Skp1 Dimerization Conceals its F-box Protein Binding Site

Hyun W. Kim[1,‡], Alexander Eletsky[2,‡], Karen J. Gonzalez[3], Hanke van der Wel[1], Eva-Maria Strauch[3], James H. Prestegard[1,2], and Christopher M. West[1,2,4,]*

[1]Dept. of Biochemistry and Molecular Biology, [2]Complex Carbohydrate Research Center, [3]Dept. of Pharmaceutical and Biomedical Sciences, [4]Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA 30602 USA

[‡] These authors contributed equally to this work.

ABSTRACT

Skp1 is an adapter that links F-box proteins to cullin-1 in the Skp1/cullin-1/F-box (SCF) protein family of E3 ubiquitin ligases that targets specific proteins for polyubiquitination and subsequent protein degradation. Skp1 from the amoebozoan *Dictyostelium* forms a stable homodimer *in vitro* with a K_d of 2.5 µM as determined by sedimentation velocity studies, yet is monomeric in crystal complexes with F-box proteins. To investigate the molecular basis for the difference, the solution NMR structure of a doubly-truncated Skp1 homodimer was determined. The dimer interface was mapped by intermolecular NOEs which identified ten contact amino acids. The resulting structure of the Skp1ΔΔ dimer has semi-parallel 2-fold symmetry with an interface that buries ~700 $Å^2$ of predominantly hydrophobic surface that partially overlaps with subsite-1 of the F-box interaction interface, explaining why only the Skp1 monomer binds FBPs. To confirm the model, Rosetta was used to identify amino acid substitutions that might disrupt that dimer, and the F97E substitution was chosen to potentially minimize interference with F-box interactions. Skp1ΔF97E behaved as a monomer at concentrations up to 500 µM and actively bound a model F-box protein, mammalian Fbs1. Therefore, the dimeric state is not required for Skp1 to carry out a basic biochemical function, and access to the monomeric state is expected to facilitate structural studies of the mechanism of regulation of protist Skp1 by posttranslational prolyl hydroxylation and glycosylation.

INTRODUCTION

The Skp1/cullin-1/F-box (SCF) protein family of E3 ubiquitin ligases is an important mediator of protein turnover in yeast/fungi, higher plants, and animals, owing to the role of polyubiquitination in serving as a signal for recognition and degradation in the 26S-proteasome. Evidence supports the importance of the SCF complex in the protist kingdom as well (West and Blader, 2015), where a novel posttranslational modification has been discovered in Skp1s from groups as diverse as amoebozoa [*Dictyostelium discoideum*], apicomplexans [*Toxoplasma gondii*], and oomycetes [*Pythium ultimum*] (van der Wel et al., 2019). Protist Skp1 is subject, in the presence of sufficient O_2-tension and a-ketoglutarate concentration, to hydroxylation of a Pro-residue that lies on the backside of subsite-2 of the F-box binding domain of Skp1. Once Skp1 is hydroxylated, the Hyp residue is recognized and glycosylated by a series of glycosyltransferases, resulting in the assembly of a stereotyped pentasaccharide. Mutational studies show that both hydroxylation and full glycosylation are required for optimal O_2-sensing in *Dictyostelium* and *Toxoplasma* (West and Blader, 2015; Mandalasi et al., 2019). Biophysical and computational studies have generated a model by which the relatively organized structure of the pentasaccharide glycan organizes the surrounding intrinsically disordered region of *Dictyostelium* Skp1 in such a way as to be more conducive to recognizing the F-box domain of FBPs (Sheikh et al., 2017), and recent studies indicate that this model is also relevant to Skp1 from *Toxoplasma* (Mandalasi et al., 2019). Confirmatory studies are hampered by the dimeric state of glycosylated Skp1 (GGFGGn-Skp1) at concentrations used for biophysical studies, because its relatively large size, 324 amino acids.

Recombinant guinea pig Skp1, whose sequence is identical across mammals, was previously reported

to dimerize with a K_d of 1.1 μM (Henzl et al., 1998). This value is substantially below the estimated concentration of Skp1 (a.k.a. OCP2) in the inner ear tissues where Skp1/OCP-2 was initially characterized, suggesting that dimerization might influence Skp1 activity in cells. Studies of *Dictyostelium* Skp1, where small angle X-ray scattering, gel filtration, and NMR studies confirmed its dimeric status at higher concentrations, indicate that glycosylation modestly inhibits dimerization (Sheikh et al., 2014). We sought to investigate the significance of Skp1 dimerization by mapping its dimer interface. The structure of Skp1 from mammals, yeast and higher plants is known when it is bound to F-box proteins (Schulman et al., 2000; Zimmerman et al., 2010), and despite the broad sequence variations in F-box domains, the sequence and structure of Skp1 in these complexes is highly though not absolutely conserved. However, free Skp1 has defied structural characterization, potentially because of intrinsically disordered regions (including its C-terminal region that contributes to F-box domain recognition) that preclude the formation of high-quality crystals for X-ray crystallography. At the same time, the Skp1 homodimer is too large for solution NMR studies without resorting ^{2}H-isotope labeling. We have found that *Dictyostelium* Skp1A remains a dimer in the absence of an internal disordered region that was originally removed to allow crystallization with F-box proteins (Schulman et al., 2000). In addition, truncation of the intrinsically disordered region that comprises the C-terminal subsite-2 of the F-box binding region is also not required for dimerization. The doubly truncated Skp1ΔΔ dimer variant (2 × 118 aa) was sufficiently small to pursue high-resolution solution NMR structure determination using uniform ^{15}N- and ^{13}C-isotope labeling. We found that the homodimer interface overlaps with F-box binding and confirmed the finding by Rosetta-guided mutagenesis, and discuss the implications for Skp1 function and future studies on its regulation.

EXPERIMENTAL PROCEDURES

Expression Plasmids. *Dictyostelium* Skp1ΔΔ was generated from pET19b-Skp1A (Xu et al., 2018) by site-directed mutagenesis, in which primers designed to bridge the deleted sequence as described in Figure S1C were used in a PCR reaction with Q5® High-Fidelity DNA Polymerase (New England Biolabs) to amplify the desired modified vector. After treatment with DpnI to destroy the original vector template, the sample was cloned into *E. coli* strain BL21-Gold(DE3).

For improved recovery and purification, the *Dictyostelium* Skp1A coding sequence was codon optimized for expression in E. *coli*, and appended with an N-terminal His_6-tag which, when excised by treatment with TEV protease, yielded the native sequence with an N-terminal extension of SMSL-, compared to the N-terminal SL- that occurs natively after removal of the start Met (West et al., 1997). The cDNA (Figure S1A) was synthesized and provided in pUC57 by GenScript, excised using NcoI and BamHI, and ligated into the restriction sites for NcoI and BamHI of pET19b, yielding pET19b-His_6DdSkp1A-optim. A second cDNA in which 12 internal amino acids (SPQGDDKKDEKR) were replaced with GGSG (Figure S1B) was synthesized and similarly ligated into pET19B yielding pET19b-His_6DdSkp1AΔLoop-optim. This plasmid was modified to generate an F97E point mutation by site-directed mutagenesis, in which the indicated primers were used in a PCR reaction to amplify the modified vector as above.

Expression and Purification of Skp1 constructs. Skp1 proteins without an affinity tag were purified from *E. coli* to near homogeneity under non-denaturing conditions (DEAE, phenyl, Q and S200 Superdex columns) as described previously (Sheikh et al., 2014). Sample was assessed by SDS-PAGE and Coomassie blue staining to be >90%.

E. coli cells expressing His_6Skp1Δ or His_6Skp1ΔF97E were incubated at 37 ºC in 2 × 1 L of Terrific Broth medium in the presence of 100 μg/ml ampicillin. Once the cell growth reached an OD_{600} of 0.6, protein expression was induced with 125 μM isopropyl 1-thio-β-D-galactopyranoside (IPTG) at 20ºC. After 12-16 h, cells were collected by centrifugation at 5000 × *g* for 10 min and resuspended in 50 mM Na^+/K^+ phosphate (pH 7.8), 300 mM NaCl, 5 μg/ml aprotinin, 5 μg/ml leupeptin at 4ºC. Cells were lysed using a probe sonicator (model 500, Thermo Fisher Scientific) for a total sonication time of 5 min. The

lysate was centrifuged at 25,000 × *g* for 45 min at 4ºC, and the supernatant was immediately applied to a 1.5 ml column of Co^{2+} Talon resin (Clontech) pre-equilibrated at 4ºC in the buffer described above. The column was washed successively with buffers supplemented with either 1 M NaCl, 10% glycerol, or 5 mM Imidazole. Protein was eluted with buffer containing 300 mM imidazole, and dialyzed against 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 1 mM EDTA, 2 mM β-mercaptoethanol and 1 mM TCEP. The sample was incubated overnight at 20°C with His_6TEV protease to cleave the His_6-tag from Skp1, and the sample was re-applied to the Talon resin. The flow-through was concentrated to 1.5 ml using a spin concentrator (Amicon) with a 3 kDa molecular weight cut-off. The concentrated sample was further purified over a Superdex 200 Hi-load 16/60 gel filtration column (GE Healthcare) pre-equilibrated with 20 mM K^+/phosphate (pH 7.4), 50 mM KCl, and 0.5 mM TCEP. The sample was estimated to be >95% pure by SDS-PAGE and staining with Coomassie blue.

Stable isotope labeled Skp1ΔΔ and His_6Skp1ΔF97E proteins for NMR studies were prepared via expression in *E. coli* using isotope enriched minimal media as previously described (Xu et al., 2018) and purified using methods described above. His_6Skp1ΔF97E was uniformly enriched with ^{15}N isotope only, and Skp1ΔΔ was labeled with ^{15}N and ^{13}C. The final Skp1ΔΔ NMR sample (105 µl in a 3-mm Shigemi tube) contained a 1:1 mixture of ^{15}N,^{13}C-Skp1ΔΔ and natural abundance Skp1ΔΔ at ~1.0 mM concentration in 20 mM MES-NaOH (pH 6.0), 50 mM NaCl, 5 mM dithiothreitol, 0.05% NaN_3. ^{15}N-labeled Skp1ΔF97E NMR samples (300 µl in 5-mm Shigemi tubes) were prepared at concentrations of 100 µM and 500 µM in the same buffer. All NMR samples contained 10% D_2O for spectrometer lock.

Analytical Ultracentrifugation. Protein was quantified based on molar absorptivity calculated from the protein sequence using ProtParam (Gastiger et al., 2005). Samples were loaded into 12-mm double-sector Epon centerpieces equipped with quartz windows and equilibrated for 2 h at 20°C in an An60 Ti rotor. Sedimentation velocity data were collected using an Optima XLA analytical ultracentrifuge (Beckman Coulter) at 50,000 rpm at 20°C. Data were recorded with absorbance optics at 280 nm, 230 nm or 215 nm in radial step sizes of 0.003 cm. SEDNTERP (Laue et al., 1992) was used to model the partial specific volume as well as the density and viscosity of the buffer. SEDFIT (Schuck, 2000) was used to analyze sedimentation data. All data were modeled as continuous *c(s)* distributions and were fit using baseline, meniscus, frictional coefficient, and systematic time-invariant and radial-invariant noise. Predicted sedimentation coefficient (s) values for Skp1 monomer and dimer were calculated using HYDROPRO (Ortega et al., 2011) with a homology model generated on the ROBETTA server (Kim & Chivian et al., 2004). Data fit and *c(s)* plots were generated using GUSSI (Brautigam, 2015). Weight-averaged S values (S_w) at each concentration were determined by integrating *c(s)* distributions. Constructed Sw isotherms were fitted with A+A<-->AA reversible self-association model using SEDPHAT (Schuck, 2003).

NMR Spectroscopy and Structure Determination of Skp1ΔΔ. NMR spectra for Skp1ΔΔ were acquired at 35°C using a Bruker AVANCE NEO 800 MHz spectrometer equipped with a 5-mm cryogenic TCI $^1H\{^{13}C,^{15}N\}$ probe, and an Agilent VNMRS 600 MHz spectrometer equipped with a 3-mm cryogenic $^1H\{^{13}C,^{15}N\}$ probe. NMR spectra for 100 µM and 500 µM Skp1ΔF97E samples were acquired using a Bruker AVANCE NEO 900 MHz spectrometer equipped with a 5-mm cryogenic TXO $^{13}C,^{15}N\{^1H\}$ probe, and the 600 MHz spectrometer equipped with a 5-mm cryogenic $^1H\{^{13}C,^{15}N\}$ probe. The acquired NMR spectra are summarized in Table S1. NOE mixing times were 70 ms for $^{13}C/^{15}N$-edited $[^1H,^1H]$-NOESY and 120 ms for $^{13}C/^{15}N$-filtered $^{13}C/^{15}N$-edited $[^1H,^1H]$-NOESY experiments. Fourier transform was performed with TopSpin (Bruker BioSpin) for Bruker NMR data, and NMRPipe (Delaglio et al., 1995) for Varian NMR data. ^1H chemical shifts were referenced relative to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), and ^{13}C and ^{15}N chemical shifts were referenced indirectly via gyromagnetic ratios. 2D and 3D NMR spectra were analyzed using CARA (Keller, 2004).

Relaxation delays were 0.1, 0.2, 0.3, 0.4, 0.7, 1.0, 1.5, and 2.0 seconds in 1D ^{15}N T_1 experiment, and 10, 30, 50, 70, 90, 110, 130, 150 and 170 ms in 1D ^{15}N T_1 experiment acquired for 500 uM Skp1ΔF97E sample. 1D T_1/T_2 relaxation data were processed and analyzed with VnmrJ v4.2 (Agilent Inc). Macro "tc"

(wiki.nesg.org) was used to integrate the regions between ^{1}H chemical shifts of 8.8 and 9.9 ppm, determine T_1 and T_2 relaxation times via exponential fitting and calculate global rotational correlation time τ_c.

Sequence-specific backbone and side-chain resonance assignments for Skp1ΔΔ were derived using CARA based on existing resonance assignments of full-length Skp1 (Xu et al., 2018). Structure calculation of the Skp1ΔΔ homodimer was performed using CYANA (Güntert et al., 2015) based on ^{1}H-^{1}H upper distance constraints derived from ^{13}C/^{15}N-edited [^{1}H,^{1}H]-NOESY, as well as backbone φ and ψ and side-chain χ_1 dihedral angle restraints from TALOS-N (Shen et al., 2013). Automated NOESY peak assignment was performed initially with CYANA, with manually assigned intermolecular ^{1}H-^{1}H upper distance constraints applied after cycle 2 of simulated annealing. These intermolecular ^{1}H-^{1}H upper distance constraints were derived from ten selected strong peaks in a ^{13}C/^{15}N-filtered ^{13}C/^{15}N-edited [^{1}H,^{1}H]-NOESY spectrum, that could be unambiguously assigned and could not be explained by intramolecular contacts. After several rounds of interactive refinement of NOE peak assignments and calibration of distance constraints, the final structure calculation was performed with CYANA. Out of 100 calculated conformers, 20 conformers with the lowest target function values were selected for subsequent refinement in explicit water bath using the program CNS (Brunger et al., 1998) with upper distance constraints relaxed by 5%.

Dimer Interface Studies. Computational analyses were performed using the conformer with lowest CYANA target function of the initial NMR structure of Skp1ΔΔ. To prevent inaccurate predictions due to small clashes in the structure, the protein was prepared using the standard Rosetta optimization protocol, "FastRelax" (Tyka et al., 2011; Das and Baker, 2008). Briefly, five cycles of rotamer packing and minimization were carried out, ramping up the repulsive weight in the scoring function within each cycle. After three rounds of symmetrical FastRelax with atom-atom pair distance constraints, the quality of the generated models was validated with the Molprobity web service (Chen et al., 2010). The lowest scoring structure based on the Rosetta energy score and the Molprobity score was selected for mutational analysis.

To identify mutations that could disrupt the Skp1 dimerization, an all amino acids-scanning mutagenesis was carried out across the homodimer interface *in silico*. The "flex ddG" protocol implemented in RosettaScripts (Barlow et al., 2018; Fleishman et al., 2011) was used to model and predict the effect of the mutations on the binding free energy of the complex. All parameters in the protocol were set up according to the default values described previously (Barlow et al., 2018). Overall, the flex ddG method takes advantage of the Rosetta backrub approach (Smith and Kortemme, 2008) to sample side-chain and backbone conformational changes around the mutated position. Once the backrub ensembles are generated, the structures are optimized by sidechain repacking and torsion minimization. The interface ΔΔG score corresponds to the average difference in binding free energy between the mutant structure and the wild-type complex. Stabilizing mutations are defined as those with interface ΔΔG scores < -1.0 Rosetta energy units (REU), while destabilizing mutations are assigned to interface ΔΔG scores > 1.0 (Barlow et al., 2018).

To determine which dimer-destabilizing substitutions do not perturb the stability of the individual monomers, the change in the total free energy of the monomer was estimated for all mutations. This analysis was performed with the current state-of-the-art Rosetta ΔΔG protocol, "cartesian_ddg" (Park et al., 2016). Prior to the simulations, the refined wild-type monomer was relaxed in Cartesian space, constraining backbone and sidechain coordinates. The model with the lowest Rosetta score was then used as input for the cartesian_ddg protocol. Within the cartesian_ddg application, the protein was relaxed again in the Cartesian space, allowing movement of only the backbone and sidechains around the mutated position (Park et al., 2016). All parameters in the method were configured as previously described (Park et al., 2016). The total ΔΔG score was finally considered as the difference in the total ΔΔG between the mutant and the wild-type monomer, multiplied by an energy scaling factor of 1.0/2.94. As above, stabilizing mutations correspond to total ΔΔG scores < -1.0, and destabilizing mutations refer to total

ΔΔG scores > 1.0.

All Rosetta commands for this report were run with the same Rosetta static executable (RosettaCommons/main.git2019-03-07, version 4ab48a76160c888257155619edb9817845bd8a67). The protocols previously described can be found at Supplementary information- Scripts.

Analytical Gel Filtration. Skp1Δ and Skp1ΔF97E with or without Fbs1 at a limiting concentration were subjected to Superdex 200 PC 3.2/30 gel filtration analysis using Pharmacia SMARTSystem HPLC as previously described (Sheikh et al., 2014).

RESULTS

Characterization of the Skp1 Dimer by Analytical Ultracentrifugation. Analytical centrifugation was used to characterize Skp1 (native sequence) in solution at various concentrations (Figure 1A). Sedimentation velocity experiments revealed peaks at 1.8 S and 2.7 S, which is consistent with predicted S-value for monomer and dimer at 1.9 S and 2.8 S. Skp1 monomer and dimer peaks are independent of loading concentrations suggesting that Skp1 dimerization undergoes a slow equilibrium on the time-scale of sedimentation. The smaller S-values indicate that the homology model used for predicting S-values were less compact. In the homology model, C-terminal region of Skp1 is modeled as a helix, as it is in complex with F-box proteins. However, the C-terminal region of free Skp1 is, as suggested by previous NMR studies, extended and unstructured having a predominantly disordered characteristic (Xu et al., 2018). A Skp1 dimer binding isotherm was constructed using weighted s-values across protein concentrations covering almost three orders of magnitude (0.5 – 45 μM), which yielded a dissociation constant for the Skp1 dimer of 2.5 μM (Figure 1B).

Characterization of Skp1ΔΔ by NMR and solution NMR structure. To reduce spectral complexity and avoid using perdeuterated samples, we utilized a truncated Skp1 variant, Skp1ΔΔ, which lacks the mainly disordered C-terminal F-box binding domain and an internal disordered loop that is frequently removed for Skp1/FBP crystallization (Figure 2A). We demonstrated that Skp1ΔΔ (118 x 2 amino acids) still forms a stable homodimer based on sedimentation velocity studies, which yielded an S-value of 1.92 compared to the 1.90 value for a dimer predicted by Hydropro (Figure S2). Also, a 2D [^{15}N,^{1}H] HSQC spectrum of Skp1ΔΔ correlated well with 2D [^{15}N,^{1}H] TROSY of full-length Skp1 (data not shown), indicating that truncations in Skp1ΔΔ do not perturb the overall structure. Using a suite of standard NMR experiments (wiki.nesg.org, Table S1) we obtained complete sequence-specific assignments of backbone and side-chain ^{1}H, ^{15}N and ^{13}C resonances.

Structure calculations were based on NOE peaks from 3D ^{15}N- and ^{13}C-edited [^{1}H,^{1}H] NOESY spectra. Intermolecular NOE peaks were identified in a separate ^{13}C/^{15}N-filtered and ^{13}C-edited [^{1}H,^{1}H] NOESY spectrum recorded with a sample of mixed U-^{15}N,^{13}C-labeled and natural abundance Skp1. An example of the NOE data is given in Figure S3. Not all intra-chain cross-peaks are filtered from the strip on the right (^{13}C/^{15}N-filtered) because the isotope incorporation (~85%) was not as high as possible. However, comparison to a strip from a more conventional set of ^{13}C-edited data (strip on the left) shows that certain peaks are of comparable intensity (inter-chain) and certain peaks are drastically reduced in the filtered set (intra-chain). Ten intermolecular NOE peaks, which could be unambiguously assigned and could not arise from intramolecular contacts, were converted to distance constraints and used to bootstrap the structure calculation with the program CYANA. Using PISA (Krissenel, 2015), ten residues involved in the packing of the interface were identified from calculated structures (Figure 2D). Cα traces of the 20 best models are shown in Figure 2B. The resulting structure of the Skp1ΔΔ dimer has a semi-parallel 2-fold symmetry. The dimer interface is organized as a four-helix bundle, with symmetrical packing contributions from residues of helices 5 & 6 of each chain, with a total surface area that buries ~700 Å^2 of predominantly hydrophobic surface (Figure 2C). A single dimer subunit was superimposed with that of human Skp1 from a crystal structure in complex with an F-box protein (βTRCP, PDB ID: 6M90), revealing a Cα atom RMSD of 1.4 Å (Figure S5). This suggests that the core structure of Skp1 in solution

is similarly folded to the functional state in a heterocomplex with an F-box protein.

The Skp1 homodimer was compared to a Skp1/FBP heterodimer complex to assess physical overlap (Figures 2E, F). An F-box domain from a crystal structure (PDB ID: 5V4B) was modeled over Skp1ΔΔ which revealed a ~600 Å overlap of interfaces in the region of helices 5 and 6. This explains why only the Skp1 monomer is found in complexes with FBPs.

Conservation of the Dimer Interface. L96 to I123 is highly conserved throughout phylogeny. Each of the dimer contact residues except K117 is almost perfectly conserved from stramenopiles to humans (Figure S7), and region is immediately surrounded by Gly or Pro residues and length variations (indels), indicating that this region represents a functional unit under selective pressure to remain intact. This is consistent with the finding that human Skp1 also dimerizes in this concentration range (Henzl et al., 2008).

Computationally-Guided Selection of a Skp1 Monomer Mutant F97E. To test the model, we searched for point mutations to destabilize the dimer. To identify residues critical for Skp1 homodimer stability, the Skp1 protein-protein interface was studied by computational scanning mutagenesis. An initial exploration conducted by alanine-scanning led to the recognition of the residues Phe97, Leu101, and Ile123, as potential destabilizing positions, as their substitution weakened the binding free energy of the complex (interface ΔΔG score > 2) (Figure 3A). To further investigate the contribution of different amino acids to this stability disruption, all 20 amino acids were mutated at each interface position. In concordance with the alanine-scanning results, the positions 97, 101, and 123 represented energetically susceptible positions, as different residue replacements impaired the binding free energy of the dimer (Figure 3C). In particular, 85% of the variations at position 97 promoted interface ΔΔG scores greater than 1.5. Overall, the mutations F97G, F97D and F97E showed the highest dimer-destabilizing effect, with interface ΔΔG scores of 5.41, 5.23 and 4.87, respectively (Figure 3C).

To determine whether the dimer-destabilizing mutations would affect monomer stability, the total monomer ΔΔG was estimated for each substitution. Remarkably, most of the mutations at positions 97, 101, and 123 presented total ΔΔG scores between 1 and -1 (Figure S6). Specifically, F97E, which could destabilize dimerization, might have a neutral effect on the monomer stability with its ΔΔG score at 0.15.

Skp1ΔF97E is a Stable Monomer in Solution. Based on the above analysis, and to allow for flexibility of the carboxyl group to avoid clashes when binding FBPs, the conserved Phe97 was changed to Glu in Skp1Δ. Skp1ΔF97E eluted as a monomer based on gel filtration and was predominantly a monomer by AUC at 100 μM (Figure 4A), which supports the interface model. A preliminary 2D [^{1}H,^{15}N]-HSQC spectrum recorded for a ^{15}N-labeled Skp1ΔF97E sample exhibited peak dispersion consistent with a well-folded protein (Figure 4B), and the peak pattern was comparable to that of wild-type Skp1 considering that there are a number of differences including an N-terminal SM- extension, replacement of 12 amino acids with 4 different amino acids in the internal loop, and truncation after amino acid 125 (not shown). Based on average ^{15}N T_1 and T_2 relaxation times measured for all ^{1}H amide resonances between 8.8 and 9.9 ppm in a 500 μM sample of Skp1ΔF97E, the effective rotational correlation time (τ_c) was estimated at 9 ns. This number is consistent with Skp1ΔF97E remaining monomeric even at high concentrations typical for solution NMR.

Skp1ΔF97E is Binding Competent with a Model FBP. Phe97 is conserved as a Phe or Tyr in known Skp1s. Analysis of Skp1 in crystal structures of complexes with 3 FBPs (Tir1, βTRCP1, Fbs1) shows that, compared to the homodimer, Phe (or Tyr) resides in a different rotamer state with solvent exposure (not shown). This suggests that the F97E replacement will, though removing a peripheral hydrophobic contact, not directly disrupt the FBP interaction. To test the functionality of Skp1ΔF97E, we used a Superdex200 gel filtration column to examine the elution profile of Skp1 in the absence and presence of Fbs1. As previously described (Sheikh et al., 2014) and replicated in Figure 4C, a mixture of Fbs1 and Skp1 eluted prior to the elution positions of either protein alone. As shown in Figure 4D,

Skp1ΔF97E exhibited similar behavior. In addition, Skp1ΔF97E clearly eluted later than native Skp1, consistent with its monomeric character as described by AUC and its rotational correlation time.

DISCUSSION

Our findings confirm that dimerization is a highly conserved property of Skp1, based on similar dissociation constant values from organisms as phylogenetically distant as *Dictyostelium* and humans. Their calculated K_d values range from 1.1 to 2.5 μM under *in vitro* conditions, though actual affinities in the cell may vary. These values are similar to the predicted total monomer Skp1 concentration in a mammalian cell line, ~2 μM (Reitsma et al., 2018). The significance of dimerization is suggested by the exceptionally high degree of conservation of the contact residues (labeled D in Fig. S7), and its ability to mask the hydrophobic character of interface.

The Skp1 dimer interface occupies ~700 $Å^2$ of predominantly hydrophobic surface, a substantial area that overlaps subsite-1 of its FBP binding site (Figure 2E). The corresponding region 99-130 in human Skp1 (aa 93-124 in *Dictyostelium* Skp1, Figure S7) exhibited a paucity of NMR resonance assignments in a recent study of human Skp1 (Kachariya et al., 2016), indicating potential effects of dimerization in human Skp1. Our dimerization model for Skp1was supported by the predicted effect of an amino acid substitution within the interface, F97E, to destabilize the interaction (Figure 4A) without unfolding the protein (Figure 4B). Indeed, monomeric Skp1 maintained its ability to bind a model FBP, Fbs1 (Figure 4C, D), and was also competent to be hydroxylated and glycosylated (unpublished data). A further contribution of the region beyond the dimer interface, which previous studies showed to possess intrinsic disorder in the free dimer (Xu et al., 2018) but becomes ordered into a stereotypical triple-helical bundle when complexed with any of several F-box proteins, seems unlikely.

Based on the new structure, the dimer interface also contributes to subsite-1 of the F-box binding region of Skp1, which explains why Skp1 is a monomer in complexes with FBPs (Zimmerman et al., 2010). Interestingly, 8 of the 13 alleles of Skp1 known to affect its function in budding and fission yeast have point mutations located on this region (Bai et al., 1996; Connelly and Heiter, 1996; Beltrao et al., 2012; Leber et al., 2018; Lehmann et al., 2004). However, the available evidence indicates that the K_d for binding to Fbs1/OCP1 is at least two orders of magnitude smaller (Tan et al., 2008), and this, together with evidence that the Skp1/FBP interaction is generally of very high affinity and indications that Skp1 is not in great excess of FBPs in cells (Reitsma et al., 2018), does not support the idea that Skp1 dimerization is regulatory in cells. Our ability to selectively perturb dimerization relative to Fbs1 binding using the F97E mutation might allow an investigation of this question *in vivo*; however, it is unclear whether interaction with other FBPs, whose sequences are distinctive, will be similarly minimally affected.

Chemical shift index analysis of assigned residues in human Skp1 (Kachariya et al., 2016) and *Dictyostelium* Skp1 (Xu et al., 2018) indicated that the overall secondary structure elements of free Skp1 in dimer state were similar to one another and to human Skp1 in complexes with FBPs, except for the C-terminal F-box subsite-2 region which was predominantly disordered in free Skp1's. The current study demonstrates, for *Dictyostelium* Skp1 up through the dimer interface (through residue 125), that the monomeric state assumes essentially the same structure as for human Skp1 bound to βTRCP with an RMSD for the Cα atoms of 1.4Å (Figure S5). Thus interactions with proteins including Cul1 (Zimmerman et al., 2010) and Sgt1 (Willhoft et al., 2017), whose crystallographically defined binding interfaces lie N-terminal to the dimer interface, might be unaffected by the dimer status of Skp1.

The two C-termini of the Skp1ΔΔ dimer lie close enough to one another in the semi-parallel arrangement of the monomers that the missing C-terminal regions might influence each other's organization in the native protein. The availability of a stable monomeric form of Skp1 will allow a direct NMR analysis of full length Skp1 and the consequences of its glycosylation, which is postulated to influence the organization of F-box subsite-2 that was deleted in the form analyzed here.

ASSOCIATED CONTENT

Supporting Information

The Supplement PDF contains Tables S1-S2, Figures S1-S7, and additional methods.

AUTHOR INFORMATION

Corresponding Author

* E-mail: westcm@uga.edu

Author Contributions

HWK and AE conducted the experiments. KJG performed calculations. AE, HWK, EMS, JHP and CMW advised on design and interpretation. HWK, AE and CMW wrote the initial manuscript draft, which was edited by all authors and finalized by CMW. All authors approved the final version. ‡These authors contributed equally.

Funding Sources

This research was supported by NIH R01-GM037539, NIH R01-AI140245, and a collaboration facilitated by the Resource for Integrated Glycotechnology, NIH P41-GM103390.

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

The authors thank John Glushka for NMR technical support.

ABBREVIATIONS

AUC, analytical ultracentrifugation; FBP, F-box protein; HNOE, heteronuclear NOE; HSQC, heteronuclear single quantum coherence spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; SCF, Skp1/cullin-1/F-box protein family of E3 ubiquitin ligases; Skp1Δ, Skp1Δinternal loop; Skp1ΔΔ, Skp1Δinternal loop/ΔC-terminus; TEV, tobacco etch virus

REFERENCES

Bai C, Sen P, Hofmann K, Ma L, Goebl M, Harper JW, Elledge SJ (1996) SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **86**:263-274.

Barlow KA, O Conchuir S, Thompson S, Suresh P, Lucas JE, Heinonen M, Kortemme T (2018) Flex ddG: Rosetta ensemble-based estimation of changes in protein-protein binding affinity upon mutation. *J Phys Chem B* **122**:5389-5399.

Beltrao P, Albanèse V, Kenner LR, Swaney DL, Burlingame A, Villén J, Lim WA, Fraser JS, Frydman J, Krogan NJ (2012) Systematic functional prioritization of protein posttranslational modifications. *Cell* **150**:413-425.

Brace EJ, Parkinson LP, Fuller RS (2006) Skp1p Regulates Soi3p/Rav1p Association with Endosomal Membranes but Is Not Required for Vacuolar ATPase Assembly. *Euk Cell* **5**:2104–2113.

Brautigam CA (2015) Calculations and publication-quality illustrations for analytical ultracentrifugation data. *Meth Enzymol* **562**:109-133.

Brunger AT, Adams PD, Clore GM, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges N,

Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system (CNS), A new software suite for macromolecular structure determination. *Acta Cryst* **54**:905-921.

Brunger AT (2007) Version 1.2 of the crystallography and NMR system. *Nature Protocols* **2**:2728-2733.

Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sec D Biol Crystallogr* **66**:12-21.

Connelly C, Hieter P (1996) Budding yeast SKP1 encodes an evolutionarily conserved kinetochore protein required for cell cycle progression. Cell 86:275-285.

Das R, Baker D. (2008) Macromolecular modeling with Rosetta. *Annu. Rev. Biochem.* **77**(1): 363-382.

Delaglio F, Grzesiek S, Vuistere GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**:277-293.

Fleishman SJ, Leaver-Fay A, Corn JE, Strauch EM, Khare SD, Koga N, Ashworth J, Murphy P, Richter F, Lemmon G, Meiler Jens, Baker D (2011) Rosettascripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One* **6**: e20161.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. pp. 571-607.

Güntert P, Buchner, L (2015) Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol NMR* **62**:453-471.

Henzl MT, Thalmann I, Thalmann R (1998) OCP2 exists as a dimer in the organ of Corti. *Hear Res* **126**:37-46.

Kachariya NN, Dantu SC, Kumar A (2016) Backbone and side chain assignments of human cell cycle regulatory protein S-phase kinase-associated protein 1. *Biomol NMR Assign* **10**:351-355.

Keller, R. (2004) The Computer aided resonance assignment tutorial, CANTINA Verlag, Goldau.

Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acid Res.* **32**(Web Server issue):W526-531.

Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**:774-797.

Krissinel E (2015) Stock-based detection of protein oligomeric states in jsPISA. *Nucleic Acids Res* **43**:W314-319.

Laue TM, Shah BD, Ridgeway TM, Pelletier SL (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In: Harding SE, Rowe AJ, Horton J, editors. Analytical Ultracentrifugation in Biochemistry and Polymer Science. Cambridge: The Royal Society of Chemistry. pp. 90–125.

Leber V, Nans A, Singleton MR (2018) Structural basis for assembly of the CBF3 kinetochore complex.

EMBO J **37**:269-281.

Lehmann A, Katayama S, Harrison C, Dhut S, Kitamura K, McDonald N, Toda T (2004) Molecular interactions of fission yeast Skp1 and its role in the DNA damage checkpoint. *Genes Cells* **9**:367-382.

Mandalasi M, Kim HW, Rahman K, Zhao P, Sheikh MO, Thieker D, van der Wel H, Gas-Pascual E, Daniel N, Ichikawa TH, Glushka JN, Wells L, Wood ZA, West CM (2019) A glycogenin homolog controls *Toxoplasma gondii* growth via glycosylation of an E3 ubiquitin ligase. *bioRxiv*.

Ortega A, D. Amorós, J. Garc1a de la Torre (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys J* **101**:892-898.

Park H, Bradley P, Greisen P, Liu Y, Mulligan VK, Kim DE, Baker D, DiMaio F (2016) Simultaneous optimization of biomolecular energy functions on features from small molecules and macromolecules. *J Chem. Theory Comput* **12**: 6201-6212.

Schuck P (2000) Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical J* **78**:1606-1619.

Schuck P (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* **320**:104-124.

Schulman BA, Carrano AC, Jeffrey PD, Bowen Z, Kinnucan ERE, Finnin MS, Elledge SJ, Harper JW, Pagano M, Pavietich NP (2000) Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* **408**:381–386

Sheikh MO, Schafer CM, Powell JT, Rodgers KK, Mooers BHM, West CM (2014) Glycosylation of Skp1 affects its conformation and promotes binding to a model F-box protein. *Biochemistry* **53**:1657-1669.

Sheikh MO, Thieker D, Chalmers G, Schafer CM, Ishihara M, Azadi P, Woods RJ, Glushka JN, Bendiak B, Prestegard JH, West CM (2017) O_2 sensing–associated glycosylation exposes the F-box–combining site of the *Dictyostelium* Skp1 subunit in E3 ubiquitin ligases. *J Biol Chem* **292**:18897-18915.

Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **56**:227-241.

Smith CA, Kortemme T (2008) Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. *J Mol Biol* **380**:742-756.

Tan A, Tanner JJ, Henzl MT (2008) Energetics of OCP1-OCP2 complex formation. *Biophys Chem* **134**:64-71.

Tyka MD, Keedy DA, Andre I, Dimaio F, Song Y, Richardson DC, Richardson JS, Baker D (2011) Alternate states of proteins revealed by detailed energy landscape mapping. *J Mol Biol* **405**:607-618.

van der Wel H, Gas-Pascual E, West CM (2019) Skp1 isoforms are differentially modified by a dual function prolyl 4-hydroxylase/N-acetylglucosaminyltransferase in a plant pathogen. *Glycobiology* Jul 8. pii: cwz049.

West CM, Kozarov E, Teng-umnuay P (1997) The cytosolic glycoprotein FP21 of *Dictyostelium*

discoideum is encoded by two genes resulting in a polymorphism at a single amino acid position. *Gene* **200**:1-10.

West CM, Blader IJ (2015) Oxygen sensing by protozoans: How they catch their breath. *Curr Opin Microbiol* **26**:41–47.

West CM, Kim HW (2019) Nucleocytoplasmic O-glycosylation in protists. *Curr Opin Struct Biol* **56**:204-212.

Willems AR, Schwab M, Tyers M (2004) A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim Biophys Acta* **1695**:133–170.

Willhoft O, Kerr R, Patel D, Zhang W, Al-Jassar C, Daviter T, Millson SH, Thalassinos K, Vaughan CK (2017) The crystal structure of the Sgt1-Skp1 complex: the link between Hsp90 and both SCF E3 ubiquitin ligases and kinetochores. *Sci Rep* **7**:41626.

Xu X, Eletsky A, Sheikh MO, Prestegard JH & West CM (2018) Glycosylation promotes the random coil to helix transition in a region of a protist Skp1 associated with F-box binding. *Biochemistry* **57**:51-515.

Zimmerman ES, Schulman BA, Zheng N (2010) Structural assembly of cullin-RING ubiquitin ligase complexes. *Curr Opin Struc Biol* **20**:714–721.

FIGURE LEGENDS

Figure 1. Sedimentation velocity analysis of *Dd*Skp1. (A) *c(s)* distribution reveals concentration dependence of dimerization. The concentration range is depicted by a rainbow spectrum with the lowest concentration in red and the highest in purple. (B) An isotherm was constructed with weighted s-values (S_w); the fitted model indicates a K_d of 2.5 μM. The color of each data point corresponds to the respective *c(s)* distribution in panel A.

Figure 2. Structure of the Skp1 dimer. (A) Domain diagrams of the constructs examined. Note that the versions derived from His_6Skp1 have a SerMet-extension beyond the native Ser- resulting from Met removal. See Figure S1 for details. (B) Superimposition of Cα-traces of 20 calculated conformers of Skp1ΔΔ. (C) Ribbon representation of the lowest energy Skp1ΔΔ conformer. Dimer subunits are colored in green and magenta. A 2-fold axis of rotational symmetry lies vertically between the subunits. (D) Ribbon representation of a single Skp1ΔΔ with the side chains of selected residues involved in close (<5 Å) intermolecular contacts in the dimer structure are shown in a stick representation with the corresponding short distances in green. (E) Surface representation of the Skp1ΔΔ dimer is shown with the rear subunit colored in green and red and the front in transparent gray. The red shading represents homodimer contact region. (F) Surface representation of a hypothetical Skp1ΔΔ/F-box heterodimer model, generated by structural alignment of a single Skp1ΔΔ subunit over Skp1 in human Skp1/FBXW7 complex (PDB ID 5V4B). Skp1ΔΔ is colored green with regions in contact with F-box in red. The transparent gray molecule in front represents FBXW7 residues 2263-2355.

Figure 3. Computational scanning mutagenesis of the Skp1 homodimer interface. (A) Alanine-scanning mutagenesis. On the y-axis are displayed the changes in binding free energy upon replacement with alanine, and on the x-axis are shown the interface residue positions. (B) Skp1 protein- protein interface. The residues at the interface are colored according to the changes in the binding free energy of the dimer upon alanine mutation. On stick representation are shown the residues with the highest energy change (Phe97 and Ile123). See panel C for color code explanation. (C) Heatmap of the changes in binding free energy upon all amino acid substitutions. On the y-axis are indicated the amino acid replacements, and on the x-axis are shown the interface positions being mutated. The colors represent the changes in the binding free energy of the dimer (interface ΔΔG score). Values greater than one (warmer colors) indicate destabilizing mutations, and values less than one (colder colors) imply stabilizing mutation (Barlow et al., 2018).

Figure 4. Skp1ΔF97E is a stable and functional monomer in solution. (A) *c(s)* distributions of 100 μM Skp1Δ or 100 μM Skp1ΔF97E are shown in cyan and black, respectively. (B) $^1H/^{15}N$-HSQC of 100 μM Skp1ΔF97E at 900 MHz, 35°C, and 4 h. The 500 μM spectrum (not shown) was indistinguishable. (C, D) Skp1ΔF97E binds the model F-box protein Fbs1. His_6Fbs1 (1.5 μM) and an estimated 2.25 μM Skp1Δ (C) or Skp1ΔLoopF97E (D) were analyzed on a Superdex 200 gel filtration column. Elution was monitored by A_{280}, which favors detection of Fbs1 relative to Skp1 because of its higher extinction coefficient.

Figure 1

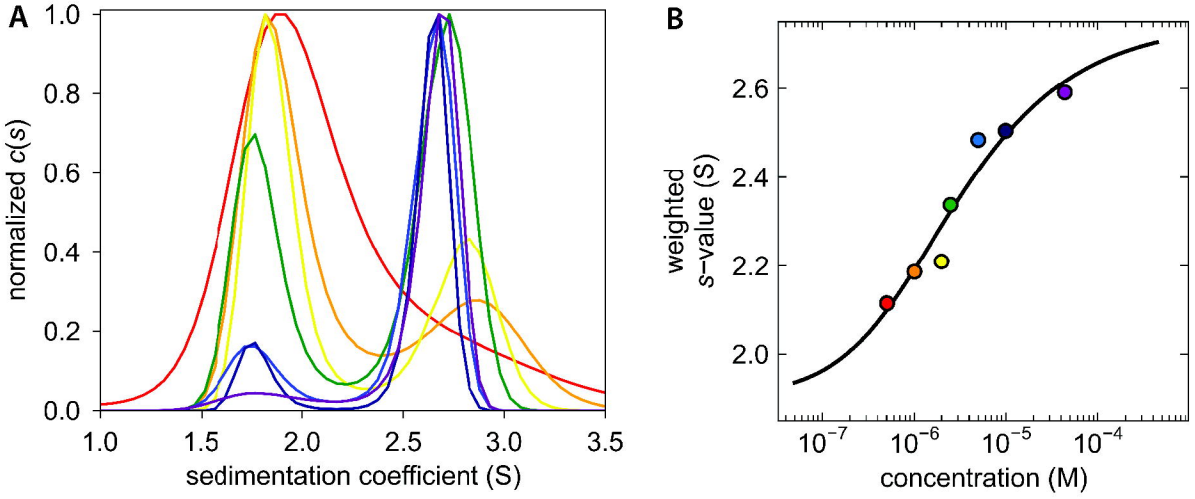

A
1.0
0.8
0.6
0.4
0.2
0.0
normalized c(s)
1.0
1.5
2.0
2.5
3.0
3.5
sedimentation coefficient (S)
B
2.6
2.4
2.2
2.0
weighted s-value (S)
10−7
10−6
10−5
10−4
concentration (M)

Figure 2

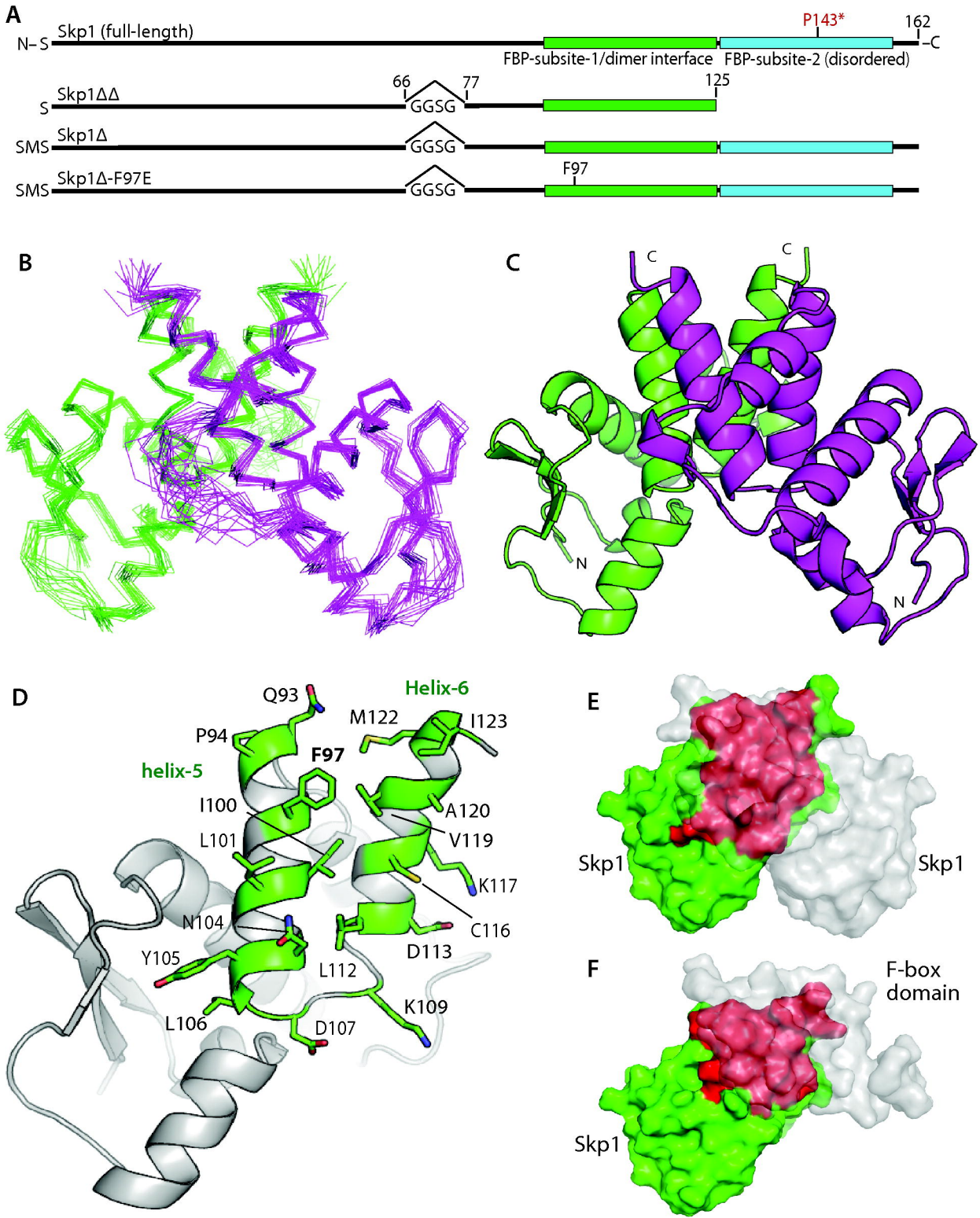

A
N– S
Skp1 (full-length)
P143*
162
–C
FBP-subsite-1/dimer interface
FBP-subsite-2 (disordered)
S
Skp1ΔΔ
66
77
125
GGSG
SMS
Skp1Δ
GGSG
SMS
Skp1Δ-F97E
F97
GGSG
B
C
C
C
N
N
D
Q93
Helix-6
M122
I123
P94
F97
helix-5
I100
A120
L101
V119
K117
N104
C116
D113
Y105
L112
K109
L106
D107
E
Skp1
Skp1
F
F-box
domain
Skp1

Figure 3

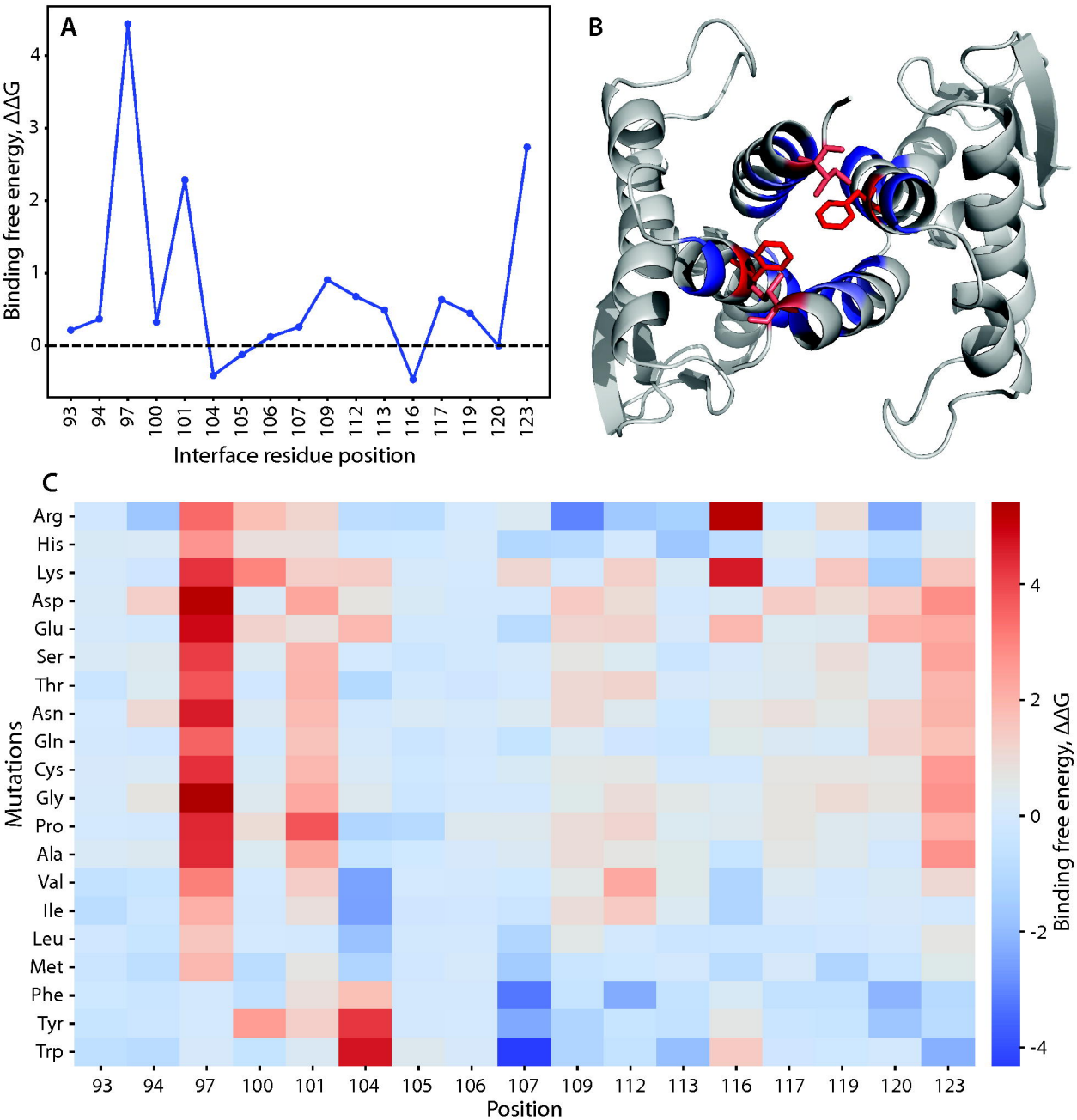
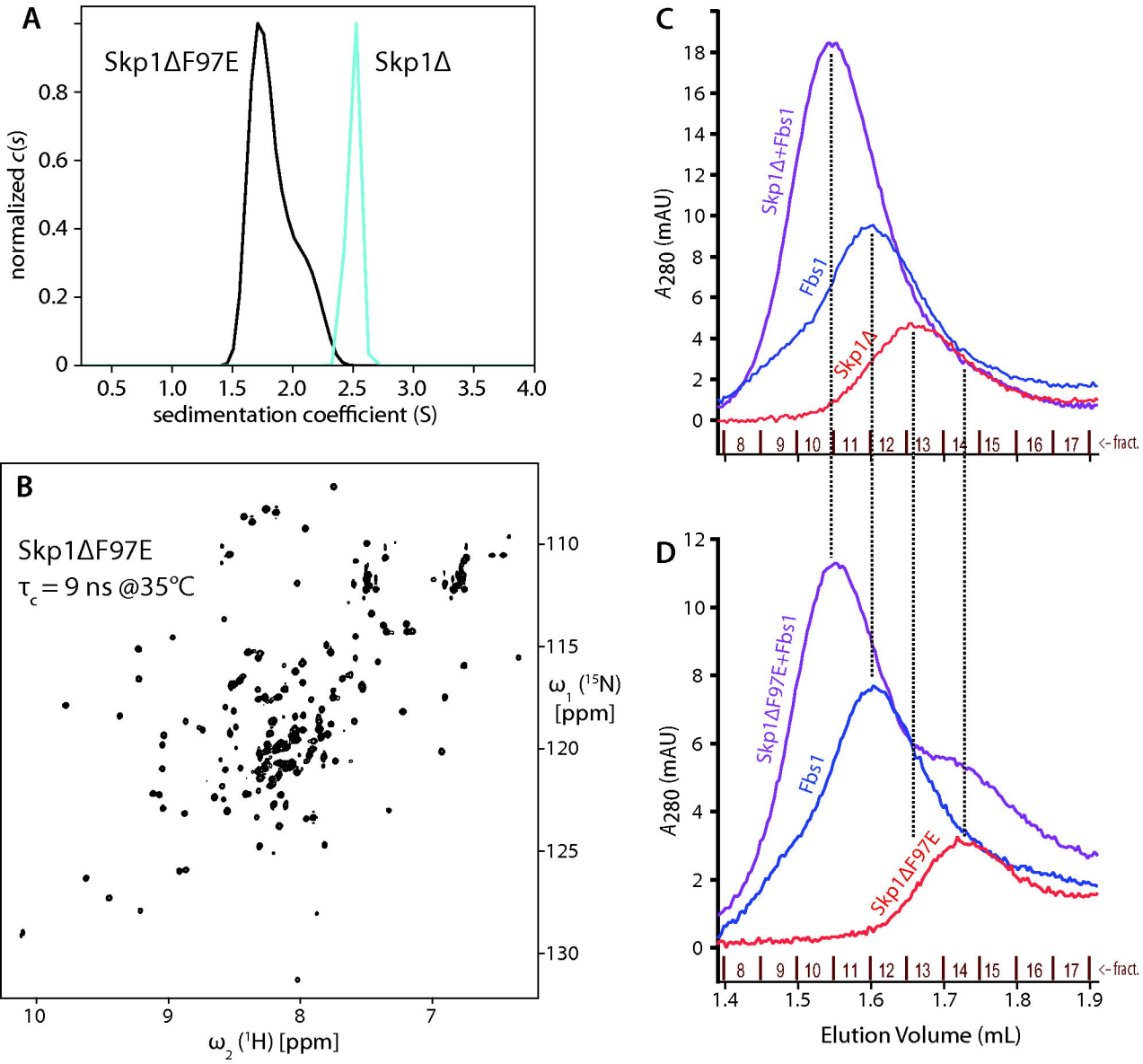

A
Binding free energy, ΔΔG
4
3
2
1
0
93
94
97
100
101
104
105
106
107
109
112
113
116
117
119
120
123
Interface residue position
B
C
Mutations
Arg
His
Lys
Asp
Glu
Ser
Thr
Asn
Gln
Cys
Gly
Pro
Ala
Val
Ile
Leu
Met
Phe
Tyr
Trp
93
94
97
100
101
104
105
106
107
109
112
113
116
117
119
120
123
Position
4
2
0
-2
-4
Binding free energy, ΔΔG

Figure 4


A
Skp1ΔF97E
Skp1Δ
normalized c(s)
sedimentation coefficient (S)
B
Skp1ΔF97E
τc = 9 ns @35°C
ω1 (15N) [ppm]
ω2 (1H) [ppm]
C
Skp1Δ+Fbs1
Fbs1
Skp1Δ
A280 (mAU)
<-fract.
D
Skp1ΔF97E+Fbs1
Fbs1
Skp1ΔF97E
A280 (mAU)
<-fract.
Elution Volume (mL)