Protein Science (1999), *8*:1268–1275. Cambridge University Press. Printed in the USA.

The Ubp6 family of deubiquitinating enzymes contains a ubiquitin-like domain: SUb

ALLISON M. WYNDHAM,[1] ROHAN T. BAKER,[1] AND GARETH CHELVANAYAGAM[2]

[1]Molecular Genetics Group, John Curtin School of Medical Research, Australian National University, Mills Rd., Acton, Canberra ACT 0200, Australia
[2]Human Genetics Group, John Curtin School of Medical Research, Australian National University, Mills Rd., Acton, Canberra ACT 0200, Australia

(RECEIVED November 20, 1998; ACCEPTED February 19, 1999)

Abstract

A sequence motif that is *S*imilar to *Ub*iquitin (SUb) has been identified in the *Saccharomyces cerevisiae* ubiquitin-specific protease Ubp6. SUb is conserved in all known Ubp6 homologues from a spectrum of eukaryotic species and is also present in a group of hypothetical proteins of unknown function (Unk1-3) present in sequence databases. An N-terminal deletion mutant of Ubp6 that lacks SUb is still capable of cleaving α-linked ubiquitin fusions, suggesting that SUb forms a separate domain to the catalytic core of Ubp6 and demonstrating that it is not required for in vitro cleavage activity. A homology model of the 78 N-terminal amino acids of human Ubp6, based on the known fold of ubiquitin, is presented. In human Ubp6, SUb shares only 20% sequence identity with ubiquitin. Even weaker similarity occurs between *S. cerevisiae* SUb and ubiquitin. The homology model supports a ubiquitin-like fold for SUb and suggests that two conserved Lys residues, corresponding to Lys48 and Lys63 of ubiquitin, are functionally important.

Keywords: deubiquitinating enzyme; SUb; Ubp; Ubp6; ubiquitin; ubiquitin-like protein

A growing number of proteins of diverse function have been shown to contain sequences similar to ubiquitin. The majority of these proteins possess an N-terminal ubiquitin-like region and an unrelated C-terminal tail of variable length. Ubiquitin-like proteins do not participate in proteolysis in the same manner as true ubiquitin, but instead appear to modulate the function of target proteins through direct protein–protein interactions. Ubiquitin-like proteins have been implicated in a multitude of roles, including DNA repair (Schauber et al., 1998), spindle pole body duplication (Biggins et al., 1996), lymphocyte differentiation (Haas & Siepmann, 1997), viral pathogenicity (Meyers et al., 1991), Hsp70 regulation (Matsuzawa et al., 1998), and transcriptional regulation (Garrett et al., 1995).

Sequence similarity between ubiquitin-like sequences and ubiquitin varies considerably, from 17% amino acid identity in yeast Smt3 (Meluh & Koshland, 1995; Johnson et al., 1997) to 76% identity in the baculovirus ubiquitin variant (Guarino, 1990). In comparison, ubiquitin itself is a highly conserved protein, exhibiting a minimum 96% sequence identity between a spectrum of eukaryotic species. Furthermore, the three-dimensional structures of human, plant, and yeast ubiquitin are almost identical (Vijay-Kumar et al., 1987b). Nevertheless, the ubiquitin fold is found in proteins with negligible sequence identity to ubiquitin and has been identified as one of nine classes of superfolds (Orengo et al., 1994). The fold is characterized by a mixed β-sheet with topology $-2+1+5-3+4$ onto which packs a regular helix, encompassing a hydrophobic core (Vijay-Kumar et al., 1987a). In addition, two single turns of 3/10 helix are found in the structure: prior to strand 3 and prior to strand 5. Structures have also been solved for a number of molecules that include a ubiquitin-like fold, including Raf (Emerson et al., 1995; Nassar et al., 1995), RalGDS (Geyer et al., 1997; Huang et al., 1997), Protein G (Achari et al., 1992; Gallagher et al., 1994), ferredoxin (Tsukihara et al., 1990), and SUMO-1 (Bayer et al., 1998). Further, based on sequence similarity, additional ubiquitin-like molecules have previously been identified that presumably also adopt a ubiquitin fold, including Fau (Kas et al., 1992), Rub-1 (Liakopoulos et al., 1998), Rad23 (Watkins et al., 1993), Bag-1 (Takayama et al., 1995), ISG15 (Haas et al., 1987; Narasimhan et al., 1996), and Dsk2p (Biggins et al., 1996). These molecules point to a diversity of functional roles for this molecular topology (Mayer et al., 1998).

The primary factor constraining sequence and structural variation in ubiquitin is thought to be its role in proteolysis. Ubiquitin is post-translationally conjugated to other proteins in the cell via an isopeptide linkage between the C-terminal Gly76 residue and the ϵ-amino group of a substrate lysine. Rounds of conjugation lead to the development of a homopolymeric ubiquitin chain on the surface of the target protein that is recognized by the 26S proteasome and results in the degradation of the target (Chau et al., 1989; Gregori et al., 1990; reviewed in Pickart, 1997). Ubiquitin-mediated

Reprint requests to: Gareth Chelvanayagam, Human Genetics Group, John Curtin School of Medical Research, Australian National University, Mills Rd., Acton, Canberra ACT 0200, Australia; e-mail: gareth.chelva@anu.edu.au.

proteolysis accounts for the degradation of the majority of short-lived proteins in the eukaryotic cell, including G1 and mitotic cyclins, oncoproteins, transcription factors, and p53, as well as structurally abnormal proteins (reviewed in Hochstrasser, 1996). A less well-understood role for ubiquitin is its conjugation either in single moieties or as polymeric chains to receptor subunits as part of receptor signaling and receptor-mediated endocytosis (reviewed in Hicke, 1997).

Ubiquitin-specific proteases (Ubps) are deubiquitinating enzymes that can cleave ubiquitin–protein peptide and isopeptide bonds and have been implicated in ubiquitin precursor processing, negative regulation of substrate targeting, and maintenance of the free ubiquitin pool (Tobias & Varshavsky, 1991; Baker et al., 1992; Papa & Hochstrasser, 1993; Amerik et al., 1997; reviewed in Hochstrasser, 1996; Wilkinson, 1997). Ubp6, one of a family of 16 Ubps present in *Saccharomyces cerevisiae*, is capable of cleaving ubiquitin from α-linked fusions in vitro; however, its role in vivo is not yet known (Park et al., 1997). Unexpectedly, the Ubp6 enzyme contains a ubiquitin-like sequence at its N-terminus. Here we show that the N-terminal motif, *S*imilar to *Ub*iquitin (SUb), is conserved throughout known Ubp6 homologues and also in a group of unrelated proteins of unknown function (Unk1-3). We have constructed a homology model of SUb from human Ubp6 (human SUb^{UBP6}) based on human ubiquitin. Potential roles for conserved residues and SUb are discussed. Activity of an *S. cerevisiae* Ubp6 N-terminal deletion mutant confirms that SUb forms a domain distinct from the catalytic core of Ubp6, and putative functions for these domains are presented. This is the first report of a ubiquitin-like region in an enzyme of the ubiquitin system itself.

Results

Database search results

BLAST searching of available databases using full-length *S. cerevisiae* Ubp6 identified weak similarity between the N-terminal region of Ubp6 and human ubiquitin. The search also identified other species homologues of Ubp6 present in humans, rabbit (*Oryctolagus cuniculus*), mouse (*Mus musculus*), *Drosophila melanogaster*, *Aspergillus nidulans*, *Caenorhabditis elegans*, and *Schizosaccharomyces pombe*. In addition, using *S. cerevisiae* SUb^{UBP6} as a search target, several putative proteins were identified that contained SUb, but were otherwise unrelated to Ubp6. We dubbed these hypothetical proteins Unk1-3 (*Unk*nown). Three sequences were identified of a hypothetical protein (Unk1) in mice, *D. melanogaster*, and the fish, *Fugu rubripes*. Additional partial sequences were identified in *Trypanasoma brucei* (Unk2), and *Arabidopsis thaliana* (Unk3). SUb^{UNK1} and SUb^{UNK2} are located at the N-terminus, as in Ubp6, while SUb^{UNK3} is located in the middle of a hypothetical protein. The N-terminal flank of SUb^{UNK3} has sequence similarity to a plant membrane-associated salt-induced protein, while the C-terminal flank is similar to Ras-binding proteins from several organisms (not shown). It is interesting to note that the Ras-binding protein RalGDS also adopts a ubiquitin fold (Geyer et al., 1997; Huang et al., 1997); however, SUb^{UNK3} and the putative Ras-binding motif are separate in the Unk3 hypothetical protein. The Unk2 protein may be the *T. brucei* Ubp6 homologue; however, in the absence of more sequence, this remains to be determined. Anomalous residues derived from EST sequences and their likelihood of error are given in Table 1.

Table 1. *Sequences derived from EST/genomic sequences*

Sequence	Acc. Nos.	Residues[a]	Conflicting residues[b]
SUb^{UBP6}	AA10665	3–87	Nil
M. mus	AA163116	3–87	Nil
	AI153244	5–87	N66 = D
	AA510768	3–87	N66 = D, K68 = I, K70 = N
	AA170034	3–87	2 FS
	AA672946	3–56	Nil
	AA086703	5–87	2 FS
	AI007125	3–40	1 FS
SUb^{UBP6}	AA695753	2–87	Nil
D. mel	AA695970	2–87	Nil
	AA949051	2–87	Nil
	AA942181	2–87	1 FS
SUb^{UBP6}	D22465	22–87	Nil (one EST only)
O. sat			
SUb^{UBP6}	Z25610	12–87	Nil (one EST only)
A. tha			
SUb^{UBP6}	AA784983	3–87	3–87 Nil (one EST only)
A. nid			
SUb^{UBP6}	Z81317	5–87	Nil (one cosmid only)
S. pom			
SUb^{UBP6}	D44559	1–87	Nil (genomic clone)
S. cer	T38669	8–87	Nil (EST)
SUb^{UNK1}	AA817183	1–87	Nil
D. mel	AA940861	1–87	Nil
	AA390894	1–87	Nil
	AA735553	1–87	Nil
	AA246382	1–87	Nil
	AI064082	1–87	Nil
SUb^{UNK1}	C89508	4–87	Nil (one EST only). All other SUbs have Q46; E46 may be error
M. mus			
SUb^{UNK1}	Z89735	4–55	Nil (cosmid; intron at codon 55)
F. rub	Z89745	4–55	Nil (cosmid; intron at codon 55)
SUb^{UNK2}	W06549	8–87	g inserted at bp 15 to obtain aa 8–12 in frame with long ORF. All other SUbs have G55; Q55 may be error
T. bru[c]			
SUb^{UNK3}	AC004165	2–87	Nil (cosmid sequence)
A. tha	AA713002	2–70	1 FS, K7 = N, Q66 = H, plus five uncertain codons due to unknown bases (EST)

[a] Numbered from the alignment in Figure 1.

[b] Discrepancies between consensus sequence in Figure 1 and ESTs/cosmids, and/or alterations made to ESTs. FS = frame shift.

[c] The SUb^{UNK2} sequence from *T. brucei* (one EST only) could represent a partial bona fide Ubp6 sequence (see Fig. 1).

Sequence alignment

Figure 1 shows an alignment of members of the SUb family against the human ubiquitin sequence. Similarity extends over the length of the ubiquitin sequence. Sequence identity between human ubiquitin and SUb ranges from 16% (*S. cerevisiae* Ubp6) to 31% (Unk3), with the majority of sequences sharing 20% or lower. This limited sequence identity provides a weak signal suggesting that SUb adopts a ubiquitin fold. In general, SUb is as distantly related

```
                1                                                                                  87
                .........+.........+.........+.........+.........+.........+.........+.........+.......
                      EEEEEE     EEEEE       HHHHHHHHHHHH   GGGEEEEE      EE        GGG       EEEEEE
Ub_H.sap            MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIF---AGKQLEDGRTLSDYNIQKESTLHLVLRLRGG---
SUbUBP6_H.sap    MPLYSVTVKWGKEKFEGVELNTDEPPMVFKAQLFALTGVQPARQKVMV---KGGTLKDD-DWGNIKIKNGMTLLMMGSADALPEE
SUbUBP6_M.mus    MPLYSVTVKWGKEKFEGVELNTDEPPMVFKAQLFALTGVQPARQKVMV---KGGTLKDD-DWGNIKMKNGMTVLMMGSADALPEE
SUbUBP6_O.cun    MPLYSVTVKWGKEKFGGVELNTDEPPMVFKAQLFALTGVQPARQRVMV---KGGTLKDD-DWGNIKIKNGMTILMMGSADALPEE
SUbUBP6_D.mel    MPAFKVKVKWGRELYTDIVVNTDEEPILFKAQLFALTGVQPDRQKVMC---KGGILKDD-QW-NLQIKDGAVVLLLGSKESVPEV
SUbUBP6_C.ele      MPIVNVKWQKEKYV-VEVDTSAPPMVFKAQLFALTQVVPERQKVVI---MGRTLGDD-DWEGITIKENMTIMMMGSVGEIPKP
SUbUBP6_O.sat                      •IDTSQPPVVFKTQLYTLTGVPPERQKIMV---KGGILKDDADWSTLGVKDGQKLMMIGTADEIVKA
SUbUBP6_A.tha             •QKKVLDGIEIDVSLPPYVFKAQLYDLTGVPPERQKIMV---KGGLLKDDGDWAAIGVKDGQKLMMMGTADEIVKA
SUbUBP6_A.nid    MASIPVIVKHQGKRYD-VELDPNSTGETFKYQLYSLTGVEPERQKILV---KGGQLKNDTLLSTINAKPNQTFMMMGTPSGDQGA
SUbUBP6_S.pom      MIPIAIRWQGKKYD-LEIEPNETGSTLKHQLYSLTQVPPERQKVIV---KGGQLKDDVLLGSVGIKPNATLLMMGTAGELPTA
SUbUBP6_S.cer MSGETFEFNIRHSGKVYP-ITLSTDATSADLKSKAEELTQVPSARQKYMV---KGGLSGEESIKIYPLIKPGSTVMLLGTPDANLIS
SUbUNK3_A.tha•MADSTIKLTVKFGGKSIP-LSVSPDCTVKDLKSQLQPITNVLPRGQKLIF---KGKVLVETSTLKQSDVGSGAKLMLMASQGLHQGE
SUbUNK2_T.bru       •VKVKWGRETFE-LTVDLRSTVKCFKEQLQQLTSVPVERQKIMGVKASQCNDNEVVTLEAAGVRAGKTLMLIGTAAEVVRA
SUbUNK1_D.mel MEVKEVVVIVKWSGKEYP-VDLTDQDTVEVLRHEIFRKTQVRPERQKLLNLKYKGKTAADNVKISALELKPNFKLMMVGSTEADIED
SUbUNK1_M.mus    MALPIIVKWGGQEYSVTTLSEDDTVLDLKQFLKTLTGVLPEREKLLGLKVKGKPAENDVKLGALKLKPNTKIMMMGTREESLED
SUbUNK1_F.rub    MSVSVIIKWGGQEYSISSLSEEDTVMDLKQSIKSLTGVLPERQKLLGLKVKG
                      +  +          ++ +     +*  +  ** * * ***++    **    +         +*    + ++*+
```

Fig. 1. Sequence alignment of SUb domains. Sequences were aligned using CLUSTAL W and manually adjusted as required (see text). Dashes represent gaps introduced to maximize alignment. Predicted secondary structures are indicated above the alignment (E = β-strand; H = α-helix; G = 3/10 helix). Dots represent missing N-terminal sequence due to partial clones. Underlined residues contain an intron in their codons. The *F. rubripes* UNK1 sequence terminates at a putative intron position in its cosmid. Asterisks (*) indicate positions that vary in at most one or two sequences, and crosses (+) indicate conservative replacements. The SUb^{UNK2} sequence from *T. brucei* (one EST only) could represent a partial bona fide Ubp6 sequence if two frame shifts and one sequence error were allowed near the 3′ end of the sequence (outside region shown in Fig. 1) to produce a Ubp6-like conserved Cys domain (Baker et al., 1992). See Table 1 for generation of sequences from EST clones. Accession numbers: *Homo sapiens* ubiquitin: X56997. SUb^{UBP6}: *H. sapiens*: U30888; *O. cuniculus*: L37420; *C. elegans*: U32223. (See Table 1 for other EST/genomic sequence accession numbers.)

to any known ubiquitin-like protein as it is to ubiquitin itself (not shown). Although there are no completely conserved positions in the alignment, roughly a third of the positions show conservative changes or differences in only one or two of the 15 SUb sequences. Lys32, Pro43, Gln46, and Gly55 appear common to both ubiquitin and SUb, while Leu35, Thr39, Val41, Arg45, Lys47, Lys54, Lys70, and Gly79 seem SUb-specific, thereby distinguishing this motif from other ubiquitin-like proteins. Residues are numbered according to their position in the alignment in Figure 1. Only Thr39 and Val41 are completely conserved in all SUb sequences, although sequencing errors may mask other positions (Table 1).

All gaps in the alignment coincide with loop regions of the ubiquitin fold, with the exception of the relative deletions found in strand 2, and in the *D. melanogaster* Ubp6 that disrupts an isolated turn of 3/10 helix in the loop between strands 4 and 5. Because the latter gap is present in four separate EST entries, it is not likely to be a sequencing error, although it is possible that sequencing is difficult in this part of the molecule. Placement of a relative gap in this vicinity requires that the helical turn partially unwind, which, providing Trp64 remains buried, will not affect the core of the model structure.

The pattern of hydrophobicity is not well maintained in strands 2 and 4. Notably, strand 2 in several sequences contains a gap at position 18, and a Pro at position 19, suggesting that this region may not form a regular β-strand. As strands 2 and 4 occur as the edge strands of the 5-stranded mixed β-sheet in the ubiquitin fold, some irregularities in these strands can be expected.

Comparing SUb^{UBP6} to SUb from Unk1-3 shows that SUb^{UNK1} and SUb^{UNK2} contain a longer loop between strands 3 and 4. Amino acids at alignment positions 9, 24, 28, 58, 62, and 66 that tend to differ between Ubp6 and the other enzymes generally cluster in this region of SUb. Any differential function of SUb in Unk1 and Unk2 compared to SUb^{UBP6} would then be expected to be largely governed by residues in and around this loop.

Model structure

A computer model for human SUb^{UBP6} is illustrated in Figure 2, highlighting the relative positions of conserved residues. The overall geometry of the model is reasonable, and there are no serious steric violations, according to the output of PROCHECK (Laskowski et al., 1993). Further, bond lengths and angles comply with those in the template structure, and none of the residues have disallowed backbone torsion angles. The model shows a compact hydrophobic core formed by Val8, Val10, Trp12, Val20, Leu22, Phe31, Leu35,

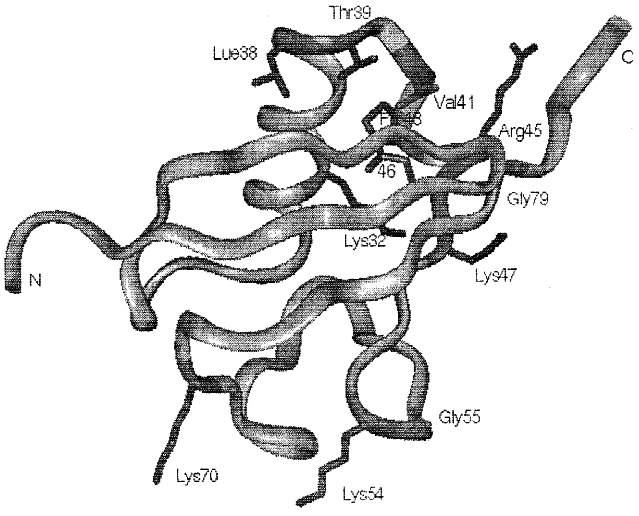

Fig. 2. Schematic diagram of the putative SUb fold. The relative locations of the more conserved residues are shown. Of particular note are Lys54 and Lys70, which are located in a similar place to Lys48 and Lys63, respectively, in ubiquitin, and Thr39, which may be phosphorylated.

Leu38, Val41, Val48, Val50, Leu58, Trp64, Ile67, Ile69, Leu75, and Met77. In some SUb sequences, residues at position 12, 20, 50, 58, 64, and 67 can be less hydrophobic. These positions, however, all lie close to the surface of the molecule where such substitutions are readily accommodated. Besides this core, two hydrophobic patches are evident on the surface of the human SUb^{UBP6} model. Located on opposite faces of the model, the larger of the two patches (Pro25, Met29, Val30, Ala33, Phe36, Ala37, Pro41, Ala44) is not conserved in either of the yeast Ubp6 enzymes or Unk3. The second hydrophobic patch (Gly11, Met49, Gly55, Gly56, Met73, Leu76, Met78) is formed mostly by the surface of strands 3 and 5, and is present in all SUb domains. Neither patch is found in ubiquitin, but it is noted that the smaller patch forms a surface adjacent to where a ubiquitin molecule might bind to Lys54 (analogous to Lys48 of ubiquitin) if it were presenting its C-terminal Gly, according to the tetra-ubiquitin (1TBE) structure.

The model also contains a number of potential covariations (Chelvanayagam et al., 1997). For example, from the alignment in Figure 1, it is noted that Trp64 usually occurs in conjunction with Pro27 and always with Pro28. In the SUb model, the Pro residues provide a tight start to the helix and allow the bulky side chain of Trp64 to pack under them. When a smaller residue is positioned at 64, more space is provided for the backbone leading into the helix. In *S. cerevisiae* SUb^{UBP6}, position 64 is Lys, and while this initially appears as a caveat to the above, further investigation of this sequence in the context of the ubiquitin fold shows that Tyr48 and Ser58, both polar residues, also occur nearby, providing suitable hydrogen bonding acceptors for Lys. Further, the increase in side-chain volume when introducing Tyr into the structure (Val/Leu/Ile $\rightarrow$ Tyr48) is concomitant with a reduction in side-chain volume at other internal neighboring positions, such as Leu $\rightarrow$ Ala35 and Phe $\rightarrow$ Leu31. Another potential covariation involves residues Thr9 and Phe17. These residues occur adjacent to one another in strands 1 and 2, respectively, in the model. Surprisingly, in the ubiquitin template, these residues are also present but are contributed by the opposite β-strands. This trend is also observed in the other mammalian Ubp6 sequences and fish SUb^{UNK1}.

Conserved residues

A total of 12 positions are highly conserved among SUb sequences, with residues varying in fewer than three sequences. Four of these positions are also conserved in ubiquitin, three of which (Lys32, Pro43, and Gln46) cluster tightly together and form a ubiquitin signature. Figure 3 shows how these residues pack together in a slight depression rendering Gln46 and, to a lesser extent, Lys32 buried. Lys32 is able to make strong hydrogen bonds with the backbone oxygen of residues Pro43 and Gln46, as well as salt link with the negatively charged residue that is almost always present at position 60. Likewise, Gln46 is able to form hydrogen bonds with the backbone oxygen of Lys32 and Val41. However, the role of the partial negative of the Oϵ1 oxygen on Gln46 is not obvious. Although not vital to the fold (e.g., SUMO-1), these residues appear able to form a stable substructure within the ubiquitin fold.

The majority of SUb conserved residues are located close to the ubiquitin signature residues. Leu38 and Thr39 are the most remote of this large cluster of conserved residues, and it is noted that the backbone ϕ angle of Thr39 is unusual for a residue in a helix, as is the equivalent residue in ubiquitin. Although Thr39, located on the surface of the model, can salt link with Lys16 and Glu18 in the human SUb^{UBP6} model, these residues are not conserved and, indeed, the loop/strand region in which they reside appears structurally diverse because a gap appears in some sequences in addition to the lack of a clear pattern of hydrophobicity for the loop. Arg45 is also surface exposed in the same vicinity as Thr39, but is about 10 Å away. Gly79 occurs at the end of the SUb domain at the end of strand 5 and may function to link SUb to the C-terminal domain.

Lys54, Gly55, and Lys70 are located at the other end of the model. Interestingly, and although not obvious from the alignment, both of the Lys residues are situated in similar spatial positions to ubiquitin residues Lys48 and Lys63 (numbered according to ubiquitin sequence). SUb^{UNK1} appears to have an extended loop between strands 3 and 4, potentially reorienting Lys54. However, it

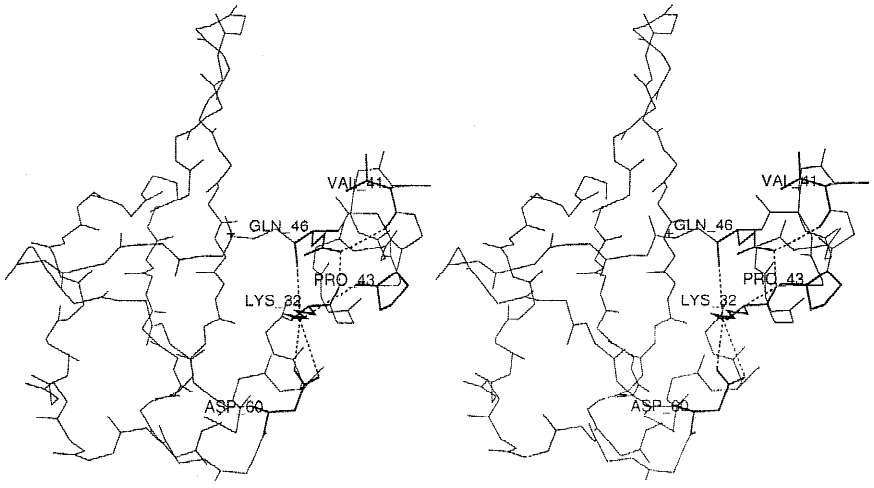

Fig. 3. Stereo pair: human SUb^{UBP6} model. Shows the network of interactions between the semi-buried residues, Lys32, Val41, Pro43, Gln46, and Asp60 that help to stabilize the structure.

is noted that these sequences also have Lys residues at positions 52 and 60, the latter being likely to superimpose perfectly with Lys48 of ubiquitin.

Automated predictions

In conjunction with the modeling done here, automated computer programs were also used to predict a structure for SUb. When tested with the human SUb^{UBP6} sequence, the program THREADER (Jones et al., 1992) suggested that the structure of human ubiquitin provided the most likely fold for the sequence of SUb among known protein structures. However, a better match was scored with the glycosidase inhibitor (1HOE) structure when the *S. cerevisiae* SUb^{UBP6} sequence was tested, although human ubiquitin still was highly ranked. Also, the secondary structure prediction program PHD (Rost & Sander, 1994; Rost et al., 1994), when given the human SUb sequence, predicted strands 1, 3, and 5 and the helix essentially as indicated in Figure 1. Strands 2 and 4 were not predicted, nor was the first 3/10 helix. The second 3/10 helix was predicted to be a strand. These predictions cannot confirm the model, but do show general support for the predicted structure.

Deubiquitinating activity of a Ubp6Δ2-79 mutant

To test the role of SUb^{UBP6} in the deubiquitinating activity of Ubp6, a deletion mutant (Ubp6Δ2-79) was constructed, where the N-terminal 79 residues of Ubp6 were removed, and a new initiation codon inserted upstream of Asn80 (see Methods). Recombinant full-length Ubp6 and recombinant Ubp6Δ2-79 were equally active in cleaving an artificial linear ubiquitin-GSTP1 fusion protein (Fig. 4). Thus, the SUb domain of Ubp6 is dispensable for its deubiquitinating activity in this in vitro assay. We conclude that the "catalytic core" of Ubp6 can fold in the absence of the SUb domain and form an active Ubp enzyme, consistent with SUb forming an independent structural domain.

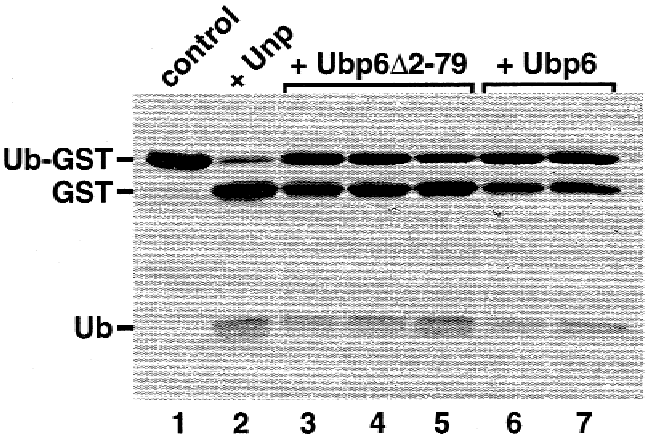

Fig. 4. The SUb domain of Ubp6 is not required for activity in vitro. Extracts of *E. coli* expressing either: no Ubp (lane 1); mouse Ubp Unp (Gilchrist et al., 1997; lane 2); Ubp6Δ2-79 (three independent clones; lanes 3–5); or full-length Ubp6 (two independent clones; lanes 7–8) were incubated at 37 °C for 1 h with purified, metabolically ^{35}S-Met-labeled Ub-GSTP1. The products were resolved by 10% SDS-PAGE and fluorographed (see Materials and methods). Bands, corresponding to Ub-GSTP1, GSTP1, and Ub, are indicated on the left.

Gene structure

Sequence database entries have allowed some comparisons of SUb gene structure and localization. For example, the human *UBP6* gene maps to chromosome 18p (Acc. U30888). The fish SUb^{UNK1} protein, derived from a cosmid/genome survey sequence, is apparently interrupted by an intron within codon 55, where a putative splice donor sequence marks a break in similarity with other SUb sequences (Fig. 1). Notably, the *A. thaliana* SUb^{UNK3} domain is interrupted by an intron at exactly the same base pair, consistent with this being a genuine intron in the fish protein. We could not identify the remainder of the SUb^{UNK1} protein in the available *F. rubripes* cosmid sequence. Of the two intron-containing genes where the complete exon/intron structure can be deduced from cosmid sequences, both have an intron near the end of SUb: within codon Ser80 in *C. elegans* SUb^{UBP6} and within codon Glu87 in *A. thaliana* SUb^{UNK3} (Fig. 1). This is consistent with SUb adopting a separate structural domain and also with the evolution of these genes by exon shuffling. We could not identify a convincing homologue with known function of either Unk1 or Unk3 in available databases.

Discussion

Ubiquitin-like proteins identified so far (Mayer et al., 1998) are generally of two types—those that possess the C-terminal Gly-Gly peptide and are cleaved/conjugated, and those that do not. None of the SUb family contain the Gly-Gly motif, consistent with the second type of ubiquitin-like protein. The level of sequence identity between the SUb regions and ubiquitin varies from 16 to 31%. A relationship between sequence identity and structural similarity has previously been described (Chothia & Lesk, 1986; Sander & Schneider, 1991; Chelvanayagam et al., 1994; Holm, 1998), and it is generally recognized that proteins with greater than 25% sequence identity share the same fold, as seen with the recently solved ubiquitin-like structures for Rub1/NEDD8 (Rao-Naik et al., 1998; Whitby et al., 1998). The length over which the identity is calculated is an important factor, and while SUb is only about 80 amino acids long, if it does not adopt a ubiquitin fold, then this example redefines the bounds on the "twilight zone" (Doolittle, 1986) within which proteins with different sequences may have the same fold, and conversely, proteins with limited similarity may adopt different folds. Nevertheless, other ubiquitin-like sequences exhibit widely differing levels of sequence identity with ubiquitin, such as Rad23 (22%), Dsk2 (36%), Rub1 (53%), and the baculoviral ubiquitin variant (76%) (Guarino, 1990; Watkins et al., 1993; Biggins et al., 1996; Liakopoulos et al., 1998). Further, the NMR structure of the ubiquitin-like protein SUMO-1 (mammalian homologue of the *S. cerevisiae* Smt3 protein) was recently solved and found to be almost identical to ubiquitin, despite SUMO-1 displaying only 18% amino acid identity with ubiquitin (Bayer et al., 1998). Therefore, the relatively low level of sequence similarity between the SUb proteins and ubiquitin need not necessarily reflect on the likelihood of SUb adopting a ubiquitin fold.

The ubiquitin-like folds that occur in the Ras-binding domains (RBD) of Raf and Ral, the B1 immunoglobulin binding domain of streptococcal protein G and [2Fe-2S] ferredoxin have evolved to interact with other proteins in ways that differ to ubiquitin. In both the Rap-Raf RBD complex (Nassar et al., 1995) and the streptococcal protein G-IgG complex (Achari et al., 1992; Gallagher

et al., 1994), an interprotein antiparallel β-sheet is formed, joining strand 2 of the ubiquitin fold with a strand in the other molecule. Figure 1 suggests that strand 2 in SUb is one of the least conserved parts of the molecule and that the loop connecting strands 1 and 2 is variable in length. Thus, even though strand 2 lies adjacent to the highly conserved C-terminal region of the helix, this is probably not a normal interaction site for SUb. The [2Fe-2S] ferredoxin domain in the oxidase-related aldehyde oxido-reductase from *Desulfovibrio gigas* (Archer et al., 1995) packs against three other domains, the ubiquitin-like helix and the loop before strand 3 forming most of the interface to the other domains. Surprisingly, few of the interface residues are hydrophobic. In the SUb model, the helix/strand 3 loop is much shorter than in the ferredoxin fold; thus it is possible that this region of the molecule is involved in packing against other domains within Ubp6 and Unk molecules.

The Thr39 residue that is absolutely conserved in the SUb family is conspicuously located on the surface of the model. Phosphorylation as a prerequisite to ubiquitination has been described for several substrates of the SCF ubiquitin protein ligase and also the IκBα kinase, the inhibitor of NFκB (Chen et al., 1996; Skowyra et al., 1997). Whether or not this property may be shared by ubiquitin-like proteins is unknown. However, the sequence surrounding Thr39 does not contain any known phosphorylation motifs (Kishimoto et al., 1985; Woodgett et al., 1986; Kreegipuu et al., 1998), although this does not rule out phosphorylation at this residue. Interestingly, the RNA splicing factor S114 (Kramer et al., 1995) and several other putative ubiquitin-like molecules have a Thr at this position.

The identification of a ubiquitin fold in an enzyme of the ubiquitin pathway is unexpected. However, there are several ways in which SUb could contribute to the function of the Ubp6 enzyme. Ubp sequences contain conserved domains that include candidate thiol protease catalytic triad residues (Baker et al., 1992; Wilkinson, 1997). There is little similarity between family members apart from these domains, and the contribution of the nonconserved sequence regions in Ubp enzymes is largely unknown. Therefore, SUb^{UBP6} represents one of only a few sequence elements so far identified in the Ubp family that is both distinct from the domains common to all Ubps and also found in other proteins. It is likely that SUb^{UBP6} exerts a function distinct from the proteolytic activity already shown for Ubp6 and that this involves the formation of a domain separate from the remainder of the enzyme. SUb^{UBP6} is not required for the Ubp activity of Ubp6, as supported by in vitro assays that show no obvious difference in Ubp activity between Ubp6 and Ubp6Δ2-79 (Fig. 4). This is further supported by the absence of any ubiquitin-like sequences in other Ubps characterized to date. Conversely, Ubp6Δ2-79 fails to rescue an *S. cerevisiae ubp6* null phenotype (R.T. Baker, unpubl. data), suggesting that Ubp6 function is mediated by SUb^{UBP6} in vivo.

It has been shown that recombinant *S. cerevisisae* Ubp6 can cleave α-linked ubiquitin fusions in vitro; however, its natural substrates remain unidentified (Park et al., 1997). Ubp6 must contribute to ubiquitin–protein degradation in yeast given that a haploid *ubp6* null strain exhibits a proteolytic defect as detected by two tests: the increased stability of a model substrate (ubiquitin–Pro-β-galactosidase) and sensitivity to the arginine analogue canavanine (A.M. Wyndham & R.T. Baker, unpubl. obs.). SUb^{UBP6} could conceivably promote the association of Ubp6 with ubiquitin-binding proteins and therefore potential substrates (including ubiquitin itself). Alternatively, SUb^{UBP6} may auto-inhibit Ubp6 by blocking its own ubiquitin binding site until displaced by the physiological substrate. The presence of a hydrophobic surface on the majority of SUb proteins is consistent with a possible role for SUb in associating with other proteins; alternatively, a hydrophobic patch may be required for interaction between SUb and the C-terminal portion of the Ubp6 or Unk proteins.

Given the role of other ubiquitin-like motifs in protein–protein interactions, it is tempting to speculate that SUb^{UBP6} contributes in some way to Ubp6 substrate recognition. A particularly interesting model is illustrated by the Rad23 protein, which binds to the 26S proteasome via a ubiquitin-like motif (Schauber et al., 1998), enabling the Rad4 excision repair enzyme, also bound to Rad23, to carry out its role in DNA repair. Therefore, Ubp6 may target substrates to the proteasome in a similar manner, a possibility that is currently being explored. The discovery of the rabbit Ubp6 homologue suggests another model. The rabbit Ubp6 homologue was first isolated from ribosomal fractions that exhibited a tRNA transglycosylase (TGT) activity (Deshpande et al., 1996). The TGT activity that was then ascribed to the rabbit protein has not been verified for it or any of the other Ubp6 homologues. Nevertheless, this led us to speculate that Ubp6 may be associated with ribosomes and could carry out processing of ubiquitin precursor proteins, a process that is thought to occur cotranslationally (Baker et al., 1994). An association of this type would also allow Ubp6 to detect "uncleavable" ubiquitin fusions as they emerged from the ribosome and target them for degradation, consistent with the *ubp6* null phenotype.

The conservation of the Lys54 and Lys70 residues and their position in the model at sites that match those for Lys48 and Lys63 of true ubiquitin has led us to speculate that one or both of these residues may be the site of multi-ubiquitin chain attachment. Lys48 in ubiquitin is the residue of primary importance for multi-ubiquitin chain formation and consequent substrate degradation by the proteasome. Lys63-linked chains are less common, but can still function to target a substrate for degradation, and this type of chain has been implicated in stress response, DNA repair, and ligand-dependent receptor internalization in yeast, and degradation of "uncleavable" ubiquitin fusions (Arnason & Ellison, 1994; Johnson et al., 1995; Galan & Hagenauer-Tsapis, 1997). Therefore, the Lys54 and Lys70 residues of SUb^{UBP6} may represent a means through which the Ubp6 protein is degraded. Interestingly, multi-ubiquitination of the N-terminal moiety is required for the degradation of "uncleavable" ubiquitin fusions such as ubiquitin-Pro-β-galactosidase and ubiquitin^{V76}-V-β-galactosidase (Johnson et al., 1995). It will be interesting to determine whether SUb, as an uncleavable ubiquitin analogue, enables Ubp6 to be targeted by a similar pathway.

Methods

Sequences and alignments

Sequence databases were searched for matches with the BLAST tool (Altschul et al., 1997; http://www.ncbi.nlm.nih.gov) using either the full length *S. cerevisiae* Ubp6 sequence or the N-terminal region (SUb^{UBP6}) as a target. Although the approach of using Expressed Sequence Tag (EST) databases can be of use in identifing new gene products (Board et al., 1997), caution must be applied so as to try to discriminate between errors and genuine sequence variation. As ESTs exhibit approximately 3% base ambiguity (Boguski et al., 1993), the quality of the sequence data in the databases is a key issue when it comes to identifying important

amino acids in the protein. To maximize accuracy, overlapping EST fragments of the same putative protein were aligned and a majority rules approach was used to designate the amino acid at any position. For those hypothetical proteins derived from a single EST (see Table 1), particular caution must be applied. Those sequences identified as containing a SUb domain were extracted and SUb regions aligned with ubiquitin in Figure 1, using the program CLUSTAL W (Thompson et al., 1994) with default settings. Subsequent minor manual adjustments were made to the alignment as suggested by computer modeling of the structure.

Computer modeling

Working from the original CLUSTAL W output, an initial model of human SUb was constructed with the Insight software package (MSI/Biosym, San Diego, California), using the human structure of ubiquitin as a template (PDB identifier 1UBQ). Starting coordinates were generated with the HOMOLOGY and BUILDER programs, which were subsequently refined with repeated cycles of energy minimization and molecular dynamics using the DISCOVER program, relaxing hydrogen atoms first before allowing side chains to move. Refinement was done in vacuum, but with a distance dependent dielectric constant, using charges but no cross or Morse terms. Due to the relative deletion in the target sequence at position 62, a random loop was inserted using Asp61 and Asn66 as anchor points. Out of the pool of loops generated, a loop with low deviation from expected backbone angles at the splice junctions and which buried Trp64 was selected. Only short minimization runs were required for refinement (200 cycles of steepest decent followed by 200 cycles of conjugate gradients). Likewise, short dynamics trajectories were calculated using 1,000 steps of equilibration followed by 5,000 steps of dynamics at each of 600, 400, and 300 K. Dynamics was only applied to the four residues at each terminus and to the residues in the loop.

To assess the quality of the model, PROCHECK (Laskowski et al., 1993) was used to check the overall geometry of the model. The burial of hydrophobic residues and exposure of hydrophilic residues were determined by visual inspection of the final model. Potential covariations were detected by computer software (G. Chelvanayagam, unpubl. obs.). The automated structure prediction computer programs THREADER (Jones et al., 1992) and PHD (Rost et al., 1994; Rost & Sander, 1994) were also used to compare results.

Construction and assay of recombinant Ubp6 proteins

The full Ubp6 open reading frame was amplified with primers UBP65 (5′ d[GCGAATTCAATATGAGCGGAGAAAAG]) and UBP63 (5′ d[GAGGATCCGACTTACAGACCAAATCC]) that add an *Eco*RI site upstream of the ATG codon, and a *Bam*HI site downstream of the stop codon, respectively. The catalytic domain was amplified with UBP6N80 (5′ d[GGGAATTCAATATGAAC CTGATTTCTAAACC]) and UBP63; the former primer adds an ATG codon and an *Eco*RI site immediately upstream of codon Asn80 of Ubp6. These primer pairs were used in conjunction with a clone containing the complete *UBP6* gene (Park et al., 1997) and a proofreading thermostable DNA polymerase (*Pfu*-Turbo, Clontech, Palo Alto, California) in the polymerase chain reaction using conditions recommended by the manufacturer. Amplification products were digested with *Eco*RI and *Bam*HI, subcloned into pUC18 and sequenced (Thermosequenase, Amersham, Little Chalfont, United Kingdom) in full to reveal no amplification-induced errors, and then re-cloned into the pKK223-3-based *Escherichia coli* expression vector pKK261 (Baker et al., 1994). Expression was induced in logarithmically growing cultures with 1 mM IPTG for 2.5 h, extracts were prepared by a lysozyme/sonication procedure (Baker et al., 1994), and insoluble material removed by centrifugation. Equal volumes of lysates were incubated with a purified, metabolically labeled ^{35}S-Ub-GSTP1 fusion protein (Baker et al., 1994), incubated for 1 h at 37 °C, resolved by 10% SDS-PAGE, and fluorographed (EnHance, NEN, Boston, Massachusetts).

Acknowledgments

This work was supported in part by a Postdoctoral Research Fellowship (to G.C.) and a Queen Elizabeth II Fellowship (to R.T.B.) from the Australian Research Council.

References

Achari A, Hale SP, Howard AJ, Clore GM, Gronenborn AM, Hardmann KD, Witlo M. 1992. 1.67 Å structure of the B2 immunoglobulin-binding domain of the streptococcal protein G and comparison to the NMR structure of the B1 domain. *Biochemistry 31*:10449–10457.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res 25*:3389–3402.

Amerik AY, Swaminathan S, Krantz BA, Wilkinson KD, Hochstrasser M. 1997. In vivo assembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome. *EMBO J 16*:4826–4838.

Archer M, Huber R, Tavares P, Moura I, Moura JJ, Carrondo MA, Sieker LC, LeGall J, Romao MJ. 1995. Crystal structure of desulforedoxin from *Desulfovibrio gigas* determined at 1.8 Å resolution: A novel non-heme iron protein structure. *J Mol Biol 251*:690–702.

Arnason T, Ellison MJ. 1994. Stress resistance in *Saccharomyces cerevisiae* is strongly correlated with assembly of a novel type of multi-ubiquitin chain. *Mol Cell Biol 14*:7876–7883.

Baker RT, Tobias JW, Varshavsky A. 1992. Ubiquitin-specific proteases of *Saccharomyces cerevisiae*. *J Biol Chem 267*:23364–23375.

Baker RT, Smith SA, Marano R, McKee J, Board PG. 1994. Protein expression using cotranslational fusion and cleavage of ubiquitin. *J Biol Chem 269*: 25381–25386.

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

Biggins S, Ivanovska I, Rose MD. 1996. Yeast ubiquitin-like genes are involved in duplication of the microtubule organizing center. *J Cell Biol 133*:1331–1346.

Board PG, Baker RT, Chelvanayagam G, Jermiin LS. 1997. Zeta, a novel class of glutathione transferases in a range of species from plants to humans. *Biochem J 328*:929–935.

Boguski MS, Lowe TMJ, Tolstoshev CM. 1993. dbEST—Database for "expressed sequence tags." *Nat Genet 4*:332–333.

Chau V, Tobias JW, Bachmair A, Marriott D, Ecker DJ, Gonda DK, Varshavsky A. 1989. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science 243*:1576–1583.

Chelvanayagam G, Eggenschwiler A, Knecht L, Gonnet GH, Benner SA. 1997. An analysis of simultaneous variation in protein structures. *Protein Eng 10*:307–316.

Chelvanayagam G, Roy G, Argos P. 1994. Easy adaptation of protein structure to sequence. *Protein Eng 7*:173–184.

Chen ZJ, Parent L, Maniatis T. 1996. Site-specific phosphorylation of IκBα by a novel ubiquitination-dependent protein kinase activity. *Cell 84*:853–862.

Chothia C, Lesk AM. 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J 5*:823–826.

Deshpande KL, Seubert PH, Tillman DM, Farkas WR, Katze JR. 1996. Cloning and characterization of cDNA encoding the rabbit tRNA-guanine transglycosylase 60-kilodalton subunit. *Arch Biochem Biophys 326*:1–7.

Doolittle RF. 1986. *Of URFs and ORFs: A primer on how to analyze derived amino acid sequences*. Mill Valley, California: University Science Books.

Emerson SD, Madison VS, Palermo RE, Waugh DS, Scheffler JE, Tsao KL, Kiefer SE, Liu SP, Fry DC. 1995. Solution structure of the Ras-binding domain of c-Raf-1 and identification of its Ras interaction surface. *Biochemistry 34*:6911–6918.

Galan J-M, Hagenauer-Tsapis R. 1997. Ubiquitin Lys63 is involved in ubiquitination of a yeast plasma membrane protein. *EMBO J 16*:5847–5854.

Gallagher T, Alexander P, Bryan P, Gilliland GL. 1994. Two crystal structure of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. *Biochemistry 33*:4721–4729.

Garrett KP, Aso T, Bradsher JN, Foundling SI, Lane WS, Conaway RC, Conaway JW. 1995. Positive regulation of general transcription factor SIII by a tailed ubiquitin homolog. *Proc Acad Natl Sci USA 92*:7172–7176.

Geyer M, Herrmann C, Wohlgemuth S, Wittinghofer A, Kalbitzer HR. 1997. Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signalling. *Nat Struct Biol 4*:694–699.

Gilchrist CA, Gray DA, Baker RT. 1997. A ubiquitin-specific protease that efficiently cleaves the ubiquitin–proline bond. *J Biol Chem 272*:32280–32285.

Gregori L, Poosch MS, Cousins G, Chau V. 1990. A uniform isopeptide-linked multiubiquitin chain is sufficient to target substrate for degradation in ubiquitin-mediated proteolysis. *J Biol Chem 265*:8354–8357.

Guarino LA. 1990. Identification of a viral gene encoding a ubiquitin-like protein. *Proc Natl Acad Sci USA 87*:409–413.

Haas AL, Ahrens P, Bright PM, Ankel H. 1987. Interferon induces a 15-kilodalton protein exhibiting marked homology to ubiquitin. *J Biol Chem 262*:11315–11323.

Haas AL, Siepmann TJ. 1997. Pathways of ubiquitin conjugation. *FASEB J 11*:1257–1268.

Hicke L. 1997. Ubiquitin-dependent internalization and down-regulation of plasma membrane proteins. *FASEB J 11*:1215–1226.

Hochstrasser M. 1996. Ubiquitin-dependent protein degradation. *Annu Rev Genet 30*:405–439.

Holm L. 1998. Unification of protein families. *Curr Opin Struct Biol 8*:372–379.

Huang L, Weng X, Hofer F, Martin GS, Kim SH. 1997. Three-dimensional structure of the Ras-interacting domain of RalGDS. *Nat Struct Biol 4*:609–615.

Johnson ES, Ma PCM, Ota IM, Varshavsky A. 1995. A proteolytic pathway that recognizes ubiquitin as a degradation signal. *J Biol Chem 270*:17442–17456.

Johnson ES, Schwienhorst I, Dohmen RJ, Blobel G. 1997. The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J 16*:5509–5519.

Jones DT, Taylor WR, Thornton JM. 1992. A new approach to protein fold recognition. *Nature 358*:86–89.

Kas K, Michiels L, Merregaert J. 1992. Genomic structure and expression of the human fau gene: Encoding the ribosomal protein S30 fused to a ubiquitin-like protein. *Biochem Biophys Res Commun 187*:927–933.

Kishimoto A, Nishiyama K, Nakanishi H, Uratsuji Y, Nomura H, Takeyama Y, Nishizuka Y. 1985. Studies on the phosphorylation of myelin basic protein by protein kinase C and adenosine 3′:5′-monophosphate-dependent protein kinase. *J Biol Chem 260*:12492–12499.

Kramer A, Mulhauser F, Wersig C, Groning K, Bilbe G. 1995. Mammalian splicing factor SFa120 represents a new member of the SURP family of proteins and is homologous to the essential splicing factor PRP21p of *Saccharomyces cerevisiae*. *RNA 1*:260–272.

Kreegipuu A, Blom N, Brunak S, Jarv J. 1998. Statistical analysis of protein kinase specificity determinants. *FEBS Lett 430*:45–50.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283–291.

Liakopoulos D, Doenges G, Matuschewski K, Jentsch S. 1998. A novel protein modification pathway related to the ubiquitin system. *EMBO J 17*:2208–2214.

Matsuzawa S, Takayama S, Froesch BA, Zapata JM, Reed JC. 1998. p53-Inducible human homologue of *Drosophila* seven in absentia (Siah) inhibits cell growth: Suppression by BAG-1. *EMBO J 17*:2736–2747.

Mayer RJ, Landon M, Layfield R. 1998. Ubiquitin superfolds: Intrinsic and attachable regulators of cellular activities? *Folding Design 3*:R97–R99.

Meluh PB, Koshland D. 1995. 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.

Meyers G, Tautz N, Dubovi EJ, Thiel HJ. 1991. Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. *Virology 180*:602–616.

Narasimhan J, Potter JL, Haas AL. 1996. Conjugation of the 15-kDa interferon-induced ubiquitin homolog is distinct from that of ubiquitin. *J Biol Chem 271*:324–330.

Nassar N, Horn G, Herrmann C, Scherer A, McCormick F, Wittinghofer A. 1995. The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. *Nature 375*:554–560.

Orengo CA, Jones DT, Thornton JM. 1994. Protein superfamilies and domain superfolds. *Nature 372*:631–634.

Papa FR, Hochstrasser M. 1993. The yeast DOA4 gene encodes a deubiquitinating enzyme related to a product of the human tre-2 oncogene. *Nature 366*:313–319.

Park KC, Woo SK, Yoo YJ, Wyndham AM, Baker RT, Chung CH. 1997. Purification and characterization of Ubp6, a new ubiquitin specific protease in *Saccharomyces cerevisiae*. *Arch Biochem Biophys 347*:78–84.

Pickart CM. 1997. Targeting of substrates to the 26S proteasome. *FASEB J 11*:1055–1066.

Rao-Naik C, delaCruz W, Laplaza JM, Tan S, Callis J, Fisher AJ. 1998. The Rub family of ubiquitin-like proteins. *J Biol Chem 273*:34976–34982.

Rost B, Sander C. 1994. Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins 19*:55–72.

Rost B, Sander C, Schneider R. 1994. PHD—An automatic mail server for protein secondary structure prediction. *Comput Appl Biosci 10*:53–60.

Sander C, Schneider R. 1991. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins 9*:56–68.

Schauber C, Chen L, Tongaonkar P, Vega I, Lambertson D, Potts W, Madura K. 1998. Rad23 links DNA repair to the ubiquitin/proteasome pathway. *Nature 391*:715–718.

Skowyra D, Craig KL, Tyers M, Elledge SJ, Harper JW. 1997. F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin–ligase complex. *Cell 91*:209–219.

Takayama S, Sato T, Krajewski S, Kochel K, Irie S, Millan JA, Reed JC. 1995. Cloning and functional analysis of BAG-1: A novel Bcl-2-binding protein with anti-cell death activity. *Cell 80*:279–284.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res 22*:4673–4680.

Tobias JW, Varshavsky A. 1991. Cloning and functional analysis of the ubiquitin-specific protease gene *UBP1* of *Saccharomyces cerevisisae*. *J Biol Chem 266*:12021–12028.

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

Vijay-Kumar S, Bugg CE, Cook WJ. 1987a. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol 194*:531–544.

Vijay-Kumar S, Bugg CE, Wilkinson KD, Vierstra RD, Hatfield PM, Cook WJ. 1987b. Comparison of the three-dimensional structures of human, yeast and oat ubiquitin. *J Biol Chem 262*:6396–6399.

Watkins JF, Sung P, Prakash L, Prakash S. 1993. The *Saccharomyces cerevisiae* DNA repair gene *RAD23* encodes a nuclear protein containing a ubiquitin-like domain required for biological function. *Mol Cell Biol 13*:7757–7765.

Whitby FG, Xia G, Pickart CM, Hill CP. 1998. Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. *J Biol Chem 273*:34983–34991.

Wilkinson KD. 1997. Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. *FASEB J 11*:1245–1256.

Woodgett JR, Gould KL, Hunter T. 1986. Substrate specificity of protein kinase C. Use of synthetic peptides corresponding to physiological sites as probes for substrate recognition requirements. *Eur J Biochem 161*:177–184.