Table 1
Classification of E3 ubiquitin-protein ligase

HECT-type:	AIP4, Ceb1, E6-AP, Herc2, Hul4, Hul5, hHYD/EDD, Itch, Nedd4, Pub1/2, Rsp5, Smurf1/2, SU/DX, Tom1, Ufd4, WWO1, WWP2, etc.
RING-type:	
(1) RING-HC (C3HC4)-finger:	BBAP, BRCA1, (Bre1), Cbls, Chfr, DTX3 (Deltex3), Efp, Hakai, HEI10, IAPs, ICP0, IE2, LNX, Mahogunin (mahoganoid), Mdm2, Mdmx, MID1, Mind Bomb (Mib), Momo, Neuralized (Neu), Nrdp1/FLRF, RAG1, Rma1, RNF2/HIPI-3, Sakura, Siah-1, SINAT5, Staring, TRAF6, etc.
(2) RING-H2 (C3H2C3)-finger:	AO7, Apc11, ARNIP, CIP8, DTX1 (Deltex1), DTX2/Deltex2, EL5, gp78, GRAIL/GREUL1, Hrd1, kf-1, NFX-1, Pirh2, Praja1/PJA1, Rbx1, RLIM, TRC8, Tul1, Ubr1
(3) RING-IBR-RING-finger:	Dorfin, HOIL-1, Parc, Parkin, etc.
(4) Atypical RING-finger:	K3/MIR1, K5/MIR2, MEKK1, Doa10, Pib1, CNOT4, etc.
U-box type:	ARC1, CHIP, CYC4, PRP19, Ufd2, Ufd2b UIP5, etc.
Others:	ICP0 HUL-1 domain, $TAF_{II}250$, (E1+E2 activity), $(UCH\text{-}L1)_2$, p300, etc.

HECT, homologous to E6-AP carboxyl terminus; RING, really interesting new gene; HUL-1, herpes virus ubiquitin ligase-1.

in the presence of E3 [15]. The fourth group of E3 consists of very unique E3s [ICPO, $TAF_{II}250$, $(UCH\text{-}L1)^2$, and p300] that have no sequence homology to known E3 enzymes. ICPO has two catalytic sites: one RING-HC and another novel HUL-1 motif [16]. $TAF_{II}250$ has intrinsic E1 and E2 activities within a single molecule, which exhibits no homology to other E3s [17]. $(UCH\text{-}L1)_2$ is the dimeric form of UCH-L1 (functioning as a de-ubiquitylation enzyme in a monomeric form) and exhibits E3 activity (see below) [18]. p300 exhibits E4-like activity in the presence of Mdm2 E3 ligase [19]. However, whether the above E3s are truly ubiquitin-ligases remains elusive at present. It is of note that all E3s except HECT-type E3s are probably not covalently bound to ubiquitin. It is plausible that certain domains, such as RING-finger or U-box, recruit E2s to the vicinity of proteins to be ubiquitylated and thus mediate ubiquitylation by facilitating the direct transfer of ubiquitin from E2-ubiquitin to the target Lys residue.

Interestingly, eukaryotic cells contain an unexpectedly large number of deubiquitylating enzymes (DUBs), which are also called Ub-specific proteases (USPs) (see Fig. 1). They belong to a family of cysteine proteases subclassified into at least two gene families that are structurally unrelated; the UCH (ubiquitin C-terminal hydrolase) family and the UBP (ubiquitin-specific processing protease) family [20]. DUBs may contribute to the production of a functional ubiquitin moiety from its precursors as well as disassembly of degradation intermediates generated by the 26S proteasome. Indeed, ubiquitin is reutilized, but not degraded in the breakdown of ubiquitylated proteins. DUBs are also thought to catalyze the reversal of the ubiquitylation reaction for "proofreading" of incorrectly ubiquitylated proteins or "trimming" of abnormal polyubiquitin structures, which play an essential role in facilitated proteolysis mediated by the 26S proteasome [21]. Intriguingly, a loss-of-function of many DUB genes by deletion or mutation or a gain-of-function by overexpression causes the induction of abnormalities in metabolism, growth, and differentiation of cells [20]. However, the reason for the presence of so many members of DUB in cells remains to be elucidated.

Ubiquitin is encoded in two types of unique genes [1]. One is a polyubiquitin gene encoding a tandemly repeated ubiquitin structure, which belongs to "heat-shock gene". Another ubiquitin gene produces an ubiquitin fused with certain ribosomal proteins, the biological significance of which is still unknown. After synthesis, these ubiquitin precursors are converted to functional molecules, the processing reaction of which is catalyzed by a set of UCH and/or certain UBPs as mentioned above. Clearly, the presence of stress-inducible polyubiquitin gene indicates that cells need a large amount of the free ubiquitin pool to respond to various environmental stresses, which serves to prevent accumulation of abnormal proteins in cells by selective proteasomal elimination after facilitated ubiquitylation of these unnecessary proteins. Thus, it becomes clear that UPS plays a pivotal role in regulating and maintaining protein homeostasis in cells.

3. Proteasome

Most cellular proteins in eukaryotic cells are targeted for degradation by the 26S proteasome, usually after they have been covalently attached to ubiquitin in the form of a polyubiquitin chain functioning as a degradation signal. The 26S proteasome is a eukaryotic ATP-dependent protease responsible for selective degradation of the polyubiquitin-tagged proteins. It is an unusually large multisubunit proteolytic complex, consisting of a central catalytic/core particle (CP, equivalent to a 20S proteasome) and two terminal regulatory particles (RP, also termed PA700 or 19S complex), which are attached to both ends of the central portion in opposite orientations to form the enzymatically active proteasome (reviewed in Refs. [7,8] and references therein). It appears to act as a highly organized apparatus designed for efficient and exhaustive hydrolysis of proteins, and can in fact be regarded as a protein-destroying machinery. The 20S proteasome is a barrel-like particle formed by the axial stacking of four rings made up of two outer α-rings and two inner β-rings, which are each made up of seven structurally similar α- and β-subunits, respectively, being associated in the order of αββα. Three β-type subunits of each inner ring have catalytically active threonine residues at

their N-terminus (in which of β1, β2, and β3 corresponding to caspase-like, trypsin-like, and chymotrypsin-like activities, respectively), and these active sites face the interior of the cylinder and reside in a chamber formed by the centers of the abutting β rings [22]. Thus, substrates gain access to the active sites only after passing through a narrow opening corresponding to the center of the α rings and the amino-termini of the α subunits form an additional physical barrier for substrates to reach the active sites.

Interestingly, the center of the α-ring of the 20S proteasome is almost closed, preventing penetration of proteins into the inner surface of the β-ring on which the proteolytically active sites are located. The regulator, PA700, was discovered as an activator of the latent 20S proteasome. PA700 (alias RP) consists of two subcomplexes, known as "base" and "lid", which, in the 26S proteasome, correspond to the portions of PA700 proximal and distal, respectively, to the 20S proteasome [23]. The base is made up of six ATPases and two large regulatory components while the lid contains multiple non-ATPase subunits. The base complex, thought to bind ATP-dependently to the outer α-ring of the central 20S proteasome, seems to be involved in opening the gate of the α-ring for entry of the protein substrate. The metabolic energy liberated by ATP consumption is probably utilized for the assembly of the base complex with the 20S proteasome, although it may also be used for unfolding target proteins, gate opening of the 20S proteasome, and substrate translocation so that they can penetrate the channel of the α- and β-rings of the 20S proteasome [24]. On the other hand, the lid-complex is thought to be involved in the recognition of target proteins, deubiquitylation for reutilization of ubiquitin, and interactions with various proteins including proteins with an ubiquitin-like domain or certain E3(s). Interestingly, it was found that Hsp90 interacts with the 26S proteasome and plays an important role in the assembly and maintenance of the 26S proteasome [25].

4. Protein quality control and Parkinson's disease

Over 30% of the newly synthesized cellular proteins are discarded without being properly folded, i.e., misfolded, and/or unassembled, even though they are normally synthesized without mutations of their genes or errors in the translation process [26]. Such inferior quality of the protein synthesizing-machine seems surprising, because it is thought that the biological system at present has been elegantly formulated during evolution. In addition, even if proteins are synthesized and folded accurately as functional proteins with normal tertiary structures, they often undergo damages due to various stresses under poor surroundings [4]. In fact, the high density of protein molecules in the intracellular space increases the spontaneous denaturation and consequently the likelihood of partially folded or unfolded proteins to undergo off-pathway reactions, such as aggregation. In addition, environmental stresses, such as heat, oxidation (i.e., formation of free radicals), and ultraviolet, could result in the production of impaired proteins. However, it is worth emphasizing that those proteins with non-native or aberrant structures are not observed in normal cells, because they are rapidly removed inside the cells. In other words, the cell is fully equipped with a surveillance system that rapidly eliminates such abnormal proteins unwanted for their presence in the cell.

In this regard, the cellular apparatus monitoring the "normality" of proteins in the cell is usually referred to as "the protein quality control system" [4,5]. Cells are equipped with two systems that prevent accumulation of abnormal proteins formed through the protein biosynthetic pathway or postsynthesis damage. One is the molecular chaperone(s) that recognizes proteins with non-native structures to prevent them from irreversible aggregation and help their conversion to a functional conformation. The other system is UPS, which is responsible for selective destruction of misfolded/unfolded and unassembled proteins, which probably fail to refold assisted by the chaperone system. Since both the chaperones and proteasomes conceptually recognize common substrates under non-native states, these two pathways act together to prevent aggregation and accumulation of such harmful proteins, thus maintaining protein homeostasis in the cell.

This monitoring machinery is considered an integral cellular component involved in maintaining cell survival and homeostasis, because it prevents the accumulation of abnormal proteins, formation of toxic inclusion bodies, with subsequent cell death, as seen in various neurodegenerative diseases [27]. It is well known that UPS is involved in the immediate destruction of proteins impaired by environmental and/or intracellular stresses. In this regard, the importance of UPS has been recently highlighted in the non-dividing cells of the brain, since neuronal intracellular inclusions are composed of ubiquitin-positive protein aggregates that have recently been described as a common ultrastructural feature of many neurodegenerative diseases, such as PD, Alzheimer's disease, Huntington's disease, and amyotrophic lateral sclerosis (ALS) and Prion disease [27]. In fact, it is also known that proteasome inhibitors increase the frequency of ubiquitin-positive intracellular inclusions in the cells carrying many neurodegenerative disorder genes. Therefore, one could assume that a critical aspect of various neuronal degenerative diseases is failure of the UPS-protein quality control system. Indeed, ample evidence has been provided for the potential link between failure of the protein quality control and neurodegeneration. In this regard, among many neurodegenerative disorders, PD is the most interesting, because there is direct evidence showing that the cause of PD is closely linked to functional abnormality of the UPS pathway [28].

Parkinson's disease, a severe neurological disorder of movement, is the second most common neurodegenerative

disorder [29]. The prevalence of PD increases markedly with age, affecting 1–2% of the population above 65 years. PD is characterized by a progressive loss of dopaminergic neurons in the substantia nigra pars compacta in the midbrain. The clinical symptoms include resting tremor, rigidity, bradykinesia, and postural instability. Sporadic PD is characterized by loss of dopaminergic neurons, i.e., selective neurodegeneration of the pigmented neurons, such as substantia nigra and locus coeruleus, in the brain stem. The pathological hallmark of PD is the presence of cytoplasmic proteinaceous inclusions known as Lewy body (LB) in surviving neurons [30]. Indeed, LB is often observed in sPD. LB displays a core and halo organization and stains strongly with anti-ubiquitin antibodies, and hence many proteins composed of LB are thought to be heavily ubiquitylated, implying that PD is associated with deviant behavior of the ubiquitin metabolism.

Approximately 5–10% of patients with PD are estimated to belong to the familial form of PD. The recent identification of several genes and gene loci linked to the familial forms of PD has significantly enhanced our understanding of the genetic mechanisms of PD [31–33]. The overlapping clinical and pathological features of idiopathic PD and familial Parkinsonism indicate a common underlying molecular pathway in all PDs. These outcomes suggest that the discovery of genes responsible for the familial forms of the disease should shed new light on the mechanism(s) leading to the selective dopaminergic cell loss in sporadic PD. Indeed, kindreds with Mendelian pattern of inheritance of PD have been known a long time ago. To date, 10 loci have been mapped in familial PD by linkage analysis and four causative genes have been cloned [31–33]. These genes of familial PD are subdivided into the autosomal dominant forms, i.e., PARK1 encoding αSynuclein (αSN) and PARK5 encoding ubiquitin C-terminal hydrolase L1 (UCH-L1), and the autosomal recessive forms, i.e., PARK2 encoding parkin and PARK7 encoding DJ-1. The biochemical features of proteins encoded by the familial PD genes are mostly unknown, but it is marvelous that parkin, a product of the causative gene PARK2 of autosomal recessive juvenile parkinsonism (ARJP), is a ubiquitin-protein ligase, indicating that ARJP is linked to impediment of proteolysis mediated by UPS. The following section will review the structure and pathophysiology of parkin.

5. Parkin

5.1. ARJP

ARJP/Park2 was first reported in 1973 by Yamamura et al. [34]. It is a levodopa-responsive parkinsonism that develops before 40 years of age (average onset around 25 years). Additional clinical festures include foot dystonia, sleep benefit, diurnal fluctuation, hyperactive tendon reflexes and less frequent resting tremor compared with sPD. Dystonic features and sleep benefit are common, particularly when the age of onset is young. In addition to these clinical features, genetic analysis of the parkin gene has expanded the clinical spectrum of ARJP. For example, Klein et al. [35] reported parkin deletions in a family with clinical features essentially identical to those of idiopathic PD and none of the family members displayed any of the clinical hallmarks described in patients with previously reported parkin mutations. The age of onset was 31 to 75 years, and showed apparently autosomal dominant inheritance. Maruyama et al. [36] also reported a Japanese family with pseudo-autosomal dominant inheritance with parkin gene mutations, although the age of onset was in the young range. These analyses broadened the clinical spectrum of ARJP, and we realize that ARJP is quite similar to adult-onset sporadic PD.

5.2. The parkin gene

The gene responsible for ARJP maps to 6q25.2–q27 chromosome, based on the indication by linkage to markers D6S305 and D6S253. The D6S305 marker was deleted in one ARJP patient [37]. By positional cloning within this microdeletion, in 1998 Kitada et al. [38] isolated a cDNA clone of 2960 bp with a 1395-bp open reading frame by positional cloning, and termed it parkin (formally PARK2). Mutations in the newly identified gene appeared to be responsible for the pathogenesis of ARJP and, therefore, the protein product was named as parkin. The *PARK2* gene is the second largest gene reported in human so far, spans over 1.4 Mb and has 12 exons. The isolated cDNA clone is 2960 bp with 1395-bp open reading frame, and encodes a protein of 465 amino acids with a molecular weight of about 52 kDa. Interestingly, the parkin gene is highly conserved across species, not only in vertebrates, such as human, rat, mouse but also in invertebrates such as *Caenorhabditis elegans* and *Drosophila melanogater* [39], suggesting that it plays a common role in various organisms. To date, various mutations such as exon deletions, exon multiplications or point mutations resulting in missense and nonsense changes of parkins have been reported in ARJP patients [40,41]. Mutations in the parkin are found in nearly 50% of patients with autosomal-recessive early-onset parkinsonism (less than 45 years of age).

5.3. The parkin protein

Parkin is characterized by the presence of an ubiquitin-like domain (UBL) at its NH_2 terminus and two RING finger motifs, flanked by one IBR (in between RING finger) motif, at its COOH-terminal region (Fig. 2). In 2000, parkin was found to be an E3 ubiquitin ligase, a critical component of the pathway that covalently attaches ubiquitin to specific proteins with a polymerization step to form a degradation signal [42–44]. Indeed, parkin catalyzes the addition of

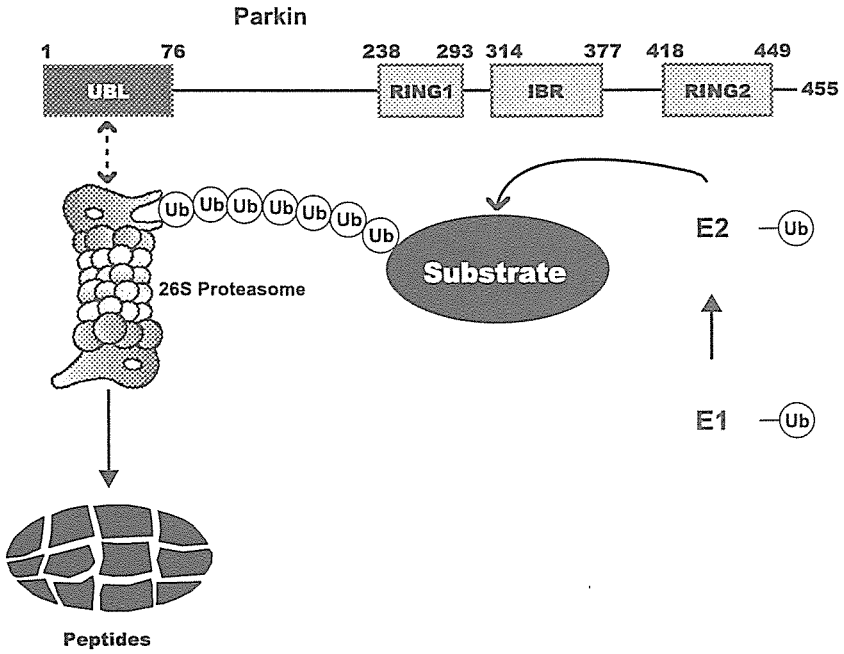

Fig. 2. Model of the parkin-directed ubiquitylation pathway. Ub, ubiquitin; E1, Ub-activating enzyme; E2, Ub-conjugating enzyme; UBL, ubiquitin-like domain. Parkin consists of two functionally distinct regions: a UBL segment responsible for interaction with the 26S proteasome and a RING-box (RING1-IBR-RING2) recruiting E2 for ubiquitylation. See text for details.

ubiquitin to target proteins prior to their destruction via the proteasome, indicating that the misregulation of proteasomal degradation of parkin substrate(s) is deleterious to dopaminergic neurons and, thus, impaired protein clearance can induce dopaminergic cell death. As mentioned, several types of mutations including deletion mutations and point mutations (both missense- and nonsense-mutations) have been detected in the parkin gene of ARJP patients. The site of these mutations spans almost all regions including the N-terminal UBL domain and the RING-IBR-RING domain, and there seems to be no difference of phenotypes among these mutation sites, indicating that the entire region of the parkin protein is essential for exerting its physiology.

As shown in Table 1, parkin belongs to a member of the RING-IBR-RING family, which is a subgroup of RING-finger type E3s. The RING-IBR-RING family is widespread in eukaryotes, with many members in animals (mammals, *Drosophila*, *Caenorhabditis*) and plants (*Arabidopsis*), although only two members exist in the budding and fission yeasts [45]. For instance, dorfin, HOIL1, and Parc are included in this family. Dorfin (double ring-finger protein) cloned from human spinal cord exhibits a RING finger-type ubiquityl ligase activity for mutant copper/zinc superoxide dismutase (SOD1) linked to ALS, but not for its wild-type [46]. Dorfin is predominantly localized in the inclusion bodies of familial ALS with SOD1 mutation as well as sporadic ALS [47]. Indeed, the cytopathological hallmark in the remaining motor neurons of ALS is the presence of ubiquitylated inclusions consisting of insoluble protein aggregates. Interestingly, dorfin is also localized with ubiquitin in LBs and colocalized in these large inclusions with ubiquitin and proteasomal components, suggesting that dorfin may be involved in LB formation and the pathogenic process of PD [48]. HOIL-1 (equivalent to XAP3/RBCK1/UIP28) resembles parkin, having the N-terminal UBL domain and C-terminal the RING-finger/IBR domain [49]. It is an E3 ligase that recognizes oxidized iron regulatory protein 2 (IRP2), a modulator of iron metabolism, suggesting that oxidation is a specific recognition signal for ubiquitylation of IRP2. Like dopamine, iron is in involved in the production of reactive oxygen species (ROS), suggesting that HOIL-1 may be responsible for neurodegenerative disorders, collaborating with or independent of parkin and/or dorfin. Parc (a parkin-like ubiquitin ligase) is identified as a cytoplasmic anchor protein in p53 (a tumor suppressor)-associated protein complexes [50]. Parc is a critical regulator of p53 subcellular localization and subsequent function. It is a large multi-domain protein containing the RING-finger/IBR domain at the C-terminal region.

The RING-IBR-RING domain is also called the TRIAD or DRILL domain, which functions by interacting with ubiquitin-conjugating enzymes (E2s), suggesting it is the catalytic site of this class of E3-enzyme family. Indeed, parkin, dorfin, and HOIL-1 bind to specific E2 species through their RING-IBR-RING motif. Intriguingly, these three RING-IBR-RING type E3s are strongly expressed in the brain.

The UBL region of parkin exhibits moderate similarity to ubiquitin, displaying approximately 20% of sequence identity. However, its role is largely unknown. The number of identified mutations of the parkin gene has recently increased in patients with early-onset parkinsonism as

described above, and a single mutation that causes Arg-Pro substitution at position 42 of the UBL domain has been identified in one family of ARJP patients [51]. This mutated parkin retains the ability to bind E2 (UbcH7), but fails to co-immunoprecipitate ubiquitylated proteins, such as *O*-glycosylated αSN [52]. Analogously, HOIL-1 interacts with oxidized IRP2 through the N-terminal UBL domain [49]. Thus, it is likely that the UBL domain contributes to the recognition of target proteins. On the other hand, accumulating evidence suggests that various proteins harboring the UBL domain, e.g., Rad23, Dsk2, and their human homologues (hHR23α/β and hPLIC1/2, respectively), provide links between the 26S proteasome and the ubiquitylation machinery [53]. In this context, it has been reported that a 50-kDa subunit Rpn10 of the human 26S proteasome, originally called S5a, could bind to polyubiquitin conjugates in vitro and, hence, could possibly function as a polyubiquitin chain-binding subunit [54]. Intriguingly, Rpn10 also binds the UBL domain of hHR23α/β and hPLIC-2 [55].

NMR studies have indicated that the three-dimensional structure of the UBL domain of parkin resembles that of ubiquitin [56]. As shown in Fig. 3, ubiquitin is a small protein with two α-helical and five β-sheet structures, which arrange in the order of ββαββαβ in the secondary structure. Overall, these structures are conserved in the UBL of parkin, indicating that both molecules appear to be structurally very similar. Intriguingly, inspection of chemical shift perturbation data revealed that UBL binds the Rpn10 subunit of the 26S proteasome with the surface area (see a model in Fig. 2). It was recently reported that one family of ARJP patients exhibited a point mutation at position 42 [51], where Arg is substituted with Pro. Intriguingly, the NMR data also indicate that Arg42 is located in the Rpn10-binding site. It is quite conceivable that this mutation induces a significant conformational change in the Rpn10-binding site of UBL, resulting in impaired proteasomal binding of parkin, which could be the structural basis of ARJP. This finding provides direct evidence that parkin is linked to cellular proteolysis, and its dysfunction presumably causes ARJP. This is of particular importance, because even if parkin is an E3 ligase, the possibility that it acts other than proteolysis could not be excluded at present.

5.4. The role of parkin in the pathogenesis of ARJP

The molecular basis of selective neuronal death in AR-JP is unknown. The biological role of parkin is also largely unknown at present. However, it is clear that a lesion in the parkin gene causes ARJP by dopaminergic neuronal loss, indicating the vital role of parkin in the survival of these neurons. Nonetheless, parkin mutations can lead to a disorder clinically similar to sporadic PD; a pathognomonic feature of ARJP is the lack of LB, which has been demonstrated in all but one reported case [57]. Why is LB missing in ARJP and how is LB formed in sporadic PD? One plausible explanation is that functional parkin plays an essential role in the formation of LBs as highly ubiquitylated inclusions.

Recent biochemical studies have clarified the type of proteins forming the LB. These include αSN, UCH-L1, parkin, synphilin-1, VPC (equivalent to p97 ATPase or CDC48), molecular chaperones, proteasome subunits [28,58,59]. The reason(s) for the formation of these proteins into LB is unknown, but they are conceptually divided into

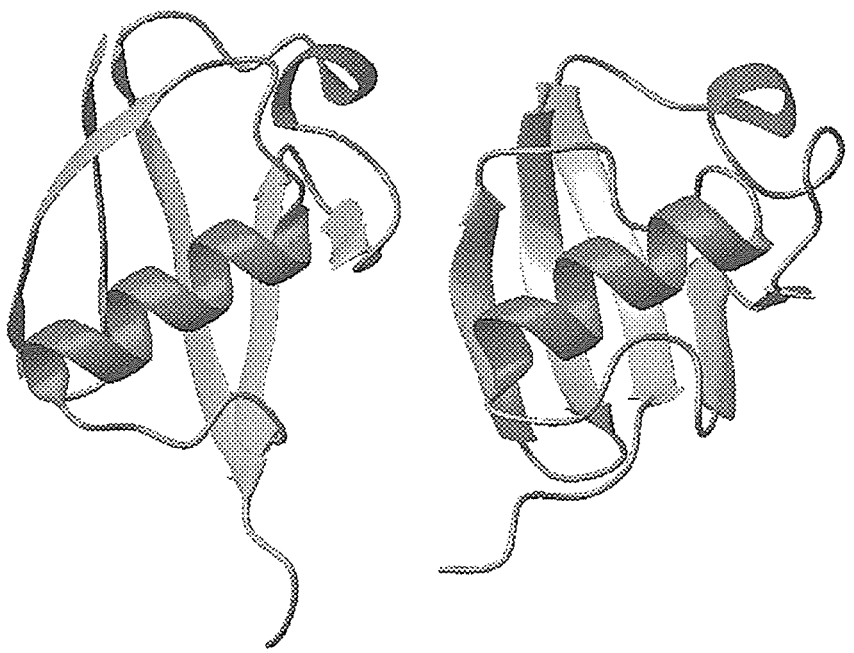

Fig. 3. The tertiary structures of ubiquitin (left) and the ubiquitin-like domain of parkin (right). α-Helices and β-sheets are shown in red and yellow, respectively.

two distinct groups based on their origin. One group belongs to specified proteins that are extraordinarily abundant in the brain and can be easily denatured and form aggregates, or other proteins associated with such proteins. The other class of proteins in the LB is involved in the cellular machinery responsible for refolding or deletion of abnormal proteins generated in cells. They include molecular chaperones and components included in UPS, respectively. Thus, the formation of LB is due to the loss of protein quality control in the cell. In other words, it is likely that certain neurons are vulnerable to the failure of UPS-mediated protein breakdown, based on the fact that multiple factors, including genetic and environmental factors, are concomitantly responsible for PD. Further studies of LB including the mechanism(s) involved in their formation and their role may shed light on the cause of PD.

LB may provide neuroprotection, because the formation of intracellular inclusions seems to be beneficial to cells. Consistently, parkin also functions in neuroprotective events (Refs. [58,59] and references therein). In considering ARJP lacking LB, parkin is thought to be an E3 ubiquitin ligase that degrades proteins with aberrant conformations [58,60]. In this context, the lack of LB in ARJP is consistent with the notion that ARJP is an early-onset disease. It is likely that failure of UPS to degrade abnormal proteins may underlie nigral degeneration and LB formation that occur in PD. Impaired protein clearance can induce dopaminergic cell death, supporting the concept that defects in the UPS may underlie nigral pathology in ARJP and perhaps sporadic forms of PD.

5.5. Parkin-interacting molecules

As mentioned, parkin interacts with E2s and Rpn10 (and hence the 26S proteasome) through the RING-IBR-RING domain and UBL domain, respectively. In addition, multiple pieces of evidence indicate that parkin interacts with various other proteins except the substrates. Of these, CHIP is an interesting molecule, because it can be regarded as "quality-control E3" that selectively ubiquitylates unfolded protein(s) by collaborating with molecular chaperones Hsp90 and Hsp70 [61]. Recently, it was reported that parkin forms a complex with CHIP, Hsp70, and Pael-R (known as parkin substrate, see below) both in vitro and in vivo [62]. The amount of CHIP in the complex is increased during stress of the endoplasmic reticulum (ER). CHIP promotes the dissociation of Hsp70 from parkin and Pael-R, thus facilitating parkin-mediated Pael-R ubiquitylation. Moreover, CHIP enhances parkin-mediated in vitro ubiquitylation of Pael-R in the absence of Hsp70. Thus, CHIP acts as a mammalian E4-like enzyme that positively regulates parkin E3 activity. On the other hand, it is also reported that parkin forms a complex with expanded poly-Gln protein, Hsp70 and the proteasome, which may be important for the elimination of the poly-Gln protein [63]. In addition, Hsp70 enhances parkin binding and ubiquitylation of poly-Gln protein, suggesting that Hsp70 may help to recruit misfolded proteins as substrates for parkin E3 ligase activity.

A recent study reported that parkin is a component of an SCF (Skp1, Cullin-1, Roc1, and F-box protein)-like ubiquitin ligase [64]. Indeed, parkin functions in a multi-protein ubiquitin ligase complex that includes the F-box/WD repeat protein hSel-10 and Cullin-1. Intriguingly, cyclin E is a substrate of this ubiquitin-ligase complex including parkin. Parkin deficiency potentiates the accumulation of cyclin E in cultured postmitotic neurons exposed to the glutamate excitotoxin kainate and promotes their apoptosis. Furthermore, overexpression of parkin attenuates the accumulation of cyclin E in toxin-treated primary neurons, including midbrain dopaminergic (DA) neurons, and protects them from apoptosis.

5.6. Putative substrates

Parkin is an E3 enzyme, hence it is conceivable that the pathogenesis of ARJP involves accumulation of unidentified neurotoxic protein(s) as a substrate of parkin. Furthermore, the abnormal accumulation of substrates caused by the loss of parkin function may be the cause of neurodegeneration in parkin-related parkinsonism. Thus, the misregulation of proteasomal degradation of parkin substrate(s) is deleterious to neurons. Therefore, accumulation of substrate(s) for ubiquitylation mediated by parkin is probably critical to our understanding of the pathogenesis of ARJP. Identification of these substrates and their role are important. Presumably, the parkin protein displays ubiquitin-ligase activity for different targets, which accumulate in the brain of ARJP patients due to parkin defect and might cause neurodegeneration. To date, accumulation of various candidate parkin-substrates has been described, including that of CDCrel-1 (a synaptic vesicle associated GTPase), Pael-R (parkin associated endothelin receptor like receptor), glycosylated αSN, synphillin-1 (an αSN interacting protein), cyclin E, α/β tubulin, and transfer RNA synthetase (p58, a key structural component of the mammalian aminoacyl-tRNA synthetase complex). Abnormal accumulation of these substrates due to loss of parkin function may be the cause of neurodegeneration in parkin-related parkinsonism, although direct links between these factors and dopaminergic cell death have not yet been established.

Pael receptor (Pael-R) identified as a protein that interacts with parkin is an attractive parkin substrate whose accumulation may account for the loss of DA neurons in ARJP [65]. Pael-R is a putative G protein-coupled transmembrane polypeptide. When overexpressed in cells, this receptor tends to become unfolded, insoluble, and ubiquitylated in vivo. The insoluble Pael-R leads to unfolded protein-induced cell death. Parkin specifically ubiquitylates this receptor in the presence of ER-resident E2s ubiquitin-conjugating enzymes and promotes the degradation of insoluble Pael-R, resulting in suppression of the cell death

induced by Pael-R overexpression. Moreover, the insoluble form of Pael-R accumulates in the brains of AR-JP patients. Accumulation of the unfolded Pael-R in the ER of dopaminergic neurons induces ER stress leading death of dopaminergic neurons in AR-JP. It was also reported in an organismal system that pan-neuronal expression of Pael-R causes age-dependent selective degeneration of *Drosophila* dopaminergic neurons [66]. Coexpression of parkin degraded Pael-R and suppressed its toxicity, whereas interfering with endogenous fly parkin function promoted Pael-R accumulation and augmented its toxicity.

Like αSN (see later), Pael-R has a propensity to misfold and aggregate. Importantly, not only the unfolded protein response (UPR) can be induced by Pael-R, but also UPR induces up-regulation of both the mRNA and protein levels of parkin. Furthermore, overexpression of parkin, but not a set of mutants without the E3 activity, specifically suppressed unfolded protein stress-induced cell death [43]. These findings suggest that parkin is involved in the ubiquitylation pathway for misfolded proteins derived from ER and contributes to protection from neurotoxicity induced by UPR. Thus, parkin functions in the ER-associated degradation (ERAD) of misfolded ER protein, and it is up-regulated by unfolded-protein stress. However, how the unfolded-protein stress response fits into a common pathway is not clear at present. In contrast, other studies indicate that parkin is not regulated by the unfolded protein response in SH-SY5Y neuroblastoma cells (Ref. [67] and references therein). Consistently, parkin levels were unaffected by thaspsiargin treatment in rat cortical neuron cultures and tunicamycin treatment in rat primary hippocampal neurons. Whether parkin truly acts as a key player in the UPR pathway awaits future studies.

αSN, consisting of 140 amino acid residues, is a highly conserved protein in vertebrates [68]. It is the primary component of the LB. It has no secondary structure and intriguingly can be degraded by the 20S proteasome as well as the 26S proteasome [69], suggesting that denaturing proteins can activate the 20S proteasome by the gate opening of the central portion of the α-ring of the 20S proteasome (see Section 3). Implicated in neurotoxicity are two αSN mutants (A53T and A30P) that cause extremely rare familial PD (PARK1). It is interesting that the *O*-glycosylated form of αSN (named αSp22) becomes a target for parkin [52]. In contrast to normal parkin, mutant parkin associated with ARJP failed to bind and ubiquitylate αSp22. Thus, αSp22 is a substrate for parkin's E3 ligase activity in normal human brain and that loss of parkin function causes pathological αSp22 accumulation. These findings demonstrate a critical biochemical reaction between the two PD-linked gene products and suggest that this reaction underlies the accumulation of ubiquitylated αSN in conventional PD. Note that nonglycosylated αSN, the major species in the brain, is not a parkin substrate in vivo or in the brain.

Overexpression of mutant αSN increases sensitivity to proteasome inhibitors by decreasing proteasome function [70]. Overexpression of parkin decreases sensitivity to proteasome inhibitors in a manner dependent on parkin's E3 ligase activity, and antisense knockdown of parkin increases sensitivity to proteasome inhibitors. Mutant αSN also causes selective toxicity to catecholaminergic neurons in primary midbrain cultures, an effect that can be mimicked by the application of proteasome inhibitors. Parkin is capable of rescuing the toxic effects of mutant αSN or proteasome inhibition in these cells. Therefore, parkin and αSN are linked by common effects on a pathway associated with selective cell death in catecholaminergic neurons.

Synphilin-1 is linked to the pathogenesis of PD since it is an αSN and parkin interacting protein. Moreover, it is a component of LB in brains of sporadic PD patients. Parkin interacts with and ubiquitylates synphilin-1 [71]. Co-expression of αSN, synphilin-1, and parkin results in the formation of LB-like ubiquitin-positive cytosolic inclusions. It was shown that familial-linked mutations in parkin disrupt the ubiquitylation of synphilin-1 and the formation of the ubiquitin-positive inclusions. These results provide a molecular basis for the ubiquitylation of LB-associated proteins and link parkin and αSN in a common pathogenic mechanism through their interaction with synphilin-1.

Several other candidate substrates for parkin E3 have been reported. For example, parkin interacts with the synaptic vesicle-associated protein, CDCrel-1, belonging to a family of spetin GTPases, through its RING-finger domains [72]. Parkin ubiquitylates and promotes the degradation of CDCrel-1, whereas its familial-linked mutations impair CDCrel-1 degradation. On the other hand, parkin is tightly bound to microtubules that are polymers of tubulin α/β heterodimers, ubiquitylates α/β tubulin and promotes their degradation [73]. Its ability to enhance the ubiquitylation and degradation of misfolded tubulins may play a significant role in protecting neurons from toxins that cause PD. Misfolded tubulin monomers are highly toxic and thus must be rapidly degraded. Parkin also interacts with, ubiquitylates and promotes the degradation of p38, a key structural component of the mammalian aminoacyl-tRNA synthetase complex [74]. Interestingly, expression of p38 in COS7 cells resulted in the formation of aggresome-like inclusions in which parkin was systematically sequestered. In the human dopaminergic neuroblastoma-derived SH-SY5Y cell line, parkin promoted the formation of ubiquitylated p38-positive inclusions. Moreover, overexpression of p38 in SH-SY5Y cells caused significant cell death against which parkin provided protection. This suggests that p38 plays a role in the pathogenesis of PD.

5.7. Physiological roles of parkin in model organisms

Several lines of evidence strongly implicate mitochondrial dysfunction as a major causative factor in PD, although the molecular mechanisms responsible for mitochondrial dysfunction are poorly understood [75]. *Drosophila* parkin null mutants exhibit reduced life span, locomotor defects,

and male sterility [76]. The tissue-specific phenotypes observed in *Drosophila* parkin mutants result from mitochondrial dysfunction and raise the possibility that similar mitochondrial impairment triggers the selective cell loss observed in ARJP. The locomotor defects are derived from apoptotic cell death of muscle subsets, whereas the male sterile phenotype is derived from a spermatid individualization defect at a late stage of spermatogenesis. Mitochondrial pathology is the earliest manifestation of muscle degeneration and a prominent characteristic of individual spermatids in parkin mutants. Thus, the tissue-specific phenotypes observed in *Drosophila* parkin mutants result from mitochondrial dysfunction and raise the possibility that similar mitochondrial impairment triggers the selective cell loss observed in AR-JP. In contrast, disruption of the parkin gene in *C. elegans* reveals no obvious phenotypes in their behaviors (Suzuki et al., unpublished results).

Inactivation of the parkin gene in mice results in motor and cognitive deficits, inhibition of amphetamine-induced dopamine release and inhibition of glutamate neurotransmission [77]. The levels of dopamine are increased in the limbic brain areas of parkin mutant mice and there is a shift towards increased metabolism of dopamine by monoamine oxidase. Although there is no evidence for a reduction of nigrostriatal dopamine neurons in the parkin mutant mice, the level of dopamine transporter protein is reduced in these animals, suggesting a decreased density of dopamine terminals or adaptive changes in the nigrostriatal dopamine system. The reduction of glutathione levels was worse in the striatum and fetal mesencephalic neurons of parkin mutant mice, suggesting that a compensatory mechanism may protect dopamine neurons from neuronal death.

6. Perspective

UPS plays a central role in the protein quality control in the cytosol as well as the ER of the cell. It is therefore conceivable that functional loss of the protein quality control system is associated with various neurodegenerative diseases. In particular, with regard to the pathogenesis of PD, parkin is an E3, indicating that the ubiquitin pathway is directly linked to the cause of ARJP [58,60]. However, loss of parkin activity due to disease-linked pathogenic mutations is not simply related to the pathogenesis of ARJP, because parkin is widely expressed in most neurons in the brain. Thus, in addition to this genetic defect, there is another factor(s) capable of inducing DA neuronal death, which may act in concert with parkin in causing the degeneration of DA neurons. Available evidence indicates that environmental factors may play a role in the diseases process. It is likely that for most cases, there is a complex interplay between these genetic and environmental influences in the causation of PD. It has been postulated that environmental conditions are closely linked to loss-of-function of certain protein(s) due to disease-linked genetic mutations. It is possible that dysfunction of the mitochondrial respiratory chain, inducing oxidative stress, could be involved in the pathogenesis of ARJP, in concert with loss of parkin activity [75]. Indeed, the mitochondrial complex I inhibitors 1-methy-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) and rotenone cause damage of nigral neurons by mechanisms involving oxidation. Oxidative stress seems to play a prominent role in ARJP and sporadic PD [28]. Compared with other brain regions, the substantia nigra pars compacta has a higher rate of basal protein oxidation due to the enzymatic- and auto-oxidation of dopamine. Thus, this brain region has a propensity to generate high levels of oxidatively and nitratively damaged proteins. At present, there is convincing evidence suggesting that proteolytic stress due to failure of the UPS to clear unwanted proteins is a common factor among the different familial and sporadic forms of PD [68]. In this regard, dopamine metabolism may be linked to induction of proteolytic stress through production of various molecules toxic for protein structure such as ROS. Disease mechanisms seem to converge around oxidative damage and impairment of protein catabolism. Further studies are required to clarify this issue molecularly.

It is interesting that parkin is associated with synaptic membranes, even if not conclusively. Moreover, αSN is also associated with synapses, indicating that it may play some role in the modulation of synaptic plasticity. Synphilin-1 is abundant in synaptic compartments. CDCrel-1 is associated with synapses. It is likely that the real link between these proteins is that they all regulate the signaling pathways at synapses. In addition, it is possible that failure of protein homeostasis, depending on the loss-of-function of parkin, in synaptic vesicles is linked to the cause of PD. However, it is still unknown whether parkin is directly linked to the function of synapses.

To date, hereditary factors have emerged as the focus of research in PD [31,32]. In addition to parkin (*PARK2*), three genes of familial PDs have been identified. The identification of pathogenic mutations of three genes, αSN (*PARK1*), parkin and UCH-L1 (*PARK5*), has elucidated UPS and its potential role as a casual pathway in PD. Mutations of αSN and UCH-L1 are extremely rare, while mutations of parkin are more common in early-onset cases but still in a small population of the total number of cases of PD, as discussed. UCH-L1 belongs to the family of DUBs, is expressed exclusively in neurons, and represents over 1% of total brain protein content and has tendency to aggregate, like αSN. In addition, both are major components of LB. Therefore, parkin, αSN, and UCH-L1 are on the same line of abnormality of protein homeostasis. Interestingly, UCH-L1 has a dimerization-dependent ligase activity that promotes addition of ubiquitin to preformed ubiquitin chains on proteins [18]. A polymorphic variant of UCH-L1 has reduced ligase activity but comparable hydrolase activity, relative to the wild-type enzyme. Thus, the ligase activity as well as the monomeric hydrolase activity of UCH-L1 may

play a role in proteasomal protein degradation, a critical process for neuronal health.

Recently, DJ-1 (PARK7) was identified as a novel early-onset recessive PD gene and found to encode DJ-1 protein, consisting of 189 amino acids (20 kDa) [78]. Overall, the symptoms of this disease are very similar to parkin and significantly overlap with idiopathic PD. However, it is difficult at this juncture to fully appreciate how mutations of the DJ-1 gene cause PD, since its function is basically unknown. DJ-1 is thought to participate in the oxidative stress response. This is of interest, because there is evidence that oxidative events occur during the course of PD.

To date, several new loci for the causative genes of other forms of familial PD have been documented. The identification of multiple genetic causes will provide further impetus to characterize the pathway leading to PD. New candidate genes are expressed to encode proteins either involved in UPS or sequestrated in intracytoplasmic protein aggregates. Future identification of disease genes is required for confirmation, thereby unifying the clinical and genetic heterogeneity of PD.

References

[1] A. Hershko, A. Ciechanover, A. Varshavsk, The ubiquitin system, Nat. Med. 6 (2000) 1073–1081.

[2] C.M. Pickart, Mechanisms underlying ubiquitination, Annu. Rev. Biochem. 70 (2001) 503–533.

[3] A.M. Weissman, Themes and variations on ubiquitylation, Nat. Rev., Mol. Cell Biol. 2 (2001) 169–178.

[4] M.Y. Sherman, A.L. Goldberg, Cellular defenses against unfolded proteins: a cell biologist thinks about neurodegenerative diseases, Neuron 29 (2001) 15–32.

[5] Y. Yoshida, A novel role for *N*-glycans in the ERAD system, J. Biochem. (2003) 183–190.

[6] J. Peng, D. Schwartz, J.E. Elias, C.C. Thoreen, D. Cheng, G. Marsischky, J. Roelofs, D. Finley, S.P. Gygi, A proteomics approach to understanding protein ubiquitination, Nat. Biotechnol. 21 (2003) 921–926.

[7] O. Coux, K. Tanaka, A.L. Goldberg, Structure and functions of the 20S and 26S proteasomes, Annu. Rev. Biochem. 65 (1996) 801–847.

[8] W. Baumeister, J. Walz, F. Zuhl, E. Seemuller, The proteasome: paradigm of a self-compartmentalizing protease, Cell 92 (1998) 367–380.

[9] M. Schneffner, U. Nuber, J.M. Huibregtse, Protein ubiquitination involving an E1–E2–E3 enzyme thioester cascade, Nature 373 (1994) 81–83.

[10] K.L.B. Borden, RING domains: master builders of molecular scaffolds? J. Mol. Biol. 295 (2000) 1103–1112.

[11] P.K. Jackson, A.G. Eldridge, E. Freed, L. Furstenthal, J.Y. Hsu, B.K. Kaiser, J.D.R. Reimann, The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases, Trends Cell Biol. 10 (2000) 429–439.

[12] H. Scheel, K. Hofmann, No evidence for PHD fingers as ubiquitin ligases, Trends Cell Biol. 13 (2003) 285–287.

[13] L. Aravind, E.V. Koonin, The U box is a modified RING finger—a common domain in ubiquitination, Curr. Biol. 10 (2000) R132–R134.

[14] M.D. Ohi, CW. Vander Kooi, J.A. Rosenberg, W.J. Chazin, K.L. Gould, Structural insights into the U-box, a domain associated with multi-ubiquitination, Nat. Struct. Biol. 10 (2003) 250–255.

[15] M. Koegl, T. Hoppe, S. Schlenker, HD. Ulrich, TU. Mayer, S. Jentsch, A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly, Cell 96 (1999) 635–644.

[16] R. Hagglund, B. Roizman, Characterization of the novel E3 ubiquitin ligase encoded in exon 3 of herpes simplex virus-1-infected cell protein 0, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 7889–7894.

[17] A.D. Pham, F. Sauer, Ubiquitin-activating/conjugating activity of $TAF_{II}250$, a mediator of activation of gene expression in *Drosophila*, Science 289 (2000) 2357–2360.

[18] Y. Liu, L. Fallon, H.A. Lashuel, Z. Liu, P.T. Lansbury Jr., The UCH-L1 gene encodes two opposing enzymatic activities that affect α-synuclein degradation and Parkinson's disease susceptibility, Cell 111 (2002) 209–218.

[19] S.R. Grossman, M.E. Deato, C. Brignone, H.M. Chan, A.L. Kung, H. Tagami, Y. Nakatani, D.M. Livingston, Polyubiquitination of p53 by a ubiquitin ligase activity of p300, Science 300 (2003) 342–344.

[20] J.H. Kim, K.C. Park, S.S. Chung, O. Bang, C.H. Chung, Deubiquitinating enzymes as cellular regulators, J. Biochem. 134 (2003) 9–18.

[21] K.D. Wilkinson, Regulation of ubiquitin-dependent processes by deubiquitinating enzymes, FASEB J. 11 (1997) 1245–1256.

[22] M. Bochtler, L. Ditzel, M. Groll, C. Hartmann, H. Huber, The proteasome, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 295–317.

[23] M.H. Glickman, A. Ciechanover, The ubiquitin–proteasome proteolytic pathway: destruction for the sake of construction, Physiol. Rev. 82 (2002) 373–428.

[24] T. Ogura, K. Tanaka, Dissecting various ATP-dependent steps in proteasomal degradation, Mol. Cell 11 (2003) 3–5.

[25] J. Imai, M. Maruya, H. Yashiroda, I. Yahara, K. Tanaka, The molecular chaperone Hsp90 interacts with 26S proteasomes and regulates their assembly, EMBO J. 22 (2003) 3557–3567.

[26] U. Schubert, L.C. Anton, J. Gibbs, C.C. Norbury, W.J. Yewdell, J.R. Bennink, Rapid degradation of a large fraction of newly synthesized proteins by proteasomes, Nature (2000) 770–774.

[27] R.J. Mayer, M. Landon, J. Lowe, Ubiquitin and the molecular pathology of human disease, in: J.-M. Peters, J.R. Harris, D. Finley (Eds.), Ubiquitin and the Biology of the Cell, Plenum Press, New York, 1998, pp. 147–189.

[28] K.S. McNaught, C.W. Olanow, Proteolytic stress: a unifying concept for the etiopathogenesis of Parkinson's disease, Ann. Neurol. 53 (Suppl. 3) (2003) S73–S84.

[29] A.E. Lang, A.M. Lozano, Medical progress: Parkinson's disease, N. Engl. J. Med. 339 (1998) 1044–1053.

[30] M.S. Pollanen, D.W. Dickson, C. Bergeron, Pathology and biology of Lewy body, J. Neuropathol. Exp. Neurol. 52 (1993) 183–191.

[31] T.T. Warner, A.H. Schapira, Genetic and environmental factors in the cause of Parkinson's disease, Ann. Neurol. 53 (Suppl. 3) (2003) S16–S23.

[32] T.M. Dawson, V.L. Dawson, Rare genetic mutations shed light on the pathogenesis of Parkinson disease, J. Clin. Invest. 111 (2003) 145–151.

[33] T. Gasser, Overview of the genetics of parkinsonism, Adv. Neurol. 91 (2003) 143–152.

[34] Y. Yamamura, S. Kuzuhara, K. Kondo, H. Matsumine, Y. Mizuno, Clinical, pathologic, and genetic studies on autosomal recessive early-onset parkinsonism with diurnal fluctuation, Parkinsonism Relat. Disord. 4 (1998) 65–72.

[35] C. Klein, K. Schumacher, H. Jacobs, J. Hagenah, B. Kis, J. Garrels, E. Schwinger, L. Ozelius, P. Pramstaller, P. Vieregge, P.L. Kramer, Association studies of Parkinson's disease and parkin polymorphisms, Ann. Neurol. 48 (2000) 126–127.

[36] M. Maruyama, T. Ikeuchi, M. Saito, A. Ishikawa, T. Yuasa, H. Tanaka, S. Hayashi, K. Wakabayashi, H. Takahashi, S. Tsuji, Novel mutations, pseudo-dominant inheritance, and possible familial affects in patients with autosomal recessive juvenile parkinsonism, Ann. Neurol. 48 (2000) 245–250.

[37] H. Matsumine, M. Saito, S. Shimoda-Matsubayshi, H. Tanaka, A. Ishikawa, Y. Nakagawa-Hattori, M. Yokochi, T. Kobayashi, S. Igarashi, H. Takano, K. Sanpei, R. Koike, H. Mori, T. Kondo, Y. Mizutani, A.A. Schaffer, Y. Yamamura, S. Nakamura, S. Kuzuhara, S. Tsuji, Y. Mizuno, Localization of a gene for autosomal recessive form of juvenile parkinsonism (AR-JP) to chromosome 6q25.2–27, Am. J. Hum. Genet. 60 (1997) 588–596.

[38] T. Kitada, S. Asakawa, N. Hattori, H. Matsumine, Y. Yamamura, S. Minoshima, M. Yokochi, Y. Mizuno, N. Shimizu, Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism, Nature 392 (1998) 605–608.

[39] P.J. Kahle, U. Leimer, C. Haass, Does failure of parkin-mediated ubiquitination cause juvenile parkinsonism? Trends Biochem. Sci. 25 (2000) 524–527.

[40] N. Abbas, C.B. Lücking, S. Ricard, A. Dürr, V. Bonifati, G. De Michele, S. Bouley, J.R. Vaughan, T. Gasser, R. Marconi, E. Broussolle, C. Brefel-Courbon, B.S. Harhangi, B.A. Oostra, E. Fabrizio, G.A. Böhme, L. Pradler, N.W. Wood, A. Filla, G. Meco, P. Denefle, Y. Agid, A. Brice, The French Parkinson's Disease Genetics Study Group and the European Consortium on Genetic susceptibility in Parkinson's disease. A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe, Hum. Mol. Genet. 8 (1999) 567–574.

[41] C.B. Lücking, A. Durr, V. Bonifati, J. Vaughan, G. De Michele, T. Gasser, B.S. Harhangi, P. Pollak, A-M. Bonnet, D. Nichol, M.D. Mari, R. Marconi, E. Broussolle, O. Rascol, M. Rosier, I. Arnould, B.A. Oostra, M.M.B. Breteler, A. Filla, G. Meco, P. Denefle, N.W. Wood, Y. Agid, A. Brice, Association between early-onset Parkinson's disease and mutations in the parkin gene, N. Engl. J. Med. 342 (2000) 1560–1567.

[42] H. Shimura, N. Hattori, S. Kubo, Y. Mizuno, S. Asakawa, S. Minoshima, N. Shimizu, K. Iwai, T. Chiba, K. Tanaka, T. Suzuki, Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase, Nat. Genet. 25 (2000) 302–305.

[43] Y. Imai, M. Soda, R. Takahashi, Parkin suppresses unfolded protein stress-induced cell death through its E3 ubiquitin-protein ligase activity, J. Biol. Chem. 275 (2000) 35661–35664.

[44] Y. Zhang, J. Gao, K.K.K. Chung, H. Huang, V.L. Dawson, T.M. Dawson, Parkin functions as an E2-dependent ubiquitin-protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 13354–13359.

[45] I. Marin, A. Ferrus, Comparative genomics of the RBR family, including the Parkinson's disease-related gene parkin and the genes of the ariadne subfamily, Mol. Biol. Evol. 19 (2002) 2039–2050.

[46] J. Niwa, S. Ishigaki, M. Doyu, K. Kato, T. Suzuki, K. Tanaka, G. Sobue, A novel human RING-finger/IBR family protein, Dorfin, resides in centrosome and mediates ubiquitin ligase activity, Biochem. Biophys. Res. Commun. 281 (2001) 706–713.

[47] J. Niwa, S. Ishigaki, N. Hishikawa, M. Yamamoto, M. Doyu, S. Murata, K. Tanaka, N. Taniguchi, G. Sobue, Orfin ubiquitylates mutant SOD1 and prevents mutant SOD1-mediated neurotoxicity, J. Biol. Chem. 277 (2002) 36793–36798.

[48] T. Ito, J. Niwa, N. Hishikawa, S. Ishigaki, M. Doyu, G. Sobue, J Dorfin localizes to Lewy bodies and ubiquitylates synphilin-1, Biol. Chem. 278 (2003) 29106–29114.

[49] K. Yamanaka, H. Ishikawa, Y. Megumi, F. Tokunaga, M. Kanie, T.A. Rouault, I. Morishima, N. Minato, K. Ishimori, K. Iwai, Identification of the ubiquitin-protein ligase that recognizes oxidized IRP2, Nat. Cell Biol. 5 (2003) 336–340.

[50] A.Y. Nikolaev, M. Li, N. Puskas, J. Qin, W. Gu, Parc: a cytoplasmic anchor for p53, Cell 112 (2003) 29–40.

[51] L. Terreni, E. Calabrese, A.M. Calella, G. Forloni, C. Mariani, New mutation (R42P) of the parkin gene in the ubiquitin-like domain associated with parkinsonism, Neurology 56 (2001) 463–466.

[52] H. Shimura, M.G. Schlossmacher, N. Hattori, M.P. Frosch, A. Trockenbacher, R. Schneider, Y. Mizuno, K.S. Kosik, D.J. Selkoe, Ubiquitination of a new form of α-synuclein by parkin from human brain: implications for Parkinson's disease, Science 293 (2001) 263–269.

[53] M.F. Kleijnen, A.H. Shih, P. Zhou, S. Kumar, R.E. Soccio, N.L. Kedersha, G. Gill, P.M. Howley, The hPLIC proteins may provide a link between the ubiquitination machinery and the proteasome, Mol. Cell 6 (2000) 409–419.

[54] H. Kawahara, M. Kasahara, A. Nishiyama, K. Ohsumi, T. Goto, T. Kishimoto, Y. Saeki, H. Yokosawa, N. Shimbara, S. Murata, T. Chiba, T. Suzuki, K. Tanaka, Developmentally regulated, alternative splicing of the Rpn10 gene generates multiple forms of 26S proteasomes, EMBO J. 19 (2000) 4144–4153.

[55] K.J. Walters, M.F. Kleijnen, A.M. Goh, G. Wagner, P.M. Howley, Structural studies of the interaction between ubiquitin family proteins and proteasome subunit S5a, Biochemistry 41 (2002) 1767–1777.

[56] E. Sakata, Y. Yamaguchi, E. Kurimoto, J. Kikuchi, S. Yokoyama, S. Yamada, H. Kawahara, H. Yokosawa, N. Hattori, Y. Mizuno, K. Tanaka, K. Kato, Parkin binds the Rpn10 subunit of 26S proteasomes through its ubiquitin-like domain, EMBO Rep. 4 (2003) 301–306.

[57] M. Farrer, P. Chan, R. Chen, L. Tan, S. Lincoln, D. Hernandez, L. Forno, K. Gwinn-Hardy, L. Petrucelli, J. Hussey, A. Singleton, C. Tanner, J. Hardy, J.W. Langston, Lewy bodies and parkinsonism in families with parkin mutations, Ann. Neurol. 50 (2001) 293–300.

[58] M.R. Cookson, Neurodegeneration: how does parkin prevent Parkinson's disease? Curr. Biol. 13 (2003) R522–R524.

[59] B.I. Giasson, V.M. Lee, Are ubiquitination pathways central to Parkinson's disease? Cell 114 (2003) 1–8.

[60] K. Tanaka, T. Suzuki, T. Chiba, H. Shimura, N. Hattori, Y. Mizuno, Parkin is linked to the ubiquitin pathway, J. Mol. Med. 79 (2001) 482–494.

[61] S. Murata, Y. Minami, M. Minami, T. Chiba, K. Tanaka, CHIP is a chaperone-dependent E3 ligase that ubiquitylates unfolded protein, EMBO Rep. 2 (2001) 1133–1138.

[62] Y. Imai, M. Soda, S. Hatakeyama, T. Akagi, T. Hashikawa, K. Nakayama, R. Takahashi, CHIP is associated with Parkin, a gene responsible for familial Parkinson's disease, and enhances its ubiquitin ligase activity, Mol. Cell 10 (2002) 55–67.

[63] Y.C. Tsai, P.S. Fishman, N.V. Thakor, G.A. Oyler, Parkin facilitates the elimination of expanded polyglutamine proteins and leads to preservation of proteasome function, J. Biol. Chem. 278 (2003) 22044–22055.

[64] J.F. Staropoli, C. McDermott, C. Martinat, B. Schulman, E. Demireva, A. Abeliovich, Parkin is a component of an SCF-like ubiquitin ligase complex and protects postmitotic neurons from kainite excitotoxicity, Neuron 37 (2003) 735–749.

[65] Y. Imai, M. Soda, H. Inoue, N. Hattori, Y. Mizuno, R. Takahashi, An unfolded putative transmembrane polypeptide, which can lead to endoplasmic reticulum stress, is a substrate of parkin, Cell 105 (2001) 891–902.

[66] Y. Yang, I. Nishimura, Y. Imai, R. Takahashi, B. Lu, Parkin suppresses dopaminergic neuron-selective neurotoxicity induced by Pael-R in *Drosophila*, Neuron 37 (2003) 911–924.

[67] A.B. West, F. Gonzalez-de-Chavez, K. Wilkes, C. O'Farrell, M.J. Farrer, Parkin is not regulated by the unfolded protein response in human neuroblastoma cells, Neurosci. Lett. 341 (2003) 139–142.

[68] L.I. Golbe, Alpha-synuclein and Parkinson's disease, Adv. Neurol. 91 (2003) 165–174.

[69] C.W. Liu, M.J. Corboy, G.N. DeMartino, P.J. Thomas, Endoproteolytic activity of the proteasome, Science 299 (2003) 408–411.

[70] L. Petrucelli, C. O' Farrell, P.J. Lockhart, M. Baptista, K. Kehoe, L. Vink, P. Choi, B. Wolozin, M. Farrer, J. Hardy, M.R. Cookson, Parkin protects against the toxicity associated with mutant alpha-synuclein: proteasome dysfunction selectively affects catecholaminergic neurons, Neuron 36 (2002) 1007–1019.

[71] K.K. Chung, Y. Zhang, K.L. Lim, Y. Tanaka, H. Huang, J. Gao, C.A. Ross, V.L. Dawson, T.M. Dawson, Parkin ubiquitinates the α-

synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease, Nat. Med. 7 (2001) 1144–1150.
[72] Y. Zhang, J. Gao, K.K. Chung, H. Huang, V.L. Dawson, T.M. Dawson, Parkin functions as an E2-dependent ubiquitin-protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 13354–13359.
[73] Y. Ren, J. Zhao, J. Feng, Parkin binds to α/β tubulin and increases their ubiquitination and degradation, J. Neurosci. 23 (2003) 3316–3324.
[74] O. Corti, C. Hampe, H. Koutnikova, F. Darios, S. Jacquier, A. Prigent, J.C. Robinson, L. Pradier, M. Ruberg, M. Mirande, E. Hirsch, T. Rooney, A. Fournier, A. Brice, The p38 subunit of the aminoacyl-tRNA synthetase complex is a Parkin substrate: linking protein biosynthesis and neurodegeneration, Hum. Mol. Genet. 12 (2003) 1427–1437.
[75] Y. Mizuno, H. Yoshino, S. Ikebe, N. Hattori, T. Kobayashi, S. Shimoda-Matsubayashi, H. Matsumine, T. Kondo, Mitochondrial dysfunction in Parkinson's disease, Ann. Neurol. 44 (1998) S99–S109.
[76] J.C. Greene, A.J. Whitworth, I. Kuo, L.A. Andrews, M.B. Feany, L.J. Pallanck, Mitochondrial pathology and apoptotic muscle degeneration in *Drosophila* parkin mutants, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 4078–4083.
[77] J-M. Itier, P. Ibáñez, M.A. Mena, N. Abbas, C. Cohen-Salmon, G.A. Bohme, M. Laville, J. Pratt, O. Corti, L. Pradier, G. Ret, C. Joubert, M. Periquet, F. Araujo, J. Negroni, M.J. Casarejos, S. Canals, R. Solano, A. Serrano, E. Gallego, M. Sánchez, P. Denèfle, J. Benavides, G. Tremp, T.A. Rooney, A. Brice, J.G. de Yébenes, Parkin gene inactivation alters behaviour and dopamine neurotransmission in the mouse, Hum. Mol. Genet. (2003) 2277–2291.
[78] M.R. Cookson, Pathways to Parkinsonism, Neuron 37 (2003) 7–10.

nature
structural &
molecular biology

Structural basis of sugar-recognizing ubiquitin ligase

Tsunehiro Mizushima[1,2], Takeshi Hirao[3], Yukiko Yoshida[2,4], Soo Jae Lee[5], Tomoki Chiba[2], Kazuhiro Iwai[4,6], Yoshiki Yamaguchi[3,4], Koichi Kato[3,4,7], Tomitake Tsukihara[5] & Keiji Tanaka[2]

Reprinted from nature structural & molecular biology, volume 11, april 2004

nature
structural &
molecular biology

Structural basis of sugar-recognizing ubiquitin ligase

Tsunehiro Mizushima[1,2], Takeshi Hirao[3], Yukiko Yoshida[2,4], Soo Jae Lee[5], Tomoki Chiba[2], Kazuhiro Iwai[4,6], Yoshiki Yamaguchi[3,4], Koichi Kato[3,4,7], Tomitake Tsukihara[5] & Keiji Tanaka[2]

SCF^{Fbs1} is a ubiquitin ligase that functions in the endoplasmic reticulum (ER)-associated degradation pathway. Fbs1/Fbx2, a member of the F-box proteins, recognizes high-mannose oligosaccharides. Efficient binding to an N-glycan requires di-*N*-acetylchitobiose (chitobiose). Here we report the crystal structures of the sugar-binding domain (SBD) of Fbs1 alone and in complex with chitobiose. The SBD is composed of a ten-stranded antiparallel β-sandwich. The structure of the SBD–chitobiose complex includes hydrogen bonds between Fbs1 and chitobiose and insertion of the methyl group of chitobiose into a small hydrophobic pocket of Fbs1. Moreover, NMR spectroscopy has demonstrated that the amino acid residues adjoining the chitobiose-binding site interact with the outer branches of the carbohydrate moiety. Considering that the innermost chitobiose moieties in N-glycans are usually involved in intramolecular interactions with the polypeptide moieties, we propose that Fbs1 interacts with the chitobiose in unfolded N-glycoprotein, pointing the protein moiety toward E2 for ubiquitination.

So far, numerous studies have emphasized the physiological importance of the ubiquitin- and proteasome-mediated proteolytic pathway[1]. The ubiquitination reaction is catalyzed by an elaborate cascade system, consisting of activating (E1), conjugating (E2) and ligating (E3) enzymes[1,2]. Of these enzymes, E3 enzymes are considered to exist as molecules with a large diversity and to have a principal role in the selection of target proteins for ubiquitination in a temporally and spatially regulated fashion[3].

One of the best-characterized E3 enzymes is the SCF complex (composed of Skp1, Cul1, Roc1 (also called Rbx1) and an F-box protein), which regulates degradation of a broad range of cellular proteins[4]. The F-box proteins consist of an F-box domain that binds to Skp1, and various C-terminal substrate recognition regions, which are subclassified into a family of proteins named Fbw and Fbl that contain WD40-repeat and leucine-rich repeat (LRR) domains, respectively[5]. In addition, the remaining groups have been provisionally classified as Fbx proteins, which show no homology to any other known proteins[5]. However, we recently discovered a third category of the F-box protein family named Fbs (F-box sugar recognition)/FBG[6], consisting of at least five structurally related proteins including Fbs1 (named originally as Fbx2)[7] and Fbs2/Fbx6b (ref. 8). Fbw and Fbl proteins usually recognize the phosphorylation status of the substrate, and the tertiary structures of some of these proteins, such as Fbw1/βTrCP, Fbw7/Cdc4 and Fbl1/Skp2, have been determined by X-ray crystallography, providing valuable information for determining the molecular recognition mechanisms of target proteins[9–12]. However, the molecular basis underlying the ability of Fbs proteins to recognize target glycoproteins remains to be clarified.

Eukaryotic cells have an abundant and diverse repertoire of N-linked oligosaccharide structures, but the role of N-glycosylation of the proteins remains largely unknown. N-glycans have recently been shown to have an important role in glycoprotein transport and sorting[13,14]. N-glycoproteins are also subjected to 'quality control,' in which aberrant proteins are distinguished from properly folded proteins and retained in the ER[15]. When the improperly folded or incompletely assembled proteins fail to restore their functional states, they are degraded by the ER-associated degradation (ERAD) system, which involves retrograde transfer of proteins from the ER to the cytosol and subsequent degradation mediated by ubiquitin and proteasomes[16]. Recently, we identified the SCF^{Fbs1} as an E3 ubiquitin–ligase complex that ubiquitinates N-linked glycoproteins, serving to clear these glycoproteins in the cytosol of the cell[7]. Fbs1 recognizes N-linked high-mannose oligosaccharides, especially the internal diacetylchitobiose structure[8]. However, the molecular mechanism of the recognition of N-glycans by Fbs1 is unknown at present. To understand the molecular basis of the interaction between Fbs1 and N-glycans, we conducted crystal structural analyses of the SBD of Fbs1 and its complex with chitobiose.

RESULTS

Overall structure of the SBD in Fbs1

The structure of the SBD of Fbs1, as determined at a resolution of 2.0 Å (Table 1), is an ellipsoid composed of a ten-stranded antiparallel β-sandwich with two α-helices (Fig. 1a,b). This structure is

[1]Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan. [2]Tokyo Metropolitan Institute of Medical Science, Bunkyo-ku, Tokyo 113-8613, Japan. [3]Department of Structural Biology and Biomolecular Engineering, Graduate School of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabe-dori, Mizuho-ku, Nagoya 467-8603, Japan. [4]Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency, Saitama 332-0012, Japan. [5]Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan. [6]Department of Molecular Cell Biology, Graduate School of Medicine, Osaka City University, Osaka 545-8585, Japan. [7]Genomic Sciences Center, RIKEN Yokohama Institute, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. Correspondence should be addressed to K.T. (tanakak@rinshoken.or.jp).

Published online 29 February 2004; doi:10.1038/nsmb732

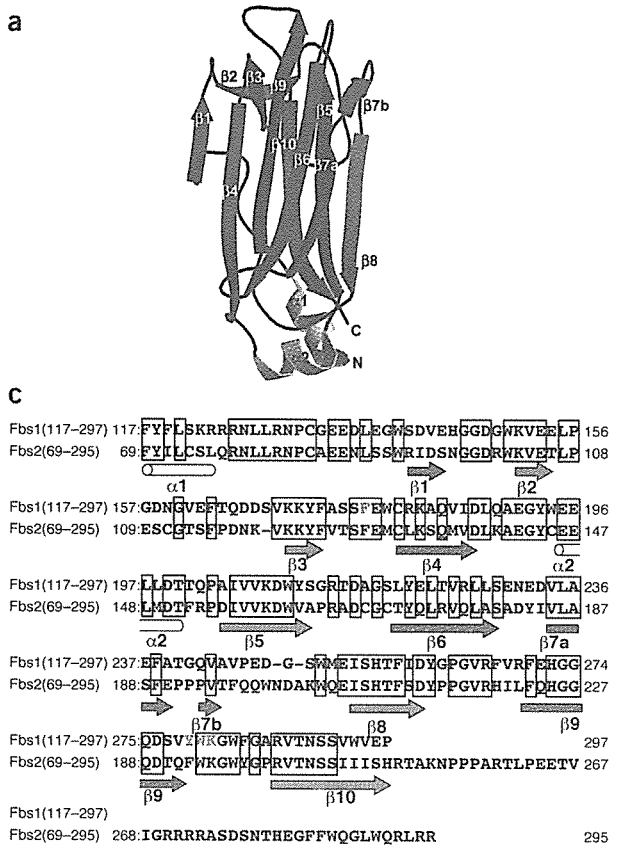

Figure 1 Tertiary structure of SBD in Fbs1. (**a**) Overall structure of SBD of Fbs1 shown as a ribbon diagram. β-strands belonging to S1 and S2 are blue and red, respectively. Loops and helices are black and yellow, respectively. (**b**) A topology diagram of SBD. The α-helices are yellow cylinders labeled α1 and α2. The β-strands are arrows labeled β1–β10. The left and right forms of β-strands correspond to S1 and S2, respectively, as in **a**. N and C, N and C termini, respectively. (**c**) Amino acid sequences of SBD in Fbs1 and corresponding region of Fbs2. Amino acid residues are numbered in the N-to-C direction, for example, from position 117 to position 297 (C-terminal end) of Fbs1, and from 69 to 295 (C-terminal end) of Fbs2. Identical residues are boxed. Secondary structure elements are colored as **a**. Substrate-binding residues are red characters.

completely different from the folds of substrate-binding regions of F-box domains so far reported, including the WD40-repeat domain of CDC4 (ref. 11) and the LRR domain of Skp2 (ref. 12). Strands β1, β4, β6, β7 and β9 form one β-sheet, whereas the other β-sheet consists of strands β2, β3, β5, β8 and β10. These two sheets are named the S1 and S2 sheets, respectively (**Fig. 1a–c**). Strand β7, which is located at one edge of the S1 sheet, is composed of two segments (β7a and β7b) separated by a bent structure. The two α-helices (α1 and α2) lie at one end of the β-sandwich; the α1 helix is at the N terminus and α2 helix is in the loop between β4 and β5. Comparisons of the SBD structure described here with the Protein Data Bank using the Dali server[17] revealed that the SBD is structurally similar to certain lectins, such as the galectin-3 carbohydrate recognition domain[18] and second family 4 carbohydrate-binding modules of xylanase 10A (ref. 19), with r.m.s. deviation values of 2.8 and 2.8 Å, respectively. Indeed, both xylanase and galectin-3 domains are composed of 11-stranded antiparallel β-sandwiches, consisting of 5- and 6-stranded β-sheets, respectively, the overall structures of which resemble that of the SBD, although, in contrast to the SBD, they lack α-helices. Comparing their primary structures, the SBD shows considerable homology to the carbohydrate-binding domain of xylanase 10A, exhibiting amino acid identity of ~20%, whereas no obvious sequence homology was found between SBD and the carbohydrate recognition of galectin-3.

The sugar-binding site

Next, we analyzed the structure of the SBD in complex with chitobiose. We could not obtain SBD crystals in the presence of chitobiose that diffracted to high resolution, so we introduced a C132A mutation in the SBD. The C132A SBD cocrystallized with chitobiose, and the structure of the C132A SBD–chitobiose complex was determined at 2.4 Å resolution (**Fig. 2a** and **Table 1**). The r.m.s. deviation values between SBD and the C132A mutant–chitobiose complex were 0.53 Å and 0.83 Å for the main chain and all atoms, respectively, indicating that the cysteine mutation hardly perturbs the conformation of the wild type SBD. It was also confirmed that this mutation did not alter the interaction and ubiquitination activities of glycoproteins (see **Fig. 3b**).

Chitobiose was clearly located in the difference electron density map (**Fig. 2a**), with *B*-factors as low as those of protein atoms around the chitobiose. The chitobiose in the SBD–chitobiose complex exhibited a *trans* conformation with respect to the *N*-acetyl groups, similar to the structures of a large number of N-glycan-binding proteins[20]. The bound chitobiose formed an intramolecular hydrogen bond between O6 of one GlcNAc(A) residue and O3 of the other GlcNAc(B) residue (**Fig. 2b,c**). The sugar-binding surface consists of two loops: L1 connects strands β3 and β4, and L2 is between strands β9 and β10. The GlcNAc(A) residue stacks on the aromatic ring of Trp280, as is often found in protein-carbohydrate interactions[21,22]. The GlcNAc-Trp280 stacking is stabilized by hydrogen bonds mediated by a water molecule between the O7 of the GlcNAc(A) and Nε1 of Trp280 (**Fig. 2c**), as well as a hydrogen bond between O6 of the GlcNAc(A) and the carbonyl oxygen atom of Lys281. The other GlcNAc(B) residue inserts the methyl group of its *N*-acetyl moiety into a small hydrophobic pocket surrounded by side chains of Phe177, Tyr279 and Lys281 (**Fig. 2a,b**) and forms a hydrogen bond between its hydroxyl and the main chain N atom of Lys281 (**Fig. 2c**). The orientation of the phenyl group of Tyr279 is stabilized by a hydrogen bond with the carboxyl group of Glu178. Upon N-glycoprotein uptake, the SBD swings the side chain of Lys281 (data not shown) and shields the methyl group from the molecular surface. The hydroxyl groups of the chitobiose almost exactly replace two binding site water molecules, which form a hydrogen bond to the backbone O and N of Lys281, respectively (**Fig. 2b,c**).

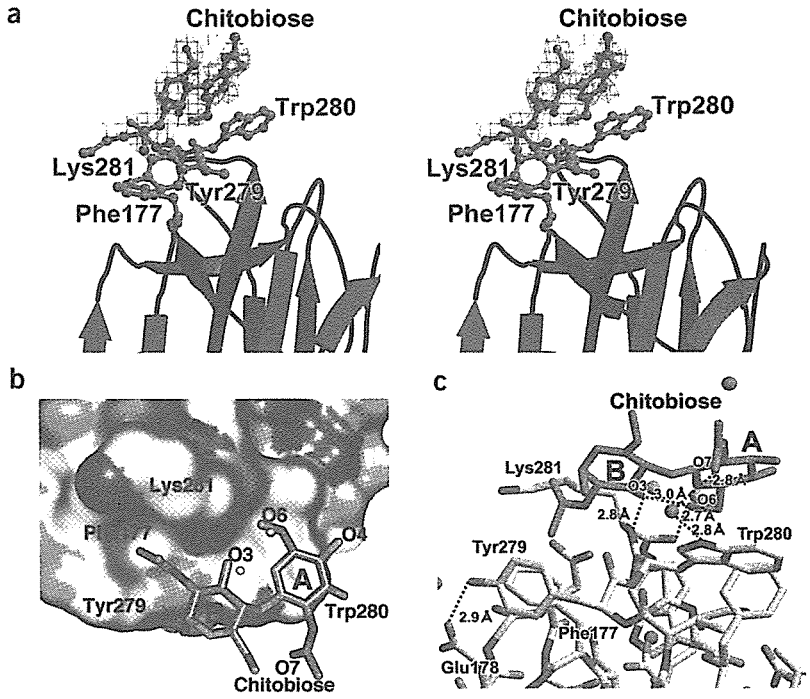

Figure 2 Structure of SBD in complex with chitobiose. (**a**) Stereo view of the difference-density map ($F_o - F_c$ with phase from the Fbs1 model) of binding chitobiose, contoured at 2.1 σ, modeled into the electron density. β-strands belonging to S1 and S2 are blue and red, respectively. Loops are black. The bound chitobiose is orange, and the residues involved in the substrate binding (FYWK, see **Fig. 1c**) are green. (**b**) Molecular surface representation of the chitobiose-binding region. The bound chitobiose is shown in ball-and-stick representation. Two GlcNAc residues are represented by A and B. Cyan spheres are two water molecules of wild type SBD that are fixed on the molecular surface through hydrogen bonds with the backbone N and O of Lys281, respectively. These water molecules are replaced by O3 and O6 of the chitobiose upon formation of the SBD–chitobiose complex. (**c**) Stick representation of the amino acids involved in binding. Hydrogen bonds are dashed lines. Oxygen and nitrogen are red and blue, respectively. Symbols of two water molecules are as in **b**.

To verify that the crystal structure accurately represents the complex formed in solution, we introduced point mutations into the residues in the pocket, and examined the *in vitro* activities in binding the ribonuclease B (RNase B) carrying a high-mannose oligosaccharide (**Fig. 3a**). Indeed, F177A, Y279A and W280A mutations reduced binding to the RNase B, whereas the K281A mutation had no effect on the binding (**Fig. 3a**, left panel). We next tested the *in vivo* activities of these mutants in binding the precursor of integrin β1, one of the *in vivo* Fbs1 targets[7] that contains high-mannose oligosaccharides. Consistent with the *in vitro* results, F177A, Y279A and W280A, but not K281A, failed to bind integrin β1 (**Fig. 3a**, right panel). These results suggest that Phe177, Tyr279 and Trp280, located in the hydrophobic pocket at the edge of the β-sandwich, are important for interaction with chitobiose in the high-mannose oligosaccharides. In contrast, although comparison of the structure of the SBD alone and that of the SBD–chitobiose complex indicated that the side chain of Lys281 underwent a conformational change upon ligand binding, the *in vivo* and *in vitro* binding studies suggested that this conformational change was not essential for the recognition of oligosaccharides. Moreover, we examined the impact of these mutants on the ubiquitinating activities of the SCF onto GlcNAc-terminated fetuin (GTF) *in vitro* (**Fig. 3b**). SCF (Fbs1-W280A), which could not bind to N-glycans, failed to ubiquitinate GTF, whereas the ubiquitinating activities of the K281A mutant were retained. Taken together, these results indicate that the hydrophobic interactions between GlcNAc(A) residue and Trp280, and of GlcNAc(B) residue with the small hydrophobic pocket, are required for substrate recognition. In addition, the hydrogen bonds between the chitobiose and Fbs1 atoms (Nε1 of Trp280 and the carbonyl oxygen atom of Lys281) are involved in selective binding to chitobiose.

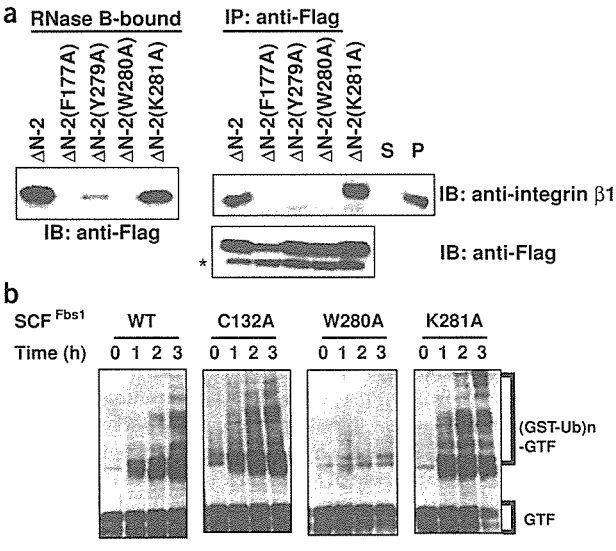

Figure 3 Residues required for interaction of Fbs1 and glycoproteins with high-mannose oligosaccharides. (**a**) Neuro2a cells were transfected with Flag-tagged Fbs1 (ΔN-2) and its listed derivatives. In pull-down assay (RNase B-bound), each ΔN-2–expressing WCE was incubated with RNase B–immobilized beads; bound proteins were eluted by 0.1 M of chitobiose and then analyzed by immunoblotting with an antibody to Flag (anti-Flag). In immunoprecipitation with anti-Flag (IP: anti-Flag), the Fbs1-binding proteins in the immune complex were analyzed by immunoblotting using the anti-Flag or anti-integrin β1 antibody. Asterisks indicate the light chain of IgG. (**b**) *In vitro* ubiquitination of GTF by the SCF^{Fbs1} E3–ligase system. The high-molecular-mass ubiquitinated GTF [(GST-Ub)n-GTF] was detected by immunoblotting with an antibody to fetuin.

NMR analyses of the SBD-sugar interactions

We have previously reported that Fbs1 shows higher affinity to $Man_{3-9}GlcNAc_2$ glycans than to chitobiose, and the number of mannose residues did not influence the affinity[8]. We conducted NMR spectroscopic analysis to determine the contribution of the outer

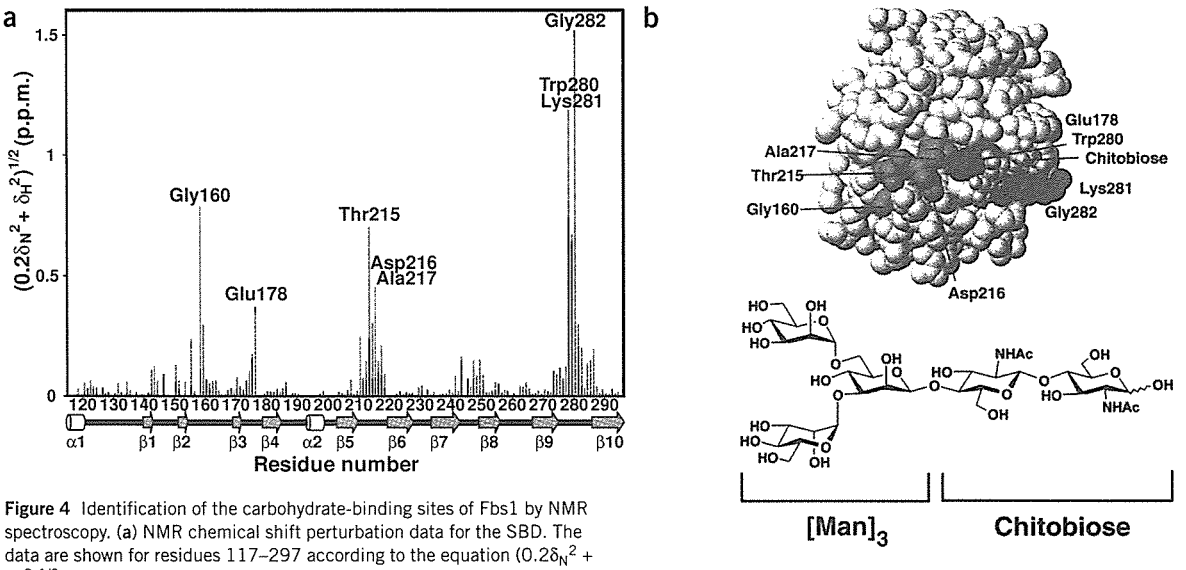

Figure 4 Identification of the carbohydrate-binding sites of Fbs1 by NMR spectroscopy. (**a**) NMR chemical shift perturbation data for the SBD. The data are shown for residues 117–297 according to the equation $(0.2\delta_N^2 + \delta_H^2)^{1/2}$, where δ_N and δ_H represent the change in nitrogen and proton chemical shifts on addition of chitobiose (blue) and $Man_3GlcNAc_2$ (red). Secondary structure elements for SBD are shown below the plot. (**b**) Mapping the perturbed residues of the SBD(97–297) ($(0.2\delta_N^2 + \delta_H^2)^{1/2} > 0.3$) upon binding to chitobiose (blue) and $Man_3GlcNAc_2$ (blue and red). The chitobiose molecule bound to SBD is yellow. The $Man_3GlcNAc_2$ structure is represented in the right panel.

mannose branches to the interaction with Fbs1, because it is generally not feasible to crystallize or to interpret the electron density of complexes of lectins with larger oligosaccharides. To identify the oligosaccharide-binding site of the SBD in solution, we compared the pattern of the chemical shift perturbation between chitobiose and$Man_3GlcNAc_2$ (Manα1→3[Manα1→6]Manβ1→4GlcNAcβ1→4GlcNAc) using the isotopically labeled SBD (**Fig. 4** and see **Supplementary Fig. 1** online). Chitobiose binding resulted in marked chemical shift perturbation for Glu178 in the L1 loop and for Trp280, Lys281 and Gly282 in the L2 loop; this is consistent with the X-ray structure of the C132A SBD–chitobiose complex. When $Man_3GlcNAc_2$ was used as a ligand, chemical shift perturbations were observed for Gly160, Thr215, Asp216 and Ala217 in addition to the residues perturbed by chitobiose, indicating that the outer branches of the carbohydrate moiety interact with the β5–β6 loop. In the crystal structure of the SBD–chitobiose complex, O4 of GlcNAc(A) is oriented toward these amino acid residues. Thus, the orientation of the sugar chain deduced from the NMR data is consistent with the expected orientation on inspection of the crystal structure. Although mannose residues of $Man_3GlcNAc_2$ seem to interact with the β5–β6 loop of Fbs1, almost no chemical shift perturbation of any consequence was observed upon addition of mannotriose (Manα1→3[Manα1→6]Man) alone (data not shown), suggesting that the affinity of $Man_3GlcNAc_2$ is dominated by the interaction with its chitobiose portion.

DISCUSSION

In the present study, we determined the SBD structure of Fbs1 and its complex with chitobiose. In general, most lectins bind nonreducing terminal sugar groups in the concave surface, which consists of several strands of β-sheets. The substrate-binding site of galectin-3 that forms a complex with *N*-acetyl-lactosamine is formed by β-strands[18]. The amino acids in these β-strands interact with the bound substrate through direct and water-mediated hydrogen bonds or through van der Waals contacts. In contrast, Fbs1 recognizes the inner chitobiose of N-linked high-mannose oligosaccharides by a specific binding surface located at one tip of the β-sandwich. To our knowledge, this is the first report of the sugar-binding mode of lectins, that is, interaction with the innermost portion of the carbohydrate moieties of glycoproteins. This is in marked contrast to the lectin chaperones in the ER, namely calnexin and calreticulin, which recognize nonreducing terminal glucose molecules of the high-mannose oligosaccharides expressed on their target glycoproteins.

In crystal structures of glycoproteins, electron densities of carbohydrate moieties are generally ambiguous because the carbohydrate moieties attached to the crystallized proteins do not necessarily exhibit a uniform chemical structure and may possess freedom of internal motion. In many cases, however, the innermost GlcNAc residue does provide unambiguous electron density because it is involved in interactions with the polypeptide moieties. It seems that those intramolecular interactions hamper the binding of Fbs1 to the chitobiose portions of glycoproteins as a result of steric hindrance in their native states. We propose that the novel sugar-binding mode embodied by Fbs1 is suitable for recognition of unfolded glycoproteins targeted in the ERAD system. RNase B used in the *in vitro* binding assay reveals an oligosaccharide that does not contact the polypeptide chain except at the covalent attachment point[23]. This glycoprotein could interact with Fbs1 even in the native form in the *in vitro* binding experiment, probably as a result of the exceptional freedom of the chitobiose portion of its carbohydrate moiety.

On inspection of the X-ray crystallographic and mutagenesis data, we conclude that the hydrophobic interaction between the GlcNAc(A) residue and Trp280, the hydrophobic interaction between the GlcNAc(B) residue and the small hydrophobic pocket composed of Phe177 and Tyr279, and the hydrogen bonds between the chitobiose and Fbs1 atoms, are essential for selective binding to this disaccharide moiety. Recently, we reported that Fbs2 also binds high-mannose oligosaccharides in a chitobiose-dependent manner, but the strength

Table 1 Data collection, phasing and refinement statistics

Data collection							
	Native	PCMBS	$NaAuCl_4$	$SmCl_3$	$OsCl_3$	$HgNO_3$	Complex
Space group	$P3_221$						$P4_32_12$
Resolution (Å)	2.0	2.5	2.7	2.5	2.7	2.2	2.4
Observations	192,305	100,271	80,395	59,131	79,742	96,045	93,981
Unique reflections	18,483	9,577	7,689	9,633	7,695	13,725	12,611
Completeness (%)[a]	99.9 (99.9)	99.8 (99.8)	99.8 (99.8)	99.8 (99.8)	99.8 (99.8)	98.6 (98.6)	99.8 (99.8)
Redundancy[a]	10.4 (10.1)	10.5 (10.6)	10.5 (10.6)	6.1 (6.2)	10.4 (10.5)	7.0 (6.9)	7.5 (7.2)
R_{sym} (%)[a,b]	9.5 (27.8)	12.5 (30.6)	9.9 (24.0)	13.5 (22.7)	14.3 (28.0)	14.4 (26.4)	8.5 (18.4)
I / σ[a]	4.4 (2.5)	4.4 (2.3)	6.0 (2.9)	3.9 (3.0)	3.7 (2.3)	3.2 (2.6)	6.4 (3.6)
MIRAS phasing							
Resolution (Å)	2.5	2.7	2.5	2.7	2.2		
Heavy atom sites	3	2	2	1	2		
Phasing power	1.37	0.89	0.76	1.71	3.31		
R_{cullis}[c]	0.8	0.84	0.88	0.69	0.46		

Refinement statistics		
	SBD	SBD C132A–chitobiose
Resolution (Å)	2.0	2.4
Reflections	17,470	11,950
Total atoms	1,602	1,562
R-factor (%)	16.2	20.0
R_{free} (%)	19.9	26.3
R.m.s. deviations		
Bond length (Å)	0.020	0.038
Bond angle (°)	1.8	2.9

[a]Values in parentheses are for the highest-resolution shell. [b]$R_{sym} = \Sigma_h \Sigma_j |I_{hj} - \langle I_h \rangle| / \Sigma_h \Sigma_j I_{hj}$, where h represents a unique reflection and j represents symmetry-equivalent indices. I is the observed intensity and $\langle I \rangle$ is the mean value of I. [c]$R_{cullis} = \Sigma ||F_{PH}| \pm |F_P|| - |F_H|| / \Sigma |F_{PH}| \pm |F_P|$.

of the glycoprotein-binding ability is weaker than that of Fbs1 (ref. 8). In Fbs2, the positions corresponding to Phe177, Tyr279 and Trp280 in Fbs1 are occupied by phenylalanine, phenylalanine, and tryptophan, respectively (**Fig. 1c**). The conservation of these residues suggests that the chitobiose-binding mode of Fbs2 is similar to that of Fbs1, and the reduced binding ability of Fbs2 may be attributed to the Tyr→Phe substitution in the chitobiose-binding pocket. Fbs1 interacts not only with chitobiose at the L1 and L2 loops but also with the outer mannose residues at the β5–β6 loop. Although binding of mannose residues to the β5–β6 loop seems to strengthen the binding affinity, mannose did not bind alone. This inner chitobiose-dependent interaction mode further restricts Fbs1 to interacting with native proteins carrying high-mannose oligosaccharides. The positions corresponding to Thr215 and Ala217, which are involved in the interaction with the mannose residues, are occupied by alanine and cysteine, respectively, in Fbs2 (**Fig. 1c**). This might be associated with the distinct affinity of Fbs1 and Fbs2 to N-glycans (ref. 8).

We attempted to model the full-length Fbs1 protein and its assembly in the SCF-E2 complex by positioning the N terminus of the SBD adjacent to the F-box domain of the reported SCF^{Skp2} complex[9]. This model places the chitobiose >50 Å from the expected position of the E2 protein that donates ubiquitin to the glycoprotein, consistent with modeling of other SCF complexes, such as SCF^{Skp2} (ref. 9) and SCF^{Cdc4} (ref. 11).

Fbs1 is a functionally unique molecule that recognizes the innermost position of N-glycans as a signal for denatured glycoproteins. Our results confirmed structurally that N-glycans act as a ubiquitination signal, thus providing new insights into the biological roles of sugar chains coupled to proteins within cells.

METHODS

Crystallization and structure determination. The SBD of murine Fbs1(117–297), with a molecular mass of 20.6 kDa, and its mutant protein C132A SBD were cloned into pET15b and expressed in *Escherichia coli*. Both proteins were purified by nickel affinity and gel filtration chromatography and concentrated to 20 mg ml^{-1}. All mutations were constructed by QuikChange mutagenesis kit (Stratagene), using synthetic oligonucleotides, and the sequences were verified in their entirety.

Crystals of SBD were grown at 25 °C by the hanging-drop vapor diffusion method. SBD crystals were grown from 1.7 M ammonium sulfate, 0.01 M nickel chloride, 0.1% (v/v) PEG400 and 0.1 M Tris-HCl, pH 8.5. The C132A SBD was cocrystallized with chitobiose (Seikagaku). C132A SBD crystals were prepared using 1.4 M sodium chloride, 1.7 M ammonium sulfate, 0.1 M PIPES, pH 7.0, and 30 mM chitobiose. SBD and its cysteine mutant crystals belonged to $P3_221$ and $P4_32_12$ space groups, with cell dimensions of $a = b = 62.4$ Å, $c = 117.2$ Å, and $a = b = 63.8$ Å, $c = 147.8$ Å, respectively. Heavy-atom soaks were done in crystallization buffer with saturated *p*-chloromercuribenzene sulfonate (PCMBS) (for 3 h), 10 mM $NaAuCl_4$ (15 h), 10 mM $SmCl_3$ (15 h), 10 mM $OsCl_3$ (15 h) and 10 mM $HgNO_3$ (15 h). Intensity data sets were collected on a Rigaku R-AXIS IV nickel-filtered double-mirror focused CuKα-radiation detector. A Rigaku RU-200 rotating-anode X-ray generator was operated at 40 kV and 100 mA. Data were processed with MOSFLM[24] and SCALA[25,26]. The structure of the SBD was determined by the multiple isomorphous replacement anomalous scattering (MIRAS) method. Phases were calculated with MLPHARE[26] to 2.2 Å. The initial electron density map was then refined by solvent flattening[27] and histogram mapping[28] using DM[26]. The initial model was constructed with Arp/WArp[29] and O[30]. The model was refined at a resolution of 2.0 Å with REFMAC[31]. Residues 43–55 of the SBD have high temperature factors and are presumably partially disordered.

Intensity data sets of the mutant protein were collected at 100 K. For cryoprotection, 20% (v/v) glycerol was added to the crystallizing solution. The

structure of C132A was determined by the molecular replacement technique using AMoRe[32] with the refined model of SBD. Structure refinement of the mutant SBD was guided by referring to the structure of the wild type. Refinement statistics of both structures are summarized in Table 1.

Binding assay. Neuro2a cells were transfected with Flag-tagged murine Fbs1 ΔN-2(95–297) and its derivatives that lack the region from the N terminus to the F-box domain[7] by lipofection (Lipofectamine Plus; Gibco BRL). At 48 h after transfection, the whole-cell extracts (WCEs), solubilized with TBS containing 0.5% (v/v) NP-40, were used for immunoprecipitation using mouse monoclonal anti-Flag (M2; Sigma), or pull-down assay using RNase B–immobilized resin, as described[7]. The (co-)immunoprecipitated proteins were detected by immunoblotting using mouse monoclonal antibodies anti-Flag (M2) and anti-integrin β1 (Transduction Laboratories).

***In vitro* ubiquitination assay.** Recombinant His-tagged human Ubc4 was produced in *E. coli*. Recombinant His-tagged mouse E1 (Uba1), His-tagged rat Fbs1 (ΔN-2) and each SCF^{Fbs1} (Flag-tagged human Skp1, human Cul1-HA/His-tagged rat Fbs1 derivatives, T7-tagged human Roc1) were produced by baculovirus-infected HiFive insect cells. Each SCF^{Fbs1} complex was obtained by simultaneously infecting four baculoviruses. These proteins were affinity-purified by a HiTrap HP column (Amersham Pharmacia Biotech), as described[7]. Each 1 μg of GTF was incubated in 50 μl of the reaction mixture containing ATP-regenerating system, 0.5 μg E1, 1 μg Ubc4 (E2), 2 μg SCF^{Fbs1}, 6.5 μg recombinant GST-ubiquitin and NEDD8 system[33] at 30 °C. After the reaction was terminated by adding 25 μl of 3× SDS-PAGE sample buffer, the proteins in 8 μl of the boiled supernatants were separated with 5–20% (w/v) SDS-PAGE, and the high-molecular-mass ubiquitinated proteins were detected by immunoblotting with anti-fetuin (Chemicon International).

NMR spectroscopy. The DNA fragment encoding residues 97–297 of Fbs1was inserted into the pGEX-6P-1 plasmid vector (Amersham Biosciences) with a N-terminal GST moiety. The protein was expressed in *E. coli* BL21(DE3) CodonPlus strain (Stratagene) in M9 minimal medium with appropriate $[^{15}N]NH_4Cl$ and $[^{13}C_6]$glucose. GST-fusion protein was purified from cell lysates on a glutathione-Sepharose column. The fusion protein was cleaved by incubation with PreScission protease (Amersham Biosciences), and GST was removed by loading a second glutathione-Sepharose column. The protein was further purified using a Superose12 gel filtration column.

NMR samples contained 0.1–1.0 mM SBD(97–297) in 10 mM NaH_2PO_4/Na_2HPO_4, pH 6.5, 50 mM KCl and 10 mM DTT. For chemical shift perturbation experiments, a ten-fold molar excess of chitobiose, ten-fold molar excess of Manα1′3(Manα1′6)Man (Calbiochem) or one molar equivalent of $Man_3GlcNAc_2$ (Sigma) was added to the protein solution. NMR spectra were acquired at 30 °C on Bruker DMX500 and DRX800 spectrometers. The ^1H, ^{15}N and ^{13}C resonances of the backbone were assigned using a standard set of double- and triple-resonance experiments[34].

Coordinates. The atomic coordinates of the SBD and C132A SBD–chitobiose complex have been deposited in the Protein Data Bank (accession codes 1UMH and 1UMI, respectively).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank Y. Wada for her help in the preparation of isotopically labeled recombinant proteins and E. Adachi for expert technical assistance. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science and Culture of Japan (K.T. and K.K.), Japan Society for the Promotion of Science and Mizutani Foundation for Glycoscience (K.K.), and in part by the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sport, Science and Technology of Japan (T.T.).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 6 October 2003; accepted 20 January 2004
Published online at http://www.nature.com/natstructmolbiol/

1. Hershko, A., Ciechanover, A. & Varshavsky, A. Basic Medical Research Award. The ubiquitin system. *Nat. Med.* **6**, 1073–1081 (2000).
2. Pickart, C.M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).
3. Weissman, A.M. Themes and variations on ubiquitylation. *Nat. Rev. Mol. Cell. Biol.* **2**, 169–178 (2001).
4. Deshaies, R.J. SCF and Cullin/Ring H2–based ubiquitin ligases. *Annu. Rev. Cell. Dev. Biol.* **15**, 435–467 (1999).
5. Winston, J.T., Koepp, D.M., Zhu, C., Elledge, S.J. & Harper, J.W. A family of mammalian F-box proteins. *Curr. Biol.* **9**, 1180–1182 (1999).
6. Ilyin, G.P. *et al.* A new subfamily of structurally related human F-box proteins. *Gene* **296**, 11–20 (2002).
7. Yoshida, Y. *et al.* E3 ubiquitin ligase that recognizes sugar chains. *Nature* **418**, 438–442 (2002).
8. Yoshida, Y. *et al.* Fbs2 is a new member of the E3 ubiquitin ligase family that recognizes sugar chains. *J. Biol. Chem.* **278**, 43877–43884 (2003).
9. Zheng, N. *et al.* Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**, 703–709 (2002).
10. Wu, G. *et al.* Structure of a β-TrCP1-Skp1–β-catenin complex: destruction motif binding and lysine specificity of the SCF(β-TrCP1) ubiquitin ligase. *Mol. Cell* **11**, 1445–1456 (2003).
11. Orlicky, S., Tang, X., Willems, A., Tyers, M. & Sicheri, F. Structural basis for phosphodependent substrate selection and orientation by the SCF^{Cdc4} ubiquitin ligase. *Cell* **112**, 243–256 (2003).
12. Schulman, B.A. *et al.* Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* **408**, 381–386 (2000).
13. Plemper, R.K. & Wolf, D.H. Retrograde protein translocation: ERADication of secretory proteins in health and disease. *Trends Biochem. Sci.* **24**, 266–270 (1999).
14. Fiedler, K. & Simons, K. The role of *N*-glycans in the secretory pathway. *Cell* **81**, 309–312 (1995).
15. Helenius, A. & Aebi, M. Intracellular functions of *N*-linked glycans. *Science* **291**, 2364–2369 (2001).
16. Ellgaard, L., Molinari, M. & Helenius, A. Setting the standards: quality control in the secretory pathway. *Science* **286**, 1882–1888 (1999).
17. Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
18. Seetharaman, J. *et al.* X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1-Å resolution. *J. Biol. Chem.* **273**, 13047–13052 (1998).
19. Simpson, P.J. *et al.* The solution structure of the CBM4-2 carbohydrate binding module from a thermostable *Rhodothermus marinus* xylanase. *Biochemistry* **41**, 5712–5719 (2002).
20. Petrescu, A.J., Petrescu, S.M., Dwek, R.A. & Wormald, M.R. A statistical analysis of *N*- and *O*-glycan linkage conformations from crystallographic data. *Glycobiology* **9**, 343–352 (1999)
21. Vyas, N.K. Atomic features of protein–carbohydrate interactions. *Curr. Opin. Struct. Biol.* **1**, 732–740 (1991).
22. Quiocho, F.A. Carbohydrate-binding proteins: tertiary structures and protein–sugar interactions. *Annu. Rev. Biochem.* **55**, 287–315 (1986).
23. Williams, R.L., Greene, S.M. & McPherson, A. The crystal structure of ribonuclease B at 2.5-Å resolution. *J. Biol. Chem.* **262**, 16020–16031 (1987).
24. Leslie, A.G.W. Molecular data processing. *Crystallographic Computing* **5**, 50–61 (1991).
25. Kabsch, W. Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallogr.* **21**, 916–924 (1988).
26. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
27. Wang, B.C. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115**, 90–112 (1985).
28. Zhang, K.Y.J. & Main P. The use of Sayre's equation with solvent flattening and histogram matching for phase extension and refinement of protein structures. *Acta Crystallogr. A* **46**, 377–381 (1990).
29. Perrakis, A., Morris, R. & Lamzin, V.S. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6**, 458–463 (1999).
30. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47** (Part 2), 110–119 (1991).
31. Murshudov, B.W. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
32. Navaza, J. An automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
33. Kawakami, T. *et al.* NEDD8 recruits E2-ubiquitin to SCF E3 ligase. *EMBO J.* **20**, 4003–4012 (2001).
34. Clore, G.M. & Gronenborn, A.M. Multidimensional heteronuclear nuclear magnetic resonance of proteins. *Methods Enzymol.* **239**, 349–363 (1994).

The EMBO Journal (2004) 23, 1977–1986 | © 2004 European Molecular Biology Organization | All Rights Reserved 0261-4189/04
www.embojournal.org

THE EMBO JOURNAL

A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier

Masaaki Komatsu[1], Tomoki Chiba[1], Kanako Tatsumi[1], Shun-ichiro Iemura[2], Isei Tanida[3], Noriko Okazaki[4], Takashi Ueno[3], Eiki Kominami[3], Tohru Natsume[2] and Keiji Tanaka[1,*]

[1]Department of Molecular Oncology, Tokyo Metropolitan Institute of Medical Science, Bunkyo-ku, Tokyo, Japan, [2]National Institutes of Advanced Industrial Science and Technology, Biological Information Research Center (JBIRC), Kohtoh-ku, Tokyo, Japan, [3]Department of Biochemistry, Juntendo University School of Medicine, Bunkyo-ku, Tokyo, Japan and [4]Kazusa DNA Research Institute, Kazusa-Kamatari, Kisarazu, Chiba, Japan

Several studies have addressed the importance of various ubiquitin-like (UBL) post-translational modifiers. These UBLs are covalently linked to most, if not all, target protein(s) through an enzymatic cascade analogous to ubiquitylation, consisting of E1 (activating), E2 (conjugating), and E3 (ligating) enzymes. In this report, we describe the identification of a novel ubiquitin-fold modifier 1 (Ufm1) with a molecular mass of 9.1 kDa, displaying apparently similar tertiary structure, although lacking obvious sequence identity, to ubiquitin. Ufm1 is first cleaved at the C-terminus to expose its conserved Gly residue. This Gly residue is essential for its subsequent conjugating reactions. The C-terminally processed Ufm1 is activated by a novel E1-like enzyme, Uba5, by forming a high-energy thioester bond. Activated Ufm1 is then transferred to its cognate E2-like enzyme, Ufc1, in a similar thioester linkage. Ufm1 forms several complexes in HEK293 cells and mouse tissues, revealing that it conjugates to the target proteins. Ufm1, Uba5, and Ufc1 are all conserved in metazoa and plants but not in yeast, suggesting its potential roles in various multicellular organisms.
The EMBO Journal (2004) **23,** 1977–1986. doi:10.1038/sj.emboj.7600205; Published online 8 April 2004
Subject Categories: proteins
Keywords: Uba5; ubiquitin; ubiquitin fold; ubiquitin-like protein; Ufm1

Introduction

Protein modification plays a pivotal role in the regulation and expansion of genetic information. In the past two decades, a new type of post-translational protein-modifying system has been identified whose uniqueness is that protein(s) is used as a ligand, that is, modification of protein, by protein, and for protein. A typical system is the ubiquitylation, a modification system in which a single or multiple ubiquitin molecules are attached to a protein, which serves as a signaling player that controls a variety of cellular functions (Hershko and Ciechanover, 1998; Pickart, 2001). Protein ubiquitylation is catalyzed by an elaborate system highly regulated in the cells, which is catalyzed by a sequential reaction of multiple enzymes consisting of activating (E1), conjugating (E2), and ligating (E3) enzymes. E1, which initiates the reaction, forms a high-energy thiolester bond with ubiquitin via adenylation in an ATP-dependent manner. The E1-activated ubiquitin is then transferred to E2 in a thioester linkage. In some cases, E2 can directly transfer the ubiquitin to substrate proteins in an isopeptide linkage; however, E2s mostly requires the participation of E3 to achieve substrate-specific ubiquitylation reaction in the cells. E3s are defined as enzymes required for recognition of specific substrates for ubiquitylation, other than E1 and E2 (Varshavsky, 1997; Bonifacino and Weissman, 1998; Glickman and Ciechanover, 2002).

A set of novel molecules called ubiquitin-like proteins (UBLs) that have structural similarities to ubiquitin has been recently identified (Jentsch and Pyrowolakis, 2000). They are divided into two subclasses: type-1 UBLs, which ligate to target proteins in a manner similar, but not identical, to the ubiquitylation pathway, such as SUMO, NEDD8, and UCRP/ISG15, and type-2 UBLs (also called UDPs, ubiquitin-domain proteins), which contain ubiquitin-like structure embedded in a variety of different classes of large proteins with apparently distinct functions, such as Rad23, Elongin B, Scythe, Parkin, and HOIL-1 (Tanaka *et al*, 1998; Jentsch and Pyrowolakis, 2000; Yeh *et al*, 2000; Schwartz and Hochstrasser, 2003).

In this report, we describe a unique human UBL-type modifier named ubiquitin-fold modifier 1 (Ufm1) that is synthesized in a precursor form consisting of 85 amino-acid residues. We also identified the human activating (Uba5) and conjugating (Ufc1) enzymes for Ufm1. Prior to activation by Uba5, the extra two amino acids at the C-terminal region of the human proUfm1 protein are removed to expose Gly whose residue is necessary for conjugation to target molecule(s). Lastly, we show that the mature Ufm1 is conjugated to yet unidentified endogenous proteins, forming ~28, 38, 47, and 70 kDa complexes in human HEK293 cells and various mouse tissues.

Results

Identification of a novel protein-activating enzyme, Uba5

Our initial plan was to identify the molecule(s) that interacts with human Atg8p homolog GATE16, a type-1 UBL modifier required for autophagy (Klionsky and Emr, 2000; Ohsumi, 2001), using a yeast two-hybrid screening. Please note that the nomenclature of the autophagy-related genes was recently unified as *ATG* (Klionsky *et al*, 2003). Among several

*Corresponding author. Department of Molecular Oncology, The Tokyo Metropolitan Institute of Medical Science, 3-18-22 Honkomagome, Bunkyo-ku, Tokyo 113-8613, Japan. Tel.: +81 3 3823 2237; Fax: +81 3 3823 2237; E-mail: tanakak@rinshoken.or.jp

Received: 1 December 2003; accepted: 15 March 2004; published online: 8 April 2004

positive clones, we identified fragments of FLJ23251 (Figure 1A), which encodes a 404-amino-acid protein highly conserved in various multicellular organisms, such as *Homo sapiens*, *Caenorhabditis elegans*, *Drosophila melanogaster*, and *Arabidopsis thaliana*, but absent in yeasts (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*) (Figure 1B). The sequence of FLJ23251 in the region containing residues 72–229 is highly homologous to the corresponding regions in Uba1 (i.e., E1 for ubiquitin) and other E1-like proteins for UBLs including the ATP-binding motif (GXGXXG) (Figure 1A and B). We named this protein Uba5, because it is a member of the E1-like enzyme family. Uba5 also has a metal-binding motif conserved in other E1-like enzymes such as Uba2, Uba3, Uba4, and Atg7. Most of E1-like enzymes have an active site Cys residue within the conserved 10–20 amino-acid residues downstream from the metal-binding

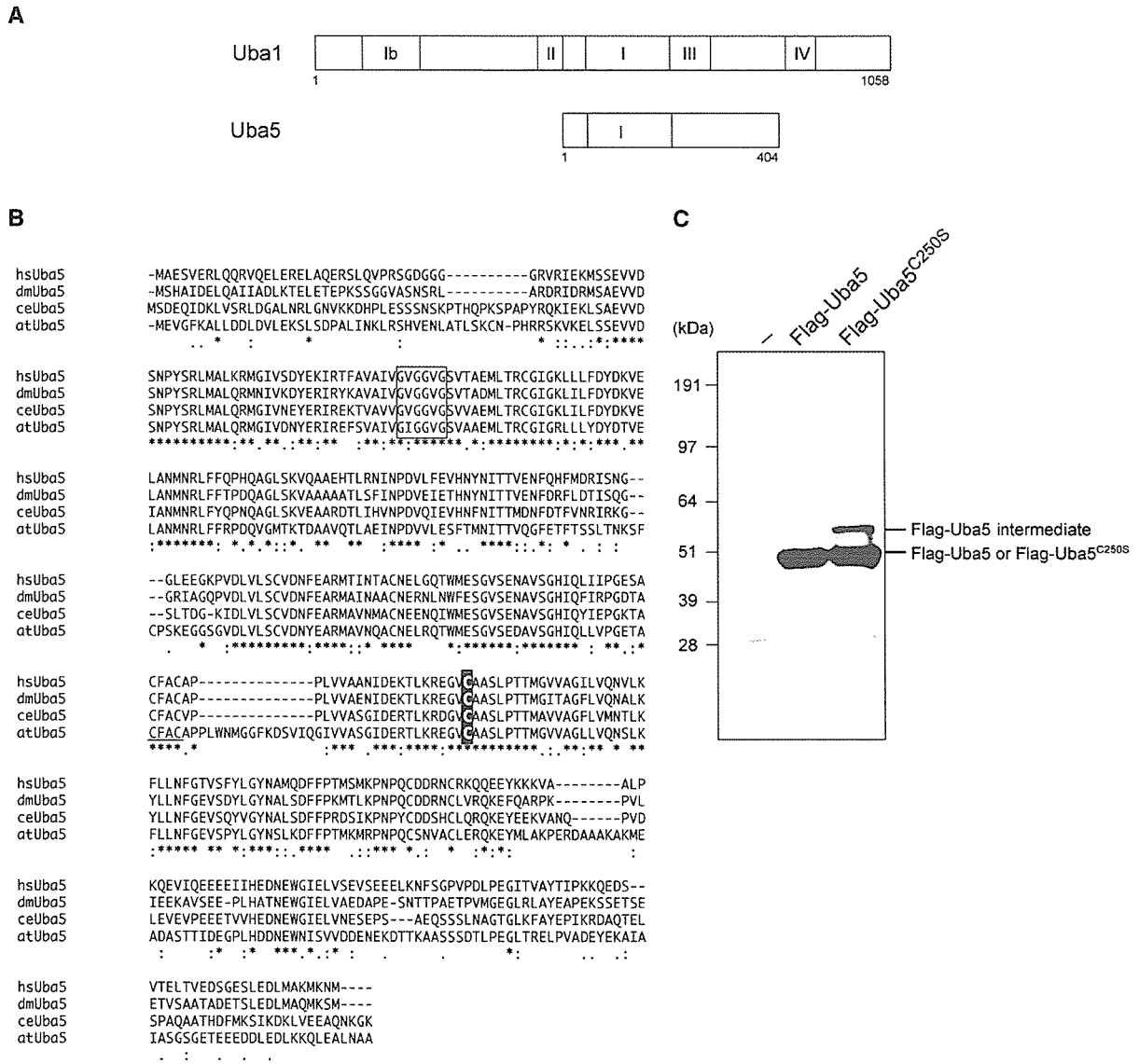

Figure 1 Uba5, a novel E1-like enzyme. (**A**) Schematic representation of Uba1 and Uba5 in *H. sapiens*. Uba1 is divided into several domains, including I, Ib, II, III, and IV boxes, which are conserved in other E1-like enzymes, and other regions without obvious similarity, described previously (Komatsu *et al*, 2001). Note that Uba5 is of a relatively small size and includes the box I and two other parts. The box I region of Uba1 (amino acids 459–611) has 48.4% similarity and 22.3% identity to amino acids 72–229 of Uba5, which includes the conserved ATP-binding motif (GXGXXG). The sequence of Uba5 is available from GenBank™ under the accession number AK026904. hs, *H. sapiens*; ce, *C. elegans*; dm, *D. melanogaster*; at, *A. thaliana*. (**B**) Sequence alignment of hsUba5 and its homologs of other species (dm, NM_132494; ce, NM_058847; at, NM_100414). The amino-acid sequence of hsUba5 is compared by the ClustalW program. *Asterisks*, identical amino acids; *single* and *double dots*, weakly and strongly similar amino acids, respectively, determined by the criteria of ClustalW program. *Open box* indicates an ATP-binding motif. The putative active site Cys residue is boxed in black. The metal-binding motif is underlined. (**C**) Identification of the intermediate linked to Uba5 in HEK293 cells. Both Uba5 and Uba5^{C250S}, in which the predicted active site Cys positioned at 250 was changed to Ser by site-directed mutagenesis, were tagged with Flag peptide at N-terminus, resulting in Flag-Uba5 and Flag-Uba5^{C250S}, respectively. Each Flag-Uba5 and Flag-Uba5^{C250S} was expressed in HEK293 cells. The cell lysates were subjected to SDS–PAGE and analyzed by immunoblotting with anti-Flag antibody.