units is a prerequisite for initiating infection (*8*). H5N1 viruses that are highly pathogenic in mice contain a stretch of basic residues adjacent to the hemagglutinin cleavage site, enabling these hemagglutinins to be cleaved by ubiquitous intracellular proteases, including furin. Recombinant H5N1 viruses lacking these basic amino acids are no longer virulent in mice (*9*), demonstrating that the presence of these amino acids, and the consequent cleavage by intracellular proteases, are required for the virulence of these viruses.

To further understand the molecular basis of virulence, Obenauer *et al*. first sequenced the genes of a large number of H5N1 viruses isolated from wild birds and poultry, providing an invaluable resource for many investigators. This analysis revealed not only the expected variability in the sequences of the two major surface proteins of the virus, hemagglutinin and neuraminidase, but also variability in the sequence of the NS1 protein. Despite variability in the latter, it was noted that the carboxyl terminus of the NS1 proteins of the vast majority of avian H5N1 viruses contains a sequence motif, Glu-Ser-Glu-Val (ESEV). These residues are predicted to mediate binding to proteins bearing a region called a PDZ domain. The multitude of human proteins that contain a PDZ domain function in diverse cellular signaling pathways including those that regulate protein traffic within the cell and those that maintain cell morphology and organization. Another PDZ-binding sequence, Glu-Pro-Glu-Val (EPEV), was identified at the carboxyl terminus of the NS1 proteins of all the virulent H5N1 viruses isolated from humans. In contrast, the carboxyl terminus of the NS1 proteins of low-virulence human influenza A usually contains a different sequence, Arg-Ser-Lys-Val (RSKV), which is not a PDZ-binding motif. Further, Obenauer *et al.* verified that the carboxyl-terminal ESEV and EPEV sequences indeed bind to PDZ domains. Consequently, the presence of a functional carboxyl-terminal PDZ-binding domain in the NS1 protein of H5N1 viruses correlates with human virulence. This supports the authors' hypothesis that the carboxyl-terminal domain of the NS1 proteins of avian H5N1 viruses acts as a virulence factor by binding cellular PDZ-containing proteins and disrupting their participation in important cellular processes.

This is an intriguing hypothesis that, however, needs to be evaluated in animal experiments with H5N1 viruses that have been altered to express a NS1 protein lacking the carboxyl-terminal ESEV/EPEV sequence. Such experiments are critical because it has already been established that this carboxyl-terminal sequence is not required for the virulence of previously isolated H5N1 viruses in ferrets (*11*). An analysis of the virulence of H5N1 viruses isolated in 2004 identified the human isolate A/Vietnam/1203/04 as the most pathogenic isolate. The NS1 protein encoded by this virus is truncated and consequently lacks the suspect carboxyl-terminal ESEV/ EPEV motif. Future experiments will establish whether eliminating the carboxyl-terminal ESEV/EPEV sequence of the NS1 protein of other H5N1 viruses has any effect on their virulence in animal models. In addition, the search for other molecular determinants of the virulence of H5N1 viruses in humans will undoubtedly continue.

References

1. T. Horimoto, Y. Kawaoka, *Nat. Rev. Microbiol.* **3**, 591 (2005).
2. D. L. Noah, R. M. Krug, in *Advances in Virus Research*, K. Maramorsch, A. J. Shatkin, Eds. (Elsevier, Amsterdam, 2005), vol. 65, pp. 121–145.
3. J. C. Obenauer *et al*., *Science* **311**, 1576 (2006); published onlne 26 January 2006 (10.1126/science.1121586).
4. C. Y. Chien *et al.*, *Biochemistry* **43**, 1950 (2004).
5. A. Garcia-Sastre, *Virology* **279**, 375 (2001).
6. R. M. Krug, W. Yuan, D. L. Noah, A. G. Latham, *Virology* **309**, 181 (2003).
7. D. L. Noah, K. Y. Twu, R. M. Krug, *Virology* **307**, 386 (2003).
8. R. A. Lamb, R. M. Krug, in *Fields Virology*, D. M. Knipe, P. M. Howley, Eds. (Lippincott, Williams, and Wilkins, Philadephia, 2001), vol. 1, pp. 1487–1531.
9. M. Hatta, P. Gao, P. Halfmann, Y. Kawaoka, *Science* **293**, 1840 (2001).
10. K. Shinya *et al.*, *Virology* **320**, 258 (2004).
11. E. A. Govorkova *et al.*, *J. Virol.* **79**, 2191 (2005).

10.1126/science.1125998

CHEMISTRY

Seamless Proteins Tie Up Their Loose Ends

Natural circular proteins found in bacteria, plants, and mammals show antimicrobial activity and exceptional stability, making them ideal templates for engineering better drugs. But just how they close into loops remains a mystery.

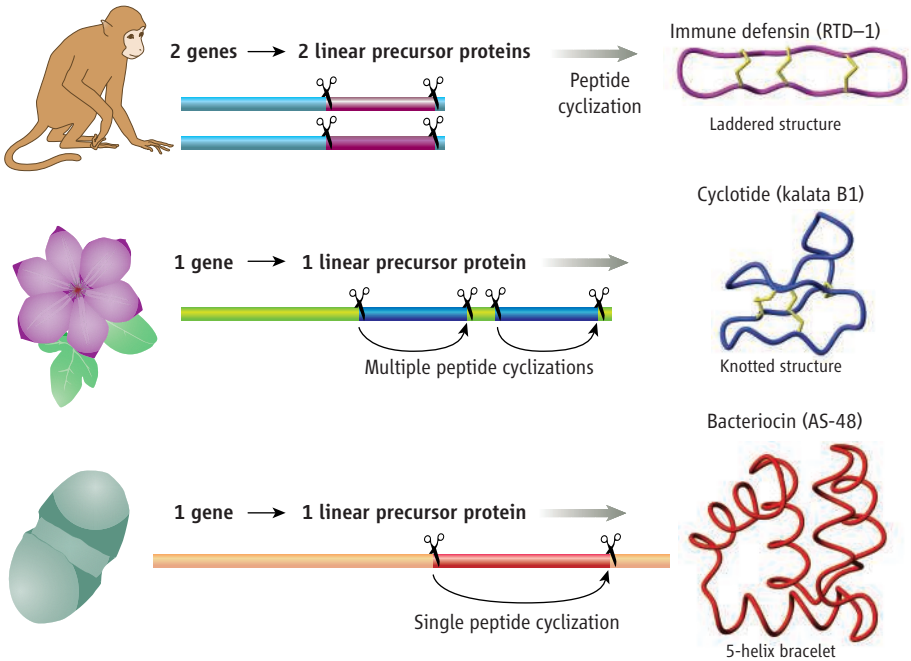
David J. Craik

In the early 1970s, tribeswomen in the Congo were reported to drink a medicinal tea made from a local plant to induce labor and facilitate childbirth (*1*). Twenty-five years later, it was discovered that the active ingredient, robust enough to withstand boiling and ingestion, is a small protein with a circular shape (*2*). It turns out that the protein, kalata B1, was not a one-off example. Many other naturally occurring circular and stable proteins have since been found in bacteria, plants, and animals from Africa, South America, Australia, and Europe (*3*). What makes them so interesting? The exceptional stability and wide range of activities of these circular proteins, from insecticidal and antimicrobial to thwarting cellular infection by HIV (*4*), may guide the development of more effective and stable drugs.

The discovery of proteins bearing two ends that are linked together, producing a circular topology (*2*), is a new and mysterious twist in protein synthesis. Most proteins are synthesized as linear chains of amino acids in which the amino terminus of one residue is linked to the carboxyl terminus of the next. Whether assembled in the cell by nature's ribosomal machinery that translates a genomic blueprint, or by the synthetic methodology of peptide chemists, a newly formed chain folds into a complex three-dimensional shape that determines the protein's function. Circular proteins have no beginning or end, and deciphering their mode of construction presents some interesting challenges. So far, we know very little about how cyclization occurs. Circular proteins appear to derive from larger precursor proteins (see the figure), but we have little knowledge of the enzymes or other processes that cleave the mature peptide from its precursor and facilitate formation of a cyclic backbone.

The diversity of sequence of the nearly 100 circular proteins known to date across species suggests that cyclization has evolved independently in vastly different organisms as a way of tidying up the loose ends of conventional proteins. Microorganisms appear to have seized upon the advantages of cyclizing peptides long ago, as has the pharmaceutical industry. For example, in the course of making the cyclic peptide cyclosporin for their own defense, fungi have provided humankind with a drug that has now revolutionized transplant therapy because of its potent immunosuppressive activity. But cyclosporin and other previously known cyclic peptides are typically small rings of fewer than a dozen amino acids and are produced not by direct gene translation but by multidomain enzymes called peptide synthetases. The excite-

The author is at the Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia. E-mail: d.craik@imb.uq.edu.au



Diced, spliced, and coming full circle. Gene-encoded circular proteins are produced as fragments of linear precursor proteins that are excised and spliced head-to-tail. In the case of rhesus θ-defensin–1 (RTD-1) (**top**), two genes each code for half of the 18–amino acid mature peptide and a double head-to-tail ligation produces the circular peptide. In the case of the plant cyclotides (**middle**), a cystine knot embedded in the circular backbone provides extra stabilization. (**Bottom**) The circular backbone of the bacterial protein AS-48 folds up to form a bracelet of five helices. Images of structures are adapted from (*3*).

ment associated with the new generation of circular proteins discovered in the last decade, and ranging in size from 14 (*5*) to 78 (*6*) amino acids, is that they are true gene products and hence can be manipulated using the tools of molecular biology. For example, genes from circular proteins that have insecticidal properties (*7*) could be transferred to crop plants to provide built-in protection against herbivorous pests, and thereby reduce the need for chemical spraying.

What are the advantages of a circular form? For one, the free ends of conventional proteins are routinely targeted by exopeptidases—enzymes whose function is to nip away at proteins to digest them. Joining the ends thus removes a major degradation pathway. Also, the ends of linear proteins are often flexible or ill-defined, in contrast to their highly structured interior. This flexibility is bad from an entropic perspective when proteins bind to their molecular receptors, leading to reduced binding affinity and biological activity. Thus, in principle, both the stability and the activity of proteins can be improved by tying up their loose ends. What is particularly impressive about circular proteins is their indestructible nature. Most proteins denature irreversibly upon heating, as exemplified by the familiar transformation when an egg is cooked. But circular proteins can be subjected to boiling, extremes of pH, and proteolytic enzymes yet still retain their structure and function—a tough crowd.

Some of the secrets to their stability have been revealed in the details of their structures. Structural determination of kalata B1 by nuclear magnetic resonance spectroscopy (*2*) revealed two surprises: Not only does it have a seamless circular backbone, but it also has a knotted arrangement of disulfide bonds that contribute to its exceptional stability (see the figure). The name "cyclotide" was coined for this family of plant proteins, which is now estimated to comprise thousands of members (*8*). The exceptional stability of the cyclotide framework suggests the possibility of using it as a template in drug design (*9*). The aim here would be to "graft" bioactive peptide sequences into the cyclotide framework. Chemical methods for the synthesis of cyclotides have been developed, so the approach is feasible. The main challenge in such studies is to ensure that the foreign peptide sequence can be grafted into the framework in such a way that it retains its biological activity

The genes for bacterial and plant circular proteins encode linear precursor proteins from which the mature peptides are excised and cylized (see the figure). The first cyclic peptide discovered in mammals, an antibacterial called rhesus θ-defensin–1 found in macaques, is in fact a product of not one but two genes, each coding for short peptides that are subsequently linked in a double head-to-tail ligation (*10*). Rhesus θ-defensin–1 is expressed in white blood cells of the macaque monkey and is part of its innate immune system. Like the cyclotides, it contains three cross-bracing disulfide bonds, but they are in a "laddered" arrangement rather than knotted. Why would organisms go to the trouble of producing cyclic peptides, and in different conformations? Again, stability and enhanced activity appear to be the answer, as cyclic rhesus θ-defensin–1 is more potent and stable than a synthetic acyclic counterpart that is active in vitro but is essentially inactive at physiological salt concentrations. The remarkable range of conformations into which the circular proteins are folded—from a ladder, to a knot, to a helix bundle—highlights the fact that circular proteins, just like conventional proteins, need to adopt diverse shapes specific to their functions.

In contrast to bacteria, plants, and some of our primate cousins, humans do not make cyclic peptides. A sequence similar to rhesus θ-defensin–1 was recently discovered in the human genome, but the gene is silenced by a premature stop codon (*11*). Not put off by this genetic impediment, Lehrer and colleagues (*11*) chemically synthesized retrocyclin, a putative defensin-like molecule, and found it to be a potent anti-HIV agent. The same group then analyzed DNA from a range of primates and showed that the stop codon emerged in the human lineage about 7 million to 10 million years ago (*12*). It is an ironic twist of fate that our evolutionary forebears acquired a mutation whose nonappearance would have left us with built-in protection against HIV.

The discovery of naturally occurring circular proteins has offered inspiration to protein engineers, as demonstrated by recent successes in the artificial cyclization of conotoxins, marine venom peptides of approximately 12 to 30 amino acids (*13*). Cyclization of a prototypic conotoxin improved its resistance to proteolytic degradation, which opens the door to enhanced applications of this class of molecules in medicine. The biggest challenge in the field of circular proteins is deciphering just how their ends are stitched together from their linear precursors: What enzymes are involved? Do cleavage and cyclization occur simultaneously? Are auxiliary proteins involved? These unanswered questions will certainly continue to drive the field forward.

References

1. L. Gran, *Acta Pharmacol. Toxicol.* **33**, 400 (1973).
2. O. Saether *et al.*, *Biochemistry* **34**, 4147 (1995).
3. M. Trabi, D. J. Craik, *Trends Biochem. Sci.* **27**, 132 (2002).
4. K. R. Gustafson, T. C. McKee, H. R. Bokesch, *Curr. Protein Pept. Sci.* **5**, 331 (2004).
5. S. Luckett *et al.*, *J. Mol. Biol.* **290**, 525 (1999).
6. R. Eisenbrandt *et al*, *J. Biol. Chem.* **274**, 22548 (1999).
7. C. Jennings *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10614 (2001).
8. S. M. Simonsen *et al.*, *Plant Cell* **17**, 3176 (2005).
9. D. J. Craik *et al.*, *Curr. Opin. Drug Discov. Devel.* **9**, 251 (2006).
10. Y.-Q. Tang *et al.*, *Science* **286**, 498 (1999).
11. A. M. Cole *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1813 (2002).
12. A. M. Cole *et al.*, *Curr. Protein Pept. Sci.* **5**, 373 (2004).
13. R. J. Clark *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13767 (2005).

10.1126/science.1125248

CREDIT: P. HUEY/*SCIENCE*

Published by AAAS