[Frontiers in Bioscience 7, a148-162, August 1, 2002]

STRUCTURAL ATTRIBUTES IN THE CONJUGATION OF UBIQUITIN, SUMO AND RUB TO PROTEIN SUBSTRATES

Sandra Goettsch, and Peter Bayer

Max-Planck Institute for Molecular Physiology, Research Group Molecular and Structural Biophysics, Otto-Hahn-Str.11, D-44227 Dortmund, Germany

TABLE OF CONTENTS

1. ABSTRACT

Many cellular and secreted proteins are chemically modified after their translation is completed. The covalent linkage of a polypeptide chain (modifier) to a substrate protein is a special case of post-translational modification. In the late seventies it was observed that ubiquitin, a small modifier, marks short-lived proteins for degradation by the 26S proteasome. Over the last decade many other ubiquitin-related proteins were discovered and isolated. Attachment of polypeptide chains onto acceptor molecules became a common feature to regulate spatially and timely organized cellular pathways of proteins. This article focuses on the structures of the three modifiers: ubiquitin, RUB and SUMO and the cognate enzymes involved in these modification pathways. We have described the homologies and differences of these proteins and indicate salient topological hallmarks common to modifier-conjugating enzymes. This characterization will help in understanding these regulatory pathways and their similarities and differences in controlling protein fate, from protein degradation signals generated by polyubiquitination to functional modification brought about by RUB and SUMO conjugation.

2. INTRODUCTION: POST-TRANSLATIONAL MODIFICATION

Many cellular and secreted proteins are chemically modified after completion of their translational

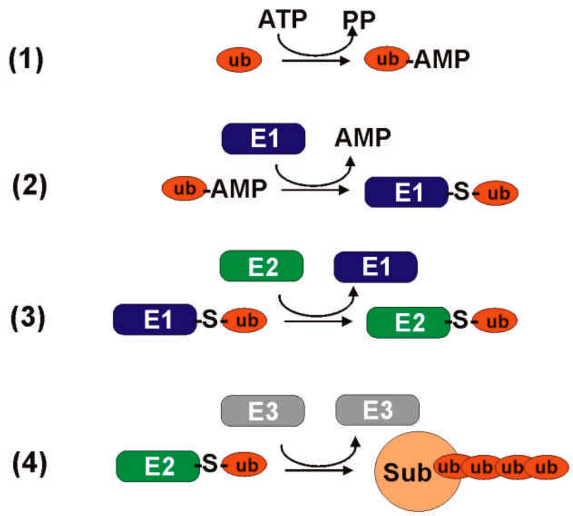

Figure 1. Ubiquitin conjugation pathway. Ub (red) stands for ubiquitin. E1 (blue), E2 (green) and E3 (grey) mark the enzymes involved in conjugation and Sub (yellow) indicates the substrate molecule.

process. A machinery of specific enzymes localized to different cellular compartments carries out these post-translational events. Many of these enzymes link a small chemical moiety like a phosphoryl, acetyl, sulfyl or methyl group to the side chain of a certain amino acid of an acceptor protein. A single enzyme, designated transferase or kinase, which exhibits a specific binding activity for both the donor and acceptor molecules, performs this transfer step. A special case of covalent protein modification is the attachment of a whole polypeptide chain (modifier) to a protein. In contrast to the transfer of a small moiety the linkage of a polypeptide chain is a multi-step process with different enzymes involved in its mechanism. The most prominent and historically oldest modifier known is ubiquitin. Over the last decade the discovery and isolation of a number of ubiquitin related modifiers was reported (reviewed in 1) implying that attachment of polypeptide chains onto acceptor molecules is a common feature to regulate spatially and timely organized cellular pathways of proteins. We focus on the structural homology and differences of the three modifiers - Ubiquitin, RUB and SUMO1 and their cognate conjugating enzymes, and discuss how these structural attributes regulate function, i.e., bring about protein modification by each of these molecules.

3. COVALENT ATTACHMENT OF PROTEINS

3.1. The players: A short description

3.1.1. Ubiquitin

Ubiquitin is a highly conserved and heat-resistant 76 amino acid protein with occurrence in apparently all eukaryotic cells and only a few prokaryotes (2). Its sequence conservation is remarkable. Ubiquitin sequences from higher plants and animals differ in only three amino acids. This is most surprising as in some organisms the protein is encoded by multiple genes (28 in the case of Arabidopsis thaliana) and is distributed among 11 different loci (3-5). The attachment of ubiquitin to target proteins is called ubiquitination or ubiquit(in)ylation. It was discovered in the late seventies of the last century (6-9; reviewed in 10) and, hence, has been extensively studied. By the action of an enzyme conjugase cascade, involving a series of thiol ester formations, the C-terminal glycine of ubiquitin is linked to the ε-amino group of lysine side chains of acceptor proteins. Mono-ubiquitination, the linkage of a single ubiquitin to one or more lysine residues of a substrate molecule plays a role in receptor-mediated endocytosis or mediates virus budding and many cellular regulatory functions (reviewed in 11). The covalent attachment of a chain consisting of several ubiquitin molecules (polyubiquitin) onto one or multiple lysine residues of a protein is called multi-ubiquitination. In polyubiquitin one lysine (either Lys29, Lys48 or Lys63) of a monomer is connected to the C-terminal glycine of a neighboring unit (reviewed in 12). A multi-ubiquination with polyubiquitin linked through Lys48 marks short-lived proteins for degradation by the 26S proteasome, the major protease in the cytosol of eukaryotic cells (13, 14). Lys63 linkages are important signals for DNA repair (15), NFκB activation (16) or ribosome function (17). Chains connected through Lys29 enhance translational activity during S-phase of cell cycling (18,19, reviewed in 20, 21). Polyubiquitin also plays a not very well defined role in neurological disorders (e.g. Alzheimer's and Huntington's diseases), where it is involved in the formation of intraneuronal inclusions like Lewy bodies or neurofibrillary tangles (reviewed in 22,23).

Ubiquitin is transcribed as an inactive precursor molecule with a C-terminal extension of several amino acids. A protease called ubiquitin carboxy-terminal hydrolase (UCH), which belongs to a family of papain-like thiol proteases, is involved in the maturation process of ubiquitin in yeast and higher eukaryotes (24). It processes the precursor C-terminal to the double-glycine motif. The C-terminal carboxyl group of ubiquitin is then activated by adenylation (figure 1). In a second step the thiol group of a cysteine residue of the activating enzyme E1 (e.g. Uba1 in yeast) attacks the carboxyl-AMP to form an E1-ubiquitin thiol ester bond. By trans-esterification E1 links the ubiquitin to the cysteine side chain of one of numerous conjugating enzymes (E2s or Ubcs). E2 now passes the modifier to a substrate lysine through formation of an amide bond between the carboxy terminal carboxylate of the modifier and the ε-NH-group of the lysine side chain of the acceptor protein. For this final step, a third enzyme, the ligase E3, is required at least in most cases. Two major types of E3-ligases are known, carrying either a HECT (25) or RING (26) domain. RING domain proteins function as scaffolds to promote ubiquitination by ensuring the appropriate arrangement of E2-ubiquitin and substrate for catalysis. HECT domain proteins participate directly in catalysis by forming a thiol ester intermediate with the modifier (reviewed in 27). Multi-ubiquitination sometimes requires a fourth enzyme E4, which binds to the ubiquitin moieties of preformed conjugates and catalyzes the chain assembly in conjunction with E1, E2 and E3 (28). The chemical details of the mechanism of ubiquitination have not been studied on an atomic level so far. However, it is reasonable to suggest, that the adenylation of ubiquitin and the thiol ester formation between E1 and the modifier may be analogous to the activation step of MoaD, an E. coli protein similar to ubiquitin (29).

Ubiquination is a reversible process that requires specific proteases for the cleavage of the isopeptide bond. Deubiquination is performed by the already mentioned ubiquitin C-terminal hydrolases (UCHs) (30).

3.1.2. SUMO

SUMO, the **s**mall **u**biquitin-related **mo**difier, is a protein of 95 to 115 amino acids, which was independently discovered in different organisms during the late nineties (31-40, for review see 41). Its sequence identity to ubiquitin is only about 18%. In contrast to yeast and invertebrates there are multiple copies of SUMO genes in vertebrates (42). Currently, there are two groups of SUMO proteins known in humans, called SUMO1 (PIC1, Ubl1, GMP1, Smt3c, hSmt3, sentrin) and SUMO2/3 (Smt3a/b, sentrin2/3). The sequence homology among the families is about 50%, whereas SUMO2 and SUMO3 are 87% identical. Like ubiquitin SUMO1 is found as covalent attachment to a number of proteins. The first target identified is RanGAP1 (36,37), the GTPase activating protein of Ran, a small GTPase involved in nucleocytoplasmic transport. Beside the effect of SUMO1 on localization and transport of proteins, the yeast homologue of SUMO1, Smt3, facilitates the disassembly of the bud neck when attached to septins (43). Meanwhile many other acceptor proteins (about 20 mammalian, four viral) are known like Mdm2 (44), c-Jun (45, 46) and p53 (46, 47, 48). The action of SUMO is diverse and yet not fully understood (reviewed in 49).

SUMO, like ubiquitin is initially synthesized as an inactive precursor, with a C-terminal extension of few amino acids (four in the case of hSUMO1), which are subsequently cleaved by cystein proteases (named Ulp1 and Ulp2 in yeast, Ulp1-Ulp7 in humans) (50). The pathway of sumoylation parallels that of ubiquitination. Like its analogue SUMO is activated by a specific E1 enzyme. While its ubiquitin directed cousine is a monomeric protein, the SUMO activating enzyme is a heterodimer in both, yeast and metazoans (51). It is composed of the two proteins AOS1 and UBA2. There is a striking sequence similarity between the carboxy-terminal region of UBA1 and UBA2 as well as between the N-terminal region of UBA1 and AOS1 (reviewed in 49). After its activation SUMO is passed to UBC9, the E2-SUMO conjugase. Sumoylation of the target substrate is catalyzed by an E3-ligase. The recently discovered PIASy is the first and yet only described E3-ligase (52). It acts as a repressor of the wnt-responsive transcription factor LEF1. Like ubiquination sumoylation is a reversible process, where cleavage of the modifier from its substrate molecule is accomplished by Ulps.

3.1.3. RUB

Rub (**r**elated to **ub**iquitin) is a small protein of 76 amino acids in length, which was first discovered in mammals in1993 (designated Nedd8) (53), two years later in Arabidopsis by homology searches (4, 54) and three years later in yeast (RUB1) (55). Like SUMO and Ubiquitin, RUB is encoded as a precurser protein and cleaved by specific proteases at the C-terminal end of its polypeptide chain to yield the active form (56-58). Interestingly, two of three RUB genes in Arabidopsis are fused to ubiquitin encoding regions. The active forms of both modifiers are released by proteolytic processing. Human forms of RUB and ubiquitin share 57% identity, whereas SUMO1 and RUB are only 21% identical. Only a few proteins are known to be modified by RUB, all of them belong to the cullin family (reviewed in 59). The best characterized target in yeast and Arabidopsis is CDC53, a part of the SCF^{Cdc4} complex that functions as an ubiquitin specific E3 ligase and is involved in cell cycle progression (57). Nedd8 seems to have a similar function in human cells (60).

The mechanism of rubylation is similar to that of ubiquitination. However, the E1 enzyme involved in the activation step is more related to the corresponding SUMO protein. The RUB specific E1 enzyme is a complex of two proteins designated AXR1 and ECR1 in Arabidopsis. After the ATP-dependent thiol ester formation of the AXR1/ECR1 complex with RUB, the modifier is passed to an E2 conjugase (Ubc12 in yeast and RCE1 in Arabidopsis) and transferred to its substrate. The last step is probably catalyzed by an E3-ligase.

3.1.4. Intertwining of pathways

Little is known about an intertwining of ubiquitin and SUMO pathways, but some observations on the activation of NFκB have demonstrated that these pathways exist (68, reviewed in 1). The inhibitor of NFκB, IκB is regulated and degraded by ubiquitination. From mutagenesis studies it has been inferred that a lysine residue of IκB is modified by both SUMO and ubiquitin. RUB and SUMO pathways are directly connected. Cullins, the targets of rubylation are part of SCF or CBC complexes, two large multimer units, which function as ubiquitin-ligases. The Nedd8 modifying-system can accelerate the formation of the E2-E3 complex and, thereby, stimulate protein polyubiquitination (60). It was shown, that conjugation of Nedd8 to ROC1-CUL1, a subcomplex of the SCF-ROC1 E3 ubiquitin ligase, selectively stimulates Cdc34-catalyzed lysine 48-linked multiubiquitin chain assembly (69)

3.2. Structural Comparison of Ubiquitin, Rub and Sumo

3.2.1. Looking at the sequence

A detailed sequence comparison of human ubiquitin with RUB and SUMO from different organisms and a scheme highlighting residues involved in binding to E2 enzymes are shown in figure 2. Red colors denote amino acids that are strictly conserved from human to yeast in all three modifiers. These residues in ubiquitin are $Val5^{ub}$, $Gly47^{ub}$, $Thr55^{ub}$, $Gly75^{ub}$ and $Gly76^{ub}$. The last two glycines are important for conjugation and reversible proteolytic deconjugation (61, 62; reviewed in 30). $Gly76^{ub}$ to Ala mutants result in derivatives that become irreversibly attached to substrate molecules, whereas deletions in

```
                                  10        20        30
                                  |         |         |
                        MQIFVKTLTGKTITLEVEPSDTIENVKAKIQ ubiq hs
                        MLIKVKTLTGKEIEIDIEPTDKVERIKERVE ned8 hs
                        MLIKVKTLTGKEIEIDIEPTDKVERIKERVE ned8 mu
                        MIVKVKTLTGKEISVELKESDLVYHIKELLE ned8 sc
                        MEIKVKTLTEKQIDIEIELTDTIERIKERIE ned8 at
MSD----QEAKPSTEDLGDKKE-GEYIKLKVIGQDSSEIHFKVKMTTHLKKLKESYC sumo hs
MSD----QEAKPSSEDLGDKKDGGDYIKLKVIGQDSSEIHFKVKMTTHLKKLKESYR sumo xl
MS------ASGGTGDEDKKPNDQMVHINLKVKGQDGNEVFFRIKRSTQMRKLMNAYC sumo le
MSDSEVNQEAKPEVKPEVKPETH---INLKV-SDGSSEIFFKIKKTTPLRRLMEAFA smt3 sc
-------------------------------xxx----------------------- u-ubc2b
--------------------------------x-------x-------x-------- u-ubc1
--------------------------------------------------------- u-ubc1t
-----------------------------xx---x---------------------- s-ubc9
--------------------------------------------------------- u-ulb1
------------------------bBBBBBBB--bBBBBBBb----hHHHHHHHHH sec

        40        50        60        70
        |         |         |         |
DKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG---------   ubiq hs
EKEGIPPQQQRLIYSGKQMNDEKTAADYKILGGSVLHLVLALRGGGGLRQ----   ned8 hs
EKEGIPPQQQRLIYSGKQMNDEKTAADYKILGGSVLHLVLALRGGGGLGQ----   ned8 mu
EKEGIPPSQQRLIFQGKQIDDKLTVTDAHLVEGMQLHLVLTLRGGN--------   ned8 sc
EKEGIPPVHQRIVYTGKQLADDLTAKHYNLERGSVLHLVLALRGGCC-------   ned8 at
QRQGVPMNSLRFLFEGQRIADNHTPKELGMEEEDVIEVYQEQTGGHSTV-----   sumo hs
QRQGVPMNSLRFLFEGQRISDHQTPKELGMEEEDVIEVYQEQTGGHSTF-----   sumo xl
DRQSVDMNSIAFLFDGRRLRAEQTPDELEMEEGDEIDAMLHQTGGSCCTCFSNF   sumo le
KRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGGATY------   smt3 sc
--------x-x-x-x-xx--------------------xxxxxxx---------   u-ubc2b
xx--------xx--------x----------------xxxxx--x---------   u-ubc1
----------xx----x---------------------xxxx--x---------   u-ubc1t
-----------xx-xx--x---------xxx-xxxx-x?????-----------   s-ubc9
----------x-x-xx-x----------x---------x-xxx-x---------   u-ulb1
HHH---hhoBBBBb—BBB----hHHHH----BBBBBBBBb--------------   sec
```

Figure 2. Sequence comparison of modifiers. Sequence numbers indicate the corresponding residues in human ubiquitin (ubiq). RUB-related proteins are abbreviated by ned8 and SUMO-related modifiers by sumo or smt3. Sequences were derived from the SWISS-PROT/TrEMBL databank or GenBank (accession codes**:** P02248, Q03968, O57686, DaQ9SMD1, Q12306, Q15843, P29595, NP_010423, AAC17623). Amino acids are classified by color code: red, identical in all modifiers; blue, homology in all groups of modifiers >75%; yellow, homology in all modifiers >40% and <75%; black, no homology or minor homology; green, homology in the RUB/ubiquitin group >80%; margenta, homology in the SUMO group >75%. Modifier residues involved in binding to enzymes or residues that show chemical shift changes in HSQC-spectra on complex formation are marked by x. The complexes used are ubiquitin/Ubc2b (u-ubc2b), ubiquitin/Ubc1(u-ubc1), thiol ester-linked ubiquitin/Ubc1(u-ubc1t), ubiquitin/Ulp1(u-ulp1) and SUMO-1/Ubc9(s-ubc9). Secondary structure elements (sec) are marked by h(low probability)/H(high probability) for helices (red) or b(low probability)/B(high probability) for β-strands (green).

$Gly75^{ub}/Gly76^{ub}$ prevent growth of modified yeast strains (63). In the same experimental approach $Gly47^{ub}$ in ubiquitin was shown to be essential for vegetative growth in yeast. The function of this residue is not clear, yet. It is often suggested that the importance of $Gly47^{ub}$ is due to its direct neighborhood to $Lys48^{ub}$. However, this seems to be unlikely considering the fact, that $Gly68^{SUMO}$ is conserved at the respective sequence position in SUMO, although $Gln69^{SUMO}$ does not play any role in conjugation. From figure 2 it is obvious that the conserved $Gly47^{ub}/68^{SUMO}$, which is part of a small loop connecting β3 and β4, is located within the modifier/E2-binding interface. Thus, it might be important for the interaction with the corresponding Ubc protein. Val5 is a central part of the hydrophobic core of the modifier, which probably explains the conservation of this residue. Sequence comparison alone gives no information on the conservation needs for Thr55.

Some amino acids in figure 2 exhibit either a high (blue color) or medium (yellow color) degree of conservation through all groups of modifiers and all species. These residues are mainly involved in the hydrophobic core or in salt bridges and, hence, are important for the structural integrity of the protein. The most interesting amino acids are either the green and

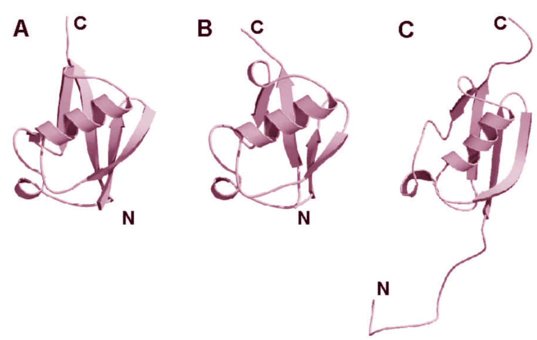

Figure 3. Molscript representations of the modifiers: a, ubiquitin (RCSB entry 1ubq), b, Nedd8 (1ndd), c, SUMO-1 (1a5r). The N- and C-termini of the proteins are labeled by capital letters.

magenta colored residues, since they determine the difference between the SUMO and the ubiquitin/RUB group of modifiers, or the black colored amino acids, since they might provide the origin of individuality (binding specificity and selectivity for enzymes). This statement is especially true for amino acids located in loop and turn regions. One very good example for this statement, which has been extensively studied is the role of amino acid 72 in RUB and ubiquitin. An Arg72^{ub} to Ala mutation was shown to alter the binding of ubiquitin-adenylate to its E1 (64) and an Ala72^{Nedd8} to Arg72 mutation in Nedd8 bound to the ubiquitin directed E1 almost as well as wild-type ubiquitin (65). Residue Ala72^{RUB} is thought to perform a key role in selecting against reaction with the ubiquitin-specific E1 enzyme and, thereby acting to prevent the inappropriate diversion of RUB into ubiquitin-specific pathways (65).

On basis of the sequence alignment the SUMO and ubiquitin/RUB families are very well separated from each other. One could hypothesize, that only few amino acids are determinants for the group character. For example the green (magenta in sumo) colored residues (figure 2) Thr7$^{ub/RUB}$-Thr9$^{ub/RUB}$, Lys11$^{ub/RUB}$, Glu/Asp16$^{ub/RUB}$, Glu/Tyr21$^{ub/RUB}$, Glu24$^{ub/RUB}$, Arg33$^{ub/RUB}$, Gln34$^{ub/RUB}$, Pro38$^{ub/RUB}$, Gln40$^{ub/RUB}$, Gln41$^{ub/RUB}$, residue 46$^{ub/RUB}$ (black color), Lys48$^{ub/RUB}$, Gln49$^{ub/RUB}$, His68$^{ub/RUB}$, Leu73$^{ub/RUB}$ and Arg74$^{ub/RUB}$ separate the ubiquitin/RUB family from the SUMO group. Clearly, the polarity (charges) of side chains at positions 9, 11, 16, 21, 24, 40, 49 and 74 of ubiquitin/RUB is changed in SUMO.

Recent studies on ubiquitin (66,67) have revealed residues Phe4^{ub} and Ile44^{ub} to be crucial for internalization and Leu8^{ub}, Ile44^{ub} and Val70^{ub} for proteasome binding and degredation of proteins. Phe4^{ub} is unique for ubiquitin, whereas RUB and SUMO proteins carry a lysine at this position. This lysine can conserve part of the hydrophobic character of Phe4^{ub}, but at the same time introduces a positive charge close to the hydrophobic patch surrounding this residue. It might prevent RUB and SUMO to compete with ubiquitin for internalization. Additionally, there are some minor differences that separate the SUMO class from the ubiquitin/RUB group. For example, sequence position 44 is occupied by isoleucine or valine in ubiquitin and RUB, but is substituted by leucine in SUMO. Also Leu8^{ub} and Val70^{ub} are replaced by various other residues in SUMO. The observation that the hydrophobic residues Leu8^{ub}, Ile44^{ub} and Val70^{ub}, which are crucial for proteasome binding, are strictly conserved among the ubiquitin/RUB family, but not within the SUMO group, kicks SUMO out of the degradation pathway, however, at the same time leaves a door open for RUB.

3.2.2. Comparison of the secondary and tertiary structures of modifiers

Ubiquitin, RUB and SUMO consist each of a five-stranded β-sheet with antiparallel aligned single strands, an α-helix connecting β2 and β3 and a short helix between β4 and β5 (figure 3). This arrangement represents the typical ubiquitin superfold (70). There are many variations on this theme and the ubiquitin superfold has become a paradigm for structural conservation in absence of significant sequence identity. It is found in proteins that are functional entirely unrelated to ubiquitin like the (2F-2S)-ferredoxin (71), the Ras-binding domains (72-74) and numerous other proteins.

The topologies of AtRUB1 and ubiquitin are very similar. The root mean square deviation (rmsd) of the two sets of α-carbon atoms is 0.612 Å (54) or 0.8 Å (DALI search). The rmsd value between SUMO-1 and ubiquitin (75) and SUMO-1 and Nedd8 (54) are 2.1 Å and 2.6 Å, respectively. The hydrophobic core of ubiquitin/RUB (SUMO) is made by side chains of residues 3$^{ub/RUB}$(24^{SUMO}), 13$^{ub/RUB}$(34^{SUMO}), 15$^{ub/RUB}$(36^{SUMO}), 43$^{ub/RUB}$(64^{SUMO}), 45$^{ub/RUB}$(66^{SUMO}), 67$^{ub/RUB}$(88^{SUMO}) of α-helix1 and residues 23$^{ub/RUB}$(44^{SUMO}), 26$^{ub/RUB}$(47^{SUMO}), 27$^{ub/RUB}$(48^{SUMO}) and 30$^{ub/RUB}$(51^{SUMO}) of the β-sheet. These amino acids are highlighted in blue color in figure 2. Leu8$^{ub/RUB}$, Ile44$^{ub/RUB}$ and Val70$^{ub/RUB}$ cluster together to form a hydrophobic patch on the surface (66), which is absent in SUMO-1.

Several salt bridges are stabilizing the topology of the modifiers. In the structures of ubiquitin and AtRUB1 Lys27$^{ub/RUB}$ and Asp52$^{ub/RUB}$ (2.74 Å and 2.90 Å) form a salt bridge and a second one exists between Lys11^{ub} and Glu34^{ub} (3.3 Å) (54). In SUMO His75 and Glu79 are involved in a salt bridge (2.6 Å) (75).

Structure comparison allows to explain the conservation of Thr55 in all three modifier groups, although mutations in this position had no effect on the viability of yeast cells (63). Thr55 is located at the surface of the protein and a detailed analysis of the structures of ubiquitin and Nedd8 reveals the side chain oxygen group to form a hydrogen bond to the amino group of Asp78$^{ub/RUB}$, thereby stabilizing helix2. Unfortunately, the SUMO structure is only weakly determined in this region (75).

A major structural difference between the ubiquitin/RUB class of proteins and its SUMO relatives is found at the N-termini of the modifiers. SUMO-1 exhibits a flexible and unstructured 21 residue long extension preceding the ubiquitin-like core domain. The function of this tail remains unclear. Homology searches (FASTA3) using the N-terminus as bait revealed 50% identity to

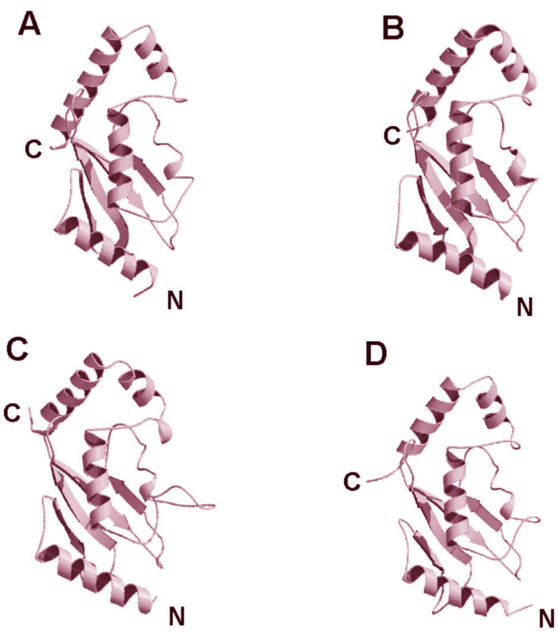

Figure 4. Molscript representation of E2 enzymes: a, ubc2 (RCSB entry 1ayz), b, ubc4 (RCSB entry 1qcq), c, ubc7 (RCSB entry 2qzq), d, ubc9 (RCSB entry 1a3s).

sequences of HMGC1 protein and chicken filamin. Ser9 and Tyr21 were detected as possible phosphorylation sites in SUMO-1 using the web engine of Blom (76). However, no post-translational modifications have been observed in SUMO until now.

3.2.3. Structure and function of E2 enzymes

E2 enzymes (Ubcs) are predominantly localized to the nucleus (77). They can be classified into four structural categories: enzymes that are comprised of a core domain (about 150 residues) carrying the active site cysteine (class I), those who additionally have either a C- or N-terminal extension (class II and III), and a group of proteins with both C- and N-terminal tails (class IV). Ubcs share a highly conserved core domain topology (figure 4) comprising a four-stranded β-sheet, one long N-terminal and three C-terminal α-helices with about 16(α1), 4(α2), 15(α3) and 35(α4) residues, respectively. The active site cysteine (Cys86^{ubc4sc}) is placed at the end of an extended loop region between β4 and α2. A common consensus sequence for all classes of ubiquitin E2 enzymes, Ubc12, the RUB directed enzyme, and Ubc9, the SUMO-1 directed conjugase, can be derived from sequence comparison (figure 5). 12 amino acids are identical in all human and yeast Ubc sequences (red color). In Ubc4 of the yeast Saccharomyces cerevisiae these residues are Gly49, Tyr61, Pro62, Pro66, His76 to Asn78, Gly83, Cys86, Leu87, Trp94 and Pro96. As these amino acids are conserved among all different Ubcs, they have to be critical for stability or function of the E2 proteins. Residues His76^{ubc4sc}-Leu87^{ubc4sc} form the binding interface and are crucial for the mechanism of conjugation by either being directly involved in the catalysis or by maintaining the proper scaffold for this action. The Ubc-interfaces are stabilized by the N-terminal α-helix of the E2 enzyme. The necessity to present a well defined binding site towards the modifier and, thus, allow specific recognition of the E1 complex, might be the reason for the conservation of amino acids Lys74^{ubc4sc}-Leu90^{ubc4sc} and Trp93^{ubc4sc}-Leu120^{ubc4sc}. A variable region of mainly three up to 16 residues, whose function could not be elucidated so far, interrupts these conserved sequence regions. In the case of Ubc13 amino acids Pro58^{ubc4sc}-Ile68^{ubc4sc}, which connect β3 and β4, were previously shown to interact with the E3-ligase Mms2 (78). Together with the RING finger protein Traf6, the Ubc13-Mms2 heterodimer catalyses the formation of Lys63^{ub}-linked ubiquitin chains (79). Amino acids 58^{ubc4sc}-68^{ubc4sc} might not only preserve the specific surface topology for this interaction, but also link and fix the modifier-binding site close to that of the corresponding ligase.

Non-conserved residues in the E2-modifier complex are the major determinants for selection of the proper ligase. Amino acids of Ubc13 critical for Mms2 binding are Tyr34^{Ubc13}, Phe57^{Ubc13} and Leu83^{Ubc13}. In the case of the E6Ap(HECT)-Ubc7h complex Phe63^{Ubc7h} is the primary determinant for E2 specificity. Together with Ala59^{Ubc7h}, Glu60^{Ubc7h}, Pro62^{Ubc7h} and the amino acids Lys96^{Ubc7h}, Ala98^{Ubc7h} and Lys100^{Ubc7h} the residue Phe63^{Ubc7h} is involved in direct contacts to the HECT domain (80). The fact that non-conserved residues are the origin of individuality of proteins also holds for other Ubcs. For example, deletion of the N-terminal nine amino acids (part of the first helix) of Rad6 (Ubc2sc) prevents the enzyme from forming a stable complex with the E3-ligase Ubr1 (81). The same was observed upon deletion of the solvent exposed residues Glu150^{Rad6}, Asp151^{Rad6}, Asp152^{Rad6} and Met153^{Rad6}, which are part of the C-terminus of Rad6.

4. INTERACTION OF E2s AND MODIFIERS

NMR titration experiments provide the necessary platform for a detailed comparative view on complex formation of Ubcs and their respective modifiers. Results of shift perturbation studies allowed mapping of the binding interfaces of ubiquitin and the two E2 conjugating enzymes Ubc1 and Ubc2b (82-84). HSQC spectra of free ubiquitin revealed changes in ^{1}H and ^{15}N chemical shifts of numerous resonances on the addition of Ubc2b. Residues belonging to these resonances cluster on one side of the protein surface (figure 6A) around the C-terminal Gly-Gly motif (red color). A collection of basic amino acids (Arg42^{ub}, Lys48^{ub}, Arg72^{ub}, Arg74^{ub}) is present in this region. Thus, the binding interface of ubiquitin is positively charged (figure 6B). The side chains of residues Leu8^{ub}, Ile44^{ub}, Val70^{ub}, Leu71^{ub} and Leu73^{ub} provide hydrophobic patches for the interaction with Ubc2b. Thr7^{ub}, Thr9^{ub}, Gln40^{ub} and Gln49^{ub} can serve as hydrogen bond donors or acceptors. Similar results could be obtained by shift perturbation experiments using an Ubc1-modifier complex linked by a thiol ester bond. The binding interface presented in figure 6C/D was mapped on the surface of ubiquitin (green color) after computational docking based on NMR data. In figure 6C/D the C-terminus of the modifier sticks out of the paper plane. The same amino

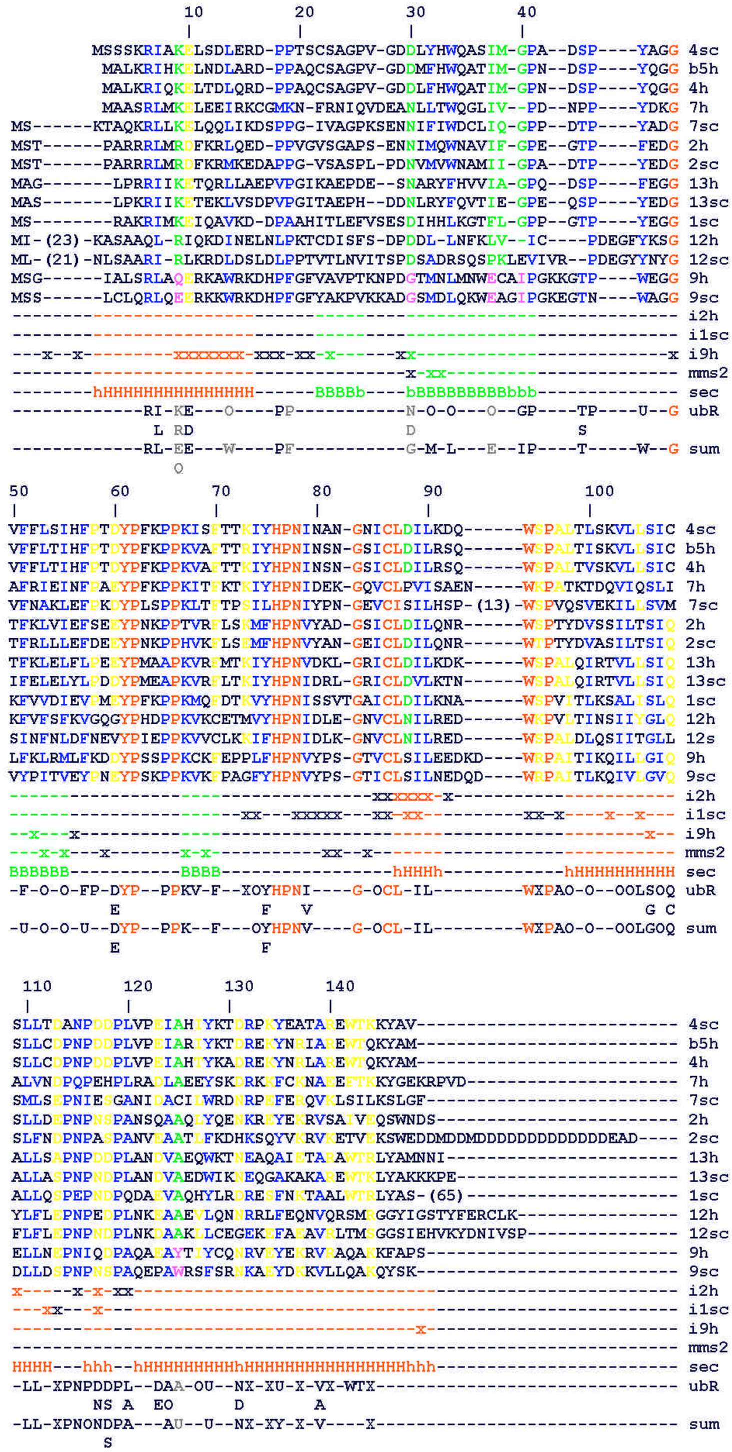

Figure 5. Sequence comparison of human (h) and yeast (sc) E2 enzymes. Sequence numbers indicate the corresponding residues in Ubc4 of Saccharomyces cerevisiae. Ubc enzymes used in this context are indicated by there Ubc number and the respective species abbreviation, e.g.12sc stands for Ubc12 of yeast. Sequences were derived from the SWISS-PROT/TrEMBL databank (accession codes**:** P15731, P51669, Q9Y2X8, P51966, Q02159, P23567, P06104, O76069, P52491, P50550, P50623, P21734, Q16781, P52490). Amino acids are classified by color code: red, identical in all E2s; blue, homology in all groups of E2s >75%; yellow, homology in all E2s <75% and >50%; black, no homology or minor homology; green, homology in Ubc12 group >75%; margenta, homology in the SUMO group >75%. Residues of Ubcs involved in binding to modifiers, residues of Ubc13 known to contact mms2 and residues that show chemical shift changes in HSQC-spectra on complex formation are marked by x. The complexes used for interface identification are ubiquitin/Ubc2b (i2h), ubiquitin/Ubc1 (i1sc), SUMO-1/Ubc9 (i9h) and Ubc13/Mms2 (mms2). Secondary structure elements (sec) are marked by h(low probability)/H(high probability) for helices (red) or b(low probability)/B(high probability) for β-strands (green). A consensus sequence of E2 (ubiquitin/RUB group) is indicated by ubR and for the SUMO group by sum. Grey color indicates residues that differ between SUMO and ubiquitin/Nedd8 family.

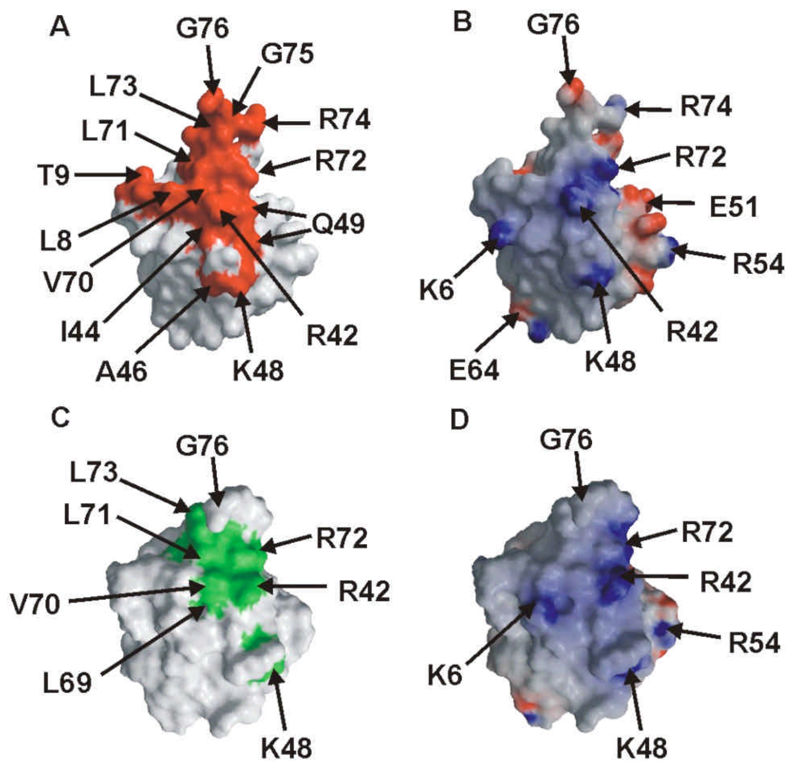

Figure 6. Binding interface and surface charges of ubiquitin. a, Molecular surface representation (GRASP) of ubiquitin. Residues, which show chemical shift changes upon addition of Ubc2b are mapped onto the surface by red color. b, Surface charges across the binding interface. Negative charges are colored in red, positive ones in blue. c, Molecular surface representation (GRASP) of ubiquitin complexed to Ubc1. Residues, which are involved in direct contacts to the thiol ester-linked Ubc1 are colored in green. d, Surface charges of complexed ubiquitin. Negative charges are colored in red, positive ones in blue.

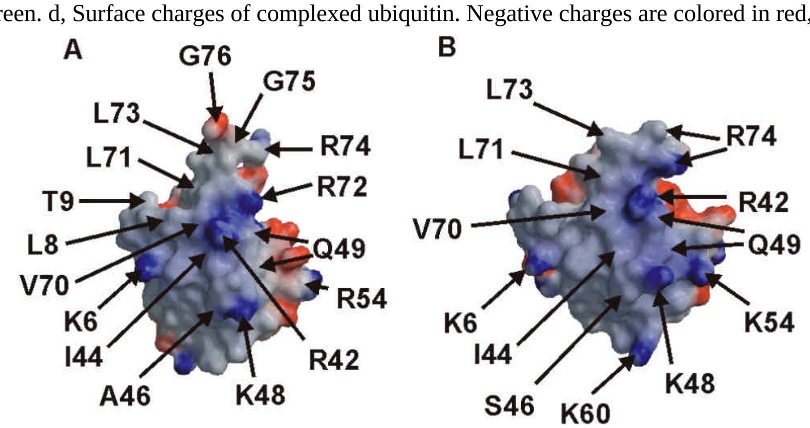

Figure 7. Comparison of the surface charges of the binding interfaces of a, ubiquitin and b, Nedd8. Negative charges are colored in red, positive ones in blue.

acids, which are involved in contacts to Ubc2b, form the modifier interface pointing towards Ubc1. These residues are mainly located in the β-sheet and the loop region connecting α3 and β5 of ubiquitin.

If ubiquitin always presents the same interface towards an E2 enzyme and only one E1 enzyme is used for activation, how can the E1-ubiquitin complex select for different Ubcs? One can currently think of two ways on how to answer this question. First, the "proper" Ubc is chosen by no selection criteria and each Ubc has the same likelihood to be accepted. This would mean, that probability of complex formation is proportional to the amount of Ubcs present in the cell. Selection would then be achieved by regulating the expression of E2 enzymes. Second, the affinities of Ubcs to the E1-ubiquitin complex are different, thereby, serving as a selection mechanism. Probably, the cell will use both criteria at the same time.

RUB closely resembles the surface properties of ubiquitin (figure 7). If one defines the binding interface of RUB on the basis of the structural homology of both

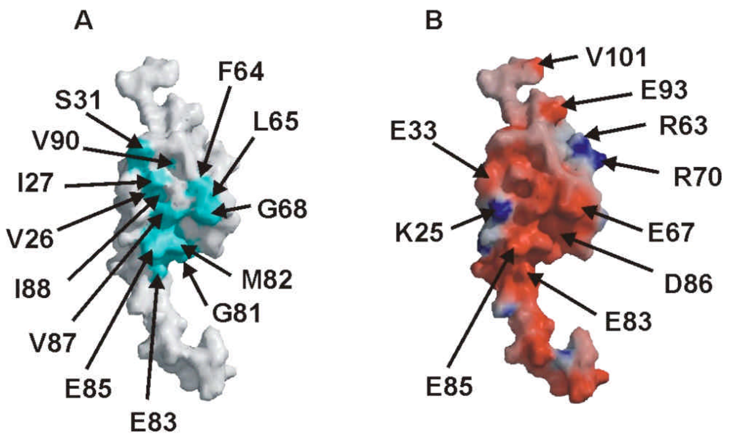

Figure 8. Binding interface and surface charges of human SUMO-1. a, Surface charges within the binding interface. Negative charges are colored in red, positive ones in blue. b, Molecular surface representation (GRASP) of SUMO-1. Residues, which show chemical shift changes upon addition of Ubc9 are marked by cyan color.

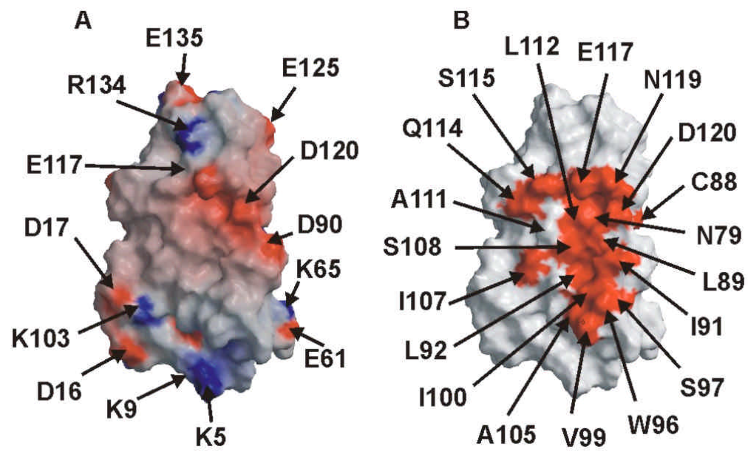

Figure 9. Binding interface and surface charges of the E2 enzyme Ubc1. a, Surface charges within the binding interface. Negative charges are colored in red, positive ones in blue. b, Molecular surface representation (GRASP) of Ubc1. Residues, which contact amino acids in ubiquitin are marked by red color.

proteins, nearly identical hydrophobic patches and charges are placed at comparable positions. Taking only homology criteria into consideration, RUBs might be able to use the ubiquitin E2-machinery. Thus, discrimination between the two pathways must occur at an earlier stage of the conjugase cascade, e.g. at the E1 branch point (see Ala72 of RUB).

The binding interface of SUMO is built up analogous to those of RUB and ubiquitin and is formed by residues located in the β-sheet and the neighboring loop regions. In contrast to RUB and ubiquitin, SUMO carries an extremely negative surface charge (figure 8 A and B). The side chains of $Glu33^{SUMO}$, $Glu67^{SUMO}$, $Glu83^{SUMO}$, $Glu85^{SUMO}$, $Glu93^{SUMO}$ and $Asp86^{SUMO}$ are directly involved in or positioned close to the interface. On the basis of NMR shift perturbation experiments no decision could be made on the functional role of $Lys25^{SUMO}$, the most prominent basic residue, which is conserved in different species (85). The same problem was faced for $rg63^{SUMO}$ and $Arg70^{SUMO}$, the other two positive charged amino acids flanking the interface.

Currently, there is no solution or crystal structure of any modifier-E2 complex available in the databank, but, as already mentioned, an ubiquitin-Ubc1 model derived from computational docking and NMR shift perturbation studies (83). Relying on this model one can find those amino acids that make direct contacts between both molecules. In figure 9 the corresponding interface is mapped on Ubc1 (figure 9B) and surface charges across the E2 binding site are marked (figure 9A). The interface area is about 1200 $Å^2$ and consists of a large hydrophobic patch formed by amino acids $Leu89^{Ubc1}$, $Ile91^{Ubc1}$, $Leu92^{Ubc1}$, $Trp96^{Ubc1}$, $Val99^{Ubc1}$, $Ile100^{Ubc1}$, $Ala105^{Ubc1}$, $Ile107^{Ubc1}$ and $Leu112^{Ubc1}$, the two negative charged residues $Glu117^{Ubc1}$ and $Asp120^{Ubc1}$ and the hydrogen bond donors (acceptors) $Asn79^{Ubc1}$, $Ser108^{Ubc1}$, $Gln114^{Ubc1}$, $Ser115^{Ubc1}$ and $Asn119^{Ubc1}$. The side chains of $Leu89^{Ubc1}$ and $Leu112^{Ubc1}$ make van der Waals-contacts to the γ-protons of $Val70^{ub}$, $Ile91^{Ubc1}$

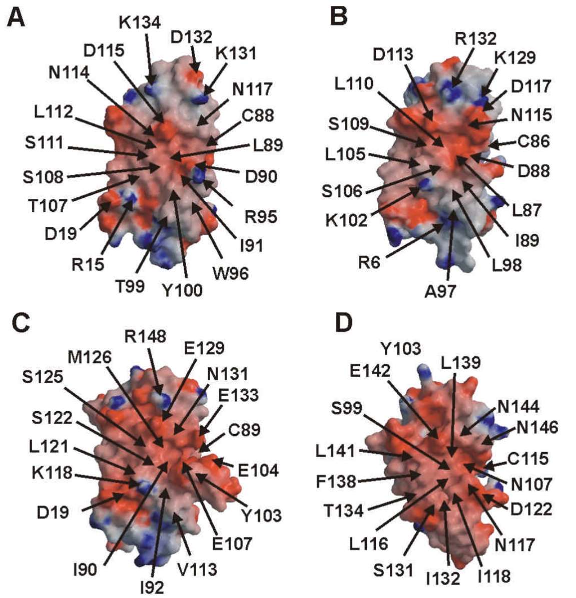

Figure 10. Comparison of surface charges of the binding interfaces of four different E2s. Negative charges are colored in red, positive ones in blue. a, Ubc2, b, Ubc4, c, Ubc7, d, Homology model of Ubc12, which based on the coordinates of Ubc4.

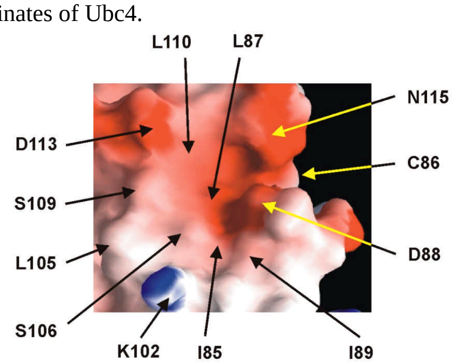

Figure 11. Conserved binding interface shown exemplary for Ubc4. Residues, which are conserved in all Ubcs are indicated.

nteracts with Leu8^{ub}, Leu7^{ub} and Leu73^{ub}, and Leu89^{Ubc1} binds to Val70^{ub} and Leu8^{ub}. There are additional contacts from Leu112^{Ubc1} to Val70^{ub} and from Ile107^{Ubc1} to Lys6^{ub}. Hydrogen bonds are formed between Gln114^{Ubc1} and Gly47^{ub} and between Glu117^{Ubc1} and Arg42^{ub} or Arg72^{ub}. Salt bridges between the argenines and the glutamic acid stabilize these interactions. The hydrogen bond between Gly47^{ub} and Gln114^{Ubc1} probably explains the sequence conservation of the glycine residue among all modifiers and species. It might be possible that hydrogen bonds exist between Asn119^{Ubc1} and Arg72^{ub} or Arg74^{ub}, but they are not formed in the model complex. Residue Ser108^{Ubc1}, which is observed in all ubiquitin and RUB directed E2s, shows no noteworthy side chain contacts, but might be involved in the formation of hydrogen bonds to Lys6^{ub}. Although Ile44^{ub} is placed in the center of the biquitin interface, there are only few contacts to hydrophobic side chains (Ala111^{Ubc1}). The C-terminal Gly76^{ub} is sandwiched between Asp90^{Ubc1} and the active site cysteine C88^{Ubc1}.

On the basis of the model and NMR titration experiments the interfaces of all structurally known Ubcs can be defined (86-89). A comparison of the ubiquitin binding sites of Ubc2, Ubc4, Ubc7 and the RUB-binding site of Ubc12 (homology model) is shown in figure 10. The interfaces of Ubc1, Ubc2, Ubc4 and Ubc7 are composed of homologous residues. There are some prominent features conserved, which can be found at the surface topology of each Ubc. These hallmarks are summarized in table 1 and highlighted in figure 11. Most outstanding is the hydrophobic groove formed of the side chains of five amino acids. Residues therein contact the hydrophobic patches found in ubiquitin and RUB. The groove turns into a small cleft, which picks up the C-terminal residues of the modifier, and finally ends in the active site cysteine. The small surface channel is flanked by the conserved negative side chain of Asp88^{Ubc4} and the asparagine residue Asn115^{Ubc4} located at the merging point of groove and cleft. A second conserved charged residue (Asp113^{Ubc4}) is found at the opposite side of the hydrophobic groove. Interestingly, the spatial arrangement of Asp113^{Ubc4} and Asn115^{Ubc4} is conserved in Ubc1, but the corresponding amino acids are shifted by two residues in the alignment when compared to the sequences of Ubc2, Ubc4 and Ubc7. The binding interface of Ubc12 shows no remarkable differences to ubiquitin directed Ubcs. The positive charged Lys102^{Ubc4} is absent in Ubc12. In the Ubc1-ubiquitin complex this residue (Lys103^{Ubc1}) is stabilizing the E2 fold, while forming hydrogen bonds to Val14^{Ubc1} and Asp17^{Ubc1}. It did not contact residues within ubiquitin. The absence of a positive charge in the corresponding region of Ubc12 is perhaps due to the lack of 26 N-terminal amino acids, which could not be attached to the protein fold by homology modeling. Nevertheless, the remarkable homology between interfaces of ubiquitin and RUB directed Ubcs mirrors the homology found among the modifier interfaces.

The situation is quite different regarding the interface of Ubc9 and SUMO (figure 12) (90,91). The major protein-protein contacts are formed by an alternative set of residues, although the topology of the binding site is conserved in Ubc9 and homologues amino acids are placed at positions proved to bind to modifier residues in ubiquitin and RUB directed enzymes (figure 2). The binding interfaces of the SUMO-Ubc9 complex are highly complementary in their electrostatic potentials and hydrophobicity (85). It is likely, that a basic patch of residues (Arg13, Lys14, Arg17, Lys18) of Ubc9 faces an acidic counterpart in SUMO (Glu83, Glu84, Glu85, Asp86). Asp19^{Ubc9} is close enough to Lys24^{SUMO} to form a salt bridge. Still, the flexible C-terminus of the modifier could span the distance (~14 Å) to the active site cysteine of the conjugating enzyme, but it is not able to saturate the hydrophobic groove at the surface of Ubc9. This would mean that residues of other molecules (E3-ligase?) are required to cover sticky side chains. It is also possible that the observed binding site at Ubc9 is only covered by a non-thiol ester-linked SUMO. The modifier might interact with different residues of the E2 enzyme while it is connected through a thiol ester bond. However, this is in contrast to experiments performed on the ubiquitin-Ubc2/1 complexes,

Table 1. Topological conserved residues[1]

Ubc1sc	Ubc2sc	Ubc4sc	Ubc7sc	Ubc9h	Ubc12h
K103	R15?	K102	K118	K110	Part missing
E117	D115	D113	E129	E122	E142
S108	S108	S106	S122	?	S131
A111	S111	S109	S125	F22?	F138?
I107	T107	L105	L121	L114	T134
L112	L112	L110	M126	L119	L139
L89	L89	L87	I90	L94	L116
I91	I91	I89	I92	I96	I118
I87	I87	I85	V88	V92	V114
N119	N117	N115	N131	N124	N144
D90	D90	D88	E107/E108?	D102?	N117 / E121/D122?
C88	C88	C86	C89	C93	C115

[1] Residues that are conserved at identical spatial position upon all E2 surfaces are indicated. The headline of the table indicates the corresponding E2 enzyme. Question marks are used for those positions that could not be assigned unambiguously to a certain residue.

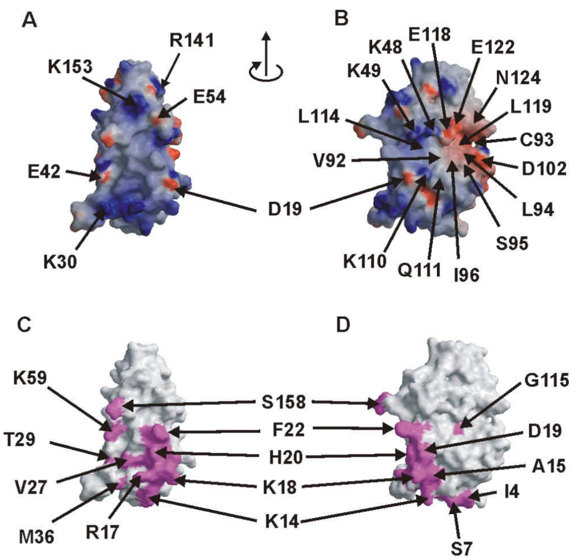

Figure 12. Binding interface and surface charges of Ubc9. a and b, Surface charges within the binding interface. Negative charges are colored in red, positive ones in blue. c and d, Molecular surface representation (GRASP) of Ubc9. Residues, which showed chemical shift changes upon addition of Ubc9 are colored in margenta. b and d are rotated with respect to A by 90° as marked by the arrow.

where it seemed to be sufficient to use non-covalently linked modifier to map specific residues within the interface of Ubc1. In the case of Ubc1/2 complex formation without the thiol ester bond no crucial differences in the interface areas to the thiol ester conjugated complex were observed (82,84).

5. PERSPECTIVES

5.1. Mechanisms of target selection

Different substrate molecules are specifically recognized and degraded by the ubiquitination machinery due to the combined action of different Ubcs and substrate specific E3-ligases. The acceptor sites for ubiquitination of target molecules send to the proteasome are quite different and depend on the recognition site of the corresponding E3 ligase. These motifs are called degrons and each degron has a corresponding E3 ligase.

In contrast to ubiquitin and RUB conjugation targets it was previously suggested that PEST sequences within the substrate might be necessary for the SUMO directed enzyme machinery to choose the right domain for sumoylation. However, there are many potential and known targets without carrying such a motif (41). Acceptor sites for sumoylation of proteins contain the single consensus motif aKxE (a is a hydrophobic amino acid and x can be anyone) that seems to be present in the majority of target molecules (41), but is not strictly conserved in all known substrate molecules (table 2).

So, how does the E2-E3-SUMO complex select for target proteins? Does it really grip a four amino acid motif? No structure of the E3 ligase or a sumoylated protein is known and thus one can only speculate on how sumoylation sites are recognized. According to a Prosite search 70% of all eukaryotic sequences carry at least one aKxE pattern, hence, all of these proteins would be potential targets for SUMO. As far as we know, this seems to be unlikely and indicates an underestimation of the actual size of the recognition motif. There are probably other aspects that have to be taken into consideration, for instance, the influence of secondary or tertiary structure of the target molecule. The motif could reside in a special element (α-helix, β-sheet etc.) or be located in close proximity to a topological feature of the protein (hydrophobic patch, charge, etc). It can either be recognized by just one of the binding partners of the E2-E3-SUMO complex or by an interface formed by two or three molecules of the multimer. To present a possible scenario what one might have to look for, the recognition sites of numerous proteins, known to be sumoylated (41), were examined for secondary structure elements. GoriV and nnPredict was used for the secondary structure prediction. Based on the algorithm the sequences can be separated into two classes. The first class contains α-helical structures and the second class contains the C- and N-termini of proteins or loop regions, which

Table 2. Sumoylation sites [1]

Number	Acceptor protein	Residue number	Sequence pattern	Topology
1	PML	K490	PRKVIKMESEE	helix
2	RanGAP1	K526	HMGLLKSEDKV	helix
3	PML	K160	HQWFLKHEARP	helix
4	IkBa	K21	PRDGLKKERLL	helix
5	p53	K386	KKLMFKTEGPD	helix-coil
6	PML	K65	CQAEAKCPKLL	helix-loop-helix
7	Cdc3	K4	MSLKEEQVS	N-terminus-helix
8	Cdc3	K11	EQVSIKQDPIQ	helix-loop-helix
9	c-Jun	K229	RLQALKEEPQT	β-loop-β
10	Mdm2	K446	CQPRPKNGCIV	β-loop-β

[1] Source: Ref 41.The sequences are numbered from 1 to 10 in column one. Second column, Names of target molecules. Third column, Number of acceptor lysine residue in the sequence. Fourth column, Sequence pattern containing the acceptor lysine. Fifth column, Secondary structure prediction of sequence. Italic letters indicate flanking sequences.

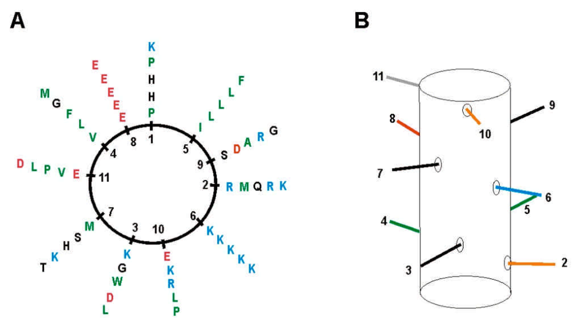

Figure 13. Helix-wheel representation of amino acid sequences. a, Sequences predicted to be helical (table 2) were ordered in circles around a helix-wheel. Green color marks hydrophobic residues, red color acidic and blue color basic ones. b, 3D-representation of a. Residues are represented by sticks.

connect α-β or β-β elements. The aKxE motif is less conserved in the latter class of sequences. The position of the aKxE motif within extended helices is close to the C-terminal end. Hence, the recognition pattern prefers to be located at an exposed site of the molecule.

In figure 13 the sequences of table 2 are shown in a helix-wheel representation. The side chains of the conserved lysine and glutamic acid residues are oriented towards opposite sides. Mainly hydrophobic residues surround the glutamic acid, whereas the helix side carrying the lysine is preferentially polar. At least one additional positively charged group is present in either position 2 or 10 of the wheel.

Although the detailed mechanism of modifier attachment onto protein targets is not known, one can suggest that the substrate molecule has to come close to the active site cysteine of Ubc9. As an example, the hydrophobic groove on the E2 surface, which is not saturated by SUMO, could pick up the hydrophobic part of the helix. The conserved glutamic acid might form a salt bridge to $Arg104^{Ubc9}$ and the positive charge on the opposite side could face $Glu118^{Ubc9}$ or $Glu122^{Ubc9}$, thereby stabilizing a conformation, wherein Lys^{motif} could be linked to the target protein. The aKXE motif of loop regions might be able to adopt a similar conformation by an induced fit mechanism. But, of course, this is just one possible hypotheses.

5.2. Polysumoylation?

It is often stated that the reason for the absence of polysumo chains is the substitution of three lysines in SUMO, which were shown to be key residues for polyubiquitin formation. However, there are no obstacles for using any other lysine for synthesis of polysumo chains. One such candidate is Lys39, which is on the opposite side of the SUMO binding interface. Also lysine residues in the flexible N-terminus might be possible target sites.

Recent studies on the structure of Ubc13-Mms2 (78), the complex catalyzing the formation of Lys63-linked ubiquitin polymers, have shed some light onto the mechanism of polyubiquitin synthesis. Mms2 is an ubiquitin E2 variant (UEV) lacking the cystein that is necessary for thiol ester formation in Ubc13. The heterodimer forms a T-shaped complex, which enables the linkage of two ubiquitin monomers head-to-tail. The efficiency of polyubiquitin synthesis is increased by binding of Rad5 (homologue of Traf6), an E3 ligase belonging to the RING domain family of proteins.

The question is now: is there a SUMO E2 variant that can form complexes with Ubc9 or a non-discovered Ubc? The answer is not straightforward, but there are some interesting findings recently published in a paper on the SUMO E3 ligase PIASy (52). Immunoblot (SDS-PAGE) studies were performed to demonstrate the activity of the enzyme by sumoylation of the wnt-dependent transcription factor LEF1. In some of the figures SUMO-LEF1 conjugation products are shown. Surprisingly, there are ladders of several different species of LEF1-SUMO conjugates visible. As the blots do not have any molecular weight markers and LEF possesses two putative sumoylation sites, one can only speculate, whether these blots might indicate the presence of multi-sumoylation.

6. ACKNOWLEDGEMENT

We like to thank Jonathan Mueller and Elena Bayer for carefully reading the manuscript. This work was supported by grants from the Deutsche Forschungs-gemeinschaft (DFG) (BA 1624/3-2 and 1624/4-1) and the Max-Planck Institute for Molecular Physiology Dortmund. We apologize to many authors that could not be cited due to space constraints.

7. REFERENCES

1. Hochstrasser M: Evolution and function of ubiquitin-like protein-conjugation systems. *Nat Cell Biol* 2, E153-E157 (2000)
2. Goldstein G**,** M. Scheid, U. Hammerling, D. H. Schlesinger, H. D. Niall & E. A. Boyse: Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universaly in living cells. *Proc.Natl.Acad.Sci USA* 72, 11-15 (1975)
3. Callis J. A, J. A. Raasch, & R. D.Vierstra: Ubiquitin extension proteins in Arabidopsis thaliana: Structure, localization and expression of their promotors in transgenic tobacco. *J.Biol.Chem.* 265, 12486-12493 (1990)
4. Callis J. A, T. B. Carpenter, C.-W. Sun & R. D. Vierstra: Structure and evolution of genes encoding polyubiquitin and ubiquitin like proteins in Arabidopsis thaliana ecotype Columbia. *Genetics* 139, 921-939 (1995)
5. Vierstra R. D & J. Callis: Polypeptide tags, ubiquitous modifiers for plant protein regulation. *Plant Mol Biol* 41, 435-442 (1999)
6. Ciechanover A, Y. Hod & A. Hershko: A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticolocytes. *Biochem.Biophys.Res.Commun*. 81, 1100-1105 (1978)
7. Wilkinson K. D, M. K. Urban & A. L. Haas: Ubiquitin is the ATP-dependent proteolysis factor of rabbit reticolocytes. *J.Biol.Chem.* 255, 7529-7532 (1980)
8. Ciechanover A, H. Heller, S. Elias, A. L. Haas & A. Hershko: ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proc.Natl.Acad.Sci USA* 77, 1365-1368 (1980)
9. Hershko A, A. Ciechanover, H. Heller, A. L. Haas & I. A. Rose: Proposed role of ATP in protein breakdown: conjugation of proteins with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc.Natl.Acad.Sci USA* 77, 1783-1786 (1980)
10. Hershko A, A. Ciechanover & A. Varshavsky: The ubiquitin system. *Nat. Med.* 6: 1073-1081 (2000)
11. Hicke L: Protein regulation by monoubiquitin. *Nat Rev Mol Cell Biol* 2, 195-201 (2001)
12. Pickart C. M: Ubiquitin in chains. *Trends Biochem Sci* 25: 544-548 (2000)
13. Chau V, J. W. Tobias, A. Bachmair, D. Marriott, D. J. Ecker, D. K. Gonda & A.Varshavsky: A multi-ubiquitin chain is confirmed to specific lysine in a targeted short-lived protein. *Science* 243, 1576-1583 (1989)
14. Hershko A & A. Ciechanover: The ubiquitin system. *Annu Rev Biochem* 67, 425-479 (1998)
15. Spence J, S. Sadis, A. L. Haas & D. Finley: A ubiquitin mutant with specific defects in DNA repair amd multiubiquitination. *Mol.Cell.Biol*. 15, 1265-1273 (1995)
16. Scheffner M, J. M. Huibregtse & P. M. Howley: Identification of a human ubiquitin-conjugating enzyme that mediates the E6-AP dependent ubiquitination of p53. *Proc.Natl.Acad.Sci.USA* 91, 8797-8801 (1994)
17. Spence, J R.R. Gali, G. Dittmar, F. Sherman, M. Karin & D. Finley: Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* 102, 67-76 (2000)
18. Feldman RM, C. C. Correll, K. B. Kaplan & R. J. Deshaies: A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p.*Cell* 91, 221-230 (1997)
19. Skowyra D, K. L. Craig, M. Tyres, S. J. Elledge & J. W. Harper: F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell* 91, 209-219 (1997)
20. DeSalle L. M. & M. Pagano: Regulation of the G1 to S transition by the ubiquitin pathway. *FEBS Lett* 490, 179-189 (2001)
21. Weissman A. M.: Themes and variations on ubiquitylation. Nat Rev Mol Cell Biol 2: 169-178, 2001.
22. Alves-Rodrigues A, L. Gregori & M. E. Figueiredo-Pereira: Ubiquitin, cellular inclusions and their role in neurodegeneration. *Trends Neurosci* 21, 516-520 (1998)
23. Vu P. K. & K. M. Sakamoto: Ubiquitin-mediated proteolysis and human disease. *Mol Genet Metab* 71, 261-266 (2000)
24. Wilkinson K. D. & M. Hochstrasser: The deubiquinating enzymes in ubiquitin. Eds: Peters J-M, Harris J.R, Finley D.J. Plenum Press, New York, 99-125 (1998)
25. Sommer, T & Jentsch, S. A protein translocation defect linked to ubiquitin conjugation at the endoplasmatic reticulum. Nature 365: 176-179 (1993)
26. Chen A, K. Wu, S. Y. Fuchs, P. Tan, C. Gomez & Z. Q. Pan. The conserved RING-H2 finger is required for ubiquitin ligation. *J.Biol.Chem*. 275, 15432-15439 (2000)
27. Chan N. L. & C. P. Hill: Defining polyubiquitin chain topology. *Nat Struct Biol* 8, 650-652 (2001)
28. Koegl M, T. Hoppe, S. Schlenker, H. D. Ulrich, T. U. Mayer & S. Jentsch: A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* 96, 635-644 (1999)
29. Lake M. W, M. M. Wuebbens, K. V. Rajagopalan, H. Schindelin: Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. *Nature* 414, 325-329 (2001)
30. Wilkinson K.D.: Ubiquitination and deubiquitination: targeting of proteins for degradation by the proteasome. *Semin Cell Dev Biol* 11, 141-148 (2000)
31. Meluh P. B. & D. Koshland: Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol.Biol.Cell*. 6, 793-807 (1995)
32. Mannen H, H. M. Tseng, C. L. Cho & S. S. Li: Cloning and expression of human homologue HSMT3 to yeast SMT3 suppressor of MIF2 mutations in a centromere protein gene. *Biochem.Biophys.Res.Commun*. 222, 178-180 (1996)
33. Chen A, H. Mannen & S. S. Li: Characterization of mouse ubiquitin-like SMT3A and SMT3B cDNAs and gene/pseudogenes. *Biochem.Mol.Biol.Int* 46, 1161-1174 (1998)
34. Boddy M. N, K. Howe, L. D. Etkin, E. Solomon & P. S. Freemont: PIC1, a novel ubiquitin-like protein which interacts with the PML component of a multiprotein complex that is disrupted in acute promyelocytic leukemia. *Oncogene* 13, 971-982 (1996)
35. Okura T, L. Gong, T. Kamitani, T. Wada, I. Okura, C. F. Wei, H. M. Chang HM. & E. T. Yeh. Protection against Fas/APO-- and tumor necrosis factor-mediated cell death by novel protein, sentrin. *J.Immunol.* 157, 4277-4281 (1996)
36. Matunis M. J, E. Coutavas & G. Blobel: A novel ubiquitin-like modification modulates the partitioning of

the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J.Cell.Biol.*135, 1457-1470 (1996)
37. Mahajan R, C. Delphin, T. Guan, L. Gerace & F. Melchior: A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* 88, 97-107 (1997)
39. Lapenta V, P. Chiurazzi, van der Spek, A. Pizzuti, F. Hanaoka & C. Brahe: SMT3A, a human homologue of the S. cerevisiae SMT3 gene, maps to chromosome 21qter and defines a novel gene family. *Genomics* 40, 362-366 (1997)
40. Shen Z, P.E. Pardington-Purtymun, J. C. Comeaux, R. K. Moyzis & D. J.Chen: UBL1, a human ubiquitin-like protein associating with human RAD51/RAD52 proteins. *Genomics* 36, 271-279 (1996)
41. Melchior F.: SUMO-nonclassical ubiquitin. *Annu Rev Cell Dev Biol* 16, 591-626, (2000)
42. Wilson V. G. & D. Rangasamy: Viral interaction with the host cell sumoylation system. *Virus Res* 81, 17-27 (2001)
43. Takahashi Y, M. Iwase, M. Konishi, M. Tanaka, A. Tohe & Y. Kikuchi: Smt3, a SUMO-1 homolog, is conjugated to Cdc3, a component of septin rings at the mother-bud neck in budding yeast. *Biochem.Biophys.Res.Com*. 259, 582-587 (1999)
44. Buschmann T, S. Y. Fuchs, C. G. Lee, Z. Q. Pan & Z. Ronai: SUMO-1 modification of Mdm2 prevents its self-ubiquitination and increases Mdm2 ability to ubiquitinate p53. *Cell* 101, 753-762 (2000)
45. Göttlicher M, S. Heck, V. Doucas, E. Wade, M. Kullmann, A. C. Cato, R. M. Evans & P. Herrlich: Interaction of the Ubc9 human homologue with c-Jun and with the glucocorticoid receptor. *Steroids* 61, 257-262 (1996)
46. Müller S, M. Berger, F. Lehembre , J. S. Seeler, Y. Haupt & A. Dejean: c-Jun and p53 activity is modulated by SUMO-1 modification. *J.Biol.Chem.* 275, 13321-13339 (2000)
47. Gostissa M, A. Hengstermann, V. Fogal, P. Sandy, S. E. Schwarz, M. Scheffner & G. Del Sal: Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1. *EMBO.J.* 18, 6455-6461 (1999)
48. Rodriguez M. S, J. M. Desterro, S. Lain, C. A. Midgley, D. P. Lane & R. T. Hay: SUMO-1 modification activates the transcriptional response of p53. *EMBO.J.* 18, 6455-6461 (1999)
49. Müller S, C. Hoege, G. Pyrowolakis & S. Jentsch: SUMO, ubiquitin's mysterious cousin. *Nat Rev Mol Cell Biol* 2, 202-210 (2001)
50. Mossessova E. & C. D. Lima: Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* 5, 865-876 (2000)
51. Johnson E. S, I. Schwienhorst, R. J. Dohmen & G. Blobel: The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J* 16, 5509-5519 (1997)
52. Sachdev S, L. Bruhn, H. Sieber, A. Pichler, F. Melchior & R. Grosschedl: PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. *Genes Dev* 15, 3088-3103 (2001)
53. Kumar S, Y. Yoshida & M. Noda: Cloning of a cDNA which encodes a novel ubiquitin-like protein. *Biochem.Biophys.Res.Com.* 195, 393-399, (1993)
54. Rao-Naik C, W. delaCruz, J. M. Laplaza, S. Tan, J. Callis & A. J. Fisher: The Rub family of ubiquitin-like proteins. *J.Biol.Chem*. 273, 34976-34982 (1998)
55. Hochstrasser M.: Ubiquitin-dependent protein degradation. *Annu Rev Genet* 30, 405-439 (1996)
56. Kamitani T, K. Kito, H. P. Nguyen & E. T. H. Yeh: Characterization of NEDD8, a developmentally down-regulated ubiquitin-like protein. *J.Biol.Chem*. 272, 28557-28562 (1997)
57. Lammer D, N. Mathias, J. M. Laplaza, W. Jiang, Y. Liu, J. Callis, M. Goebl & M. Estelle: Modification of yeast Cdc53p by the ubiquitin-related protein rub1p affects function of the SCFCdc4 complex. *Gen.Dev*. 12, 914-926 (1998)
58. Liakopoulos D, G. Doenges, K. Matuschewski, & S. Jentsch: A novel protein modification pathway related to the ubiquitin system. *EMBO J* 17, 2208-2214 (1998)
59. Hochstrasser M.: There's the rub: a novel ubiquitin-like modification linked to cell cycle regulation. *Genes Dev* 12, 901-907 (1998)
60. Kawakami T., T. Chiba, T. Suzuki, K. Iwai, K. Yamanaka, N. Minato, H. Suzuki, N. Shimbara, Y. Hidaka, F. Osaka, M. Omata & K. Tanaka: Nedd8 recruits E2-ubiquitin to SCF E3 ligase. *EMBO J.* 20, 4003-4012 (2001)
61. Hodgins R. R, K. S. Ellison & M. J. Ellison: Expression of a ubiquitin derivative that conjugates to protein irreversibly produces phenotypes consistent with a ubiquitin deficiency. *J Biol Chem* 267, 8807-8812 (1992)
62. Pickart C. M, E. M. Kasperek, R. Beal & A. Kim: Substrate properties of site-specific mutant ubiquitin protein (G76A) reveal unexpected mechanistic features of ubiquitin-activating enzyme (E1). *J Biol Chem* 269, 7115-7123 (1994)
63. Sloper-Mould K. E, J. C. Jemc, C. M. Pickart & L. Hicke: Distinct functional surface regions on ubiquitin. *J Biol Chem* 276, 30483-30489 (2001)
64. Burch T. J. & A. L. Haas: Site-directed mutagenesis of ubiquitin. Differential roles for arginine in the interaction with ubiquitin-activating enzyme. *Biochemistry* 33, 7300-7308 (1994)
65. Whitby F. G, G. Xia, C. M. Pickart & C. P. Hill: Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. *J Biol Chem* 273, 34983-34991 (1998)
66. Beal R, Q. Deveraux, G. Xia, M. Rechsteiner & C. Pickart: Surface hydrophobic residues of multiubiquitin chains essential for proteolytic targeting. *Proc Natl Acad Sci U S A* 93, 861-866 (1996)
67. Shih S. C, K. E. Sloper-Mould & L. Hicke: Monoubiquitin carries a novel internalization signal that is appended to activated receptors. *EMBO J* 19, 187-198 (2000)
68. Desterro J. M, M. S. Rodriguez & R. T. Hay: SUMO-1 modification of IκBα inhibits NF-κB activation. *Mol Cell* 2, 233-239 (1998)
69. Wu K, A. Chen, P. L. Tan & Z. Q. Pan: The Nedd8-conjugated ROC1-CUL1 core ubiquitin ligase utilizes Nedd8 charged surface residues for efficient polyubiquitin

chain assembly catalyzed by Cdc34. *J.Biol.Chem* 277, 516-527 (2002)

70. Orengo C. A, D. T. Jones & J. M. Thornton: Protein superfamilies and domain superfolds. *Nature* 372, 631-634 (1994)

71. Tsukihara T, K. Fukuyama, M. Mizushima, T. Harioka, M. Kusunoki, Y. Katsube, T. Hase & H. Matsubara: Structure of the (2Fe-2S) ferredoxin I from blue-green algae Aphanothece sacrum at 2.2A resolution. *J.Mol.Biol.* 216, 399-410 (1990)

72. Nassar N, G. Horn, C. Herrmann, A. Scherer, F. McCormick & A. Wittinghofer: The 2.2A crystal structure of the ras-binding domain of the serine/threonine kinasec-Raf1 in complex with Rap1A and a GTP-analogue. *Nature* 375, 554-560 (1995)

73. Geyer M, C. Herrmann, S. Wohlgemuth, A. Wittinghofer & H. R. Kalbitzer: Structure of the Ras-binding domain of RalGEF and implications for Ras-binding and signalling. *Nature Struc. Biol.* 4, 694-699 (1997)

74. Esser D, B. Bauer, R. M. F. Wolthuis, A. Wittinghofer, R. H. Cool & P. Bayer: Structure determination of the Ras-binding domain of the Ral-specific Guanine nucleotide exchange factor Rlf. *Biochemistry* 37, 13453-13462 (1998)

75. Bayer P, A. Arndt, S. Metzger, R. Mahajan, F. Melchior, R. Jaenicke & J. Becker: Structure determination of the small ubiquitin-related modifier SUMO-1. *J.Mol.Biol.* 280, 275-286 (1998)

76. Blom N., S. Gammeltoft & S. Brunak: Sequence- and Structure-Based Prediction of Eukaryotic Protein Phosphorylation Sites. *J. Mol. Biol.* 294, 1351-1362 (1999)

77. Jentsch S, J. P. McGrath & A. Varshavsky: The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating enzyme. *Nature* 329, 131-134 (1987)

78. Moraes T. F, R. A. Edwards, S. McKenna, L. Pastushok, W. Xiao, J. N. Glover & M. J. Ellison: Crystal structure of the human ubiquitin conjugating enzyme complex, hMms2-hUbc13. *Nat Struct Biol* 8, 669-673 (2001)

79. Deng L, C. Wang, E. Spencer, L. Yang, A. Braun, J. You, C. Slaughter, C. Pickart & Z. J. Chen: Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* 103, 351-361 (2000)

80. Huang L, E. Kinnucan, G. Wang, S. Beaudenon, P. M. Howley, J. M. Huibregtse & N. P. Pavletich: Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. *Science* 286, 1321-1326 (1999)

81. Watkins J. F, P. Sung, S. Prakash & L. Prakash: The extremely conserved amino terminus of RAD6 ubiquitin-conjugating enzyme is essential for amino-end rule-dependent protein degradation. *Genes Dev* 7, 250-261 (1993)

82. Hamilton K. S, M. J. Ellison & G. S. Shaw: Identification of the ubiquitin interfacial residues in a ubiquitin-E2 covalent complex. *J Biomol NMR* 18, 319-327 (2000)

83. Hamilton K. S, M. J. Ellison, K. R. Barber, R. S. Williams, J. T. Huzil, S. McKenna, C. Ptak, M. Glover M & G. S. Shaw: Structure of a conjugating enzyme-ubiquitin thiolester intermediate reveals a novel role for the ubiquitin tail. *Structure* 9, 897-904 (2001)

84. Miura T, W. Klaus, B. Gsell, C. Miyamoto & H. Senn: Characterization of the binding interface between ubiquitin and class I human ubiquitin-conjugating enzyme 2b by multidimensional heteronuclear NMR spectroscopy in solution. *J Mol Biol* 290, 213-228 (1999)

85. Liu Q, C. Jin, X. Liao, Z. Shen, D. J. Chen & Y. Chen: The binding interface between an E2 (UBC9) and a ubiquitin homologue (UBL1). *J Biol Chem* 274, 16979-16987 (1999)

86. Cook W. J, L.C. Jeffrey, M. L. Sullivan & R. D. Vierstra: Three-dimensional structure of a ubiquitin-conjugating enzyme (E2). *J Biol Chem* 267, 15116-15121 (1992)

87. Cook W. J, L. C. Jeffrey, Y. Xu & V. Chau: Tertiary structures of class I ubiquitin-conjugating enzymes are highly conserved: crystal structure of yeast Ubc4. *Biochemistry* 32, 13809-13817 (1993)

88. Cook W. J, P. D. Martin, B. F. Edwards, R. K. Yamazaki & V. Chau: Crystal structure of a class I ubiquitin conjugating enzyme (Ubc7) from Saccharomyces cerevisiae at 2.9 angstroms resolution. *Biochemistry* 36: 1621-1627 (1997)

89. Worthylake D. K, S. Prakash, L. Prakash & C. P. Hill: Crystal structure of the Saccharomyces cerevisiae ubiquitin-conjugating enzyme Rad6 at 2.6 A resolution. *J Biol Chem* 273, 6271-6276 (1998)

90. Giraud M. F, J. M. Desterro & J. H. Naismith: Structure of ubiquitin-conjugating enzyme 9 displays significant differences with other ubiquitin-conjugating enzymes which may reflect its specificity for sumo rather than ubiquitin. *Acta Crystallogr D Biol Crystallogr* 54 (Pt 5), 891-898 (1998)

91. Tong H, G. Hateboer, A. Perrakis, R. Bernards R & T. K. Sixma: Crystal structure of murine/human Ubc9 provides insight into the variability of the ubiquitin-conjugating system. *J Biol Chem* 272, 21381-21387 (1997)

Key Words: RUB, SUMO, Ubiquitin, E2, Protein, Structure, Function

Send correspondence to: Dr. Peter Bayer, Max-Planck Institute for Molecular Physiology, Head of Resaerch Group, Molecular and Structural Biophysics, Otto Hahn Straße 11, D-44227 Dortmund, Germany, Tel.: ++49-231-133-2222, Fax: ++49-231-133-2699, E-mail: peter.bayer@mpi-dortmund.mpg.de