2006**, *1764* (10), 1568–1576.
17. Göransson, U., Craik, D. J., *J. Biol. Chem*. **2003**, *278* (48), 48188–48196.
18. Wang, C. K., Shu-Hong, H., Martin, J. L., Sjögren, T., Hajdu, J., Bohlin, L., Claeson, P., Göransson, U., Rosengren, K. J., Tang, J., Ning-Hua, T., Craik, D. J., *J. Biol. Chem*. **2009**, *284* (16), 10672–10683.
19. Poth, A. G., Mylne, J. S., Grassl, J., Lyons, R. E., Millar, A. H., Colgrave, M. L., Craik, D. J., *J. Biol. Chem*. **2012**, *287* (32), 27033–27046.
20. Aaron G. Poth., Michelle L. Colgrave., Russell E. Lyons., Norelle L. Daly., David J. Craik., *Proc. Natl. Acad. Sci. U. S. A*. **2011**, *108* (25), 10127–10132.
21. Jean-François Hernandez, Jean Gagnon, Laurent Chiche, Tuyet Mai Nguyen, Jean-Pierre Andrieu, Annie Heitz, Thai Trinh Hong, T. Trân Châu Pham, Nguyen, D. Le., *Biochemistry*. **2000**, *39* (19), 5722–5730.
22. Schoepke, T., Hasan Agha, M. I., Kraft, R., Otto, A., Hiller, K., *Sci. Pharm*. **1993**, *61*, 145–153.
23. Ravipati, A. S., Poth, A. G., Troeira Henriques, S., Bhandari, M., Huang, Y. H., Nino, J., Colgrave, M. L., Craik, D. J., *J. Nat. Prod*. **2017**, *80* (5), 1522–1530.
24. Burman, R., Yeshak, M. Y., Larsson, S., Craik, D. J., Rosengren, K. J., Göransson, U., *Front. Plant Sci*. **2015**, *6*
25. Wahlert, G. A., Marcussen, T., de Paula-Souza, J., Feng, M., Ballard, H. E., *Syst. Bot*. **2014**, *39* (1), 239–252.
26. Marcussen, T., Jakobsen, K. S., Danihelka, J., Ballard, H. E., Blaxland, K., Brysting, A. K., Oxelman, B., *Syst. Biol.* **2012**, *61* (1), 107–126.
27. Daly, N. L., Rosengren, K. J., Craik, D. J., *Advanced Drug Delivery Reviews*. **2009**, 918–930.
28. Zhang, J., Liao, B., Craik, D. J., Li, J. T., Hu, M., Shu, W. S., *Gene* **2009**, *431* (1–2), 23–32.
29. Simonsen, S. M., *Plant cell online*. **2005**, *17* (11), 3176–3189.
30. Saska, I., Gillon, A. D., Hatsugai, N., Dietzgen, R. G., Hara-Nishimura, I., Anderson, M. A., Craik, D. J., *J. Biol. Chem*. **2007**, *282* (40), 29721–29728.
31. Gruber, C. W., Čemažar, M., Heras, B., Martin, J. L., Craik, D. J., *Trends in Biochemical Sciences*. **2006**,

455–464.

32. Witherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., Sardana, M., *J. Nat. Prod.* **1994**, *57* (12), 1619–1625.

33. Tam, J. P., Lu, Y.-A., Yang, J.-L., Chiu, K.-W., *Proc. Natl. Acad. Sci.* **1999**, *96* (16), 8913–8918.

Professor Chamari Hettiarachchi *obtained her BSc (Hons) Degree from University of Colombo and PhD in Biochemistry Molecular Biology from Jawaharlala Nehru University, New Delhi. Currently she serves as a Professor in Biochemistry & Molecular Biology at the Department of Chemistry, University of Colombo.*

Sri Lankan students shine at the 55th International Chemistry Olympiad 2023

The four triumphant students, Kuragodage Thisarindu Adeepa Lakshan (Richmond College, Galle), Robolge Pahanma Upani Lenora (Sanghamitta Balika Vidyalaya, Galle), Isitha Dinujaya Arachchi (Ananda College, Colombo) and Thuvarahan Chandrakumar (S. Thomas' College, Mt. Lavinia) won four bronze medals at the 55th International Chemistry Olympiad (IChO) 2023 held from 16th to 25th July 2023 in Zurich, Switzerland.

IChO is a competition organized for secondary school students from all over the globe with the objective of promoting their creativity and cognitive skills in solving chemistry related problems. This year a total of 348 highly talented high school students from 89 countries participated in this prestigious competition. The chemistry knowledge of these talented students was tested in five-hour-long practical and theoretical examinations. The Sri Lankan team was accompanied by the mentors, Dr Ireshika De Silva (University of Colombo) and Dr Chinthaka Ratnaweera (University of Ruhuna).

The Chemistry Olympiad Sri Lanka (COSL) committee which is a sub-committee under the Council of the Institute of Chemistry Ceylon (IChemC) organizes the Sri Lanka National Chemistry Olympiad (SLNChO) competition annually and trains the team that represents Sri Lanaka in this prestigious IChO competition. First, a pool of around 30 students is selected from the all-island preliminary selection examination held usually in January each year and thereafter, COSL committee conducts several practical and theory workshops for them. Subsequently, four representatives of the Sri Lankan team are selected from very competitive practical and theory examinations.

This is indeed a remarkable achievement by the Sri Lankan team as Sri Lanka physically participated in the IChO competition for the first time. From 2020 to 2023 the IChO competitions were held as remote examinations due to the global pandemic situation. Sri Lankan team won three bronze medals in 2020 and 2021, and one silver and three bronze medals in 2022.

The champions of the IChO 2023 were recognized at a felicitation ceremony held on 19th August 2023 at the Institute of Chemistry Ceylon (IChemC) headquarters, Adamantane House in Rajagiriya. The event was attended by students, their parents, Council members of the Institute, members of the COSL committee, and invitees.

The applications for the preliminary selection examination of the SLNChO 2024 will be called in October 2023. For more details visit the COSL website https://www.web.ichemc.edu.lk/cosl/ and COSL Facebook page . Information regarding the application process will be uploaded soon on the website.

(please see inner back cover for photographs)