Crystal Structure of the Cyclin-Specific Ubiquitin-Conjugating Enzyme from Clam, E2-C, at 2.0 Å Resolution[†,‡]

Fan Jiang and Ravi Basavappa*

Department of Biochemistry and Biophysics, University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, New York 14642

Received January 20, 1999; Revised Manuscript Received March 23, 1999

ABSTRACT: The destruction of the cyclin B protein is necessary for the cell to exit from mitosis. The destruction of cyclin B occurs via the ubiquitin/proteasome system and involves a specific ubiquitin-conjugating enzyme (Ubc) that donates ubiquitin to cyclin B. Here we present the crystal structure of the cyclin-specific Ubc from clam, E2-C, determined at 2.0 Å resolution. The E2-C enzyme contains an N-terminal extension in addition to the Ubc core domain. The N-terminal extension is disordered, perhaps reflecting a need for flexibility as it interacts with various partners in the ubiquitination system. The overall structure of the E2-C core domain is quite similar to those in previously determined Ubc proteins. The interaction between particular pairs of E2-C proteins in the crystal has some of the hallmarks of a functional dimer, though solution studies suggest that the E2-C protein exists as a monomer. Comparison of the E2-C structure with that of the other available Ubc structures indicates conserved surface residues that may interact with common components of the ubiquitination system. Such comparison also reveals a remarkable spine of conserved hydrophobic residues in the center of the protein that may drive the protein to fold and stabilize the protein once folded. Comparison of residues conserved only among E2-C and its homologues indicates surface areas that may be involved in mitotic-specific ubiquitination.

The ubiquitin/proteasome system provides a means for the rapid and selective destruction of proteins. In this system, ubiquitin molecules are conjugated selectively to the target protein. The 26S proteasome then recognizes and hydrolyzes the ubiquitinated protein. The ubiquitin/proteasome system is important in diverse areas of cellular function (for reviews, see refs *1*−*4*). For example, this system destroys ephemeral and misfolded proteins (*5*, *6*), cleaves proteins for presentation as antigens to the MHC (*7*, *8*), matures transcription activators from precursor forms (*9*), and activates phytochrome (*10*). The system also controls several key transitions in the cell cycle by removing, when appropriately cued, proteins that block further progression. For example, the ubiquitin/proteasome system destroys the cyclin-dependent kinase inhibitor proteins to allow entry into S phase (*11*, *12*), the "glue" proteins binding the sister chromatids together to permit entry into anaphase (*13*, *14*), and cyclin B to permit exit from mitosis (*15*, *16*). Thus, it has become abundantly clear that the regulated destruction of proteins is of paramount importance for proper cell function.

Ubiquitin is conjugated to the target protein through the coordinated action of three enzyme activities designated E1, E2, and E3. The E1 or ubiquitin-activating enzyme forms, in an ATP-dependent manner, a thioester linkage between its active site cysteine and the carboxy terminus of ubiquitin. The ubiquitin then is transferred from E1 to the active site cysteine in the E2 or ubiquitin-conjugating enzyme (Ubc)[1] through a trans-thiol esterification reaction. The Ubc protein in turn contributes the ubiquitin to the target protein. The ubiquitin molecule is attached to the target protein through an isopeptide linkage between the carboxy terminus of ubiquitin and the ϵ-amino group of a lysine in the target protein. Some Ubc proteins may transfer the ubiquitin to the target protein directly. However, most Ubc proteins require the intervention of a separate E3 ligase activity. Formation of polyubiquitin chains on the target protein, perhaps by the iterative covalent addition of ubiquitin to previously attached ubiquitin molecules, is necessary for the efficient recognition and hydrolysis of the targeted protein by the proteasome. The specificity in targeting a protein for ubiquitination resides primarily in cognate pairs of E2 and E3 enzymes. Different E2/E3 pairs exist for ubiquitinating different sets of target proteins.

In the ubiquitination/destruction machinery, the Ubc proteins are quite remarkable. Despite their relatively modest size (usually 14−35 kDa), they must interact with at least three and perhaps four different proteins: ubiquitin, E1 protein, and E3 protein and/or the target protein. Even more remarkable, all these interactions must necessarily involve the active site cysteine in the E2 protein. Further, some E2 proteins have terminal extensions. This apparently reflects the need to bind to diverse E3/targets as well as the need for modulating other aspects of function such as intracellular localization. Those E2 proteins that lack terminal extensions

[†] This work was supported in part by NIH Grant GM 57536 to R.B.

[‡] The coordinates of the E2-C crystal structure have been deposited in the Protein Data Bank, Brookhaven National Laboratories, under file name 2E2C.

* Corresponding author.

[1] Abbreviations: Ubc, ubiquitin-conjugating enzyme; APC, anaphase-promoting complex.

10.1021/bi9901329 CCC: $18.00
Published on Web 04/28/1999

are designated class I, whereas those with extensions at the C-terminus, N-terminus, or both termini are designated class II, class III, and class IV, respectively.

The E2-C protein from clam is a class III Ubc enzyme. Of its 177 amino acid residues, the first 25 residues comprise an N-terminal extension and the remaining residues form the E2 core domain. The E2-C protein and its cognate ligase, termed the cyclosome or the anaphase promoting complex (APC), conjugate ubiquitin to cyclin B, thereby tagging the cyclin B for destruction and permitting subsequent exit from mitosis (*17*). Dominant-negative mutants of E2-C block mammalian cells in metaphase (*18*). This observation suggests that the E2-C protein also may be required for entry into anaphase, perhaps by ubiquitinating the glue proteins that must be destroyed to allow sister chromatid segregation. Indeed, the "destruction box" sequence necessary for recognition by the E2-C/APC is present as a single copy in mitotic cyclins and as a double copy in the glue proteins. Several homologues of E2-C have been identified. The E2-C homologue in *Xenopus*, Ubc-x, has been implicated in cyclin B destruction (*19*). The homologue in humans, UbcH10, can substitute for E2-C in clam extracts. Dominant-negative UbcH10 mutants, like dominant-negative E2-C mutants, arrest mammalian cells in metaphase. The homologue in *Schizosaccharomyces pombe*, UbcP4, is essential for entry into anaphase (*20*).

Here we report the crystal structure analysis of the E2-C enzyme. Analysis of sequence features conserved in all the Ubc proteins with available crystal structures and in E2-C homologues reveals surface patches that may be involved in interactions with the conserved E1 and ubiquitin proteins. This comparison also reveals a striking buried spine of hydrophobic residues that may drive the protein to fold. Comparison of sequence features with E2-C homologues has allowed the tentative identification of surfaces interacting with the APC/cyclosome and/or the target protein.

EXPERIMENTAL PROCEDURES

An BL21(DE3) pLysS *Escherichia coli* strain harboring the expression plasmid with the E2-C gene was provided by J. V. Ruderman (Harvard Medical School, Boston, MA). The cultures were grown at 37 °C while they were constantly shaken in a rotor in LB broth with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. Cultures for protein overexpression were begun by inoculating with overnight cultures at a volume ratio of 1:100. The cultures, upon growing to an optical density of approximately 0.6 A_{595}, were induced by adding IPTG to a concentration of 1.0 mM. The induced cultures were incubated for 3 h, harvested by centrifugation (10 min at 8000*g*), and frozen overnight at −80 °C.

Cells were lysed by sonication; the cell debris was pelleted by centrifugation and the supernatant collected. The protein was purified by a three-step process. In the first step, much of the nucleic acid component and many of the proteins were precipitated by adding polyethyleneimine (PEI) (Sigma) to a concentration of 0.2% (v/v), incubating on ice for 30 min, centrifuging at ∼10000*g* for 10 min at 4 °C, and collecting the supernatant. The PEI precipitation was repeated once to improve the removal of nucleic acids and unwanted proteins. In the second step, supernatant from the second PEI precipitation was applied to a POROS HQ20 anion exchange column (1.7 mL) on a BioCad Sprint chromatography system (PerSeptive Biosystems) equilibrated with 20 mM Tris and bis-Tris Propane at pH 8.5, and the E2-C protein was collected in the flow-through. Since subsequent IEF analysis indicated oxidation of the cysteine residues in the protein, 1.0 M DTT was added to the fraction tubes prior to collection such that final concentration of DTT in the fraction was 1.0 mM. In the third step, the flow-through from the anion exchange column was loaded onto a POROS HS20 cation exchange column (1.7 mL) equilibrated in 20 mM MES and HEPES at pH 4.5. A NaCl gradient was developed from 0 to 1.5 M in 40 column volumes at a flow rate of 10 mL/min. The E2-C protein elutes at ∼100 mM NaCl. DTT was added as above to the fraction tubes. The sample at this stage was >98% pure as assessed by SDS−PAGE and silver staining. Scaleup of the purification was accomplished by employing the POROS columns multiple times. In preparation for crystallization, the E2-C purified as above was concentrated to ∼14 mg/mL with a Centricon 10 instrument (Amicon).

The E2-C was crystallized using the vapor diffusion/hanging drop method. Crystallization conditions were screened by a sparse matrix (*21*) using precipitation reagents provided in Hampton Screens I and II (Hampton Research). The droplet was prepared by mixing 1 μL of the protein solution with 1 μL of the precipitant solution and then suspended over 750 μL of the precipitant solution. Crystals were observed in the droplet suspended over the reservoir containing 30% PEG 4000, 0.1 M Tris-HCl (pH 8.5), and 0.2 M LiSulfate. The crystals were observed first after 2 days and stopped growing after ∼1 week.

A crystal with an approximate size of 0.2 mm × 0.2 mm × 0.1 mm was used for collecting the diffraction data. The crystal was mounted in a loop and flash-frozen in the mother liquor. The entire data set was collected from a single crystal at −176 °C. X-rays were produced using a rotating copper anode on a Rigaku RU200 generator operating at 50 kV and 100 mA and focused with a Yale-design mirror system (Molecular Structure Corp.). The oscillation method with 1° sweeps was used to collect the data. Diffraction data up to a d_{min} of 2.0 Å were collected using an R-axis II imaging phosphor system (Molecular Structure Corp.). The data were reduced with DENZO and SCALEPACK (*22*). Two possible space groups assignments for this crystal were $P3_121$ ($a = b = 46.334$ Å and $c = 150.52$ Å, with one molecule in the asymmetric unit, and weak violations of the systematic absences) and $C2$ ($a = 80.35$ Å, $b = 46.0$ Å, $c = 150.56$ Å, and $\beta = 90.01°$, with three molecules in the asymmetric unit, and no violations of the systematic absences). The higher-symmetry space group ($P3_121$) was chosen since the systematic absence violations were quite small (<2.5% of the allowed 00*l* reflections). The diffraction data statistics are summarized in Table 1.

The phase problem for the E2-C crystal structure was solved by the molecular replacement method. The structure of the *Arabidopsis thaliana* Ubc1 protein (PDB file name 1AAK) was chosen as the basis for the molecular replacement model. The *A. thaliana* Ubc1 sequence is 44% identical to that of the E2-C when the sequences are aligned with FASTA (*23*). The molecular replacement model was prepared from the 1AAK model based on the sequence alignment. Identical residues were left unchanged; homolo-

Table 1: Summary of Crystallographic and Model Data

unit cell dimensions (space group $P3_121$)	
a and b (Å)	46.33
c (Å)	150.52
resolution (Å)	5.0−2.0
data quality	
no. of unique reflections	13327 (789)[a]
completeness (%)	99.4 (92.7)
average $I/\sigma(I)$	11.7 (3.7)
R_{sym}[b] (%)	6.1 (35.8)
no. of residues in model monomer	22−177
no. of protein atoms in the model monomer	1250
no. of solvent atoms in the asymmetric unit	172
R-factors[c]	
$R_{working}$	0.2156
R_{free}	0.2972
stereochemistry	
rms deviation from ideal values	
bond lengths (Å)	0.007
bond angles (deg)	1.343
dihedrals (deg)	23.2

[a] The numbers in parentheses are for the highest-resolution shell (2.05−2.00 Å). [b] $R_{sym} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|/\Sigma_h\Sigma_i I(h)$, where $I_i(h)$ is the ith observation of the intensity of reflection with index h and $\langle I(h)\rangle$ is the average value of all observations of $I(h)$ including its Friedel and symmetry-equivalent reflections. [c] R-factor $= \Sigma_h|F_o(h) - F_c(h)|/\Sigma_h F_o(h)$, where $F_o(h)$ and $F_c(h)$ are the observed and calculated structure factor magnitudes, respectively. Prior to refinement, 10% of the data were randomly selected and extracted from the data set. $R_{working}$ is the R-factor for the data used in the refinement. R_{free} is the R-factor for the extracted data.

gous residues were changed to the corresponding residue in E2-C, and nonhomologous residues were changed to alanine. The rotation and translation functions solutions were found by using REPLACE (*24*) and X-PLOR (*25*).

The initial R-factor using reflections from 10 to 4 Å was 46.9%. At this stage, an $F \geq 2\sigma(F)$ cutoff was applied and 10% of the remaining diffraction data were randomly extracted from the data set and set aside for later use in R_{free} cross-validation (*26*). The refinement was performed against the remaining 90% of the data ($R_{working}$). Positional refinement using X-PLOR reduced the $R_{working}$ to 36.5% and the R_{free} to 40.1%. An electron density map using sigmaA-weighted $2F_o - F_c$ coefficients (*27*) and model-derived phases then was calculated. Map display and model building were performed with O (*28*) on a SGI Indigo2 graphics workstation. The model was changed to correspond to the sequence of the E2-C protein (residues 27−175). Further refinement proceeded by cycles of simulated annealing using X-PLOR followed by interactive graphics sessions for inspecting the model and electron density and also for rebuilding portions of the model as necessary. During the course of the refinement, residues 22−26, 176, and 177 were added to the model. These residues have no counterparts in the Ubc1 search model. Attempts to convincingly model residues 1−21 were not successful despite extensive efforts. In the later stages of refinement, 172 solvent atoms were added to the model. The quality of the model was assessed using PROCHECK (*31*) and the WHAT_CHECK subset of WHATIF (*32*). The statistics of the final model are given in Table 1.

Solution Experiments for Determining the Quaternary State. Gel filtration experiments were performed with a Superdex 75 HR 10/30 column (Pharmacia Biotech) attached to a FPLC II system. The column was calibrated using four proteins with known molecular masses (ribonuclease A, chymotrypsinogen A, ovalabumin, and albumin with specified molecular masses of 13.7, 25.0, 43.0, and 67.0 kDa, respectively; Pharmacia Biotech). A 100 μL sample of the E2-C protein [initially 5 mg/mL in 150 mM NaCl, 20 mM Tris, and 2 mM DTT (pH 7.6)] was applied to the column equilibrated in the same buffer. The oligomeric state of E2-C in solution also was probed by NMR self-diffusion measurements as developed by Byrd and colleagues (*33*). In this technique, pulsed field gradient stimulated echo experiments are used to determine the translational diffusion coefficient of the protein, which can then be used to calculate the apparent molecular mass. The experiment was calibrated by first measuring the translation diffusion coefficients of monomeric proteins with known molecular masses (fibronectin type III, lysozyme, ospA, and ovalbumin). The diffusion coefficient of E2-C (at 5 mg/mL) was then measured and its apparent molecular mass calculated by interpolation using the calibration curve.

FIGURE 1: Overall fold of the E2-C structure (stereoview). The α-helices and β-strands are shown in light and dark shades, respectively. The active site cysteine (Cys-114) is shown as a ball-and-stick rendering. The modeled portion of the amino-terminal extension, seen at the top of model, makes contacts with a symmetry-related molecule. This figure was created with Insight II (MSI Corp.).

RESULTS AND DISCUSSION

The crystal structure of the mitotic-specific ubiquitin-conjugating enzyme has been determined using data to 2.0 Å resolution. All residues other than the first 21 N-terminal residues have been modeled. The E2-C protein is an α/β protein containing one four-stranded antiparallel β-sheet and four α-helices (Figure 1). Three of these helices flank two opposite edges of the sheet, and one helix lays diagonally across one broad face of the sheet. The other face of the sheet is exposed to solvent. One turn of a 3_{10}-helix (residues 116−118) is located between the fourth strand of the sheet and the second α-helix. The active site cysteine (residue 114) is situated in a segment between the fourth strand of the sheet and the 3_{10}-helix. The overall shape of the E2-C protein is roughly that of an elongated triangular prism with the

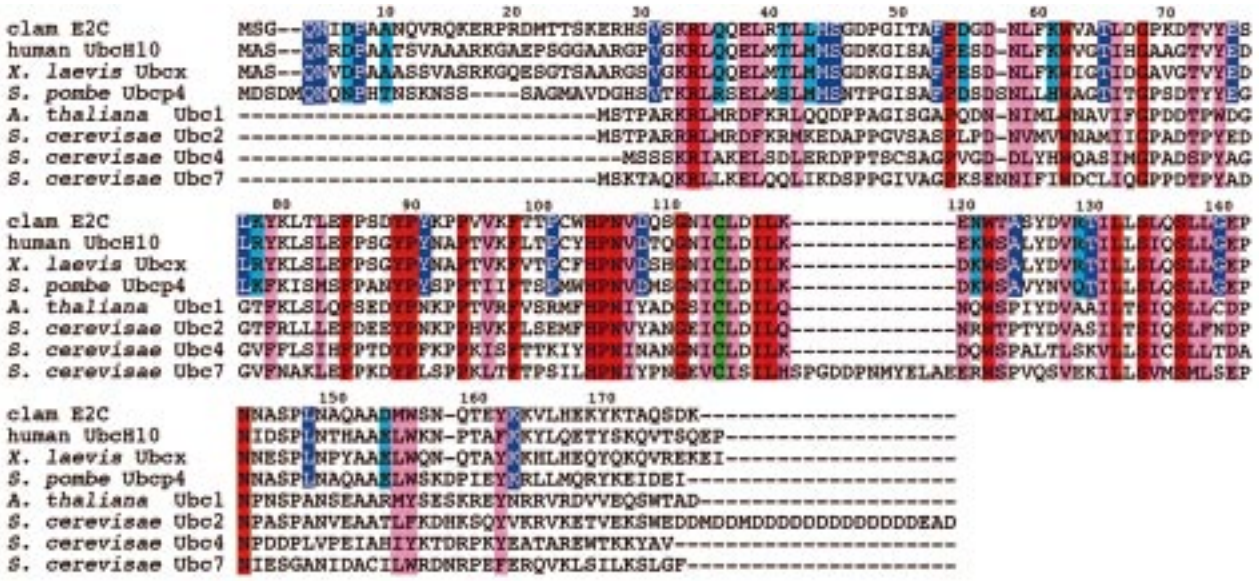

FIGURE 2: Sequence alignment of E2-C with its homologues and with the Ubc enzymes for which the crystal structures are available. The color coding is as follows: red, identical in all aligned sequences; light purple, highly conserved in all sequences; blue, identical within the E2-C family and no more than one nonfamily sequence contains that residue at that position; cyan, highly conserved only within the E2-C family (changes are conservative, and no more than one nonfamily sequence contains that residue at that position); and green, active site cysteine. The sequences were aligned using CLUSTALW (*61*).

N-terminal arm extending from one triangular face. The sulfur atom of the active site cysteine sits highly exposed midway along one long edge of the prism and in a slight local depression.

Disorder of the N-Terminal Region. The first 21 N-terminal residues of the protein are disordered in the X-ray structure. It is not yet clear whether the N-terminal extension is necessary for the cell cycle-specific function of E2-C. The *Xenopus laevis* homologue of yeast Ubc4 does not have an N-terminal extension but does mediate cyclin ubiquitination (*34*), though with a much reduced activity relative to that of E2-C in clam extracts (*17*). This suggests that the N-terminal extension is not essential for ubiquitination activity per se. The lack of a clearly established function for the N-terminal extension is in contrast to class II enzymes in which various carboxyl-terminal segments have been demonstrated to be structured (*35*), provide target specificity (*36*, *37*), possess ligase activity (*38*), or determine intracellular location (*39*).

The N-terminal domain of E2-C may become ordered upon binding to other proteins. Such induced local folding upon interaction with another molecule may be critically important to function by allowing the domain to recognize a variety of targets. This role has been suggested in the case of, for example, the disordered segment in $p21^{Waf1/Cip1/Sdi1}$ (*40*) and the disordered domain in the apical region of GroEL (*41*). Mitotic-specific protein degradation occurs through a complex network of interactions that controls both the spatial and temporal aspects of destruction. The role of the flexible N-terminal segment in E2-C should become clearer as the nature of these interactions becomes better understood.

Possibility of a Dimeric State. A point of some contention has been whether the Ubc enzymes form functional dimers. Some laboratories have reported certain E2 proteins to exist as dimers on the basis of gel filtration chromatography (*38*, *42*−*44*), genetic analysis (*37*), two-hybrid system assays (*45*, *46*), chemical cross-linking (*35*, *47*), or the presence of specific autoubiquitination (*48*). However, most other studies

FIGURE 3: Hydrophobic spine formed by conserved buried residues in the available Ubc crystal structures (stereoview). The location of these residues is shown in the context of the E2-C crystal structure. The side chains of the conserved internal residues are shown as a space-filling rendering. Nitrogen and oxygen atoms are shown as dark spheres. The active site cysteine also is shown in space-filling mode with the sulfur atom depicted as a light sphere. This figure was prepared using SETOR (*62*).

either characterize E2 proteins as or assume E2 proteins are monomers. The crystal structures of the class I E2 proteins exhibit interaction surfaces that are more typical of crystal packing interactions than of functional interfaces, suggesting that these proteins are monomers. This includes the yeast Ubc4 protein, which has been suggested to be dimeric on the basis of cross-linking and gel filtration experiments.

The E2-C forms what may be a functional dimer in the crystal. This putative dimer is formed by two monomers related by a crystallographic 2-fold axis and involves residues

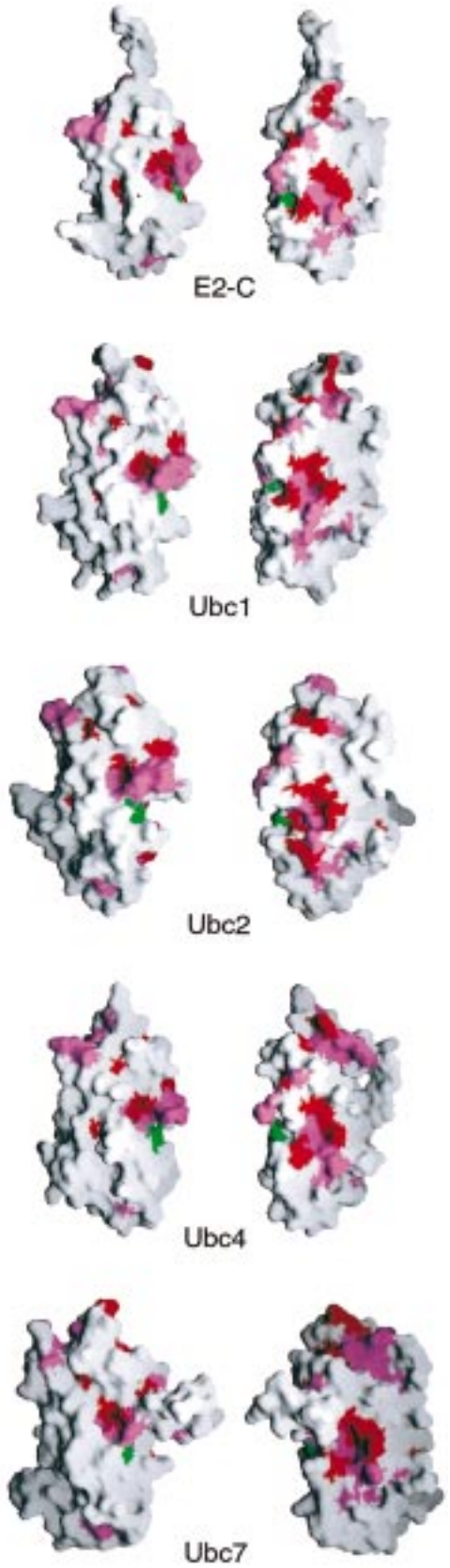

FIGURE 4: Comparison of conserved surface features in the available Ubc crystal structures. The two views for each protein are related by a 120° rotation about a vertical axis in the plane of the paper. The color coding is the same as that described in the legend of Figure 2: identical residues are red, highly conserved residues are light purple, and the active site cysteine is green. This figure was prepared using GRASP (*63*).

36, 39, 40, 43, 44, 53−57, 59, 63, 65, and 80. Several observations suggest that this interaction may be of functional importance. First, the surface area of interaction is approximately 1318 Å^2, which is within the range (1200−1500 Å^2) for known functional dimers (*49*). Second, the interface surface is formed by a patch of residues that is primarily hydrophobic in the center and hydrophilic at the periphery. Such an interface can be indicative of a functionally dimeric protein (*50*). Third, part of the interface consists of one helix from one monomer packing against its symmetry mate in the other monomer with a "knobs into holes" motif. A high degree of complementarity also is indicative of a functional interface (*50*).

Despite the crystal structure having the hallmarks of a functional dimer, two solution state experiments suggest that the E2-C protein is monomeric. First, gel filtration experiments at relatively high initial E2-C concentrations (5 mg/mL) indicate an apparent molecular mass of 30 kDa. The calculated molecular mass of the monomer is 20.1 kDa. Second, NMR self-diffusion experiments at 5 mg/mL E2-C also indicate an apparent molecular mass of 30 kDa. The high apparent molecular mass likely is due to the somewhat extended shape of the molecule and to the disorder of the first 20 N-terminal residues. Therefore, the putative dimer in the crystal system may represent a weak specific interaction of unknown functional significance. Experiments, such as functional and biophysical characterization of appropriate site-directed mutants chosen on the basis of this structure, should help to resolve the question of the quaternary state of the Ubc.

Comparison to Other E2 Structures. The crystal structures of several other E2 proteins have been reported: Ubc1 from *A. thaliana* (*51*), presumably involved in DNA repair; Ubc4 from *Saccharomyces cerevisae* (*52*), primarily involved in ubiquitinating short-lived and misfolded proteins; Ubc7 from *S. cerevisiae* (*53*), a nonessential gene linked to the ubiquitination of a variety of substrates; and Rad6 (also known as Ubc2) from *S. cerevisiae*, involved in the degradation of short-lived proteins and in DNA repair (*58*). For these proteins and E2-C, the sequence homology is relatively low, within the range of 35−45%. Despite the disparity in protein targets and the relative lack of sequence homology, the overall fold adopted by these proteins is remarkably similar. The structure of a Ubc-like protein, Ubc9 from human (100% identical to Ubc9 from mouse), has also been determined (*54*). The Ubc9 protein initially was presumed to be involved in the ubiquitination of a diverse set of target proteins but recently was found to be involved in the conjugation of the ubiquitin-like protein SUMO-1 (*55*−*57*). The Ubc9 structure is not included in the comparison since it is not a bona fide ubiquitin-conjugating enzyme. When the Ubc proteins are superimposed using corresponding secondary structural elements, the rms deviation in corresponding α-carbon atoms is in the range of 0.8−1.25 Å. When the positions of corresponding α-carbons are compared between E2-C and the other Ubc structures, the greatest deviations are observed to occur in the α-helices at the N- and C-termini and the loop (residues 144−148) near the active site. This close similarity among Ubc protein structures allows us to identify potentially important conserved structural features.

Analysis of residues that are identical or highly conserved in all available Ubc crystal structures and E2-C homologues

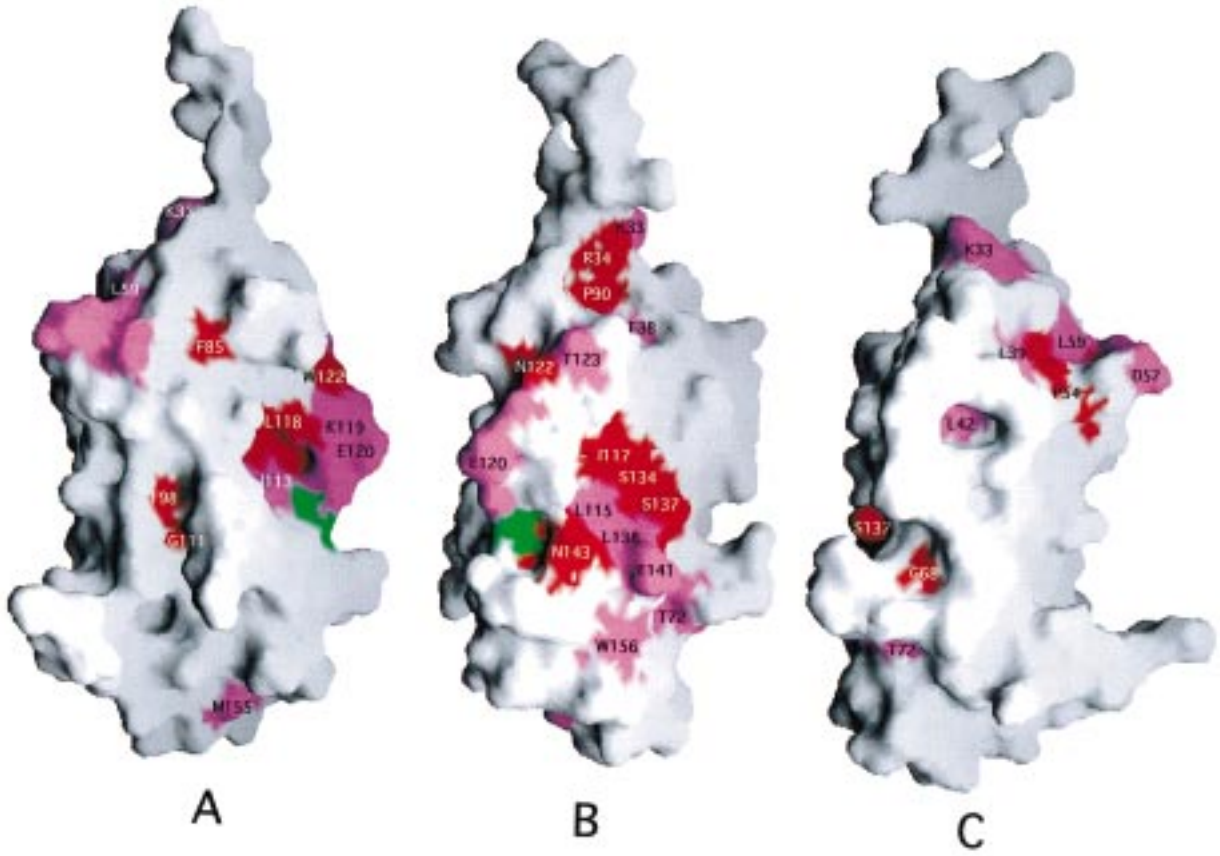

FIGURE 5: Mapping of generally conserved residues that contribute to the surface of the E2-C protein. The views in panels B and C are related to the view depicted in panel A by rotations of 120° and 240°, respectively, about a vertical axis in the plane of the page. The areas shaded red and light purple are contributed by residues that are identical or highly conserved, respectively, in the available crystal structures. This figure was prepared using GRASP (*63*).

(Figure 2) reveals that these residues partition into two sets on the basis of their general location in the protein. One set contains residues that are totally buried in the protein, and the other set contains residues that have atoms on the surface of the protein. The conserved buried residues are primarily hydrophobic (Figure 3). The residues in this set form a remarkable spine down the center of the protein. The formation of the hydrophobic spine may drive the folding of the protein, and the spine may stabilize the protein once it is folded. The residues are widely dispersed in the linear sequence and are located primarily in the secondary structural elements. The particular location of these residues may guide the packing of the secondary structural elements against each other. The very highly conserved HPN triplet (residues 104−106) does not seem to be an integral part of the spine but is adjacent to it. Its close proximity to the active site cysteine suggests that it may play a more direct role in ubiquitin transfer and not merely stabilize the protein as a whole.

Comparison of the disposition of the conserved surface residues indicates these residues form similar surface patches on their respective proteins (Figure 4). If the protein is viewed as forming an extended triangular prism with the active site on one long edge, then most of the conserved patches are situated in the two long faces forming the edge (Figure 5A,B). The remaining face, opposite the active site cysteine, is fairly bereft of conserved surface residues (Figure 5C). This finding is in overall agreement with observations made previously with regard to the other Ubc crystal structures. The conserved surface patches may be involved in interacting with the conserved elements of the ubiquitination system, namely, ubiquitin and E1. Indeed, the two most striking of these conserved patches together straddle the active site cysteine. The surface of interaction with E1 and ubiquitin would be expected to include areas near the active site. These two patches possess a significant hydrophobic character. The large patch shown in Figure 5B is rather hydrophobic in the center and contains negatively charged atoms along much of the periphery (not shown). The remaining significant patch is contributed by a pair of basic residues in the N-terminal helix, Lys-33 and Arg-34 in E2-C. The importance of this patch in intermolecular interactions is corroborated by molecular genetic and biochemical experiments that implicate the equivalent residues in Ubc1 from *A. thaliana* (*59*) and Ubc2b from human (*60*) in binding to E1 protein.

Mapping of Residues Conserved among E2-C and Its Functional Homologues. Analysis of residues conserved only among E2-C and its homologues and not among Ubcs in general can provide clues as to what features are specific to the function of the E2-C family. These conserved residues are shown in Figure 6. It is notable that all of these residues contain atoms that contribute to the protein surface. None of these residues is totally buried in the protein. This finding is quite consistent with the notion that these residues are conserved to provide a shared molecular interface (such as the second set of generally conserved residues discussed above) and not to provide conserved internal scaffolding (such as the first set discussed above). As expected, patches of E2-C family conserved residues exist near the active site

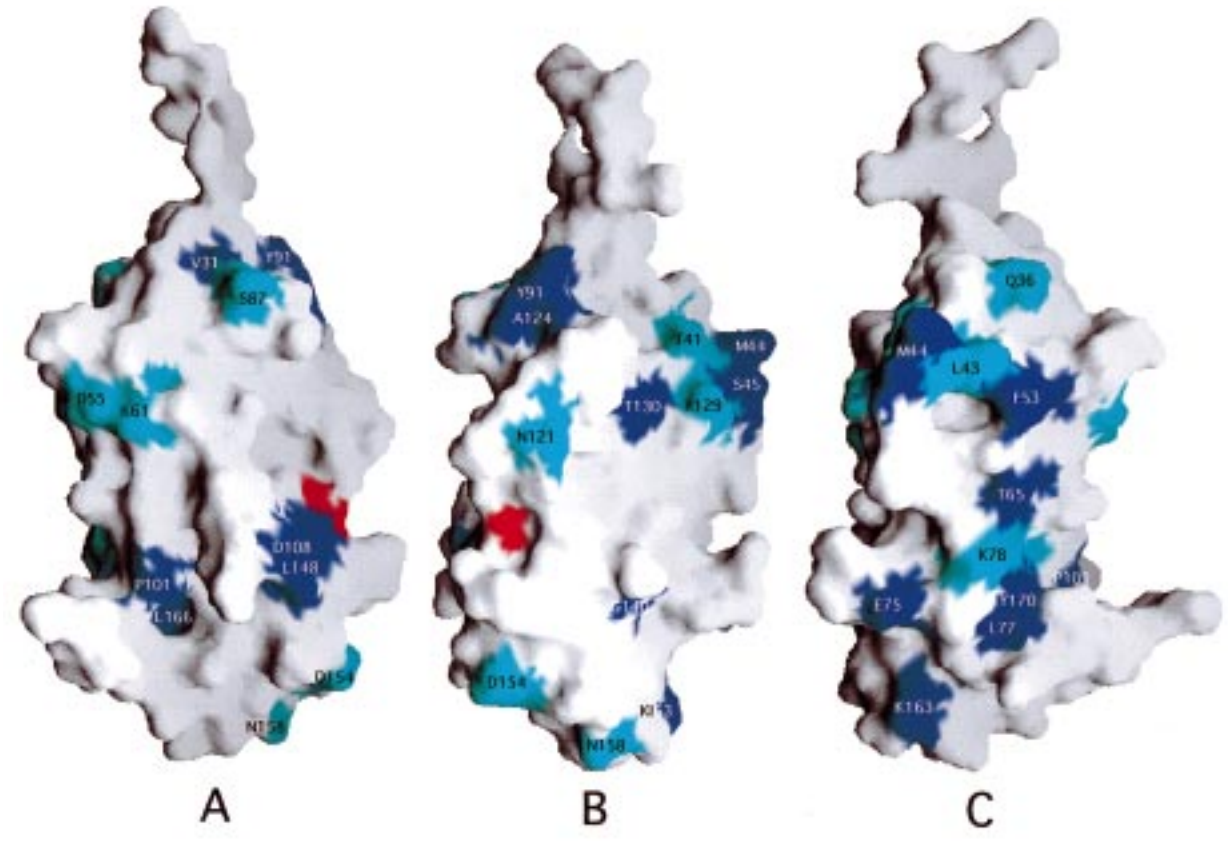

FIGURE 6: Mapping of residues conserved exclusively in the E2-C family onto the surface of the E2-C protein. The views in panels B and C are related to the view in panel A by rotations of 120° and 240°, respectively, about a vertical axis in the plane of the page. The areas shaded blue and cyan are contributed by residues that are identical or highly conserved, respectively, in E2-C and its known homologues. The red area corresponds to the active site cysteine. This figure was prepared using GRASP (*63*).

cysteine (Figure 6A), as observed for the more generally conserved Ubc residues. The other two faces also contain significant conserved surfaces (Figure 6B,C). In particular, the side opposite the active site cysteine contains an intriguing track of conserved residues along the length of the face (Figure 6C). This face also contains a significant hydrophobic patch contributed by residues Met-44, Phe-53, and Leu-143. Thus, the E2-C protein contains significant conserved surface features on all three major faces. This may be relevant to the observation that the APC/cyclosome is a large, multimeric complex and could, in principle, contact the protein on different faces. The E2-C crystal structure provides a guide for designing molecular genetic, biochemical, and biophysical experiments that will help us better understand the stereochemistry of targeted protein ubiquitination.

ACKNOWLEDGMENT

We thank Dr. Joan Ruderman (Harvard Medical School) for the bacterial strain harboring the E2-C plasmid and for helpful discussions, Andrey Mezhiba, Dr. Scott Kennedy, and Dr. Shohei Koide for performing the NMR self-diffusion experiment, Tom Colby and Michael Strickler for help with data collection, and Dr. Barry Goldstein and Dr. Cecile Pickart (Johns Hopkins University) for a critical reading of the manuscript.

REFERENCES

1. Jennissen, H. P. (1995) *Eur. J. Biochem. 231*, 1−30.
2. Wilkinson, K. D. (1995) *Annu. Rev. Nutr. 15*, 161−189.
3. Hochstrasser, M. (1996) *Cell 84*, 813−815.
4. Hershko, A., and Ceichanover, A. (1998) *Annu. Rev. Biochem. 67*, 425−479.
5. Ciechanover, A., Finley, D., and Varshavsky, A. (1984) *Cell 37*, 57−66.
6. Seufert, W., and Jentsch, S. (1990) *EMBO J. 9*, 543−550.
7. Michalek, M. T., Grant, E. P., Gramm, C., Goldberg, A. L., and Rock, K. L. (1993) *Nature 363*, 552−554.
8. Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A. L. (1994) *Cell 78*, 761−777.
9. Pahl, H. L., and Baeuerle, P. A. (1996) *Curr. Opin. Cell Biol. 8*, 340−347.
10. Shanklin, J., Jabben, M., and Vierstra, R. D. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 359−363.
11. Schwob, E., Bohm, T., Mendenhall, M. D., and Nasmyth, K. (1994) *Cell 79*, 233−244.
12. Pagano, M., Tam, S. W., Theodoras, A. M., Beer-Romero, R., Del Sal, G., Chau, V., Yew, P. R., Draetta, G. F., and Rolfe, M. (1995) *Science 269*, 682−685.
13. Cohen-Fix, O., Peters, J. M., Kirschner, M. W., and Koshland, D. (1996) *Genes Dev. 10*, 3081−3093.
14. Funabiki, H., Yamano, H., Kumada, K., Nagao, K., Hunt, T., and Yanagida, M. (1996) *Nature 381*, 438−441.
15. Murray, A. W., Solomon, M. J., and Kirschner, M. W. (1989) *Nature 339*, 280−286.
16. Glotzer, M., Murray, A. W., and Kirschner, M. W. (1991) *Nature 349*, 132−138.

17. Aristarkhov, A., Eytan, E., Moghe, A., Admon, A., Hershko, A., and Ruderman, J. V. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 4294−4299.
18. Townsley, F. M., Aristarkhov, A., Beck, S., Hershko, A., and Ruderman, J. V. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 2362−2367.
19. Yu, H., King, R. W., Peters, J. M., and Kirschner, M. W. (1996) *Curr. Biol. 6*, 455−466.
20. Osaka, F., Seino, H., Seno, T., and Yamo, F. (1997) *Mol. Cell. Biol. 17*, 3388−3397.
21. Jancarik, J., and Kim, S.-H. (1991) *J. Appl. Crystallogr. 24*, 409−411.
22. Otwinowski, J. (1993) in *Data Collection and Processing* (Sawyer, L., Issacs, N., and Bailey, S., Eds.) pp 56−62, Daresbury Laboratory, Warrington, U.K.
23. Pearson, W. R. (1990) *Methods Enzymol. 183*, 63−98.
24. Tong, L. (1993) *J. Appl. Crystallogr. 26*, 748−751.
25. Brünger, A. T. (1992) *X-PLOR. A System for Crystallography and NMR*, Yale University Press, New Haven, CT.
26. Brünger, A. T. (1992) *Nature 355*, 472−474.
27. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
28. Jones, T. A., Zhou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D53*, 240−255.
30. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
31. Laskwoski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
32. Vriend, G. (1990) *J. Mol. Graphics 8*, 52−56.
33. Altieri, A. S., Hinton, D. P., and Byrd, R. A. (1995) *J. Am. Chem. Soc. 117*, 7566−7567.
34. King, R. W., Peters, J. M., Tugendreich, S., Rolfe, M., Hieter, P., and Kirschner, M. W. (1995) *Cell 81*, 279−288.
35. Ptak, C., Prendergast, J. A., Hodgins, R., Kay, C. M., Chau, V., and Ellison, M. J. (1994) *J. Biol. Chem. 269*, 26539−26545.
36. Kolman, C. J., Toth, J., and Gonda, D. K. (1992) *EMBO J. 11*, 3081−3090.
37. Silver, E. T., Gwozd, T. J., Ptak, C., Goebl, M., and Ellison, M. J. (1992) *EMBO J. 11*, 3091−3098.
38. Haldeman, M. T., Xia, G., Kasperek, E. M., and Pickart, C. M. (1997) *Biochemistry 36*, 10526−10537.
39. Sommer, T., and Jentsch, S. (1993) *Nature 365*, 176−179.
40. Kriwacki, R. W., Hengst, L., Tennant, L., Reed, S. I., and Wright, P. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 11504−11509.
41. Braig, K., Adams, P. D., and Brunger, A. T. (1995) *Nat. Struct. Biol. 2*, 1083−1094.
42. Pickart, C. M., and Rose, I. A. (1985) *J. Biol. Chem. 260*, 1573−1581.
43. Haas, A. L., and Bright, P. M. (1988) *J. Biol. Chem. 263*, 13258−13267.
44. Girod, P. A., and Viestra, R. D. (1993) *J. Biol. Chem. 268*, 955−960.
45. Kovalenko, O. V., Plug, A. W., Haaf, T., Gonda, D. K., Ashley, T., Ward, D. C., Radding, C. M., and Golub, E. I. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2958−2963.
46. Jiang, W., and Koltin, Y. (1996) *Mol. Gen. Genet. 251*, 153−160.
47. Gwozd, C. S., Arnason, T. G., Cook, W. J., Chau, V., and Ellison, M. J. (1995) *Biochemistry 34*, 6296−6302.
48. Banerjee, A., Gregori, L., Xu, Y., and Chau, V. (1993) *J. Biol. Chem. 268*, 5668−5675.
49. Janin, J., and Rodier, F. (1995) *Proteins 23*, 580−587.
50. Jones, S., and Thornton, J. M. (1995) *Prog. Biophys. Mol. Biol. 63*, 31−65.
51. Cook, W. J., Jeffrey, L. C., Sullivan, M. L., and Vierstra, R. D. (1992) *J. Biol. Chem. 267*, 15116−15121.
52. Cook, W. J., Jeffrey, L. C., Xu, Y., and Chau, V. (1993) *Biochemistry 32*, 13809−13817.
53. Cook, W. J., Martin, P. D., Edwards, B. F., Yamazaki, R. K., and Chau, V. (1997) *Biochemistry 36*, 1621−1627.
54. Tong, H., Hateboer, G., Perrakis, A., Bernards, R., and Sixma, T. K. (1997) *J. Biol. Chem. 272*, 21381−21387.
55. Saitoh, H., Pu, R., Cavenagh, M., and Dasso, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 3736−3741.
56. Schwarz, S. E., Matuschewski, K., Liakopoulos, D., Scheffner, M., and Jentsch, S. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 560−564.
57. Desterro, J. M., Thomson, J., and Hay, R. T. (1997) *FEBS Lett. 417*, 297−300.
58. Worthylake, D. K., Prakash, S., Prakash, L., and Hill, C. P. (1998) *J. Biol. Chem. 273*, 6271−6276.
59. Sullivan, M. L., and Vierstra, R. D. (1991) *J. Biol. Chem. 266*, 23878−23885.
60. Haas, A. L., and Siepmann, T. J. (1997) *FASEB J. 11*, 1257−1268.
61. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
62. Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.
63. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.

BI9901329