THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 34, Issue of August 22, pp. 21381–21387, 1997
Printed in U.S.A.
© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.

Crystal Structure of Murine/Human Ubc9 Provides Insight into the Variability of the Ubiquitin-conjugating System*

(Received for publication, April 6, 1997, and in revised form, May 27, 1997)

Harry Tong, Guus Hateboer‡, Anastassis Perrakis§, René Bernards, and Titia K. Sixma¶

From the Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands

Murine/human ubiquitin-conjugating enzyme Ubc9 is a functional homolog of *Saccharomyces cerevisiae* Ubc9 that is essential for the viability of yeast cells with a specific role in the G_2-M transition of the cell cycle. The structure of recombinant mammalian Ubc9 has been determined from two crystal forms at 2.0 Å resolution. Like *Arabidopsis thaliana* Ubc1 and *S. cerevisiae* Ubc4, murine/human Ubc9 was crystallized as a monomer, suggesting that previously reported hetero- and homo-interactions among Ubcs may be relatively weak or indirect. Compared with the known crystal structures of Ubc1 and Ubc4, which regulate different cellular processes, Ubc9 has a 5-residue insertion that forms a very exposed tight β-hairpin and a 2-residue insertion that forms a bulge in a loop close to the active site. Mammalian Ubc9 also possesses a distinct electrostatic potential distribution that may provide possible clues to its remarkable ability to interact with other proteins. The 2-residue insertion and other sequence and structural heterogeneity observed at the catalytic site suggest that different Ubcs may utilize catalytic mechanisms of varying efficiency and substrate specificity.

Conjugation of ubiquitin to various eukaryotic cellular proteins regulates their activities by controlling protein concentration through ubiquitin-directed degradation (1–4) or by directly modifying protein function through the attached ubiquitin molecules (5). The formation of ubiquitin-protein conjugates proceeds via a cascade of reactions that involve two, or often three, enzymes: the ubiquitin-activating enzyme E1,[1] the ubiquitin-conjugating enzyme Ubc (E2), and the ubiquitin-ligating enzyme E3 (6–8). The first step is the ATP-assisted formation of a high energy thioester bond between ubiquitin and E1. Ubiquitin is then transferred to a conserved cysteine group of Ubc. In some ubiquitin pathways, Ubc alone, or in cooperation with E3, attaches ubiquitin to the ϵ-amino group of a lysine residue of a substrate via an isopeptide bond. In others, Ubc first passes ubiquitin to a thiol group of E3, and then E3 attaches it to the substrate. Repeated conjugation of ubiquitin to lysine residues of previously bound ubiquitin moieties is required for proteolysis of the substrates by the 26 S proteasome (9). A large body of genetic and biochemical evidence indicates that the Ubcs, together with E3s, are the primary determinants of the specificity of individual ubiquitin pathways (10).

Based on amino acid sequence comparison, Ubcs can be divided broadly into four classes (10). Class I enzymes consist of a relatively conserved catalytic core domain of about 150 residues showing at least 25% sequence identity. Class II and III enzymes have either extra C-terminal or extra N-terminal extensions attached to the core domain, respectively. Class IV enzymes have both C- and N-terminal extensions. Some of these extensions to the core domain confer a certain degree of specificity for enzyme-substrate recognition or provide a localization signal. However, both specificity and localization signals also reside within the core domain itself. Different subsets of Ubcs, comprising either a single member, or multiple members from the same or different classes, are involved in different cellular processes and thereby constitute distinct functional subfamilies (10).

Murine/human Ubc9 (18 kDa) has been cloned in yeast two hybrid assays as a class I ubiquitin-conjugating enzyme that interacts with a large variety of proteins, including the adenovirus E1A oncoprotein (11), the human Rad51 recombinase (12), the human papillomavirus type 16 E1 replication protein (13), the *Saccharomyces cerevisiae* centromere DNA-binding core complex (14), the negative regulatory domain of the Wilms' tumor gene product (WT1) (15), and the Fas antigen (CD95) (16). These interactions have been further confirmed in the cases of adenovirus E1A, WT1, and Fas by glutathione *S*-transferase (GST) pull-down assays, although their physiological relevance remains to be established. The amino acid sequence of Ubc9 is found to be 100% identical between mice and humans (11, 12). It is closely related to *S. cerevisiae* Ubc9 (17) (56% identity) and *Schizosaccharomyces pombe* Hus5 (18) (66% identity). Both yeast Ubc9 enzymes are essential for cell viability with a role in regulating cell cycle progression at the G_2 or early M phase. *S. cerevisiae* Ubc9 has been shown to target the degradation of the M-phase cyclin Clb5, the S-phase cyclin Clb2 (17), as well as the G_1 cyclins Cln1 and Cln2 (19). Mammalian Ubc9 can complement a *S. cerevisiae* Ubc9 temperature sensitive defect but not a similar mutation in the apparently more similar Hus5 of *S. pombe* (11–16). Here we describe the x-ray crystal structure of recombinant murine Ubc9 and compare it with the known crystal structures of *Arabidopsis thaliana* Ubc1 (20) and *S. cerevisiae* Ubc4 (21).

* This work was supported in part by EMBL Hamburg through the HCMP Acess to Large Facilities Grant CHGE-CT93-0040. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors for Ubc9 (1U9A and R1U9ASF for crystal form I; 1U9B and R1U9BSF for crystal form II) have been deposited in the Protein Data Bank, Brookhaven National Laboratory, Upton, NY.

‡ Funded by a grant from the Netherlands Organization for Scientific Research (NWO). Present address: European Institute of Oncology, Department of Experimental Oncology, Via Ripamonti 435, 20141 Milano, Italy

§ Recipient of a European Molecular Biology Organization Fellowship (ALTF-215, 1995).

¶ To whom correspondence should be addressed. Tel.: 31-20-512-1959; Fax: 31-20-512-1954; E-mail: sixma@nki.nl.

[1] The abbreviations used are: E1, ubiquitin-activating enzyme; E2, ubiquitin-conjugating enzyme; E3, ubiquitin-ligating enzyme; GST, glutathione *S*-transferase; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)-propane-1,3-diol; Mes, 4-morpholineethanesulfonic acid; MR, molecular replacement.

Downloaded from www.jbc.org at UNIVERSITEIT UTRECHT on December 5, 2006

TABLE I
Diffraction data and final model statistics

Crystal form	I222	$P2_1$
Data		
Unit cell		
a (Å)	35.39	52.04
b (Å)	93.95	35.18
c (Å)	115.89	58.10
β (°)		111.2
Resolution range (Å)	10–2.0	10–2.0
Completeness (%)	98 (99)[a]	91 (94)[a]
Redundance	3.2	3.0
No. of unique reflections	12854	12295
I/σ (I)	17.0 (3.0)[a]	8.9 (3.8)[a]
R_{merge}[b] (%)	5.8 (41)[a]	12 (42)[a]
Solvent content (%)	52	53
Model		
No. of amino acids	159	160
No. of proteins atoms	1276	1283
No. of water molecules	96	103
R-factor[c] (%)	18.5	16.0
Free R-factor[d] (%)	25.2	25.5
Mean B-factor (Å^2)		
Protein	25.1	21.9
Solvent	39.1	36.0
Overall B-factor (Å^2)	26.1	22.9
Wilson B-factor (Å^2)	29.7	25.1
Root mean square deviations from ideal stereochemistry		
Bond distances (Å)	0.015	0.012
Bond angles (°)	1.44	1.29
Torsion angles (°)	16.7	15.2

[a] The numbers given in parentheses denote the respective values of the highest resolution shell (2.05 − 2.02 Å for crystal form I222; 2.03 − 2.00 Å for $P2_1$).

[b] $R_{merge} = \Sigma|I - \langle I\rangle|\Sigma\langle I\rangle$, where the summation is over all the equivalent measurements including Friedel pairs.

[c] R-factor = $\Sigma||F_o| - |F_c||/\Sigma|F_o|$.

[d] A subset (5%) of the data was set aside exclusively for free R-factor calculation.

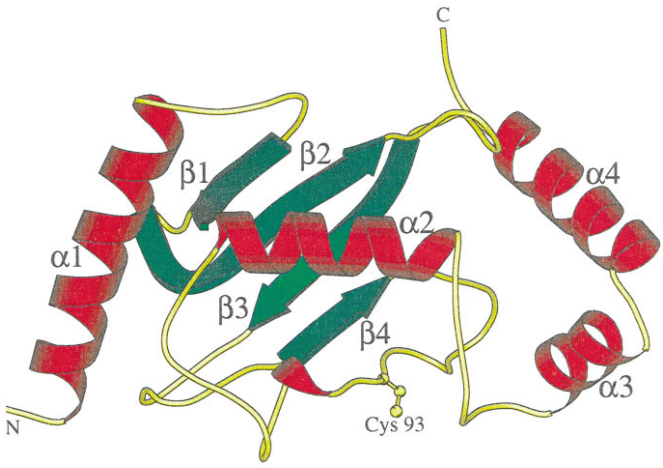

FIG. 1. **Schematic representation of the Ubc9 model in the $P2_1$ crystal form.** Helices are drawn in *red* and strands are drawn in *green*. The side chain atoms of the ubiquitin-accepting cysteine Cys[93] are represented as *yellow ball-and-sticks*. The first α-helix (residues 1–18) and the four β-strands (residues 25–30, 36–46, 57–63, and 74–77) consist of amino acids from the N-terminal half of the molecule, whereas the three other α-helices (residues 109–121, 131–139, and 141–154) consist of residues from the C-terminal half. These secondary structure elements comprise approximately 50% of the structure (17% β-strands and 33% α-helices). This figure was prepared using MOLSCRIPT (55).

EXPERIMENTAL PROCEDURES

Overexpression and Purification—The cloning of the mouse UBC9 gene was as described previously (11). Ubc9 was expressed and purified using a GST fusion system (22). The complete UBC9 gene was subcloned into the pGEX-2T expression vector (Pharmacia Biotech Inc.) and transformed into *Escherichia coli* strain DH5α (Life Technologies Inc.). This expression system allows Ubc9 to be expressed as the C-terminal part of a GST fusion protein with a thrombin cleavage site in the linker region. Upon cleavage of the fusion protein with thrombin, the recombinant Ubc9 acquires eight extra residues at the N terminus: GSPGISLN. Transformed *E. coli* cells were grown to an A_{600} of 0.8-1.2 in LB medium in the presence of 34 μg/liter carbenicillin. Protein expression was then induced by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside at 30 °C. Cells were harvested 4–5 h after induction. Lysis of the cell pellets was carried out by sonication in lysis buffer (100 mM NaCl, 0.5% Nonidet P-40, 5 mM EDTA, 1 mM EGTA, 60 mM 1,4-dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 25 mM Tris-Cl buffer, pH 8.0). The fusion protein was found in the soluble part of the lysate and was first purified over a glutathione-Sepharose 4B column (Pharmacia) by extensive washing with dissociation buffer (23) (3 mM ATP, 1 mM dithiothreitol, 10 mM $MgSO_4$, 150 mM NaCl, 10 mM Tris-Cl buffer, pH 8.0) prewarmed at 37 °C. Ubc9 was then separated from GST by bovine thrombin cleavage on the column at room temperature (22 °C). The cleavage solution was 1.5 units/ml bovine thrombin (Boehringer Mannheim) in cleavage buffer (2.5 mM $CaCl_2$, 150 mM NaCl, 20 mM Tris-Cl, pH 8.5). Upon elution from the glutathione-Sepharose 4B column, Ubc9 was passed through a desalting column to exchange into 50 mM Bis-Tris buffer, pH 6.5. The protein was further purified over a cation exchange column (MonoS from Pharmacia) using a 0 to 1 M NaCl gradient. Ubc9 eluted between 200 and 310 mM NaCl. The protein was finally purified to apparent homogeneity through a gel filtration column (preparation grade Superdex 75 from Pharmacia). The apparent molecular weight of Ubc9 is 25 kDa as estimated by comparing its elution profile with those of protein markers of known molecular masses, which is somewhat larger than the calculated molecular mass of 18 kDa for a monomer, but well short of the size predicted for a dimer. The final yield of Ubc9 was 3 mg for each liter of LB culture, as estimated by measuring the UV absorbance of the purified protein at 280 nm and using the molar extinction coefficient (ϵ_M = 29400 M^{-1} cm^{-1}) calculated with the GCG package (24).

Crystallization and Data Collection—The protein was concentrated to 11 mg/ml in storage buffer (150 mM NaCl, 0.5 mM DTT, 10 mM Hepes buffer, pH 7.5) before crystallization trials. Crystals of two different forms were grown from hanging drops by the technique of vapor diffusion at room temperature. Both types of crystals have the shape of elongated parallelepiped rods but belong to different space groups: I222 (a = 35.4 Å, b = 93.9 Å, c = 115.9 Å) for crystal form I, and $P2_1$ (a = 52.0 Å, b = 35.2 Å, c = 58.1 Å, β = 111.2°) for crystal form II. From crystallization drops with an initial volume of 10 μl (5 μl of protein solution plus 5 μl of precipitant solution), type I crystals grew to a full size of 1.6 mm × 0.1 mm × 0.06 mm within 2 days in the presence of 23% polyethylene glycol monomethyl ether 5000, 9% isopropanol, 0.1 M $(NH_4)_2SO_4$, 0.1 M MES buffer, pH 6.5; type II crystals grew up to a size of 1.8 mm × 0.06 mm × 0.04 mm over the period of one week in the presence of 9% polyethylene glycol 4000, 9% isopropanol, 0.1 M Hepes buffer, pH 7.5. Diffraction data up to 2.2 Å resolution were initially recorded on a Macscience DIP2020 image-plate system (Enraf-Nonius, Delft) at room temperature (22 °C). High resolution data were collected using synchrotron radiation on a MAR image-plate system (MAR-Research, Hamburg) at the EMBL outstation DESY, Hamburg. Both types of crystals were cooled to 8 °C in cold air streams and diffracted beyond 2.0 Å resolution (Table I). All data were indexed with DENZO and scaled with SCALEPACK (25).

Structure Determination and Analysis—Initial phase information for Ubc9 in the I222 crystal form was obtained by molecular replacement (MR) using AMoRe (26) from CCP4 (27). Atomic coordinates of plant Ubc1 (33) and yeast Ubc4 (34) were alternately used as search models against an initial 2.7 Å data set. The rotation and translation function searches performed with both search models yielded one distinct solution. The correlation coefficients of the MR solution were 0.31 for Ubc1, and 0.26 based on Ubc4, whereas the corresponding next highest peaks were 0.22 and 0.20, respectively. After rigid body fitting (AMoRe) of the MR solution using the Ubc1 model, the crystallographic R-factor was 48.2% with data in the range 8–3 Å. The correctness of the MR solution was further confirmed by the identification of a number of structural features unique to Ubc9 in a difference Fourier map. An initial structure solution was also found for the $P2_1$ crystal form by the MR method using the Ubc9 model that was already refined in the I222 space group. The R_{cryst} was 31.9% for this MR solution.

Refinement for the structures in both crystal forms followed similar protocols. From each of the data sets used, 5% was set aside for the R_{free} (28) calculation, which was used to monitor the progress of the refine-

Downloaded from www.jbc.org at UNIVERSITEIT UTRECHT on December 5, 2006

ment along with the conventional R_{cryst} and stereochemical criteria. Several rounds of refinement were carried out by subjecting the model alternately to simulated annealing refinement with X-PLOR (29), and manual adjustment based on $3F_o$-$2F_c$ and F_o-F_c difference electron density maps using O (30). Some electron density maps used for model building were calculated with models generated using ARP (31) in combination with PROLSQ (32). Waters were modeled and checked by ASIR (33) in combination with TNT (34). Final rounds of refinement using a conjugate-direction algorithm and bulk solvent correction in the TNT program resulted in R_{cryst} of 18.5% (R_{free} = 25.2%) for Ubc9 in the I222 crystal form and R_{cryst} of 16.0% (R_{free} = 25.5%) in the $P2_1$ crystal form. Both Ubc9 models include the full 158 residues of the native protein (Fig. 5). Of the 8-residue N-terminal extension introduced as a cloning artifact, 1 (in I222) or 2 ($P2_1$) residues were observed in electron density. Some modeled structural features, including the 5-residue insertion that is unique to mammalian Ubc9 and its yeast homologs, are better defined in the $P2_1$ crystal form than in I222, or vice versa. The I222 model and the $P2_1$ model include, respectively, 96 and 103 water molecules with a few of them extending beyond the first water shell. Analysis with PROCHECK (35) indicated that the final models for both types of structure have good stereochemistry (Table I). General structure analysis was carried out using WHATIF (36).

Sequence Analysis—Multiple sequence alignments of ubiquitin conjugating enzymes were constructed with CLUSTALW (37). Homology-derived structure prediction values of relative entropy of variability of amino acids (38) were calculated using the PHD server (39).

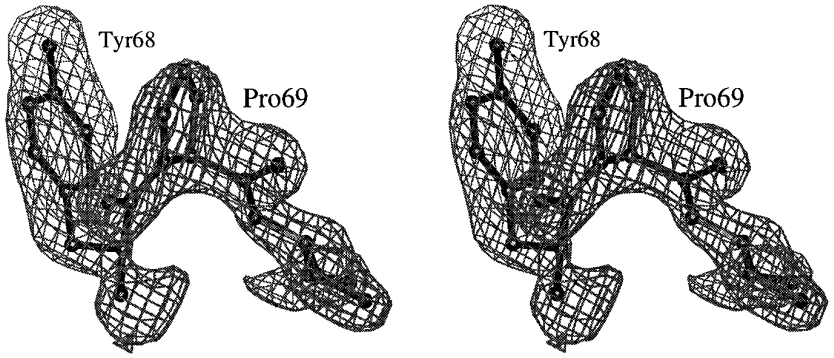

FIG. 2. **Final $2F_o$-F_c map for the region around a conserved *cis*-proline: Pro^{69}.** Mutating Pro^{69} to Ser in *S. cerevisiae* Ubc9 resulted in a temperature sensitive phenotype (17). The contour level is set at 2.0σ.

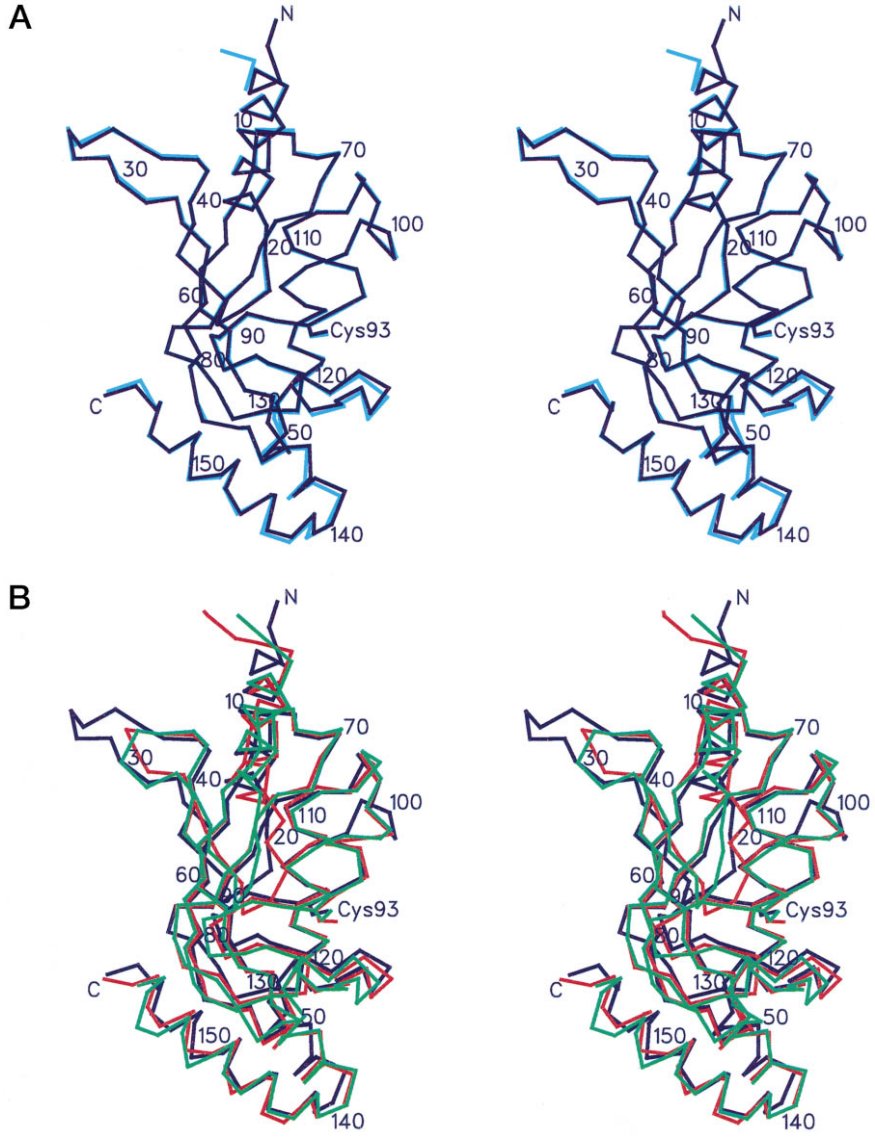

FIG. 3. **Stereoviews of the superimposed C$^{\alpha}$-atoms of the known Ubc crystal structures.** *A*, mammalian Ubc9 from crystal form I (I222, *cyan*) and II ($P2_1$, *blue*). *B*, mammalian Ubc9 ($P2_1$ (*blue*), *Arabidopsis* Ubc1 (*red*), and *S. cerevisiae* Ubc4 (*green*).

Downloaded from www.jbc.org at UNIVERSITEIT UTRECHT on December 5, 2006

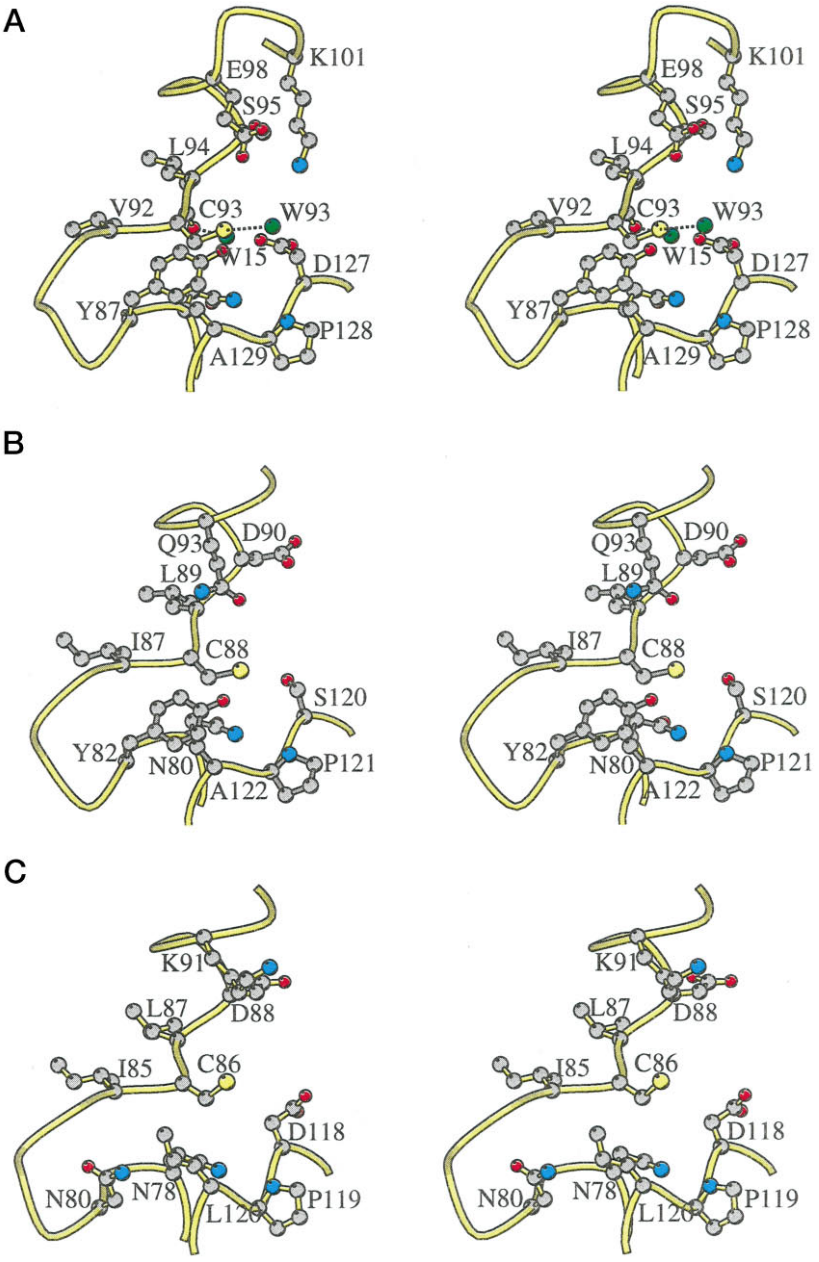

FIG. 4. **Comparison of the active site for mammalian (*A*) Ubc9, *Arabidopsis* Ubc1 (*B*), and S. cerevisiae Ubc4 (*C*).** All three stereoviews in this figure are oriented similarly to those of Fig. 3. Side chains with atoms within 6 Å of the S atom of Cys^{93} are represented by *ball-and-sticks*. Two active site water molecules (*W15* and *W93*) are shown as *green balls* with *dashed lines* indicating the hydrogen bonds between them and Cys^{93}. This diagram was generated using MOLSCRIPT (55).

RESULTS AND DISCUSSION

Overall Structure—The structure of recombinant murine Ubc9 has been determined to 2.0 Å resolution in two crystal forms ($P2_1$ and I222). Ubc9 forms a single domain $\alpha + \beta$ structure that is typical of the Ubc core domain (21). The molecule is asymmetric with overall dimensions of approximately 20 Å × 30 Å × 50 Å. The structure (Fig. 1) contains an antiparallel β-sheet with four strands (β1 to β4, these and all subsequent secondary structure assignments were carried out by the program DSSP, Ref. 40) bound on one side and at both ends by four α-helices (α1 to α4).

The active site residue Cys^{93} is situated close to the middle of a long extended stretch of 31 residues (78–108) found between the fourth β-strand and the second α-helix. This polypeptide segment contains five tight turns and one turn of 3_{10} helix. Part of this long loop (residues 85–102) and another loop between the α2 and the α3 helix (residues 122–130) form a crevice with the active site cysteine in between. The structure contains two *cis*-prolines: Pro^{69} (Fig. 2), and Pro^{79}.

The N terminus is situated at one end of the long axis of the molecule, whereas the C terminus is located opposite to the catalytic site (Fig. 1). The observed portion of the artificial N-terminal extension indicates that it extends away from the core domain, and does not appear to affect the folding of the rest of the protein.

Comparison between the Two Ubc9 Models—The two Ubc9 models from crystal form I (I222) and crystal form II ($P2_1$) are essentially identical except for some minor differences (Fig. 3*A*). The root mean square difference for all 158 C^{α} atoms corresponding to the native protein is 0.44 Å after superposition. Protein atoms displaying the largest differences include those of the first three N-terminal residues, an exposed 5-residue insertion (^{32}PDGTM^{36}) which is barely visible in the I222 crystal form, and a large segment from residue Asn^{121} to Lys^{146}, including the active crevice forming loop between the α2 and the α3 helix (residues 122–130). Compared with the rest of the protein, this segment rotates by 3°, resulting in a slight widening of the active crevice in the I222 model compared with the $P2_1$ model.

Genetic and biochemical data indicate that some Ubcs inter-

Downloaded from www.jbc.org at UNIVERSITEIT UTRECHT on December 5, 2006

act with one another in homo- and heterocomplexes (41). In some cases these interactions may play an important functional role (42, 43). We and others have also observed the self-association of Ubc9 in yeast two-hybrid assays (12, 14). However, our gel filtration experiments did not show Ubc9 as dimers, even in high concentration (10 mg/ml) at close to physiological ionic strength (150 mM NaCl, 10 mM Tris-Cl, pH 7.5). Crystal packing analysis of the two crystal forms of Ubc9 revealed only a small set of residues involved in the formation of intermolecular contacts that are conserved in both crystal forms but they are not more extensive than can be expected from normal crystal packing (~150 Å^2). They certainly do not represent a general mechanism for dimer formation among Ubcs because they have no equivalents in the crystal packing of plant Ubc1 or in yeast Ubc4, which were both crystallized as monomers as well (21, 44). It appears that the widely observed interactions among Ubcs may be either relatively weak or indirect.

Comparison with the Structures of Ubc1 and Ubc4—Arabidopsis Ubc1 is highly similar to *S. cerevisiae* Ubc2 (Rad6) involved in DNA repair (45), whereas *S. cerevisiae* Ubc4 is involved in the degradation of abnormal and short lived proteins, especially in stress conditions (46). Murine/human Ubc9 shares 39% sequence identity with *Arabidopsis* Ubc1 and 35% identity with *S. cerevisiae* Ubc4. Despite their involvement in distinct functional pathways and the limited sequence similarity for the three Ubcs, the tertiary structure of Ubc9 is similar to those of Ubc1 and Ubc4 (Fig. 3*B*). Murine/human Ubc9 has 6 more residues than *Arabidopsis* Ubc1 and 10 more than *S. cerevisiae* Ubc4. These differences in amino acid sequence are primarily accommodated by two insertions in the structure of Ubc9. The first insertion occurs at residues 32–36 and these 5 residues form most of a very exposed β-hairpin that connects strand β1 and β2. The second insertion occurs at residues 100–101, and forms a bulge in a loop (residues 94–102) close to Cys^{93}. Overall, the root mean square differences are 2.4 Å for 150 equivalent C$^\alpha$ atoms between Ubc9 and Ubc1, and 2.0 Å for 148 equivalent C$^\alpha$ atoms between Ubc9 and Ubc4.

The Catalytic Site—Despite the overall similarity in the folding of the three Ubcs, there are considerable differences in the detailed features of the molecule, especially at the active site. There are 10 residues in Ubc9 within 6 Å of the sulfhydryl group of the ubiquitin-accepting cysteine, Cys^{93}. Among these residues, Asn^{85}, Tyr^{87}, Glu^{98}, Lys^{101}, and Asp^{127} are the most likely to mediate in the catalytic action as their side chains are orientated toward the ubiquitin-accepting sulfhydryl group (Fig. 4). Only Asn^{85}, Leu^{94}, and Pro^{128} are conserved compared with both Ubc1 and Ubc4. In all three Ubc crystal structures, the carbonyl O atom of Asn^{85} is hydrogen bonded to the backbone N atom of Cys^{93}, whereas the side chain of Asn^{85} makes hydrogen bonds to the main chain of residues 124 and 127, keeping this relatively mobile loop between the α2 and α3 helix in position. Therefore it is not obvious that Asn^{85} can participate in the catalytic mechanism unless this loop moves away upon binding of Ubc to the E1 ubiquitin adduct. Otherwise the catalytic site displays considerable sequence and structural heterogeneity among the three Ubcs.

A major structural difference is created by the 2-residue insertion in Ubc9: Asp^{100} and Lys^{101}. These 2 residues form a small protruding loop near the active site cysteine. Ubc3 (cdc34), another ubiquitin-conjugating enzyme with known cell cycle function, has a 12-residue insertion at the equivalent position. It appears that such inserted loops can provide additional binding sites for substrates without blocking access to the active site cysteine. Charge-to-alanine scanning mutagenesis indicated that charged residues of the 12-residue insertion

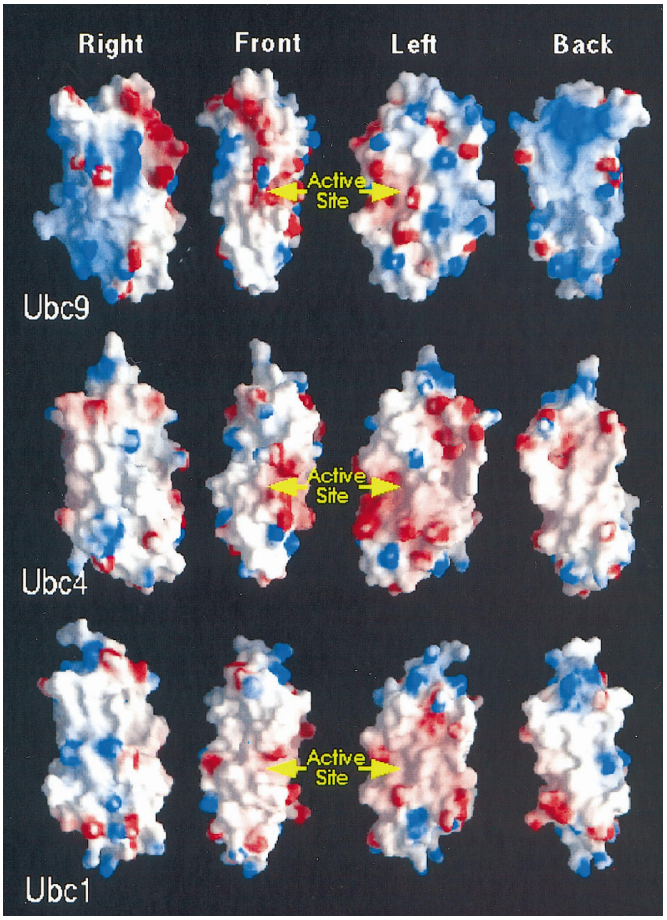

FIG. 5. **Electrostatic potentials mapped onto the surface of the mammalian Ubc9, *Arabidopsis* Ubc1, and *S. cerevisiae* Ubc4 structures.** Among the four different orientations, the "right" view corresponds to the active site cysteine on the right-hand side of the figure. This view is similar to that of Fig. 3*A*. The rest are generated by successive rotations of 90° around the vertical axis. The color spectrum from *red* to *blue* corresponds to changes from negative to positive potential (−10 to 10 K_BT, where K_B is the Boltzmann constant). This diagram was produced using GRASP (50).

in Ubc3 are important to its *in vivo* function without affecting its enzymatic competence with respect to unfacilitated (E3-independent) ubiquitination (47). The 2-residue insertion in Ubc9 could similarly contribute to Ubc9-specific functions.

There are two ordered water molecules found in the vicinity of Cys^{93}: Wat^{15} and Wat^{93}. It is possible that these active-site water molecules play some role in the catalytic mechanism for ubiquitin-conjugating enzymes. Wat^{15} is attached to the carbonyl oxygen of Cys^{93}, and Wat^{93} is directly attached to the sulfhydryl group of the active site cysteine. Whether Ubc1 and Ubc4 also possess similarly positioned water molecules is not clear, since few or no water molecules were modeled due to the lower resolution limits of these two structures.

Structural differences at the catalytic site may provide some clues to why different subfamilies of Ubc's are involved in different functional pathways. In addition to substrate specificity and requirement for different E3s, recent discoveries of ubiquitin-like proteins (48, 49) suggest another possible reason for the variability of the catalytic machinery of the ubiquitin-conjugating system, that some ubiquitin-conjugating enzymes may be involved in conjugating the ubiquitin-like proteins rather than ubiquitin itself to substrates.

Surface Electrostatic Potentials—Large variations are found in the surface electrostatic potentials of the three Ubc structures (Fig. 5), with the exception of a negative patch surround-

Downloaded from www.jbc.org at UNIVERSITEIT UTRECHT on December 5, 2006

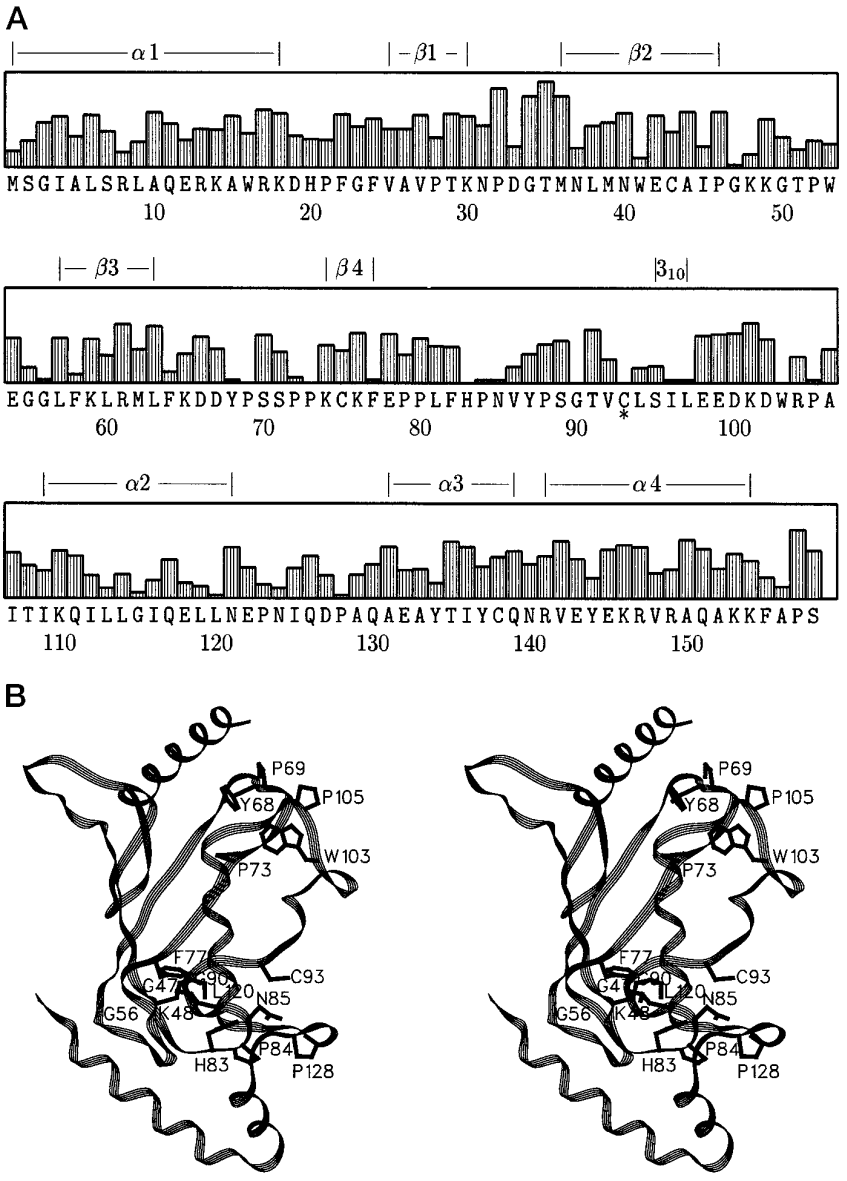

FIG. 6. **Sequence consensus analysis for the catalytic domain of 66 ubiquitin-conjugating enzymes from SWISS-PROT (March, 1997).** *A*, histogram representation of homology-derived structure prediction-derived relative entropies of variability (*vertical axis*, relative scale from 0 to 100) for all amino acids occurring in mammalian Ubc9 (*horizontal axis*). A small relative entropy indicates a high degree of sequence homology. For residues with deletions in other family members (1–2, 30–36, 99–101, and 158), relative entropies have been assigned a value of 100. The secondary structure elements indicated are for the Ubc9 structure. *B*, stereoview of the most conserved residues (with relative entropies of variability <15%) mapped onto the three-dimensional structure of Ubc9. This figure was drawn using SETOR (56).

ing the active site. While this negative patch may be important in orienting common interaction partners such as E1 or ubiquitin, the varied electrostatic features probably reflect the need to recognize different E3s and substrates. Overall, Ubc9 possesses a considerably stronger electrostatic dipole (541 Debye, calculated by GRASP (50) at pH 7.0) than either Ubc1 (310 Debye) or Ubc4 (149 Debye). Positive patches are scattered on the "back face" of Ubc9 (Fig. 5), including the N-terminal region composed of a segment of basic residues separated by nonpolar residues (^{13}RKAWRK^{18}) that is highly conserved among Ubc9s (51). Notably, this conserved negative patch is close to the highly exposed β-hairpin, another distinguishing structural feature of Ubc9. The spatial proximity of these two "specifically" conserved structural features suggests that this region could be important for Ubc9 function. The lack of these two structural features in *Arabidopsis* Ubc9 can be explained because it belongs to another subfamily, yeast Ubc4 (52).

Sequence analysis of the known interaction partners of mammalian Ubc9 indicates that they mostly have a strong overall negative charge, or at least possess a large region of about 100 residues with a particularly low isoelectric point (pI). In particular, the region on adenovirus E1A responsible for interacting with Ubc9 has been mapped to a polypeptide segment of 70 residues with a predicted pI of 4.6 (11). These observations suggest an important role for electrostatic attractions in the liaison between mammalian Ubc9 and its multiple interaction partners. However, these interactions must also have a relatively strong hydrophobic component, as evidenced by the fact that the association between Ubc9 and adenovirus E1A is sensitive to a single Leu→Ile mutation in the transformation-relevant conserved region 2 of E1A (11).

Conserved Residues—We have extended the comparison of the amino acid sequences to the currently available sequences of the catalytic domain of the entire Ubc family (66 annotated sequences from SWISS-PROT (53), March 1997). We have used the homology-derived structure prediction program to analyze variability for amino acids with equivalents in Ubc9. A histogram representation of these values shows that residues comprising the loops are better conserved than those forming the regular secondary structure elements in general (Fig. 6*A*). This probably reflects the importance of the "unstructured" stretch (from residue 78 to 108 in Ubc9) of Ubc structures in forming the active site, as well as the important structural role of some other loops. Apart from the Cys^{93} residue essential to the catalytic action, there are 15 residues which are particularly well conserved (Gly^{47}, Lys^{48}, Gly^{56}, Tyr^{68}, Pro^{69}, Pro^{73},

Downloaded from www.jbc.org at UNIVERSITEIT UTRECHT on December 5, 2006

Phe[77], His[83], Pro[84], Asn[85], Gly[90], Trp[103], Pro[105], Leu[120], and Pro[128]) with a relative entropy of variability less than 15% of the value for the most variable residue (Fig. 6*A*). These residues are clustered around both ends of the long irregular loop containing the active-site (within 18 Å of Cys[93], Fig. 6*B*). However, with the exception of Asn[85], whose possible involvement in the catalytic reaction has been discussed, and Pro[128], they are outside the immediate vicinity of the ubiquitin-accepting cysteine (>6 Å to the sulfhydryl group of Cys[93]). Notably, 11 of the 15 highly conserved residues are nonpolar residues. It is unlikely that any of these is directly involved in the catalytic action but most are positioned to maintain the special conformation of the active site. The important structural role for such conserved residues is demonstrated in the case of *cis*-Pro[69] (Fig. 2). A Pro[69] → Ser mutation has been shown to cause a temperature sensitive defect in *S. cerevisiae* Ubc9 (17). Loss of this proline makes Ubc9 sensitive to proteolysis by a ubiquitin- and proteasome-dependent pathway at the restrictive temperature (54), indicating that this mutation destabilizes the protein fold.

Mapping of the amino acid variability values onto the surface of the three-dimensional structure of Ubc9 indicates that one side of the active site cysteine (the "front" side of Fig. 1) displays a higher degree of conservation than the other, as was shown in the comparison between the Ubc1 and the Ubc4 structure (21). Our analysis provides further support to the hypothesis by Cook *et al.* (21) that this better conserved side may contain possible binding sites for the E1-ubiquitin adduct, although the conserved regions do not appear to be as contiguous as when only two structures were compared.

CONCLUSIONS

The overall similarity of the high resolution mammalian Ubc9 structure to those of plant Ubc1 and yeast Ubc4 suggests that the folding of the catalytic domain of the family of Ubc enzymes is conserved in all eukaryotes. Mapping of amino acid variability onto the surface of the three-dimensional structure of Ubc9 shows a better conserved surface on one side of the ubiquitin-accepting cysteine that may serve as possible recognition surface regions on Ubcs for their common physiological partners, E1 and ubiquitin.

There is considerable structural heterogeneity observed in the catalytic crevice among the Ubcs with known crystal structures. Sequence consensus analysis for the entire Ubc family also shows a lack of conserved residues close to the active site cysteine. A variable catalytic machinery might account for some of the differences among Ubcs in their efficiency and in their requirement for E3s to ubiquitinate different sets of target proteins.

A number of features unique to mammalian Ubc9, such as a protruding surface loop and a strong overall electrostatic dipole, may have a role in conferring the distinctive property to Ubc9 for interacting with an exceptionally large variety of proteins. Understanding such interactions may also provide insight into the *modus operandi* for some of the biologically important interaction partners.

Acknowledgments—We thank Alex Teplyakov and Zbigniew Dauter for assistance during synchrotron data collection at beamlines BW7B and X11 of the EMBL/Hamburg outstation. We also thank Pim van Dijk for general assistance.

REFERENCES

1. Hochstrasser, M. (1995) *Curr. Opin. Cell Biol.* **7,** 215–223
2. Jentsch, S. & Schlenker, S. (1995) *Cell* **82,** 881–884
3. Jentsch, S. (1992) *Trends Cell Biol.* **2,** 98–103
4. Ciechanover, A. (1994) *Cell* **79,** 13–21
5. Hochstrasser, M. (1996) *Cell* **84,** 813–815
6. Ciechanover, A., Heller, H., Katz-Etzion, R. & Hershko, A. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 761–765
7. Hershko, A., Heller, H., Elias, S. & Ciechanover, A. (1983) *J. Biol. Chem.* **258,** 8206–8214
8. Pickart, C. M. & Rose, I. A. (1985) *J. Biol. Chem.* **260,** 1573–1581
9. Goldberg, A. L. (1995) *Science* **268,** 522–523
10. Jentsch, S. (1992) *Annu. Rev. Genet.* **26,** 179–207
11. Hateboer, G., Hijmans, E. M., Nooij, J. B. D., Schlenker, S., Jentsch, S. & Bernards, R. (1996) *J. Biol. Chem.* **271,** 25906–25911
12. Kovalenko, O. V., Plug, A. W., Haaf, T., Gonda, D. K., Ashley, T., Ward, D. C., Radding, C. M. & Golub, E. I. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 2958–2963
13. Yasugi, T. & Howley, P. M. (1996) *Nucleic Acids Res.* **24,** 2005–2010
14. Jiang, W. & Koltin, Y. (1996) *Mol. Gen. Genet.* **251,** 153–160
15. Wang, Z.-Y., Qiu Q.-Q., Seufert, W., Taguchi, T., Testa, J. R., Whitmore, S. A., Callen, D. F., Welsh, D., Shenk, T. & Deuel T. F. (1996) *J. Biol. Chem.* **271,** 24811–24816
16. Wright, D. A., Futcher, B., Ghosh, P. & Geha, R. S. (1996) *J. Biol. Chem.* **271,** 31037–31043
17. Seufert, W., Futcher, B. & Jentsch, S. (1995) *Nature* **373,** 78–81
18. Al-Khodairy, F., Enoch, T., Hagan, I. M. & Carr, A. M. (1995) *J. Cell Sci.* **108,** 475–486
19. Blondel, M. & Mann, C. (1996) *Nature* **384,** 279–282
20. Cook, W. J., Jeffrey, L. C., Sullivan, M. L. & Vierstra, R. D. (1992) *J. Biol. Chem.* **267,** 15116–15121
21. Cook, W. J., Jeffrey, L. C., Xu, Y. P. & Chau, V. (1993) *Biochemistry* **32,** 13809–13817
22. Smith, D. B. & Johnson, K. S. (1988) *Gene* (*Amst.*) **67,** 31–40
23. Yu-Sherman, M. & Goldberg, A. L. (1992) *EMBO J.* **11,** 71–77
24. GCG version 7 (1991) Genetics Computer Group, Madison, WI
25. Otwinowski, Z. (1993) in *Proceedings of the CCP 4 Study Weekend: Data Collection and Processing.* (Sawyer, L., Issacs, N. & Bailey, S., ed), pp. 56–62, SERC Daresbury Laboratory, England
26. Navaza, J. (1994) *Acta Crystallogr. Sec. A* **50,** 157–163
27. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. Sec. D* **50,** 760–763
28. Brünger, A. T. (1992) *Nature* **355,** 472–474
29. Brünger, A. T. (1992) *X-PLOR: Version 3.1. A System for Crystallography and NMR*, Yale University Press, New Haven, CT
30. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. Sec. A* **47,** 110–119
31. Lamzin, V. S. & Wilson, K. S. (1993) *Acta Crystallogr. Sec. D* **49,** 129–147
32. Hendrickson, W. (1995) *Methods Enzymol.* **115,** 252–270
33. Tong, H., Berghuis, A. M., Chen, J., Luo Y. G., Guss, J. M., Freeman, H. C. & Brayer, G. D. (1994) *J. Appl. Crystallogr.* **27,** 421–426
34. Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987) *Acta Crystallogr. Sec. A* **43,** 489–501
35. Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
36. Vriend, G. (1990) *J. Mol. Graph.* **8,** 52–56
37. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
38. Sander C. & Schneider R. (1991) *Proteins* **9,** 56–68
39. Rost B. (1996) *Methods Enzymol.* **266,** 525–539
40. Kabsch W. & Sander, C. (1983) *Biopolymers* **22,** 2577–2637
41. Gwozd, C. S., Arnason, T. G., Cook, W. J., Chau, V. & Ellison, M. J. (1995) *Biochemistry* **34,** 6296–6302
42. Silver, E. T., Gwozd, T. J., Ptak, C., Goebl, M. & Ellison, M. J. (1992) *EMBO J.* **11,** 3091–3098
43. Chen, P., Johnson, P., Sommer, T., Jentsch, S. & Hochstrasser, M. (1993) *Cell* **74,** 357–369
44. Cook, W. J., Jeffrey, L. C., Sullivan, M. L. & Vierstra, R. D. (1992) *J. Mol. Biol.* **223,** 1183–1186
45. Sullivan, M. L. & Vierstra, R. D. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 9861–9865
46. Seufert, W. & Jentsch, S. (1990) *EMBO J.* **9,** 543–550
47. Pitluk, Z. W., McDonough, M., Sangan, P. & Gonda, D. K. (1995) *Mol. Cell. Biol.* **15,** 1210–1219
48. Shen, Z. Y., Pardington-Purtymun, P. E., Comeaux, J. C., Moyzis, R. K. & Chen, D. (1996) *Genomics* **36,** 271–279
49. Mahajan, R., Delphin, C., Guan, T. L., Gerace, L. & Melchior, F. (1997) *Cell* **88,** 97–107
50. Nicholls, A., Sharp, K. A. & Honig, B. (1991) *Proteins* **11,** 281–296
51. Watanabe, T. K., Fujiwara, T., Kawai, A., Shimizu, F., Takami, S., Hirano, H., Okuno, S., Ozaki, K., Takeda, S., Shimada, Y., Nagata, M., Takaichi, A., Takahashi, E., Nakamura, Y. & Shin, S. (1996) *Cytogenet. Cell Genet.* **72,** 86–89
52. Girod, P. A., Carpenter, T. B., van Nockers, S., Sullivan, M. L. & Vierstra, R. D. (1993) *Plant J.* **3,** 545–552
53. Bairoch, A. & Boekmann, B. (1994) *Nucleic Acids Res.* **22,** 3578–3580
54. Betting, J. & Seufert, W. (1996) *J. Biol. Chem.* **271,** 25790–25796
55. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
56. Evans, S. V. (1993) *J. Mol. Graph.* **11,** 134–138

Downloaded from www.jbc.org at UNIVERSITEIT UTRECHT on December 5, 2006