



HAL
open science

The logic of protein post-translational modifications (PTMs): Chemistry, mechanisms and evolution of protein regulation through covalent attachments

Marcin J Suskiewicz

► To cite this version:

Marcin J Suskiewicz. The logic of protein post-translational modifications (PTMs): Chemistry, mechanisms and evolution of protein regulation through covalent attachments. *BioEssays*, 2024, 10.1002/bies.202300178 . hal-04438237

HAL Id: hal-04438237

<https://hal.science/hal-04438237v1>

Submitted on 5 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

REVIEW

The logic of protein post-translational modifications (PTMs): Chemistry, mechanisms and evolution of protein regulation through covalent attachments

Marcin J. Suskiewicz 

Centre de Biophysique Moléculaire, CNRS – Orléans, UPR 4301, affiliated with Université d'Orléans, Orléans, France

Correspondence

Marcin J. Suskiewicz, Centre de Biophysique Moléculaire, CNRS – Orléans, UPR 4301, Orléans, France.
Email: marcin.suskiewicz@cnrs-orleans.fr

Funding information

European Union's Horizon Europe Research and Innovation Programme, Grant/Award Number: 101078837; La Ligue contre le Cancer; Centre National de la Recherche Scientifique (CNRS)

Abstract

Protein post-translational modifications (PTMs) play a crucial role in all cellular functions by regulating protein activity, interactions and half-life. Despite the enormous diversity of modifications, various PTM systems show parallels in their chemical and catalytic underpinnings. Here, focussing on modifications that involve the addition of new elements to amino-acid sidechains, I describe historical milestones and fundamental concepts that support the current understanding of PTMs. The historical survey covers selected key research programmes, including the study of protein phosphorylation as a regulatory switch, protein ubiquitylation as a degradation signal and histone modifications as a functional code. The contribution of crucial techniques for studying PTMs is also discussed. The central part of the essay explores shared chemical principles and catalytic strategies observed across diverse PTM systems, together with mechanisms of substrate selection, the reversibility of PTMs by erasers and the recognition of PTMs by reader domains. Similarities in the basic chemical mechanism are highlighted and their implications are discussed. The final part is dedicated to the evolutionary trajectories of PTM systems, beginning with their possible emergence in the context of rivalry in the prokaryotic world. Together, the essay provides a unified perspective on the diverse world of major protein modifications.

KEYWORDS

enzymatic mechanisms, enzyme evolution, history of biochemistry, post-translational modifications (PTMs), protein phosphorylation, protein ubiquitylation

INTRODUCTION

The term protein post-translational modification (PTM), technically speaking, refers to any covalent addition to, transformation of, or subtraction from a protein's chemical structure that occurs in living organisms post-translationally, that is, after ribosome-catalysed biosynthesis of a given protein. However, especially in recent years, the usage has often been narrowed down to covalent addition to

proteins, or, more specifically, the enzymatic covalent addition of elements such as small chemical groups, nucleotides, sugars, lipids or proteinaceous modifiers to protein amino-acid sidechains. This article will focus on such 'canonical PTMs'. Cleavage of protein backbones by proteases (proteolytic processing),^[1] which arguably constitutes the second most important source of covalent post-translational proteome diversification, will not be discussed. Less 'canonical' reactions that result in irreversible transformations of the core part of the amino-acid

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *BioEssays* published by Wiley Periodicals LLC.

sidechains (for example, citrullination^[21]) or affect the protein backbone (for instance, proline isomerisation^[31]) will not be included in the scope of the text. Neither will this review cover non-enzymatic reactions between proteins and reactive molecules such as reactive oxygen species (ROS), isocyanic acid, some thioesters, and so forth, which are thought to happen primarily in disease but might also, in some cases, be physiological and regulatory.^[4] For more details on the modifications excluded from the scope of this text – some of which involve interesting chemical and molecular mechanisms – I refer the reader to the cited reviews.

In some, perhaps most, cases, canonical PTMs on which this essay focusses are reversible and – by affecting protein structure, dynamics and interactions – offer a way of temporarily or locally switching a protein molecule to a state with altered localisation, interaction network, activity or half-life. The high proportion of protein-coding human genes devoted to enzymes that produce, remove and recognise such modification (including several hundred protein kinase and even more protein E3 ubiquitin ligase genes), as well as many known cases of pathological consequences of their dysregulation, highlight their crucial physiological importance. Numerous examples of the links between PTM defects and various diseases are provided in a recent review.^[5]

Here, in the search of ways of understanding canonical PTMs, I will look in turn to historical milestones in detecting and conceptualising PTMs, chemical and mechanistic principles of the PTM systems, and – lastly – the emergence and evolution of such systems. While it is not possible to give an exhaustive account of all these topics in a single review, my aim is to provide a coherent narrative that covers various aspects of PTMs and illustrates these aspects with selected examples. The narrative highlights a central point, which is that major protein PTMs, although diverse, share underlying similarities. I hope that my essay will serve as a useful introduction to the field, offer a unifying framework and stimulate new hypotheses.

MILESTONES IN CHARACTERISING AND UNDERSTANDING PTMS

Our current understanding of scientific phenomena is shaped by the history of research into a given domain and filtered through concepts and paradigms that have emerged throughout that history. With this in mind, the best introduction to protein PTMs is, perhaps, by means of a brief (and necessarily selective) historical survey of key ‘research programmes’ (to borrow Imre Lakatos’ phrase) related to PTMs, highlighting key concepts that originated from these projects and still inform the field, as well as the technical advancements that have shaped it (see also the timeline in Figure 1).

Protein phosphorylation as a regulatory switch

The first relevant line of research is the study of reversible protein phosphorylation as a mechanism for regulating enzyme function^[6,7] (Figure 2A). This programme emerged from the study of sugar metabolism pioneered in the 1940s by Carl and Gerti Cori. The Coris,

among their other contributions, identified glycogen phosphorylase as the enzyme that catalyses the rate-limiting step of glycogen breakdown and demonstrated that it exists in two interconvertible states,^[8,9] a discovery that laid the foundation for the subsequent work.

It took another decade before Edwin Krebs and Edmond Fischer and, in parallel, Earl Sutherland and co-workers showed that these two states correspond to the phosphorylated and dephosphorylated forms of glycogen phosphorylase, establishing phosphorylation as a regulatory switch.^[10–13] The interconversion was enzymatic, catalysed in each direction by a specific enzyme and involved a transfer of the γ phosphoryl of ATP onto one major serine residue^[14] and, in reverse, hydrolysis of the phosphorylated serine. These phosphorylation/dephosphorylation events directly influenced the catalytic activity. The underlying structural mechanism was later elucidated by means of a crystal structure of phosphorylated glycogen phosphorylase by Louise Johnson’s laboratory.^[15] While Krebs, Fischer and Sutherland were not the first to observe protein phosphorylation – one should mention the detection of phosphate^[16] and phosphoserine^[17] in the protein vitellin by Phoebus A. Levene and co-workers and, later, the study of enzymatic casein phosphorylation by Burnett and Kennedy^[18] – the attractiveness of their findings stemmed from assigning to phosphorylation clear functional relevance that was corroborated through subsequent work.

The kinase that regulates glycogen phosphorylase (phosphorylase kinase) was later shown to be itself activated by another kinase (cAMP-dependent protein kinase or PKA), which provided the first example of a kinase ‘cascade’.^[19] Around the same time, it was demonstrated that not only glycogen phosphorylase (which breaks down glycogen), but also glycogen synthase (which produces it), is regulated by phosphorylation, in the latter case, the more active form being the dephosphorylated one.^[20] Ultimately, both enzymes are regulated by hormones insulin and glucagon, through signalling pathways that – as has been gradually revealed – rely heavily on both serine/threonine and tyrosine phosphorylation, with the insulin receptor itself being a receptor tyrosine kinase.^[21] It was not until 1970s and 1980s that protein phosphorylation began to be appreciated as a general regulatory mechanism implicated in, among other processes, the fundamental cellular synchronisation system: the cell cycle. As Paul Nurse, Tim Hunt and their co-workers showed, the transitions and checkpoints during the cell cycle depend in part on protein phosphorylation by cyclin-dependent kinases (CDKs).^[22]

The role of protein phosphorylation in the regulation of sugar metabolism by hormones and its extension to other signalling pathways and the cell cycle provide a powerful paradigm for thinking about PTMs as a regulatory switch. A related but somewhat separate field of research evolved around the structure,^[23] allosteric regulation and small-molecule inhibition of canonical protein kinases,^[24] which arguably remain the best-studied protein-modifying enzymes.

Ubiquitin as a protein degradation signal

The second relevant research programme has been the study of protein ubiquitylation (also known as ubiquitination) as a eukaryotic signal

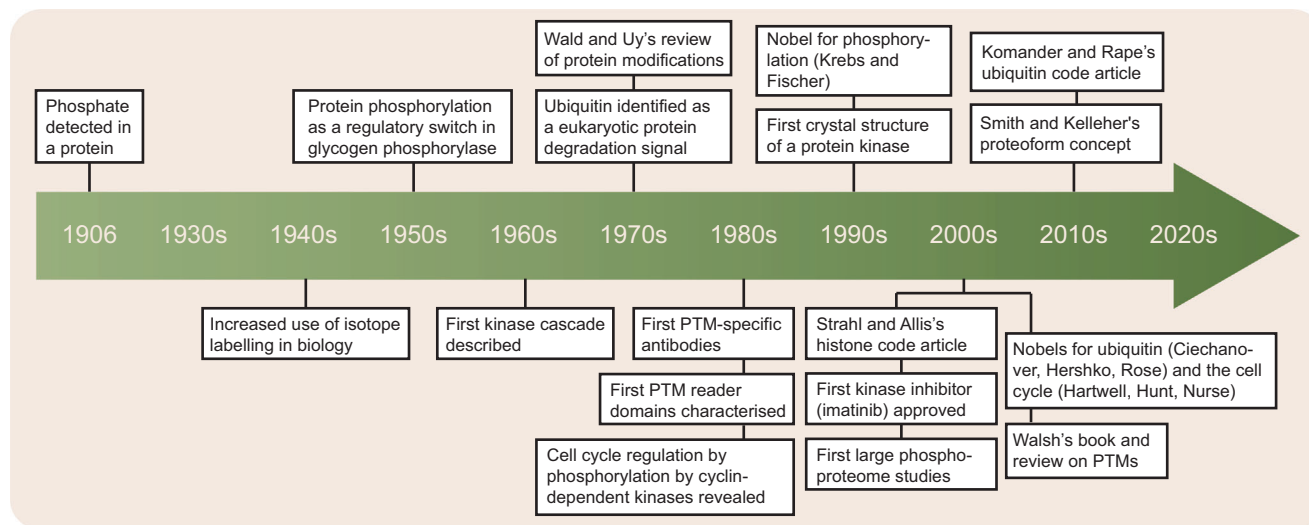


FIGURE 1 Selected milestones in characterising and understanding protein PTMs. The discussion of the indicated discoveries and references to publications can be found in the main text.

for intra-cellular protein degradation^[25–27] (Figure 2B). This field was inaugurated in the late 1970s by Avram Hershko, Aaron Ciechanover and Irwin A. Rose, who, through their work on protein degradation in cell extracts, demonstrated that proteins mixed with an extract become covalently linked to a small protein, which promotes their digestion.^[28–31] This protein was soon identified, by Keith D. Wilkinson and colleagues,^[32] as ubiquitin, which by that time had already been detected covalently linked to histones through an isopeptide bond to a lysine.^[33]

Avram Hershko and colleagues subsequently purified a set of E1, E2 and E3 enzymes required for efficient ubiquitylation^[34,35] and showed that ubiquitin can form polymeric chains on a substrate^[36] – soon recognised by Alexander Varshavsky and co-workers to be the actual signal required for degradation.^[37] It was also Varshavsky's group that was instrumental in identifying 'degrons' – motifs on the substrate that trigger its ubiquitylation and degradation.^[38–40] The biological importance of these findings lies in showing that, contrary to what had been thought before, protein degradation, like protein synthesis, can be a highly *specific* process dependent on a *specific* signal ('specificity' being one of the keywords of early molecular biology, as the historian of biology Michel Morange argued^[41]). In later decades, other PTMs – notably Pupylation and arginine phosphorylation – have been demonstrated to serve analogous roles as degradation signals in some prokaryotes.^[42,43] At the same time, ubiquitylation has been shown to have multiple other functions in eukaryotes beyond promoting proteasomal degradation,^[44] and homologues of ubiquitin, including small ubiquitin-like modifier (SUMO) paralogues and other ubiquitin-like proteins, have been implicated in various functions as well, mainly related to stress and immune responses.

The gradual realisation that phosphorylation and ubiquitylation are analogous in some ways (despite one employing a small inorganic group and the other a protein as a modifier) and that the umbrella term under which both of them fall is that of a protein PTM has arguably been a key

factor in the current appreciation of PTMs' importance.^[45–49] These two modifications have also been at the forefront of PTM-related pharmaceutical research, with kinase inhibitors and small-molecule degraders (that is molecules, including PROTACs and 'molecular glues', that enhance degradation of desired substrates by physically linking them with E3 ligases^[50]) as two major modalities in current drug research.

Detecting and cataloguing protein modifications

In parallel to these two research programmes – which started from a function and identified a PTM behind the function – protein chemists and biochemists have long reported various noncanonical amino acids in proteins without necessarily knowing their functional importance. We see the fruit of these efforts summarised in early reviews such as that by Finn Wold and Rosa Uy published in 1977, which catalogued 140 possible 'amino-acid derivatives', gathering them (together with proteolysis) in a single category of 'covalent posttranslational modifications'.^[51] Wold and Uy proposed to see the modification of specific amino-acid sidechains in a protein as a maturation step, on a par with proteolytic processing, that occurs after translation to yield the final protein product. This is reminiscent of the recent concept of a 'proteoform' that is discussed in Section 'Techniques for studying PTMs' below.

The 2006 seminal book on PTMs^[52] and a popular review article^[53] published around the same time by the recently deceased Christopher T. Walsh (1944–2023) can be seen as a continuation of Wold's systematising project. Walsh, a biological chemist who had previously published a textbook on catalytic mechanisms of metabolic enzymes,^[54] conceived of PTM reactions in chemical terms, as, for the most part, instances of transfer of electrophilic groups onto nucleophilic protein side-chains^[55] – a realisation that I take up and develop below

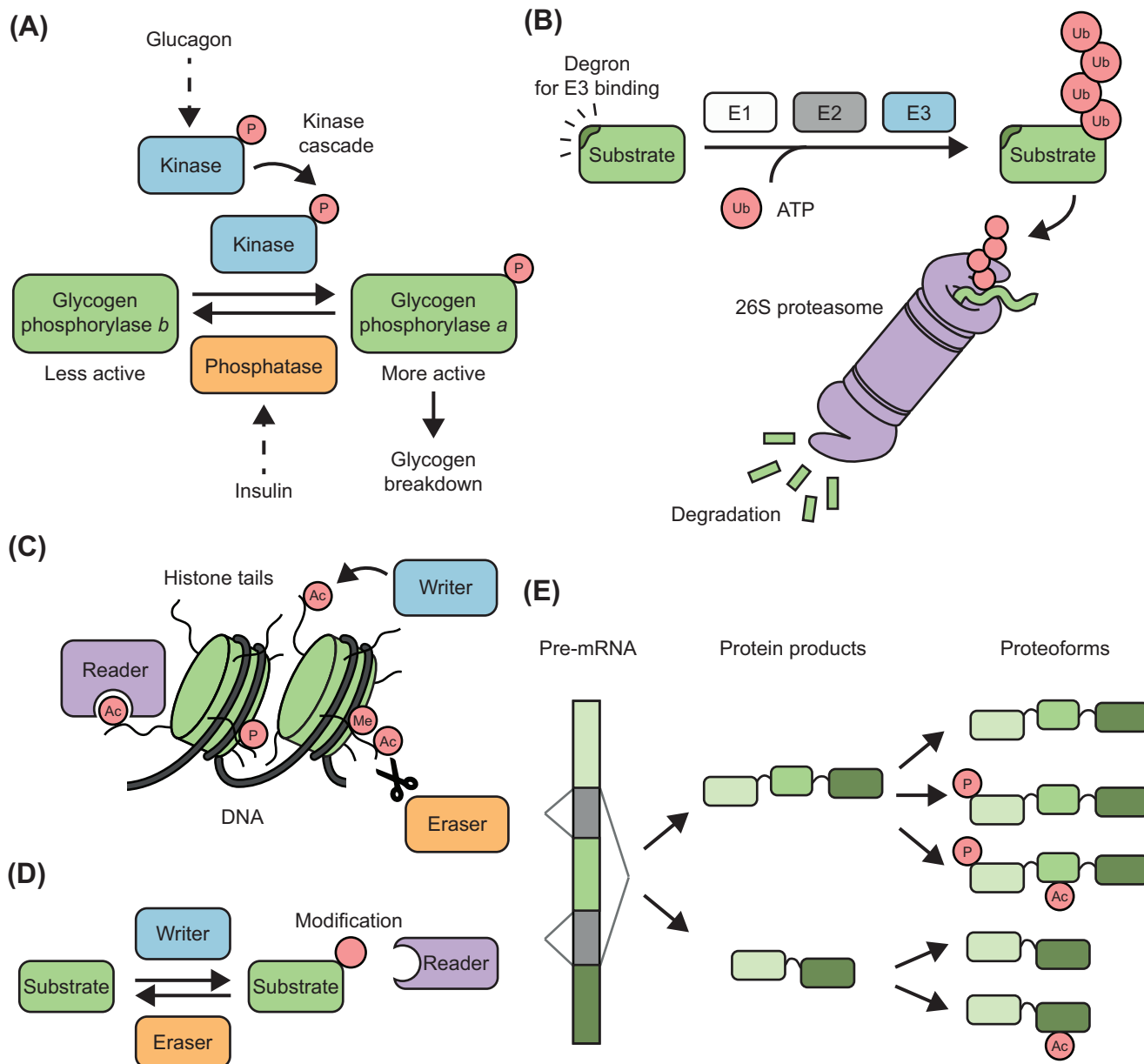


FIGURE 2 Key paradigms in PTM research. In all panels (and other figures in this article), modifications are shown in light red, protein substrates in green, writers in blue, erasers in yellow and readers in violet. (A) Regulation of the glycogen-degrading activity of the enzyme glycogen phosphorylase by protein phosphorylation. Phosphorylation and dephosphorylation of this enzyme are ultimately regulated by the hormones glucagon and insulin, through signalling pathways indicated schematically with dashed arrows. (B) Protein ubiquitylation as a signal for degradation by the 26S proteasome. The ubiquitylation reaction is catalysed by an enzymatic cascade composed of E1, E2 and E3 proteins and requires ATP. A degron motif on the substrate promotes ubiquitylation by physically interacting with the E3 ligase. A poly(ubiquitylated) substrate is recognised by receptor proteins within the 26S proteasome, unfolded and degraded. (C) Regulation of chromatin structure and gene expression by the histone code. Protein modifications on histone tails are installed by writer enzymes, removed by eraser enzymes and recognised by reader proteins. (D) A general scheme of protein regulation by PTMs based on panel C. (E) Sources of variation that produce multiple proteoforms from a single protein-coding gene. A single gene can be alternatively spliced to produce several isoforms, which can be further diversified through differential PTM patterns. Additional sources of proteoform diversity omitted from the figure include, for example, single-nucleotide polymorphisms and alternative translation start sites. Ac, acetylation; Me, methylation; P, phosphorylation; Ub, ubiquitin.

when describing the ‘chemical logic’ of PTMs. My review aims to commemorate Walsh’s extraordinary contribution in the wake of his passing.

Since the early reviews by Wold and co-workers, the project of cataloguing PTMs has progressed, reaching the current total count

of over 650 known PTM types and steadily increasing.^[5] It is worth noting that, although PTMs have been traditionally more studied in eukaryotes and especially in human cells, recent years show that bacteria are a fertile ground for the discovery of interesting PTM systems, some of which – such as Pupylation or histidine and arginine

phosphorylation – are apparently unique to, or particularly abundant in, bacteria.^[56]

The histone code

The last major thread that greatly contributed to the history of PTM research and the way we currently understand PTMs is the concept, put forward by Strahl and Allis in 2000, of a histone ‘language’ or ‘code’^[57] (Figure 2C). The ‘code’ in question refers to a set of different PTM signals (especially methylation, acetylation and phosphorylation) on specific sites within histone proteins, which could act sequentially or together to determine particular functional outcomes, primarily related to gene expression.^[58] The idea that histone modifications could have regulatory roles is much older,^[59] but its conceptualisation by Strahl and Allis as a ‘code’ – and the rapid adoption of this idea by other authors^[60,61] – has had a crucial impact of its own.

According to this framework, the various PTM ‘letters’ of the code are supposed to be ‘written’ (i.e., produced), ‘erased’ (i.e., removed), and ‘read’ (i.e., recognised via a noncovalent interaction) by dedicated protein classes, which soon began to be called ‘writers’, ‘erasers’ and ‘readers’, respectively, a terminology that is now broadly followed in the PTM field (Figure 2D). The effects of histone PTMs on chromatin state are thought to be either direct (the PTM itself affects histone:histone and/or histone:DNA interactions) or mediated by readers such as energy-consuming chromatin remodellers. Moreover, a coupling between reading and writing can lead to one modification promoting the installation of another.

A vindication of sorts of the writer–eraser–reader paradigm for histone modifications came with the discovery of erasers of protein methylation,^[62,63] which happened after the paradigm was first proposed, methylation initially appearing exceptional in having no known erasers. Furthermore, the attractiveness of this framework lies in it gathering together many diverse chromatin-associated factors whose mutations are known to be associated with disease and which are targets in the ongoing efforts of drug development and clinical trials.

Conceptually closest to the ‘histone code’ is the so-called RNA polymerase II CTD code, which refers to the functional link between changing patterns of PTMs in the C-terminal domain (CTD) of RNA polymerase II and various (co-)transcriptional events.^[64,65] Moreover, the use of the term ‘code’ – undoubtedly tracing its roots back to the ‘genetic code’ – is a recurring theme in the PTM field, evident in concepts such as the ‘tubulin code’ (regulation of tubulin by its different PTMs^[66]) or ‘ubiquitin code’ (distinct functional roles of different ubiquitin chains^[67–69]).

Techniques for studying PTMs

The famous biologist Sydney Brenner argued that ‘progress in science depends on new techniques, new discoveries and new ideas, probably in that order’.^[70] Above, I have reviewed some key discoveries and ideas in the PTM field, but they would not have happened without appropriate technology.

The birth of the protein phosphorylation field was made possible, in part, by the use of ³²P as a radioactive tracer, which had been pioneered, alongside that of some other radioactive as well as non-radioactive but heavy isotopes, by Georg de Hevesy in the 1920s. As the technique became more widely used on both sides of the Atlantic from the late 1930s, it played a significant role in the study of metabolism and early developments in molecular biology.^[71] Likewise, the use of ³²P, and specifically ³²P-labelled ATP, was instrumental in the early days of protein phosphorylation and remains widely used in the field to this day, particularly for in vitro kinase assays, where it is typically coupled with a sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE), a now standard technique introduced in 1970s.^[72] Other forms of isotope labelling (for example, ³H, ¹²⁵I or ³⁵S) have been important for tracking the fate of specific proteins in the ubiquitylation field.

Like the early studies of serine phosphorylation, the discovery of protein tyrosine phosphorylation was also indebted to the use of ³²P.^[73] However, the key role in the later development of the tyrosine phosphorylation field has been played by high-quality general anti-phosphotyrosine antibodies introduced in the 1980s,^[74,75] all the more important for tyrosine phosphorylation being less abundant than canonical serine/threonine phosphorylation. Around the same time, the first anti-ubiquitin antibody was also introduced.^[76] Indeed, throughout the PTM field, anti-PTM antibodies that can either detect a given modification on all/many substrates or be specific for one site in a particular substrate, continue to play a key role through their use in immunoprecipitation, immunoblotting and immunofluorescence. For complex modifications such as ubiquitylation, antibodies with preferences for different chain types exist^[77], and for some PTMs, non-covalent and covalent probes other than antibodies are also widely used.^[78] Aside from the tools for detecting and enriching PTMs, the study of protein PTMs has also benefited from the use of small-molecule inhibitors targeting eraser enzymes, exemplified by the early serine/threonine phosphatase inhibitor, okadaic acid.^[79] These inhibitors not only increase the stability and abundance of PTMs in extracts but also allow observing the effects of dysregulating specific PTMs in living cells.

Research into PTMs has received a powerful boost from the development of mass spectrometric (MS) methods, which provide an unparalleled tool for detecting protein modifications, both those obtained through enzymatic reactions reconstituted in vitro and those occurring natively within cells.^[80,81] As PTMs are relatively lowly abundant and vary in their enzymatic, chemical and fragmentation sensitivity, new advances in the field have been associated with tailored extraction, enrichment and fragmentation protocols, in addition to the steady general improvement in MS instrumentation and software. For example, proteomics-scale studies of protein phosphorylation in cell extracts were performed in early 2000s following the development of suitable enrichment approaches.^[82–84] In some cases, a PTM is transformed using chemicals or enzymes prior to enrichment and detection. Moreover, various label-based or label-free quantitative proteomics approaches allow comparing PTM levels and detecting changes in PTMs related to particular conditions or stimuli.^[85]

The MS-based analysis comes with several caveats. The differential sensitivity and detectability of different PTM types and different modified peptides pose a challenge, as it introduces a bias against those modifications and sites that are less amenable to the available technical pipelines. Moreover, it might sometimes be challenging to distinguish native PTMs from chemical protein modification that occurred during sample preparation. A further important caveat of MS-based identification of PTMs is that, unless we actively look for a given PTM based on the prior knowledge (or assumption) of its existence, it might become lost in unassigned spectral peaks, making it conceptually challenging to detect yet unidentified modifications. These qualifications notwithstanding, the astonishing progress in MS methods over the years played a key role in the identification of PTM substrates, yielding high-quality profiles of major PTMs in various organisms and cell types. This includes deep-probing profiling of, for example, phosphorylation,^[86,87] acetylation,^[88,89] ubiquitylation,^[90] SUMOylation^[91,92] and ADP-ribosylation^[93,94] in human or other mammalian cells. Databases grouping PTM sites from both low- and high-throughput (the latter all MS-based) studies are available, including PhosphoSitePlus,^[95] iPTMnet^[96] and dbPTM.^[97]

While the mentioned datasets were predominantly obtained with bottom-up MS approaches, in which protein samples are digested into peptides prior to the analysis, an alternative top-down strategy involves injecting intact proteins into a mass spectrometer. This allows the identification of distinct intact covalent protein products – known as ‘proteoforms’ – each having a signature molecular weight^[98] (Figure 2E). The weight of a proteoform is a function of its amino-acid sequence (determined by the DNA coding sequence and the way it is spliced), as well as any combination of PTMs that might be simultaneously present on the same polypeptide. While the number of human protein-coding genes is relatively limited (estimated to be around 20 000), the combined effect of alternative splicing and PTMs gives rise to a much larger number of proteoforms. The Human Proteoform Project has so far detected over 60 000 distinct species originating from around 5000 human genes, but the total number of proteoforms is likely much larger, reaching the range of hundreds of thousands or more.^[99,100] It has been proposed that such diversity could explain the stunning complexity of the human organism, which is difficult to square with the limited gene number. However, it remains challenging to distinguish functionally distinct proteoforms from variation that might be the result of PTM ‘noise’ with limited functional bearing.

The current understanding of PTMs would not be possible without MS, which arguably remains the most powerful technique for probing the PTM world. Ultimately, however, the research into PTMs – like in other areas – has advanced through a fruitful mixture of different techniques and approaches.

From the past to the future

The above-mentioned lines of research are of course interdependent. For example, there is a clear link between the idea of writing and erasing the ‘histone code’ and the pioneering work on phosphorylation as a reversible process. The concept of ‘PTM reading’ is also ultimately

indebted to the phosphorylation field, where first domains capable of recognising phosphorylated peptides were reported in the early 1990s,^[101–104] which motivated the search for analogous specific domains for the recognition of other major protein PTMs. Similarly, the concept of a proteoform is related not only to the project of cataloguing different PTMs and to the technical developments in MS but also to the idea of a functional ‘code’ potentially produced by various PTMs simultaneously present on the same polypeptide.

The history of the PTM research has furnished powerful paradigms for thinking about PTMs as a regulatory mechanism, but all paradigm-based thinking can both facilitate and distort understanding. For example, the classical case of glycogen phosphorylase phosphorylation (by a single kinase, on a single main site, with a direct effect on activity and structure) is not – as Krebs and Fischer already realised^[7] – representative of all protein phosphorylation, let alone other PTMs. Similarly, not all PTM instances are ‘read’ by specific domains, or are necessarily reversible.

As a matter of fact, we are now aware that PTMs are quite widely installed on different proteins (but sometimes only on a fraction of copies of a given protein) and most detected sites might not have any clear functional relevance, at least when the site or the protein is considered in isolation from others.^[105] To overcome this ambiguity, efforts to distinguish sites that are more likely to be functionally important – based on evolutionary conservation and other criteria – are made.^[106,107] Explanations drawing on evolutionary and systems biology are likely needed to fully appreciate the biological role of PTMs, in addition to more traditional biochemical, structural, proteomics and cell biology analyses.

CHEMICAL AND CATALYTIC LOGIC OF PTM REACTIONS

PTMs come in different shapes and sizes, ranging from small chemical groups such as methyl, acetyl or phosphoryl, through lipids, nucleotides and sugars, to proteinaceous modifications such as ubiquitin and its homologues (Figure 3A). Some PTMs can be attached as chains, which further increases their size. Due to the inherent specificity of main writer enzymes, each PTM type is predominantly associated with particular amino-acid acceptors (in a given group of organisms), as in – in eukaryotes – serine, threonine and tyrosine phosphorylation; lysine and arginine methylation; lysine acetylation, ubiquitylation and SUMOylation or serine ADP-ribosylation. While phosphorylation, acetylation and ubiquitylation appear to be among the most abundant modifications in eukaryotes,^[108] comparisons of abundance are confounded by differential stability and detectability of, and unequal investment of the scientific community in, different PTMs.

Modification reaction

While PTMs are chemically diverse at the level of the modification itself, the modification reaction tends to, at least for some of the most common PTMs, follow a unified scheme, whereby the modification (in

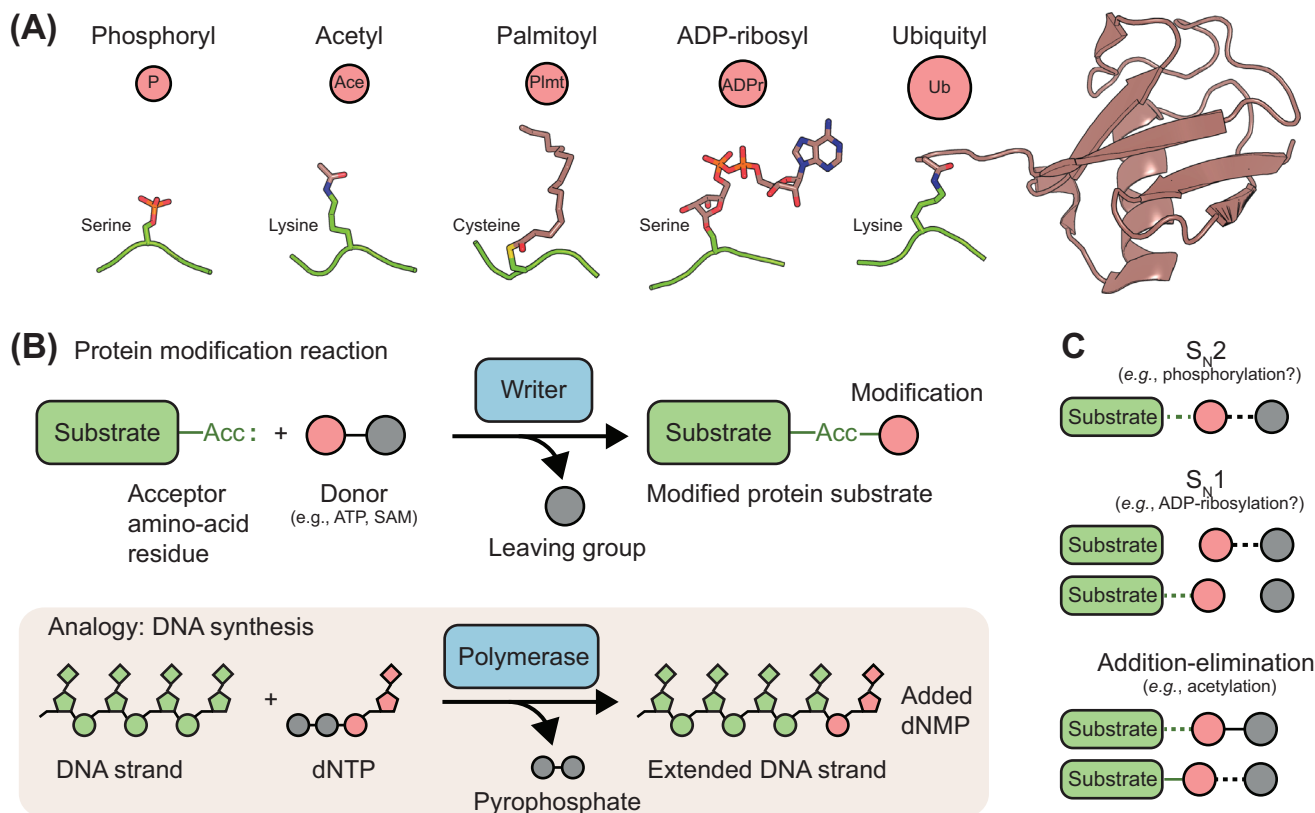


FIGURE 3 Chemical diversity and similarity among PTMs. (A) A selection of PTMs differing in chemical character and size. Three-dimensional structural representations of different modifications together with the modified amino-acid residue and the flanking backbone are shown at the same scale based on PDB entries: 1T2V (phosphoryl), 2WP1 (acetyl), 5EMW (palmitoyl), 7AKS (ADP-ribosyl), 5LN1 (ubiquityl). (B) A schematic representation of a protein modification reaction. PTM reactions are typically nucleophilic substitution reactions involving a nucleophilic acceptor amino-acid residue on the substrate and a donor molecule, from which a leaving group is released upon reaction. The box shows a simplified representation of a DNA synthesis reaction, in which dNMP is transferred from dNTP to the 3' end of a growing DNA strand with concomitant release of pyrophosphate. Removal of a proton from the acceptor amino-acid residue, which is frequent in PTM reactions, is not shown. (C) Variation in nucleophilic substitution types observed among PTM reactions. In S_N2 -type reaction, the bond between the substrate and the modification is being created at the same time as the bond between the modification and the leaving group is being broken. In S_N1 and addition-elimination reactions, these two events happen in two separate events shown in a simplified manner one after the other. Bonds that are being broken or formed in a given step are indicated with dashed lines. Acc, acceptor.

light red) is transferred from the donor to an acceptor residue in the substrate (in green) with the concomitant departure of a leaving group (in grey) (Figure 3B).

To help the readers grasp this process more easily, I will draw a parallel between a protein modification reaction and a classical example of covalent addition in biology: that of a new nucleotide to the 3' hydroxyl group on a growing DNA chain during DNA synthesis (Figure 3B, box). In the latter reaction, whether occurring in a cell or a PCR tube, DNA polymerase employs deoxynucleotide triphosphates (dNTPs) as nucleotide donors. When a deoxynucleotide monophosphate (dNMP) becomes incorporated into DNA, the remaining two phosphates (the pyrophosphate moiety) from dNTP depart as a leaving group. Pyrophosphate is a good leaving group, which means, in simplified terms, that it is adept at accepting electrons and dissociating, thus making the group to which it is/was attached (in this case, dNMP) an easy prey for the attacking nucleophile (in this case, the 3' hydroxyl). A dNTP molecule can, therefore, be seen as an 'activated form' of a

deoxynucleotide, the pyrophosphate making the dNMP more reactive by depriving it of electrons. We can observe the same logic during main PTM reactions.

A PTM reaction that is chemically very similar to the DNA synthesis reaction is protein AMPylation (also known as adenylation),^[109] which involves the attachment of adenosine monophosphate (AMP) to protein residues. Here, the donor molecule is again a triphosphate (ATP), and the role of the pyrophosphate moiety is to 'activate' AMP by serving as a good leaving group. Other PTM reactions are less directly similar to DNA synthesis but still analogous. Protein phosphorylation, interestingly, uses the same donor as AMPylation, the triphosphate ATP, but a different part of the donor molecule (the γ phosphoryl) becomes attached to the protein, while ADP departs as the leaving group. For other modification reactions, donors are chemically different, but it is still possible to distinguish, in each case, a modification and a (good) leaving group, as discussed in Section 'Donor properties and enzymatic donor activation' below.

Chemically speaking, the described processes conform to a nucleophilic substitution mechanism in which an electron pair from a nucleophilic acceptor makes a bond to an electrophilic modification in the donor, displacing a leaving group. The leaving group can depart before, during or after the formation of the bond between the acceptor and the modification, which corresponds to S_N1 -type, S_N2 -type or addition–elimination substitution mechanisms, respectively (Figure 3C). The S_N2 mechanism, which is thought to be dominant in biological systems,^[110] might also be predominant among PTMs, but the mechanism depends on the group that is being transferred. Phosphorylation most likely proceeds via an S_N2 -like mechanism,^[111] with a particular loose or ‘dissociative’ transition state,^[112] but the issue is not settled yet. In contrast, acyl-based modifications, such as acetylation, palmitoylation or ubiquitylation, generally proceed by addition–elimination. Lastly, ADP-ribosylation likely has some S_N1 character, where the departure of nicotinamide slightly precedes the formation of a bond between a protein residue and ADP-ribose.^[113]

PTM catalysis

The difference between a biological PTM reaction and an idealised chemical mechanism is that the first is a biochemical process that takes place in an active site of a specific enzyme. PTM acceptors and donors would be expected to have some tendency to spontaneously react – at least when brought into a close proximity – as this can facilitate the emergence of an enzymatically controlled reaction (for reasons elaborated in Section ‘Chemical driving forces’ below). Generally speaking, however, many of the key chemical groups encountered in biological systems tend to strike a balance between being relatively inert in the absence of catalysts at physiological pH and getting easily activated in the active site of an appropriate enzyme – and this is also the case for PTM acceptors and donors.

Since chemical reactions are similar at some level for different PTMs, there are also analogies between different writer enzymes in terms of how they exert their catalytic role. In addition to the proper orientation of the acceptor and donor groups relative to each other – which is particularly important for S_N2 -type substitution mechanism that requires good orbital alignment – there are also other recurrent ‘tricks’ by which enzymes potentiate PTM acceptors and donors.

Acceptor properties and enzymatic acceptor activation

PTMs can be installed on various amino-acid side-chains, the most common being serine, lysine, threonine, tyrosine, cysteine and arginine.^[108] Enzymes that produce (and also those that erase) PTMs tend to have a strong preference for a specific amino-acid residue or, at most, closely similar residues, as an acceptor, which allows the existence, in the same cell, of enzymatically orthogonal and potentially functionally distinct systems for the same PTM (for example, serine/threonine vs. tyrosine phosphorylation).

PTM acceptors, chemically speaking, behave as nucleophiles that use their electrons to attack – and attach – a PTM. The common feature of the most frequently used acceptors of PTMs is that they exist predominantly in their poorly nucleophilic protonated forms at neutral pH, but can become more highly nucleophilic upon the loss of a proton. Active sites of writer enzymes, therefore, tend to encourage the deprotonated form and sometimes perform direct proton abstraction from (‘deprotonation’ of) the acceptor by a suitably positioned catalytic base. This mechanism might be particularly important for serine, threonine and arginine modification due to their extremely low intrinsic tendency to lose a proton (approximate pK_a value of 13, compared to 10 for tyrosine and lysine, and 8 for cysteine).

While it is difficult to conclusively prove a residue’s role in deprotonation (as opposed to, for instance, an electrostatic or an orientation effect), a potential catalytic base (typically an aspartate, a glutamate, or a histidine) has been proposed in many writer enzymes, including, among others, protein kinases (an aspartate, Figure 4A),^[23,114–117] PRMT (protein arginine methyltransferase)-family protein arginine methyltransferases (a glutamate),^[118] some GNAT (GCN5-related N-acetyltransferase)-family protein lysine acetyltransferases (a glutamate),^[119] Fic domain-containing AMPylating enzymes (a histidine)^[120] and hydroxyl-specific ubiquitin E3 ligases (a histidine).^[121–123]

An interesting case is provided by the histone poly(ADP-ribosylation) factor 1 (HPF1):poly(ADP-ribose)polymerase 1 (PARP1) and HPF1:PARP2 complexes, which are responsible for the bulk of DNA damage-induced protein ADP-ribosylation in human cells (Figure 4B). In this case, the active site catalysing the PTM reaction is built of two proteins, where one (HPF1) provides a potential catalytic base (a glutamate) for the deprotonation of a serine acceptor, while the other (PARP1 or PARP2) binds and activates the donor NAD.^[124–129] In the absence of HPF1, PARP1 or PARP2 appear unable to significantly modify serine residues and instead slowly ADP-ribosylates glutamate acceptors in proteins, which are predominantly deprotonated at a neutral pH (thus not requiring a catalytic base for deprotonation) but are not strongly nucleophilic and rarely serve as PTM acceptors.^[124,128,130] In HPF1’s presence, the complex efficiently targets serine acceptors as the main physiological ADP-ribosylation targets.

In the case of some protein methyltransferases, there is no explicit suggestion of a specific catalytic base residue, but the active site of the enzyme is thought to create an environment that promotes proton loss by the bound lysine, thereby effectively lowering its pK_a value. In these enzymes, a channel has been postulated to provide access to bulk water, which would accept the proton.^[131]

Donor properties and enzymatic donor activation

In this general framework, the donor molecules are composed of the transferable group that will become the protein modification (in light red in figures) and a part that will become the leaving group (in grey), both elements being important for the reaction.

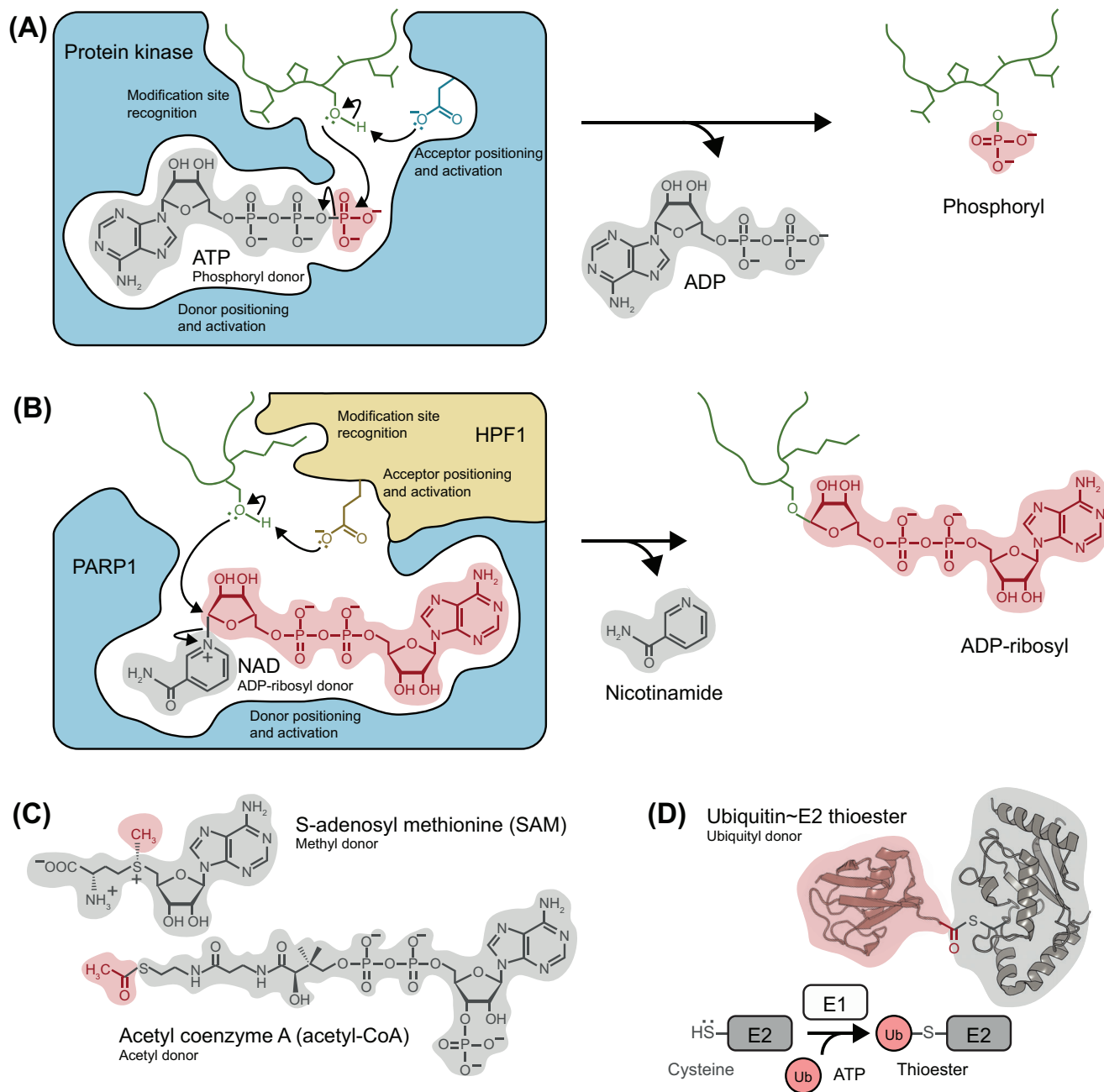


FIGURE 4 Donors of PTMs and catalytic mechanisms of protein modification. In all panels, the modifications are shown in light red, while the leaving groups are indicated in grey. (A) A simplified catalytic mechanism of a PTM reaction using protein phosphorylation as an example. The enzyme recognises the donor and the modified motif in the substrate and positions and activates both elements. The acceptor amino-acid residue in the substrate can be activated through deprotonation by a catalytic base. (B) A simplified catalytic mechanism of protein ADP-ribosylation by the HPF1:PARP1 complex showing analogy to the phosphorylation mechanism shown in (A). (C) Chemical structures of donors of protein methylation and acetylation. (D) A schematic chemical structure of the donor of ubiquitylation, the ubiquitin~E2 thioester molecule. Protein structures are not shown to scale compared to chemical bonds. A schematic below provides a simplified picture of the E1-catalysed reaction in which Ub and E2 become covalently joined to form the ubiquitin~E2 thioester.

We can distinguish these two elements in major PTM donors including ATP (phosphorylation – see Figure 4A – and AMPylation/adenylation), SAM (methylation, Figure 4C), NAD (ADP-ribosylation, Figure 4B), acetyl-coA (acetylation, Figure 4C) and lipid-CoA (various forms of lipidation) and UDP-sugars (various forms of glycosylation). All these molecules serve also other roles in cellular

metabolism^[132] and cells have pathways for maintaining their appropriate (often relatively high) levels through regeneration from the leaving group or other synthetic pathways. In fact, it is the cell's ability to maintain a high ratio of these donors relative to their corresponding leaving groups (for example ATP relative to ADP) – rather than any specific 'high-energy bonds' that they might contain – that is

the key to these donors' status as 'high-energy molecules' that can thermodynamically drive the reactions in which they are involved in the forward direction. The favourable chemical properties of these molecules, including the types of bonds involved, can account for their favourable reactivity, but it is the displacement of their concentration from equilibrium relative to the concentration of the leaving group that 'stores' free energy and makes many PTM writing reactions effectively unidirectional. As a rule, the demodification reaction has to be catalysed using a different mechanism and a different enzyme – an eraser – as discussed in Section 'Reversibility' below.

In the case of ubiquitylation and ubiquitin-like modifications, the donor is a covalent complex of two proteins, including an enzyme (typically an E2) and the modifier, linked through a thioester bond (Figure 4D). The thioester between a cysteine in E2 and the modifier's C terminus is created in an ATP-dependent reaction catalysed by an E1 enzyme. Indeed, the need to create a dedicated donor molecule for ubiquitylation explains why this and related PTMs, unlike others, require a complex catalytic cascade composed of E1, E2 and E3 proteins instead of a single writer. Within this cascade, an E2 enzyme functions both as the modification carrier (the 'leaving group') and also, in part, the writer responsible for catalysis. In the latter task, E2 is assisted by an E3, typically a scaffolding protein that stabilises the active conformation of the modifier~E2 molecule and recruits a protein substrate. In some cases – that is, homologous to the E6AP carboxyl terminus (HECT) and RING-between-RING (RBR) families and some more recently identified instances – the E3 ligase contains a reactive cysteine, or a relay of cysteines, on which it can temporarily accept ubiquitin from the E2.^[133] In these cases, it is the E3 protein that serves the role analogous to an E2: both the 'leaving group' and a writer. Despite these special aspects, the ubiquitylation and related reactions can be said to broadly conform to the same basic scheme as other PTMs.

For all major PTMs, the part of the donor that will become the protein modification contains an atom that becomes the target of the nucleophilic attack by the PTM acceptor. This atom tends to have electrophilic properties, that is to say, be relatively electron-poor, for example, due to covalent bonds to electron-withdrawing groups. For instance, in ATP, all three phosphorus atoms are relatively electrophilic due to bonds to surrounding oxygens, and we know that two of them (γ and α) can mediate protein PTMs (phosphorylation and AMPylation/adenylation, respectively), as mentioned above. The choice between these two modifications is governed, in this case, by the catalysing enzyme, through the positioning of the appropriate phosphorus atom for the nucleophilic attack. Interestingly, examples of evolutionary plasticity where a kinase-like protein (SELENOO) catalyses protein AMPylation^[134] while an AMPylator-like protein (catalyses protein phosphorylation^[135] – due to a shift in ATP orientation – have been reported, highlighting the chemical suitability of various phosphorus atoms in ATP as targets of nucleophilic attack and the decisive role of positioning by the writer enzyme. In the case of the methylation and ADP-ribosylation donors, SAM and NAD, the central carbon atoms of the modification parts – which are not inherently electrophilic – are rendered electrophilic through their linkage to pow-

erfully electron-withdrawing positively charged sulphur or nitrogen atoms.

As argued above in the parallel drawn between DNA synthesis and protein modification, the part of the donor that becomes the leaving group plays an important role in the reaction by accepting the electron pair that previously made the bond between it and the modification, thus pulling electrons from the modification and 'liberating' it for transfer. A chemical group that is adept at these tasks can promote the overall substitution reaction. The quality of the leaving group depends on its ability to accommodate extra electrons, which can be done, for example, by using them to neutralise a positive charge or by distributing them within a resonance system. Again, SAM – in which the leaving part (SAH) turns from positively charged to neutral upon methyl transfer – is an excellent example of a donor with a potent leaving group, as is NAD, in which nicotinamide leaves as a neutral aromatic compound (Figure 4B). On the other hand, the pyrophosphate group (in inorganic pyrophosphate or ADP) provides a case of a resonance system (Figure 4A).

While canonical donor molecules are well suited for their role – explaining their recurrent use not only for PTMs but also in various other cellular reactions – they do tend to remain relatively inert in the absence of catalysis, as otherwise, they would dissipate through spontaneous non-specific reactions. Therefore, the donor molecules have to combine their non-equilibrium status with high kinetic stability in the absence of catalysts. For phosphate-containing molecules such as ATP, this has been proposed to be related to the electrostatic effects such as shielding of the central electrophilic phosphorus atoms from nucleophilic attacks by neighbouring negatively charged oxygens.^[136,137] In active sites of enzymes, such electrostatic protection can be overcome by proximity/orientation effects combined with a neutralising electrostatic environment, and both these aspects might be fulfilled by a tight network of positively charged amino-acid side-chains and metal ions observed for canonical and noncanonical protein kinases and AMPylating enzymes.^[23,109,138]

More generally, donor activation might involve binding of the donor in a conformation where the modification moiety is exposed to the nucleophilic attack. This is well established for ubiquitylation and other ubiquitin-like modifications, where E3 proteins typically fulfil their activatory role by chaperoning the so-called 'closed' conformation of the otherwise flexible modifier~E2 thioester, this closed conformation appearing the most conducive to a nucleophilic attack by a lysine.^[139–143] Analogously, NAD appears to be always bound in a similar conformation by ADP-ribosylating writer enzymes, which possibly not only exposes the anomeric carbon atom to the attack by the acceptor residue but also introduces a strain into the donor molecule, the release of which could help drive the reaction.^[144]

MECHANISTIC LOGIC OF PTM SYSTEMS

While the active site catalysing a PTM reaction can be likened to an 'engine' of a PTM system, other elements and mechanisms are needed to turn it into effective regulatory machinery. The 'steering wheel'

is provided by mechanisms of substrate recruitment that direct the modification to particular proteins and sites, while eraser enzymes furnish a necessary 'brake' (or, indeed, a 'reverse gear') to regulate the output. These features are discussed in turns below. Finally, the mechanisms by which the modifications exert their regulatory roles will be briefly discussed, with a particular focus on the propensity of PTMs for promoting new protein:protein interactions.

Substrate specificity

Traditionally, the question of PTM specificity has been reduced to that of the targeted amino-acid residue and the surrounding amino-acid sequence motif (Figure 5A). For writers as diverse as protein kinases,^[145,146] the SUMOylating E2 enzyme UBC9^[147] or the ADP-ribosylating complex HPF1:PARP1,^[148] consensus motifs ranging from a dozen or so down to two residues flanking the modified site can be detected, although the requirement for them is rarely absolute.

Recent research suggests that it is not always desirable for a substrate to have an optimal consensus modification motif; instead, by diverging from ideal sequence patterns, substrates can be modified to the extent and with kinetics that are appropriate for their function.^[149] As a result, sequence motifs can determine the order of substrate modification during the cell cycle^[150] and are responsible for varying sensitivity of substrates to the inhibition of kinase activity.^[151] In one recently described example, artificially improving a suboptimal phosphorylation motif found in a substrate important for T-cell activation resulted in detrimental excessive reactivity to self-antigens due to enhanced signalling.^[152] Similar logic of a spectrum of 'better' and 'worse' substrates can apply to PTM removal, where some proteins and sites are faster demodified than others in cells by eraser enzymes.^[153]

While motifs surrounding the modification site dominated attention in the past, more recent research suggests that such motifs – where they exist – are rarely the sole or even the main determinant of which proteins become modified, although they can guide where the modification will be within a protein. In many cases, another mechanism has to first ensure co-localisation of the substrate and the writer. One way this can be achieved is by a secondary docking element – typically also a linear motif, but sometimes a structural one – on the substrate that binds to a substrate-recruiting region on the writer distinct from the active site (Figures 5B and 5C). In protein kinase substrates, these two types of motifs often synergise in promoting phosphorylation.^[149] In substrates of some other writers – for example, the ADP-ribosylating enzyme tankyrase^[154] or many ubiquitin E3 ligases^[39,155] – the docking motif actually plays the dominant role, whereas there is little apparent preference regarding the actual modification site other than that it is the correct amino-acid type and is accessible. The secondary docking site can be found on the same protein domain of the writer as the active site (for example, in some kinases; Figure 5B) or a different domain within the writer (in tankyrase and some single-protein ubiquitin E3 ligases; Figure 5C). Alternatively, substrate recruitment is sometimes performed by a different protein that physically associates with the writer and acts as its substrate receptor (Figure 5D). A good

example of the last-mentioned mechanism is provided by some multi-protein ubiquitin E3 ligase complexes, such as cullin-RING E3 ubiquitin ligases.^[133,155,156] Interestingly, viruses can derail the host's ubiquitin signalling by encoding an alternative substrate receptor that directs an endogenous E3 ubiquitin ligase to a novel substrate.^[157]

One particularly well-studied class of docking motifs are degrons, which target proteins for ubiquitylation by the combined action of E3 and E2 enzymes.^[40] Degrons are often located on protein termini,^[38,156,158,159] although they can be internal as well. The interaction of some degrons with ubiquitin E3 ligases is regulated by other PTMs such as phosphorylation or acetylation, which can either prevent or promote binding to the ligase,^[158,160] providing an example of PTM interdependence. An unusual mechanism for substrate selection has been reported for the bacterial protein arginine kinase McsB, which marks proteins for degradation in some bacteria. This non-canonical kinase does not appear to select substrates via degron sequences but instead forms an oligomeric cage with a narrow entrance, which excludes bulky, folded proteins.^[161]

Another interesting case of substrate targeting is represented by histone-modifying and demodifying enzymes, which are often first recruited to the nucleosome, for instance by binding to the H2A-H2B acidic patch using an arginine-containing anchoring sequence^[162] (Figure 5E). Here, it is the enzyme that develops a simple motif to engage a pocket on the substrate – a reverse of the above examples where the enzyme recognised simple motifs in a substrate. Additionally, histone writers and erasers can contain reader domains for recognising PTM signals present on the nucleosome,^[163] resulting in reading–writing coupling that again exemplifies a functional interplay between different PTMs.

In addition to mechanisms reliant on (direct or indirect) physical contacts between a writer and a specific substrate, some enzymes define their range of targets simply through co-localisation in the same cellular location. Such 'spray-like' modification of pools of proximal proteins has been postulated for SUMOylation, acetylation and ADP-ribosylation writers.^[164–166] Co-localisation to the same locale can be achieved through the binding of both the writer and its substrates to the membrane or DNA (Figure 5F). Thus, the main ADP-ribosylating writer PARP1, which physically interacts with DNA breaks, appears to modify substrates that associate with DNA or nucleosomes in the vicinity of DNA breaks.

Lastly, it should be mentioned that, at the structural level, the majority of the mechanisms mentioned above rely on structural complementarity between linear motifs embedded in flexible or intrinsically disordered regions of substrates and either the active site or dedicated pockets within writer complexes. This can be illustrated by the SUMOylation consensus sequence found within SUMOylation substrates, which binds to the active site of the SUMOylation writer UBC9 (Figure 5G), and by docking motifs bound to a substrate-recruiting pocket distinct from the active site found in the protein kinase ERK2 (Figure 5H). In (relatively rare) cases where the recognised motifs are in a folded substrate region, they might not be linear but instead be assembled in space out of non-contiguous residues.^[167] While PTM writers tend to target multiple substrates by recognising relatively

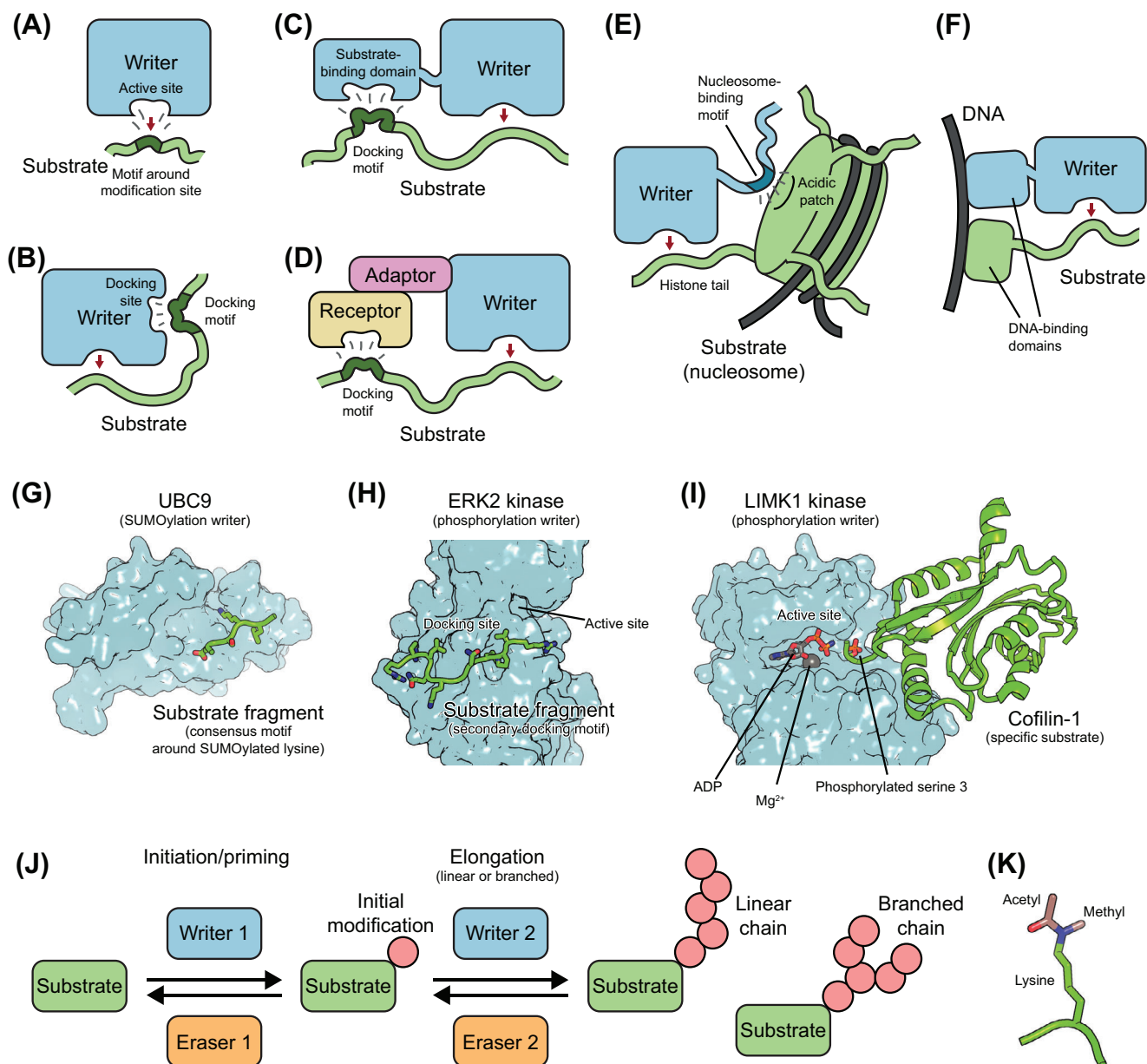


FIGURE 5 Mechanisms of substrate selection by PTM writers. (A)–(D) Schematic illustrations of selected mechanisms through which a writer enzyme recognises its substrates. The writer can use its active site to recognise a specific sequence motif surrounding the modified residue (A). Additionally or alternatively, it can utilise a secondary site located on the same domain as the active site (B), on another domain (C) or on a separate receptor subunit (D). The secondary substrate-docking site recognises a specific docking motif on the substrate distinct from the modified site. The modification site is indicated with a red arrow. (E) A schematic representation of recognition of nucleosome by a PTM writer through a motif that docks to the acidic patch composed of residues from histones H2A and H2B. (F) A schematic representation of co-localisation of a PTM writer and its substrate through the binding of both elements to DNA. (G) A three-dimensional structural representation of a SUMOylation consensus motif derived from a model SUMOylation substrate, RANGAP1, bound to the active site of the SUMOylation writer, the E2 protein UBC9. The figure was made using a fragment of PDB entry 2GRN. (H) A three-dimensional structural representation of a docking motif ('kinase-interaction motif') of MKP3 bound to a secondary docking site at the back of the protein kinase domain of ERK2. PDB entry 2FYS. (I) A three-dimensional structural representation of the interaction between the protein kinase LIMK1 and its specific substrate, cofilin-1, which results in the positioning of the main modification site, serine 3, in the kinase active site. The structure shows a post-reaction state in which serine is phosphorylated and ADP is still bound to the active site. Note an extensive interaction surface that ensures specificity for this particular substrate and site. PDB entry 5HVK. (J) A schematic representation of two steps needed for the formation of chains of modifiers such as ADP-ribosyl, ubiquityl or SUMOyl. The initiation/priming and elongation steps may be catalysed and reversed by distinct writers and erasers. The formation of linear or branched chains is shown. (K) A three-dimensional structural representation of the composite acetyl-methyl modification of a lysine residue. PDB entry 8SB6.

simple motifs, there are also more extreme examples of very specific enzymes that target a single main substrate or a small set of substrates. As a rule, such enzymes tend to develop a larger binding surface that facilitates discrimination of their cognate substrate and specific site within them^[149,168,169] (Figure 5I).

Chains and hybrids

Unlike single-step modifications such as phosphorylation, some PTMs – including glycosylation, ubiquitylation and other ubiquitin-like modifications, and ADP-ribosylation – can be attached to a protein substrate as linear or branched polymers, which are assembled prior to an *en bloc* transfer onto a protein or, more commonly, produced by repeated rounds of ligation.^[170–172]

In the latter case, chain formation can be divided into two distinct steps: initiation or priming (attachment of the first PTM unit to a protein) and elongation (attachment of succeeding units to preceding ones), which are sometimes controlled by distinct enzymatic activities (both in terms of reading and writing) (Figure 5J). This appears to be the case for serine ADP-ribosylation, where the initial attachment is performed by the HPF1:PARP1 complex (and reversed by the eraser ARH3), whereas elongation is catalysed by PARP1 alone (and reversed primarily by PARG).^[173] Similarly, for ubiquitin, distinct E2s can be responsible for the initiation and elongation stages.^[174] Erasers that process chains can employ either *endo* (within the chain) or *exo* (at chain termini) cleavage.^[175]

Importantly, different linkages between repeating units in chains can lead to structurally and functionally distinct signals, as best understood for poly(ubiquityl)ation and encapsulated in the idea of the ubiquitin code^[67] mentioned earlier.

As a final addition to this trend, recent research has provided examples of composite or hybrid modifications in which a modification attached to a protein substrate is itself modified by another modification, as in the case of phosphorylated or acetylated ubiquitin,^[68] hybrid chains of various ubiquitin-like proteins^[176] or mixed ADP-ribose-ubiquitin signals^[123] (the last ones detected only *in vitro* so far). A recently reported combination of acetylation and methylation on the same lysine represents a further case of a combined PTM^[177] (Figure 5K). Such PTM hybrids – more of which likely remain to be discovered – can potentially increase the complexity of functional signals encoded by PTMs.

Reversibility

PTMs differ in how transient or stable they are in the cell. Some stable modifications – notably many canonical cases of lipidation and glycosylation – might be best described as final maturation steps needed for proper localisation and constitutive function of a protein. Other PTMs – including phosphorylation, methylation, acetylation, ADP-ribosylation, ubiquitylation, SUMOylation or a specific nuclear/cytoplasmic type of glycosylation termed O-GlcNAcylation

(the attachment of O-linked N-acetylglucosamine) – tend to be reversible.

Enzymatic reversibility of a PTM relies on the existence of opposing eraser enzymes that counteract the action of writers. Importantly, although we often speak of erasers ‘reversing’ the PTMs, they generally do not recreate the donor molecule used in the forward reaction. It could be argued that the chemical nature of canonical PTMs explained above – where an electrophilic element is added to a nucleophilic protein side-chain – already implies the chemical possibility of detachment, most simply by hydrolysis, which is the common mechanism used by eraser enzymes (Figure 6A). The enormous concentration of water molecules thermodynamically drives the reaction. Hydrolysis can take place in a single step – as in the case of canonical protein serine/threonine phosphatases^[178] and some ADP-ribosyl hydrolases^[179] – the water molecule directly attacking the protein-ligated modification. Catalysis of such hydrolytic reactions usually consists of exposing the modification to the attack and activating the water molecule with metal ions and catalytic residues. Alternatively, the modification can be transferred onto a nucleophilic residue on the eraser (for example, a cysteine on tyrosine phosphatases^[178]) prior to being attacked by water (Figure 6B). Both types of mechanisms are also observed for deubiquitinating enzymes, where ubiquityl cleavage proceeds either in a single step, using a metal-activated water molecule, or through a serine/cysteine protease-like mechanism involving a covalent intermediate analogous to that for tyrosine phosphatases.^[175]

Modifications that are not intrinsically electrophilic and are efficiently installed owing to a particularly good leaving group in the donor – which is the case for ADP-ribosylation, but even more so, for methylation – might be more difficult to detach, possibly explaining why ADP-ribosylation removal sometimes involves unusual hydrolytic pathways^[179] and demethylation proceeds via complex, oxidation-dependent mechanisms.^[180] In fact, as mentioned earlier, protein methylation has initially been considered irreversible and it was not until early 2000s that first protein demethylases have been discovered.^[62,63] Furthermore, acetylation can be removed both through metal-catalysed hydrolysis and – in the case of acetylation erasers called sirtuins – through a more complex mechanism reliant on NAD as a co-substrate.^[181]

Although, at the molecular level, PTM erasers counteract the reaction catalysed by writers, their role is not simply limited to extinguishing responses that were elicited by the modification. Writers and erasers are often active, at least to some extent, simultaneously, which might appear as a ‘futile cycle’ that wastes donor molecules, but can, in fact, serve positive roles, such as fine-tuning the responsiveness of the system or allowing proofreading.^[182] The proofreading role, fulfilled in this case by deubiquitinases, might be particularly important for regulating ubiquitylation because of its potential to commit a protein to irreversible degradation.^[183,184]

Together with other mechanisms, such as sequential activation or inhibition of one enzyme by another, writer–reader competition can create sophisticated regulatory tools (switches, feedback and feedforward loops, clocks), as best seen during the cell cycle, where networks

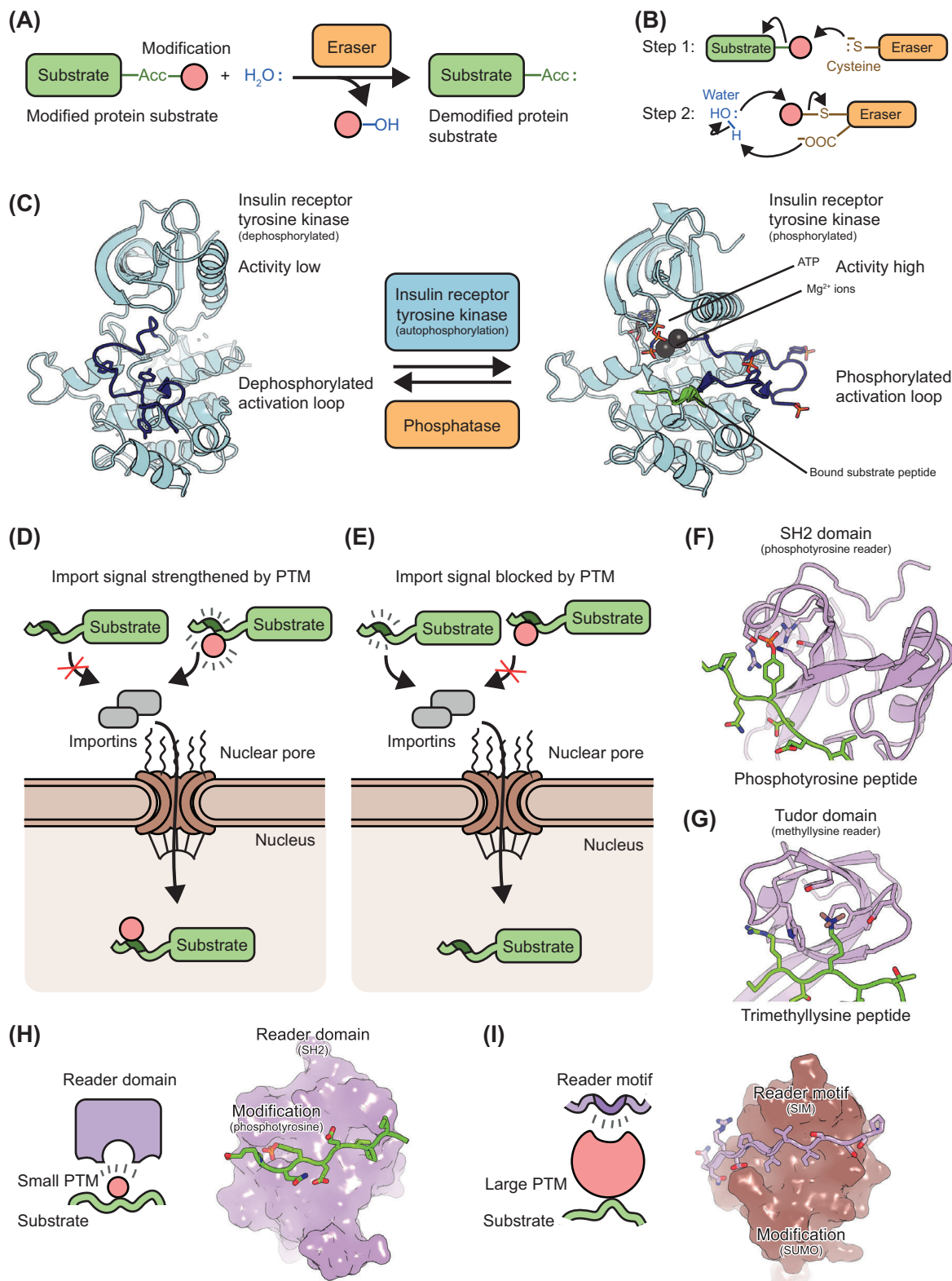


FIGURE 6 Regulation and functional consequences of PTMs. (A) A schematic representation of protein demodification through hydrolysis. A hydroxyl moiety from water becomes attached to the modification instead of the acceptor protein amino-acid residue. Proton abstraction from water is not shown. Note that not all PTMs are erased through this simple mechanism. (B) A schematic representation of a two-step hydrolysis mechanism used by some erasers (protein tyrosine phosphatases and some deubiquitinases), in which the modification is first transferred onto a reactive cysteine or serine residue in the eraser before being hydrolysed. (C) A three-dimensional structural representation of a phosphorylation-dependent activity switch in protein kinases. The protein tyrosine kinase domain of insulin receptor in its dephosphorylated and phosphorylated forms is shown based on PDB entries 1IRK and 1IR3. The activation loop is indicated in dark blue. Three tyrosine residues on this loop are seen modified in the phosphorylated form, which primes the kinase for substrate binding (a substrate peptide and ATP are shown) and

of kinases and phosphatases drive dynamic transitions between subsequent stages.^[185]

Molecular effects of PTMs

Proteins are remarkably sensitive systems. We know from countless mutational studies that even subtle alterations to a protein's chemical make-up, such as a single well-placed amino-acid substitution, can, in some cases, tip the delicate balance between different protein conformations, catalytically proficient/deficient enzymatic environments, or the presence/absence of an interaction with a partner. This sensitivity makes proteins inherently regulatable. Indeed, similarly to what is observed with mutations, the introduction or removal of even the smallest covalent PTM can profoundly alter protein properties. However, unlike mutations, PTMs are reversible and tightly controlled.

It is worth noting that bulkier and charged groups are predicted to have a more substantial impact on structure and function than smaller or neutral ones, which could explain the existence of many large PTMs and, among small chemical groups, the prominence of charged protein phosphoryl.^[186] Disrupting interactions through steric or electrostatic repulsion is arguably 'easier' with a larger or charged modification, and the same goes for creating new intra- or intermolecular interactions, with larger modifications potentially creating more new contacts and burying larger surfaces upon binding. Particularly large PTMs – especially those that can make chains, such as ubiquitylation and ADP-ribosylation – can provide platforms for the simultaneous recruitment of multiple components.^[48]

The basic molecular mechanism by which PTMs can exert their functions boils down to induction or disruption of either intra-molecular interactions (altering the protein's structure and dynamics) or intermolecular interactions. In the latter case, the affected interactions could be homotypic (oligomerisation of the modified protein) or heterotypic (interactions of the modified protein with other proteins or non-proteinaceous molecules such as DNA, membranes or small-molecule ligands). A PTM's effect on intra- and inter-molecular contacts can in turn translate into at least three different PTM-dependent functional outcomes: modulation of enzymatic activity, changes in sub-cellular localisation (including localisation to specific compartments and formation of molecular complexes) or altered half-life.

Two examples of PTM-dependent regulation of enzymatic activity were already discussed in Section 'Protein phosphorylation as a regulatory switch' in the context of sugar metabolism, where both glycogen phosphorylase (Figure 2A) and glycogen synthase change their enzy-

matic activity as a function of their phosphorylation status. Another example of enzyme regulation by phosphorylation concerns canonical protein kinases themselves, which are often activated through phosphorylation of the so-called activation loop or segment, which is located near the active site. Activation loop phosphorylation typically happens on a serine or threonine residue in serine/threonine kinases and tyrosine residues in tyrosine kinases and can proceed either in *cis* (which is controversial), where a kinase molecule modifies its own activation loop, or in *trans* (which is more established). In *trans*, phosphorylation can be catalysed by another copy of the same kinase or by an upstream kinase in a cascade. Upon phosphorylation, the activation loop adopts a conformation that supports substrate binding and catalytic activity. A classic example of this mechanism is provided by the insulin tyrosine kinase receptor, for which crystal structures in the dephosphorylated and phosphorylated state have long been available, allowing visualisation of this functional switch^[187,188] (Figure 6C).

PTM-induced changes in sub-cellular localisation can be both dramatic (the modified protein is found preferentially in a different compartment than the unmodified one) and subtle (the modified protein is found within the same compartment, but co-localised or not with different specific partners). A change in the global pattern of sub-cellular distribution is often related to a PTM's effect on the modified protein's interaction with cellular machinery that regulates passage between compartments, especially nuclear import and export pathways. Many instances of phosphorylation of residues located in the vicinity of nuclear import/export signals resulting in either increased or decreased nuclear localisation are known^[189,190] (Figure 6D and 6E). In some cases, phosphorylation does not directly affect the signal but instead triggers a conformational change that exposes an otherwise inaccessible signalling sequence. An alternative mechanism for regulating compartmental localisation is by affecting interactions that are crucial not for transport but for retaining the protein in a given compartment. A further potential mode of regulation is related to changes in solubility and condensation properties of a protein upon modification.^[191]

PTMs often trigger more subtle changes in localisation, such as recruitment or not to particular binding partners. While disruption of intermolecular interactions by PTMs typically relies on a steric or charge conflict between the modification and the would-be binding partner, induction of new interactions by PTMs is a more specific process where the binding partner has to preferentially recognise the modified protein form. This recognition is often mediated by specific reader domains that specialise in binding to a particular PTM, such as SH2 (Src homology 2) domains for phosphotyrosine, 14-3-3,

catalytic activity. (D) and (E) A schematic representation of selected mechanisms by which PTMs can regulate nuclear import. The nuclear import signal can be PTM-dependent, requiring strengthening by a PTM to confer efficient translocation through the pore (D). Alternatively, the nuclear import signal can be constitutively active but be weakened or blocked upon modification (E). (F) and (G) A three-dimensional structural representation of 'reading' of PTMs by specific domains. The recognition of a phosphotyrosine-containing peptide by the SH2 domain of LCK and of a trimethyllysine-containing histone H3 peptide by the double tudor domain of KDM4A is shown. PDB entries 1LCJ and 2GFA. (H) and (I) A comparison of domain-mediated and peptide-mediated PTM 'reading'. For each situation, a schematic and a corresponding illustrative structure are shown. A modification that is small in size is typically recognised by folded domains that contain a pocket which can surround the modification and its flanking sequence, increasing the interaction surface (H). In contrast, a large modification such as SUMO can itself contain a pocket for binding a linear motif on a reader (I). PDB entries 1LCJ and 6JXW. Acc, acceptor.

WW (named after conserved tryptophane residues), or FHA (forkhead-associated) domains for serine or threonine phosphorylation, tudor domains for methylated lysine or arginine, chromodomains for methylated lysine, bromodomains for acetylated lysine, macrodomains for ADP-ribosylation and so forth.^[192-194] The reader domains specific for small modifications typically recognise not only the modification itself but also the rest of the modified amino-acid sidechain, sometimes together with the surrounding sequence. Nonetheless, the modification makes a key contribution to binding, rendering the interaction PTM-dependent. This can be seen for the interaction between an SH2 domain and a phosphotyrosine peptide, where the anionic phosphoryl moiety in the modified peptide is surrounded by three positively charged arginine or lysine residues and two serine hydrogen-bond donors in the reader (Figure 6F). Methyllysine-recognising tudor domains, on the other hand, surround the methylated lysine side-chain with an aromatic cage (Figure 6G). In the case of poly(ubiquityl), different reader domains or motifs can distinguish between different types of ubiquitin chains (different inter-ubiquitin linkages).^[67] The specificity of readers not only for the PTM type but also for the particular modified substrate can come from the reader domains favouring particular sequence motifs around the PTM site or from secondary interactions between other elements in the reader protein and in the modified substrate. In addition to folded reader domains, some modifications can be recognised by simple reader motifs, as discussed at the end of Section 'Evolutionary development of PTM systems' below.

In the cases where a modification promotes physical interaction with proteins that have proteolytic functions – as with ubiquitin-dependent targeting to the 26S proteasome discussed in Section 'Protein phosphorylation as a regulatory switch' (Figure 2B) or similar mechanisms dependent on Pup (prokaryotic ubiquitin-like protein; note that this is an analogue, not a homologue, of ubiquitin) and phosphoarginine in bacteria – modification can act as a trigger of protein degradation. The self-compartmentalised proteases involved in these processes contain reader domains that recognise the appropriate PTMs, with at least three different ubiquitin receptors present within the eukaryotic 26S proteasome^[195] and 12 repeats of a similar phosphoarginine-binding site in the bacterial protease ClpCP.^[43] One step upstream of the degradation signal itself, eukaryotic proteolysis can be promoted or inhibited by PTMs that either increase or decrease a substrate's interaction with an E3 ubiquitin ligase.^[160]

EVOLUTIONARY LOGIC OF EMERGENCE AND EXPANSION OF PTM SYSTEMS

When discussing evolutionary 'rationale' for the emergence of specific PTMs and for various aspects of the PTM reactions (for instance, the use of specific donors), it is important to bear in mind that the evolution of a new process does not simply tend towards a theoretically optimal solution to a specific problem. Thus, it is not possible to claim that the existing PTMs are an all-optimal tool for intracellular signalling or, more generally, protein regulation (although it should be possible to argue that they are well-suited for this task – as I have in Section 'Molecu-

lar effects of PTMs' above). Evolving processes, shaped as they are by historical contingency and entrenchment, should be seen in the context of specific evolutionary histories and potentialities and limitations of resources at hand. As François Jacob remarked, evolution proceeds by 'tinkering' and is 'a matter of using the same elements, of adjusting them, of altering here or there, of arranging various combinations to produce new objects.'^[196] It is within such a framework that I discuss some hypotheses about PTM emergence and expansion below.

The conflict context hypothesis

L. Aravind Iyer and co-workers trace back the evolutionary origin of several key regulatory PTMs, alongside nucleic acid modifications, to a phase in the history of prokaryotes, which might overlap with a period in Earth's history known as the great oxygenation event and be a key to the origin of eukaryotes.^[197,198] According to this scenario, following the development of more basic metabolic processes, competition within and between prokaryotic species as well as between prokaryotes and phages led to the rapid development of systems for attacking and defending, which relied to a large extent on enzymatic modification of small-molecules, nucleic acids and proteins. Modifications could serve to meddle with the enemy's physiology (for example by blocking essential activities or interaction) as well as to inactivate the enemy's weapons (by modifying and thus inactivating an antibiotic, a protein toxin or an infectious nucleic acid). In Aravind's account, increasing evolutionary pressure characteristic of the conflict context could account for the explosive innovation in terms of new catalytic activities, many of which catalysed the transfer of chemical groups as a means of interfering with an opponent or protecting from their interference. Such a context could also explain the emergence of PTM erasers as 'anti-toxins' that counteract PTMs. The rich enzymatic repertoire of writers and erasers would later be re-used and developed for regulatory, and especially epigenetic, purposes by prokaryotes themselves and especially emerging eukaryotes as an example of a '«peacetime» use of «wartime» inventions.'^[197] It could be added that the use of PTMs for attack and defence is still widely observed in the living world. For instance, some of the notorious bacterial toxins (including cholera and diphtheria) are protein ADP-ribosylating enzymes that target the essential processes of a host. In another, more recent, example, a human E3 ligase RNF213 is involved in ubiquitylating invading intra-cellular bacteria as part of a defensive response.^[199]

The ancient conflict context could conceivably explain not only *that* many potential writer activities developed but also some of the features of the protein writer enzymes. For example, these enzymes, while quite specific for donors, can often target various substrates and relatively rapidly change their substrate preference on an evolutionary timescale, even switching between protein and non-protein substrates. In fact, we begin to see that modifications traditionally thought to be confined to proteins such as ADP-ribosylation or ubiquitylation do also target non-protein substrates.^[199-209] It is conceivable that, on the whole, similar modifications can be ligated to proteins, nucleic acids, and possibly, lipids or sugars, but we fail to appreciate this commonality

due to insufficient research into non-protein molecules and technical limitations.

New use for ‘old’ donors and mechanisms

Whether the conflict scenario of PTM origin is true or not, PTMs did not develop in a vacuum but in cells that already contained certain resources (genes, chemicals), and these ‘starting conditions’ must be included in an evolutionary account of PTM origins. In particular, cells already possessed suitable donor molecules, which were present there for other reasons. Indeed, as already mentioned, donors used for many key PTMs (ATP, SAM, NAD, acetyl-coenzyme A, lipid-coenzyme A, GDP-sugars etc.) are also used in primary metabolism, whether as donors for metabolite modification or as ‘energy molecules’ (as with ATP and NADH/NAD).^[132] Some of these core compounds might have originally emerged through spontaneous reactions in a primordial environment (as proposed for SAM^[210]), but by the time enzymatic PTMs became prominent regulatory mechanisms, these molecules were likely maintained at relatively high levels in the cell through dedicated biosynthetic and salvage pathways. Relying on the same compounds as the ones used for energy homeostasis has an added advantage of allowing for a direct crosstalk between the metabolic state of the cell and protein regulation by PTMs.^[211]

One prominent exception to the repurposing of ancient metabolic donors for PTMs is ubiquitylation and other ubiquitin-like PTMs. But even here, a suitable donor – a C-terminally activated protein modifier – arguably first evolved for another purpose, to serve as a ‘sulphur carrier protein’ in sulphur transfer reactions. This hypothesis is supported by the similarity between components of the eukaryotic ubiquitin and prokaryotic sulphur carrier systems^[212–215] and, even more strikingly, by the case of the eukaryotic protein Urm1, a ubiquitin-like ‘molecular fossil’ that can act both as a sulphur carrier and a protein modification.^[216–220]

A different example of repurposing an existing resource (this time – an enzymatic activity) for a PTM reaction is provided by the Pupylation pathway of *Mycobacterium tuberculosis*.^[42] Here, the writer enzyme PafA, which ligates a small intrinsically disordered protein called Pup to substrates, is related to metabolic enzymes including glutamine synthetase, which catalyse the attachment of a glutamyl moiety to an amino group in a biosynthetic pathway.^[221] PafA uses the same chemistry to link the Pup modification, via a C-terminal glutamate, to a lysine side-chain on a substrate.^[222–224] A further case of potential repurposing is represented by the possible evolution of enzymatic activities for histone lysine methylation and acetylation from enzymes that catalyse corresponding modifications of polyamines such as spermine and spermidine, cationic compounds widely found in living cells.^[225]

Chemical driving forces

Apart from the historical context and the available resources, the course of evolution is shaped by the inherent reactivities of potential

PTM acceptors and donors. Although – as stressed above – chemical groups found in biological systems are relatively inert by conventional chemical standards, some are more likely than others to engage in chemical reactions. At initial stages of PTM evolution, when relevant enzymatic activities only began to emerge or were switching from a non-protein to a protein substrate, catalysis was likely weak, making the intrinsic substrate reactivities relatively more important in the process of emergence of a new PTM. Thus, acceptor residues that are more inherently nucleophilic and donor molecules that contain an electrophilic modification and/or a good leaving group are, all other things being equal, more likely to become substrates of an enzymatic reaction than do less reactive acceptors and donors. Indeed, some canonical PTM reactions are observed to proceed even without enzymes, and we could envision some of the most ancient PTM events to have been *quasi*-spontaneous reactions, happening, for example, in active sites of metabolic enzymes that interact with reactive donors.^[211] From the point of view of thermodynamics, the maintenance of high levels of the donor compounds and low levels of the corresponding free leaving groups by the cell – which likely preceded or coincided with the emergence of PTMs – could drive modification reactions as soon as they were kinetically permitted.

The final, most ‘classical’ part of the search for an evolutionary ‘logic’ of PTM emergence has to do with the physiological advantage of PTMs. The supposed existence of such an advantage is what promoted the retention and expansion of accidental genetic changes that supported PTM emergence and development. Here, the answer to the question whether PTM systems could be advantageous is the enormous regulatory potential of even the smallest covalent addition, discussed in more detail in Section ‘Molecular effects of PTMs’ above.

Evolutionary development of PTM systems

In addition to the initial emergence of PTMs, the current widespread use of protein modification by living organisms reflects the expansion and fine-tuning of PTM system in the course of natural history. Some major PTMs, notably phosphorylation and acetylation, are universally present across bacteria, archaea and eukaryotes.^[105,211] Other modifications – and some specific sub-types of those mentioned – emerged after the common ancestor of these three domains of life and exist only in some of the lineages, the best example being ubiquitylation, which is present in archaea and eukaryotes but not bacteria.^[105] Lastly, instances of PTM loss in particular lineages and horizontal transfer between lineages muddle a simple narrative of PTM evolutionary history. An example of a PTM with a complex history is protein ADP-ribosylation, which is present in species from all kingdoms of life, but apparently lacking in some notable species (such as yeast), and has spread at least in part through horizontal gene transfer.^[226]

Across kingdoms of life, PTMs tend to show two distinct expansion trends. On the one hand, we observe a diversity of PTMs related to the range of lifestyles and occupied habitats. This is particularly evident in bacteria, where various species differ widely in their PTM repertoires and represent a large PTM diversity when considered

collectively. On the other hand, we observe the expansion of PTM systems that is seemingly correlated with increasing organismal complexity and which might be linked to the challenging development and signal processing in complex, multi-cellular organisms. The latter trend has been at play in higher eukaryotes. Both trends can be illustrated by the evolution of protein phosphorylation. On the one hand, phosphorylation exhibits large diversity in bacteria, with not only canonical protein kinases that modify serine and threonine amino acids, but also non-canonical protein tyrosine, histidine and arginine kinases, which, generally, are not found in eukaryotes.^[56,227,228] On the other hand, the canonical protein kinase domain and the associated substrate, eraser and reader pools for hydroxyl amino-acid phosphorylation have expanded tremendously during the emergence and later evolution of eukaryotes. By the time the last eukaryotic common ancestor appeared, the set of different canonical protein kinases – presumably mostly targeting serine and threonine residues – seems to have already expanded to almost a hundred members,^[229] and later has further expanded, reaching several hundred in humans. One particular branch of the canonical protein kinase family whose development has been much studied is protein tyrosine kinases, which are not present in bacteria or yeast but emerged in higher eukaryotes by switching their acceptor specificity from serine/threonine to tyrosine (they are different from non-canonical bacterial tyrosine kinases). Wendell Lim and Tony Pawson have discussed how, following the individual emergence of three tyrosine-specific functions – a kinase, a phosphatase and a reader domain (SH2) – this new complete signalling ‘toolkit’ dramatically expanded, becoming a crucial means of cell-to-cell communication in metazoans.^[230] Interestingly, the expansion of phosphotyrosine signalling seems to have gone hand in hand with the decrease in tyrosine frequency in the proteome, possibly to avoid unintended phosphorylation.^[231] Another fascinating example related to the evolution of protein phosphorylation in eukaryotes is provided by the replacement, in the course of evolution, of some glutamate or aspartate residues in key positions of certain proteins with phosphorylatable serine, threonine or tyrosine residues.^[232] Phosphorylated versions of these residues can functionally mimic negatively charged glutamate/aspartate residues, allowing phosphorylation-dependent regulation of protein structure and function.

As illustrated by the development of phosphotyrosine signalling, the expansion of PTM systems in eukaryotes is closely related to the modular nature of eukaryotic proteins. Over the course of evolution, protein domains responsible for writing, erasing and reading modifications – once all three types are found together in one species, creating an advantageous functional toolkit – can be individually duplicated and re-combined in new arrangements with other domains, thus achieving specialisation for new functions and sub-cellular niches. This is powerfully illustrated by the human family of 17 PARP proteins, most of which are active as ADP-ribosylation writers.^[233] Each of these proteins contains the same writer domain, related to bacterial ADP-ribosylating toxins such as cholera and (especially) diphtheria toxin, but, in PARPs, these ancient domains have diverged in their catalytic and allosteric properties and became combined with other domains and motifs. As a result, distinct PARPs catalyse ADP-ribosylation on different acceptor residues and in different functional contexts ranging from DNA repair

to anti-viral immunity and localise to various possible cellular compartments. One of the human PARPs, PARP14, has recently been shown to be not only a writer, but also a substrate, a reader and an eraser of ADP-ribosylation, each function being associated with a different region or domain in this large multi-domain protein.^[234,235]

In addition to the evolution based on domain duplication, the expansion of PTM systems in eukaryotes benefits from some elements of these systems relying on short linear motifs (SLiMs), which are widespread in eukaryotic proteins.^[236] Due to their simplicity and location in intrinsically disordered regions that are tolerant to mutations, such motifs can relatively rapidly emerge both *ex nihilo* and through duplication, and then further adjust or disappear in the course of evolution, thus facilitating the expansion and fine-tuning of PTM systems of which they are part. In particular, this applies to PTM sites on substrates, which are enriched in loops or disordered regions,^[237] but it is also relevant to secondary docking sites on substrates such as degrons.

Fast evolving linear motifs might also, in some cases, be responsible for PTM reader or even writer functions, provided the modification is sufficiently large. This is related to the idea that, to achieve sufficient binding affinity, heterodimeric interactions typically require at least one relatively large partner, which – when physically interacting with a smaller molecule or motif – is able to increase the interaction surface by surrounding the small partner. Following this logic, PTMs that consist of small chemical groups tend to be recognised by protein domains with dedicated pockets (reader domains) that encapsulate the modification and its surrounding sequence (Figure 6H). In contrast, large modifications – particularly those involving a proteinaceous modifier – are able to form sufficiently strong interactions with short protein motifs. The best-studied example of this phenomenon is the ubiquitin-like protein SUMO, which can interact with short hydrophobic SUMO-interacting motifs or SIMs^[238–240] (Figure 6I). SIM motifs can, therefore, function as very simple SUMO reader modules. The distinction between domain- and motif-based PTM reading is important from an evolutionary point of view, considering the favourable evolutionary properties of linear motifs. It also has implications for reader detection, as domains are easily found through sequence or structural homology, but functional linear motifs are more challenging to conclusively identify with bioinformatic tools alone. Hydrophobic linear motifs such as SIMs might emerge particularly rapidly during protein evolution, as hydrophobic substitutions are apparently favoured during random mutagenesis due to their codon composition.^[241] Linear motifs have also been reported for ADP-ribosylation binding^[194] and helical, but simple, motifs are known for ubiquitin recognition.^[192] Of note, a particular SIM-linker-SIM region found in zinc-finger protein 451 (ZNF451) is sufficient to confer on this human protein a SUMO E3 ligase activity,^[242] showing that short motifs can also participate in writing of some PTMs.

Concluding remarks

The field of protein PTMs has evolved from small areas of research dedicated to the regulation of specific processes (initially sugar metabolism) to an enormous and diverse discipline that touches on

every aspect of molecular biology across all kingdoms of life. With the number of known PTMs exceeding 650, there is apparently no end to nature's creativity in modifying protein structure with additional elements that can serve regulatory roles. The mere fact that so many PTMs exist already suggests that they might serve nonredundant functions and cannot be exhaustively described in general terms. In the end, PTMs are not generic 'on' and 'off' signals but particular chemical groups, with their physicochemical properties, and, similarly, proteins that regulate them are complex entities with individual structural and mechanistic features. The purpose of the above essay was not to obscure this diversity, but to point to some analogies that can be found, despite the diversity, due to convergent evolution. The appreciation of analogies between diverse systems might help relate the knowledge obtained for one system to open questions in another. Indeed, the history of the PTM field is a good example of how concepts developed in one context stimulate formulating hypotheses in another, and how different experimental and theoretical approaches combine in formulating a more complete picture. With the current tremendous advancements in available techniques, it can only be wished that this cross-fertilisation should continue and develop in the future.

ACKNOWLEDGEMENTS

I acknowledge the contribution of, and apologise to, all researchers in the PTM community whose research and ideas have not been covered in this text or have been unsatisfactorily discussed. I thank Philip Cohen, Tim Clausen and Ivan Ahel, who have introduced me to the world of different PTMs. I thank the anonymous reviewers for their valuable feedback, Ivo Alexander Hendriks for advice on PTM proteomics, and my colleagues Bertrand Castaing, Stéphane Goffinont, Franck Coste, Lucija Mance and Aanchal Mishra for the discussions on related topics. The responsibility for any remaining errors lies with me. The initial version of the essay was based on seminars that I have given in various places and I thank the participants for all the questions asked on these occasions. My research is financially supported by the European Union's Horizon Europe Research and Innovation Programme (ERC Starting Grant 'SUMOwriteNread', no 101078837), La Ligue contre le Cancer and the Centre National de la Recherche Scientifique (CNRS). I am an associate fellow of Le Studium Loire Valley Institute of Advanced Studies and the ATIP-Avenir programme.

CONFLICT OF INTEREST STATEMENT

The author declares no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new datasets were generated or analysed during the current study.

ORCID

Marcin J. Suskiewicz  <https://orcid.org/0000-0002-3279-6571>

REFERENCES

- Klein, T., Eckhard, U., Dufour, A., Solis, N., & Overall, C. M. (2018). Proteolytic cleavage—mechanisms, function, and “omic” approaches for a near-ubiquitous posttranslational modification. *Chemical Reviews*, *118*, 1137–1168.
- Mondal, S., & Thompson, P. R. (2021). Chemical biology of protein citrullination by the protein A arginine deiminases. *Current Opinion in Chemical Biology*, *63*, 19–27.
- Schmidpeter, P. A. M., & Schmid, F. X. (2015). Prolyl isomerization and its catalysis in protein folding and protein function. *Journal of Molecular Biology*, *427*, 1609–1631.
- Harmel, R., & Fiedler, D. (2018). Features and regulation of non-enzymatic post-translational modifications. *Nature Chemical Biology*, *14*, 244–252.
- Zhong, Q., Xiao, X., Qiu, Y., Xu, Z., Chen, C., Chong, B., Zhao, X., Hai, S., Li, S., An, Z., & Dai, L. (2023). Protein posttranslational modifications in health and diseases: Functions, regulatory mechanisms, and therapeutic implications. *MedComm*, *4*, e261.
- Krebs, E. G., Perry, S. V., & Cohen, P. (1997). Historical perspectives on protein phosphorylation and a classification system for protein kinases. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *302*, 3–11.
- Cohen, P. (2002). The origins of protein phosphorylation. *Nature Cell Biology*, *4*, E127–E130.
- Cori, G. T., & Green, A. A. (1943). Crystalline muscle phosphorylase: II. prosthetic group. *Journal of Biological Chemistry*, *151*, 31–38.
- Cori, G. T., & Cori, C. F. (1945). The enzymatic conversion of phosphorylase a to b. *Journal of Biological Chemistry*, *158*, 321–332.
- Krebs, E. G., & Fischer, E. H. (1956). The phosphorylase b to a converting enzyme of rabbit skeletal muscle. *Biochimica et Biophysica Acta*, *20*, 150–157.
- Krebs, E. G., Graves, D. J., & Fischer, E. H. (1959). Factors affecting the activity of muscle phosphorylase b kinase. *Journal of Biological Chemistry*, *234*, 2867–2873.
- Sutherland, E. W., & Wosilait, W. D. (1955). Inactivation and activation of liver phosphorylase. *Nature*, *175*, 169–170.
- Rall, T. W., Sutherland, E. W., & Berthet, J. (1957). The relationship of epinephrine and glucagon to liver phosphorylase: IV. effect of epinephrine and glucagon on the reactivation of phosphorylase in liver homogenates. *Journal of Biological Chemistry*, *224*, 463–475.
- Fischer, E. H., Graves, D. J., Crittenden, E. R. S., & Krebs, E. G. (1959). Structure of the site phosphorylated in the phosphorylase b to a reaction. *Journal of Biological Chemistry*, *234*, 1698–1704.
- Sprang, S. R., Acharya, K. R., Goldsmith, E. J., Stuart, D. I., Varvill, K., Fletterick, R. J., Madsen, N. B., & Johnson, L. N. (1988). Structural changes in glycogen phosphorylase induced by phosphorylation. *Nature*, *336*, 215–221.
- Levene, P. A., & Alsberg, C. L. (1906). The cleavage products of vitellin. *Journal of Biological Chemistry*, *2*, 127–133.
- Lipmann, F. A., & Levene, P. A. (1932). Serinephosphoric acid obtained on hydrolysis of vitellinic acid. *Journal of Biological Chemistry*, *98*, 109–114.
- Burnett, G., & Kennedy, E. P. (1954). The enzymatic phosphorylation of proteins. *Journal of Biological Chemistry*, *211*, 969–980.
- Walsh, D. A., Perkins, J. P., & Krebs, E. G. (1968). An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle. *Journal of Biological Chemistry*, *243*, 3763–3765.
- Friedman, D. L., & Lerner, J. (1963). Studies on UDPG- α -glucan transglucosylase. III. Interconversion of two forms of muscle UDPG- α -glucan transglucosylase by a phosphorylation-dephosphorylation reaction sequence*. *Biochemistry*, *2*, 669–675.
- Cohen, P. (2006). The twentieth century struggle to decipher insulin signalling. *Nature Reviews Molecular Cell Biology*, *7*, 867–873.
- Hunt, T. (1993). *The cell cycle: An introduction*. WH Freeman.
- Knighton, D. R., Zheng, J., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S., & Sowadski, J. M. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*, *253*, 407–414.

24. Mingione, V. R., Paung, Y., Outhwaite, I. R., & Seeliger, M. A. (2023). Allosteric regulation and inhibition of protein kinases. *Biochemical Society Transactions*, 51, 373–385.
25. Hershko, A., Ciechanover, A., & Varshavsky, A. (2000). The ubiquitin system. *Nature Medicine*, 6, 1073–1081.
26. Ciechanover, A. (2009). Tracing the history of the ubiquitin proteolytic system: The pioneering article. *Biochemical and Biophysical Research Communications*, 387, 1–10.
27. Varshavsky, A. (2006). The early history of the ubiquitin field. *Protein Science*, 15, 647–654.
28. Ciechanover, A., Hod, Y., & Hershko, A. (1978). A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochemical and Biophysical Research Communications*, 81, 1100–1105.
29. Hershko, A., Ciechanover, A., & Rose, I. A. (1979). Resolution of the ATP-dependent proteolytic system from reticulocytes: A component that interacts with ATP. *Proceedings of the National Academy of Sciences*, 76, 3107–3110.
30. Hershko, A., Ciechanover, A., Heller, H., Haas, A. L., & Rose, I. A. (1980). Proposed role of ATP in protein breakdown: Conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proceedings of the National Academy of Sciences*, 77, 1783–1786.
31. Ciechanover, A., Elias, S., Heller, H., Ferber, S., & Hershko, A. (1980). Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes. *Journal of Biological Chemistry*, 255, 7525–7528.
32. Wilkinson, K. D., Urban, M. K., & Haas, A. L. (1980). Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *Journal of Biological Chemistry*, 255, 7529–7532.
33. Goldknopf, I. L., & Busch, H. (1977). Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. *Proceedings of the National Academy of Sciences*, 74, 864–868.
34. Ciechanover, A., Elias, S., Heller, H., & Hershko, A. (1982). 'Covalent affinity' purification of ubiquitin-activating enzyme. *Journal of Biological Chemistry*, 257, 2537–2542.
35. Hershko, A., Heller, H., Elias, S., & Ciechanover, A. (1983). Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *Journal of Biological Chemistry*, 258, 8206–8214.
36. Hershko, A., & Heller, H. (1985). Occurrence of a polyubiquitin structure in ubiquitin-protein conjugates. *Biochemical and Biophysical Research Communications*, 128, 1079–1086.
37. Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., & Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science*, 243, 1576–1583.
38. Bachmair, A., Finley, D., & Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science*, 234, 179–186.
39. Bachmair, A., & Varshavsky, A. (1989). The degradation signal in a short-lived protein. *Cell*, 56, 1019–1032.
40. Varshavsky, A. (1991). Naming a targeting signal. *Cell*, 64, 13–15.
41. Morange, M. (2020). *The black box of biology: A history of the molecular revolution*. Harvard University Press.
42. Pearce, M. J., Mintseris, J., Ferreyra, J., Gygi, S. P., & Darwin, K. H. (2008). Ubiquitin-like protein involved in the proteasome pathway of mycobacterium tuberculosis. *Science*, 322, 1104–1107.
43. Trentini, D. B., Suskiewicz, M. J., Heuck, A., Kurzbauer, R., Deszcz, L., Mechtler, K., & Clausen, T. (2016). Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature*, 539, 48–53.
44. Liao, Y., Sumara, I., & Pangou, E. (2022). Non-proteolytic ubiquitylation in cellular signaling and human disease. *Communications Biology*, 5, 1–15.
45. Rechsteiner, M. (1987). Ubiquitin-mediated pathways for intracellular proteolysis. *Annual Review of Cell Biology*, 3, 1–30.
46. Finley, D., & Chau, V. (1991). Ubiquitination. *Annual Review of Cell Biology*, 7, 25–69.
47. D'Andrea, A., & Pellman, D. (1998). Deubiquitinating enzymes: A new class of biological regulators. *Critical Reviews in Biochemistry and Molecular Biology*, 33, 337–352.
48. Hochstrasser, M. (2000). Evolution and function of ubiquitin-like protein-conjugation systems. *Nature Cell Biology*, 2, E153–E157.
49. Cohen, P., & Tcherpakov, M. (2010). Will the ubiquitin system furnish as many drug targets as protein kinases? *Cell*, 143, 686–693.
50. Békés, M., Langley, D. R., & Crews, C. M. (2022). PROTAC targeted protein degraders: The past is prologue. *Nature Reviews Drug Discovery*, 21, 181–200.
51. Uy, R., & Wold, F. (1977). Posttranslational covalent modification of proteins. *Science*, 198, 890–896.
52. Walsh, C. (2006). *Posttranslational modification of proteins: Expanding nature's inventory*. Roberts and Company Publishers.
53. Walsh, C. T., Garneau-Tsodikova, S., & Gatto, G. J. (2005). Protein posttranslational modifications: The chemistry of proteome diversifications. *Angewandte Chemie International Edition*, 44, 7342–7372.
54. Walsh, C. (1979). *Enzymatic reaction mechanisms*. W. H. Freeman.
55. Walsh, C. T. (2010). Catalysis at the intersection of biology, chemistry, and medicine. *Journal of Biological Chemistry*, 285, 29681–29689.
56. Macek, B., Forchhammer, K., Hardouin, J., Weber-Ban, E., Grangeasse, C., & Mijakovic, I. (2019). Protein post-translational modifications in bacteria. *Nature Reviews Microbiology*, 17, 651–664.
57. Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, 403, 41–45.
58. Rando, O. J. (2012). Combinatorial complexity in chromatin structure and function: Revisiting the histone code. *Current Opinion in Genetics & Development*, 22, 148–155.
59. Allfrey, V. G., Faulkner, R., & Mirsky, A. E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis*. *Proceedings of the National Academy of Sciences*, 51, 786–794.
60. Turner, B. M. (2000). Histone acetylation and an epigenetic code. *BioEssays*, 22, 836–845.
61. Jenuwein, T., & Allis, C. D. (2001). Translating the Histone Code. *Science*, 293, 1074–1080.
62. Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstone, J. R., Cole, P. A., Casero, R. A., & Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, 119, 941–953.
63. Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., & Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, 439, 811–816.
64. Egloff, S., & Murphy, S. (2008). Cracking the RNA polymerase II CTD code. *Trends in Genetics*, 24, 280–288.
65. Harlen, K. M., & Churchman, L. S. (2017). The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nature Reviews Molecular Cell Biology*, 18, 263–273.
66. Janke, C. (2014). The tubulin code: Molecular components, readout mechanisms, and functions. *Journal of Cell Biology*, 206, 461–472.
67. Komander, D., & Rape, M. (2012). The ubiquitin code. *Annual Review of Biochemistry*, 81, 203–229.
68. Yau, R., & Rape, M. (2016). The increasing complexity of the ubiquitin code. *Nature Cell Biology*, 18, 579–586.
69. Dikic, I., & Schulman, B. A. (2023). An expanded lexicon for the ubiquitin code. *Nature Reviews Molecular Cell Biology*, 24, 273–287.
70. Brenner, S. (2002). Life sentences: Detective Rummage investigates. *Genome Biology*, 3, comment1013.1–comment1013.2.
71. Gest, H. (2005). The early history of ³²P as a radioactive tracer in biochemical research: A personal memoir. *Biochemistry and Molecular Biology Education*, 33, 159–164.

72. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680–685.
73. Hunter, T., & Sefton, B. M. (1980). Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proceedings of the National Academy of Sciences*, 77, 1311–1315.
74. Ross, A. H., Baltimore, D., & Eisen, H. N. (1981). Phosphotyrosine-containing proteins isolated by affinity chromatography with antibodies to a synthetic hapten. *Nature*, 294, 654–656.
75. Comoglio, P. M., Di-Renzo, M. F., Tarone, G., Giancotti, F. G., Naldini, L., & Marchisio, P. C. (1984). Detection of phosphotyrosine-containing proteins in the detergent-insoluble fraction of RSV-transformed fibroblasts by azobenzene phosphonate antibodies. *Embo Journal*, 3, 483–489.
76. Hershko, A., Eytan, E., Ciechanover, A., & Haas, A. L. (1982). Immunochemical analysis of the turnover of ubiquitin-protein conjugates in intact cells. Relationship to the breakdown of abnormal proteins. *Journal of Biological Chemistry*, 257, 13964–13970.
77. van Wijk, S. J., Fulda, S., Dikic, I., & Heilemann, M. (2019). Visualizing ubiquitination in mammalian cells. *EMBO Reports*, 20, e46520.
78. Dasovich, M., & Leung, A. K. L. (2023). PARPs and ADP-ribosylation: Deciphering the complexity with molecular tools. *Molecular Cell*, 83, 1552–1572.
79. Haystead, T. a. J., Sim, A. T. R., Carling, D., Honnor, R. C., Tsukitani, Y., Cohen, P., & Hardie, D. G. (1989). Effects of the tumour promoter okadaic acid on intracellular protein phosphorylation and metabolism. *Nature*, 337, 78–81.
80. Witze, E. S., Old, W. M., Resing, K. A., & Ahn, N. G. (2007). Mapping protein post-translational modifications with mass spectrometry. *Nature Methods*, 4, 798–806.
81. Doll, S., & Burlingame, A. L. (2015). Mass spectrometry-based detection and assignment of protein posttranslational modifications. *ACS Chemical Biology*, 10, 63–71.
82. Oda, Y., Nagasu, T., & Chait, B. T. (2001). Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nature Biotechnology*, 19, 379–382.
83. Zhou, H., Watts, J. D., & Aebersold, R. (2001). A systematic approach to the analysis of protein phosphorylation. *Nature Biotechnology*, 19, 375–378.
84. Ficarro, S. B., McClelland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., & White, F. M. (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nature Biotechnology*, 20, 301–305.
85. Ordureau, A., Münch, C., & Harper, J. W. (2015). Quantifying ubiquitin signaling. *Molecular Cell*, 58, 660–676.
86. Sharma, K., D'Souza, R. C. J., Tyanova, S., Schaab, C., Wiśniewski, J. R., Cox, J., & Mann, M. (2014). Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Reports*, 8, 1583–1594.
87. Bekker-Jensen, D. B., Bernhardt, O. M., Hogrebe, A., Martinez-Val, A., Verbeke, L., Gandhi, T., Kelstrup, C. D., Reiter, L., & Olsen, J. V. (2020). Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nature Communications*, 11, 787.
88. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., & Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*, 325, 834–840.
89. Lundby, A., Lage, K., Weinert, B. T., Bekker-Jensen, D. B., Secher, A., Skovgaard, T., Kelstrup, C. D., Dmytriiev, A., Choudhary, C., Lundby, C., & Olsen, J. V. (2012). Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. *Cell Reports*, 2, 419–431.
90. Hansen, F. M., Tanzer, M. C., Brüning, F., Bludau, I., Stafford, C., Schulman, B. A., Robles, M. S., Karayel, O., & Mann, M. (2021). Data-independent acquisition method for ubiquitinome analysis reveals regulation of circadian biology. *Nature Communications*, 12, 254.
91. Hendriks, I. A., & Vertegaal, A. C. O. (2016). A comprehensive compilation of SUMO proteomics. *Nature Reviews Molecular Cell Biology*, 17, 581–595.
92. Hendriks, I. A., Lyon, D., Su, D., Skotte, N. H., Daniel, J. A., Jensen, L. J., & Nielsen, M. L. (2018). Site-specific characterization of endogenous SUMOylation across species and organs. *Nature Communications*, 9, 2456.
93. Hendriks, I. A., Larsen, S. C., & Nielsen, M. L. (2019). An advanced strategy for comprehensive profiling of ADP-ribosylation sites using mass spectrometry-based proteomics. *Molecular & Cellular Proteomics*, 18, 1010–1026.
94. Hendriks, I. A., Buch-Larsen, S. C., Prokhorova, E., Elsborg, J. D., Rebak, A. K. L. F. S., Zhu, K., Ahel, D., Lukas, C., Ahel, I., & Nielsen, M. L. (2021). The regulatory landscape of the human HPF1- and ARH3-dependent ADP-ribosylome. *Nature Communications*, 12, 5893.
95. Hornbeck, P. V., Kornhauser, J. M., Latham, V., Murray, B., Nandhikonda, V., Nord, A., Skrzypek, E., Wheeler, T., Zhang, B., & Gnad, F. (2019). 15 years of PhosphoSitePlus®: Integrating post-translationally modified sites, disease variants and isoforms. *Nucleic Acids Research*, 47, D433–D441.
96. Huang, H., Arighi, C. N., Ross, K. E., Ren, J., Li, G., Chen, S. C., Wang, Q., Cowart, J., Vijay-Shanker, K., & Wu, C. H. (2018). iPTMnet: An integrated resource for protein post-translational modification network discovery. *Nucleic Acids Research*, 46, D542–D550.
97. Li, Z., Li, S., Luo, M., Jhong, J. H., Li, W., Yao, L., Pang, Y., Wang, Z., Wang, R., Ma, R., Yu, J., Huang, Y., Zhu, X., Cheng, Q., Feng, H., Zhang, J., Wang, C., Hsu, J. B. K., Chang, W. C., ... Lee, T. Y. (2022). dbPTM in 2022: An updated database for exploring regulatory networks and functional associations of protein post-translational modifications. *Nucleic Acids Research*, 50, D471–D479.
98. Smith, L. M., & Kelleher, N. L. (2013). Proteoform: A single term describing protein complexity. *Nature Methods*, 10, 186–187.
99. Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bertozzi, C. R., Boja, E. S., Costello, C. E., Cravatt, B. F., Fenselau, C., Garcia, B. A., Ge, Y., Gunawardena, J., Hendrickson, R. C., Hergenrother, P. J., Huber, C. G., Ivanov, A. R., Jensen, O. N., Jewett, M. C., Kelleher, N. L., ... Zhang, B. (2018). How many human proteoforms are there? *Nature Chemical Biology*, 14, 206–214.
100. Smith, L. M., Agar, J. N., Chamot-Rooke, J., Danis, P. O., Ge, Y., Loo, J. A., Paša-Tolić, L., Tsybin, Y. O., Kelleher, N. L., & The Consortium for Top-down Proteomics. (2021). The Human Proteoform Project: Defining the human proteome. *Science Advances*, 7, eabk0734.
101. Matsuda, M., Mayer, B. J., Fukui, Y., & Hanafusa, H. (1990). Binding of transforming protein, P47gag-crk, to a broad range of phosphotyrosine-containing proteins. *Science*, 248, 1537–1539.
102. Feng, G. S., Hui, C. C., & Pawson, T. (1993). SH2-Containing phosphotyrosine phosphatase as a target of protein-tyrosine kinases. *Science*, 259, 1607–1611.
103. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., & Lechleider, R. J. (1993). SH2 domains recognize specific phosphopeptide sequences. *Cell*, 72, 767–778.
104. Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D., & Kuriyan, J. (1993). Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: Crystal structures of the complexed and peptide-free forms. *Cell*, 72, 779–790.
105. Bradley, D. (2022). The evolution of post-translational modifications. *Current Opinion in Genetics & Development*, 76, 101956.
106. Ochoa, D., Jarnuczak, A. F., Viéitez, C., Gehre, M., Soucheray, M., Mateus, A., Kleefeldt, A. A., Hill, A., Garcia-Alonso, L., Stein, F., Krogan, N. J., Savitski, M. M., Swaney, D. L., Vizcaíno, J. A., Noh, K. M., & Beltrao, P. (2020). The functional landscape of the human phosphoproteome. *Nature Biotechnology*, 38, 365–373.

107. Liang, Z., Liu, T., Li, Q., Zhang, G., Zhang, B., Du, X., Liu, J., Chen, Z., Ding, H., Hu, G., Lin, H., Zhu, F., & Luo, C. (2023). Deciphering the functional landscape of phosphosites with deep neural network. *Cell Reports*, 42, 113048.
108. Ramazi, S., & Zahiri, J. (2021). Post-translational modifications in proteins: Resources, tools and prediction methods. *Database: The Journal of Biological Databases and Curation*, 2021, baab012.
109. Hedberg, C., & Itzen, A. (2015). Molecular perspectives on protein adenylylation. *ACS Chemical Biology*, 10, 12–21.
110. Fersht, A. (1999). *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*. Macmillan.
111. Ho, M., Bramson, H. N., Hansen, D. E., Knowles, J. R., & Kaiser, E. T. (1988). Stereochemical course of the phospho group transfer catalyzed by cAMP-dependent protein kinase. *Journal of the American Chemical Society*, 110, 2680–2681.
112. Wang, Z., & Cole, P. A. (2014). *Methods in enzymology*. (Vol. 548, pp. 1–21) (K. M. Shokat, Ed.). Academic Press.
113. Berti, P. J., Blanke, S. R., & Schramm, V. L. (1997). Transition state structure for the hydrolysis of NAD⁺ catalyzed by diphtheria toxin. *Journal of the American Chemical Society*, 119, 12079–12088.
114. Yoon, M. Y., & Cook, P. F. (1987). Chemical mechanism of the adenosine cyclic 3',5'-monophosphate dependent protein kinase from pH studies. *Biochemistry*, 26, 4118–4125.
115. Gibbs, C. S., & Zoller, M. J. (1991). Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions. *Journal of Biological Chemistry*, 266, 8923–8931.
116. Skamnaki, V. T., Owen, D. J., Noble, M. E. M., Lowe, E. D., Lowe, G., Oikonomakos, N. G., & Johnson, L. N. (1999). Catalytic mechanism of phosphorylase kinase probed by mutational studies. *Biochemistry*, 38, 14718–14730.
117. Gerlits, O., Weiss, K. L., Blakeley, M. P., Veglia, G., Taylor, S. S., & Kovalevsky, A. (2020). *Methods in enzymology* (Vol. 634, pp. 311–331) (P. C. E. Moody, Ed.). Academic Press.
118. Zhang, X., Zhou, L., & Cheng, X. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *Embo Journal*, 19, 3509–3519.
119. Tanner, K. G., Trievel, R. C., Kuo, M. H., Howard, R. M., Berger, S. L., Allis, C. D., Marmorstein, R., & Denu, J. M. (1999). Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator *. *Journal of Biological Chemistry*, 274, 18157–18160.
120. Luong, P., Kinch, L. N., Brautigam, C. A., Grishin, N. V., Tomchick, D. R., & Orth, K. (2010). Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain *. *Journal of Biological Chemistry*, 285, 20155–20163.
121. Pao, K. C., Wood, N. T., Knebel, A., Rafie, K., Stanley, M., Mabbitt, P. D., Sundaramoorthy, R., Hofmann, K., van Aalten, D. M. F., & Virdee, S. (2018). Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity. *Nature*, 556, 381–385.
122. Ahel, J., Fletcher, A., Grabarczyk, D. B., Roitinger, E., Deszcz, L., Lehner, A., Virdee, S., & Clausen, T. (2021). E3 ubiquitin ligase RNF213 employs a non-canonical zinc finger active site and is allosterically regulated by ATP. *BioRxiv*, Preprint at <https://doi.org/10.1101/2021.05.10.443411>
123. Zhu, K., Suskiewicz, M. J., Hloušek-Kasun, A., Meudal, H., Mikoč, A., Aucagne, V., Ahel, D., & Ahel, I. (2022). DELTEX E3 ligases ubiquitylate ADP-ribosyl modification on protein substrates. *Science Advances*, 8, eadd4253.
124. Suskiewicz, M. J., Zobel, F., Ogden, T. E. H., Fontana, P., Ariza, A., Yang, J. C., Zhu, K., Bracken, L., Hawthorne, W. J., Ahel, D., Neuhaus, D., & Ahel, I. (2020). HPF1 completes the PARP active site for DNA damage-induced ADP-ribosylation. *Nature*, 579, 598–602.
125. Sun, F. H., Zhao, P., Zhang, N., Kong, L. L., Wong, C. C. L., & Yun, C. H. (2021). HPF1 remodels the active site of PARP1 to enable the serine ADP-ribosylation of histones. *Nature Communications*, 12, 1028.
126. Rudolph, J., Roberts, G., Muthurajan, U. M., & Luger, K. (2021). HPF1 and nucleosomes mediate a dramatic switch in activity of PARP1 from polymerase to hydrolase. *Elife*, 10, e65773.
127. Suskiewicz, M. J., Palazzo, L., Hughes, R., & Ahel, I. (2021). Progress and outlook in studying the substrate specificities of PARPs and related enzymes. *FEBS Journal*, 288, 2131–2142.
128. Langelier, M. F., Billur, R., Sverzhinsky, A., Black, B. E., & Pascal, J. M. (2021). HPF1 dynamically controls the PARP1/2 balance between initiating and elongating ADP-ribose modifications. *Nature Communications*, 12, 6675.
129. Suskiewicz, M. J., Prokhorova, E., Rack, J. G. M., & Ahel, I. (2023). ADP-ribosylation from molecular mechanisms to therapeutic implications. *Cell*, 186, 4475–4495.
130. Palazzo, L., Leidecker, O., Prokhorova, E., Dauben, H., Matic, I., & Ahel, I. (2018). Serine is the major residue for ADP-ribosylation upon DNA damage. *Elife*, 7, e34334.
131. Luo, M. (2018). Chemical and biochemical perspectives of protein lysine methylation. *Chemical Reviews*, 118, 6656–6705.
132. Walsh, C. T., Tu, B. P., & Tang, Y. (2018). Eight kinetically stable but thermodynamically activated molecules that power cell metabolism. *Chemical Reviews*, 118, 1460–1494.
133. Zheng, N., & Shabek, N. (2017). Ubiquitin ligases: Structure, function, and regulation. *Annual Review of Biochemistry*, 86, 129–157.
134. Sreelatha, A., Yee, S. S., Lopez, V. A., Park, B. C., Kinch, L. N., Pilch, S., Servage, K. A., Zhang, J., Jiou, J., Karasiewicz-Urbańska, M., Łobocka, M., Grishin, N. V., Orth, K., Kucharczyk, R., Pawłowski, K., Tomchick, D. R., & Tagliabracchi, V. S. (2018). Protein AMPylation by an evolutionarily conserved pseudokinase. *Cell*, 175, 809–821.e19.
135. Castro-Roa, D., Garcia-Pino, A., De Gieter, S., van Nuland, N. A. J., Loris, R., & Zenkin, N. (2013). The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. *Nature Chemical Biology*, 9, 811–817.
136. Westheimer, F. H. (1987). Why nature chose phosphates. *Science*, 235, 1173–1178.
137. Kamerlin, S. C. L., Sharma, P. K., Prasad, R. B., & Warshel, A. (2013). Why nature really chose phosphate. *Quarterly Reviews of Biophysics*, 46, 1–132.
138. Suskiewicz, M. J., Hajdusits, B., Beveridge, R., Heuck, A., Vu, L. D., Kurzbauer, R., Hauer, K., Thoeny, V., Rumpel, K., & Mechtler, K. (2019). Structure of McsB, a protein kinase for regulated arginine phosphorylation. *Nature Chemical Biology*, 15, 510–518.
139. Reverter, D., & Lima, C. D. (2005). Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature*, 435, 687–692.
140. Plechanovová, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H., & Hay, R. T. (2012). Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature*, 489, 115–120.
141. Pruneda, J. N., Littlefield, P. J., Soss, S. E., Nordquist, K. A., Chazin, W. J., Brzovic, P. S., & Klevit, R. E. (2012). Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Molecular Cell*, 47, 933–942.
142. Branigan, E., Carlos Penedo, J., & Hay, R. T. (2020). Ubiquitin transfer by a RING E3 ligase occurs from a closed E2~ubiquitin conformation. *Nature Communications*, 11, 2846.
143. Goffinont, S., Coste, F., Prieu-Serandon, P., Mance, L., Gaudon, V., Garnier, N., Castaing, B., & Suskiewicz, M. J. (2023). Structural insights into the regulation of the human E2~SUMO conjugate through analysis of its stable mimetic. *Journal of Biological Chemistry*, 299, 104870.
144. Tsuge, H., Nagahama, M., Oda, M., Iwamoto, S., Utsunomiya, H., Marquez, V. E., Katunuma, N., Nishizawa, M., & Sakurai, J. (2008). Structural basis of actin recognition and arginine ADP-ribosylation

- by *Clostridium perfringens* ι -toxin. *Proceedings of the National Academy of Sciences*, 105, 7399–7404.
145. Kemp, B. E., & Pearson, R. B. (1990). Protein kinase recognition sequence motifs. *Trends in Biochemical Sciences*, 15, 342–346.
 146. Johnson, J. L., Yaron, T. M., Huntsman, E. M., Kerelsky, A., Song, J., Regev, A., Lin, T. Y., Liberatore, K., Cizin, D. M., Cohen, B. M., Vasan, N., Ma, Y., Krismer, K., Robles, J. T., van de Kooij, B., van Vlimmeren, A. E., Andrée-Busch, N., Käufer, N. F., Dorovkov, M. V., ... Cantley, L. C. (2023). An atlas of substrate specificities for the human serine/threonine kinome. *Nature*, 613, 759–766.
 147. Rodriguez, M. S., Dargemont, C., & Hay, R. T. (2001). SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting *. *Journal of Biological Chemistry*, 276, 12654–12659.
 148. Bonfiglio, J. J., Fontana, P., Zhang, Q., Colby, T., Gibbs-Seymour, I., Atanassov, I., Bartlett, E., Zaja, R., Ahel, I., & Matic, I. (2017). Serine ADP-ribosylation depends on HPF1. *Molecular Cell*, 65, 932–940.e6.
 149. Miller, C. J., & Turk, B. E. (2018). Homing in: Mechanisms of substrate targeting by protein kinases. *Trends in Biochemical Sciences*, 43, 380–394.
 150. Swaffer, M. P., Jones, A. W., Flynn, H. R., Snijders, A. P., & Nurse, P. (2016). CDK substrate phosphorylation and ordering the cell cycle. *Cell*, 167, 1750–1761.e16.
 151. Kang, S. A., Pacold, M. E., Cervantes, C. L., Lim, D., Lou, H. J., Ottina, K., Gray, N. S., Turk, B. E., Yaffe, M. B., & Sabatini, D. M. (2013). mTORC1 phosphorylation sites encode their sensitivity to starvation and rapamycin. *Science*, 341, 1236566.
 152. Lo, W. L., Shah, N. H., Rubin, S. A., Zhang, W., Horkova, V., Fallahee, I. R., Stepanek, O., Zon, L. I., Kuriyan, J., & Weiss, A. (2019). Slow phosphorylation of a tyrosine residue in LAT optimizes T cell ligand discrimination. *Nature Immunology*, 20, 1481–1493.
 153. Holder, J., Mohammed, S., & Barr, F. A. (2020). Ordered dephosphorylation initiated by the selective proteolysis of cyclin B drives mitotic exit. *Elife*, 9, e59885.
 154. Guettler, S., LaRose, J., Petsalaki, E., Gish, G., Scotter, A., Pawson, T., Rottapel, R., & Sicheri, F. (2011). Structural basis and sequence rules for substrate recognition by tankyrase explain the basis for cherubism disease. *Cell*, 147, 1340–1354.
 155. Lucas, X., & Ciulli, A. (2017). Recognition of substrate degrons by E3 ubiquitin ligases and modulation by small-molecule mimicry strategies. *Current Opinion in Structural Biology*, 44, 101–110.
 156. Sherpa, D., Chrustowicz, J., & Schulman, B. A. (2022). How the ends signal the end: Regulation by E3 ubiquitin ligases recognizing protein termini. *Molecular Cell*, 82, 1424–1438.
 157. Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., & Yu, X. F. (2003). Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science*, 302, 1056–1060.
 158. Varland, S., Silva, R. D., Kjosås, I., Faustino, A., Bogaert, A., Billmann, M., Boukhatmi, H., Kellen, B., Costanzo, M., Drazic, A., Osberg, C., Chan, K., Zhang, X., Tong, A. H. Y., Andrezza, S., Lee, J. J., Nedyalkova, L., Ušaj, M., Whitworth, A. J., ... Arnesen, T. (2023). N-terminal acetylation shields proteins from degradation and promotes age-dependent motility and longevity. *Nature Communications*, 14, 6774.
 159. Varshavsky, A. (2019). N-degron and C-degron pathways of protein degradation. *Proceedings of the National Academy of Sciences*, 116, 358–366.
 160. Lee, J. M., Hammarén, H. M., Savitski, M. M., & Baek, S. H. (2023). Control of protein stability by post-translational modifications. *Nature Communications*, 14, 201.
 161. Hajdusits, B., Suskiewicz, M. J., Hundt, N., Meinhart, A., Kurzbauer, R., Leodolter, J., Kukura, P., & Clausen, T. (2021). McsB forms a gated kinase chamber to mark aberrant bacterial proteins for degradation. *Elife*, 10, e63505.
 162. McGinty, R. K., & Tan, S. (2021). Principles of nucleosome recognition by chromatin factors and enzymes. *Current Opinion in Structural Biology*, 71, 16–26.
 163. Zhang, T., Cooper, S., & Brockdorff, N. (2015). The interplay of histone modifications – writers that read. *Embo Reports*, 16, 1467–1481.
 164. Jentsch, S., & Psakhye, I. (2013). Control of nuclear activities by substrate-selective and protein-group SUMOylation. *Annual Review of Genetics*, 47, 167–186.
 165. Weinert, B. T., Narita, T., Satpathy, S., Srinivasan, B., Hansen, B. K., Schölz, C., Hamilton, W. B., Zucconi, B. E., Wang, W. W., Liu, W. R., Brickman, J. M., Kesicki, E. A., Lai, A., Bromberg, K. D., Cole, P. A., & Choudhary, C. (2018). Time-resolved analysis reveals rapid dynamics and broad scope of the CBP/p300 acetylome. *Cell*, 174, 231–244.e12.
 166. Huang, D., & Kraus, W. L. (2022). The expanding universe of PARP1-mediated molecular and therapeutic mechanisms. *Molecular Cell*, 82, 2315–2334.
 167. Duarte, M. L., Pena, D. A., Nunes-Ferraz, F. A., Berti, D. A., Paschoal-Sobreira, T. J., Costa-Junior, H. M., Abdel-Baqui, M. M., Disatnik, M. H., Xavier-Neto, J., Lopes-de-Oliveira, P. S., & Schechtman, D. (2014). Protein folding creates structure-based, noncontiguous consensus phosphorylation motifs recognized by kinases. *Science Signaling*, 7, ra105.
 168. Dar, A. C., Dever, T. E., & Sicheri, F. (2005). Higher-order substrate recognition of eIF2 α by the RNA-dependent protein kinase PKR. *Cell*, 122, 887–900.
 169. Hamill, S., Lou, H. J., Turk, B. E., & Boggon, T. J. (2016). Structural basis for noncanonical substrate recognition of cofilin/ADF proteins by LIM kinases. *Molecular Cell*, 62, 397–408.
 170. Deol, K. K., Lorenz, S., & Strieter, E. R. (2019). Enzymatic logic of ubiquitin chain assembly. *Frontiers in Physiology*, 10, 835.
 171. Reber, J. M., & Mangerich, A. (2021). Why structure and chain length matter: On the biological significance underlying the structural heterogeneity of poly(ADP-ribose). *Nucleic Acids Research*, 49, 8432–8448.
 172. Schjoldager, K. T., Narimatsu, Y., Joshi, H. J., & Clausen, H. (2020). Global view of human protein glycosylation pathways and functions. *Nature Reviews Molecular Cell Biology*, 21, 729–749.
 173. Prokhorova, E., Agnew, T., Wondisford, A. R., Tellier, M., Kaminski, N., Beijer, D., Holder, J., Gros Lambert, J., Suskiewicz, M. J., Zhu, K., Reber, J. M., Krassnig, S. C., Palazzo, L., Murphy, S., Nielsen, M. L., Mangerich, A., Ahel, D., Baets, J., O'Sullivan, R. J., & Ahel, I. (2021). Unrestrained poly-ADP-ribosylation provides insights into chromatin regulation and human disease. *Molecular Cell*, 81, 2640–2655.e8.
 174. Lips, C., Ritterhoff, T., Weber, A., Janowska, M. K., Mustroph, M., Sommer, T., & Klevit, R. E. (2020). Who with whom: Functional coordination of E2 enzymes by RING E3 ligases during poly-ubiquitylation. *Embo Journal*, 39, e104863.
 175. Mevissen, T. E. T., & Komander, D. (2017). Mechanisms of deubiquitinase specificity and regulation. *Annual Review of Biochemistry*, 86, 159–192.
 176. Pérez-Berrocal, D. A., Witting, K. F., Ovaa, H., & Mulder, M. P. C. (2020). Hybrid chains: A collaboration of ubiquitin and ubiquitin-like modifiers introducing cross-functionality to the ubiquitin code. *Frontiers in Chemistry*, 7, 931.
 177. Lu-Culligan, W. J., Connor, L. J., Xie, Y., Ekundayo, B. E., Rose, B. T., Machyna, M., Pintado-Urbanc, A. P., Zimmer, J. T., Vock, I. W., Bhanu, N. V., King, M. C., Garcia, B. A., Bleichert, F., & Simon, M. D. (2023). Acetyl-methyllysine marks chromatin at active transcription start sites. *Nature*, 622, 173–179.
 178. Barford, D., Das, A. K., & Egloff, M. P. (1998). The structure and mechanism of protein phosphatases: Insights into catalysis and regulation. *Annual Review of Biophysics and Biomolecular Structure*, 27, 133–164.
 179. Rack, J. G. M., Palazzo, L., & Ahel, I. (2020). (ADP-ribosyl)hydrolases: Structure, function, and biology. *Genes & Development*, 34, 263–284.

180. Ng, S. S., Yue, W. W., Oppermann, U., & Klose, R. J. (2008). Dynamic protein methylation in chromatin biology. *Cellular and Molecular Life Sciences*, *66*, 407–422.
181. Sauve, A. A., & Youn, D. Y. (2012). Sirtuins: NAD⁺-dependent deacetylase mechanism and regulation. *Current Opinion in Chemical Biology*, *16*, 535–543.
182. Lemmon, M. A., Freed, D. M., Schlessinger, J., & Kiyatkin, A. (2016). The dark side of cell signaling: Positive roles for negative regulators. *Cell*, *164*, 1172–1184.
183. Wilkinson, K. D. (2000). Ubiquitination and deubiquitination: Targeting of proteins for degradation by the proteasome. *Seminars in Cell & Developmental Biology*, *11*, 141–148.
184. Glickman, M. H., & Adir, N. (2004). The proteasome and the delicate balance between destruction and rescue. *PLoS Biology*, *2*, e13.
185. Fisher, D., Krasinska, L., Coudreuse, D., & Novák, B. (2012). Phosphorylation network dynamics in the control of cell cycle transitions. *Journal of Cell Science*, *125*, 4703–4711.
186. Hunter, T. (2012). Why nature chose phosphate to modify proteins. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 2513–2516.
187. Hubbard, S. R., Wei, L., Ellis, L., & Hendrickson, W. A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature*, *372*, 746–754.
188. Hubbard, S. R. (1997). Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *Embo Journal*, *16*, 5572–5581.
189. Chuderland, D., Konson, A., & Seger, R. (2008). Identification and characterization of a general nuclear translocation signal in signaling proteins. *Molecular Cell*, *31*, 850–861.
190. Nardozi, J. D., Lott, K., & Cingolani, G. (2010). Phosphorylation meets nuclear import: A review. *Cell Communication and Signaling*, *8*, 32.
191. Sridharan, S., Hernandez-Armendariz, A., Kurzawa, N., Potel, C. M., Memon, D., Beltrao, P., Bantscheff, M., Huber, W., Cuylen-Haering, S., & Savitski, M. M. (2022). Systematic discovery of biomolecular condensate-specific protein phosphorylation. *Nature Chemical Biology*, *18*, 1104–1114.
192. Seet, B. T., Dikic, I., Zhou, M. M., & Pawson, T. (2006). Reading protein modifications with interaction domains. *Nature Reviews Molecular Cell Biology*, *7*, 473–483.
193. Musselman, C. A., Lalonde, M. E., Côté, J., & Kutateladze, T. G. (2012). Perceiving the epigenetic landscape through histone readers. *Nature Structural & Molecular Biology*, *19*, 1218–1227.
194. Teloni, F., & Altmeyer, M. (2016). Readers of poly(ADP-ribose): Designed to be fit for purpose. *Nucleic Acids Research*, *44*, 993–1006.
195. Osei-Amponsa, V., & Walters, K. J. (2022). Proteasome substrate receptors and their therapeutic potential. *Trends in Biochemical Sciences*, *47*, 950–964.
196. Jacob, F. (1977). Evolution and tinkering. *Science*, *196*, 1161–1166.
197. Aravind, L., Burroughs, A. M., Zhang, D., & Iyer, L. M. (2014). Protein and DNA modifications: Evolutionary imprints of bacterial biochemical diversification and geochemistry on the provenance of eukaryotic epigenetics. *Cold Spring Harbor Perspectives in Biology*, *6*, a016063.
198. Aravind, L., Anantharaman, V., Zhang, D., De Souza, R., & Iyer, L. (2012). Gene flow and biological conflict systems in the origin and evolution of eukaryotes. *Frontiers in Cellular and Infection Microbiology*, *2*, 89.
199. Otten, E. G., Werner, E., Crespillo-Casado, A., Boyle, K. B., Dharamdasani, V., Pathe, C., Santhanam, B., & Randow, F. (2021). Ubiquitylation of lipopolysaccharide by RNF213 during bacterial infection. *Nature*, *594*, 111–116.
200. Ichimura, Y., Kirisako, T., Takao, T., Satomi, Y., Shimonishi, Y., Ishihara, N., Mizushima, N., Tanida, I., Kominami, E., Ohsumi, M., Noda, T., & Ohsumi, Y. (2000). A ubiquitin-like system mediates protein lipidation. *Nature*, *408*, 488–492.
201. Sakamaki, J., Ode, K. L., Kurikawa, Y., Ueda, H. R., Yamamoto, H., & Mizushima, N. (2022). Ubiquitination of phosphatidylethanolamine in organellar membranes. *Molecular Cell*, *82*, 3677–3692.e11.
202. Talhaoui, I., Lebedeva, N. A., Zarkovic, G., Saint-Pierre, C., Kutuzov, M. M., Sukhanova, M. V., Matkarimov, B. T., Gasparutto, D., Saparbaev, M. K., Lavrik, O. I., & Ishchenko, A. A. (2016). Poly(ADP-ribose) polymerases covalently modify strand break termini in DNA fragments in vitro. *Nucleic Acids Research*, *44*, 9279–9295.
203. Jankevicius, G., Ariza, A., Ahel, M., & Ahel, I. (2016). The toxin-antitoxin system DarTG catalyzes reversible ADP-Ribosylation of DNA. *Molecular Cell*, *64*, 1109–1116.
204. Schuller, M., Butler, R. E., Ariza, A., Tromans-Coia, C., Jankevicius, G., Claridge, T. D. W., Kendall, S. L., Goh, S., Stewart, G. R., & Ahel, I. (2021). Molecular basis for DarT ADP-ribosylation of a DNA base. *Nature*, *596*, 597–602.
205. Bullen, N. P., Sychantha, D., Thang, S. S., Culviner, P. H., Rudzite, M., Ahmad, S., Shah, V. S., Filloux, A., Prehna, G., & Whitney, J. C. (2022). An ADP-ribosyltransferase toxin kills bacterial cells by modifying structured non-coding RNAs. *Molecular Cell*, *82*, 3484–3498.e11.
206. Weixler, L., Feijs, K. L. H., & Zaja, R. (2022). ADP-ribosylation of RNA in mammalian cells is mediated by TRPT1 and multiple PARPs. *Nucleic Acids Research*, *50*, 9426–9441.
207. Sakamaki, J., & Mizushima, N. (2023). Ubiquitination of non-protein substrates. *Trends in Cell Biology*, *33*, 991–1003.
208. Squair, D. R., & Virdee, S. (2022). A new dawn beyond lysine ubiquitination. *Nature Chemical Biology*, *18*, 802–811.
209. Kelsall, I. R. (2022). Non-lysine ubiquitylation: Doing things differently. *Frontiers in Molecular Biosciences*, *9*, 1008175.
210. Laurino, P., & Tawfik, D. S. (2017). Spontaneous emergence of S-adenosylmethionine and the evolution of methylation. *Angewandte Chemie International Edition*, *56*, 343–345.
211. Beltrao, P., Bork, P., Krogan, N. J., & van Noort, V. (2013). Evolution and functional cross-talk of protein post-translational modifications. *Molecular Systems Biology*, *9*, 714.
212. Rajagopalan, K. V. (1997). Biosynthesis and processing of the molybdenum cofactors. *Biochemical Society Transactions*, *25*, 757–761.
213. Taylor, S. V., Kelleher, N. L., Kinsland, C., Chiu, H. J., Costello, C. A., Backstrom, A. D., McLafferty, F. W., & Begley, T. P. (1998). Thiamin biosynthesis in *Escherichia coli*: Identification of this thiocarboxylate as the immediate sulfur donor in the thiazole formation *. *Journal of Biological Chemistry*, *273*, 16555–16560.
214. Lake, M. W., Wuebbens, M. M., Rajagopalan, K. V., & Schindelin, H. (2001). Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB–MoaD complex. *Nature*, *414*, 325–329.
215. Iyer, L. M., Burroughs, A. M., & Aravind, L. (2006). The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like β -grasp domains. *Genome Biology*, *7*, R60.
216. Furukawa, K., Mizushima, N., Noda, T., & Ohsumi, Y. (2000). A protein conjugation system in yeast with homology to biosynthetic enzyme reaction of prokaryotes*. *Journal of Biological Chemistry*, *275*, 7462–7465.
217. Schmitz, J., Chowdhury, M. M., Hänzelmann, P., Nimitz, M., Lee, E. Y., Schindelin, H., & Leimkühler, S. (2008). The sulfurtransferase activity of Uba4 presents a link between ubiquitin-like protein conjugation and activation of sulfur carrier proteins. *Biochemistry*, *47*, 6479–6489.
218. Leidel, S., Pedrioli, P. G. A., Bucher, T., Brost, R., Costanzo, M., Schmidt, A., Aebersold, R., Boone, C., Hofmann, K., & Peter, M. (2009). Ubiquitin-related modifier Urm1 acts as a sulphur carrier in thiolation of eukaryotic transfer RNA. *Nature*, *458*, 228–232.
219. Pabis, M., Termathe, M., Ravichandran, K. E., Kienast, S. D., Krutyholowa, R., Sokotowski, M., Jankowska, U., Grudnik, P., Leidel, S. A., & Glatt, S. (2020). Molecular basis for the bifunctional Uba4–Urm1 sulfur-relay system in tRNA thiolation and ubiquitin-like conjugation. *EMBO Journal*, *39*, e105087.

220. Ravichandran, K. E., Kaduhr, L., Skupien-Rabian, B., Shvetsova, E., Sokołowski, M., Krutyholowa, R., Kwasna, D., Brachmann, C., Lin, S., Guzman Perez, S., Wilk, P., Kösters, M., Grudnik, P., Jankowska, U., Leidel, S. A., Schaffrath, R., & Glatt, S. (2022). E2/E3-independent ubiquitin-like protein conjugation by Urm1 is directly coupled to cysteine persulfidation. *EMBO Journal*, *41*, e111318.
221. Iyer, L. M., Burroughs, A., & Aravind, L. (2008). Unraveling the biochemistry and provenance of pupylation: A prokaryotic analog of ubiquitination. *Biology Direct*, *3*, 45.
222. Özcelik, D., Barandun, J., Schmitz, N., Sutter, M., Guth, E., Damberger, F. F., Allain, F. H. T., Ban, N., & Weber-Ban, E. (2012). Structures of Pup ligase PafA and depupylase Dop from the prokaryotic ubiquitin-like modification pathway. *Nature Communications*, *3*, 1014.
223. Hecht, N., Regev, O., Dovrat, D., Aharoni, A., & Gur, E. (2018). Proteasome accessory factor A (PafA) transferase activity makes sense in the light of its homology with glutamine synthetase. *Journal of Molecular Biology*, *430*, 668–681.
224. Guth, E., Thommen, M., & Weber-Ban, E. (2011). Mycobacterial ubiquitin-like protein ligase PafA follows a two-step reaction pathway with a phosphorylated pup intermediate*. *Journal of Biological Chemistry*, *286*, 4412–4419.
225. Hang, H. C., Pratt, M. R., & Prescher, J. A. (Eds.) (2023). *Advanced chemical biology: Chemical dissection and reprogramming of biological systems*. John Wiley & Sons.
226. Perina, D., Mikoč, A., Ahel, J., Četković, H., Žaja, R., & Ahel, I. (2014). Distribution of protein poly(ADP-ribosylation) systems across all domains of life. *DNA Repair*, *23*, 4–16.
227. Hajredini, F., Alphonse, S., & Ghose, R. (2023). BY-kinases: Protein tyrosine kinases like no other. *Journal of Biological Chemistry*, *299*, 102737.
228. Stancik, I. A., Šestak, M. S., Ji, B., Axelson-Fisk, M., Franjevic, D., Jers, C., Domazet-Lošo, T., & Mijakovic, I. (2018). Serine/Threonine protein kinases from bacteria, archaea and eukarya share a common evolutionary origin deeply rooted in the tree of life. *Journal of Molecular Biology*, *430*, 27–32.
229. van Wijk, L. M., & Snel, B. (2020). The first eukaryotic kinome tree illuminates the dynamic history of present-day kinases. *BioRxiv*, Preprint at <https://doi.org/10.1101/2020.01.27.920793>
230. Lim, W. A., & Pawson, T. (2010). Phosphotyrosine signaling: Evolving a new cellular communication system. *Cell*, *142*, 661–667.
231. Tan, C. S. H., Pasulescu, A., Lim, W. A., Pawson, T., Bader, G. D., & Linding, R. (2009). Positive selection of tyrosine loss in metazoan evolution. *Science*, *325*, 1686–1688.
232. Pearlman, S. M., Serber, Z., & Ferrell, J. E. (2011). A mechanism for the evolution of phosphorylation sites. *Cell*, *147*, 934–946.
233. Suskiewicz, M. J., Munnur, D., Strømmland, Ø., Yang, J. C., Easton, L. E., Chatrin, C., Zhu, K., Baretić, D., Goffinont, S., & Schuller, M. (2023). Updated protein domain annotation of the PARP protein family sheds new light on biological function. *Nucleic Acids Research*, *51*, 8217–8236.
234. Đukić, N., Strømmland, Ø., Elsborg, J. D., Munnur, D., Zhu, K., Schuller, M., Chatrin, C., Kar, P., Duma, L., Suyari, O., Rack, J. G. M., Baretić, D., Crudgington, D. R. K., Gros Lambert, J., Fowler, G., Wijngaarden, S., Prokhorova, E., Rehwinkel, J., Schüler, H., ... Ahel, I. (2023). PARP14 is a PARP with both ADP-ribosyl transferase and hydrolase activities. *Science Advances*, *9*, eadi2687.
235. Torretta, A., Chatzicharalampous, C., Ebenwaldner, C., & Schüler, H. (2023). PARP14 is a writer, reader, and eraser of mono-ADP-ribosylation. *Journal of Biological Chemistry*, *299*, 105096.
236. Davey, N. E., Cyert, M. S., & Moses, A. M. (2015). Short linear motifs – ex nihilo evolution of protein regulation. *Cell Communication and Signaling*, *13*, 43.
237. Bludau, I., Willems, S., Zeng, W. F., Strauss, M. T., Hansen, F. M., Tanzer, M. C., Karayel, O., Schulman, B. A., & Mann, M. (2022). The structural context of posttranslational modifications at a proteome-wide scale. *PLoS Biology*, *20*, e3001636.
238. Hannich, J. T., Lewis, A., Kroetz, M. B., Li, S. J., Heide, H., Emili, A., & Hochstrasser, M. (2005). Defining the SUMO-modified proteome by multiple approaches in *Saccharomyces cerevisiae**. *Journal of Biological Chemistry*, *280*, 4102–4110.
239. Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P., & Dikic, I. (2006). Specification of SUMO1- and SUMO2-interacting motifs*. *Journal of Biological Chemistry*, *281*, 16117–16127.
240. Kerscher, O. (2007). SUMO junction--what's your function? *EMBO Reports*, *8*, 550–555.
241. Hochberg, G. K. A., Liu, Y., Marklund, E. G., Metzger, B. P. H., Laganowsky, A., & Thornton, J. W. (2020). A hydrophobic ratchet entrenches molecular complexes. *Nature*, *588*, 503–508.
242. Eisenhardt, N., Chaugule, V. K., Koidl, S., Droescher, M., Dogan, E., Rettich, J., Sutinen, P., Imanishi, S. Y., Hofmann, K., Palvimo, J. J., & Pichler, A. (2015). A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. *Nature Structural & Molecular Biology*, *22*, 959–967.

How to cite this article: Suskiewicz, M. J. (2024). The logic of protein post-translational modifications (PTMs): Chemistry, mechanisms and evolution of protein regulation through covalent attachments. *BioEssays*, e2300178. <https://doi.org/10.1002/bies.202300178>

AUTHOR BIOGRAPHY

Marcin J. Suskiewicz is a researcher and a co-leader of a research group at the Centre de Biophysique Moléculaire in Orléans, France. He specialises in the structural biology and biochemistry of protein post-translational modifications (PTMs). Polish by birth, he studied Biochemistry in the UK, followed by a PhD and a postdoc with Tim Clausen at the IMP in Vienna, where he worked on protein arginine phosphorylation, and a postdoc with Ivan Ahel at the Dunn School of Pathology in Oxford, where he focussed on protein ADP-ribosylation. His main current research interests are protein SUMOylation, protein ubiquitylation, and protein self-assembly.