Tertiary Structures of Class I Ubiquitin-Conjugating Enzymes Are Highly Conserved: Crystal Structure of Yeast Ubc4[†,‡]

William J. Cook,[*,§,∥] Leigh C. Jeffrey,[∥] Yuping Xu,[⊥] and Vincent Chau[⊥]

Department of Pathology and Center for Macromolecular Crystallography, University of Alabama at Birmingham, Birmingham, Alabama 35294, and Department of Pharmacology, Wayne State University School of Medicine, Detroit, Michigan 48201

Received July 15, 1993; Revised Manuscript Received September 3, 1993[●]

ABSTRACT: The three-dimensional structure of a yeast ubiquitin-conjugating enzyme, encoded by the *Saccharomyces cerevisiae UBC4* gene, has been determined at 2.7 Å. The structure was solved using molecular replacement techniques and refined by simulated annealing to an *R*-factor of 0.198. Bond lengths and angles in the molecule have root mean square deviations from ideal values of 0.018 Å and 4.0°, respectively. Ubc4 is an α/β protein with four α-helices and a four-stranded antiparallel β-sheet. The ubiquitin-accepting cysteine is located in a cleft between two loops. Comparison with the recently determined structure of a different plant enzyme suggests that class I ubiquitin-conjugating enzymes are highly conserved in their three-dimensional folding. Except for two extra residues at the N- and the C-terminus of the plant enzyme, the Cα atoms of the two enzymes can be superimposed with a root mean square deviation of only 1.52 A. Greater variations are found between the surfaces of the two molecules, as most of the identical residues between the two enzymes are either buried or clustered on one surface that lies adjacent to the ubiquitin-accepting cysteine. We suggest that this conserved surface functions in protein–protein binding during ubiquitin thiol ester formation.

Ubiquitin-conjugating enzymes (E2) are a family of closely related proteins which function in the posttranslational addition of ubiquitin to a variety of cellular proteins (Finley & Chau, 1991; Jentsch, 1992a,b). These enzymes can be empirically defined by their ability, in the presence of the ubiquitin-activating enzyme (E1) and ATP, to form a thiol ester adduct with the C-terminal carboxyl group of ubiquitin (Ub) in the following reactions (Haas et al., 1982; Pickart, 1988):

$$E1_{SH} + Ub + ATP \leftrightarrow E1_{S\text{-}Ub} + AMP + PP_i$$

$$E1_{S\text{-}Ub} + E2_{SH} \leftrightarrow E1_{SH} + E2_{S\text{-}Ub}$$

Ubiquitin–protein conjugates are formed by transferring ubiquitin from $E2_{S\text{-}Ub}$ to an acceptor protein, resulting in the formation of an isopeptide bond between the C-terminal carboxyl group of ubiquitin and an ϵ-amino group in the acceptor protein. The latter reaction may result directly from interactions of a specific E2 with its cognate substrate protein or may require the E2 enzyme to be complexed with an additional substrate-binding protein known as E3 (Hershko & Ciechanover, 1992).

E2 gene families have been identified from many eukaryotic species, ranging from yeast to plants and mammals (Jentsch, 1992b). Comparisons of their primary sequences indicate that the ubiquitin-accepting cysteine in all E2 enzymes is located within a core domain of about 150 amino acids that show at least 25% sequence identities. E2 enzymes have been grouped into four distinct classes, depending on the presence and the location of unrelated sequences. Class I enzymes are the smallest E2 enzymes and consist almost entirely of the conserved core domain. Those that contain either extra C-terminal or extra N-terminal extensions from the core domain are called class II and III enzymes, respectively. Class IV enzymes contain both N- and C-terminal extensions. Two class I enzymes have been shown to require an E3 or E3-like protein for substrate ubiquitination (Wing et al., 1992; Girod & Vierstra, 1992), suggesting that the core domain is insufficient for enzyme–substrate recognition. Whether this is the function of the extra sequences in class II, III, and IV enzymes remains to be rigorously tested.

UBC4 is one of ten presently known genes in *Saccharomyces cerevisiae* that encode ubiquitin-conjugating enzymes (Jentsch, 1992b; Seufert & Jentsch, 1990). This gene encodes a class I enzyme and has been shown to function in ubiquitin-mediated proteolysis. In this degradative pathway, the ubiquitin moiety in ubiquitin–protein conjugates serves to target acceptor proteins for selective proteolysis. Substrates in this degradative pathway include constitutively or conditionally short-lived proteins (Ciechanover et al., 1984; Glotzer et al., 1991; Hochstrasser et al., 1991; Shanklin et al., 1987) as well as abnormal proteins (Seufert & Jentsch, 1990) and free subunits of oligomeric protein complexes (Finley & Chau, 1991). *UBC4* has been shown to function in the degradation of abnormal proteins and many newly synthesized proteins that can be rapidly labeled *in vivo* (Seufert & Jentsch, 1990) as well as stable ubiquitin-fusion proteins (Johnson et al., 1992). This gene is expressed in exponentially grown cells and can be induced by heat shock or by the presence of amino acid analogs in the growth medium. An almost identical protein is also encoded by *UBC5*, which is expressed predominantly during stationary growth phase. E2 enzymes that share about 80% protein sequence identity with Ubc4/Ubc5 have been identified from E2-encoding genes in *Drosophila melanogaster* (Trier et al., 1992) and *Caenorhabditis elegans* (Zhen et al., 1993). This degree of protein sequence identity is also evident for two

[†] Supported by NASA Grant NAGW813 and NIH Grant NS-30374 (W.J.C.) and by NIH Grants GM-35803 and GM-47604 (V.C.).

[‡] Coordinates for *Saccharomyces cerevisiae* Ubc4 have been deposited with the Brookhaven Protein Data Bank, Accession Number 1UCE.

[§] Department of Pathology, University of Alabama at Birmingham.

[∥] Center for Macromolecular Crystallography, University of Alabama at Birmingham.

[⊥] Wayne State University School of Medicine.

[●] Abstract published in *Advance ACS Abstracts,* November 1, 1993.

0006-2960/93/0432-13809$04.00/0

Table I: Summary of Data Collection for Yeast Ubc4[a]

data set	resolution (Å)	no. of observations: total	unique	% complete	R_{sym}
crystal no. 1	2.9	4145	2768	55.6	0.117
crystal no. 2	2.7	9226	4262	69.9	0.116
crystal no. 3	2.7	5770	3521	57.2	0.079
crystal no. 4	2.8	4450	3238	58.4	0.085
merged native	2.7	22 434	5940	96.9	0.134

[a] R_{sym} is defined as $\sum_{h,k,l} \sum_i^N |\bar{I} - I_i| / \sum \bar{I}$, where $\bar{I}$ is the mean intensity of the N reflections with intensities I_i and common indices h, k, and l.

Table II: Summary of the Refinement Statistics[a]

resolution (Å)	no. of reflections: total	$F > 2\sigma(F)$	% complete	R
4.47–6.00	823	784	95.3	0.185
3.83–4.47	793	753	95.0	0.176
3.45–3.83	778	711	91.4	0.181
3.18–3.45	771	661	85.7	0.213
2.98–3.18	768	628	81.8	0.224
2.83–2.98	740	492	66.5	0.239
2.70–2.83	655	386	58.9	0.256

[a] R is defined as $\sum||F_o| - |F_c|| / \sum|F_o|$. The % completeness and R values are based on the data with $F > 2\sigma(F)$.

Arabidopsis thaliana E2 genes (GeneBank translations Z17692 and Z18473) for which only partial coding sequences for 113 and 82 amino acid residues are available. The *Drosophila* and *C. elegans* genes have been shown to complement *UBC4* functions in a *ubc4* null mutant (Treier et al., 1992; Zhen et al., 1993), suggesting that functional homologs of Ubc4/Ubc5 in other species can be identified on the basis of their degree of sequence identity with these two E2 enzymes.

We report here the determination of the three-dimensional structure of the yeast Ubc4 protein. The overall folding of this molecule is similar to that of a recently determined *Arabidopsis* class I E2 enzyme (Cook et al., 1992) which only shares 42% sequence identity with Ubc4/Ubc5 and apparently functions in a different ubiquitination pathway (Sullivan & Vierstra, 1991). Comparisons between these two structures

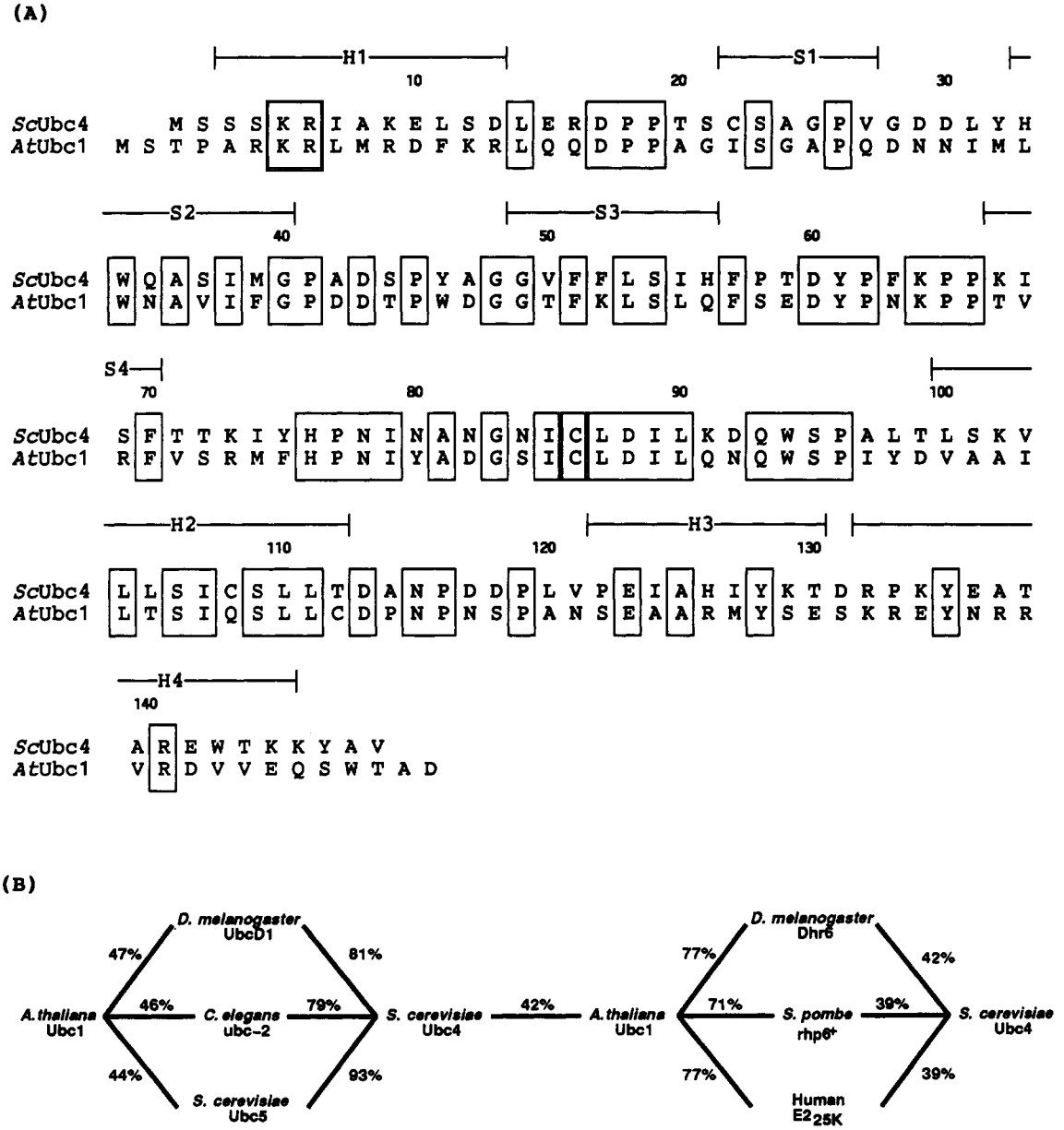

FIGURE 1: (A) Amino acid sequences of *S. cerevisiae* Ubc4 and *A. thaliana* Ubc1. Secondary structure assignments based on the Ubc4 structure are noted above the residues. The ubiquitin-accepting cysteine in each sequence is boxed by a double line. (B) Percent identical amino acid residues shared among *S. cerevisiae* Ubc4, *A. thaliana* Ubc1, and homologs of each.

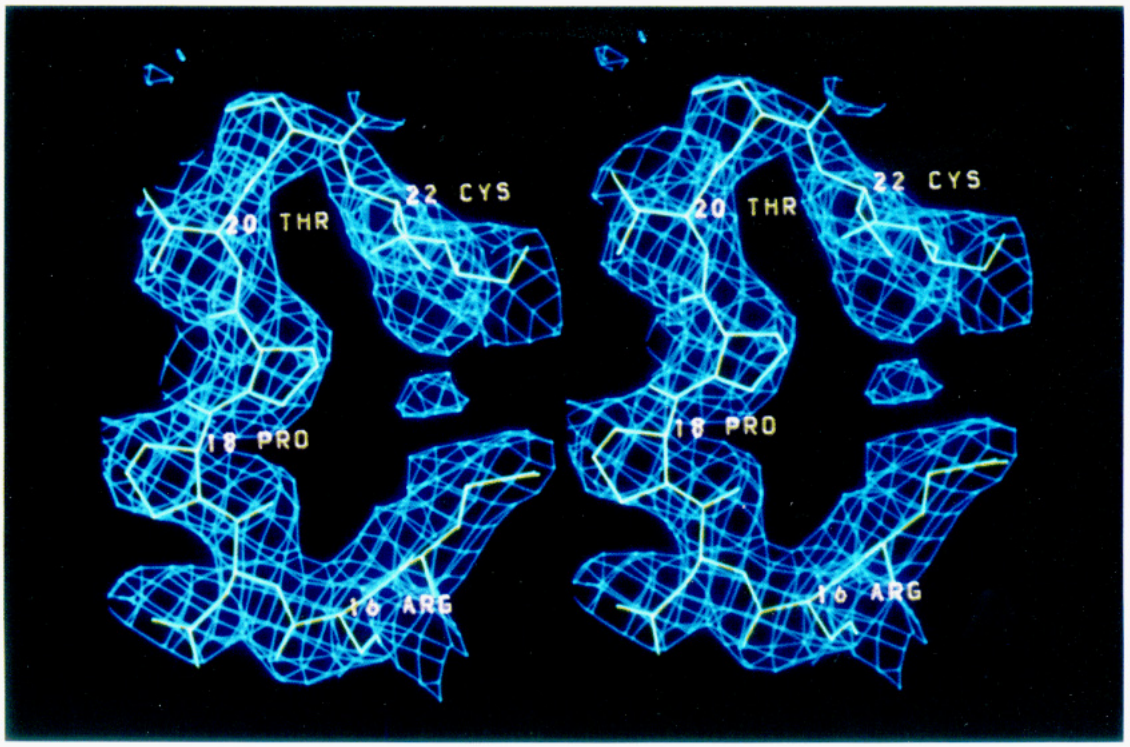

FIGURE 2: Stereoview of the loop involving residues 15–23 with the associated 2.7-Å electron density from the final $2F_o - F_c$ map. The localized density in the center of the hole would accommodate a water molecule at a reasonable hydrogen-bonding distance from Asp17. [Prepared using FRODO (Jones, 1978).]

reveal a surface region that may be the E1-binding site common to all E2 enzymes as well as divergent surface regions that may enable an individual class I enzyme to interact with its respective substrates or E3.

EXPERIMENTAL PROCEDURES

Cloning and expression of recombinant *S.cerevisiae UBC4* in *Escherichia coli* will be described elsewhere (Y. Xu, A. Banerjee, and V. Chau, unpublished results). The purified protein is functional as judged by its ability to form the ubiquitin thiol ester adduct. Crystals of Ubc4 were initially obtained by vapor-diffusion equilibration using a screening kit from Hampton Crystal that is based on a modified sparse matrix sampling method (Jancarik & Kim, 1991). These drops, which consisted of 1 μL of a solution containing 10 mg of protein/mL of distilled water plus 1 μL of the reservoir solution, were equilibrated against 0.7 mL of each trial solution at room temperature (23 °C) and at 4 °C. Within 24 h several sets of conditions at 23 °C gave similar-appearing aggregates of orthorhombic crystals. Further screening around these conditions produced crystals suitable for X-ray diffraction studies. The best crystals were grown from 10% poly(ethylene glycol) 8000 containing 0.2 M magnesium acetate at pH 5.8 and 0.1 M cacodylate buffer.

Intensity data to 2.7-Å resolution were collected at room temperature with a Nicolet X-100A area detector at 22 °C using Cu Kα radiation from a Rigaku RU-300 rotating-anode generator operating at 40 kV and 100 mA. In order to obtain a complete data set with multiple measurements of all reflections, multiple data sets were collected from four crystals. The detector-to-crystal distance was 12 cm, and the detector 2θ value was 12°. Oscillation frames covered 0.25° and were measured for 450 s. The crystals diffract rather weakly and suffer significant decay within 24 h. Therefore, a limited number of frames were collected from each crystal. A total of 22 434 reflections were processed and merged into 5940 unique reflections (97% complete). The R_{sym} value (based on I) for the data to 2.7 Å was 0.134. Of the 514 reflections in the highest resolution range (2.7–2.8 Å), only 173 reflections (34%) had $I > 2\sigma(I)$. Table I gives statistics for the data processing. Indexing and integration of intensity data were carried out using the XENGEN processing programs (Howard et al., 1987).

Of the potential solutions generated by the indexing program, one had all angles approximately equal to 90° and axial lengths consistent with a tetragonal space group. This was the only solution that successfully indexed all reflections. The cell parameters refined to $a = b = 66.90$ Å and $c = 91.81$ Å. The unit cell was tested for the two possible tetragonal Laue symmetry groups by comparing the integrated intensities of potentially equivalent reflections. This analysis confirmed the Laue group $4/mmm$. The systematic absence of reflections $h00$ with $h \neq 2n$ and $00l$ with $l \neq 4n$ indicated either space group $P4_12_12$ or space group $P4_32_12$. There is one molecule in the asymmetric unit, and the solvent volume fraction is 44%.

The crystal structure was solved using the molecular replacement routines in XPLOR (Brünger et al., 1987), which utilize the real-space Patterson search method. The search model was the 2.4-Å crystal structure of a ubiquitin-conjugating enzyme (Ubc1) from the plant *Arabidopsis thaliana* (Cook et al., 1992). *Arabidopsis* Ubc1 has 152 amino acids and 42% sequence identity with yeast Ubc4 (Figure 1A). Because residues 151 and 152 in *Arabidopsis* Ubc1 were never identified in the crystal structure, residues 1–150 were used as the search model. A cross-rotation function was calculated using a radius of integration of 25 Å and data between 12 and 4 Å. The rotation search values did not show much discrimination; the highest peak was 11.3 times the σ value of the map, and the fifth highest was 10.9 times σ. The five highest peaks from the rotation search were subjected to Patterson coefficient (PC) refinement (Brünger, 1990). The

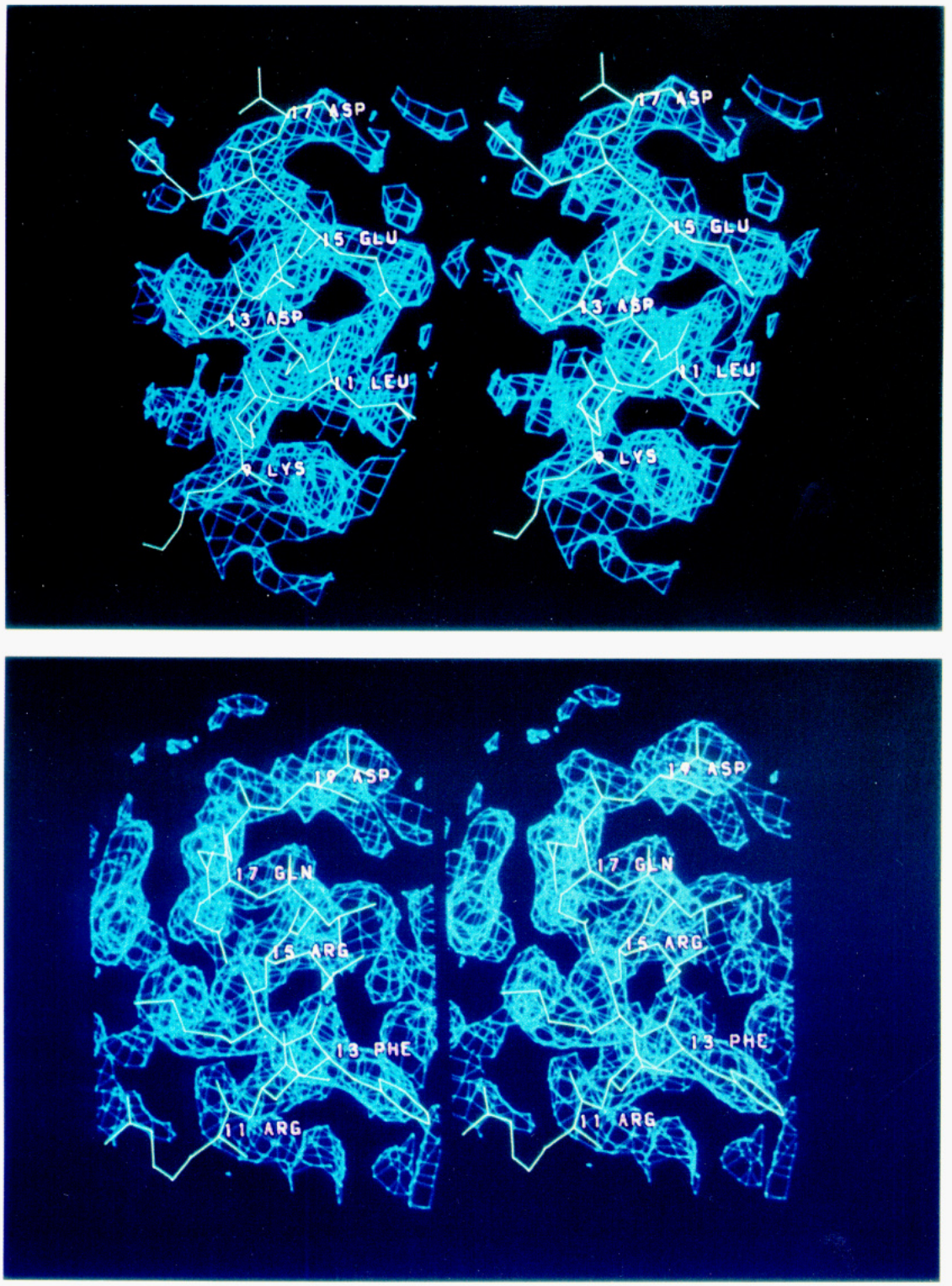

FIGURE 3: Stereoview of helix A in Ubc4 (A, top) and Ubc1 (B, bottom) with the associated electron density. The $2F_o - F_c$ maps were calculated with phases based on the refined structures after removal of residues 8–21 for Ubc4 and 10–23 for Ubc1. [Prepared using FRODO (Jones, 1978).]

four helices in Ubc1 were treated as independent entities within the PC refinement rigid-body refinement procedure. PC refinement produced a clear single maximum peak (solution 1 from the cross-rotation search). The correlation coefficient of this solution was 0.095, with the next highest peak being only 0.068. This solution corresponded to a rotation of the model by the Eulerian angles $\theta_1 = 343.6°$, $\theta_2 = 42.8°$, and $\theta_3 = 323.0°$.

The model was rotated by the angles corresponding to the rotation solution, and the translation function from XPLOR was used with data from 15 to 4 Å for each of the two enantiomorphic space groups. For space group $P4_12_12$ the highest translation function value was 0.487 (7.4σ), while for space group $P4_32_12$ the highest value was 0.413 (6.4σ). The R-factors for the best solution in each case, using data from 8 to 4 Å, were 0.502 and 0.526, respectively. On the basis of these results, the space group $P4_12_12$ was chosen.

Rigid-body and simulated annealing refinement were done with XPLOR (Brünger et al., 1987). Using data from 15- to 3-Å resolution, rigid-body refinement of the best solution

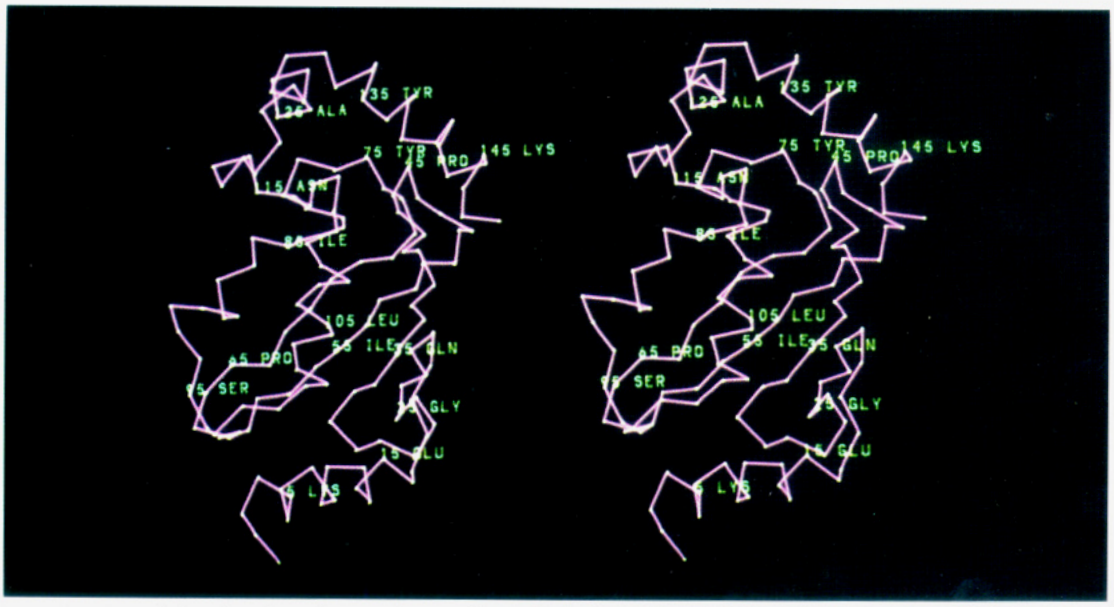

FIGURE 4: Stereo drawing of Ubc4 based on the α-carbon positions. Beginning with residue Lys5, every tenth residue is numbered. [Prepared using FRODO (Jones, 1978).]

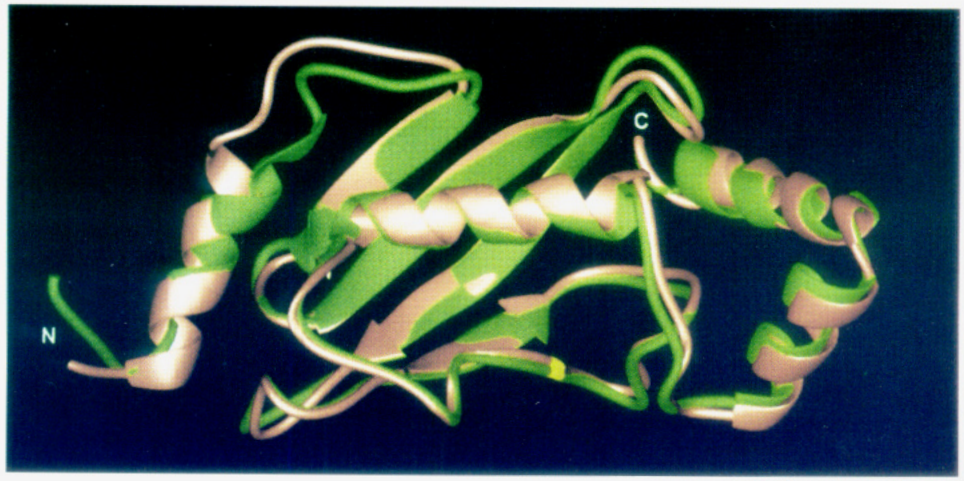

FIGURE 5: Ribbon drawing of the *A. thaliana* Ubc1 molecule (green) superimposed on the *S. cerevisiae* Ubc4 molecule (bronze). The superposition is based on the Cα positions for residues 1–148 from *S. cerevisiae* Ubc4 and residues 3–150 from *A. thaliana* Ubc1. The ubiquitin-accepting cysteine in each protein is yellow. [Prepared using Ribbons 2.0 (Carson, 1987).]

gave an R-factor of 0.460. At this point conservative changes (e.g., Gln to Glu) were made in the sequence, and non-conservative differences in the sequence were replaced with Ala (52 replacements). Refinement by simulated annealing, using data from 6 to 3 Å with $F > 2\sigma(F)$, lowered the *R*-factor to 0.277. Using the computer graphics program FRODO (Jones, 1978), a $2F_o - F_c$ map allowed us to replace all but 10 of the Ala substitutions with the correct residues. Several more rounds of simulated annealing alternating with examination of $2F_o - F_c$ maps and rebuilding allowed modeling of the entire polypeptide. The resolution was gradually extended to 2.7 Å, and individual residues were rebuilt as necessary from $2F_o - F_c$ maps, using computer graphics. Data above 6 Å were not included in the refinement, since they would be especially influenced by disordered solvent.

The *R*-factor at 2.7-Å resolution using isotropic temperature factors of 15.0 Å^2 for all atoms was 0.217. Individual thermal parameters were introduced with restraints on the differences of temperature factors between connected atoms. The *R*-factor after refinement of individual temperature factors was 0.195. As the last step of refinement, restrained least-squares refinement using PROLSQ (Hendrickson, 1985) was used to improve the geometry, although this slightly increased the *R*-factor. The final model includes 148 residues with 1161 protein atoms and no water molecules. (Coordinates for *S. cerevisiae* Ubc4 have been deposited with the Brookhaven Protein Data Bank, Accession Number 1UCE.) The *R*-factor, based on 4415 reflections in the range $2.7 \leq d \leq 6.0$ Å with $F > 2\sigma(F)$, is 0.198. The *R*-index for all nonzero reflections in this range (5024 reflections) with no σ cutoff is 0.226. Table II gives the *R*-factors tabulated as a function of resolution. The final $2F_o - F_c$ electron density map is generally of good quality (Figures 2 and 3A), although the electron density for some of the side chains of hydrophilic residues on the surface is poor, presumably due to large thermal motion or disorder.

Superposition of the *R*-factor curve with theoretical curves for different mean positional errors gives an estimated error of 0.3 Å in the atomic coordinates (Luzzati, 1952). The final coordinates of Ubc4 deviate from ideal bond lengths and angles by 0.018 Å and 4.0°, respectively. All the $\phi - \psi$ pairs for non-glycine residues fall within or are close to allowed regions. The root mean square ΔB values are 1.32 Å^2 for bonded atoms of the main chain and 2.34 Å^2 for bonded atoms of the side chains. Averaged temperature factors for residues range from 9 to 33 Å^2, with an average *B*-factor of 18.3 Å^2 for all protein

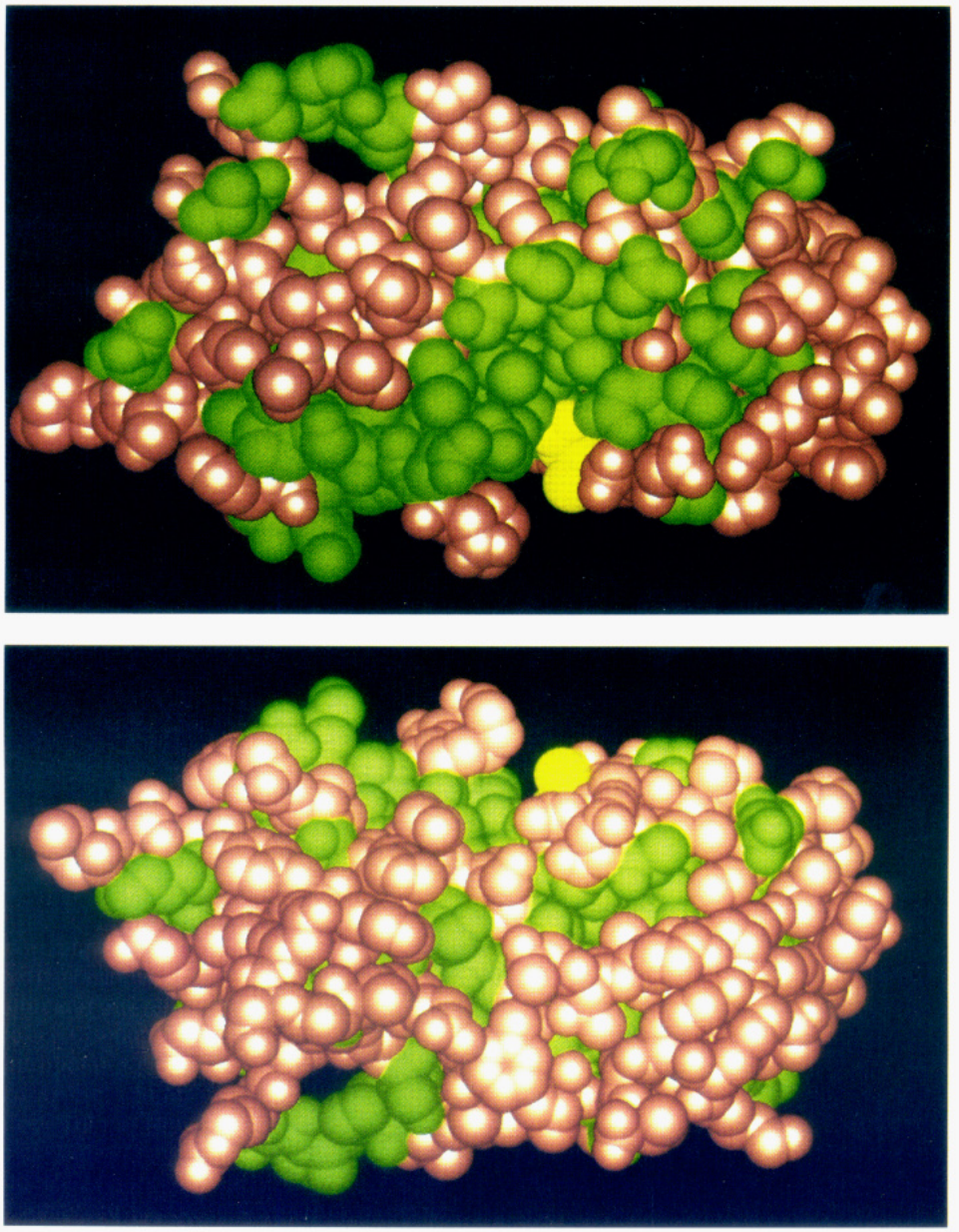

FIGURE 6: Space-filling models of *S. cerevisiae* Ubc4. The orientation of the model in panel A (top) is the same as Figure 4. The model in panel B (bottom) has been rotated by approximately 180° from that in panel a. The color codes are as follows: identical residues between *S. cerevisiae* Ubc4 and *A. thaliana* Ubc1, green; ubiquitin-accepting Cys, yellow; all others, bronze. [Prepared using Ribbons 2.0 (Carson, 1987).]

atoms. In general, there is a good correlation between the *B*-factor and the solvent accessibility of the chain. The highest temperature factors in the molecule occur at the N- and the C-terminus and in large exposed loops on the surface of the protein.

RESULTS AND DISCUSSION

S. cerevisiae Ubc4 contains 148 amino acids and belongs to the class of E2s that consists entirely of the relatively conserved globular "core" domain (class I). The molecule is highly asymmetric with overall dimensions of approximately 21 Å x 25 × 53 Å^3. Ubc4 is an α/β protein (Figures 4 and 5) that contains one antiparallel β-sheet with four strands (designated S1 to S4), bounded on each end and on one side by four helices (designated A to D). One face of the β-sheet is on the surface of the protein. The four β-strands include residues 22–27, 33–40, 49–56, and 67–70, and the four helices consist of residues 3–13, 100–112, 122–130, and 132–145. Thus, Ubc4 contains approximately 18% β-sheet and 32% α-helix. On the basis of the location of the exon boundaries (Seufert & Jentsch, 1990), one would predict a loop at or near residue 16, which agrees well with the observed secondary structure.

The three-dimensional structure of another E2 enzyme has been determined recently (Cook et al., 1992). This enzyme is encoded by the *UBC1* gene from *Arabidopsis thaliana* and also belongs to the class I enzyme group (Sullivan & Vierstra, 1991). Yeast Ubc4 and *Arabidopsis* Ubc1 share only 42% identity in primary sequence and are apparently not functional homologs of the same enzymes in different species (Figure 1B). The *Arabidopsis* protein shares about 80% sequence identity with a group of enzymes that function apparently in DNA repair (Schneider et al., 1990; Koken et al., 1991; Reynolds et al., 1990), while yeast Ubc4 shares the same degree of identity with another E2 group that functions in abnormal protein degradation (Seufert & Jentsch, 1990; Treier et al., 1992; Zhen et al., 1993). Despite their limited sequence

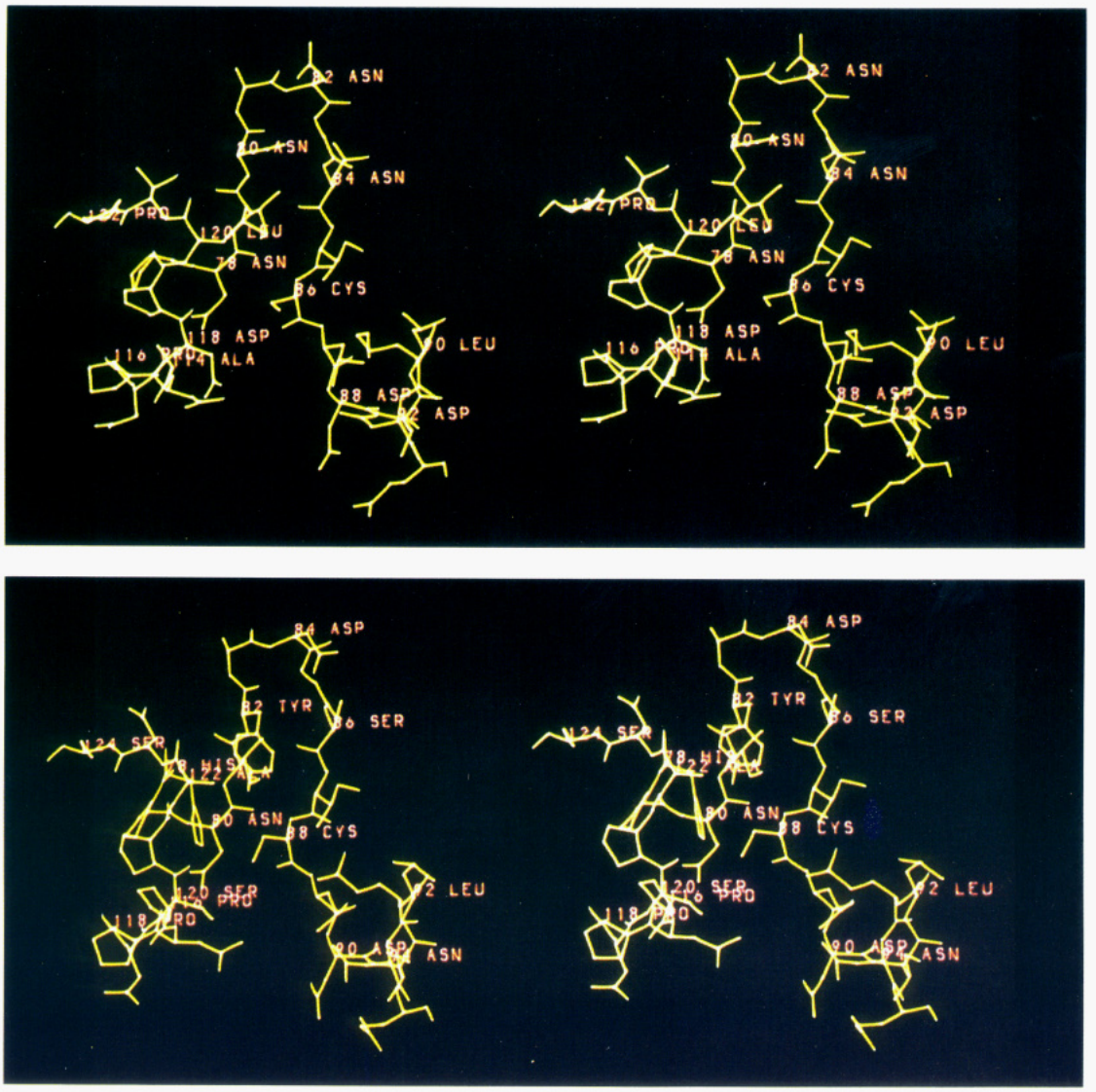

FIGURE 7: Stereo drawings of the environment around the ubiquitin-accepting cysteine in *S. cerevisiae Ubc4* (A, top) and *A. thaliana* Ubc1 (B, bottom). [Prepared using FRODO (Jones, 1978).]

identity and their difference in function, the overall folding of the yeast Ubc4 and *Arabidopsis* Ubc1 molecules was found to be remarkably similar. *Arabidopsis* Ubc1 has four more residues than yeast Ubc4, but the major secondary structural features can be aligned with no insertions or deletions. The four additional residues occur in two-residue stretches at the N- and the C-terminus of *Arabidopsis* Ubc1. Even though there are small variations in the loops between major secondary structural features, the overall topology is quite similar. Superposition of 148 Cα atoms in Ubc4 and Ubc1 gives a root mean square deviation of 1.52 Å.

Most of the differences between the two structures occur in the loops and are fairly small, but there are three regions that are notably different. The first is in helix A, which forms an amphipathic helix along one end of the molecule. This helix in Ubc4 is slightly longer than the corresponding helix in *Arabidopsis* Ubc1. Unlike Ubc1, only the first two residues at the N-terminus are nonhelical, and there is no tight turn at the beginning of the helix. Another change occurs at the C-terminal end of helix A, which contains a short stretch of 3_{10} helix in *Arabidopsis* Ubc1. The helix in Ubc4 is entirely α-helical.

The other major structural differences between the two molecules occur in the loops consisting of residues 14–21 and 41–46 in Ubc4 (residues 16–23 and 43–48 in *Arabidopsis* Ubc1). Omitting these two loops (14 residues) from the superposition calculation lowers the root mean square difference to 1.23 Å. The loop from residues 14–21 in Ubc4 extends further away from the core structure of the rest of molecule, and in fact space-filling models show a small hole between the loop and the rest of the structure (Figure 6). However, the final $2F_o - F_c$ electron density map shows electron density that probably represents a water molecule in the center of the loop (Figure 2). A water molecule placed in this density would be within hydrogen-bonding distance of the main-chain N atom of Asp17.

The ubiquitin-accepting cysteine in both enzymes is located in a long extended stretch between the fourth strand of the β-sheet (S4) and the second α-helix (B) (Figures 4 and 5). An identical view of the Ubc4 molecule with a space-filling model is shown in Figure 6A. When identical residues between yeast Ubc4 and *Arabidopsis* Ubc1 are highlighted (green), it is clear that the 59 identical residues in these two enzymes are not randomly distributed but are located mainly on one surface (Figure 6). Thirty of these 59 residues are located within a 15-Å distance from the ubiquitin-accepting cysteine (Table III). Of these 30 residues, 10 are essentially buried (<1 Å^2 accessible area), and the remaining 20 residues are clustered on a surface that is also shared by the ubiquitin-accepting cysteine. One possible function of this conserved

Table III: Accessible Surface Area of Identical Residues in the Sequences of *S. cerevisiae* Ubc4 and *Arabidopsis* Ubc1[a]

	residue	surface (Å²)		residue	surface (Å²)
5	Lys	49.5	6	Arg	150.6
14	Leu	122.6	17	Asp	107.4
18	Pro	136.9	19	Pro	66.9
23	Ser	63.1	26	Pro	44.7
34	Trp	0.0	36	Ala	1.8
38	Ile	0.0	40	Gly	3.0
41	Pro	24.9	43	Asp	127.9
45	Pro	7.9	48	Gly	75.9
49	Gly	0.0	51	Phe	0.0
53	Leu*	0.0	54	Ser	15.4
57	Phe	3.4	60	Asp	72.5
61	Tyr	0.0	62	Pro	9.8
64	Lys	116.7	65	Pro*	31.3
66	Pro*	0.0	70	Phe*	2.1
76	His*	0.0	77	Pro*	0.0
78	Asn*	0.0	79	Ile*	0.0
81	Ala*	91.6	83	Gly*	0.9
85	Ile*	12.7	86	Cys*	32.5
87	Leu*	13.5	88	Asp*	62.6
89	Ile*	21.6	90	Leu*	17.5
93	Gln*	102.4	94	Trp*	24.3
95	Ser	51.8	96	Pro	48.8
104	Leu*	0.2	106	Ser*	21.1
107	Ile*	0.0	109	Ser*	61.6
110	Leu*	5.5	111	Leu*	0.6
113	Asp*	114.7	115	Asn*	47.8
116	Pro*	25.5	119	Pro*	56.1
123	Glu	146.9	125	Ala*	2.1
128	Tyr*	7.9	135	Tyr	4.9
140	Arg	123.0			

[a] A probe radius of 1.6 Å was used. Asterisks denote residues within 15 Å of the active site cysteine residue.

surface region is to provide a binding site for E1 and/or ubiquitin during the formation of $E2_{S-Ub}$. In the yeast *S. cerevisiae*, E1 is encoded by a single *UBA1* gene, while genes encoding ten different ubiquitin-conjugating enzymes have been identified. The ability of a single E1 enzyme to act on different E2s suggests that a common E1-binding site is present in all yeast E2 enzymes. Since mammalian E1 can also act on yeast E2 enzymes (Banerjee et al., 1993; Haas et al., 1991), a conserved E1-binding site is likely present in E2 enzymes from all species. The finding that E1 is also a highly conserved enzyme (Handley et al., 1991; Hatfield et al., 1990; McGrath et al., 1991) is consistent with this notion. The identification of a conserved surface area that may function in E1 binding leads to testable predictions, and experiments are in progress to address this possibility.

The ubiquitin-accepting cysteine in Ubc4 can also be approached from non-conserved surfaces. As in *Arabidopsis* Ubc1, the side chain of Cys86 is exposed and sits in a slight depression on the surface. There are loops containing a number of hydrophilic residues on either side of Cys86, consisting of residues 88–94 and 115–123. The first loop shows a high degree of sequence identity between the two proteins, but there is considerably less in the second. It is noteworthy that the loop consisting of residues 88–94 adjacent to this crucial cysteine residue can be almost completely superimposed on its counterpart in *Arabidopsis* Ubc1. However, the loop containing residues 115–123 is generally further from the cysteine side chain (1–2 Å). Examination of the regions around the cysteine reveals that the cysteine side chain is slightly more exposed in yeast Ubc4 than in *Arabidopsis* Ubc1 (Figure 7). For example, there is a tyrosine residue in Ubc1 that is about 4.5 Å from the cysteine side chain. The corresponding residue in Ubc4 is Asn, which has a much smaller side chain.

In summary, the similarity of the structure of yeast Ubc4 to *Arabidopsis* Ubc1 suggests that all class I ubiquitin-conjugating enzymes will have the same general conformation, even though there is limited amino acid sequence identity among them. Since the tertiary structure around the ubiquitin-accepting cysteine residue is so highly conserved, it is likely that these regions are specific for ubiquitin and/or E1 binding, while the surfaces with few conserved residues are utilized in binding the protein targeted for ubiquitination. Although reasonable guesses for the residues specifically involved in ubiquitin binding may be derived from mutagenesis studies, definitive localization will require determination of the structure of a ubiquitin–E2 conjugate. The preparation and purification of such a stable adduct between ubiquitin and *Arabidopsis* Ubc1 has been recently reported (Sullivan & Vierstra; 1993).

REFERENCES

Banerjee, A., Gregori, L., Xu, Y., & Chau, V. (1993) *J. Biol. Chem. 268*, 5668–5675.

Brünger, A. T. (1990) *Acta Crystallogr., Sect. A 46*, 46–57.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458–460.

Carson, M. (1987) *J. Mol. Graphics 5*, 103–106.

Ciechanover, A., Finley, D., & Varshavsky, A. (1984) *Cell 37*, 57–66.

Cook, W. J., Jeffrey, L. C., Sullivan, M. L., & Vierstra, R. D. (1992) *J. Biol Chem. 267*, 15116–15121.

Finley, D., & Chau, V. (1991) *Annu. Rev. Cell Biol. 7*, 25–69.

Girod, P.-A. & Vierstra, R. D. (1992) *J. Biol. Chem. 268*, 955–960.

Glotzer, M., Murray, A. W., & Kirschner, M. W. (1991) *Nature 349*, 132–138.

Haas, A. L., & Rose, I. A. (1982) *J. Biol. Chem. 257*, 10329–10337.

Haas, A. L., Reback, P. B., & Chau, V. (1991) *J. Biol. Chem. 266*, 5104–5112.

Handley, P. M., Mueckler, M., Siegel, N. R., Ciechanover, A., & Schwartz, A. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 258–262.

Hatfield, P. M., Callis, J., & Vierstra, R. D. (1990) *J. Biol. Chem. 265*, 15813–15817.

Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.

Hershko, A., & Ciechanover, A. (1992) *Annu. Rev. Biochem. 61*, 761–807.

Hochstrasser, M., Ellison, M. J., Chau, V., & Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4606–4610.

Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. R. (1987) *J. Appl. Crystallogr. 20*, 383–387.

Jancarik, J., & Kim, S.-H. (1991) *J. Appl. Crystalllogr. 24*, 409–411.

Jentsch, S. (1992a) *Trends Cell Biol. 2*, 98–103.

Jentsch, S. (1992b) *Annu. Rev. Genet. 26*, 179–207.

Johnson, E. S., Bartel, B., Seufert, W., & Varshavsky, A. (1992) *EMBO J. 11*, 497–505.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Koken, M., Reynolds, P., Bootsma, D., Hoeijmakers, J., Prakash, S., & Prakash, L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3832–3836.

Luzzati, V. (1952) *Acta Crystallogr. 5*, 802–810.

McGrath, J. P., Jentsch, S., & Varshavsky, A. (1991) *EMBO J. 10*, 277–237.

Pickart, C. M. (1988) in *Ubiquitin* (Rechsteiner, M., Ed.) pp 77–99, Plenum Press, New York.

Reynolds, P., Koken, M. H. M., Hoeijmakers, J. H. J., Prakash, S., & Prakash, L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3832–3836.

Schneider, R., Eckerskorn, C., Lottspeich, F., & Schweiger, M. (1990) *EMBO J. 9*, 1431–1435.

Seufert, W., & Jentsch, S. (1990) *EMBO J. 9*, 543–550

Shanklin, J., Jabben, M., & Vierstra, R. D. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 359–363.

Sullivan, M. L., & Vierstra, R. D. (1991) *J. Biol Chem. 266*, 23878–23885.

Sullivan, M. L., & Vierstra, R. D. (1993) *J. Biol. Chem. 268*, 8777–8780.

Treier, M., Seufert, W., & Jentsch, S. (1992) *EMBO J. 11*, 367–372.

Wing, S. S., Dumas, F., & Banville, D. (1992) *J. Biol. Chem. 267*, 6495–6501.

Zhen, M., Heinlein, R., Jones, D., Jentsch, S., & Candido, E. P. M. (1993) *Mol. Cell. Biol. 13*, 1371–1377.