Post-Translational Modifications of Protein Backbones: Unique Functions, Mechanisms, and Challenges

Manuel M. Müller*

Department of Chemistry, King’s College London, 7 Trinity Street, London SE1 1DB, United Kingdom

ABSTRACT: Post-translational modifications (PTMs) dramatically enhance the capabilities of proteins. They introduce new functionalities and dynamically control protein activity by modulating intra- and intermolecular interactions. Traditionally, PTMs have been considered as reversible attachments to nucleophilic functional groups on amino acid side chains, whereas the polypeptide backbone is often thought to be inert. This paradigm is shifting as chemically and functionally diverse alterations of the protein backbone are discovered. Importantly, backbone PTMs can control protein structure and function just as side chain modifications do and operate through unique mechanisms to achieve these features. In this Perspective, I outline the various types of protein backbone modifications discovered so far and highlight their contributions to biology as well as the challenges in studying this versatile yet poorly characterized class of PTMs.

Proteins are assembled from a set of typically 20 α-L-amino acids by iterative formation of amide bonds within the ribosome. Following their biosynthesis, proteins are further tailored through covalent modifications that introduce new reactive groups not inherently present in the standard amino acid building blocks, thus enabling new chemistries. Of equal importance are modifications that dynamically or irreversibly control the localization and activities of most natural proteins and thus serve as regulators or even on/off switches. These so-called post-translational modifications (PTMs) are typically considered to occur primarily on amino acid side chains that present a variety of nucleophilic groups that are easily targeted by electrophilic cofactors. The peptide backbone, which makes up approximately 50% of every protein’s mass, is often thought to remain as forged by the ribosome. The rationale for this conjecture is the inherently low reactivity of amides and the notion that the backbone merely holds proteins together in one (sequence) and three dimensions (structure). However, the polypeptide backbone plays an active role in shaping protein structure, and the properties of backbone atoms depend critically on the surrounding amino acid sequence and local conformation, which can profoundly alter amide reactivity. These features make backbone atoms ideal targets for spontaneous and enzymatic modifications. Indeed, peptides and proteins featuring alterations to all components of the backbone have been discovered (Figure 1), introducing new functional groups and unique capabilities to fine-tune protein structures on demand. This Perspective is aimed at showcasing the chemical and functional diversity of backbone PTMs (bbPTMs), focusing on covalent modifications within proteins. Modification of termini, targeted proteolysis, splicing, and proline isomerization have been excellently reviewed elsewhere and are beyond the scope of this work. I will begin by providing a glimpse into backbone modifications found in peptide natural products to illustrate the rich palette of biochemical possibilities for bbPTMs, followed by a more in-depth treatment of bbPTMs that occur in large proteins. Subsequently, I will highlight a series of examples to illustrate how bbPTMs can (i) endow proteins with novel properties, (ii) constitutively enhance protein stability and activity, and (iii) serve as dynamic regulators of activity. To conclude, I will discuss the unique mechanisms of action of bbPTMs and the tools and challenges for their discovery, systematic cataloguing, and functional characterization.

Backbone Modifications Are Ubiquitous in Peptide Natural Products

Ribosomally synthesized and post-translationally modified peptide natural products (RiPPs) are prime examples of the rich biochemistries that Nature harnesses to modify the polypeptide backbone. These peptides are often deployed by their producers, organisms from all domains of life, as toxins for targeted chemical warfare against prey, predators, and competitors for resources. Crucial to the activities of many RiPPs is the occurrence of PTMs that occur on the polypeptide backbone (Figure 1). Masking of amide bonds that would otherwise be susceptible to attack by proteases increases the biochemical stability of peptides and can improve pharmacological properties such as membrane permeability. In addition, backbone modifications can control local and global conformation and thus drive the formation of well-defined three-dimensional structures even in short peptides. bbPTMs comprise chemically conservative modifications (such as the conversion of an L- to a D-amino acid or the methylation of the amide nitrogen) and substantial alterations to the backbone (including the formation ofazole heterocycles). Collectively, RiPPs serve as an inspiration for the types of
bbPTMs that may be present in larger proteins and for enzymes that install them.

PROTEIN BACKBONE MODIFICATIONS INTRODUCE UNIQUE PROTEIN FUNCTIONS

In contrast to peptide natural products, the majority of proteins benefit from extensive tertiary interactions to stabilize their folds, and many are not exposed to the harsh environments of the extracellular space. Are bbPTMs nevertheless exploited by proteins, as well? Indeed, bbPTMs have been discovered in a range of proteins, in microorganisms and animals alike, where they fine-tune protein properties and even introduce unique chemical motifs that confer novel functions. Perhaps most prominently, fluorescent proteins such as GFP from the jellyfish Aequorea victoria mature through a series of bbPTMs that establish the fluorophore core (Figure 2a,b).14 Nucleophilic attack of the amide nitrogen of Gly67 on the preceding peptide bond results in the formation of a five-membered ring. Subsequent elimination of water and oxidation of Tyr66 into a C_o−C_β unsaturated derivative yield a conjugated π-system spanning backbone atoms from three adjacent residues that forms the basis for fluorescence. What allows this reaction to occur is that when the protein folds, the amide bonds involved are juxtaposed and their reactivities tuned by adjacent functional groups. The cofactor 4-methylidene-5-imidazole-5-one (MIO) is formed in a similar manner. Harnessed by amino acid ammonia lyases and mutases from prokaryotes to mammals, MIO is generated by cyclization and dehydration of a tripeptide motif to introduce an electrophile within their active sites (Figure 2c,d).15 These examples represent spontaneous, constitutive maturation processes that demonstrate the malleability of the protein backbone and the types of novel structures that can be accessed through its modification.

Figure 1. Examples of post-translational modifications of the polypeptide backbone (bbPTMs). This Perspective focuses on covalent modifications at Cα (blue dotted lines), the amide N (green), and C=O (orange) as well as backbone extensions (purple) on proteins (yellow shaded areas). Selected protein examples for the depicted modifications are listed below. Additional bbPTMs found in ribosomally synthesized and post-translationally modified peptides (RiPPs) are shaded with a blue background. MCR represents methyl-enzyme M reductase.

Figure 2. Novel functions provided by backbone modifications. (a and b) Maturation of the GFP fluorophore via a backbone cyclization, dehydration, and oxidation (Protein Data Bank (PDB) entry 1EMA). (c and d) Formation of the electrophilic cofactor 4-methylidene-5-imidazole-5-one (MIO) via backbone cyclization and dehydration steps (PDB entry 1GKJ).
PROTEIN BACKBONE MODIFICATIONS MODULATE PROTEIN ACTIVITY AND STABILITY

Backbone modifications can also improve existing functions by increasing enzyme stability and activity. Such bPTMs are introduced via spontaneous maturation processes in cis (as discussed above for GFP and MIO) or by specific enzymes in trans. A case in point for autocatalytic bPTMs is the Cα hydroxylation of an active site proline in a bacterial polysaccharide deacetylase (Figure 3a).18 The additional OH group, installed under aerobic conditions, provides a hydrogen bond to stabilize the tetrahedral intermediate and thus accelerates deacetylase activity.18 Another series of spontaneous bPTMs involves asparagine and aspartate residues. These reactions are initiated by nucleophilic attack of a backbone amide nitrogen on the side chain amide (Asn) or, to a lesser extent, acid (Asp).19,20 The resulting succinimide is typically unstable and is hydrolyzed to provide either Asp or its β-linked derivative, isoAsp (Figure 4a). Succinimides are also prone to epimerization, leading to the formation of d-Asp and d-isoAsp. Accordingly, isoAsp and its mirror image accumulate over time and are thus often detected in naturally and artificially aged proteins, where their formation has been implicated in protein damage and age-related human pathologies.19 Importantly, however, there are cases in which Asn-related backbone modifications occur constitutively in "young" proteins and confer desirable properties. A recent study of glutaminase from a hyperthermophilic archaeon revealed that a stable succinimide moiety confers extraordinary stability on this enzyme (Figure 3b).21 The succinimide, shielded from hydrolysis by surrounding anionic residues, introduces a conformational constraint that is important for protein stability. Consequently, disrupting the succinimide by site-directed mutagenesis diminishes the protein’s resilience to chemical and thermal denaturation.21 Similarly, an isoAsp residue has been observed in MurA,22 an enzyme involved in bacterial cell wall biosynthesis (Figure 3c). It is likely that the additional methylene group of isoAsp triggers the formation of an unusual turn structure, which might also enhance protein stability, although this conjecture has not been experimentally tested.

A suite of interesting, enzymatically installed bPTMs has been discovered in archaeal methyl-coenzyme M reductase (Figure 3d).23 Didehydroaspartate,24 α-methylglutamine,24 and thioglycine24 (substitution of the carbonyl oxygen of a glycine residue with a sulfur atom) have been detected in several homologues near the...
active site, attesting to their importance. A candidate oxidoreductase for didehydroaspartate formation has been postulated,25 and it has been demonstrated that the Cα-methyl group of α-methylglutamine originates from the cofactor S-adenosylmethionine, but a specific methyltransferase has not been identified.26 Recently, the enzymes responsible for thioglycine formation, a backbone kinase (YcaO) and an auxiliary protein of unknown function (TfuA), have been identified through homologues involved in natural product thioamide biosynthesis pathways (Figure 4a).13,27 Deleting the genes encoding the modifying enzymes in a methanogenic auxiliary protein of unknown function (TfuA), have been identified through homologues of several repair cycles (Figure 4a).30 Recently, it was also found that the formation of isoAsp from Asp can be enzymatically promoted by glycosyltransferases [through the formation of a glycosyl aspartate intermediate (Figure 4a)] and possibly other transferases, as well.29 The existence of enzymes that catalyze the formation and reversal of this bbPTM suggests that it is harnessed for dynamic regulatory purposes.

The spontaneous nature of isoAsp formation provides an additional regulatory mechanism. This process can act as a molecular timer, turning on or off protein function in a time-dependent manner. The rate of isoAsp formation is strongly influenced by the surrounding sequence as well as the local structure, with half-times ranging from hours to centuries.20 These time scales represent typical lifetimes of proteins and organisms. However, given that the time scales of this reaction are evolutionarily (protein sequence and structure) and biochemically controlled,20 it is possible that isoAsp formation is actively harnessed for signaling. Indeed, several components of apoptotic pathways,31,32 chromatin,33,34 and the cell adhesion machinery35 are believed to signal through this bbPTM, and more detailed structure–function studies will be required to illuminate their mechanism of action.

Backbone modifications are also exploited to regulate proteins directly in active sites. For example, human protein tyrosine phosphatase 1B is switched off by masking the catalytic cysteine residue as a sulfenyl-amide with the adjacent backbone thioamide nitrogen (Figure 5a).36 Cyclization is controlled by the cellular redox state; reactive oxygen species promote S–N bond formation, whereas the cellular reducing agent glutathione reverses it. Another case in point is the phosphothreonine lyase OspF, deployed by pathogens to inactivate their hosts immune signaling pathways (Figure 5b). This enzyme catalyzes the elimination of phosphate from phosphothreonine residues in key signal transduction proteins, thereby inactivating the target and leaving an α→β unsaturated residue at the site of action.37 In this case, the backbone modification does not provide a unique function but is merely a byproduct in the inactivation of a signaling pathway. It is worth noting that homologous enzymes install α→β unsaturated residues in peptide natural products,38 lending further support to the notion that peptide-modifying enzymes can inspire the identification of bbPTMs in proteins.

**REGULATORY ROLES OF PROTEIN BACKBONE MODIFICATIONS**

Side chain PTMs are often harnessed to switch proteins on or off, as illustrated by the well-known role of protein phosphorylation in signaling cascades. This type of control is mediated by enzymes to dynamically install and remove PTMs (e.g., kinases and phosphatases, respectively, in the case of phosphorylation), but little is known about reversibility in the context of backbone modifications. An intriguing exception is the formation and reversal of isoAsp linkages. A dedicated protein isoaspartate methyltransferase (PIMT) catalyzes the methylation of isoAsp at the free α-carboxylate to accelerate succinimide formation, leading to the accumulation of Asp over several repair cycles (Figure 4a).30 Recently, it was also found that the formation of isoAsp from Asp can be enzymatically promoted by glycosyltransferases [through the formation of a glycosyl aspartate intermediate (Figure 4a)] and possibly other
The examples described above demonstrate that bbPTMs exert their effects not only by contributing new reactivities but also through unique conformational mechanisms. Specifically, they can fine-tune the position of existing functional groups (Figure 6a) and dramatically alter the trajectory of the polypeptide backbone (Figure 6b). For instance, d-amino acids repurpose their own side chain and provide access to dihedral angles that are otherwise accessible only to glycine, which is particularly favorable for inducing turns (Figure 6b). Cα methylation stabilizes α-helices by reducing the favorable ϕ and ψ angle space (Figure 6b). This effect possibly explains the presence of the α-methyl-Gln residue in a helix that packs against the active site of methyl-coenzyme M reductase (Figure 3b). Cyclically constrained residues such as succinimides rigidify the backbone and provide access to unusual dihedral angles (Figure 3b). Reduced flexibility is a hallmark of thermophilic enzymes, and it is conceivable that succinimide formation, which has been shown to contribute to the thermostability of glutaminase (Figure 3b), these changes can propagate to alter global protein properties.

**IDENTIFICATION OF PROTEIN BACKBONE PTMS, FROM SERENDIPITOUS DISCOVERIES TO “OMICS”**

To date, most protein backbone modifications were discovered serendipitously. Careful interpretations of electron density maps from X-ray crystallography studies have revealed the existence of most of the bbPTMs discussed above (see, for example, Figure 3). Structural biology represents an important tool kit in the identification and characterization of bbPTMs but is inherently limited in throughput and favors autocatalytic modifications that manifest during recombinant protein production. To gain a systematic understanding of the distribution of bbPTMs, new approaches are required that can also address regulatory PTMs, which occur in a spatially, temporally, and stoichiometrically restricted manner. Protein mass spectrometry, in conjunction with enrichment strategies (commonly employed for side chain modifications), will be instrumental in this process (Figure 7a). In the context of bbPTMs, such proteomics strategies have been developed and successfully applied only to search for new isoAsp- and succinimide-containing proteins (Figure 7b). Current strategies hinge upon identifying isoAsp by labeling this residue with the specific “repair” enzyme PIMT or unique fragmentation patterns of β-peptidic linkages observed through high-end mass spectrometry. Such efforts were used to catalogue the “isoAspartome” for example in bacteria, as well as in cell lines, tissues, and fluids from mice lacking PIMT. Of particular interest is the finding that ribosomal protein S11 in *Escherichia coli* exhibits close to stoichiometric amounts of isoAsp, suggesting that this modification is functionally important, although its specific role is still unclear. Moreover, enrichment protocols based on nucleophilic trapping of the methylester or succinimide intermediate have been developed to increase the sensitivity of detection (Figure 7b).

Similar technologies for enriching and identifying other bbPTMs will be needed to fully appreciate the role of these modifications in biology. As for side chain PTMs, affinity reagents such as antibodies (or synthetic analogues thereof) will be needed to fully appreciate the role of these modifications in biology. As for side chain PTMs, affinity reagents such as antibodies (or synthetic analogues thereof) will be needed to fully appreciate the role of these modifications in biology.
as well as chemical reactivity-based strategies will likely be instrumental in this process. In this regard, antibodies that recognize specific isoAsp residues, for example, in amyloids, eye lens crystallin, and histones have been described, but monoclonal versions that bind bbPTMs regardless of sequence context are still lacking. Presumably, such pan-specific reagents will be easier to elicit for modifications that involve major changes to the backbone (e.g.,azole heterocycles) and more challenging for more subtle modifications (e.g., N-methyl and D-amino acids). Where known, the enzymes that are responsible for modifying the polypeptide backbone might also find application in the identification of unknown substrates as discussed above for PIMT. Chemical genetic strategies originally developed for interrogating kinases via engineered kinase–ATP analogue pairs (“bump-hole method”), which allows covalent capture of substrates, might be harnessed for similar quests involving bbPTM–enzyme systems. It is worth noting that lysine methyltransferases have recently been repurposed to transfer alkyne handles enabling isolation of their products. These modified cofactor analogues might also be accommodated by suitably engineered backbone methyltransferases.

■ TOOLS FOR DETERMINING PROTEIN BACKBONE PTM FUNCTIONS

In many cases, the exact role that backbone modifications play is unclear, and there is a pressing need for effective tools to characterize them. Many technologies developed for side chain PTMs can be adapted, but additional innovations are necessary to accommodate the specific requirements of bbPTMs. Insight into the function of PTMs in cells is typically gleaned from the constitutive produce or are devoid of a given modification. Where known, genetic or chemical manipulation of the modifying enzymes or mutation of the modified residues is exploited to interrogate PTMs. Currently, however, few such enzymes are known in the context of bbPTMs, and backbone atoms cannot be changed directly by site-directed mutagenesis.

To understand how bbPTMs control proteins mechanistically, biochemical and biophysical dissection of side-specifically modified proteins will be crucial. Where available, enzymes that install PTMs can be deployed to install modifications. Often, however, insufficient activity and specificity of enzymes in vitro or in recombinant hosts can hamper these efforts. Synthetic proteins have been harnessed extensively to explore the roles of side chain PTMs, for example, in chromatin biochemistry, and are ripe for investigating protein backbone modifications. In fact, a variety of building blocks with unnatural backbones have already been installed in proteins using synthetic chemical and biological methods, frequently for biophysical investigations of protein folding. To infiltrate ribosomal protein synthesis with backbone-modified building blocks, chemical and biological methods for charging tRNAs with unnatural monomers such as N-alkyl- or C$_\text{\-dialkyl}$ amino acids have been developed (Figure 7c). In certain cases, the ribosome itself and other elements of the translation machinery must be reengineered to promote the synthesis of modified backbones in addition to standard peptide bonds. This challenging feature has been achieved, for example, through selecting ribosomes that are sensitive to backbone-modified versions of the translation-inhibiting antibiotic puromycin. Chemical methods, i.e., solid phase peptide synthesis, are ideally suited for the incorporation of backbone-modified building blocks, and specialized protocols for accommodating a variety of bbPTMs have been developed. Amide bonds involving D-amino acids, C$_\text{\-dialkyl}$ amino acids, $\beta$-amino acids (i.e., isoAsp), and N-methyl amino acids can be formed using standard coupling chemistries, although the added bulk of some of these modifications can pose a challenge. Thiouamides can be accessed through the substitution of sulfur for oxygen at the building block level, and cyclic structures such as oxazoles can be incorporated via dipeptide analogues. Importantly, synthetic peptides can be elaborated into full-length proteins bearing site-specific bbPTMs through the use of convergent, chemo-selective strategies. Native chemical ligation is an ideal method for this purpose because both synthetic and recombinant fragments can be employed and its mild conditions are compatible with most backbone modifications (Figure 7c). A few backbone modifications can even be installed by site-specific modification of full-length proteins. Dehydroalanine, for example, can be accessed selectively from cysteine residues via mild elimination reactions. Thus, a suite of chemical biology tools are available to probe in detail the functional consequences and mechanism of action of protein backbone modifications, but adapting these methods for specific backbone chemistries as well as the development of new technologies will be required to fully appreciate the role of protein backbone modifications in biology.

■ SUMMARY

The examples presented throughout this Perspective demonstrate that site-specific PTMs of protein backbones are chemically feasible and biologically relevant. The functions and mechanisms of action of bbPTMs are as diverse as their chemistries. Autocatalytic condensation reactions on the backbone provide unique conjugated systems with valuable photophysical and catalytic properties. The spontaneous nature of the formation of some bbPTMs can be harnessed for molecular timers as well as pH or redox sensors and thus provide elegant ways to regulate signaling pathways. In addition, bbPTMs contribute to pushing the limits of protein stability and activity by fine-tuning the positioning of functional groups and through new structural motifs. Such features are particularly important for proteins that operate under extreme conditions, including hyperthermophilic proteins and enzymes situated in high-flux metabolic pathways.

Despite these fascinating examples, there is little insight into the distribution of protein backbone modifications. While bbPTMs have been observed in proteins from all domains of life, to date, enzymes that specifically introduce bbPTMs into proteins have been identified only in prokaryotes. Nevertheless, enzymes with backbone-modifying activities also exist in multicellular eukaryotes. For example, animals harbor enzymes to install D-amino acids in peptides; fungi exploit backbone amide N-methyltransferases that produce bioactive peptides, and catalysis of formation of isoAsp by a human enzyme has been demonstrated in vitro. Moreover, it is currently unclear whether specific binding domains for recognizing bbPTMs exist. Such domains are common for side chain modifications and orchestrate tunable protein–protein interaction networks. Certainly, bbPTMs can modulate protein–protein interactions by reshaping linear motifs and binding surfaces, yet the extent of their use in signaling and their interplay with side chain PTMs remain to be explored. Recent improvements in all aspects of protein science, including structural biology and analytical and synthetic
methodology, as well as the availability of genetic tools in diverse organisms make the study of bbPTMs a timely pursuit. Systematic application of this refined tool kit and the development of novel strategies will be required to tie the isolated case studies of bbPTMs into a cohesive field of biochemistry. The resulting understanding of the properties of naturally occurring backbone modifications will likely also inspire new techniques for artificially manipulating proteins. Following the success of backbone engineering in peptides, protein backbone engineering may pave the way to therapeutic proteins and biocatalysts with extraordinary activities and durability.

AUTHOR INFORMATION

Corresponding Author

*E-mail: manuel.muller@kcl.ac.uk.

ACKNOWLEDGMENTS

The author thanks Dr. Sarah Barry, Dr. Roger Morris, and Karola Gerecht for valuable discussions.

REFERENCES

deamidation is a critical switch in the regulation of the response to DNA damage. Cell 111, 51–62.


