

Variations on complexity

Back in the early days of molecular biology, life was a simple equation: a segment of DNA was transcribed into messenger RNA that was, in turn, translated into protein. The combination and quantities of all the proteins expressed in a cell thus defined its characteristics. To understand life, we believed, we would need only to identify every protein and its function. With time, research demonstrated that this was far too simplistic a view of the complex processes of the cell. First, it became clear that genes sometimes have many promoters and therefore different controls on their expression. Similarly, the discovery of alternative splicing and RNA editing did away with the linear 'gene-to-protein' concept: now we know that a single gene can give rise to a variety of alternative products.

However, the biggest complications come from post-translational protein modifications. Reversible phosphorylation was first discovered by Edmond Fisher and Edwin Krebs more than 40 years ago—they were awarded a Nobel Prize in 1992 for their discovery and its impact—and since then, kinase cascades have been at the heart of molecular biology, with phosphatases 'completing the circuit'. The level of complexity was quickly increased by the discovery that a protein can have many phosphorylation sites, each with a binary on/off state that a cell can use to integrate various physiological cues.

In addition, work on the glycosylation of proteins has uncovered another unimaginably complicated modification system that regulates a protein's final destination—within or outside the cell—its function and fate. Unlike phosphorylation, which involves a single phosphate group, glycosylation can tag a protein with linear and/or branched moieties, generating as many or even more implicit possible outcomes. Physiologically, glycosylation carries many

messages that define, among other things, a protein's half-life as it circulates through the body, which has had a direct impact on drug research and development.

Ubiquitination is another example of the complex world of protein modifications. The original simple message—that polyubiquitination targets a protein for degradation—is now recognized as a mere first approximation of reality. Limited ubiquitination modifies the activity of some proteins and can even be reversible. Alternatively, ubiquitination can act as a timed suicide mechanism: the initial activation of the protein by the addition of the first ubiquitins is followed by a 'switch-off' as the proteasomal system recognizes the protein as a target for destruction. Moreover, it is now clear that ubiquitin can be modified at any of its seven lysine residues, generating different linear and branched chains that might lead to fairly diverse functions.

It is already clear that these insights and further developments will affect the work of nearly all biologists. To highlight this important field of research, *EMBO reports* has commissioned a series of review articles that will appear over the next months to present some less familiar examples of post-translational modifications. The theme, 'Protein modifications: beyond the usual suspects', will cover O-linked-N-acetylglucosamination (O-GlcNAcylation), poly(ADP-ribosylation), four types of amino acid or protein conjugations and autophagy—a process that requires both lipidation and protein-protein conjugation. A review of atypical ubiquitin-chain formation (see page 536, this issue) opens the series, followed by less trodden roads of protein or amino-acid modification including polyglutamylation, conjugation of the ubiquitin-like protein Nedd8, and the recently described Urm1.

Some of these modifications are old friends with new functions—only discovered

after technological breakthroughs allowed researchers to uncover new targets and/or selectively inactivate/knock down the modifying enzymes—and some are brand new discoveries with as yet unclear functions. Ironically, the latest unusual suspect—the ubiquitin-like Urm1-Uba4 system—is actually likely to be the most ancient protein-conjugation system, as it has mechanistic and structural homologues among prokaryotes.

Some of these modifications seem to be as abundant as protein phosphorylation. Moreover, they can conjugate to many sites on a single target and, in many cases, form chains of varying lengths. To make matters even more complex, they seem to frequently modulate each other. It is now clear that post-translational modifications generate enormous molecular diversity that allows for extensive fine-tuning of protein regulation and stability, making the task of understanding molecular processes a daunting one (Gannon F (2007) *EMBO Rep* 8, 705). Inevitably, researchers will need a systems-biology approach to handle the enormous amount of data about post-translational modifications and their effects and to integrate them into useful hypotheses.

Often, as scientists, we reduce a problem to a simple equation that excludes other, less relevant information. This is useful to achieve short-term goals, but eventually, such artificial boundaries hinder our quest for knowledge. *EMBO reports* therefore hopes that this series of reviews will help our readers to gain a more comprehensive overview of the immensely complex web of protein modifications. It is, without a doubt, a challenging theme, but it is also clear that it will have to be addressed if we ever want to understand the complexity of life that, ultimately, is the goal of biological research.

Frank Gannon & Nonia Pariente

doi:10.1038/embor.2008.85