UBIQUITIN: A single ubiquitin molecule, with thicker tubes representing the more flexible parts of the molecule

On the Chain Gang

More than simply helping haul out a cell's garbage, ubiquitin, with its panoply of chain lengths and shapes, marks and regulates many unrelated cellular processes.

BY KEITH D. WILKINSON AND DAVID FUSHMAN

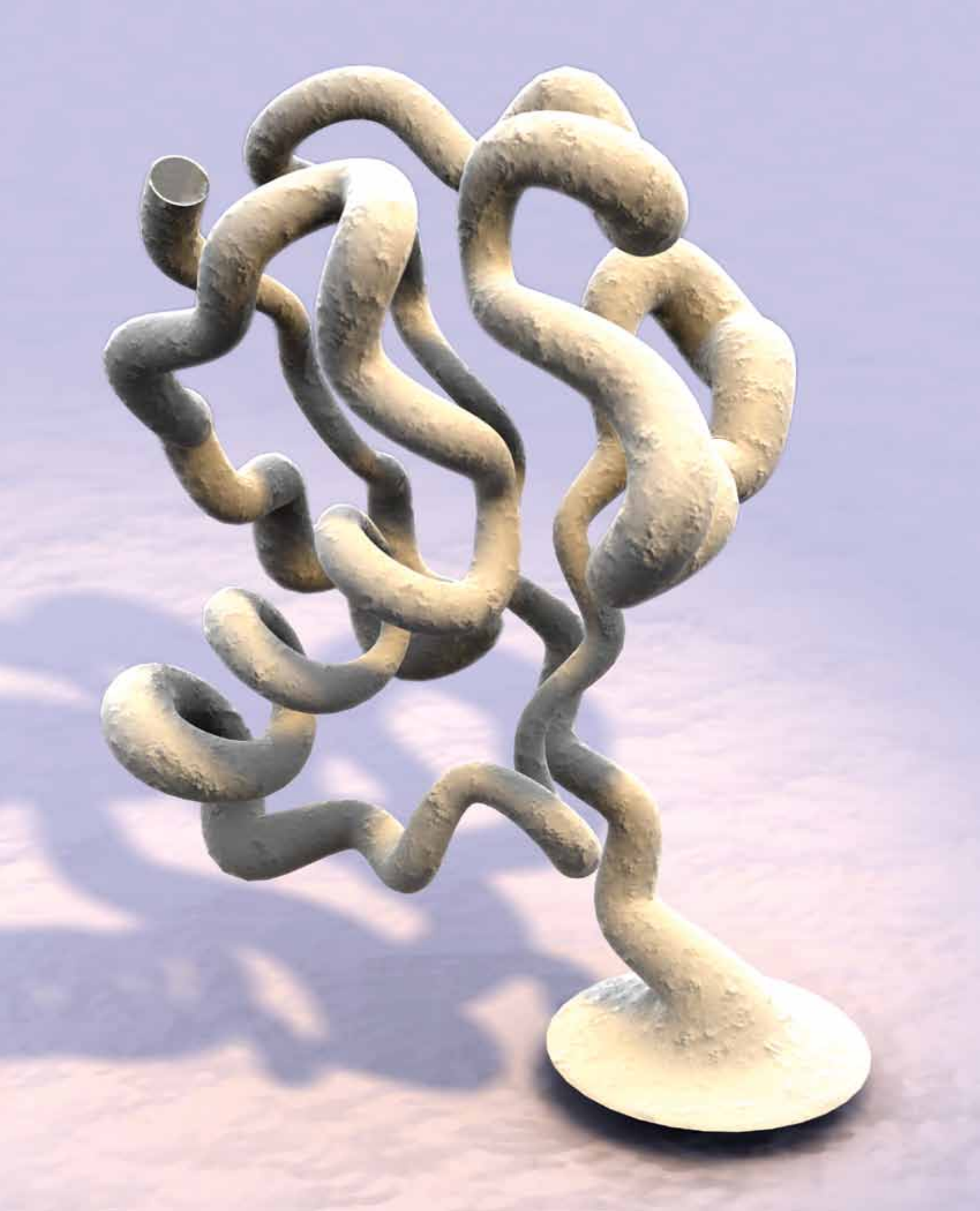
In 1974 and 1975, a group led by Gideon Goldstein at New York University discovered and sequenced a 76-amino-acid protein from bovine thymus cells that appeared to be important in stimulating immune-cell function. But as they continued to characterize the protein, like a bad contaminant, they found it everywhere—in every tissue of the human body, and in cell cultures from worms, other animals, plants, and even bacteria. The authors surmised that the protein must be "a universal constituent of living cells," and consequently named it ubiquitin. Later it became apparent that the ubiquitin found in bacterial cultures came from the yeast extracts on which they were cultured, leading to the realization that ubiquitin was limited to eukaryotic cells. For several years, little more was learned about the protein's structure or function. In fact, a National Institutes of Health panel, reviewing William Cook's proposal to determine the crystal structure of ubiquitin, concluded that the project was not interesting, since the protein was found everywhere and had no known function.

Despite such an inauspicious start, ubiquitin was soon recognized as a constituent of histone proteins (through work by Ira Goldknopf and Harris Busch) and later as a necessary cofactor in a vital cellular process—the degradation of proteins. Work by Avram Hershko, Aaron Ciechanover, and Irwin Rose (for which they received the 2004 Nobel Prize in Chemistry) showed that covalent attachment of a small protein, which turned out to be ubiquitin, provided proteins with a tag or label that directed them to the cell's degradation machinery. While degradation is essential for normal cellular function, such as helping clear damaged proteins, it always seemed as though a protein so well conserved and ever-present must play an even larger role in cell biology.

Although the biochemical studies done by the Nobel Prize winners were strongly suggestive, it was only after Alex Varshavsky began to define the genetics of the ubiquitin system in 1984 that the multifaceted cellular role of the little protein became more obvious. Varshavsky, an eminent histone biochemist who defected to the United States from the former Soviet Union, had become intrigued by this molecule that tagged both histones and damaged cellular proteins. His early genetic studies led to the discovery of a dozen or so ubiquitin-like proteins.

Soon researchers discovered new roles for ubiquitin in addition to protein degradation, and learned that the ubiquitin protein's structure and the architecture of its polymeric forms have more to do with its function than does its mere presence on a protein substrate. In contrast to modifications such as phosphorylation, methylation, and acetylation, the attachment of one or more ubiquitin monomers provides a large interaction surface by which the modification can be encoded, and results in a vast number of poten-

PHANTATOMIX/PHOTO RESEARCHERS, INC.



tial signals by virtue of the varied architectures linking the multiple ubiquitin molecules. Recent studies have begun to define the roles of different polyubiquitin signals in physiology and disease, and it has become obvious that the manipulation of these signals and of their recognition will be important in developing new treatments.[1]

Diversity of polyubiquitin chains and linkages

Polyubiquitin chains (polyUb) consist of ubiquitin (Ub) monomers linked to each other covalently. These ubiquitin chains are attached to protein substrates with the help of several accessory proteins called E1, E2 and E3. These accessory proteins select the appropriate protein and then recruit Ub tagging machinery to build a chain of Ub molecules on the target protein.

It always seemed as though a protein so well conserved and ever-present must play an even larger role in cell biology.

While the attachment of a single molecule of ubiquitin to a protein resembles simple modifications such as phosphorylation or acetylation, the attachment of polyubiquitin chains more closely resembles glycosylation. Both provide a much broader functional canvas because of the immense variation in length and linkage architecture. Monomers of ubiquitin can be built into chains at multiple attachment sites to assemble a huge number of different targeting signals. Importantly, all of these polyUb forms appear to serve diverse functions, not only tagging a protein for transport to a particular location, but also aiding in the assembly of protein complexes that modulate a protein's function or stability.

Proteins tagged for degradation, for example, are recognized and degraded by the proteasome, a large multicatalytic protease that degrades the target protein into small peptides, which can be then broken down into free amino acids, and that also disassembles the polyubiquitin tag. This canonical mechanism is responsible for maintaining the temporal order of the cell cycle, wiping the cell clean of one type of cyclin protein after the next, allowing each subsequent wave of cyclins to push the cell further along the path to cell division. The canonical pathway was also found to be behind the cycling of the circadian clock. A master circadian protein that accumulates throughout the day is completely degraded when its levels reach a threshold, thus resetting the clock for the new cycle. Similarly, this same mechanism rapidly tags and degrades damaged proteins that arise due to aging, stress, or oxidative damage.

In the past 2 decades, researchers have found much variability in both the shape and the function of ubiquitin chains. They've discovered Ub chains of varying length and linkage architecture; chains that include other ubiquitin-like proteins; solitary Ub chains not attached to any proteins; and proteins with multiple Ub chains attached to them. Each new configuration of polyubiquitin hints at new functions, many of which are yet to be discovered, and confirms that these ubiquitous polymers are indeed essential for a wider variety of cellular processes than we had imagined.

Ubiquitin is often attached to proteins as a chain of various lengths, and the number of links appears, in part, to determine the ultimate destination of the protein; four links, for example, are sufficient for delivery to the proteasome. Polyubiquitin linkages can be made between the C-terminal end of one ubiquitin and any one of eight primary amino groups on the next one: an amine from one of the monomer's seven lysines or the N-terminal methionine. Imagine ubiquitin as a dreidel with flat faces and straight edges and a hole in each of the faces. The peg of one dreidel (the C-terminus) can be inserted into any hole of another. If each face is a different color, then every combination has a unique surface architecture. Each successive ubiquitin monomer can attach at a different linkage site, and any one ubiquitin can contain more than one linkage. The variability can run the gamut from linear chains with "homogeneous" linkages (all using the same lysine) to "heterogeneous" linkages (using different lysines) to observe "branched" chains with multiple distal termini. (See illustration on opposite page.) Byzantine indeed!

Chains can also be "mixed," made up not only of ubiquitin monomers, but also of other members of the ubiquitin-like family of proteins that are similar to ubiquitin in shape, but not in sequence. These ubiquitin-like proteins, such as the small ubiquitin-related modifier (SUMO), are themselves attached to proteins to target them to various locations in the cell, regulating such processes as apoptosis or transcriptional control.[2]

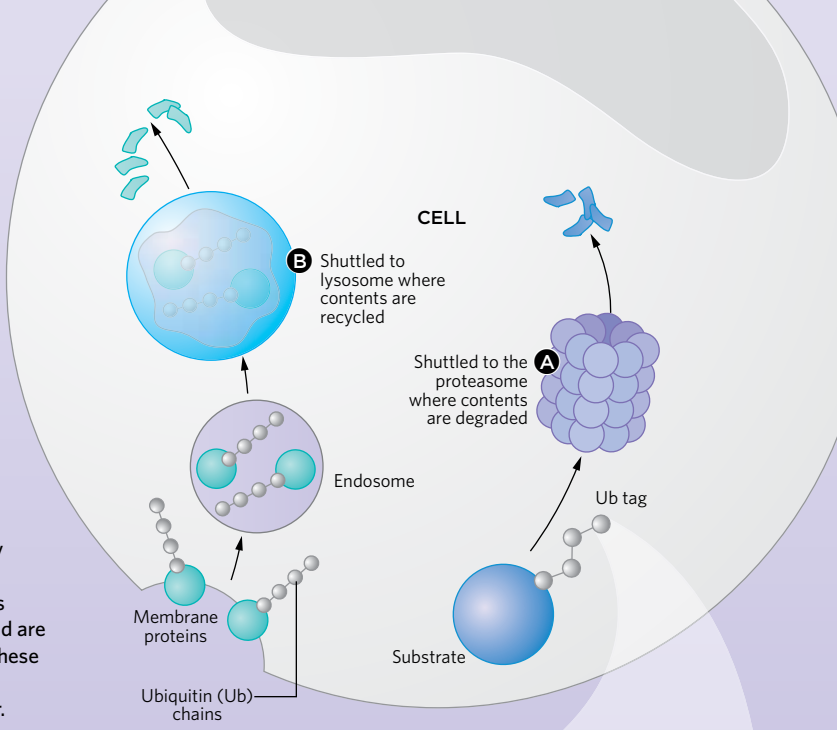
Very recently, researchers have also noticed chains of unanchored ubiquitin in the cell, and have found evidence that these free chains play a role in cell signalling, specifically by activating protein kinases and other pathways involved in antiviral innate immunity, although they may have other functions as well.[3]

In addition to the diversity generated by the manifold linkage patterns, a polyubiquitin chain can also exhibit a unique three-dimensional structure. It can fold back on itself, creating kinks or knobs at various points. These shapes alter how ubiquitin "receptor" proteins bind and thereby read the message. Receptors can be any proteins that specifically recognize and bind to mono- or polyubiquitin. Thus, the chain's structure defines its ability to interact with specific receptors that perform various functions, from shuttling the ubiquitinated protein to a new location to hydrolyzing the ubiquitin chain to degrading the substrate protein.[4,5]

The critical question that researchers in this field are still teasing apart concerns how the three-dimensional structure of polyUb chains specifies its binding partner, i.e. how does the receptor recognize and decode the signal? Two features of the chains appear to be essential in determining binding. First, the hydrophobic patch on one face of each ubiquitin monomer can interact with the most common ubiquitin-binding domains on receptor proteins and with other ubiquitins in the chain. Second, ubiquitin's C-terminus tail, which links monomers in the chain, is highly flexible, making possible a variety of conformations. For example, linking either two or four Ubs together by attaching the C-terminal tail of one monomer to lysine at position 48 on the next monomer creates a chain that is in equilibrium between two or more forms; a tightly-packed, closed conforma-

UBIQUITIN BASICS

Despite its discovery as a protein that seems to show up everywhere, at least in eukaryotic cells, researchers are only beginning to scratch the surface of all of the cellular functions that involve ubiquitin. Ubiquitin can bind to proteins as a monomer, or in long chains that bend or branch. Not much is known, however, about the receptors that decode these various shapes and relay their messages.

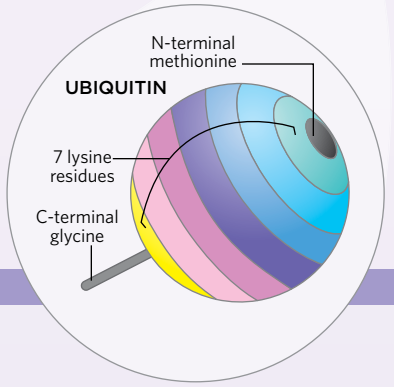


LITTLE MESSENGERS

Ubiquitin (Ub) is best known for tagging proteins for degradation. A four-monomer-long ubiquitin chain connected via the lysine-48 of each Ub is used to mark proteins destined for proteasomal degradation A. Accessory proteins called E1, E2, and E3 help choose which protein should be tagged and recruit ligases to link the Ub monomers into a chain. Ubiquitin also tags membrane proteins, which then pinched inward into endosomal vesicles and are trafficked to the lysosome for digestion B. In addition to degradation via these two pathways, however, ubiquitin appears to also play a role in DNA repair, apoptosis, and the transport of proteins from one part of the cell to another.

DIFFERENT LINKS FOR DIFFERENT KINKS

Ubiquitin owes much of its diversity of function to the simplicity and flexibility of its linkages. Not only can it form long chains or polymers, those polymers can take on many shapes: branching, buckling, and even including other ubiquitin-like molecules. The C-terminus of each ubiquitin molecule can bind to one of the seven exposed lysine residues (Lys) on the neighboring Ub, as well as to its N-terminus. How the molecules link up determines the shape of the chain and the message it conveys to the receptors. (Artistic representation of ubiquitin molecule at right depicts how lysine binding affects chain shape. The C-terminus of Ub will bind one of the lysine residues, which in real life sit in specific positions on the molecule, rather than encircle it.)

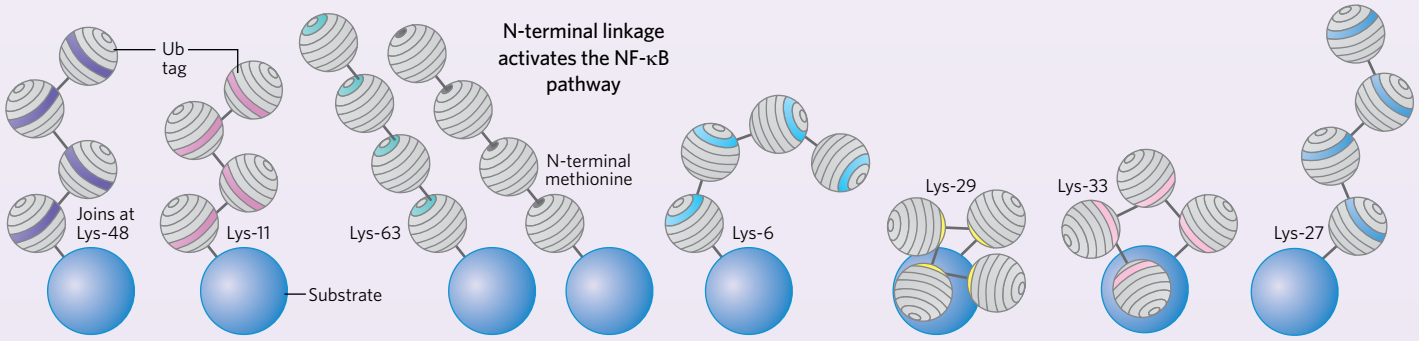


HOMOGENEOUS POLYBUIQUITIN

Linkage via Lys-48 and Lys-11 targets tagged proteins to the proteasome

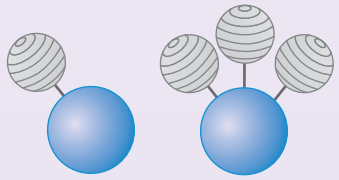
Lys-63–linked chains are involved in NF-κB activation, DNA repair, and targeting proteins to the lysosome

Lys-6, Lys-29, Lys-33 and Lys-27 linkages: function unknown



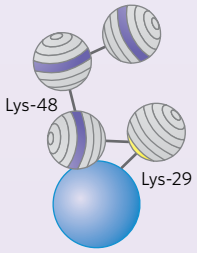
MONOUBIQUITINATION

Proteins tagged with one or many single ubiquitin molecules are involved in protein interactions and localization



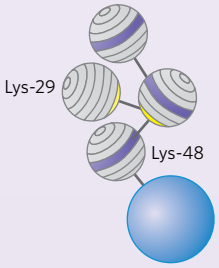
HETEROGENOUS

Unknown



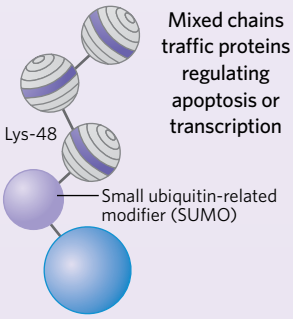
BRANCHED/FORKED

Unknown



MIXED CHAIN

Mixed chains traffic proteins regulating apoptosis or transcription



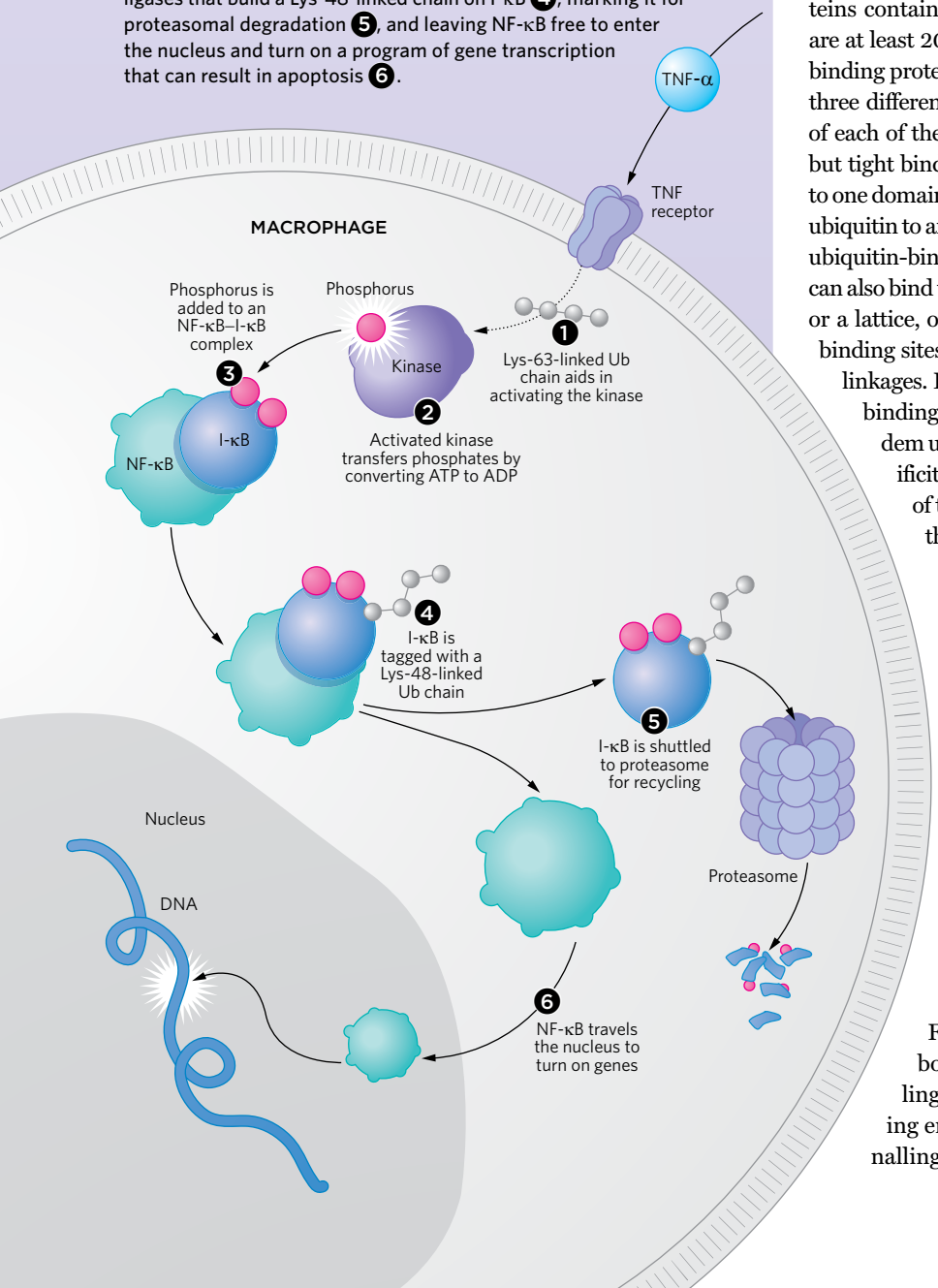
LUCY READING-IKKANDA

UBIQUITIN CHAINS IN ACTION

Present in every tissue of the body, ubiquitin appears to be involved in a dizzying array of functions, from cell cycle and division to organelle and ribosome biogenesis, as well as the response to viral infection. The protein plays at least two roles in turning on signals that destroy virally infected cells from within.

DOUBLE ACTION

TNF-α is a inflammatory cytokine that can induce cell death or apoptosis in a virus-infected cell. When the cytokine's receptor detects TNF-α, it initiates a signaling cascade in which a ubiquitin chain linked via lysine-63, which may either be free floating or bound to a substrate ❶, activates a kinase ❷. This kinase converts ATP to ADP, and phosphorylates I-κB ❸, which keeps NF-κB inactive as long as the two are bound together. The phosphorylation activates ligases that build a Lys-48-linked chain on I-κB ❹, marking it for proteasomal degradation ❺, and leaving NF-κB free to enter the nucleus and turn on a program of gene transcription that can result in apoptosis ❻.



tion that conceals the monomers' hydrophobic patches and one or more open conformations.[6] Chains linked at lysine-63 or the amino terminus, on the other hand, predominantly adopt an extended, open structure that exposes the hydrophobic patches, making them readily available for interactions with receptors. Structural data and computer modelling indicate an even greater structural variability for polyubiquitin chains connected through other lysines.

Polyubiquitin receptors

While the evidence for the existence of ubiquitin receptors is strong, little is known about the molecular details of most. The two exceptions are the enzymes that disassemble polyubiquitin (deubiquitinating enzymes) and shuttling proteins that ferry polyubiquitinated proteins to the proteasome.

Specific recognition of polyubiquitin is accomplished by proteins containing one or more ubiquitin-binding domains. There are at least 20 families of these domains, and many polyubiquitin-binding proteins, or receptors, contain multiple copies, with two or three different domains connected by flexible linkers.[7] The affinity of each of these individual domains for ubiquitin is modest (μM), but tight binding is achieved because the binding of polyubiquitin to one domain lowers the entropic barrier for binding of an adjacent ubiquitin to another domain. Some shuttling receptors have not only ubiquitin-binding domains but also ubiquitin-*like* domains, so they can also bind to each other in networks that assemble into oligomers or a lattice, offering a highly selective array of available ubiquitin binding sites exhibiting specificity for certain polyubiquitin chain linkages. Indeed, studies using artificial oligomers of ubiquitin-binding domains, such as GST-UBA fusions or TUBES (tandem ubiquitin-binding entities), have emphasized that specificity is determined more by the oligomeric arrangement of these domains than by the weak specificity inherent in the individual domains.

Specific recognition of chain linkage

The chain linkage architecture is important for determining the shape of the chain and also appears to inform the fate of the tagged protein by determining which receptors bind the chain. For instance, a linkage at lysine 6, 11, 29, or 48 directs proteins to the proteasome, while linking at lysine-63 or methionine-1 (M1) serves to mark the protein for a role in DNA-damage response or NF-κB–mediated inflammatory pathways.

Finally, receptors can distinguish between polyubiquitin chains bound to different target proteins if the receptors contain a ubiquitin-binding domain as well as a site for binding to the target protein. For instance, the A20 deubiquitinating enzyme binds both ubiquitin and RIP1, a polyubiquitinated signalling protein in the NF-κB pathway. This deubiquitinating enzyme then removes the ubiquitin tag from the signalling protein RIP1, converting RIP1 from a complex that

LUCY READING-IKKANDA

What we know about ubiquitin domain binding specificity and chain architectures suggests that a deeper understanding awaits.

prevents cell death to one that drives it forward, helping destroy virus-infected cells from within. There must be many of these types of receptors that recognize both ubiquitin and the tagged protein, since the cell must distinguish among the numerous proteins that have similar polyubiquitin chains attached.[7]

The importance of length

Ever since Cecile Pickart at Johns Hopkins University initially observed that a four-ubiquitin chain was the minimal effective length to deliver proteins to the proteasome, the question of how chain length affects the fate of a ubiquitinated protein has been debated. Chains must achieve a length that provides sufficient binding affinity for a Ub receptor through binding at multiple sites. However, long chains can change conformation, perhaps folding together so tightly that the dissociation of catalytic intermediates, or "hand off" from one receptor to the next, is prevented.[8]

In part, chain length can be controlled by how long the enzymes, or ligases, that link ubiquitin monomers together can remain on the chain before falling off—a property called processivity. The longer the enzyme and substrate remain associated, the more ubiquitins can be attached. Length can also be affected by deubiquitinating as they trim or disassemble chains. The modular nature of receptors containing multiple binding domains and the ability of longer polyubiquitin chains to bind multiple receptors may serve as length sensors. For instance, the deubiquitinating enzyme USP5 selectively binds a tetra-ubiquitin chain, which it then severs using an ensemble of four ubiquitin binding sites. Longer polyubiquitin chains can also be "handed off" from one receptor to another, as exemplified by the trafficking of ubiquitinated proteins through the endosomal sorting complex required for transport (ESCRT), which delivers ubiquitinated proteins into the cell's vesicles.[8]

Localization of the polyubiquitin signal

Recent observations show that both free-floating and attached methonine-1–linked polyubiquitin chains can activate signaling of the innate immune response mediated by the NF-κB pathway, protecting cells from invading viruses. (See illustration on preceding page.) These chains directly activate kinases that drive the signaling cascade.[9] Unanchored M1-linked ubiquitin chains are also the primary gene product of several genes transcribed in response to genotoxic stress. Normally, however, levels of M1-linked ubiquitin chains in cells are very low, in part because the primary gene product is cleaved to monomeric ubiquitin as it's being transcribed at the ribosome and because of the presence of a large amount of USP5, the enzyme responsible for disassembling polyubiquitin intermediates that might otherwise accumulate in the cell. Thus, it is unlikely that chains with an M1 linkage are widely distributed in the cell. Rather, they may be locally generated at the site, or sites, of signaling. A similar mechanism may be at play in the case of ubiquitinated proteins that accumulate in other signaling cascades. A great deal of cellular specificity in the ubiquitin pathway seems to depend on the use of adaptors and scaffolds that colocalize polyubiquitin and the enzymes that metabolize it. For instance, deubiquitinating enzymes are very often found in the same protein complex as the ubiquitin ligases that synthesize polyubiquitin.[10] This suggests that if a polyubiquitin chain or polyubiquitinated protein is not properly channeled to its target by ubiquitin receptors, it can be disassembled before it leaves the site of synthesis.

Future directions

The incredible diversity of polyubiquitin chains observed in vivo suggests a similar complexity in the receptors that recognize the chains. It seems likely that additional ubiquitin-binding motifs and domains remain to be discovered. More importantly, what we know about ubiquitin domain binding specificity and chain architectures suggests that a deeper understanding awaits studies of binding specificity in the context of the full-length receptors. As more engineered and synthetic polyubiquitins become available, structure determination of polyubiquitin-receptor complexes will be vital to understanding the decoding of the polyubiquitin signals. Finally, we need a more sophisticated understanding of the "hand off" of a receptor-bound polyubiquitin to the next receptor in a sequence. It is still a mystery how shuttling receptors pass ubiquitin from the ligases and chaperone complexes to the proteasome, or how sequential ESCRT complexes can direct endocytic cargos carrying the ubiquitin signal. ■

Keith D. Wilkinson is at the Emory University School of Medicine and David Fushman is at the University of Maryland.

References

1. A. Ciechanover, K. Iwai, "The ubiquitin system: from basic mechanisms to the patient bed," *IUBMB Life*, 56:193-201, 2004.
2. M.C. Geoffroy, R.T. Hay, "An additional role for SUMO in ubiquitin-mediated proteolysis," *Nat Rev Mol Cell Biol*, 10:564-68, 2009.
3. K. Parvatiyar, E.W. Harhaj, "Cell signaling. Anchors away for ubiquitin chains," *Science*, 328:1244-45, 2010.
4. F.E. Reyes-Turcu et al., "Recognition of polyubiquitin isoforms by the multiple ubiquitin binding modules of isopeptidase T," *J Biol Chem*, 283:19581-92, 2008.
5. C.M. Pickart, D. Fushman, "Polyubiquitin chains: polymeric protein signals," *Curr Opin Chem Biol*, 8:610-16, 2004.
6. R. Varadan et al., "Structural properties of polyubiquitin chains in solution," *J Mol Biol*, 324:637-47, 2002.
7. H. Wu et al., "Recent advances in polyubiquitin chain recognition," *F1000 Biol Rep*, 2:1-5, 2010.
8. J.H. Hurley, "The ESCRT complexes," *Crit Rev Biochem Mol Biol*, 45:463-87, 2010.
9. C. Behrends, J.W. Harper, "Constructing and decoding unconventional ubiquitin chains," *Nat Struct Mol Biol*, 18:520-28, 2011.
10. K.H. Ventii, K.D. Wilkinson,"Protein partners of deubiquitinating enzymes," *Biochem J*, 414:161-75, 2008.

This article is adapted from a review in F1000 Biology Reports, *DOI: 10.3410/B3-26 (open access at http://f1000.com/reports/b/3/26).*