Received: 2 April 2018 | Revised: 20 June 2018 | Accepted: 22 June 2018

DOI: 10.1002/pep2.24091

WILEY PeptideScience
THE AMERICAN PEPTIDE SOCIETY JOURNAL

REVIEW

Peptides meet ubiquitin: Simple interactions regulating complex cell signaling

Gianluca Veggiani[1,2] | Sachdev S. Sidhu[1,2,3]

[1]Donnelly Centre for Cellular and Biomolecular Research, Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario, Canada

[2]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

[3]Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada

Correspondence
*Sachdev Sidhu, Department of Molecular Genetics, University of Toronto, Toronto, ON M5S3E1, Canada.
Email: sachdev.sidhu@utoronto.ca

Funding information
Canadian Institutes of Health Research, Grant/Award Numbers: FDN- MOP-136956, FDN-MOP-136956

Abstract

The interplay between the ubiquitin proteasome system (UPS) and diverse peptide motifs controls almost every aspect of cell homeostasis. To achieve such an impressive functional diversity nature has evolved hundreds of different peptide motifs that interact with proteins in the UPS and ubiquitin itself to generate an immense network of interactions. Short peptides embedded in proteins are involved in controlling the intracellular levels of proteins as well as in the translation of ubiquitin signals into biochemical events. Therefore, it is not surprising that dysregulation of such interactions is associated with many diseases, including metabolic syndromes, neurodegenerative disorders, and cancer. We review the structural and functional features of peptide motifs interacting with the UPS and their use for generating protein-protein interaction inhibitors.

KEYWORDS

degrons, protein engineering, ubiquitin, ubiquitin-binding domains

1 | INTRODUCTION

Ubiquitin is a highly conserved and abundant 76-amino acid regulatory protein that is covalently conjugated to cellular proteins to modulate virtually every cellular process, including protein degradation, protein trafficking, cell-cycle regulation, signal transduction, and DNA repair.[1] Ubiquitination, the process of ubiquitin conjugation to proteins, involves the formation of an isopeptide bond between the ubiquitin carboxy-terminal glycine and the N-terminus or ε-amino groups of lysine residues in target proteins.[2] The ubiquitination reaction is catalyzed by the coordinated activity of 3 enzymes: a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2), and a ubiquitin ligase (E3). Upon adenosine triphosphate (ATP)-dependent activation by the E1 enzyme, ubiquitin is transferred to an E2 conjugating enzyme, and through the interaction of E2 with substrate-specific E3 enzymes, it is covalently attached to target proteins.[3–5]

Ubiquitin contains 7 lysine residues (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63) that, in addition to its N-terminus, can act as target sites for the conjugation of other ubiquitin molecules, allowing the formation of homotypic, heterotypic, or branched polyubiquitin chains to generate a highly complex molecular code (Figure 1).[6] Ubiquitination can be reversed by the activity of deubiquitinating enzymes (DUBs), ubiquitin-specific proteases that cleave mono and polyubiquitin chains with preferred linkage specificity.[5,7]

Early research was focused on the role of ubiquitin as a molecular tag for protein degradation by the proteasome 26S.[8] More recently, the non-degradative roles of ubiquitin have emerged as major regulators of cell homeostasis, including DNA repair and endocytosis.[9,10] Such activity is mediated by the non-covalent interaction of ubiquitin-modified proteins with diverse and divergent peptide motifs that are involved in the regulation of signal transduction networks. Among these peptide motifs, ubiquitin-binding domains (UBDs) act as readers of the complex ubiquitin code.[11–13] UBDs are modular domains (20-150 amino acids) embedded in proteins that bind to ubiquitinated proteins to translate ubiquitin signals into cellular outcomes.

This review will describe the role of peptides in the ubiquitin proteasome system (UPS) and their role in regulating signaling networks and maintaining cell homeostasis. We will describe how short peptide sequences encoding degradation signals, known as degrons, upon interaction with E3 ligases, control the half-life of a protein in vivo.[14] Moreover, we will focus on the interplay between peptide motifs and the UPS, with a particular focus on the readers of the ubiquitin code. We will confine this review to simple UBDs shorter than 50 residues, and we will describe the role of peptide motifs in regulating protein homeostasis and their implications for cell biology and biotechnology.

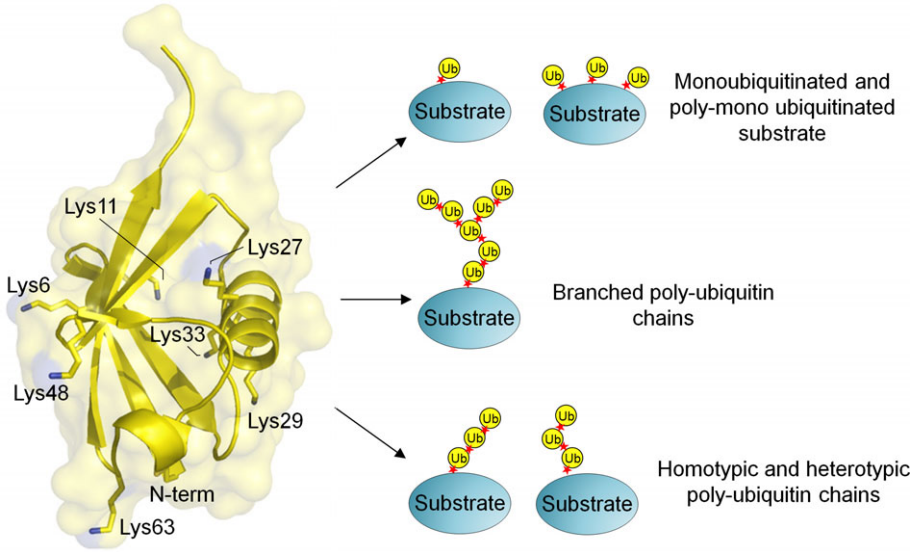


FIGURE 1 Structural features of the ubiquitin code. Ubiquitin (represented as cartoon and surface, PDB entry: 1UBQ) can act as a scaffold for the generation of complex polyubiquitin chains. Seven lysine residues (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63, rendered as sticks) and the N-terminal methionine can act as sites for ubiquitin modifications

2 | DEGRONS: PEPTIDES FOR DESTRUCTION

Removal of misfolded and/or damaged cellular proteins is a fundamental process for regulating cell homeostasis.[15,16] The UPS is the principal pathway for proteolysis, with the ubiquitination of protein targets by substrate-specific E3 ligases acting as the main specificity determinant for determining protein fate.[17] Recently, a more detailed understanding of the degradation machinery has revealed that specific and localized peptide sequences in proteins, called degrons, are responsible for fine-tuning components of the ubiquitin ligase machinery for their substrates.[14]

Degrons are short peptide sequences often found in N- or C-terminal regions of proteins, and they encode spatiotemporally controlled signals for the recognition of protein substrates by the ubiquitin ligase machinery and the proteasome.[14] A comprehensive and curated list of degrons can be found in the eukaryotic linear motif (ELM; http://elm.eu.org)[18] and the Dosztányi laboratory databases (http://dosztanyi.web.elte.hu/CANCER/DEGRON), and these analyses have revealed that amino acids involved in the interaction with the proteolytic machinery are evolutionary conserved.[19,20]

Due to their small interaction surfaces (typically 5-10 residues), degrons usually interact with specific E3 ligases with low affinity.[21,22] However, the presence of multiple degrons in close proximity has been shown to confer stronger binding.[23] An example of such an interaction is presented by the recognition of Cyclin E by the Fbw7 (F-box and WD repeat domain-containing 7) E3 ligase. Cyclin E contains 2 degron motifs with low affinity, which act in a cooperative manner to increase the affinity for Fbw7 dimers, thus improving Cyclin E turnover (Figure 2A).[23] In addition to multivalency, increased affinity of degrons for specific E3 ligases is also achieved using interaction surfaces that extend beyond the degron motif.[24]

Furthermore, the activity of degrons is often regulated by post-translational modifications (PTMs) that act as molecular switches governing degron accessibility in response to cellular signals.[18] For example, phosphorylation of the degron motif of the cyclin-dependent kinase inhibitor p27 is required for its recognition by the F-box SKP2 (S-phase kinase associated protein 2) E3 ligase (Figure 2B).[25]

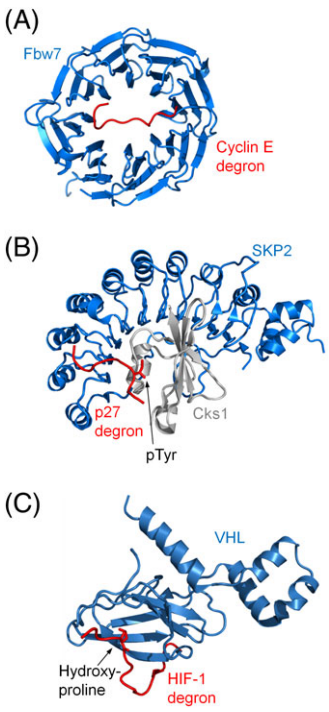


FIGURE 2 Structures of E3 ligase:degron complexes. A, the C-terminal degron motif of Cyclin E (red) binds in an extended fashion to the 8-bladed β-propeller of Fbw7 SCF E3 ligase (blue; PDB entry: 2OVQ). B, The degron motif of p27 (red) interacts with Cks1 (gray) and the E3 ligase Skp2 (blue). A phosphorylated tyrosine (pTyr) in the degron (shown as sticks) inserts into the Cks1 protein (PDB entry: 2AST). C, The oxygen-dependent degron (ODD) of HIF-1 (red) binds in an extended conformation to the E3 ligase VHL (blue, PBD entry: 1LQB). A hydroxyproline residue in the degron (shown as sticks) is almost entirely buried in a binding pocket on VHL

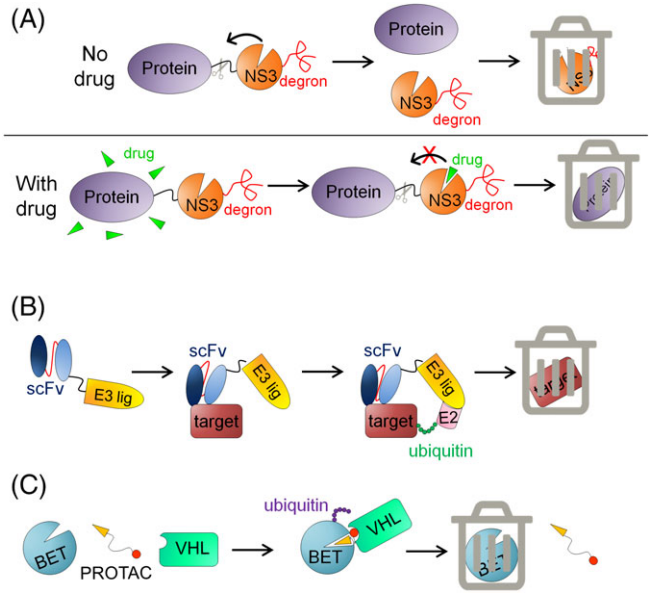


FIGURE 3 Control of protein stability with engineered degrons. A, schematic representation of the SMASh system. A protein of interest is fused to the NS3 protease from hepatitis C virus (HCV) carrying a degron motif, via a linker containing a NS3 proteolytic site. In the absence of NS3 inhibitor (upper panel), the protease can cleave the linker connecting NS3 to the protein, avoiding its degron-dependent ubiquitination and degradation. In the presence of inhibitor, the proteolytic activity of NS3 is blocked, thus inducing the rapid degradation of the protein due to the degron activity. B, Schematic representation of the ubiquibody system. A ubiquibody is a fusion protein with a single-chain Fv (scFv) antibody linked to an E3 ligase. Upon recognition of a target protein by the scFv, the E3 ligase conjugates ubiquitin to the target, thus activating protein degradation. (C) Schematic representation of the PROTAC technology. PROTACs are heterodimeric molecules composed of a protein-targeting module linked to a molecule capable of engaging an E3 ligase. By simultaneously interacting with the BET protein and the VHL E3 ligase, the BET PROTAC enables the recruitment of VHL, which ubiquitinates BET and causes its degradation

Phosphorylation of degrons can also abolish binding of E3 ligases and thus prevent proteolysis. For example, phosphorylation of degrons in proximity to the D-box of the prereplicative complex protein Cdc6 (cell division cycle 6) prevents recognition by the cyclin-dependent kinase 1 (Cdk1), thereby stabilizing Cdc6.[26]

In addition to phosphorylation, other PTMs and cellular stimuli also regulate ubiquitination and degradation of proteins.[27–30]

An example of stimuli-responsive degradation is provided by the hypoxia-inducible factor-1 (HIF-1) under normoxic conditions. Hydroxylation of Pro564 in the oxygen-dependent degron of HIF-1 enables tight binding of the cullin-RING ubiquitin ligase von Hippel-Lindau protein (VHL) to HIF-1, thereby facilitating HIF-1 turnover (Figure 2C).[31]

2.1 | Degron engineering to modulate protein expression

An important feature of degrons is that they are transferable elements that can be genetically fused to the termini of proteins to destabilize otherwise long-lived proteins. Several groups have exploited this property to develop new therapeutic strategies and cellular biosensors.

Chung et al. developed a general method (SMASh system) for achieving the drug-induced suppression of proteins of interest by genetically fusing the drug-sensitive NS3 protease and a degron to fine-tune the expression of proteins using inhibitors of the protease.[32] Intramolecular proteolytic activity of NS3 results in removal of the degron element from the protein of interest, thus enabling its expression. Conversely, inhibition of NS3 protease prevented the removal of the degron motif from the fusion protein, thereby promoting its degradation (Figure 3A).[32]

Fusion of single-domain antibodies to E3 ligases proved fruitful in achieving targeted protein knockout (Figure 3B),[33,34] paving the route to efficient post-translational protein knockout. For example, by replacing the natural substrate-binding domain of the human E3 ubiquitin ligase CHIP (carboxyl terminus of Hsc70-interacting protein) with a single-chain Fv (scFv) targeting β-galactosidase or with a fibronectin type III domain monobody specific for maltose binding protein, Portnoff et al. generated ubiquibodies able to efficiently deplete their respective targets in vivo.

Another approach exploiting degron-encoded destruction signals is exemplified by the PROTAC (PROteolysis TArgeting Chimera) technology.[35] PROTACs are heterodimeric small molecules composed of a targeting module conjugated to a recognition element for an E3 ligase. A PROTAC binds to a protein of interest and acts as a bridge linking the target protein to an E3 ligase, resulting in ubiquitination and degradation of the target.[36,37] For example, a PROTAC, capable of binding the bromodomains of the bromodomain and extra-terminal (BET) protein family and recruiting the VHL E3 ligase (Figure 3C), efficiently induced tumor regression in mouse models of castration-resistant prostate cancer.[38]

3 | UBDS: READERS OF THE UBIQUITIN CODE

The highly complex code of ubiquitin signals is written by E3 ligases, erased by DUBs, and read by a collection of modular protein domains, known as UBDs. These small folded domains (20-150 residues) are found embedded in numerous intracellular ubiquitin-binding proteins with highly diverse cellular localization, function, and structure, thus highlighting the major role of UBDs in interpreting and translating ubiquitin signals. In fact, as a consequence of their importance in regulating cell physiology, an increasing number of reports have correlated dysregulation of UBDs with a variety of diseases, including autoimmune disorders,[39] skeletal defects,[40] and cancer.[41]

More than 20 UBD families adopting structurally diverse folds have been identified through computational motif analysis and biochemical methods (Table 1). The affinities of UBDs for ubiquitin is generally weak (10-500 μM), enabling the generation of a transient and dynamic network of interactions, which translate the ubiquitin code into molecular events in cells.

Most UBDs interact with a common surface of ubiquitin centered on Ile44 (Ile44 patch),[42] but they use a variety of mechanisms to achieve affinity and specificity. UBDs are often found in tandem to

facilitate cooperative binding, resulting in improved binding affinity and selectivity. For example, analysis of the ubiquitin-interacting domains (UIMs) of the receptor-associated protein 80 (RAP80) revealed weak affinities of individual UIMs for Lys63-linked di-ubiquitin, but binding improved more than an order of magnitude when UIMs were used in tandem.[43] Moreover, the tandem arrangement of RAP80 UIMs resulted in a 20-fold increase in binding preference for Lys63-linked di-ubiquitin over monomeric ubiquitin, indicating that UBDs exploit avidity to achieve specificity.[43] In addition to multivalent binding of UBDs in the same protein, avidity and cooperativity have also been observed for UBDs within different proteins in multiprotein complexes.[44]

Structural analysis of UBDs bound to ubiquitin chains highlighted striking differences in such interactions. In fact, the intrinsic structural heterogeneity of ubiquitin allows the chains to adopt distinct and diverse conformations, thus controlling the binding specificity of each individual UBD for a particular ubiquitin chain.[45–47] Binding specificity of ubiquitin readers for particular chain linkages is further regulated by the nature of the linker connecting tandem UBDs. For example, the 8-residue linker between the two UIMs of AIRAPL (arsenite inducible RNA-associated protein like protein) confers high flexibility that enables simultaneous binding to Lys48-linked tri-ubiquitin chains.[48] A similar recognition mechanism has been observed in RAP80,[49] where two UIMs connected by a highly flexible 7-residue linker enable precise positioning of the UIMs for high-affinity interactions with Lys63-linked ubiquitin chains.[50] Conversely, the dipeptide linker between the first and second UIMs of the DUB Ataxin-3 defines the binding preference of the protein for Lys48-linked di-ubiquitin.[43] Notably, grafting of the Ataxin-3 linker between the UIMs of RAP80 altered the binding selectivity of the domains, resulting in a binding preference for Lys48-linked di-ubiquitin over Lys63-linked di-ubiquitin, highlighting the importance of the linkers between UIMs for conferring specificity.[43]

In addition to the linker spacing tandem UBDs and the flexibility of ubiquitin chains, the binding specificity of a UBD is determined by its structural-fold. Alpha helical UBDs (eg, UBA, CUE, UIM, VHS, UBAN, GAT) bind a common face of ubiquitin centered on the Ile44 patch, whereas domains consisting predominantly of beta sheets (NZF, UBC, WD40) can interact with more diverse surfaces (Table 1).[51] For example, the zinc-finger UBD (ZnF UBP) of isopeptidase T (IsoT) does not interact with Ile44, but rather, makes

TABLE 1 UBD structural families

UBD	Length (amino acids)	Binding surface on Ub	K_d mono-Ub (μM)	Examples (PDB ID)
UBA	~45	Ile44 patch	~14-500	1WR1
CUE	~45	Ile44 patch	~2-160	2LVO
UIM	~20	Ile44 patch	~100-500	4XKH
MIU	~30	Ile44 patch	~29	2C7M
DIUM	~20	Ile44 patch	~200	2D3G
VHS	~150	Ile44 patch	~50-200	3LDZ
UBAN	~90	Ile44 patch, Phe4	100	5B83
GAT	~120	Ile44 patch	~180	1O3X
NFZ	~35	Ile44 patch	~100-400	1Q5W
Ubc	146	Ile44 patch	~300	3FUH
WD40	~300	Ile44 patch	~200	3ODT
UBZ	~30	Ile44 patch	~100	2MRE
UEV	~145	Ile44 patch	~100-500	1S1Q
PFU	~100	Ile44 patch	~900	3PSP
GLUE	~145	Ile44 patch	~500	2DX5
ZnF A20	~70	Asp58	~20	2C7N
ZnF UBP	~110	Ile36 patch, Ub C-term	2.8	2G43
PRU	150	Ile44 patch	65	2Z59
SH3	~60	Ile44 patch, G47	~12-200	2K6D
UBM	~35	Ile44 patch	~70	2KWV

Affinity values are for the interaction of UBDs with monomeric ubiquitin (mono-Ub). Protein data bank identification codes (PDB ID) are provided for examples of UBDs in complex with ubiquitin. Abbreviations: CUE, coupling of ubiquitin conjugation to endoplasmatic reticulum degradation; DIUM, double-sided ubiquitin-interacting motif; GAT, GGA1 and Tom1 domains; GLUE, GRAM-like ubiquitin-binding in Eap45; MIU, motif interacting with ubiquitin; NZF, Npl4 zinc finger; PFU, PLAA (phospholipase A2 activating protein) family ubiquitin-binding domain; PRU, plekstrin homology receptor for ubiquitin; SH3, Src homology 3; UBA, ubiquitin-associated domain; UBAN, ubiquitin-binding in ABIN and NEMO; Ubc, ubiquitin-conjugating; UBM, ubiquitin-binding motif; UBZ, ubiquitin-binding zinc finger; UEV, ubiquitin E2 variant; UIM, ubiquitin-interacting motif; VHS, Vps27, HRS, STAM; ZnF A20, tumor necrosis factor-inducible zinc-finger protein A20; ZnF UBP, ubiquitin C-terminal hydrolase-like zinc finger.

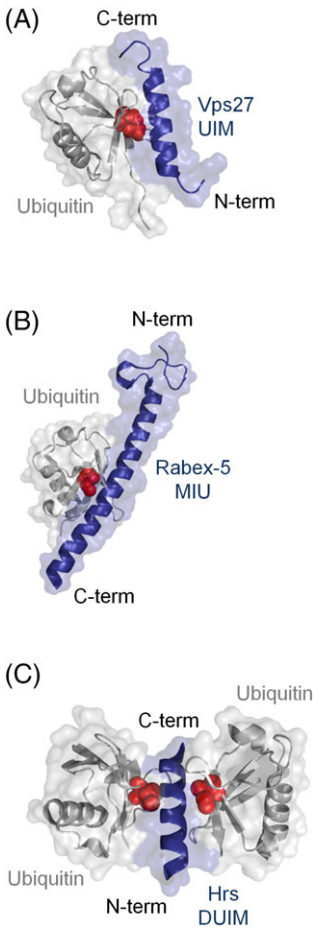


FIGURE 4 Recognition of ubiquitin by UIMs, DIUMs, and MIUs. Ubiquitin is depicted as a gray cartoon and surface with Ile44 shown as a red sphere. The helical domains (blue) are depicted as cartoons and surface for (A) the UIM of Vps27 (PDB entry: 1Q0W), (B) the MIU domain of Rabex-5 (PDB entry: 2C7M), and (C) the DIUM of endosome-associated protein Hrs (PDB entry: 2D3G)

hydrophobic interactions with a patch centered on Ile36 and with the C-terminus, which penetrates into a pocket on the UBD.[51]

A further demonstration of the versatility of ubiquitin readers is provided by structural analyses of UIMs, MIUs (motifs interacting with ubiquitin), and DUIMs (double-sided ubiquitin-interacting motifs) (Figure 4). All these UBDs consist of a single alpha-helix that makes extensive contacts with the Ile44 patch.[52–54] MIUs bind ubiquitin in a fashion similar to UIMs, but in the opposite orientation, enabling the formation of additional hydrogen bonds and electrostatic interactions, which increase affinity by an order of magnitude relative to most UIMs (Figure 4A,B).[53] The structures of DUIMs and UIMs are highly similar and the domains have analogous binding modes for recognition of the Ile44 patch.[54] However, whereas in UIMs one side of the alpha-helix is solvent-exposed (Figure 4A), in DUIMs both faces of the helix are capable of binding to ubiquitin (Figure 4C).[54]

In addition to binding ubiquitinated proteins, some UBDs can mediate the ubiquitination of proteins containing them, through a mechanism known as coupled ubiquitination.[55] In fact, UBDs enable the recruitment of the ubiquitination machinery, thus promoting protein ubiquitination and degradation. For example, Oldham et al. demonstrated that the UIMs of epsin, a protein involved in creating cell membrane curvature, are necessary for its ubiquitination, and that fusion of the UIMs to glutathione-S-transferase (GST) was sufficient to mediate GST ubiquitination.[56] In another example, the CUE (coupling of ubiquitin conjugation to endoplasmatic reticulum degradation) domain of Vps29 (vacuolar protein sorting-associated protein 29) recruits the Rsp5 ubiquitin ligase to promote ubiquitination of Vps29.[57] Coupled ubiquitination can also occur in an E3 ligase-independent manner, as in the case of GST-UBD fusions that were ubiquitinated in vitro in the presence of E1 and E2 enzymes without an E3 ligase.[55] Thus, UBD-containing proteins can be regulated in multiple ways, as their coupled ubiquitination promotes the intramolecular binding of UBDs to the attached ubiquitin, thus inhibiting their interaction with other ubiquitinated proteins.[58]

3.1 | UBDs engineered as molecular tools

Inspired by the large variety of naturally occurring peptide motifs interacting with ubiquitin, several groups have developed molecular tools to enhance our understanding of the UPS. TUBEs (tandem ubiquitin-binding entities) were developed by exploiting the UBA1 (ubiquitin-associated domain-1) domain of the human HR23A protein and the UBA domain of ubiquilin (Figure 5A).[59] Fusing tandem repeats of UBDs spaced by a flexible linker to GST enabled tight binding to both Lys48 and Lys63 polyubiquitin chains, as well as the ability to isolate ubiquitinated proteins from cell extracts.[59] Moreover, TUBEs protected ubiquitinated proteins from proteolysis by DUBs and thus preserved them for downstream analysis.[59]

A similar approach has been used to isolate naturally occurring unanchored ubiquitin chains not conjugated to proteins.[60] Although unanchored polyubiquitin chains have emerged as important regulators of the immune response,[61] autophagy,[62] and kinase activity,[63] the role of free polyubiquitin chains in controlling cell physiology is still elusive. Scott et al. constructed a chimeric protein (t-UBD) by fusing the UBA2 motif of HR23A protein to the ZnF UBP of USP16 (Figure 5A).[60] By combining the preference of the UBA2 motif for Lys48-linked ubiquitin with the ability of the ZnF UBP domain to bind the free C-terminus of ubiquitin, the authors developed a tool for efficient and selective purification of Lys48-linked unanchored polyubiquitin chains.[60]

A peptide mimetic of the UIM of epsins has been developed as a potential cancer therapy.[64] Epsin UIMs are involved in promoting the endocytosis and degradation of VEGFR2 (vascular endothelial growth factor receptor 2), and their deletion in mouse models surprisingly resulted in impaired tumor angiogenesis and retarded tumor growth, thus providing a potential new therapeutic strategy.[65] By fusing a mimetic peptide of epsin UIMs to a plasma membrane-anchoring peptide and a tumor-homing peptide (Figure 5B), Dong et al. were able to stimulate aberrant tumor angiogenesis in vivo, which reduced cancer growth and increased survival rates in glioma murine models.[64]

Tandemly arranged UIMs have been engineered to develop high-affinity polyubiquitin-sensor proteins for in vivo tracking of linkage-specific ubiquitin signals (Figure 5C).[66] By optimizing the rigidity and the length of the linker spacing three Vps27 (vacuolar protein sorting-associated protein 27) UIMs, Sims et al. obtained a 1000-fold binding preference for Lys63-linked ubiquitin chains over Lys48-linked chains.[66] Moreover, these sensors proved to be useful probes of the Lys63-linked polyubiquitin dependence of NF-kB (nuclear factor kappa-light-chain-enhancer of activated B cells) activation and Interleukin-1β (IL-1β) signaling.[66]

3.2 | Ubiquitin variants as probes of the UPS

Rather than engineering UBDs to interact selectively with ubiquitin chains, useful probes of the UPS can also be developed by engineering ubiquitin variants (UbVs) to interact specifically with components of the UPS. Large phage-displayed libraries of UbVs, designed by us and others through computational and structural analysis of ubiquitin interactions, yielded specific inhibitors of USPs (ubiquitin-specific proteases).[67,68] Similarly, by randomizing solvent-exposed residues, we generated potent and selective UbV inhibitors of the SCF (Skp1-Cullin-F-box) E3 ligase family[69] and DUBs from pathogenic viruses,[70] and we developed both inhibitors and activators of the HECT (homologous to the E6AP carboxyl terminus) E3 ligase family.[71]

Recently, we obtained a UbV (UbV.v27.1) having unprecedented affinity for the N-terminal UIM domain of yeast Vps27.[72] Affinity analysis and specificity profiling revealed that UbV.v27.1 specifically bound to the target UIM domain more than 2 orders of magnitude more tightly than did wild-type ubiquitin (Ub.wt) (Figure 5D).[72] Structural analysis highlighted the molecular determinants of the improved affinity of UbV.v27.1 for the UIM. UbV.v27.1 contacted the UIM α-helix through the same surface used by Ub.wt, with four substituted residues forming an extended network of hydrophobic and hydrophilic interactions (Figure 5D).[72] However, mutagenesis studies revealed that only one of these contact substitutions contributed significantly to the enhanced affinity, and substitutions remote from the binding interface contributed substantially to the interaction.[72] Thus, the conformational plasticity of ubiquitin can be exploited to develop UbVs that bind with high-affinity and specificity even to short peptides.

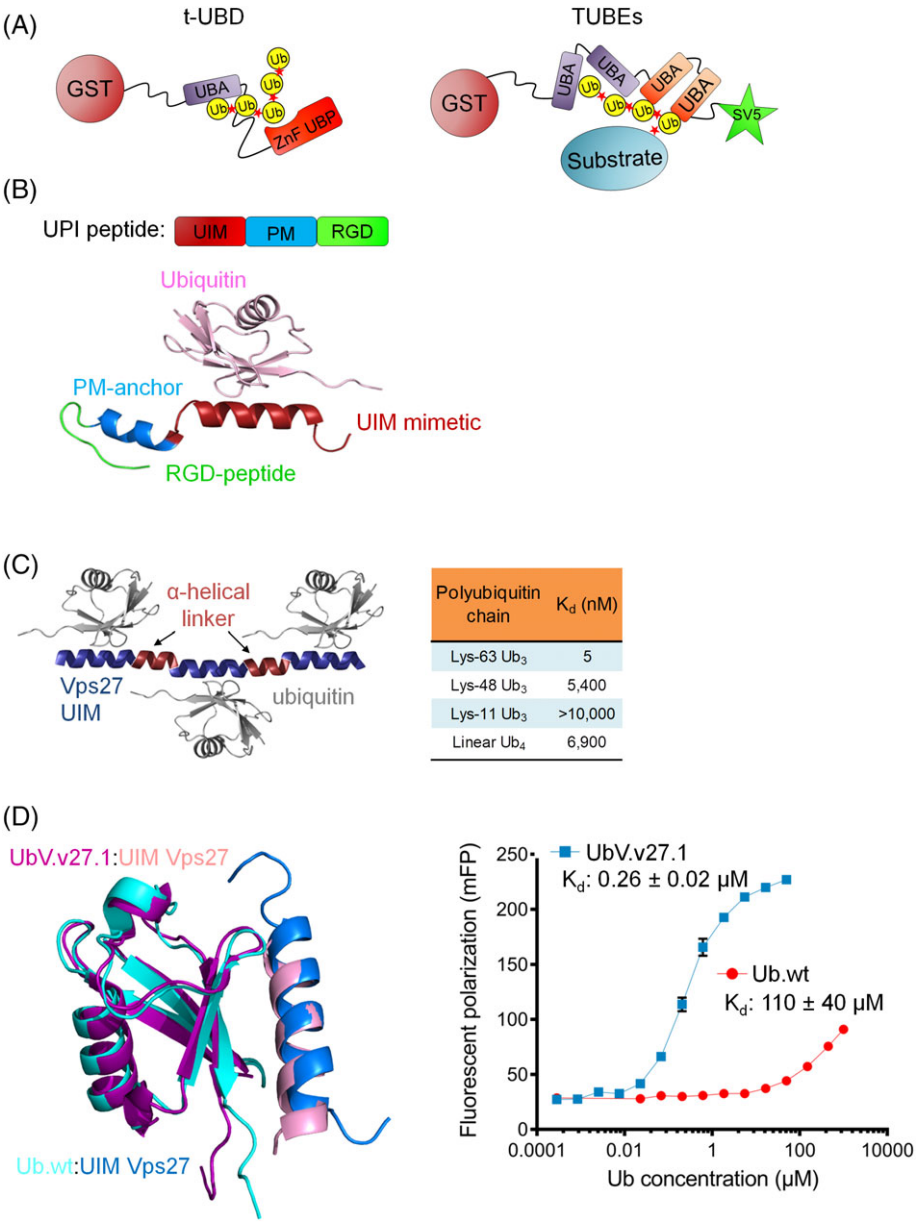


FIGURE 5 Applications of engineered UBDs and UbVs. A, Engineered UBDs have been developed to purify linkage-specific ubiquitin chains. TUBEs exploit the binding preference of UBA domains of ubiquilin and HR23A protein to isolate both Lys48 and Lys63 polyubiquitin chains conjugated to proteins. By combining the ability of the ZnF UBP domain to bind the ubiquitin free C-terminus with the Lys48-linked ubiquitin chain specificity of the UBA2 motif of HR23A protein, t-UBDs enable the purification of Lys48-linked unanchored polyubiquitin chains. B, Cartoon representation of the UPI mimetic peptide predicted using the PEP-FOLD tool.[73] The UPI mimetic peptide was obtained by fusing the epsin UIM domain to a plasma membrane-anchoring peptide (PM) and a tumor-homing peptide (RGD). C, To engineer a polyubiquitin sensor, 3 copies of the second UIM domain of Vps27 (blue, PDB entry: 1Q0W) were connected by linkers from the sequence spacing the two UIM domains of RAP80 (light red, PDB entry: 3A1Q). Binding affinity of the polyubiquitin-sensor Rx3(A7) was determined by fluorescence anisotropy. D, Superposition of the crystal structure of the complex of UbV.v27.1 (magenta) bound to the first UIM of Vps27 (pink, PDB entry: 5UCL) with the NMR structure of wild-type ubiquitin (Ub.Wt, cyan) bound to the same UIM (blue, PDB entry: 1Q0W). UbV.v27.1 and Ub.wt bind in a very similar fashion, but the UbV binds 500-fold more tightly

4 | CONCLUSIONS

Ubiquitination modulates numerous cell signaling pathways to elicit diverse biochemical responses that regulate a large number of cellular processes in health and disease. Such functional diversity based on the simple and highly conserved ubiquitin protein relies on interactions with thousands of enzymes, domains, and peptide motifs. Degrons are emerging as key regulators of the levels of intracellular proteins. Moreover, a growing number of studies are reporting the presence of mutations in the degrons of oncogenes, highlighting their relevance in the maintenance of cell homeostasis. Notwithstanding recent advances in our understanding of the regulatory activities of degrons, significant challenges lie ahead. So far, only ~25 classes of degrons have been characterized. Identification of new degrons and the understanding of their conformational dynamics will enhance our comprehension of the decision-making mechanisms of cells and will

enlarge the palette of biotechnological approaches for therapeutic interventions.

Hundreds of UBDs cooperate in the recognition of the ubiquitin code and its translation into diverse cellular events. Our knowledge of UBD structure and function continues to increase, but there are still many questions waiting to be addressed. Although the binding preference of some UBDs has been investigated extensively, the ability to distinguish between different polyubiquitin chains still remains elusive in the context of full-length proteins. In this regard, engineering of UBDs and ubiquitin has provided powerful strategies for elucidating ubiquitin-UBD interactions and for understanding how UBDs regulate protein localization and activity.

ACKNOWLEDGMENTS

This work was supported by Canadian Institutes of Health Research (CIHR) grant FDN-MOP-136956. The authors thank Drs. Joan Teyra, Kevin M. Davis, Mart Ustav Jr. and Jan Tykvart, and Mr. Gregory D. Martyn for helpful discussions.

ORCID

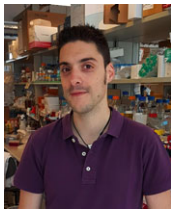
Gianluca Veggiani https://orcid.org/0000-0002-4064-0112

REFERENCES

[1] D. Komander, M. Rape, *Annu. Rev. Biochem.* **2012**, *81*, 203.
[2] L. Buetow, D. T. Huang, *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 626.
[3] I. Lee, H. Schindelin, *Cell* **2008**, *134*, 268.
[4] I. Valimberti, M. Tiberti, M. Lambrughi, B. Sarcevic, E. Papaleo, *Sci. Rep.* **2015**, *5*, 14849.
[5] A. Y. Amerik, M. Hochstrasser, *Biochim. Biophys. Acta Mol. Cell Res.* **1695**, *2004*, 189.
[6] R. Yau, M. Rape, *Nat. Cell Biol.* **2016**, *18*, 579.
[7] X. Zhang, A. H. Smits, G. B. A. van Tilburg, P. W. T. C. Jansen, M. M. Makowski, H. Ovaa, M. Vermeulen, *Mol. Cell* **2017**, *65*, 941.
[8] J. Smalle, R. D. Vierstra, *Annu. Rev. Plant Biol.* **2004**, *55*, 555.
[9] P. Xu, D. M. Duong, N. T. Seyfried, D. Cheng, Y. Xie, J. Robert, J. Rush, M. Hochstrasser, D. Finley, J. Peng, *Cell* **2009**, *137*, 133.
[10] K. A. Ball, J. R. Johnson, M. K. Lewinski, J. Guatelli, E. Verschueren, N. J. Krogan, M. P. Jacobson, *PLoS Comput. Biol.* **2016**, *12*, e1004898.
[11] I. Dikic, S. Wakatsuki, K. J. Walters, *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 659.
[12] L. Hicke, H. L. Schubert, C. P. Hill, *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 610.
[13] D. Scott, N. J. Oldham, J. Strachan, M. S. Searle, R. Layfield, *Proteomics* **2015**, *15*, 844.
[14] M. Guharoy, P. Bhowmick, M. Sallam, P. Tompa, *Nat. Commun.* **2016**, *7*, 10239.
[15] E. Reinstein, A. Ciechanover, *Ann. Intern. Med.* **2006**, *145*, 676.
[16] A. Ciechanover, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 5944.
[17] B. Sander, W. Xu, M. Eilers, N. Popov, S. Lorenz, *eLife* **2017**, *6*, e21036.
[18] K. Van Roey, H. Dinkel, R. J. Weatheritt, T. J. Gibson, N. E. Davey, *Sci. Signal.* **2013**, *6*, rs7.
[19] B. Mészáros, M. Kumar, T. J. Gibson, B. Uyar, Z. Dosztányi, *Sci. Signal.* **2017**, *10*, eaak9982.
[20] E. K. Schrader, K. G. Harstad, A. Matouschek, *Nat. Chem. Biol.* **2009**, *5*, 815.
[21] N. E. Davey, D. O. Morgan, *Mol. Cell* **2016**, *64*, 12.
[22] E. P. T. Hertz, T. Kruse, N. E. Davey, B. López-Méndez, J. O. Sigurðsson, G. Montoya, J. V. Olsen, J. Nilsson, *Mol. Cell* **2016**, *63*, 686.
[23] M. Welcker, E. A. Larimore, J. Swanger, M. T. Bengoechea-Alonso, J. E. Grim, J. Ericsson, N. Zheng, B. E. Clurman, *Genes Dev.* **2013**, *27*, 2531.
[24] W. Gan, X. Dai, A. Lunardi, Z. Li, H. Inuzuka, P. Liu, S. Varmeh, J. Zhang, L. Cheng, Y. Sun, J. M. Asara, A. H. Beck, J. Huang, P. P. Pandolfi, W. Wei, *Mol. Cell* **2015**, *59*, 917.
[25] B. Hao, N. Zheng, B. A. Schulman, G. Wu, J. J. Miller, M. Pagano, N. P. Pavletich, *Mol. Cell* **2005**, *20*, 9.
[26] R. W. King, M. Glotzer, M. W. Kirschner, *Mol. Biol. Cell* **1996**, *7*, 1343.
[27] X. Huang, V. M. Dixit, *Mol. Cell. Biol.* **2011**, *31*, 3682.
[28] T. Hunter, *Mol. Cell* **2007**, *28*, 730.
[29] M. Li, J. Luo, C. L. Brooks, W. Gu, *J. Biol. Chem.* **2002**, *277*, 50607.
[30] H.-B. Ruan, Y. Nie, X. Yang, *Mol. Cell. Proteomics MCP* **2013**, *12*, 3489.
[31] M. Ivan, K. Kondo, H. Yang, W. Kim, J. Valiando, M. Ohh, A. Salic, J. M. Asara, W. S. Lane, W. G. Kaelin, *Science* **2001**, *292*, 464.
[32] H. K. Chung, C. L. Jacobs, Y. Huo, J. Yang, S. A. Krumm, R. K. Plemper, R. Y. Tsien, M. Z. Lin, *Nat. Chem. Biol.* **2015**, *11*, 713.
[33] A. D. Portnoff, E. A. Stephens, J. D. Varner, M. P. DeLisa, *J. Biol. Chem.* **2014**, *289*, 7844.
[34] E. Caussinus, O. Kanca, M. Affolter, *Nat. Struct. Mol. Biol.* **2011**, *19*, 117.
[35] M. Toure, C. M. Crews, *Angew. Chem. Int. Ed Engl.* **2016**, *55*, 1966.
[36] G. M. Burslem, B. E. Smith, A. C. Lai, S. Jaime-Figueroa, D. C. McQuaid, D. P. Bondeson, M. Toure, H. Dong, Y. Qian, J. Wang, A. P. Crew, J. Hines, C. M. Crews, *Cell Chem. Biol.* **2018**, *25*, 67.
[37] T. K. Neklesa, J. D. Winkler, C. M. Crews, *Pharmacol. Ther.* **2017**, *174*, 138.
[38] K. Raina, J. Lu, Y. Qian, M. Altieri, D. Gordon, A. M. K. Rossi, J. Wang, X. Chen, H. Dong, K. Siu, J. D. Winkler, A. P. Crew, C. M. Crews, K. G. Coleman, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 7124.
[39] M. Kool, G. van Loo, W. Waelput, S. De Prijck, F. Muskens, M. Sze, J. van Praet, F. Branco-Madeira, S. Janssens, B. Reizis, D. Elewaut, R. Beyaert, H. Hammad, B. N. Lambrecht, *Immunity* **2011**, *35*, 82.
[40] R. Layfield, B. Ciani, S. H. Ralston, L. J. Hocking, P. W. Sheppard, M. S. Searle, J. R. Cavey, *Biochem. Soc. Trans.* **2004**, *32*, 728.
[41] M. Gyrd-Hansen, M. Darding, M. Miasari, M. M. Santoro, L. Zender, W. Xue, T. Tenev, P. C. A. da Fonseca, M. Zvelebil, J. M. Bujnicki, S. Lowe, J. Silke, P. Meier, *Nat. Cell Biol.* **2008**, *10*, 1309.
[42] J. H. Hurley, S. Lee, G. Prag, *Biochem. J.* **2006**, *399*, 361.
[43] J. J. Sims, R. E. Cohen, *Mol. Cell* **2009**, *33*, 775.
[44] T. P. Garner, J. Strachan, E. C. Shedden, J. E. Long, J. R. Cavey, B. Shaw, R. Layfield, M. S. Searle, *Biochemistry* **2011**, *50*, 9076.
[45] O. F. Lange, N.-A. Lakomek, C. Farès, G. F. Schröder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, B. L. de Groot, *Science* **2008**, *320*, 1471.
[46] T. T. Thach, D. Shin, S. Han, S. Lee, *Acta Crystallogr.* **2016**, *72*, 524.
[47] A. B. Datta, G. L. Hura, C. Wolberger, *J. Mol. Biol.* **2009**, *392*, 1117.
[48] S. Rahighi, I. Braunstein, N. Ternette, B. Kessler, M. Kawasaki, R. Kato, T. Matsui, T. M. Weiss, A. Stanhill, S. Wakatsuki, *Structure* **2016**, *24*, 412.
[49] J. Yan, Y.-S. Kim, X.-P. Yang, L.-P. Li, G. Liao, F. Xia, A. M. Jetten, *Cancer Res.* **2007**, *67*, 6647.
[50] Y. Sato, A. Yoshikawa, H. Mimura, M. Yamashita, A. Yamagata, S. Fukai, *EMBO J.* **2009**, *28*, 2461.
[51] F. E. Reyes-Turcu, J. R. Horton, J. E. Mullally, A. Heroux, X. Cheng, K. D. Wilkinson, *Cell* **2006**, *124*, 1197.
[52] D. J. Katzmann, C. J. Stefan, M. Babst, S. D. Emr, *J. Cell Biol.* **2003**, *162*, 413.
[53] L. Penengo, M. Mapelli, A. G. Murachelli, S. Confalonieri, L. Magri, A. Musacchio, P. P. Di Fiore, S. Polo, T. R. Schneider, *Cell* **2006**, *124*, 1183.
[54] S. Hirano, M. Kawasaki, H. Ura, R. Kato, C. Raiborg, H. Stenmark, S. Wakatsuki, *Nat. Struct. Mol. Biol.* **2006**, *13*, 272.
[55] D. Hoeller, C.-M. Hecker, S. Wagner, V. Rogov, V. Dötsch, I. Dikic, *Mol. Cell* **2007**, *26*, 891.
[56] C. E. Oldham, R. P. Mohney, S. L. H. Miller, R. N. Hanes, J. P. O'Bryan, *Curr. Biol.* **2002**, *12*, 1112.
[57] S. C. Shih, G. Prag, S. A. Francis, M. A. Sutanto, J. H. Hurley, L. Hicke, *EMBO J.* **2003**, *22*, 1273.
[58] S. L. H. Miller, E. Malotky, J. P. O'Bryan, *J. Biol. Chem.* **2004**, *279*, 33528.
[59] R. Hjerpe, F. Aillet, F. Lopitz-Otsoa, V. Lang, P. England, M. S. Rodriguez, *EMBO Rep.* **2009**, *10*, 1250.
[60] D. Scott, T. P. Garner, J. Long, J. Strachan, S. C. Mistry, A. R. Bottrill, D. J. Tooth, M. S. Searle, N. J. Oldham, R. Layfield, *Proteomics* **2016**, *16*, 1961.

[61] W. Zeng, L. Sun, X. Jiang, X. Chen, F. Hou, A. Adhikari, M. Xu, Z. J. Chen, *Cell* **2010**, *141*, 315.
[62] P. Nanduri, R. Hao, T. Fitzpatrick, T.-P. Yao, *J. Biol. Chem.* **2015**, *290*, 9455.
[63] Z.-P. Xia, L. Sun, X. Chen, G. Pineda, X. Jiang, A. Adhikari, W. Zeng, Z. J. Chen, *Nature* **2009**, *461*, 114.
[64] Y. Dong, H. Wu, H. N. A. Rahman, Y. Liu, S. Pasula, K. L. Tessneer, X. Cai, X. Liu, B. Chang, J. McManus, S. Hahn, J. Dong, M. L. Brophy, L. Yu, K. Song, R. Silasi-Mansat, D. Saunders, C. Njoku, H. Song, P. Mehta-D'Souza, R. Towner, F. Lupu, R. P. McEver, L. Xia, D. Boerboom, R. S. Srinivasan, H. Chen, *J. Clin. Invest.* **2015**, *125*, 4349.
[65] S. Pasula, X. Cai, Y. Dong, M. Messa, J. McManus, B. Chang, X. Liu, H. Zhu, R. S. Mansat, S.-J. Yoon, S. Hahn, J. Keeling, D. Saunders, G. Ko, J. Knight, G. Newton, F. Luscinskas, X. Sun, R. Towner, F. Lupu, L. Xia, O. Cremona, P. De Camilli, W. Min, H. Chen, *J. Clin. Invest.* **2012**, *122*, 4424.
[66] J. J. Sims, F. Scavone, E. M. Cooper, L. A. Kane, R. J. Youle, J. D. Boeke, R. E. Cohen, *Nat. Methods* **2012**, *9*, 303.
[67] Y. Zhang, L. Zhou, L. Rouge, A. H. Phillips, C. Lam, P. Liu, W. Sandoval, E. Helgason, J. M. Murray, I. E. Wertz, J. E. Corn, *Nat. Chem. Biol.* **2013**, *9*, 51.
[68] A. Ernst, G. Avvakumov, J. Tong, Y. Fan, Y. Zhao, P. Alberts, A. Persaud, J. R. Walker, A.-M. Neculai, D. Neculai, A. Vorobyov, P. Garg, L. Beatty, P.-K. Chan, Y.-C. Juang, M.-C. Landry, C. Yeh, E. Zeqiraj, K. Karamboulas, A. Allali-Hassani, M. Vedadi, M. Tyers, J. Moffat, F. Sicheri, L. Pelletier, D. Durocher, B. Raught, D. Rotin, J. Yang, M. F. Moran, S. Dhe-Paganon, S. S. Sidhu, *Science* **2013**, *339*, 590.
[69] M. Gorelik, S. Orlicky, M. A. Sartori, X. Tang, E. Marcon, I. Kurinov, J. F. Greenblatt, M. Tyers, J. Moffat, F. Sicheri, S. S. Sidhu, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 3527.
[70] W. Zhang, B. A. Bailey-Elkin, R. C. M. Knaap, B. Khare, T. J. Dalebout, G. G. Johnson, P. B. van Kasteren, N. J. McLeish, J. Gu, W. He, M. Kikkert, B. L. Mark, S. S. Sidhu, *PLoS Pathog.* **2017**, *13*, e1006372.
[71] W. Zhang, K.-P. Wu, M. A. Sartori, H. B. Kamadurai, A. Ordureau, C. Jiang, P. Y. Mercredi, R. Murchie, J. Hu, A. Persaud, M. Mukherjee, N. Li, A. Doye, J. R. Walker, Y. Sheng, Z. Hao, Y. Li, K. R. Brown, E. Lemichez, J. Chen, Y. Tong, J. W. Harper, J. Moffat, D. Rotin, B. A. Schulman, S. S. Sidhu, *Mol. Cell* **2016**, *62*, 121.
[72] N. Manczyk, B. P. Yates, G. Veggiani, A. Ernst, F. Sicheri, S. S. Sidhu, *Protein Sci. Publ. Protein Soc.* **2017**, *26*, 1060.
[73] Y. Shen, J. Maupetit, P. Derreumaux, P. Tufféry, *J. Chem. Theory Comput.* **2014**, *10*, 4745.

AUTHOR BIOGRAPHIES



Dr Gianluca Veggiani obtained his PhD in Biochemistry from the University of Oxford in 2016, under the mentorship of Professor Mark Howarth. His work focused on protein nanotechnologies for cancer analysis. Since 2016, he has worked as a postdoctoral fellow in the group of Dr Sachdev Sidhu at the University of Toronto, where his research focuses on engineering ubiquitin and readers of post-translational modifications to modulate cell behavior.

Dr Sachdev Sidhu obtained his PhD from Simon Fraser University in 1996, and subsequently worked at Genentech, Inc. from 1996 to 2008. Since 2008, he has been a Professor in the Department of Molecular Genetics at the Donnelly Centre in the University of Toronto. Dr Sidhu's research aims to develop novel protein engineering technologies for understanding and modulating cellular function, and for developing therapies for unmet medical needs.

How to cite this article: Veggiani G, Sidhu SS. Peptides meet ubiquitin: Simple interactions regulating complex cell signaling. *Peptide Science*. 2019;111:e24091. https://doi.org/10.1002/pep2.24091