



King's Research Portal

DOI:

[10.1016/j.tibs.2016.01.006](https://doi.org/10.1016/j.tibs.2016.01.006)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Alfano, C., Faggiano, S., & Pastore, A. (2016). The Ball and Chain of Polyubiquitin Structures. *TRENDS IN BIOCHEMICAL SCIENCES*, 41(4), 371-385. <https://doi.org/10.1016/j.tibs.2016.01.006>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Trends in Biochemical Sciences
The ball and chain of polyubiquitin structural studies
--Manuscript Draft--

Manuscript Number:	TIBS-D-15-00167R2
Article Type:	Review
Corresponding Author:	Annalisa Pastore London, UNITED KINGDOM
First Author:	Caterina Alfano
Order of Authors:	Caterina Alfano Serena Faggiano Annalisa Pastore
Abstract:	Ubiquitylation is a post-translational modification implicated in several different cellular pathways. The possibility to forming chains through covalent cross-linking between any of the seven lysines or the initial methionine to the C terminus of another moiety provides ubiquitin (Ub) with special flexibility in its function in signalling. Here, we review the knowledge accumulated over the last years about the functions and structural features of polyUb chains. This analysis reveals the need to understand further the functional role of some of the linkages and the structural code that determines recognition of polyUbs by protein partners.

Trend Box

- Ubiquitin (Ub), a small signalling protein, forms chains thanks to covalent linkage between any of the 7 lysines or the N-terminal methionine of one subunit and the C terminus of another Ub.
- PolyUb chains both with the same or mixed linkages are involved in several cellular functions from proteasomal targeting to protein regulation and have thus acquired an enormous importance in cellular signalling.
- PolyUbs exhibit a unique repertoire of conformational states that is dependent on the specific linkages, which have different flexibilities.
- Further complexity is added by the interaction of polyUbs with cellular partners, which modulate their structure and functions.
- Understanding the structural and functional aspects of the polyUb code continues to offer an important challenge.

Review

The ball and chain of polyubiquitin structures

Caterina Alfano¹, Serena Faggiano², Annalisa Pastore^{1*}

¹Department of Clinical and Basic Neuroscience, King's College London, London, United Kingdom

²Department of Pharmacy, University of Parma, Parma, Italy

*To whom correspondence should be addressed.

Email: annalisa.pastore@crick.ac.uk

Keywords: biophysics, post-translational modifications, proteasomal pathway, signalling, structure, ubiquitin

Running title: Structural studies of polyubiquitin chains

Ubiquitylation is a post-translational modification implicated in several different cellular pathways. The possibility of forming chains through covalent cross-linking between any of the seven lysines, or the initial methionine, and the C terminus of another moiety provides ubiquitin (Ub) with special flexibility in its function in signalling. Here, we review the knowledge accumulated over the past several years about the functions and structural features of polyUb chains. This analysis reveals the need to understand further the functional role of some of the linkages and the structural code that determines recognition of polyUbs by protein partners.

Why are polyUb chains interesting?

Post-translational modifications are one of the keys to understanding the complexity of eukaryotic genomes: capable of considerably changing protein surfaces and binding properties of a protein, they act as a common mechanism for modulating and regulating protein function. Eukaryotes have a large repertoire of post-translational modifications, which include phosphorylation, acetylation,

1 methylation, ubiquitylation, sumoylation, neddylation, glycosylation and glycation [1]. Among these,
2 ubiquitylation is a modification that brings particular flexibility both from structural and functional
3 perspectives. Ub (76 amino acid residues) is a ubiquitous globular protein (**Box 1**) that is evolutionarily
4 stable, being highly conserved among eukaryotes and sharing a sequence identity between human and
5 yeast as high as 96% [2].
6
7
8
9
10

11 Ubiquitylation occurs by the covalent attachment of the C-terminal carboxyl group of Ub to the ϵ -
12 amine of lysine residues of a substrate protein through an isopeptide bond. The acceptor can be either
13 a different protein or Ub itself. When the latter occurs, this leads to formation of chains with different
14 linkages that may involve Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 or Lys63 [3]. Chains can also
15 contain linear (via the N-terminal Met1 of Ub) or mixed linkages [4].
16
17
18
19
20
21
22
23

24 The seven lysine residues are present on different surfaces of the Ub monomer and have different
25 degrees of exposures to the solvent. Lys6, Lys33 and Lys63 are fully exposed and do not form
26 intramolecular contacts. Lys11 and Lys27 are the least exposed to the solvent, as they point towards
27 the hydrophobic core and are involved in salt-bridges with the carboxylate oxygen atoms of Glu34 and
28 Asp52, respectively. Finally, Lys29 and Lys48 form hydrogen bonds from their ϵ -amino nitrogen
29 atoms and the carbonyl oxygen atoms of Glu16 and Ala46, respectively, and are partially exposed to
30 the solvent. The amino acid position involved in the linkage thus dictates the structural properties of
31 the corresponding chain and differences in the biological function. All chain types are present in both
32 yeast and mammalian cells but the relative abundance of the individual chains depends on the specific
33 cell state [5-7].
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 The richness of conformations that the different linkages confer has been the subject of intense studies
49 over the past 1-2 decades, following advancements in producing ubiquitin chains *in vitro*. In this
50 review, we summarize our knowledge of the function(s) of polyUb and describe the conformational
51 landscape of polyUb chains both in isolation and in complexes with other proteins. Although our list
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 of examples will not be exhaustive given the continuous flow of publications, it is hoped that this
2 review will provide a useful guideline to discuss the different aspects of polyUb versatility.
3
4
5
6

7 **The functional role of polyUb**

8

9
10 The different polyUb chains are involved in different metabolic pathways, where they seem to have
11 an essential role in regulation. All linkages of polyUb homopolymers can target proteins for
12 proteasomal degradation. However, Lys48 polyUb chains are most clearly involved in targeting a
13 substrate for degradation [8-10]. PolyUbs with Lys48 and Lys63 linkages are also involved in cellular
14 signalling [11]. Lys11, Lys63 and Met1 linkages seem to be involved in the regulation of cell cycle
15 and nuclear factor- κ B activation [12-14]. Lys11 chains also mediate protein degradation during cell
16 division and bind to the S5 subunit of the proteasome [12]. The complexity of the polyUb code does
17 not end here: mixed chains with linkages from different lysines also seem to target proteins for
18 proteasomal degradation [15, 16]. Branched chains, in which more than one Ub domain is attached to
19 the same subunit, have also been described [4]. Since these chains bind to the proteasome but with low
20 affinity, they are thought to be involved in other pathways. Mixed, multiply branched chains with
21 Lys6, Lys27 and Lys48 linkages are attached to the monoubiquitinating ligase RING1b. This
22 modification is needed to stimulate the RING1b enzyme activity toward the substrate histone H2A
23 [17, 18].
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43 Ub function can in turn be modulated by post-translational modifications such as acetylation or
44 phosphorylation [19-21]. Phosphorylation of eight out of the eleven possible sites (Thr7, Thr12, Thr14,
45 Ser20, Ser57, Tyr59, Ser65, and Thr66) has been observed in mammals [20, 22-26]. Preferential
46 phosphorylation of Ser65 is favored in Lys6 and Lys11 polyUb chains, while it is disfavored in Lys27
47 chains [27]. Parkin E3 ligase is activated by pSer65-Ub initiating a complex chain of events [28, 29],
48 whereas the deubiquitylating efficiency of deubiquitinating (DUB) enzymes USP2, USP5, USP8,
49 USP15, USP30, ataxin-3, and AMSH is inhibited by phosphorylation at this position [21, 27, 30].
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 Interestingly, the site of phosphorylation (distal or proximal domain; see Box 1) is relevant: USP2 is
2 inhibited by phosphorylation of the distal Ser65, while AMSH is inhibited by phosphorylation of the
3 proximal Ser65 [30].
4
5
6

7 Acetylation also significantly affects the ubiquitylation cascade, apparently by repressing polyUb
8 elongation *in vitro* while stabilizing the monoubiquitylation state in cells [20]. The known acetylated
9 sites are Lys6 and Lys48 [20]. Ub acetylation can occur after Ub is conjugated to partners and affects
10 Ub interaction with proteins including DUBs, Ub binding domains (UBDs) and enzymes involved in
11 Ub chain assembly [20, 31, 32].
12
13
14
15
16
17
18
19
20
21

22 **The kaleidoscopic world of polyUb structure**

23
24 Ubiquitylation is catalyzed *in vivo* by a three-step enzymatic cascade. First, an E1 activating enzyme
25 uses ATP to link the C-terminal glycine of Ub via a thioester bond to a cysteine in the E1 active site
26 [33-35]. The activated Ub intermediate is then transferred to the catalytic cysteine of an E2 enzyme.
27
28 The E3 is finally responsible for further transfer of Ub to the substrate [36]. Target specificity and
29 independent regulation of different linkages occurs through the E2 and E3 enzymes [37]. E3 ligases
30 are divided into four families, HECT, RING-finger, U-box, and PHD-finger, based on the mechanism
31 of catalysis. HECT E3 ligases temporarily accept activated Ub, whereas RING E3 ligases catalyse
32 direct transfer of Ub from the E2 to the substrate [38].
33
34
35
36
37
38
39
40
41
42
43

44 A prerequisite to allow *in vitro* structural and functional studies of polyUb chains is to produce suitable
45 quantities of pure samples. This is often not an easy task given that some of the enzymes responsible
46 for linkage are unknown. Several strategies have thus been designed to circumvent the problem, some
47 of which rely on enzymatic approaches (**Box 2**). Other methods are based on chemical synthesis (**Box**
48 **3**).
49
50
51
52
53
54

55 The first Lys48 diUb structure appeared in 1992 [39]. Since then, several crystallographic and NMR
56 structures of polyUb chains have been solved: more than 100 structures are available in the PDB data
57
58
59
60
61
62

1 repository. Most of these studies relate to Lys48 and Lys63 linkage chains, whereas the first structure
2 of a Met1 (or linear) diUb was released only in 2009 [40] followed by the structures of Lys6 and Lys11
3 diUbs in 2010 [41, 42]. We had to wait until 2015 to see structures of Lys29 [43, 44] and Lys33 diUb
4 [44, 45]. Notably, structural information on Lys27 linkage chains is still missing. Two leitmotifs
5 common to these structures are open/closed conformations and flexibility (**Figure 1**). The linker
6 geometry determines whether the chain forms a more compact ‘closed’ assembly, in which the
7 domains pack against each other typically through various hydrophobic patches, or forms a more
8 extended open structure. Open structures are usually flexible and cover a large number of isoenergetic
9 states of the energy landscape. Next, we briefly discuss the most important features of these structures.
10 The order chosen matches the order of publication for each particular linkage.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 **Lys48 linked chains: open or closed?**

28
29 There are several structures of isolated Lys48 di and tetraUb [39, 46-54] (**Table 1**). The predominant
30 state of Lys48 diUb chains at near-physiological conditions is a ‘closed’ conformation in which the
31 Ile44-centred hydrophobic interfaces (**Box 1**) of each subunit pack against each other (pdb code: 1aar
32 and 3m3j), rendering this closed conformation incompetent for interaction with protein partners [39,
33
34 48]. However, this is thought of as only one of the several possible conformations of the Lys48 diUb
35 energy landscape. Solution studies demonstrated that open and closed conformations are in fast
36 exchange with each other [54, 55] and that this equilibrium is critical for recognition of Lys48-linked
37 polyUb chains (2pea, 2pe9, 3ns8 and 3aul). Lys48 tetraUbs form a pseudo-tetragonal structure in
38 which pairs of hydrophobic patches interact, suggesting that diUb is the minimal structural unit. As
39 for diUb chains, tetraUbs are thought to visit closed and open conformations in fast equilibrium and
40 use this dynamical conformation state for partner recognition (1tbe, 1f9j, 2o6v and 3alb) [46, 47, 50,
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Lys63 and Met1 linked chains: it could not be more extended

1
2 In all available structures (which are from X-ray diffraction), Lys63 chains have an entirely different
3
4 three-dimensional structure from Lys48-linked chains [40, 49, 56-58]. The conformation is extended
5
6 and open, with no well-defined non-covalent interface between moieties. The inter-domain interaction
7
8 between the Ile44-centred hydrophobic patches of the two moieties seen in Lys48 chains is not
9
10 sterically possible because the Lys63 side chain is on the opposite side of the subunit compared to the
11
12 Ile44-centred hydrophobic patch, while the linker is not long enough to allow inter-domain interactions
13
14 between the two patches. This leaves the hydrophobic patches exposed and competent for binding. In
15
16 the crystal structures (2jf5, 3h7p, 3hm3), the hydrophobic patches do not lie on the same surface of the
17
18 Ub molecule but are sequentially rotated by $\sim 90^\circ$ [40, 57, 58]. In solution it is more likely that the
19
20 linker is flexible, thus allowing the individual monomers to undergo free rotational movements that
21
22 would change the relative orientation of the patches. What has been named the ‘fluidity’ or dynamics
23
24 of these chains may be important for function [58].
25
26
27
28
29
30

31 The overall conformation of Met1 chains is virtually equivalent to Lys63 chains and consists of
32
33 structures in which individual Ub moieties are rotationally unrestrained and highly flexible (3axc and
34
35 2w9n). This is expected because Met1 is spatially close to Lys63 in the Ub structure (**Box 1**). The two
36
37 types of chains thus differ mainly by the chemistry of the linkage between domains [40].
38
39
40
41
42
43

Non-canonical chains: Lys6, Lys11, Lys29 and Lys33-linked chains

44
45 The only known diUb structure of the Lys6 linkage is a crystal structure (2xk5) [41]. This reveals an
46
47 asymmetric compact conformation in which the proximal Ub domain commits only two of its
48
49 canonical hydrophobic residues (Ile44 and Val70) to bind a slightly different but close hydrophobic
50
51 patch in the distal Ub domain containing Ile36, Leu71 and Leu73 (referred to as the Ile36 patch). This
52
53 asymmetric interaction leaves part of the canonical Ile44-centred patch of the distal domain exposed
54
55
56
57
58
59
60
61
62
63
64
65

1 and available to binding to other partners. It is, however, possible that in solution the dynamic nature
2 of polyUb chains allows other interfaces.
3

4 Three different structures of Lys11 chains have been published (3nob, 2xew, 2mbo), which were
5 determined under different conditions [42, 59, 60]. Notably, the orientation of the two Ub moieties is
6 different in all three structures and leads to a different degree exposure of the Ile44-centred
7 hydrophobic patches. This could however be partially the consequence of crystal packing in the crystal
8 structure.
9

10 The crystal structure of Lys29 diUb (4s22) adopts an extended conformation, distinct from that
11 observed in Lys63 and in Met1 linked chains. The inter-domain contacts are mainly polar and involve
12 Arg42 and Arg72 of the distal Ub domain and Glu16, Glu18 and Asp21 of the proximal domain [43,
13 44].
14

15 The crystal structures of Lys33 di- and triUb were solved recently (4xyz, 4y1h, 5af4 and 5af5) [44,
16 45]. Lys33-linked diUb adopts a compact structure where the two monomeric subunits interact
17 extensively through two hydrophobic surfaces and involving the interdigitation of patches formed by
18 Ile36, Leu71, Leu73, Leu8, Ile13 and Leu69. This arrangement leaves the Ile44-centred patch exposed
19 to the solvent. This compact conformation is similar to the closed conformation adopted by one of the
20 Lys11 diUb structures (3nob) [59] but distinct from that observed in Lys6 and Lys48 diUbs. DUB
21 enzymes can distinguish between these linkages [61]. The crystal structure of Lys33 triUb reveals an
22 open conformation where the monomeric subunits form no mutual interactions except for the
23 isopeptide linkage [44, 45].
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 **PolyUb chains in complex with UBDs**

52 To understand the mechanisms by which proteins recognise different Ub linkages, it is essential to
53 look at the structures of polyUb complexes with other partners. One of the first structures and the
54 simplest model for polyUb interactions was that of the Ub associated 2 (UBA2) domain of hHR23A
55
56
57
58
59
60
61
62
63
64
65

1 with Lys48 polyUb. UBA2 forms asymmetric interactions with the proximal and distal Ub domains
2 (1zo6) [62]. The proximal Ub has higher-affinity binding than the distal domain, as also seen in the
3
4 S5a/Lys48 diUb complex (2kde) [63, 64]. The interacting surface of Ub is centred on I44, implying a
5
6 closed-to-open conformation transition of the diUb upon binding to UBA2 that results sandwiching of
7
8 UBA2 between the two Ub subunits. The authors suggested that selectivity is given by the entropic
9
10 cost of opening the interface and immobilizing the UBA2 domain.
11
12

13 After this seminal work, several other papers have reported structures of polyUb chains with other
14
15 proteins. The currently available structures (**Tables 2,3**) can be divided into two families: the structures
16
17 of complexes with UBDs and those of complexes with DUBs. From them we can address questions
18
19 such as: Are there general rules that dictate Ub interactions with other proteins? Can we explain
20
21 through structural determinants chain-selectivity? It is perhaps still too soon to answer the question of
22
23 whether complexes of short polyUb chains follow the same rules as longer ones.
24
25
26

27 There seems to be two distinct strategies for interaction and linkage specificity. One, preferentially
28
29 adopted by UBDs, in which linkage specificity is not mediated by direct interactions with the linkage.
30
31 In the second, adopted by several DUBs, the isopeptide linkage participates to recognition. Another
32
33 common theme to Ub complexes is the importance of the distal Ub in recognition, which almost always
34
35 establishes direct contacts with the UBD (**Figure 2 and Table 2**).
36
37
38

39 Representative examples of UBD/polyUb complexes are the crystal structures of the RAP80 tandem
40
41 Ub interacting motifs (UIMs) in complex with Lys63 diUb (3a1q) [65], the UBAN (Ubiquitin Binding
42
43 of ABIN and NEMO) motif of NEMO in complex with Met1 diUb (2zvn) [66], the NMR-based
44
45 structural model of the second UBA of hHR23A in complex with Lys48 diUb (1zo6) [62] and a Fab
46
47 fragment Apu2.16 in complex with Lys63 diUb (3dvn) [67].
48
49
50

51 Recently, the structure of the nitrate zinc finger 1 (NZF1) domain of TRABID in complex with Lys29
52
53 diUb (4s1z) was published [43]. In this case, there is no direct interaction between the linkage and the
54
55 UBDs. Instead, the NZF1 domain of TRABID interacts both with the distal and proximal Ub domains
56
57
58
59
60
61

1 and recognises the Ile44-centred patch on the distal domain and residues around Lys29 of the proximal
2 domain. The overall conformation and relative orientation of the two Ub subunits change in the
3 complex with respect to the arrangement on the unbound Lys29 diUb. Structural remodelling enables
4 the NZF1 to bind simultaneously both the distal and the proximal domains. This conformation is
5 thought to confer binding selectivity.
6
7
8
9
10

11 The accumulated structural knowledge has revealed that recognition of polyUb is often achieved
12 through rather weak binding of different domains to more than one Ub position in the polyUb chain.
13 This allows the formation of tight complexes through a mechanism of avidity and/or cooperativity
14 even when interaction between individual sites and monoUb might be weak (**Box 4**). The complex of
15 the individual tandem UIMs of RAP80 with monoUb is, for instance, too weak (Kd value 500 μ M) to
16 form a stable complex (3a1q) [68] but the simultaneous interactions of multiple UIMs with the multiple
17 subunits of a Lys63 polyUb chain significantly strengthen the complex. A possible advantage of having
18 several weak interactions is the possibility of gaining linkage specificity when the arrangement of the
19 domains favors avid binding to one polyUb linkage over others.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 **The critical role of spacers between UBDs**

38
39 The length and the flexibility of the linkers between UBDs are crucial for an avidity mechanism that
40 could be sensitive to the orientation of the binding subunits [65, 68]. The binding sites must be
41 optimally placed for simultaneous binding. A similar mechanism is observed in the structures of the
42 complexes of hHR23A with Lys48 diUb, of NEMO with Met1 diUb [62, 66], and of antibodies against
43 Lys63-linked Ub chains in complex with Lys63 diUb (3dvn) [67]. RAP80 recognises the hydrophobic
44 patches centred on Ile44 of both Ub moieties by the tandem UIM regions: UIM1 interacts with the
45 proximal Ub domain whereas UIM2 interacts solely with the distal domain. The authors suggested that
46 selectivity could be ensured by the inter-UIM region which has the right geometry and length to adopt
47 a conformation suitable for this type of recognition. Indeed, the linker forms a twelve amino acid long
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

alpha helix, which functions as a ‘molecular spacer’ and dictates the distance and the relative orientation between the two UIMs.

In the UBAN motif of NEMO in complex with Met1 diUb (2zvn) [66], the UBAN forms a typical parallel coiled-coil homodimer held together by hydrophobic interactions. This homodimer binds two diUb chains, one at each side of the coiled-coil. The distal Ub domain interacts mainly by its hydrophobic patch centred on Ile44, while the proximal domain interacts through polar interactions with NEMO residues. Linkage specificity is granted by the fact that the simultaneous binding of both proximal and distal domains can occur only with Met1 chains.

Linkage specificity is also dictated by the spacing between the distal and proximal Ub domains in the complex of Met1 diUb and the HOIL-1L NZF domain of LUBAC (linear Ub chain assembly complex; the enzyme responsible for the formation of Met1 chains) (3b0a) [69]. HOIL-1L NZF contains two Ub binding regions arranged to simultaneously bind both the proximal and the distal domains while having no direct contact between the HOIL-1L NZF domain and the linkage.

The NZF domains of TAB2 and TAB3 in complex with Lys63 diUb (3a9j, 3a9k and 2wwz) [70, 71] are almost identical and present the same structural features. Both proteins use two different well-conserved hydrophobic regions of the NZF domain to interact with the hydrophobic patches of the distal and proximal Ub domains and have no direct interaction with the Lys63 isopeptide bond. Like RAP80, the two Ub-binding sites of the TAB proteins are arranged to optimise binding affinity. The distance between them enables the simultaneous binding of the distal and the proximal domains, enhancing the affinity.

Interaction of polyUb with DUB enzymes

Much fewer structures are available of polyUb complexes with DUBs. The available structures are representative of only three of the linkage types (Met1, Lys11 and Lys63) and only three of the five DUB subfamilies (**Table 3**). Additionally, the structure of OTUB1 (an OTU DUB) in complex with

1 the E2s UbcH5b/UBE2D2 and Ubc13 (4dhz) indirectly suggests the arrangement of a Lys48 diUb
2 bound to an OTU DUB protein [72]. In this structure, an unbound monoUb was observed in the crystal
3
4 in addition to the bound Ub. The two Ubs are roughly positioned to allow a Lys48 linkage. Likewise,
5
6 a model of the Josephin domain of ataxin-3 (Josephin subfamily) with Lys48 diUb was suggested [73]
7
8 but it remains unclear whether this has physiological meaning.
9

10
11 DUB recognition seems much more differentiated and strongly dependent on the specific DUB
12
13 subfamily, some of which do not have linkage specificity (reviewed in [38]), such as the Ub-specific
14
15 protease (USP) subfamily. The OTU subfamily can have different specificities: OTUB1 and Cezanne,
16
17 for instance, prefer Lys48 and Lys11 linkages, respectively, and OTULIN is specific for Met1
18
19 linkages. TRABID and CYLD are examples of dual specificity, the former cleaving Lys29 and Lys33
20
21 linkages, the latter being specific for Met1 and Lys63. Also DUBs of the OTU family exhibit dual
22
23 specificity [31].
24
25
26
27

28
29 Overall, two features seem to be common (**Figure 3**). In all complexes the distal Ub makes contacts
30
31 through the Ile44 patch or nearby residues. This is true also for the structure of USP21 in complex with
32
33 linear diUb-aldehyde (**Figure 3c**) where the proximal Ub is covalently attached to the catalytic
34
35 cysteine. There are several examples in which specificity is dictated by direct recognition of the
36
37 (iso)peptide linkage. In the complex between AMSH-LP DUB (JAMM family) and Lys63 diUb (2znv)
38
39 [74], for instance, Lys63 diUb keeps an extended conformation as in the unbound form. AMSH-LP is
40
41 a zinc dependent DUB enzyme that specifically cleaves Lys63 polyUb chains. The binding affinity
42
43 mainly relies on the interaction of AMSH-LP with two hydrophobic patches of the distal Ub domain
44
45 centred on Ile44 (Leu8, Ile44, Val70) and on Ile36 (Ile36, Leu71, Leu73). While the interaction
46
47 between AMSH-LP with distal Ub domain is critical for binding affinity, the interaction of AMSH-LP
48
49 with proximal Ub is critical for catalysis. To allow cleavage of the isopeptide bond, Lys63 of the
50
51 proximal Ub has to be kept in the right orientation for catalysis. Thus, the surface around the Lys63 of
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 the proximal Ub is located inside a concave surface made by two regions of AMSH-LP where the
2 tripeptide Gln62-Lys63-Glu64 is correctly aligned for cleavage.
3

4 OTULIN (OTU family), also known as FAM105B or Gumby, seems to be the only DUB specific for
5 Met1 linkages. The catalytic domain of OTULIN interacts with Met1 diUb (3znz and 4ksl) [75, 76]
6 through the Ile44 centred hydrophobic patch of the distal Ub domain, while the proximal domain binds
7 OTULIN by a different surface. The binding interface between OTULIN and the proximal Ub is quite
8 extensive and is very important for binding specificity as it orients the proximal Ub to place the Met1
9 linkage in the catalytic pocket and positions it for cleavage. OTULIN is the first DUB enzyme to show
10 a substrate-assisted catalytic mechanism: the proximal Ub participates directly in the organization of
11 the catalytic site with Met1 displacing the catalytic His by steric hindrance from its unbound position.
12 This movement results in activation of the otherwise inactive enzyme. Glu16 of the proximal Ub also
13 participates in restricting the movement of the catalytic His and in introducing a negative charge.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31 **Concluding remarks**

32 We have seen here how complex the repertoire of structures allowed by chains of a protein as small as
33 Ub can be. The possibility of so many different linkages and the interactions with other partners
34 amplifies the modularity of Ub behaviour, allowing the adoption of very different conformational
35 landscapes. Our understanding of the structural determinants of polyUb functions and interactions
36 continues to expand; we have tried to summarize the current knowledge. We have also discussed the
37 challenges and questions that remain to be addressed in the near future, before a more accurate
38 understanding is available that clarifies the role of polyUb in signalling and provides new insights into
39 its importance for all cellular functions. An important challenge for the future is to solve the structure
40 of Lys27-linked polyUb. We will also need to understand if there is a more universal code for polyUb
41 interactions and whether and how longer chains differ from diUbs. To answer to these questions, more
42 structures of Ub complexes are needed. When they become available, we may also be in the position
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 of understanding the significance that the flexibility of different linkages plays for partner recognition
2 and the role of each linkage. Finally, a fascinating open question is how post-translational
3
4
5 modifications modulate the Ub code.
6
7
8

9 **Acknowledgements**

10 We thank Pierandrea Temussi and David J. Thomas for critical reading and Dorian Lamba for help
11
12
13
14 with Figure 2 preparation. This review was possible thanks to the financial support of MRC.
15
16
17

18 **References**

- 19 1 Lothrop, A.P., *et al.* (2013) Deciphering post-translational modification codes. *FEBS Lett.* 587, 1247-1257
- 20 2 Goldstein, G., *et al.* (1975) Isolation of a polypeptide that has lymphocyte-differentiating properties and is
21 probably represented universally in living cells. *Proc Natl Acad Sci U S A* 72, 11-15
- 22 3 Ikeda, F. and Dikic, I. (2008) Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond
23 the Usual Suspects' review series. *EMBO Rep.* 9, 536-542
- 24 4 Kim, H.T., *et al.* (2007) Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases
25 (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. *J. Biol.*
26 *Chem.* 282, 17375-17386
- 27 5 Xu, P., *et al.* (2009) Quantitative proteomics reveals the function of unconventional ubiquitin chains in
28 proteasomal degradation. *Cell* 137, 133-145
- 29 6 Dammer, E.B., *et al.* (2011) Polyubiquitin linkage profiles in three models of proteolytic stress suggest the
30 etiology of Alzheimer disease. *J. Biol. Chem.* 286, 10457-10465
- 31 7 Peng, J., *et al.* (2003) A proteomics approach to understanding protein ubiquitination. *Nat. Biotechnol.* 21,
32 921-926
- 33 8 Hershko, A. and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* 67, 425-479
- 34 9 Thrower, J.S., *et al.* (2000) Recognition of the polyubiquitin proteolytic signal. *EMBO J.* 19, 94-102
- 35 10 Finley, D. (2009) Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev.*
36 *Biochem.* 78, 477-513
- 37 11 Chen, Z.J. and Sun, L.J. (2009) Nonproteolytic functions of ubiquitin in cell signaling. *Mol. Cell* 33, 275-286
- 38 12 Wickliffe, K.E., *et al.* (2011) K11-linked ubiquitin chains as novel regulators of cell division. *Trends Cell Biol.*
39 21, 656-663
- 40 13 Iwai, K. (2012) Diverse ubiquitin signaling in NF-kappaB activation. *Trends Cell Biol.* 22, 355-364
- 41 14 Iwai, K., *et al.* (2014) Linear ubiquitin chains: NF-kappaB signalling, cell death and beyond. *Nat. Rev. Mol.*
42 *Cell Biol.* 15, 503-508
- 43 15 Meyer, H.J. and Rape, M. (2014) Enhanced protein degradation by branched ubiquitin chains. *Cell* 157, 910-
44 921
- 45 16 Grice, G.L., *et al.* (2015) The Proteasome Distinguishes between Heterotypic and Homotypic Lysine-11-
46 Linked Polyubiquitin Chains. *Cell Rep* 12, 545-553
- 47 17 Nguyen, L.K., *et al.* (2011) Switches, excitable responses and oscillations in the Ring1B/Bmi1 ubiquitination
48 system. *PLoS Comput Biol* 7, e1002317
- 49 18 Ben-Saadon, R., *et al.* (2006) The polycomb protein Ring1B generates self atypical mixed ubiquitin chains
50 required for its in vitro histone H2A ligase activity. *Mol. Cell* 24, 701-711

19 Herhaus, L. and Dikic, I. (2015) Expanding the ubiquitin code through post-translational modification. *EMBO Rep*

20 Ohtake, F., *et al.* (2015) Ubiquitin acetylation inhibits polyubiquitin chain elongation. *EMBO Rep* 16, 192-201

21 Wauer, T., *et al.* (2015) Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis. *EMBO J.* 34, 307-325

22 Lee, H.J., *et al.* (2009) Quantitative analysis of phosphopeptides in search of the disease biomarker from the hepatocellular carcinoma specimen. *Proteomics* 9, 3395-3408

23 Shiromizu, T., *et al.* (2013) Identification of missing proteins in the neXtProt database and unregistered phosphopeptides in the PhosphoSitePlus database as part of the Chromosome-centric Human Proteome Project. *J Proteome Res* 12, 2414-2421

24 Bian, Y., *et al.* (2014) An enzyme assisted RP-RPLC approach for in-depth analysis of human liver phosphoproteome. *J Proteomics* 96, 253-262

25 Zhou, H., *et al.* (2013) Toward a comprehensive characterization of a human cancer cell phosphoproteome. *J Proteome Res* 12, 260-271

26 Sharma, K., *et al.* (2014) Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep* 8, 1583-1594

27 Swaney, D.L., *et al.* (2015) Phosphorylation of ubiquitin at Ser65 affects its polymerization, targets, and proteome-wide turnover. *EMBO Rep* 16, 1131-1144

28 Ordureau, A., *et al.* (2014) Quantitative proteomics reveal a feedforward mechanism for mitochondrial PARKIN translocation and ubiquitin chain synthesis. *Mol. Cell* 56, 360-375

29 Wauer, T., *et al.* (2015) Mechanism of phospho-ubiquitin-induced PARKIN activation. *Nature* 524, 370-374

30 Bondalapati, S., *et al.* (2015) Chemical synthesis of phosphorylated ubiquitin and diubiquitin exposes positional sensitivities of e1-e2 enzymes and deubiquitinases. *Chemistry* 21, 7360-7364

31 Mevissen, T.E., *et al.* (2013) OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell* 154, 169-184

32 Husnjak, K. and Dikic, I. (2012) Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu. Rev. Biochem.* 81, 291-322

33 Pickart, C.M. and Eddins, M.J. (2004) Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* 1695, 55-72

34 Schulman, B.A. and Harper, J.W. (2009) Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat. Rev. Mol. Cell Biol.* 10, 319-331

35 Ye, Y. and Rape, M. (2009) Building ubiquitin chains: E2 enzymes at work. *Nat. Rev. Mol. Cell Biol.* 10, 755-764

36 Berndsen, C.E. and Wolberger, C. (2014) New insights into ubiquitin E3 ligase mechanism. *Nat. Struct. Mol. Biol.* 21, 301-307

37 Wenzel, D.M., *et al.* (2010) E2s: Structurally Economical and Functionally Replete. *Biochem. J.* 433, 31-42

38 Komander, D. and Rape, M. (2012) The ubiquitin code. *Annu. Rev. Biochem.* 81, 203-229

39 Cook, W.J., *et al.* (1992) Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *J. Biol. Chem.* 267, 16467-16471

40 Komander, D., *et al.* (2009) Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep* 10, 466-473

41 Virdee, S., *et al.* (2010) Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat. Chem. Biol.* 6, 750-757

42 Bremm, A., *et al.* (2010) Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nat. Struct. Mol. Biol.* 17, 939-947

43 Kristariyanto, Y.A., *et al.* (2015) K29-selective ubiquitin binding domain reveals structural basis of specificity and heterotypic nature of k29 polyubiquitin. *Mol. Cell* 58, 83-94

44 Michel, M.A., *et al.* (2015) Assembly and specific recognition of k29- and k33-linked polyubiquitin. *Mol. Cell* 58, 95-109

1 45 Kristariyanto, Y.A., *et al.* (2015) Assembly and structure of Lys33-linked polyubiquitin reveals distinct
2 conformations. *Biochem. J.* 467, 345-352
3 46 Cook, W.J., *et al.* (1994) Structure of Tetraubiquitin Shows How Multiubiquitin Chains Can Be Formed. *J.*
4 *Mol. Biol.* 236, 601-609
5 47 Phillips, C.L., *et al.* (2001) Structure of a new crystal form of tetraubiquitin. *Acta Crystallogr D Biol Crystallogr*
6 57, 341-344
7 48 Varadan, R., *et al.* (2002) Structural Properties of Polyubiquitin Chains in Solution. *J. Mol. Biol.* 324, 637-647
8 49 Tenno, T., *et al.* (2004) Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains.
9 *Genes Cells* 9, 865-875
10 50 Eddins, M.J., *et al.* (2007) Crystal Structure and Solution NMR Studies of Lys48-linked Tetraubiquitin at
11 Neutral pH. *J. Mol. Biol.* 367, 204-211
12 51 Trempe, J.F., *et al.* (2010) A new crystal form of Lys48-linked diubiquitin. *Acta Crystallogr Sect F Struct Biol*
13 *Cryst Commun* 66, 994-998
14 52 Satoh, T., *et al.* (2010) Crystal structure of cyclic Lys48-linked tetraubiquitin. *Biochem. Biophys. Res.*
15 *Commun.* 400, 329-333
16 53 Hirano, T., *et al.* (2011) Conformational dynamics of wild-type Lys-48-linked diubiquitin in solution. *J. Biol.*
17 *Chem.* 286, 37496-37502
18 54 Lai, M.Y., *et al.* (2012) Structural and biochemical studies of the open state of Lys48-linked diubiquitin.
19 *Biochim. Biophys. Acta* 1823, 2046-2056
20 55 Ryabov, Y. and Fushman, D. (2006) Interdomain mobility in di-ubiquitin revealed by NMR. *Proteins* 63, 787-
21 796
22 56 Varadan, R., *et al.* (2004) Solution conformation of Lys63-linked di-ubiquitin chain provides clues to
23 functional diversity of polyubiquitin signaling. *J. Biol. Chem.* 279, 7055-7063
24 57 Weeks, S.D., *et al.* (2009) Crystal structures of Lys-63-linked tri- and di-ubiquitin reveal a highly extended
25 chain architecture. *Proteins* 77, 753-759
26 58 Datta, A.B., *et al.* (2009) The Structure and Conformation of Lys63-Linked Tetraubiquitin. *J. Mol. Biol.* 392,
27 1117-1124
28 59 Matsumoto, M.L., *et al.* (2010) K11-Linked Polyubiquitination in Cell Cycle Control Revealed by a K11
29 Linkage-Specific Antibody. *Mol. Cell* 39, 477-484
30 60 Castañeda, Carlos A., *et al.* (2013) Unique Structural, Dynamical, and Functional Properties of K11-Linked
31 Polyubiquitin Chains. *Structure* 21, 1168-1181
32 61 Ritorto, M.S. *et al.* (2014) Screening of DUB activity and specificity by MALDI-TOF mass spectrometry.
33 *Nature Comm.* 5, 4763.
34 62 Varadan, R., *et al.* (2005) Structural determinants for selective recognition of a Lys48-linked polyubiquitin
35 chain by a UBA domain. *Mol. Cell* 18, 687-698
36 63 Wang, Q., *et al.* (2005) Structure of S5a bound to monoubiquitin provides a model for polyubiquitin
37 recognition. *J. Mol. Biol.* 348, 727-739
38 64 Zhang, N., *et al.* (2009) Structure of the s5a:k48-linked diubiquitin complex and its interactions with rpn13.
39 *Mol. Cell* 35, 280-290
40 65 Sato, Y., *et al.* (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem
41 UIMs of RAP80. *EMBO J.* 28, 2461-2468
42 66 Rahighi, S., *et al.* (2009) Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB
43 activation. *Cell* 136, 1098-1109
44 67 Newton, K., *et al.* (2008) Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. *Cell*
45 134, 668-678
46 68 Sims, J.J. and Cohen, R.E. (2009) Linkage-specific avidity defines the lysine 63-linked polyubiquitin-binding
47 preference of rap80. *Mol. Cell* 33, 775-783
48 69 Sato, Y., *et al.* (2011) Specific recognition of linear ubiquitin chains by the Npl4 zinc finger (NZF) domain of
49 the HOIL-1L subunit of the linear ubiquitin chain assembly complex. *Proc Natl Acad Sci U S A* 108, 20520-20525
50 70 Sato, Y., *et al.* (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by NZF
51 domains of TAB2 and TAB3. *EMBO J.* 28, 3903-3909
52
53
54
55
56
57
58
59
60
61
62
63
64
65

71 Kulathu, Y., *et al.* (2009) Two-sided ubiquitin binding explains specificity of the TAB2 NZF domain. *Nat. Struct. Mol. Biol.* 16, 1328-1330

72 Wiener, R., *et al.* (2012) The mechanism of OTUB1-mediated inhibition of ubiquitination. *Nature* 483, 618-622

73 Nicastro, G., *et al.* (2010) Understanding the role of the Josephin domain in the PolyUb binding and cleavage properties of ataxin-3. *PLoS One* 5, e12430

74 Sato, Y., *et al.* (2008) Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. *Nature* 455, 358-362

75 Keusekotten, K., *et al.* (2013) OTULIN Antagonizes LUBAC Signaling by Specifically Hydrolyzing Met1-Linked Polyubiquitin. *Cell* 153, 1312-1326

76 Rivkin, E., *et al.* (2013) The linear ubiquitin-specific deubiquitinase gumbly regulates angiogenesis. *Nature* 498, 318-324

77 Archibald, J.M., *et al.* (2003) Novel ubiquitin fusion proteins: ribosomal protein P1 and actin. *J. Mol. Biol.* 328, 771-778

78 Kimura, Y. and Tanaka, K. (2010) Regulatory mechanisms involved in the control of ubiquitin homeostasis. *J Biochem* 147, 793-798

79 Vijay-Kumar, S., *et al.* (1987) Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* 194, 531-544

80 Schlesinger, D.H., *et al.* (1975) The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. *Biochemistry* 14, 2214-2218

81 Hurley, J.H., *et al.* (2006) Ubiquitin-binding domains. *Biochem. J.* 399, 361-372

82 Komander, D. (2009) The emerging complexity of protein ubiquitination. *Biochem. Soc. Trans.* 37, 937-953

83 Faggiano, S.A., C; Pastore, A (2015) Methods for the enzymatic synthesis and purification of ubiquitin chains. *Anal. Biochem.* (in press)

84 Hospenthal, M.K., *et al.* (2013) Assembly, analysis and architecture of atypical ubiquitin chains. *Nat. Struct. Mol. Biol.* 20, 555-565

85 Dong, Ken C., *et al.* (2011) Preparation of Distinct Ubiquitin Chain Reagents of High Purity and Yield. *Structure* 19, 1053-1063

86 Komander, D., *et al.* (2008) The Structure of the CYLD USP Domain Explains Its Specificity for Lys63-Linked Polyubiquitin and Reveals a B Box Module. *Mol. Cell* 29, 451-464

87 Bremm, A. and Komander, D. (2012) Synthesis and analysis of K11-linked ubiquitin chains. *Methods Mol. Biol.* 832, 219-228

88 Reyes-Turcu, F.E., *et al.* (2008) Recognition of Polyubiquitin Isoforms by the Multiple Ubiquitin Binding Modules of Isopeptidase T. *J. Biol. Chem.* 283, 19581-19592

89 Spasser, L. and Brik, A. (2012) Chemistry and biology of the ubiquitin signal. *Angew. Chem., Int. Ed. Engl.* 51, 6840-6862

90 Strieter, E.R. and Korasick, D.A. (2012) Unraveling the complexity of ubiquitin signaling. *ACS Chem. Biol.* 7, 52-63

91 Hemantha, H.P. and Brik, A. (2013) Non-enzymatic synthesis of ubiquitin chains: Where chemistry makes a difference. *Bioorg. Med. Chem.* 21, 3411-3420

92 Yang, R., *et al.* (2010) Synthesis of K48-linked diubiquitin using dual native chemical ligation at lysine. *Chem Commun (Camb)* 46, 7199-7201

93 Kumar, K.S., *et al.* (2011) Total chemical synthesis of a 304 amino acid K48-linked tetraubiquitin protein. *Angew. Chem., Int. Ed. Engl.* 50, 6137-6141

94 Oualid, F.E., *et al.* (2012) Synthesis of atypical diubiquitin chains. *Methods Mol Biol* 832, 597-609

95 Castañeda, Carlos A., *et al.* Nonenzymatic Assembly of Natural Polyubiquitin Chains of Any Linkage Composition and Isotopic Labeling Scheme. *J. Am. Chem. Soc.* 133, 17855-17868

96 Eger, S., *et al.* (2010) Synthesis of defined ubiquitin dimers. *J. Am. Chem. Soc.* 132, 16337-16339

97 Valkevich, E.M., *et al.* (2012) Forging isopeptide bonds using thiol-ene chemistry: site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases. *J. Am. Chem. Soc.* 134, 6916-6919

98 Rohaim, A., *et al.* (2012) Structure of a compact conformation of linear diubiquitin. *Acta Crystallogr D Biol Crystallogr* 68, 102-108

1 99 Bolton, D., *et al.* (2001) Structure and properties of a dimeric N-terminal fragment of human ubiquitin. *J. Mol. Biol.* 314, 773-787

2 100 van Dijk, A.D., *et al.* (2005) Various strategies of using residual dipolar couplings in NMR-driven protein

3 docking: application to Lys48-linked di-ubiquitin and validation against 15N-relaxation data. *Proteins* 60, 367-

4 381

5 101 Ryabov, Y. and Fushman, D. (2007) Structural assembly of multidomain proteins and protein complexes

6 guided by the overall rotational diffusion tensor. *J. Am. Chem. Soc.* 129, 7894-7902

7 102 Matsumoto, M.L., *et al.* (2012) Engineering and structural characterization of a linear polyubiquitin-

8 specific antibody. *J. Mol. Biol.* 418, 134-144

9 103 Tokunaga, F., *et al.* (2012) Specific recognition of linear polyubiquitin by A20 zinc finger 7 is involved in

10 NF-kappaB regulation. *EMBO J.* 31, 3856-3870

11 104 Liu, S., *et al.* (2012) Promiscuous interactions of gp78 E3 ligase CUE domain with polyubiquitin chains.

12 *Structure* 20, 2138-2150

13 105 Toma, A., *et al.* (2015) Structural basis for ubiquitin recognition by ubiquitin-binding zinc finger of FAAP20.

14 *PLoS One* 10, e0120887

15 106 Sekiyama, N., *et al.* (2012) NMR analysis of Lys63-linked polyubiquitin recognition by the tandem

16 ubiquitin-interacting motifs of Rap80. *J. Biomol. NMR* 52, 339-350

17 107 Yoshikawa, A., *et al.* (2009) Crystal structure of the NEMO ubiquitin-binding domain in complex with Lys

18 63-linked di-ubiquitin. *FEBS Lett.* 583, 3317-3322

19 108 Peisley, A., *et al.* (2014) Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I. *Nature*

20 509, 110-114

21 109 Sato, Y., *et al.* (2015) Structures of CYLD USP with Met1- or Lys63-linked diubiquitin reveal mechanisms

22 for dual specificity. *Nat. Struct. Mol. Biol.* 22, 222-229

23 110 Ye, Y., *et al.* (2011) Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21.

24 *EMBO Rep* 12, 350-357

25 111 Shrestha, R.K., *et al.* (2014) Insights into the mechanism of deubiquitination by JAMM deubiquitinases

26 from cocrystal structures of the enzyme with the substrate and product. *Biochemistry* 53, 3199-3217

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

Box 1: The Ub identity card

1
2 Ub is a small protein (76 amino acid residues) with a globular, compact and tightly hydrogen-bonded fold.
3
4 Mammalian cells have four Ub genes, which encode polyUb or fusions with ribosomal proteins. The polyUb
5
6 precursor is a head-to-tail linear polymer with the carboxyl group of the C-terminal Gly76 of one subunit being
7
8 covalently bound to the α -amino group of the N-terminal Met1 of the next molecule. Cleavage of these linear
9
10 chains and of Ub fusion proteins produces the mature monomeric protein in which the N terminus, no longer
11
12 involved in the peptide bond, changes orientation and packs against the structural core forming a compact
13
14 globular structure [77, 78]. The fold consists of a mixed β -sheet with five strands, a short 310-helix and 3.5
15
16 turns of an α -helix, which packs against the cavity of the β -sheet (Figure I) [79]. The first seven residues in the
17
18 N-terminal region of the protein are tightly constrained. This renders the N terminus not easily accessible,
19
20 preventing degradation. Ub is highly stable over a wide range of pH and temperature values and against
21
22 protease digestion despite the presence of seven lysine residues and four arginine residues, which could easily
23
24 be attacked by trypsin [80]. The only part of the protein protruding out from the structure is the C-terminal
25
26 Leu-Arg-Gly-Gly motif, which does not interact with the rest of the molecule and is thus accessible to enzymes
27
28 involved in the formation of both isopeptide and peptide bonds. Other small proteins with folds similar to Ub
29
30 have also been identified, among which are NEDD8, SUMO-1, SUMO-2, SUMO-3, ATG8, ISG15, and FAT10.
31
32 These proteins are collectively called Ubl. A key feature of Ub is the solvent-exposed hydrophobic-patch
33
34 centred on Leu8, Ile44, and Val70 (Figure I, in red). This is the preferential target of most UBDs and DUBs
35
36 despite their diversity in function and structure [81, 82]. Alternative hydrophobic surfaces are also possible,
37
38 such as the patch centred on Ile36 (Ile36, Leu71, Leu73, in orange in Figure I). In polyUbs, the C terminus of
39
40 one subunit (conventionally named the distal domain) can covalently attach to one of each of the seven lysine
41
42 residues (Figure I, in blue) or to the N terminus Met1 (Figure I, in yellow) of a second subunit (conventionally
43
44 named the proximal domain).
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

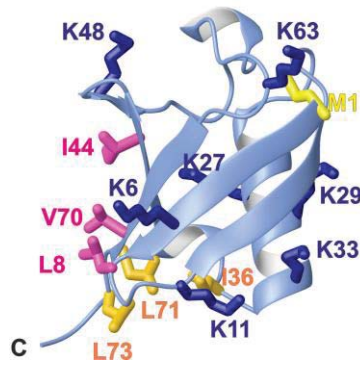


Figure I. Ribbon representation of monoUb (1ubq).

Box 2: Producing polyUbs in the laboratory in large quantities

Enzymatic production can be achieved in vitro either using wild-type or a mutated Ub as a substrate (for a recent extensive review see [83]) (Figure I). Mutants are more modular as chain elongation can be controlled step by step. This allows the possibility of differentiating between the different positions of the chain and obtaining chains in which each position can be differentially isotope labelled as required, such as in NMR studies [48, 60, 84]. However, producing mutants is more laborious than using the commercial wild-type protein. More recently, protocols for the production of Lys48, Lys63, and Lys11 linked chains using wild-type Ub have been developed [42, 85-87]. The method is based on the formation of unanchored Ub chains by E1 and a specific E2/E3 enzyme and separation of the products by cation exchange chromatography. It requires the use of an E2/E3 enzyme specific for a target chain type. A limitation is, however, that no selective isotope labelling of a particular position is possible, as the reagents are indistinguishable. Also, chains longer than tetraUb can be separated with difficulty by cation exchange chromatography. Other protocols allow the production of Lys29 and Lys33 chains exploiting DUB enzymes to cleave unwanted chain types formed as secondary products [43-45]. Finally, production of milligrams of Met1 chains can be achieved using suitable plasmids for expression of the required chain length. Several of these plasmids for the expression of di, tri and tetraUb are already available [85, 88].

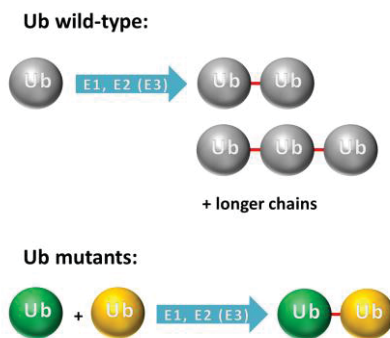
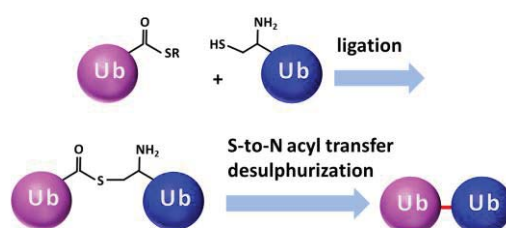


Figure I. The two approaches for the enzymatic synthesis of Ub chains.

Box 3: Chemical synthesis of polyUb chains

Chemical methods for the production of Ub chains have been developed to circumvent the need of identification of the specific E2/E3 system for the ligation of a selected chain type (for previous reviews see [89-91]). Another advantage is their suitability to produce Ub chains with selective isotope labelling in a desired position of the chain. The most exploited approach for the chemical synthesis of an isopeptide bond between two Ub building blocks is expressed protein ligation (EPL) (Figure I). This technique involves native chemical ligation (NCL) between a Ub subunit carrying a C-terminal thioester (α -thioester) and a second Ub with a free thiol adjacent to an amine group (mercaptolysine) [92-94]. α -thioesters are generated from thiolysis of an intein fusion protein, while the mercaptolysine is obtained through solid-phase peptide synthesis (SPPS). EPL consists in a transthioesterification between the thiol and the α -thioester, followed by a spontaneous intramolecular S-to-N acyl transfer and a desulphurization step, resulting in a native isopeptide bond. An evolution of NCL is genetically encoded orthogonal protection and activated ligation (GOPAL) [41]. As in NCL, a Ub α -thioester is obtained from intein fusion thiolysis. All lysines are blocked with a protecting group. The second ligand has a genetically encoded modified lysine, protected with a different protecting group, which is removed to expose the selected amine group to ligation. Finally, the other protecting groups are also removed. GOPAL has allowed synthesis of Lys6 and Lys29 diubiquitin [41].

1 An extended GOPAL strategy was applied to the preparation of longer chains [95]. Di and triUb α -thioesters
2 were prepared using intein chemistry, and a different protecting group for lysines was used that could be
3 removed in milder conditions. Other methods were developed for synthesis of Ub chains formed via non-
4 native linkages. For example, Ubs were linked via a triazole bond exploiting a copper-catalyzed azide-alkyne
5 “click” reaction [96], or using thiolene coupling chemistry [97]. Recent advances include the chemical synthesis
6 of phosphorylated diUb [30]. The chemical synthesis of chains remains the preferred method for the
7 production of Lys27 linked chains, since the E2/E3 necessary for enzymatic ligation are not known.
8
9
10
11
12
13
14
15
16



17
18
19
20
21
22
23
24
25
26
27 Figure 1. Expressed protein ligation (EPL)

31 **Box 4: Avidity versus affinity**

32
33 Affinity measures the strength of binding between two molecules, each of which contain a single binding site.
34
35 Avidity is instead a term originally derived from immunology to refer to the properties of antibody-
36 antigen binding. It was meant as a measure of the overall strength of binding of an antigen with many antigenic
37 determinants and multivalent antibodies. It is dependent on three major parameters: affinity of the antibody
38 for the epitope, valence of both the antibody and antigen and structural rearrangements of the parts that
39 interact. As a result we measure more than the sum of the individual affinities. For instance, IgM antibodies
40 may have low affinity but they have high avidity due to their 10 weak binding sites contrary to the two strong
41 binding sites of IgG antibodies. In the cited example, the authors tested whether Rap80, which contains
42 tandem UIMs, could achieve linkage selectivity by exploiting the distinct orientation and spacing of Ub subunits
43 of particular polyUb linkages. They hypothesised that multiple UBDs could be arrayed in space to optimize
44 simultaneous interactions with both Ub subunits and adopt a configuration which could selectively bind only
45 one type of Ub-Ub linkage. Binding by the first of multiple UBDs to one Ub molecule would then position the
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

second UBD more favorably for the interaction with a second Ub in the chain. This mechanism can be seen as a way to increase the local concentration of the interactor by restricting its conformational space thus making the binding even more favorable.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

TABLE 1: The structures of unbound polyUb chains.

Chain type	PDB	Comments	Ref.
Met1 diUb	2w9n	X-ray, 2.25 Å	[40]
Met1 diUb	3axc	X-ray, 2.19 Å	[98]
Lys6 diUb	2xk5	X-ray, 3.00 Å	[41]
Lys6 triUb	3zlz	X-ray, 2.90 Å	[84]
Lys11 diUb	2mbo/2mbq	NMR, inconsistent with the crystal structures	[60]
Lys11 diUb	2xew	X-ray, 2.20 Å	[42]
Lys11 diUb	3nob	X-ray, 2.19 Å	[59]
Lys29 diUb	4s22	X-ray, 2.30 Å	[43]
Lys33 diUb	4xyz	X-ray, 1.65 Å	[45]
Lys33 diUb	5af4	X-ray, 1.85 Å	[44]
Lys33 triUb	4y1h	X-ray, 1.40 Å	[45]
Lys33 triUb	5af5	X-ray, 1.68 Å	[44]
Lys48 diUb	1aar	X-ray, 2.3 Å	[39]
Lys48 diUb	1gjz	NMR, truncated construct 1-51	[99]
Lys48 diUb	3aul	X-ray, 2.39 Å	[53]
Lys48 diUb	3m3j	X-ray, 1.60 Å	[51]
Lys48 diUb	3ns8	X-ray, 1.71 Å	[54]
Lys48 diUb	2bgf	NMR, used as a model to implement methods	[100]
Lys48 diUb	2pe9/2pea	NMR, used as a model to implement methods	[101]
Lys63 diUb	2jf5	X-ray, 1.95 Å	[40]
Lys63 diUb	3h7p/3h7s	X-ray bound to Cd or Zn, 1.90 Å and 2.30 Å	[57]
Lys48 tetraUb	1f9j	X-ray, 2.70 Å	[47]
Lys48 tetraUb	1tbe	X-ray, 2.40 Å	[46]
Lys48 tetraUb	2o6v	X-ray, 2.2 Å validated by NMR data	[50]
Lys48 tetraUb	3alb	X-ray, 1.85 Å	[52]
Lys63 tetraUb	3hm3	X-ray, 1.96 Å	[58]

TABLE 2: Structures of polyUb chains in complexes with UBDs.

Chain type	Partner	PDB ID*	UBD	Ref.
Met1 diUb	NEMO	2zvn/2zvo	UBAN	[66]
Met1 diUb	HOIL-1L	3b08/3b0a	NZF	[69]
Met1 diUb	Met1 linkage specific antibody	3u30	Fab	[102]
Met1 tetraUb	A20	3vuy	A20 ZnF	[103]
Lys29 diUb	TRABID	4s1z	NZF	[43]
Lys48 diUb	gp78CUE	2lvp/2lvq	CUE	[104]
Lys48 diUb	S5a	2kde/2kdf	UIM	[64]
Lys48 diUb	hHR23A	1zo6	UBA	[62]
Lys63 diUb	FAAP20	3wwq	NZF	[105]
Lys63 diUb	RAP80	2rr9	UIM	[106]
Lys63 diUb	RAP80	3a1q	UIM	[65]
Lys63 diUb	NEMO	3jsv	UBAN	[107]
Lys63 diUb	RIG-I	4nqk	2CARD	[108]
Lys63 diUb	Lys63 linkage specific antibody	3dvg/3dvn	Fab	[67]
Lys63 diUb	TAB2 and TAB3	3a9j/3a9k	NZF	[70]
Lys63 diUb	TAB2	2wx0/2wwz	NZF	[71]
Lys63 triUb	TAB2	2wx1	NZF	[71]

^aThey are all X-ray structures with the exception of 2kde/2kdf, 2rr9 and 1zo6..

1
2 **Table 3 – Crystal structure of complexes of polyUb chains with DUBs.**
3

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Chain type	Partner	PDB	DUB family	Ref.
Met1 diUb	CYLD	3wxe/3wxf	USP	[109]
Met1 diUb	OTULIN/Gumby/Fam105	4ksl	OTU	[76]
Met1 diUb	OTULIN/Gumby/Fam105	3znz	OTU	[75]
Met1 diUb aldehyde	USP21	2y5b	USP	[110]
Lys11 diUb	OTUD2	4boz	OTU	[31]
Lys63 diUb	CYLD	3wxg	USP	[109]
Lys63 diUb	Sst2	4nql	JAMM	[111]
Lys63 diUb	AMSH-LP	2znv	JAMM	[74]

Figure Legends

Figure 1. Representative structures of diUbs with different linkers. Top: The structures of distal Ub were superposed on monoUb (1ubq) and then displaced. The side chain of the proximal linker is explicitly indicated in magenta. As a reference, the Ile44 patch (Leu8, Ile44 and Val70) is indicated by warm red spheres corresponding to the C α atom. When relevant, alternative hydrophobic patches as indicated in the text are added as orange spheres. Bottom: A schematic representation of the structures using the same colour code used above with indications of the position of the Ile44 patch (in warm red). The relative PDB codes are indicated. The structures of longer chains are not shown for sake of simplicity.

Figure 2. Structures of representative complexes of diUbs with UBD partners. Distal Ub domains (always shown on the top) are indicated with lighter blue ribbons, while the proximal Ubs are shown in dark violet. The Ile44 and Ile36 centred patches are indicated as red and orange spheres respectively. We also added His68 to the Ile44 patch as this residue is often also involved in interaction. The various partners are indicated as black ribbons. The structures correspond to PDB entries: a) 3b08, b) 4s1z, c) 3a9j, d) 3wwq, e) 2kde, f) 2rr9, g) 2zvn, h) 3jsv, i) 1zo6, j) 2lvp, k) 2lvq, l) 4nqk, m) 3u30 and n) 3dvg. The linkage and the partner name are reported. When more than one structure was reported in the same article for the same linkage only one is reported. The only exception is CUE because the two structures significantly differ for the packing of the diUb against the partner. Please note that we have reported the structures available in the PDB Databank; this does not necessary correspond to specificity of binding for a certain chain type. For example, FAAP20 also binds Met1- and Lys48-linked chains. NEMO is also not specific for Lys63-linked chains.

Figure 3. Structures of representative complexes of diUbs with DUB partners. The distal and proximal Ub domains and the hydrophobic patches are indicated as in Figure 2. The structures correspond to PDB entries: a) 4ksl, b) 4boz, c) 2y5b, d) 3wxe, e) 3wxg, f) 4nql and g) 2znv. The Ub linkage and the

partner name are indicated. Note that in c) the proximal Ub of Met1 diUb is covalently linked to the catalytic cysteine of USP21 through its C-terminal Gly.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Outstanding Questions Box

- What is the structure of Ly27 polyUbs? This information is currently missing because we do not know the enzymes responsible for these linkages, which makes their production difficult. Answering this question would complete the 'periodic table' of polyUb structure. To do this, we must design a strategy for the chemical or enzymatic formation of these chains or identify the specific enzymes.
- Also missing are the structures of representative members of different DUB families and many more structures of longer chains. Also in the latter case a bottleneck remains the production of suitable quantities of these polyUb chains. This remains a challenging task.
- Is there a general code of polyUb interaction? PolyUbs are very promiscuous and interact with several very different cellular partners. Is there any general structural determinant that would allow us to predict the structure of polyUbs in complexes?
- How does the intrinsic flexibility of most linkages relate to their functional properties? What is the gain from these dynamical properties?
- **What is the exact role of each polyUb linkage in the proteasomal pathway? It was originally thought that only Lys48 linkages are involved.** More recently this view has changed with the realization that chains with all linkages interact with proteasomal components. Is there any specificity or selectivity?
- How is the Ub/polyUb code further modulated by post-translational modifications such as acetylation and phosphorylation? This is an emerging theme that adds a further layer of complexity but, at the same time, increases the repertoire of possible conformational states.

Figure 1

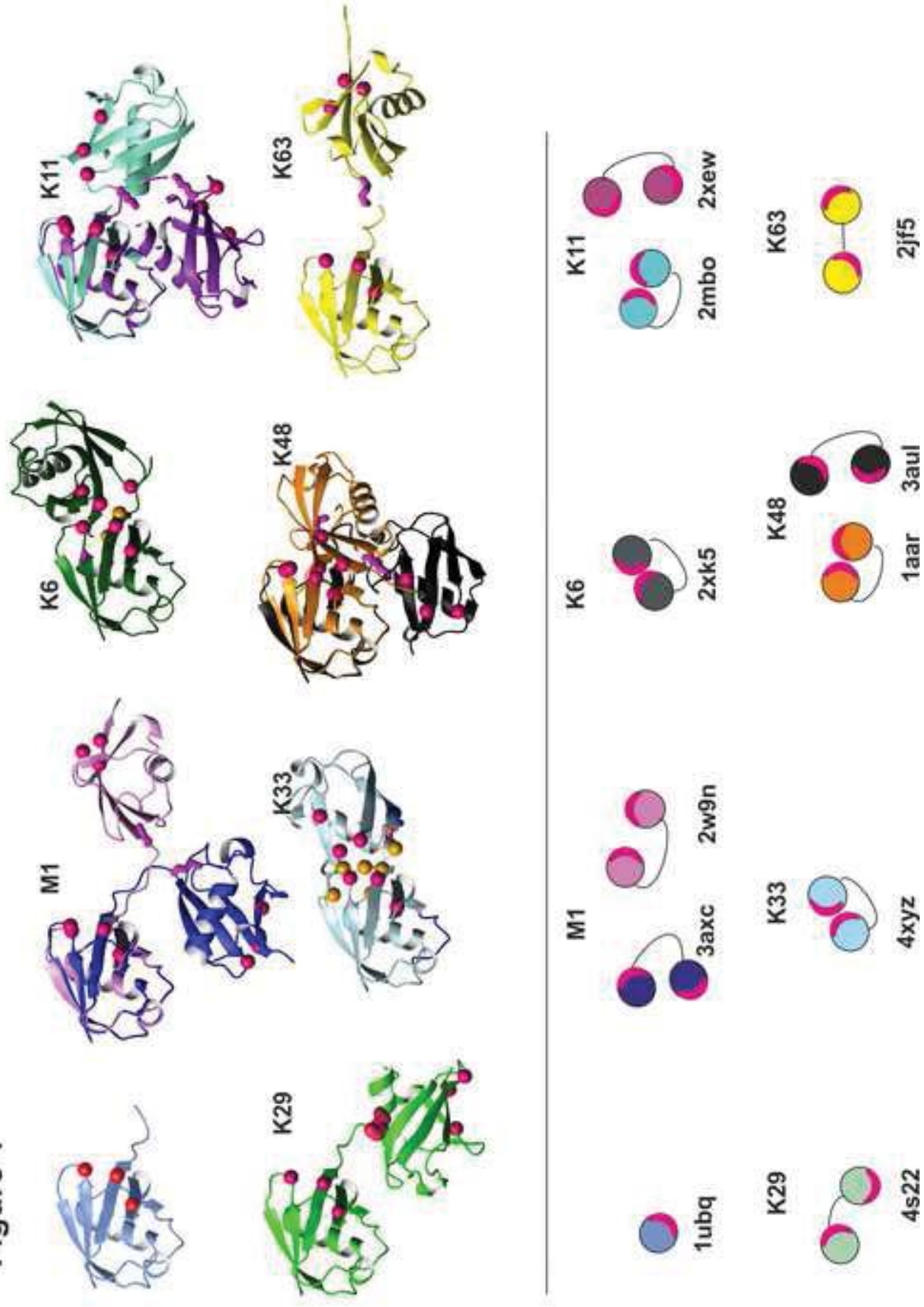


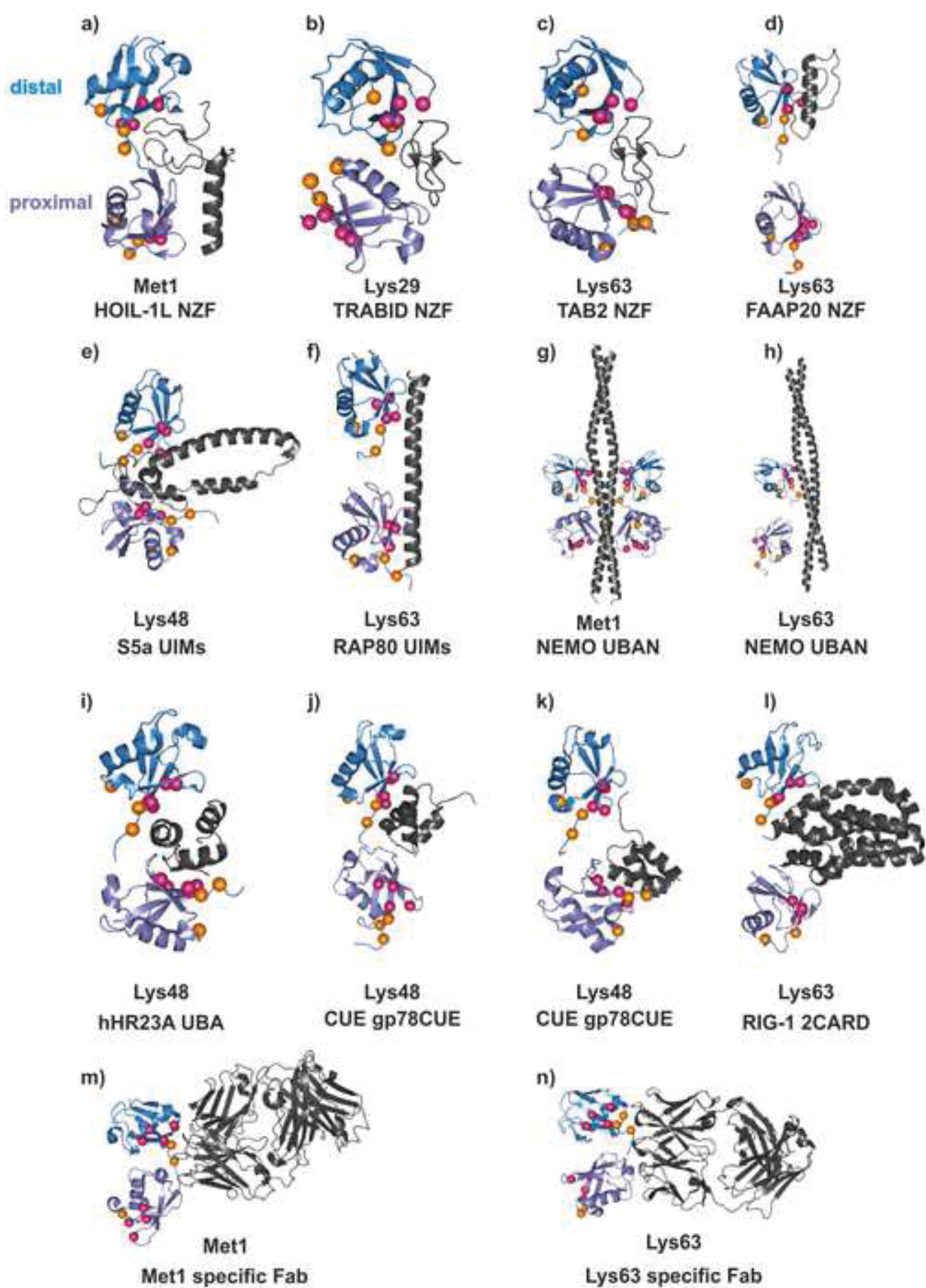
Figure 2

Figure 3

