Structure and Biochemical Function of a Prototypical Arabidopsis U-box Domain*

Received for publication, May 6, 2004, and in revised form, June 25, 2004
Published, JBC Papers in Press, June 30, 2004, DOI 10.1074/jbc.M405057200

Pernille Andersen‡, Birthe B. Kraglund‡, Addie N. Olsen‡, Flemming H. Larsen‡, Nam-Hai Chua‡, Flemming M. Poulsen‡,§ and Karen Skriver‡,¶

From the ‡Institute of Molecular Biology, University of Copenhagen, DK-1353 Copenhagen K, Denmark and the §Laboratory of Plant Molecular Biology, The Rockefeller University, New York, New York 10021

U-box proteins, as well as other proteins involved in regulated protein degradation, are apparently over-represented in Arabidopsis compared with other model eukaryotes. The Arabidopsis protein AtPUB14 contains a typical U-box domain followed by an Armadillo repeat region, a domain organization that is frequently found in plant U-box proteins. In vitro ubiquitination assays demonstrated that AtPUB14 functions as an E3 ubiquitin ligase with specific E2 ubiquitin-conjugating enzymes. The structure of the AtPUB14 U-box domain was determined by NMR spectroscopy. It adopts the βαβαβ fold of the Prp19p U-box and RING finger domains. In these proteins, conserved hydrophobic residues form a putative E2-binding cleft. By contrast, they contain no common polar E2 binding site motif. Two hydrophobic cores stabilize the AtPUB14 U-box fold, and hydrogen bonds and salt bridges interconnect the residues corresponding to zinc-coordinating residues in RING domains. Residues from a C-terminal α-helix interact with the core domain and contribute to stabilization. The Prp19p U-box lacks a corresponding C-terminal α-helix. Chemical shift analysis suggested that aromatic residues exposed at the N terminus and the C-terminal α-helix of the AtPUB14 U-box participate in dimerization. Thus, AtPUB14 may form a biologically relevant dimer. This is the first plant U-box structure to be determined, and it provides a model for studies of the many plant U-box proteins and their interactions. Structural insight into these interactions is important, because ubiquitin-dependent protein degradation is a prevalent regulatory mechanism in plants.

The ubiquitin proteolytic pathway plays an important role in regulated protein degradation (1). Proteins designated for degradation are covalently modified by attachment of a ubiquitin polymer and degraded by the 26 S proteasome. A ubiquitin-activating enzyme (E1) catalyzes ATP-dependent formation of a thioester bond between ubiquitin and itself and transfers the activated ubiquitin to a ubiquitin-conjugating enzyme (E2). Formation of an isopeptide bond between ubiquitin and a subunit is facilitated by a ubiquitin ligase (E3) that can bind both the E2-ubiquitin complex and the substrate. Members of the HECT and RING protein families are the best characterized E3 ligases, but recently U-box proteins have also been shown to function as E3s (2, 3). This may be the general function of U-box proteins, although they were initially suggested to function as ubiquitin chain assembly factors (E4s) (4). Recently, research on U-box proteins, especially the carboxyl terminus of Hsc70-interacting proteins (CHIP), has focused on their ability to interact with molecular chaperones and selectively ubiquitinate unfolded proteins. Thus ubiquitination can also function in protein quality control (5).

The U-box motif is a peptide chain that contains ~70 amino acid residues, with characteristics suggesting that it is a structural variant of the RING fold but lacks the signature zinc-binding amino acids of the RING domain (6). Recently, the first three-dimensional structure of a U-box domain was published (7). The structure of the yeast pre-mRNA splicing factor Prp19p U-box was determined by NMR spectroscopy and verified the structural similarity to the RING domain stabilized by an extensive hydrogen-bonding network as a replacement for the zinc binding sites of the RING domains. Only two U-box proteins are present in yeast, and typically six are present in animals (5). By contrast, Arabidopsis contains many U-box proteins (8). The largest class of Arabidopsis U-box proteins contains an ARAMIDILLO (ARM) repeat region linked to the U-box. ARM repeats are short leucine-rich protein-interacting motifs (9), first identified in the segment polarity protein, armadillo, from Drosophila melanogaster (10). Only a few plant U-box proteins have been characterized biochemically, and so far three-dimensional structures have not been available. The Brassica U-box protein ARC1 binds to the S-locus receptor kinase, the female determinant of pollen self-incompatibility (11), via its ARM repeats (12). It was suggested that ARC1 promotes ubiquitination and degradation of compatibility factors in the pistil leading to pollen rejection (13). Arabidopsis also contains a CHIP orthologue that functions as an E3 in vitro (14). AtCHIP is up-regulated by certain stress conditions and overexpression of AtCHIP rendered Arabidopsis...
more sensitive to both low and high temperatures, suggesting a link between protein ubiquitination and stress responses in plants.

Ubiquitin-dependent protein degradation has been shown to play important roles in plant growth and development (15). The abundance of U-box proteins, their ability to interact functionally with E2 enzymes to ubiquitinate protein substrates (16), and the importance of their roles in the regulation of important cellular processes (17) make structure determination of an Arabidopsis U-box domain of interest. Here we report the NMR solution structure of a prototypical Arabidopsis U-box domain from the AtPUB14 E3 protein (8).

EXPERIMENTAL PROCEDURES
Identification and Analysis of Predicted Arabidopsis U-box Proteins—Arabidopsis U-box proteins were identified by BLAST homology searches (17) and Inter Pro (18) and SMART (19) database searches. Redundant sequences were identified by ClustalW sequence alignments (20). Non-redundant U-box proteins were analyzed for domain architecture and intrinsic sequence features by SMART and BLAST domain analysis.

Cloning, Expression, and Purification of Recombinant Proteins—AtPUB14 (GenBank™ accession AT3g54850) DNA fragments were amplified from the expressed sequence tag clone RAFL09-40-H12 obtained from The Experimental Plant Division, BioResource Center (BRC), RIKEN Tsukuba Institute, Tsukuba, Japan. Fragments encoding the following regions were cloned into pGEX-4T-1 (Amersham Biosciences) to obtain glutathione S-transferase (GST) fusion proteins: full-length AtPUB14 (amino acids 1–632) using the primers 5′-GGAGAATTCATG-GGTTAC, the U-box domain (amino acids 249–321) using the primers 5′-GGAGAATTCCCGAGTATTCCTCGGGGTGTT and 5′-TCATCCTTGAGGATGAGAATTCCCAGAGTATTTTCGGTGT and the U-box and ARM repeat region (amino acids 249–632) using the primers 5′-GGAGAATTCATG-GGTTAC. A full-length fragment was also cloned into pMal-c2 (New England Biolabs) to obtain maltose-binding protein (MBP) fusion proteins: full-length AtPUB14 using the primers 5′-GGGAATTCATGGGATTAACGAATTGT and 5′-GGGAATTCATG-GGTTAC from database matching using TALOS (25).

Cloning, Expression, and Purification of Recombinant Proteins—Arabidopsis U-box Structure—Arabidopsis U-box Structure—

NMR Spectroscopy and Resonance Assignments—Samples for NMR spectroscopy typically contained 0.6 mM PUB14-(249–321), 20 mM sodium phosphate, 0.15 mM NaCl, 1 mM DTT, and 10% or 99.96% D2O (v/v), pH 7.5. NMR experiments were carried out at 25 °C using Varian Unity Inova 750 or 800 spectrometers. For resonance assignment 15N-1H HSQC, HNCA, CBCA(CO)NH, HNHC0, HCNC0, 13C-TOCSY, 13C-NOESY-HSQC, and (HB)CB(CGC-HCCH)HSQC spectra were recorded using the standard set-up provided by Varian. The spectra were processed using NMRPipe and analyzed using Proton3D (23). Autoseg (24) was used to assign the backbone atoms. These assignments were confirmed by interactive computer spectrum analysis. Torsion angle restraints were derived from database matching using TALOS (25).

NMR Spectroscopy and Resonance Assignments—Samples for NMR spectroscopy typically contained 0.6 mM PUB14-(249–321), 20 mM sodium phosphate, 0.15 mM NaCl, 1 mM DTT, and 10% or 99.96% D2O (v/v), pH 7.5. NMR experiments were carried out at 25 °C using Varian Unity Inova 750 or 800 spectrometers. For resonance assignment 15N-1H HSQC, HNCA, CBCA(CO)NH, HNHC0, HCNC0, 13C-TOCSY, 13C-NOESY-HSQC, and (HB)CB(CGC-HCCH)HSQC spectra were recorded using the standard set-up provided by Varian. The spectra were processed using NMRPipe and analyzed using Proton3D (23). Autoseg (24) was used to assign the backbone atoms. These assignments were confirmed by interactive computer spectrum analysis. Torsion angle restraints were derived from database matching using TALOS (25).

Cloning, Expression, and Purification of Recombinant Proteins—Arabidopsis U-box Structure—

From the sample of 30 water-refined structures an ensemble of 20 structures comprising those with the lowest number of violations with the lowest energy was chosen to represent the PUB14-(249–321) U-box structure. These structures were analyzed and checked using Procheck-NMR (28), XPLOR (27), and NACCESS (31). For a comparison of related structures a set of six RING structures and one U-box structure was selected using CE (32) and MODELLER 6.2 (33). The selection criterion was an alignment of more than 80% of their Cα atoms within 2.5 Å of the structure of PUB14-(249–321). The structures in the set determined by NMR spectroscopy, only model 1 was considered. The secondary structures were displayed using PyMOL, an open-source computer graphics program Web Lab Viewer and MOLMOL (34).

PG-SLED NMR Experiments—To a 1.1 mM 15N-1H/C-labeled PUB14-(249–321) sample in 99.6% (v/v) D2O was added 0.07% 1,4-dioxane (Fluka). The Stokes radius of the U-box was determined as described previously (35).

HYDROPRO Prediction of the Stokes Radius—HYDROPRO was used to compute the Stokes radius of the calculated PUB14-(249–321) U-box structure (36). The three lowest energy structures of the ensemble of 20 structures representing the PUB14-(249–321) U-box were used for the calculations, and an average of the three resulting Stokes radii was calculated. Standard parameters were used, and molecular masses and temperatures were set to 5754 Da and 298 K, respectively.

Coordinates—The coordinates for each of the 20 structures in the selected ensemble of PUB14-(249–321) U-box structures together with all NMR assignments have been deposited at the Protein Data Bank under the accession code 1TH.
RESULTS

Arabidopsis U-box Proteins—The Arabidopsis thaliana genome encodes for many U-box proteins (8). To select a representative U-box protein for structural and functional studies we analyzed the Arabidopsis genome for predicted U-box proteins using BLAST (17) homology searches and keyword searches of the SMART (19) and InterPro (18) protein domain databases. This analysis uncovered 62 Arabidopsis proteins with a U-box domain and domain compositions as shown in Fig. 1a. The SMART, BLAST, and InterPro analyses showed that a
The structure of the complex between the RING domain from c-Cbl and the E2 UbcH7 (37) serves as a model for interactions between RING domain E3 and E2 enzymes. According to this, Ile383 and Trp408 of c-Cbl form a hydrophobic groove that in- teracts with a phenylalanine in UbcH7, and Trp108 is involved in determining the specificity of the E2/E3 interactions (37). Identical amino acids are present in corresponding positions in c-Cbl and the E2 UbcH7 (37) serves as a model for interactions between RING domain E3 and E2 enzymes. According to this, Ile383 and Trp408 of c-Cbl form a hydrophobic groove that interacts with a phenylalanine in UbcH7, and Trp108 is involved in determining the specificity of the E2/E3 interactions (37). Identical amino acids are present in corresponding positions in c-Cbl and the E2 UbcH7 (37) serves as a model for interactions between RING domain E3 and E2 enzymes. According to this, Ile383 and Trp408 of c-Cbl form a hydrophobic groove that interacts with a phenylalanine in UbcH7, and Trp108 is involved in determining the specificity of the E2/E3 interactions (37). Identical amino acids are present in corresponding positions in c-Cbl and the E2 UbcH7 (37) serves as a model for interactions between RING domain E3 and E2 enzymes. 根据这一模型，Ile383和Trp408的c-Cbl与UbcH7中的苯丙氨酸相互作用，而Trp108参与决定E2/E3相互作用的特异性（37）。在c-Cbl和E2 UbcH7中的相应位置存在相同的氨基酸。
AtPUB14 fused to GST (GST-PUB14-(249–632)) or MBP (MBP-PUB14-(219–632)) were soluble and contained putative substrate-binding ARM repeats. For the assays GST-PUB14-(249–632) was incubated with recombinant ubiquitin, rabbit E1, and human recombinant E2 (UbcH5b or UbcH13) in the presence of E. coli proteins. UbcH5b was used, because it interacts functionally with several U-box and RING proteins (3, 21). The reactions were probed by Western blotting using anti-ubiquitin or anti-GST antibodies for the detection. When the reaction mixture contained E. coli proteins and was analyzed using anti-ubiquitin antibodies, GST-PUB14-(249–632) mediated ubiquitination of protein substrates in the reaction mixture (Fig. 3a). This modification was dependent on both ubiquitin, ATP, an E1, the type of E2, and GST-PUB14-(249–632). UbcH5b, but not UbcH13, functioned with GST-PUB14-(249–632) showing that AtPUB14 requires a specific E2 for activity. In addition, the results showed that the UND domain is not required for in vitro ubiquitination activity of the Arabidopsis UND/U-box/ARM proteins, as suggested previously (16).

Auto-ubiquitination—E3s often undergo auto-ubiquitination, and auto-ubiquitination of U-box proteins has also been reported (2). Therefore, the ability of GST-PUB14-(249–632) to undergo auto-ubiquitination was examined. When detection was performed using an antibody against GST, a substantial proportion of the 68-kDa GST-PUB14-(249–632) protein exhibited a shift in electrophoretic mobility toward the top of the gel indicative of polyubiquitination. This reconfirmed the ability of the U-box domain to function as an E3 ubiquitin ligase.

Structure Determination of the PUB14-(249–321) U-box Domain by NMR Spectroscopy—The ability of AtPUB14 to function as an E3 ubiquitin ligase has made structure determination of its U-box domain important for an in-depth understanding of how this function is executed. Analysis of the NMR spectra of PUB14-(249–321) U-box resulted in the assignment of the NMR signals from more than 95% of the NMR active nuclei in the peptide backbone and the amino acid side chains. A total of 1083 non-redundant NOEs was assigned and applied in the structure calculations together with 88 backbone dihedral angle constraints derived from secondary chemical shift database mining using TALOS (Table I). The PUB14-(249–321) U-box domain has a well defined structure in the region between residues Tyr 251 and Glu321. The 20 structures of this part of the U-box domain superimpose well with an r.m.s.d. to the average structure of 0.58 ± 0.13 Å for the backbone atoms and 1.0 ± 0.1 Å for all heavy atoms, and for the entire ensemble 98.2% of the (ϕ, ψ) angle combinations fall in the allowed regions of the Ramachandran plot (Table I).

**Tertiary Structure of the PUB14-(249–321) U-box Domain**—The AtPUB14 U-box domain contains two α-helices, helix1 (Glu274-Gly285) and helix2 (Tyr305-Asn318), an antiparallel...
β-sheet of three strands formed by two strands in a β-hairpin (Pro264–Tyr277) and an additional β-strand (Thr269–Asn284), and of two long loops, loop1 (Tyr281–Asp283) and loop2 (His296, Leu301) (Figs. 1a and 4b). The structure is globular with helix1 packed against the antiparallel β-sheet and helix2 packed against the N-terminal loop. The structure is stabilized by a number of long range interactions (Fig. 4b). Residues in the N-terminal loop (Leu260, Met261, and Pro264) and residues in the C-terminal part of the β-hairpin (Val267 and Tyr273). Finally, the Lys262–Glu274 salt bridge forms a contact between loop1 and helix1. Hydrogen Bond Networks Instead of Zinc Ion Binding Ligands in the AtPUB14 U-box Domain—In the Prp19p U-box a hydrogen bond network replaces the stabilization of the zinc ion binding sites in the RING motif (7). The residues in the AtPUB14 U-box in the positions corresponding to the first zinc ion binding site in the RING motif are Cys264, Ser267, Glu274, and Ser276. In the RING motif the zinc ion binding site forms interactions between loop1 and helix1. Similarly, in the AtPUB14 U-box three of these residues, Cys264, Ser267, and Ser277, and an additional residue, Glu259, form a network of hydrogen bonds between these two elements of the structure. This is very similar to the Prp19p U-box where the corresponding four residues (Cys3, Ser6, Glu24, and Ser26) are involved in the hydrogen bond network, which also includes Lys8. Together with the backbone HN of Ser267 (in the AtPUB14 U-box) the side chains of these four residues appear to form a network of hydrogen bonds in most of the individual structures of the NMR ensemble (15 out of 20). Central to this network are the two oxygen atoms of the carboxylate group of Glu259. These can form hydrogen bonds to the hydrogen of the sulphydryl group of Cys264, to the hydrogen of the hydroxyl group of Ser277, and to the hydrogen of the peptide group of Ser267. A tentative picture of the hydrogen bond network can be drawn from the ensemble of NMR structures (Fig. 4c). The network apparently consists of at least three hydrogen bond donors from two side chains and one from the peptide backbone. A single carboxylate group (Glu259) provides a pair of oxygen acceptors one of which may form a bifurcated hydrogen bond to two of the hydrogen donors. The residues in AtPUB14 in the positions corresponding to the second zinc ion binding site in the RING motif are Thr269, Thr272, Cys269, and Ser292. A network of hydrogen bonds and salt bridges in the AtPUB14 U-box domain involves two of these residues, Cys269 and Ser292, which together with Lys291 and Glu294 seem to replace the interactions in the second zinc ion binding site in the RING motif. This network is not defined very well in the ensemble of NMR structures. However, slow hydrogen exchange of the HN both in Thr269 and Thr272 suggests that these form hydrogen bonds that stabilize the network. The analysis of the hydrogen bonds in the structures (Fig. 4d) indicates the presence of a salt bridge between the carboxylate of Glu294 and the amino group of Lys291, which may stabilize the loop2 conformation, and a hydrogen bond between the side chain carboxyl of Glu271 and the hydrogen of the sulphydryl group of Cys269, which is the only interaction seen that connects the C-terminal string of the hairpin and loop2 as in the zinc ion site in the RING motif. There are a number of differences, between the ScPrp19p U-box and the AtPUB14 U-box in those residue positions, that are involved in the hydrogen bond network and that replace the corresponding zinc ion ligation sites in the RING motifs (Fig. 1c). It is noteworthy, however, that the residues surrounding the conserved E2 binding site residues in the two loops are engaged in the hydrogen bond network in the two U-box structures. Structural Comparison of RING and U-box Folds—The typical mixed α+β fold of the RING finger, found both in Prp19p and AtPUB14 U-box, is seen in a number of protein structures in the Protein Data Bank. The scaffold of these structures is remarkably similar to ten RING finger structures, all of which are classified as belonging to the RING finger domain, C3HC4 family, of the SCOP database (38). An alignment of these 12 three-dimensional protein structures using combinatorial extension (32) and MODELLER 6.2 (33) revealed a total of seven proteins that align more than 80% of their Cα atoms within 2.5 Å of the lowest energy structure of the PUB14-(249–321) U-box domain (Table II). A superimposition of these eight structures is shown in Fig. 5 and includes the RING and U-box domains of c-Cbl, Rbx1, BRCA1, RAG1, EL5, Prp19p, MAT1, and AtPUB14. All of these proteins, except MAT1, have been shown to function as ubiquitin protein ligases (7, 39–43). The similarity of the orientation of the secondary structures consisting of one α-helix (helix1) and three anti-parallel β-strands is obvious, but also the orientation and packing of the two loops against the α-helix is equivalent. Interactions with E2 Enzymes—In the structure of the complex between c-Cbl and UbcH7, three residues, Ile283, Trp285, and Pro347, have been proposed to form a hydrophobic binding groove for the E2 interaction (37). In six of the seven structures aligning with AtPUB14, the proline and a large aliphatic residue are conserved, and they all contain a hydrophobic groove (Table II and Fig. 6). Substitution of Ile285 in BRCA1 with an alanine abolished the ubiquitin ligase activity of the BRCA1–BRCA1-associated RING domain protein complex (44) demonstrating the importance of this aliphatic position. MAT1 contains an arginine in this position, which, besides being engaged in the hydrophobic groove, protrudes its guanidino group right in the middle of the putative E2 binding site. This could explain why no E3 activity has been demonstrated for MAT1. The
U-box structure collapses upon mutation of the proline to alanine (7). This suggests that, besides its role in the E2 binding site, this proline is also mandatory for stabilizing the loop and the structural scaffold. In AtPUB14, van der Waals interactions of Pro290 in loop2 are found mainly to occur in residues Tyr273 in the \( \beta \)-hairpin, Trp281 in helix1, and Cys289 in loop2, but also in Ile256 in loop1. The three residues in the putative E2 binding site in the AtPUB14 U-box are Ile256 in loop1, Trp281 in helix2, and Pro290 in loop2. To establish the appropriate geometry for a binding site of this type the three elements that provide the residues have to be fixed in a well defined way relative to each other. In the RING motif this is in part ensured by the two zinc ion binding sites, both of which link the loop regions to a well defined secondary structure element. Furthermore, the two binding site residues are placed right between zinc ion binding site ligand residues in the two loops (Fig. 1c). In the AtPUB14 U-box the two hydrophobic cores and the long range hydrogen bonds and salt bridges may serve this purpose together with the network of local and long range hydrogen bonds that is observed in the structures. In comparison to the Prp19p U-box the loop1 position in the AtPUB14 U-box is further stabilized by the interactions to the C-terminal helix, which is not observed in the Prp19p U-box structure.

In addition to the hydrophobic binding surface in the E2 binding site it has been proposed that the exposed side of the common \( \alpha \)-helix (helix1) in the RING and U-box domains is part of the binding site (Fig. 6) (7, 37, 43, 44). Several of the residues in this site are polar and charged, however there is no clear sequence homology in this region (Fig. 1b). In c-Cbl Cys404, Ser407, and Ser411 are part of the polar E2 interface, whereas in AtPUB14 the corresponding residues are Ser277, Lys280, and Ala284. In BRCA1 NMR studies have shown that several polar residues are in the proximity of the binding site for Ubc5Hc (44), and similar indications were seen for E2-binding by EL5 (43). However, a comparison, which involves eight highly familiar U-box and RING structures, reveals that there is no common surface charge pattern in this part of the structure (Fig. 6). Hence, there is no common polar E2 binding site motif in these eight structures or in the Arabidopsis U-box proteins (Fig. 1b).

**Fig. 6. Comparison of RING and U-box domains similar to the AtPUB14 U-box.** The domains of AtPUB14-(249–321), c-Cbl-(377–434), Rbx1-(34–106), BRCA1-(13–90), RAG1-(18–91), EL5-(127–181), Prp19p-(1–56), and MAT1-(1–65) are shown in ribbon format and as molecular surfaces with electrostatic potentials calculated using GRASP. Side chains of residues corresponding to the E2-interacting Ile280, Trp281, and Pro417 of c-Cbl are highlighted in black, and the \( \alpha \)-helix, also contributing to the interaction surface, is shown in front of the structures.
Dimer Formation of the AtPUB14 U-box Domain—The C-terminal α-helix (helix2) in the AtPUB14 U-box domain is not present in the Prp19p U-box domain. However, a corresponding C-terminal α-helix is also found in the RING domain of RAG1, where it is involved in dimerization (45), and in the RING domain of BRCA1, where it forms intramolecular interactions with an N-terminal helix (Fig. 6) (46). The AtPUB14 U-box domain presents solvent-accessible non-polar residues from the N terminus (Tyr251 and Phe252) and from the C-terminal α-helix (Leu313 and Trp314). In RAG1, similar residues (Phe280, Phe284, Phe288, and Ile297) form a homodimerization interface (45). This similarity may suggest that the AtPUB14 protein can form a dimer. Therefore, the molecular weight of the AtPUB14 U-box was determined by two different methods using gel filtration (data not shown) and pulsed field gradient NMR techniques, respectively. Both techniques measured a Stokes radius of 21 ± 1 Å for the AtPUB14 U-box domain. This is different from the predicted Stokes radius of 17.6 Å of the monomeric AtPUB14 U-box structure determined by NMR spectroscopy and suggests that the AtPUB14 U-box might form a dimer in solution. An attempt to dissociate the dimer by dilution and monitoring this by chemical shift changes in the NMR spectrum was performed. At low concentration, 0.032 mM, of AtPUB14 U-box a very small but significant chemical shift changes were observed as well as specific line broadening effects (Fig. 7). It is of interest to note that, although these effects are very small, they coincide with the position of the dimerization site in RAG1, and they involve non-polar residues in the N-terminal region and in the C-terminal α-helix.

Although there is a strong indication from several different methods that the AtPUB14 U-box forms a dimer in solution, the present study offers no result to describe the structure of the interface between two monomeric structures. The 15N-H HSQC NMR spectrum has only one peak per residue, and there are no signs of signal doubling for any atoms in the spectra. This suggests that if the molecule forms a dimer it is highly symmetrical. The set of NOEs that has been used for the structure calculations defines the monomer structure well. Attempts to model a dimer structure, using otherwise assigned intramolecular NOEs from the suggested dimerization region as intermolecular NOEs between two molecules, never resulted in structures that complied with the new set of restraints. Because there are no additional unassigned NOEs in the spectra, this suggests strongly that the set of NOEs used in the structure calculations contains all intramolecular NOEs, and apparently there are no intermolecular NOEs in the set that can be used to describe the dimer interface. The similarity of the monomer structure to seven other RING and U-box structures (Fig. 5 and Table II) is also in support of the structure determination of the monomeric form of AtPUB14 U-box being a representative structure of this even as a part of the suggested dimer.

DISCUSSION

By far the largest subgroup of Arabidopsis U-box proteins contains one or several ARM repeats C-terminal to a highly conserved U-box domain (Fig. 1a) (47). Important physiological functions have been assigned to these proteins (13, 48), and the origin of the different functions may be due to differences in the binding specificities of the ARM repeats (47). There is increasing evidence that these proteins function as ubiquitin ligases (16) and thus participate in regulated protein degradation. The structure of the AtPUB14 U-box domain reveals similarities, as well as differences, to other RING/U-box domains and based on predictions allows considerations of the activity of the many plant U-box proteins and their specificity toward E2s.

UbcH5b, but not UbcH13, functions with AtPUB14 to ubiquitinate protein substrates. UbcH5b contains a phenylalanine in a position that is likely to interact with Trp281 of the AtPUB14 U-box domain (37). Most of the Arabidopsis U-box domains contain a tryptophan in this position, but in some domains the position is occupied either by a histidine, a tyrosine, or a cysteine (Fig. 1b) (8). The same residues, and leucine, are also found in other eukaryotic U-box domains (Fig. 16). This position was initially proposed to be a specificity determinant for the interaction with E2 enzymes based on the c-Cbl-UbcH7 complex structure (37). This has since been supported by independent studies. For example, Prp19p contains a tyrosine and functions specifically with Ubc3 (7), and U-box domains with a histidine functionally interact with Ubc4 and Ubc5 (2, 3). Whereas the identity of this position is essential for the activity of c-Cbl (39), containing a tryptophan, and Prp19p, containing a tyrosine (7), the leucine in BRCA1 is dispensable.
for the ubiquitination activity of the BRCA1-BRCA1-associated RING domain protein ubiquitin-ligase complex (44). This could be explained by the differential importance of this position depending on the chemical nature of the residue. Whereas leucine is one of several hydrophobic contributors to the hydrophobic E2-binding groove, the effect of the structurally more characteristic tryptophan and tyrosine could be more specific. However, it was recently demonstrated that five different Arabidopsis U-box proteins, all containing a tryptophan, were active with different E2s having a preference that did not correlate with their phylogenetic relationship (16). Thus, additional residues and regions must be of importance to specificity. The structure of AtPUB14 and comparisons suggest that the distribution of exposed charged and polar residues of the central α-helix of AtPUB14 may act as specificity determinants. In this structure element, highly conserved positions are mixed with positions showing sequence variance required for specificity (Fig. 1b). This structural information now allows rational examination of the specificity of the many Arabidopsis U-box domains and E2 enzymes.

Based on sequence alignments, the consensus sequence of the Arabidopsis U-box domains was expanded at the C terminus with U-box domains from other organisms. This region was predicted (49) to form an α-helix, which was confirmed by the determined structure of AtPUB14. In the RAG1 RING domain a similar C-terminal and an N-terminal α-helix exposes hydrophobic residues that form a dimerization surface (45). An N-terminal α-helix was not predicted for the AtPUB14 U-box fragment does not contain an N-terminal α-helix. This is in support of its absence from the Arabidopsis U-box protein motif. However, hydrophobic amino acids are exposed from both the N-terminal loop region and the C-terminal α-helix (helix2) of the AtPUB14 U-box domain and can explain the ability of the domain to form a dimer, in accordance with the chemical shift analysis (Fig. 7).

The C-terminal α-helix is also likely to play an important role in stabilization of the U-domain. The interactions between residues in loop1 and helix2 add, not only, to the stabilization of the molecule but also to formation of an important part of the hydrophobic binding site. The Arabidopsis U-box domains are assumed to be stabilized compared with what was reported with their phylogenetic relationship (16). Thus, additional U-box proteins, many of which contain several (45). An N-terminal residues in loop1 and helix2 add, not only, to the stabilization role in stabilization of the U-domain. The interactions between folding of this small and commonly found protein fold.

REFERENCES

1. Pickart, C. M. (2001) Annu. Rev. Biochem. 70, 503–533
2. Hatakeyama, S., Yada, M., Matsumoto, M., Ishida, N., and Nakayama, K. I. (2001) J. Biol. Chem. 276, 33111–33120
3. Jiang, J., Ballinger, C. A., Wu, Y., Dai, Q., Cyr, D. M., Hohfeld, J., and Patterson, C. (2001) J. Biol. Chem. 276, 42938–42944
4. Johnson, E. S., Ma, P. C., Ota, I. M., and Varshavsky, A. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 5651–5656
5. Brzovic, P. S., Reeve, J. R., Nishikawa, H., Miyamoto, K., Fox, D., 3rd, Fukuda, M., Ohta, T., and Klevecz, R. (2001) J. Mol. Biol. 312, 307–320
6. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) J. Mol. Biol. 321, 1043–1057
7. Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M. C., and Klevecz, R. E. (2001) Nat. Struct. Biol. 8, 833–837
8. Coates, J. C. (2003) Trends Cell Biol. 13, 463–471
9. Lomax, A., Heginbotham, P., Hoyt, D. W., and Klevecz, R. E. (2001) Nat. Struct. Biol. 8, 833–837
10. Coates, J. C. (2003) Trends Cell Biol. 13, 463–471
11. Amador, V., Monte, E., Garcia-Martinez, J. L., and Prat, S. (2001) Cell 106, 343–354
12. Jorgensen, P., Lamb, B., and Klevecz, R. E. (2001) Structure 9, 586–591
Structure and Biochemical Function of a Prototypical Arabidopsis U-box Domain
Pernille Andersen, Birthe B. Kragelund, Addie N. Olsen, Flemming H. Larsen, Nam-Hai Chua, Flemming M. Poulsen and Karen Skriver

J. Biol. Chem. 2004, 279:40053-40061.
doi: 10.1074/jbc.M405057200 originally published online June 30, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M405057200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 48 references, 14 of which can be accessed free at http://www.jbc.org/content/279/38/40053.full.html#ref-list-1