Review Article

The Biological Implications of Threaded Protein Topologies

Ellinor Haglund*

Department of Chemistry, University of Hawaii at Manoa, USA

Abstract

Protein drugs are used to treat various diseases including cancer, autoimmune disorders, infectious diseases, and genetic disorders. They have an advantage to small molecules as they directly target specific proteins and receptors with great precision, minimizing off-target effects. This specificity can reduce the risk of side effects compared to small molecules, which often have broader interactions. One of the major challenges in protein-based therapeutics is the cost of production, administration, and storage limiting their accessibility. However, proteins can be engineered and modified to enhance their properties, such as increasing their stability, altering their pharmacokinetics, or adding specific functionalities. Cyclic peptides have gained significant attention in drug design and development due to their unique properties and potential therapeutic benefits. Cyclic peptides are small proteins where the backbone forms a macrocycle. The novel class of proteins called pierced lasso topology (PLTs), may have an untapped potential in drug design as the biological activity is linked to conformational dynamics controlled by the threaded topology. Interestingly, the threaded topology is controlled by the redox potential fine-tuning biological activity on and off. In this review, we will discuss the role of PLTs in human health and their possible pharmaceutical applications.

INTRODUCTION

New technologies have revolutionized modern drug research and development to incorporate knowledge-based approaches to improve human health. Modern computational and experimental techniques together with the human genome project have expanded the molecular understanding of protein function, or lack of, aiding in the design of recombinantly engineered proteins. Therapeutic proteins can replace abnormal or deficient proteins in genetic disorders as seen for example in the treatment of diabetes using insulin replacement therapy [1] and genetic obesity using leptin replacement therapy [2,3] as well as in other rare genetic disorders [4].

These proteins may be genetically modified to resemble the natural proteins they replace, or they can be enhanced by small molecules that extend the protein's duration of activity and stability *in vivo*. Based on their pharmacological activity, therapeutic proteins can be divided into groups: (*i*) replacement therapy when a protein is deficient or abnormal, (*ii*), supplementing existing pathway(s), (*iii*) providing a novel function or activity, (*iv*) interfering with a molecule or organism, and (*v*) delivering other compounds or proteins [5]. Protein-based drug design has revolutionized the treatment of many diseases by offering targeted and often more effective treatments. For example, the COVID-19 pandemic caused by the

SARS-CoV-2 virus has significantly contributed to the vast burst of mRNA and protein-based vaccines developed. Understanding the molecular interactions between SARS-CoV-2 and human cell receptor proteins were crucial in the development of vaccines targeting the protein: protein interactions responsible for the virus: host interactions [6]. Protein-based therapeutics have unique challenges related to production, administration, and cost. To overcome these obstacles, the special features of disulfide bonds were explored in protein engineering and drug design. They are mainly introduced to stabilize proteins and/ or "trap" topologies in an active/inactive state. This is seen in for instance ins: (*i*) Circular proteins seem to have a common role in host defense mechanisms [7], (*ii*) Cystine knots are exceptionally stable and protected against degradation, and have been implemented as therapeutics acting as ion channel blockers, haemolytic agents, and antiviral and antibacterial activities [8], (*iii*) Lasso peptides have high stability and protected against proteolysis due to the compact structure and the lack of an N-terminus [7]. Lasso peptides possess a wide range of biological activities including receptor antagonist, enzyme inhibition, and antiviral antimicrobial activity [9], (*iv*) Protein catenanes have a versatile role and holds a great promise in for example industrial enzyme engineering and therapeutic proteins [10]. The remarkable stability, protection against degradation, and diverse pharmacologically relevant bioactivities makes these proteins great therapeutic models [11-13]. Ongoing research aims to

address these challenges and expand the use of protein drugs in medicine. In this review, we will discuss the role of proteins with a pierced lasso topology (PLT) in human health and their possible pharmaceutical applications.

ENTANGLED PROTEINS IN THERAPEUTICS

It is hypothesized that entanglements in proteins play a role against protein degradation [14], provide structural stability in transporter proteins [15], enhance the structural rigidity of the native state [16], help shape and form the binding site of enzymes [17-19], enhance thermal [20,21] and mechanical [21-23] stability, or even alter enzymatic activity [22], or act as molecular switches [24,25]. Specifically, directed evolution technologies and rational design aid in the engineering of topologically complex proteins, further expanding the scope and diversity of the rapeutic proteins. Cyclic peptides are a closed loop utilizing peptide bonds, creating a cyclic structure in which the amino acid sequence can thread though. These peptides provide an interesting topology due to their unique properties and potential therapeutic benefits. They are stable protein structures due to their closed-loop structure which can prolong the therapeutic effect and may increase the binding affinity for their target molecules [26]. Cyclic peptides have also shown to be effective in targeting protein-protein interactions, which are often difficult to disrupt using small molecules. By designing cyclic peptides that mimic specific protein surfaces or interfaces, the peptides can target a variety of biomolecules, including proteins, receptors, and enzymes and potentially modulate disease pathways seen in cancer biology, infectious diseases, inflammation, and central nervous system disorders [27]. Protein engineering of cyclic peptides have been able to enhance their oral bioavailability, making them more suitable for administration as pills or capsules.

Specificity can also be designed from varying the amino acid sequence adding specificity to develop cyclic peptides tailored to target. different biological molecules. Several cyclic peptides are already used as for example in antibiotic (vancomycin) and immunosuppressive (cyclosporine) treatments. Thus, they represent a promising class of molecules in drug design. Ongoing research and technological advancements are likely to lead to the development of more cyclic peptide-based therapeutics in the future.PLTs in drug development. The novel class of PLTs was discovered almost a decade ago, where a single disulfide bond forms a closed-loop where part of the backbone is threaded through [28-30] (Figure 1). The novel topology was shown to affect the conformational dynamics controlling biological activity [28]. Thus, PLTs may act as molecular switches *in vivo* [24,25,29], i.e., acting as an on/off switch controlled by the chemical environment. Since its discovery, more than 600 proteins with a disulfide-linked PLT have been found in nature [25,30,31]. PLTs are represented in all kingdoms of life and care classified with 14 different protein biological functions, from hydrolases and immune system proteins to membrane proteins [25] [Figures 1,2]. Human proteins are represented in 10 out of the 14 classes [Table 1], found in many cell-signaling, bacterial and viral proteins (Figure 2B). This novel class of threaded topologies have been less explored in drug design and therapeutic treatments. PLTs form a threaded topology where the backbone is threaded through a covalent-loop formed by either a disulfide bond or an amide bond [9,25,32,33]. Amide-linked PLTs are found in 50 proteins and classified as lasso peptides [9,32,33]. These peptides are a class of natural products synthesized by bacteria. Due to their unusually high stability, lasso peptides have been used for their therapeutic function of antimicrobial activity [9,34]. This review focuses on disulfide-
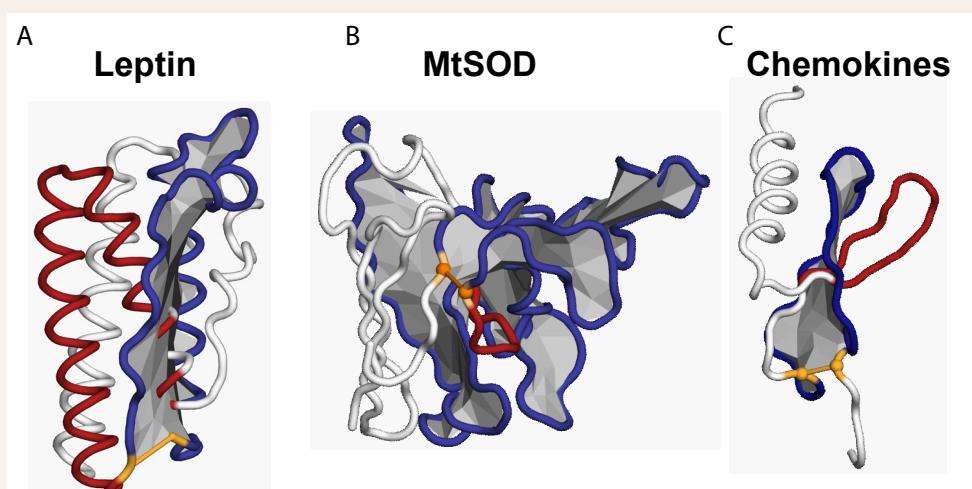

Figure 1 A cartoon representation of three PLT proteins. The structures are color coded with the covalent loop in deep blue, the threaded element in red and the N-terminus in white. The disulfide is highlighted in yellow, and the plane of the covalent-loop is gray for visualization. A) The satiety hormone leptin has an L1-topology, where the backbone threads through the covalent loop one time (PDB ID: 1AX8). B) Superoxide dismutase (SOD) from Mycobacterium tuberculosis (MtSOD) has an L1-topol (PDB ID: 1PZS). C) Chemokines have an L2-topology, where the backbone threads through the covalent loop to loop back and thread it again, thus threading the covalent-loop twice (PDB ID: 2MGS). discoveries in protein folding. The timeline describes the major contributions to the field of protein folding. The figure was created using the PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, and the plug in PyLasso (50).
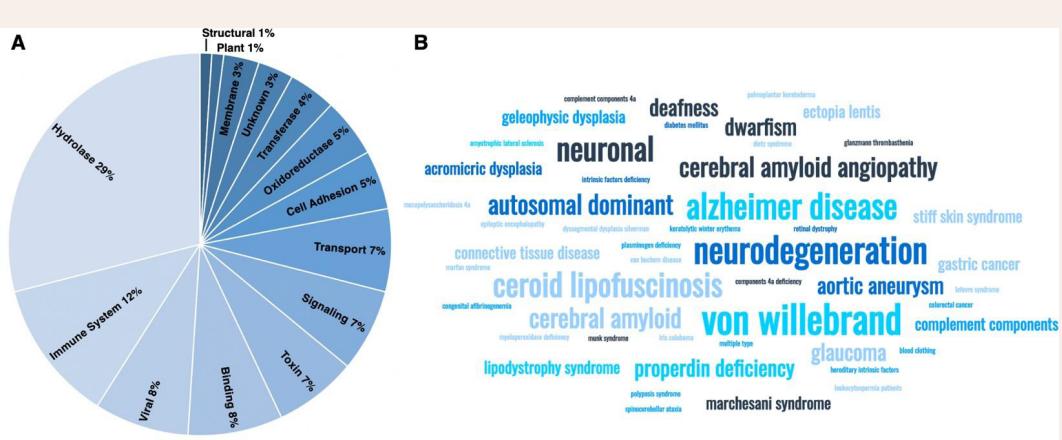
Figure 2 PLT classification. PLTs are represented in 14 different biological functions important in human health. They can be classified into 14 different classes according to the PDB (51-53). 10 out of these 14 classes contain human proteins. Additionally, viral proteins represent 8% of PLTs. Additionally, many PLTs are found in bacterial-, toxin-, and viral proteins that may invade a foreign host, important in human health. B) There are more than 200 human PLTs in nature. Misfolding and or aggregation of these proteins may lead to disease. The word cloud plot is created using the classification of disease of human PLTs from Uniprot Pathology classifications (51-53).
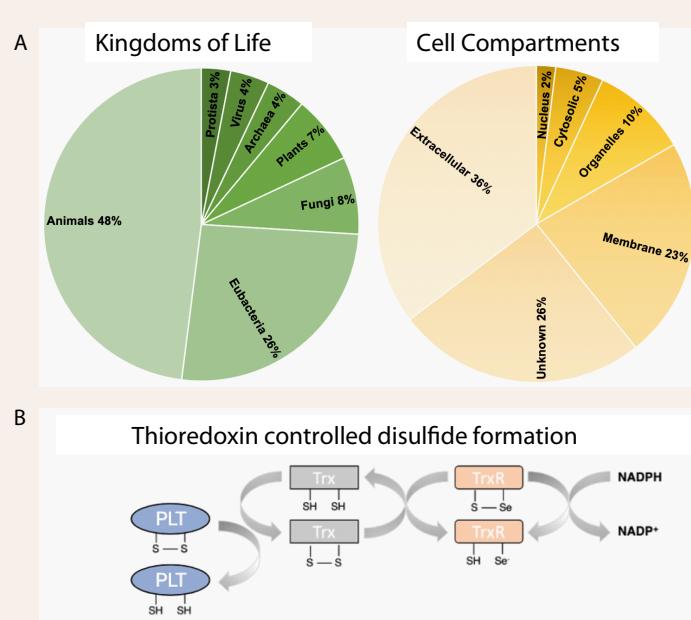
Figure 3 Ts in nature. Proteins with a pierced lasso topology are found in all kingdoms of life in various cell compartments. A) They exist in different cell compartments predominantly in the extracellular matrix and the membrane with an oxidizing environment. However, PLTs exist in the reducing environment of the cytosol where the disulfide bond is promoted in the presence of thioredoxin proteins (TXN) or the glutathione (GSH) (38). B) The redox relay is affected by the "health-state" driven by molecular oxygen.

linked PLTs, the advantage of the threaded topology that can be controlled by the redox potential *in vivo*, and their role in drug design and therapeutics.The pierced lasso topology. There are over 600 proteins with a disulfide linked PLT [24,25,29,30]. These biologically diverse proteins have different secondary structures representing a variety of protein motifs from β - barrel proteins to four-helix bundles and mixed α/β protein motifs [25]. The number of loop crossings varies from one threaded element to closed-loops with up to six crossings [30,31]. The

complexity of the threaded topology may play a role in protein degradation. Proteasomal degradation occurs in the cytosol [25], where proteins in the endoplasmic reticulum (ER) have to pass through a narrow pore to translocate into the cytosol. Thus, the threaded topology may physically protect them from re-entering the cytosol to undergo a proteolytic cleavage. This indicates that degradation and conformational dynamics play an essential role in survival and invasion of host cells. This is seen for example in severe acute respiratory syndrome virus spike protein, where

Table 1: Human proteins are represented in 10 out of the different 14 classes of PLTs.

Classification[1]	% Human proteins[2]
Immune system proteins	69
Signaling proteins	55
Structural proteins	40
Cell adhesion proteins	38
Transport proteins	27
Binding proteins	25
Hydrolases	19
Toxins	18
Membrane proteins	13
Oxidoreductases	10
Transferases	-
Viral proteins[3]	-
Plant proteins	-
Unknown function	-

[1]The biological function is taken from the Protein Data Bank (PDB) function [50 52].
[2]Percentage calculated for each individual protein classification [25].
[3]Many of the viral proteins, like for example the SARS coronavirus protein may attack humans.

the threaded topology may facilitate the invasion of the host cell by interacting with specific receptors on the cell surface to deliver the nucleic acid [35,36] or for superoxide dismutase in Mycobacterium tuberculosis, where SOD can convert reactive oxygen species from the host to non-toxic spices and there for survive and invade the host [37]. Many PLTs are also found within the cytokine protein family, i.e., in interleukins and chemokines, important in inflammation and oxidative stress. These are examples of PLTs that may be modulated as a molecular switch by the chemical environment.

Redox Chemistry in human health: The process of disulfide bond formation *in vivo* is often catalyzed by enzymes, such as protein disulfide isomerases (PDIs) or oxidoreductases, which facilitate the oxidation of cysteine residues. Various biochemical processes occur within cells and tissues involving both oxidizing and reducing environments. These processes are essential for maintaining the body's overall function, including energy production, detoxification, and cell signaling. Different cell compartments provide different chemical environments where the mitochondria, extracellular matrix and the cell membrane provide an oxidizing environment promoting the formation of disulfide bonds. The immune cells, such as macrophages and neutrophils, the liver, and the cytosol of the cell provides a reducing environment obstructing the formation of disulfide bonds. This involves the production of reactive oxygen species (ROS) to destroy bacteria, viruses, and other harmful substances. The redox potential in these environments are controlled by thioredoxin proteins (TXN) and the tripeptide glutathione (GSH) [38]. TXN and GSH can promote/obstruct the formation of disulfide bonds though a redox relay controlled by molecular oxygen as the electron acceptor (Figure 3B). In an unhealthy condition, such as inflammation and oxidative stress a dysregulation of ROS may occur. TXN and GSH and SOD proteins can catalyze and neutralize the excess ROS and maintain cellular integrity. The human body maintains a delicate balance between oxidizing and reducing environments to ensure the proper functioning of various cellular processes. Imbalances, such as excessive oxidative stress, hypoxia and hyperoxia, can lead to cellular damage associated with cancer, neurodegenerative disorders, and cardiovascular diseases. The body's ability to regulate these redox processes is crucial for maintaining health and homeostasis. In a cancer induced hypoxia state, the access of molecular oxygen is limited obscuring the redox really of TXN and GSH. Furthermore, certain bacteria of the gut are able to produce GSH which is secreted to the extracellular matrix. Thus, affecting the redox really of extracellular proteins of the host [39]. The modifications of disulfide bonds are affected by various health states. This can be utilized in protein-based therapeutics using proteins with a PLT, where the biological activity is controlled by the threaded topology maintained by a single disulfide bond. PLTs may provide a platform for protein-based therapeutics as they may act as molecular switches. However, further work is needed in this area to establish the biological implications of PLT proteins in biology.

CONCLUSION

Protein-based drug development and replacement therapy are highly successful in cancer therapy, immune disorders, autoimmune disorders, and rare genetic disorders [40,41]. Protein engineering plays a crucial role in overcoming challenges due to proteins' larger size, protection against proteolytic cleavage, solubility at neutral pH, and protein stability. Topologically complex natural proteins such as circular proteins, cystine knots [42], and lasso peptides [43] have shown to be successful in the treatment of cancer, asthma, and other anti-inflammatory disorders [44-49]. The advantage of these therapeutic proteins is their small size, unusually high stability, and easy to genetically modify. Less is known about the possible role of Pierced Lasso Topologies (PLTs) as a protein therapeutics. Although PLTs necessarily do not increase the stability and may be larger in size, they may play an important role in drug design important in human health as PLTs are involved in human disorders such as obesity, cancer, and neurodegenerative disorders.

Protein based therapeutics have an enormous untapped market potential as a drug target and/or drug delivery, and it is reasonable to anticipate that disulfide-linked PLT proteins will be more extensively engineered in the future. The advantage of a PLT is that they may act as molecular switches initiated by changes in redox potential making/breaking the threaded topology controlling the biological activity. We propose that the topological twist introduced by the threaded may be used to enhance/suppress biological activity due to changes in conformational dynamics. Furthermore, the threaded topology may block access to break the disulfide bond or inhibit degradation by proteases; supported by the presence of PLTs in bacterial-, toxin-, and viral proteins that must survive in a foreign host. We propose that PLTs should be considered in future drug design for their ability to fine-tune the activity of a protein. Designed PLTs may create new functions within largely the same 3-dimensional structure. Depending on the redox potential, the activity of target PLT proteins may be turned on and off, suggesting that PLT proteins may act as a molecular switch *in vivo*.

⊘SciMedCentral

REFERENCES

1. Goeddel DV, Kleid DG, Bolivar F, Heyneker HL, Yansura DG, Crea R, et al. Expression in Escherichia coli of chemically synthesized genes for human insulin. Proc Natl Acad Sci U S A. 1979; 76: 106-10.

2. Blüher S, Shah S, Mantzoros CS. Leptin deficiency: clinical implications and opportunities for therapeutic interventions. J Investig Med. 2009; 57: 784-8.

3. Montague CT, Farooqi IS, Whitehead JP, Soos MA, Rau H, Wareham NJ, et al. Congenital leptin deficiency is associated with severe early-onset obesity in humans. Nature. 1997; 387: 903-8.

4. Ekins S. Industrializing rare disease therapy discovery and development. Nat Biotechnol. 2017; 35: 117-8.

5. Dimitrov DS. Therapeutic proteins. Methods Mol Biol. 2012; 899: 1-26.

6. Santos-Mendoza T. The Envelope(E) Protein of SARS-CoV-2 as a Pharmacological Target. Viruses. 2023; 15: 1000.

7. Xu L, Zhang W-B. Topology: a unique dimension in protein engineering. Science China Chemistry. 2018; 3–16.

8. Craik DJ, Daly NL, Waine C. The cystine knot motif in toxins and implications for drug design. Toxicon. 2001; 39: 43-60.

9. Hegemann JD, Zimmermann M, Xie X, Marahiel MA. Lasso peptides: an intriguing class of bacterial natural products. Acc Chem Res. 2015; 48: 1909-19.

10. 10.Da XD, Zhang WB. Active Template Synthesis of Protein Heterocatenanes. Angew Chem Int Ed Engl. 2019; 58: 11097-104.

11. Craik DJ, Fairlie DP, Liras S, Price D. The future of peptide-based drugs. Chem Biol Drug Des. 2013; 81: 136-47.

12. Craik DJ, Du J. Cyclotides as drug design scaffolds. Curr Opin Chem Biol. 2017; 38: 8-16.

13. Wang CK, Craik DJ. Designing macrocyclic disulfide-rich peptides for biotechnological applications. Nat Chem Biol. 2018; 14: 417-27.

14. 14.Virnau P, Mirny LA, Kardar M. Intricate knots in proteins: Function and evolution. PLoS Comput Biol. 2006; 2: e122.

15. Sułkowska JI, Rawdon EJ, Millett KC, Onuchic JN, Stasiak A. Conservation of complex knotting and slipknotting patterns in proteins. Proc Natl Acad Sci U S A. 2012; 109: E1715-23.

16. Soler MA, Faísca PF. Effects of knots on protein folding properties. PLoS One. 2013; 8: e74755.

17. Nureki O, Shirouzu M, Hashimoto K, Ishitani R, Terada T, Tamakoshi M, et al. An enzyme with a deep trefoil knot for the active-site architecture. Acta Crystallogr D Biol Crystallogr. 2002; 58: 1129-37.

18. Nureki O, Watanabe K, Fukai S, Ishii R, Endo Y, Hori H, et al. Deep knot structure for construction of active site and cofactor binding site of tRNA modification enzyme. Structure. 2004; 12: 593-602.

19. Jacobs SA, Harp JM, Devarakonda S, Kim Y, Rastinejad F, Khorasanizadeh S. The active site of the SET domain is constructed on a knot. Nat Struct Biol. 2002; 9: 833-8.

20. King NP, Yeates EO, Yeates TO. Identification of rare slipknots in proteins and their implications for stability and folding. J Mol Biol. 2007; 373: 153-66.

21. Sułkowska JI, Sulkowski P, Szymczak P, Cieplak M. Stabilizing effect of knots on proteins. Proc Natl Acad Sci U S A. 2008; 105: 19714-9.

22. Alam MT, Yamada T, Carlsson U, Ikai A. The importance of being knotted: effects of the C-terminal knot structure on enzymatic and mechanical properties of bovine carbonic anhydrase II. FEBS Lett. 2002; 519: 35-40.

23. Bornschlögl T, Anstrom DM, Mey E, Dzubiella J, Rief M, Forest KT. Tightening the knot in phytochrome by single-molecule atomic force microscopy. Biophys J. 2009; 96: 1508-14.

24. Haglund E. Engineering covalent loops in proteins can serve as an on/off switch to regulate threaded topologies. J Phys Condens Matter. 2015; 27: 354107.

25. Simien JM, Haglund E. Topological Twists in Nature. Trends Biochem Sci. 2021; 46: 461-71.

26. Zhang H, Chen S. Cyclic peptide drugs approved in the last two decades (2001-2021). RSC Chem Biol. 2022; 3: 18-31.

27. Santini BL, Zacharias M. Rapid Rational Design of Cyclic Peptides Mimicking Protein-Protein Interfaces. Methods Mol Biol. 2022; 2405: 231-44.

28. Haglund E, Sulkowska JI, He Z, Feng GS, Jennings PA, Onuchic JN. The unique cysteine knot regulates the pleotropic hormone leptin. PLoS One. 2012; 7: e45654.

29. Haglund E, Sulkowska JI, Noel JK, Lammert H, Onuchic JN, Jennings PA. Pierced Lasso Bundles are a new class of knot-like motifs. PLoS Comput Biol. 2014; 10: e1003613.

30. Niemyska W, Dabrowski-Tumanski P, Kadlof M, Haglund E, Sulkowski P, Sulkowska JI. Complex lasso: new entangled motifs in proteins. Sci Rep. 2016; 6: 36895.

31. Dabrowski-Tumanski P, Niemyska W, Pasznik P, Sulkowska JI. LassoProt: server to analyze biopolymers with lassos. Nucleic Acids Res. 2016; 44: W383-9.

32. Duquesne S, Destoumieux-Garzón D, Zirah S, Goulard C, Peduzzi J, Rebuffat S. Two enzymes catalyze the maturation of a lasso peptide in Escherichia coli. Chem Biol. 2007; 14: 793-803.

33. Cheung-Lee WL, Link AJ. Genome mining for lasso peptides: past, present, and future. J Ind Microbiol Biotechnol. 2019; 46: 1371-1379.

34. Weber W, Fischli W, Hochuli E, Kupfer E, Weibel EK. Anantin--a peptide antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation and biological activity. J Antibiot (Tokyo). 1991; 44: 164-71.

35. Belouzard S, Millet JK, Licitra BN, Whittaker GR. Mechanisms of coronavirus cell entry mediated by the viral spike protein. Viruses. 2012; 4: 1011-33.

36. Hwang WC, Lin Y, Santelli E, Sui J, Jaroszewski L, Stec B, et al. Structural basis of neutralization by a human anti-severe acute respiratory syndrome spike protein antibody, 80R. J Biol Chem. 2006; 281(45): 34610- 6.

37. Spagnolo L, Törö I, D'Orazio M, O'Neill P, Pedersen JZ, Carugo O, et al. Unique features of the sodC-encoded superoxide dismutase from Mycobacterium tuberculosis, a fully functional copper-containing enzyme lacking zinc in the active site. J Biol Chem. 2004; 279: 33447-55.

38. Wu G, Fang YZ, Yang S, Lupton JR, Turner ND. Glutathione metabolism and its implications for health. J Nutr. 2004; 134: 489-92.

39. Mardinoglu A, Shoaie S, Bergentall M, Ghaffari P, Zhang C, Larsson E, et al. The gut microbiota modulates host amino acid and glutathione metabolism in mice. Mol Syst Biol. 2015; 11: 834.

40. Lagasse HA, Alexaki A, Simhadri VL, Katagiri NH, Jankowski W, Sauna ZE, et al. Recent advances in (therapeutic protein) drug development. F1000Res. 2017; 6: 113.

41. Ebrahimi SB, Samanta D. Engineering protein-based therapeutics through structural and chemical design. Nat Commun. 2023; 14: 2411.

42. Craik DJ, Daly NL, Bond T, Waine C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. J Mol Biol. 1999; 294: 1327-36.

43. Wyss DF, Lahm HW, Manneberg M, Labhardt AM. Anantin--a peptide antagonist of the atrial natriuretic factor (ANF). II. Determination of the primary sequence by NMR on the basis of proton assignments. J Antibiot (Tokyo). 1991; 44: 172-80.

44. Depuydt AS, Rihon J, Cheneval O, Vanmeert M, Schroeder CI, Craik DJ, et al. Cyclic Peptides as T-Type Calcium Channel Blockers: Characterization and Molecular Mapping of the Binding Site. ACS Pharmacol Transl Sci. 2021; 4: 1379-1389.

45. Hellinger R, Muratspahic E, Devi S, Koehbach J, Vasileva M, Harvey PJ, et al. Importance of the Cyclic Cystine Knot Structural Motif for Immunosuppressive Effects of Cyclotides. ACS Chem Biol. 2021; 16: 2373-86.

46. Jackson MA, Xie J, Nguyen LTT, Wang X, Yap K, Harvey PJ, et al. Plant-based production of an orally active cyclotide for the treatment of multiple sclerosis. Transgenic Res. 2023; 32: 121-33.

47. Liu X, Henriques ST, Craik DJ, Chan LY. Unlocking the Potential of the Antimicrobial Peptide Gomesin: From Discovery and Structure-Activity Relationships to Therapeutic Applications. Int J Mol Sci. 2023; 24: 5893

48. Muratspahic E, Koehbach J, Gruber CW, Craik DJ. Harnessing cyclotides to design and develop novel peptide GPCR ligands. RSC Chem Biol. 2020; 1: 177-91.

49. Tyler TJ, Durek T, Craik DJ. Native and Engineered Cyclic Disulfide-Rich Peptides as Drug Leads. Molecules. 2023; 28: 3189.

50. Berman H, Henrick K, Nakamura H. Announcing the worldwide Protein Data Bank. Nat Struct Biol. 2003; 10: 980.

51. Berman H, Henrick K, Nakamura H, Markley JL. The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. Nucleic Acids Res. 2007; 35: D301-3.

52. Consortium U. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019; 47: D506- D515.