Structure of SAP18: A Ubiquitin Fold in Histone Deacetylase Complex Assembly‡

Scott A. McCallum,§,⊥ J. Fernando Bazan,§ Mark Merchant,|| JianPing Yin,§ Borlan Pan,§ Frederic J. de Sauvage,|| and Wayne J. Fairbrother*,§

Departments of Protein Engineering and Molecular Biology, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080-4990

Received April 7, 2006; Revised Manuscript Received July 20, 2006

ABSTRACT: Signal transduction pathways are frequently found to repress transcription of target genes in the absence of stimulation and, conversely, to upregulate transcription in the presence of a signal. Transcription factors are central in this dual regulatory mechanism and widely use a generalized mechanism to repress transcription through recruitment of a Sin3−histone deacetylase (HDAC) complex to their binding sites on DNA. The protein SAP18 (Sin3-associated polypeptide of 18 kDa) has been shown to play a key role in gene-specific recruitment of the HDAC complex by a number of transcription factors including Gli, GAGA, and Bicoid. The solution structure of SAP18 reveals a ubiquitin-like fold with several large loop insertions relative to other family members. This fold supports the functional role of SAP18 as a protein−protein adapter module and provides insight for how SAP18 may bridge the Sin3−HDAC complex to transcription factors.

The Hedgehog (Hh) pathway is involved in many aspects of development, and misregulation of its activity leads to developmental disorders and cancer. Its activity is mediated for the most part by the Gli zinc-finger family of transcription factors (*1*−*3*). In the absence of the Hh ligand, the Hh signaling pathway represses transcription of target genes that include important regulators of cell growth and differentiation. This inhibitory role is mediated, in part, through the recruitment of the Sin3−histone deacetylase (HDAC)[1] complex. Gene-specific recruitment of HDAC activity is recognized as a major regulatory mechanism of gene expression (for review, see refs *4* and *5*). HDAC activity drives the packaging of DNA into condensed chromatin and leads to transcriptional repression by making the DNA inaccessible to transcription machinery. Gli-dependent HDAC recruitment and function requires the assemblage of a number of protein cofactors, among which is the accessory protein Sin3-associated polypeptide of 18 kDa (SAP18) (*6*).

SAP18 was first identified by immunopurification of Sin3-associated proteins; it was further demonstrated to be a component of the Sin3−HDAC complex and to enhance Sin3-mediated repression of transcription (*7*). Human SAP18 is 153 amino acids in length and contains a single cysteine residue. SAP18 tends to be highly conserved among higher eukaryotes with an identity of approximately 98, 60, and 54% in sequence alignments between the human polypeptide and homologues from mouse, *Drosophila melanogaster*, and *Caenorhabditis elegans*, respectively; a distant orthologue of SAP18 is also detected in fission but not fusion yeast genomes (see Figure S3 in the Supporting Information). However, sequence alignments reveal no significant domain links to non-SAP18 homologues and therefore no matches to known folding motifs (*8*).

SAP18 has been reported to bind a number of transcription factors (*6*, *9*, *10*). While the functional role of SAP18 has not been defined clearly, Gli and Bicoid have been shown to require SAP18 to repress transcription of their respective target genes through recruitment of the Sin3−HDAC complex (*6*, *11*). In a study reported by Cheng and Bishop, SAP18 was shown to function as a protein−protein adapter that bridges the transcription factor Gli, the transcriptional repressor suppressor of fused [Su(fu)], and the Sin3−HDAC complex (*6*). Using a biotinylated DNA oligonucleotide containing the binding site of Gli as bait, Gli, SAP18, and the Gli-binding protein Su(fu) were required to pull down the Sin3−HDAC complex. Additional *in vitro* and yeast two-hybrid analyses demonstrated specific and direct binding between SAP18 and Su(fu). Results from cell-based assays show that SAP18 and Su(fu) act cooperatively with Gli to recruit histone deacetylation machinery and to repress the transcription of Hh target genes in the absence of Hh stimulation (*6*, *12*).

The association of SAP18 with a select number of transcription factors supports the hypothesis that the HDAC complex utilizes a modular architecture that includes several adapter proteins. In this model, adapter proteins such as SAP18 organize the assembly of the HDAC enzyme and DNA-bound transcription factors to provide gene-specific targeting of HDAC activity. Therefore, the temporal and spatial distribution of transcription factors and HDAC adapter proteins are important factors in this transcriptional regulatory mechanism. Here, we report the nuclear magnetic resonance (NMR)-based structure of SAP18 and speculate, on the basis

‡ The coordinates of SAP18 have been deposited in the RCSB Protein Data Bank under accession code 2HDE.

* To whom correspondence should be addressed. Telephone: 650-225-6372. Fax: 650-225-3734. E-mail: fairbro @gene.com.

§ Department of Protein Engineering.

⊥ Current address: Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180-3590.

|| Department of Molecular Biology.

[1] Abbreviations: SAP18, Sin3-associated polypeptide of 18 kDa; HDAC, histone deacetylase; RDC, residual dipolar coupling; BTB, BR-C, ttk and bab.

10.1021/bi060687l CCC: $33.50
Published on Web 09/07/2006

of the SAP18 structure and structure−function relationships of structurally related proteins, the functional role of SAP18 as a protein adapter module within several larger molecular assemblies including the present work with Su(fu)−Gli in Hh pathway signaling and the Sin3−HDAC complex in chromatin regulation.

EXPERIMENTAL PROCEDURES

NMR Sample Preparation. ^{15}N- and ^{15}N/^{13}C-labeled SAP18 (amino acids 6−149, C26S) was overexpressed in *Escherichia coli* BL21 (DE3) RIL-codon plus (Stratagene) as a thrombin-cleavable amino-terminal His_6-tagged protein, purified, and cleaved specifically as described (*13*). NMR samples typically contained approximately 1 mM protein in buffer composed of 50 mM sodium phosphate (pH 5.5), 50 mM NaCl, 20 μM ethylenediaminetetraacetic acid (EDTA), 3 mM NaN_3, 0.1 mM DSS, and 8% D_2O. A "100%" D_2O sample was prepared by resuspension of lyophilized SAP18 in 99.995% D_2O.

NMR Spectroscopy. NMR spectra were acquired at 25 °C on Bruker DRX-600 and DRX-800 spectrometers equipped with a 5 mm inverse triple-resonance cryoprobe with a *z*-axis gradient and a 5 mm inverse triple-resonance probe with three-axis gradient coils, respectively. Chemical shifts were referenced to internal DSS (*14*). Near complete backbone and side-chain resonance assignments were obtained using standard methods (*15*) and were reported previously (BMRB accession number 6371) (*13*). The following nuclear Overhauser effect spectrometry (NOESY) spectra were acquired using the ^{15}N-labeled sample in 92% H_2O/8% D_2O: 3D ^{15}N-heteronuclear single-quantum coherence (HSQC) NOESY (mixing times of 30 and 150 ms) and a 3D ^{15}N-resolved ^{15}N-^{1}H-HSQC NOESY (mixing time of 150 ms). The following NOESY spectra were acquired using the ^{15}N/^{13}C-labeled sample: 3D ^{13}C-resolved ^{15}N-HSQC NOESY (CNH-NOESY; mixing time of 150 ms) in 92% H_2O/8% D_2O and a 3D ^{13}C-HSQC NOESY (mixing times of 30 and 150 ms) in D_2O. Steady-state ^{1}H-^{15}N nuclear Overhauser effects (NOEs) were measured on the ^{15}N-labeled sample using pulse sequences described previously that were modified to simultaneously evolve ^{1}H-^{15}N coupling and frequency label ^{15}N nuclei during a semiconstant time period (*16*, *17*). The spectra were acquired with and without saturation of amide proton magnetization in an interleaved manner. Saturation was achieved through applying a train of pulses with an effective tilt angle of 120 separated by 5 ms during the 6 s recycle delay period. Amide hydrogen-exchange kinetics were measured during the preparation of a "100%" D_2O sample. After lyophilization, ^{15}N/^{13}C-labeled SAP18 was resuspended in 99.995% D_2O and a series of 2D ^{1}H-^{15}N-HSQC spectra were acquired. Spectral processing and analysis were performed using FELIX (version 2000.1; Accelrys, San Diego, CA).

Distance and Dihedral Restraints. Distance restraints were derived from the intensities of manually picked peaks in the 3D ^{13}C-edited NOESY (mixing time of 30 ms) and the 3D ^{15}N-edited NOESY (mixing time of 30 ms). Automated NOE assignments were obtained in combination with structure calculations using a standard CANDID (version 1.1) (*18*) protocol and the CYANA program (version 1.06) (*19*). Manual restraint checking was performed in repetitive rounds of structure/assignment calculations by analyzing the NOESY spectra acquired with both the 30 and 150 ms NOE mixing times and in 3D NNH, CNH, and CCH NOESY spectra. Distance restraints were calibrated using the structure-dependent calibration module of CYANA and preliminary structural models also refined within CYANA. Pseudo-atom corrections were added to the upper bound limits when necessary. Distance restraints involving methylene and aromatic protons with degenerate chemical shifts and from methyl groups were calculated as sums of the individual contributions. A 0.5 Å correction was added to the upper bounds for restraints involving methyl groups with the lower bounds for all distance restraints set to the sum of the van der Waals radii. Hydrogen-bond restraints were inferred on the basis of protection from amide hydrogen−deuterium exchange, inter-residue ^{1}HN−^{1}Hα, ^{1}Hα−^{1}Hα, and ^{1}HN−^{1}HN NOE patterns, dihedral angle restraints, and structural models calculated in the absence of hydrogen bonds. TALOS was used to provide 103 ϕ and 100 ψ backbone torsion angle restraints based on ^{1}Hα, ^{13}Cα, ^{13}Cβ, ^{13}CO, and ^{15}N secondary chemical shifts (*20*). Angles were typically restrained to ±2 or 3 times the standard deviation for matches classified as "good" and "warn", respectively.

Residual Dipolar Coupling (RDC) Restraints. Partial molecular alignment of SAP18 was achieved using 6.8% (w/v) stretched polyacrylamide gel matrix as described by Bax and co-workers (*21*). Partially dried gel (83:1 acrylamide to bisacrylamide) was rehydrated with uniformly ^{15}N-labeled SAP18 to provide final concentrations of 0.5 mM protein, 50 mM sodium phosphate (pH 5.5), 50 mM NaCl, 20 μM EDTA, 3 mM NaN_3, 0.1 mM DSS, and 8% D_2O (see the Supporting Information for details). The gel was then extruded from the 5.4 mm casting chamber into a NMR tube with an inside diameter of 4.24 mm using gel stretching accessories from New Era Enterprises. Residual dipolar couplings (RDCs) were measured for backbone amide groups in partially aligned and gel-free states using a modified in-phase/anti-phase (IPAP) HSQC that suppresses natural-abundance ^{15}N background signals from amide groups in polyacrylamide (*22*). RDC data were not used from amide groups with significantly overlapped resonance peaks or from structural regions assessed to be dynamic via heteronuclear NOE analysis. Because of broad amide linewidths and poor spectral dispersion, RDC restraints were applied using a square-well penalty function with a ±4 Hz width. A grid search procedure was used to determine the axial (D_a) and rhombic (D_r) components of the alignment tensor in Xplor-NIH (*23*, *24*). For each combination of D_a and R (defined as D_a/D_r) values in the grid search, SAP18 structures, calculated in the absence of RDC restraints, were rotated as single rigid structural units to a fixed tensor axis system to obtain the best agreement between the predicted and measured RDC values. An optimal D_a of 13.0 ± 1.1 Hz and R of 0.31 ± 0.09 were calculated as the average values in fits for the five lowest-energy solutions.

Structure Calculations. Structural refinement was performed using a simulated annealing protocol with restrained torsion angle dynamics using CNX (version 2002.02; Accelrys). A total of 100 structures were calculated using a refinement protocol that started with high-temperature restrained torsion angle dynamics (15 ps at 50 000 K) followed by a cooling stage (120 ps), Cartesian dynamics (20 ps), and conjugate-gradient minimization. The force constant for RDC

restraints increased from 0.01 to 2.0 kcal mol^{-1} Hz^{-2} during the final 5000 K of the cooling stage. The 20 structures with the lowest restraint violation energies were chosen to be included in the final solution structure ensemble. The quality of the structural models was evaluated and visualized using Procheck-NMR (version 3.5.4) (*25*), INSIGHT (version 2000; Accelrys), and PyMOL (version 98.5; DeLano Scientific).

Identification of a Minimal SAP18-Binding Fragment of Su(fu). Su(fu) plasmids used in this study have been described previously (*26*). Human SAP18 was cloned into a pRK5-based mammalian expression vector that encodes a N-terminal Myc-epiptope tag, encoded by synthetic oligonucleotide linkers (see the Supporting Information). Clones were sequence-verified and found to express an 18−19-kD-sized Myc-tagged protein as analyzed by Western blot when transiently transfected by GeneJuice (Novagen) into CHO-7 cells and HEK293 cells.

Transfections, immunoprecipitations, and Western blot analysis were performed as described previously (*26*). COS-7 cells were grown in 50:50 media (half Ham's F-12-low glucose and half Delbecco's modified Eagle's media) containing 10% fetal bovine serum, 2 mM glutamine, and 10 mM HEPES. The cells were transiently transfected with various expression plasmids in 10 cm plates with GeneJuice (Novagen). A total of 24 h after transfection, cells were lysed for 20 min at 4 °C with rotation in 1 mL of 1% NP-40 lysis buffer [50 mM Tris (pH 8.0), 150 mM NaCl, 1 mM EDTA, 1 mM PMSF, and a cocktail of additional protease inhibitors (Complete, Boehringer Mannheim)]. Samples were centrifuged at 20000*g* at 4 °C for 20 min to remove NP-40 insoluble proteins. Alternatively, constructs were expressed alone, and lysates were mixed after normalization. Reaction mixtures were then rotated at 4 °C for 4 h to allow for binding, after which immunoprecipitations were performed.

For immunoprecipitations, lysates were precleared for 1 h at 4 °C with 30 μL of Protein A Sepharose (Amersham/Pharmacia Biotechnology, Inc.). Lysates were transferred to a new tube containing 3 μg of anti-Myc (9E10, Boehringer Mannheim), anti-FLAG-M2 (Sigma), or anti-GST (Amersham/Pharmacia Biotechnology, Inc.) monoclonal antibodies and rotated for 2 h at 4 °C. Protein A Sepharose beads (30 μL) were then added to each sample and allowed to rotate for 1 h. The immune-complexed beads were then washed 4 times with 1 mL of lysis buffer. To ensure that binding was specific, some immunoprecipitations were subjected to two high-salt washes (500 mM NaCl). The immune-complexed beads were resuspended in 30 μL of sodium dodecyl sulfate (SDS) loading buffer (Invitrogen, Inc.) containing β-mercaptoethanol and boiled at 100 °C for 5 min. Samples were analyzed using denaturing 4−20% SDS−PAGE, transferred to 0.2 μm nitrocellulose membranes, and blocked for 1 h with 5% milk and 2% BSA in phosphate-buffered saline (PBS). Membranes were then probed with anti-Myc (9E10)-peroxidase (Boehringer Mannheim, Inc.) at 1:2500 in blocking solution, anti-FLAG-M2-peroxidase (Sigma) at 1:1500 in blocking solution, or anti-GST (Amersham/Pharmacia Biotechnology, Inc.) at 1:3000 in PBS containing 0.05% Tween-20 (PBST) for 1 h at room temperature. Blots were washed 4 times with 10 mL of PBST for 5 min, each followed by chemiluminescent development using Chemiglow West (Alpha Innotech, Inc.).

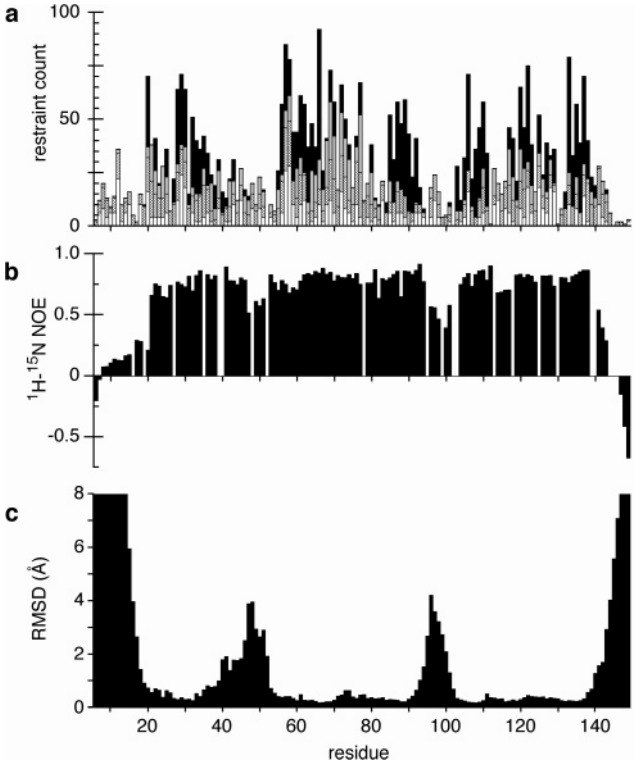

FIGURE 1: Residue-based structural statistics for residues 6−149 of SAP18. (a) Residue-based summary of NOE distance retraints. Intraresidue, sequential, medium-range ($1 < |i - j| < 5$), and long-range ($|i - j| \geq 5$) NOEs are displayed as open, hatched, gray, and black bars, respectively. (b) ^{1}H-^{15}N NOE values plotted as a function of the residue number. (c) Atomic pairwise rmsd values determined for the backbone heavy atoms (N, C$^\alpha$, and C) of the 20 final structures (superposing residues 21−139).

RESULTS

Structure Determination. The manually assisted CANDID approach was effective at combined autoassignment of NOESY cross-peaks and structure determination of SAP18. Here, rounds of generating automated assignments were interleaved with manual interactive checking of both assigned but violated and nonassigned peaks from ^{15}N- and ^{13}C-edited NOESY spectra. During this series of calculations, vetted NOE crosspeak data were supplemented with restraints for 32 hydrogen bonds and 203 dihedral angle restraints. As a result, a total of 2165 nonredundant and structurally restraining NOE-derived distance restraints were generated. A residue-based summary of the 377 intraresidue, 569 sequential, 382 medium-range ($1 < |i - j| < 5$), and 837 long-range ($|i - j| \geq 5$) distance restraints is provided in Figure 1a. SAP18 was refined using CNX with 2165 interproton distance restraints, 203 ϕ and ψ dihedral angle restraints, 32 hydrogen bonds (employed as 2 distance restraints per hydrogen bond), and 62 ^{1}H-^{15}N RDC restraints as input (see Table 1).

Structural statistics for the final 20-member ensemble (Figure 2) are summarized in Table 1. The calculated structures all converged to solutions with low violations and with similar atomic coordinates for residues 20−36, 54−91, and 102−139. In contrast, residues 1−19 (N terminus) and 140−149 (C terminus) and residues within two extended internal loops between β strands 1 and 2 (residues N36−S53) and β strands 3 and 4 (residues V92−V103), respectively, are poorly defined by the data and appear conforma-

tionally unrestricted in the calculated structures (see Figures 1c and 2). ^{1}H-^{15}N heteronuclear NOE ratios suggest the presence of high-frequency structural dynamics in most of these disordered regions (Figure 1b).

Figure 1 highlights the strong correlation between residues with high atomic root-mean-square deviations (rmsd values) and residues with few long-range NOEs and low ^{1}H-^{15}N heteronuclear NOE ratios. More specifically, the N and C termini and residues V96−Y101 of the β3−β4 loop are determined to be dynamically disordered on the basis of relatively low ^{1}H-^{15}N heteronuclear NOE ratios, resonance line shape, and the absence of long-range NOE connectivities (see parts a−c of Figure 1). Amide peaks were not observed for residues D95, R102, and V103, a possible result of line broadening associated with loop dynamics or fast amide proton exchange with the solvent. In contrast, many residues within the β1−β2 loop have relatively high ^{1}H-^{15}N heteronuclear NOE ratios and NOE patterns consistent with the helical structure. Furthermore, Procheck-NMR detects a helical turn formed by residues R42−F46 in all of the final ensemble members, yet loop residues have poor superpositions when using all core residues (see Figure 1c) or just those from within the loop (not shown). The ability to structurally define the residues within the helical turn beyond the secondary structure is clearly limited by the lack of long-range distance restraints. Residues after the helical turn in the β1−β2 loop have markedly reduced ^{1}H-^{15}N heteronuclear NOE values and virtually no long-range NOEs. Therefore, this portion of the β1−β2 loop is likely structurally disordered and flexible on the pico−nanosecond time scale, as are residues V96−Y101 in the β3−β4 loop and in the N- and C-terminal regions.

The structures satisfy the experimental restraints well, as evident by the low restraint violation energies (see Table 1), with the greatest NOE and dihedral violations within structured regions being 0.15 Å and 2.5°, respectively. The protein exhibits good covalent geometry with approximately 84, 16, and 0.2% of all ψ, ϕ angles in the favored, allowed, and generous regions of the Ramachandran plot, respectively, and none in disallowed regions. The atomic rmsd values from the mean coordinates for core residues (20−36, 54−91, 102−139) in the final 20 structures are 0.37 ± 0.07 and 0.95 ± 0.11 Å for backbone and all heavy atoms, respectively. When loop residues 37−53 and 92−101 are included in the superposition, atomic rms differences from the mean coordinates increase to 1.03 ± 0.29 and 1.66 ± 0.24 Å for backbone and all heavy atoms, respectively.

Description of Structure. A central fragment of SAP18 (spanning amino acid residues 15−149) was used in an automated search of the Protein Data Bank (PDB) structure repository with the DALI program (*27*), revealing a decided topological similarity to ubiquitin-like proteins and other members of the large β-grasp fold class. The SAP18 β-grasp architecture comprises a mostly antiparallel five-stranded β sheet (with a strand order of 2−1−5−3−4, where β1 and β5 have a parallel orientation) curled around a central α-helix H2, as shown in Figure 3. In addition to the close topological match, particular structural features shared by SAP18 with ubiquitin-like folds include the overall β-sheet dimensions, orientation of the helix within the curvature of the β sheet, and conservation of a well-ordered turn that includes both a positive ϕ angle for K126 (N60 in ubiquitin) and a short helix (residues 122−125) in the linker between β4 and β5. For instance, 73 Cα atoms of SAP18 superimpose to corresponding positions within the archetype ubiquitin fold

Table 1: Structural Statistics for the Solution Structure of SAP18[a]

parameter	ensemble
input restraints	
NOE total	2165
intraresidue	377
sequential	569
medium range	382
long range	837
dihedral angles total	203
ϕ	103
ψ	100
hydrogen bonds[b]	32
RDCs	62
violations	
rmsd from experimental restraints	
NOE distance (Å)	0.01 ± 0.0005
dihedral (deg)	0.20 ± 0.06
NOE distance violations	
number > 0.1 Å	3.6 ± 1.6
number > 0.2 Å	0.0
mean maximum violation (Å)	0.13 ± 0.03
dihedral violations	
number > 0.1°	6.9 ± 2.2
mean maximum violation (deg)	1.6 ± 0.6
RDC violations	
number > 0.1 Hz	0.0
mean maximum violation (Hz)	0.01 ± 0.01
rmsd from idealized geometry	
bonds (Å)	0.0018 ± 0.0002
angles (deg)	0.33 ± 0.01
impropers (deg)	0.21 ± 0.02
energies[c]	
energy components (kcal mol^{-1})	
NOE (2165)	11.2 ± 1.7
CDIH (203)	0.5 ± 0.2
RDC (62)	14.5 ± 16.5
bonds	7.3 ± 1.7
angles	71.2 ± 5.3
impropers	8.5 ± 1.6
van der Waals	−413.3 ± 20.0
stereochemistry	
Ramachandran (%)	
favored	83.9
allowed	15.8
generous	0.2
disallowed	0.0
structural precision	
atomic rmsd values (Å)[d]	
backbone[e]	
residues 20−139	1.03 ± 0.29
core residues[f]	0.37 ± 0.07
heavy[g]	
residues 20−139	1.66 ± 0.24
core residues[f]	0.95 ± 0.11

[a] All terms are calculated for the 20 low-energy structures. [b] Each hydrogen bond was employed as two distance restraints with r_{HN-O} = 1.8−2.4 Å and r_{N-O} = 2.7−3.5 Å. [c] The final force constants of the target functions used were 50 kcal mol^{-1} Å^{-2} for NOE, 200 kcal mol^{-1} rad^{-2} for dihedral angle, 2.0 kcal mol^{-1} Hz^{-2} for RDC, 1000 kcal mol^{-1} Å^{-2} for bond length, 500 kcal mol^{-1} rad^{-2} for angle and improper dihedral, 4 kcal mol^{-1} Å^{-4} for the Lennard−Jones potential function with a 10 Å cutoff (employing effective van der Waals radii set to 0.80 times their value in the CHARMM parameters). [d] Atomic rmsd values were calculated with respect to mean coordinates. Mean coordinates were determined using a best fit superposition of the specified atoms among all ensemble members. [e] Atomic rmsd values were determined for backbone (N, C$^{\alpha}$, and C) atoms. [f] Atomic rmsd values were determined for the structurally ordered core residues 20−36, 54−91, and 102−139. [g] Atomic rmsd values were determined using all heavy atoms from the specified region.

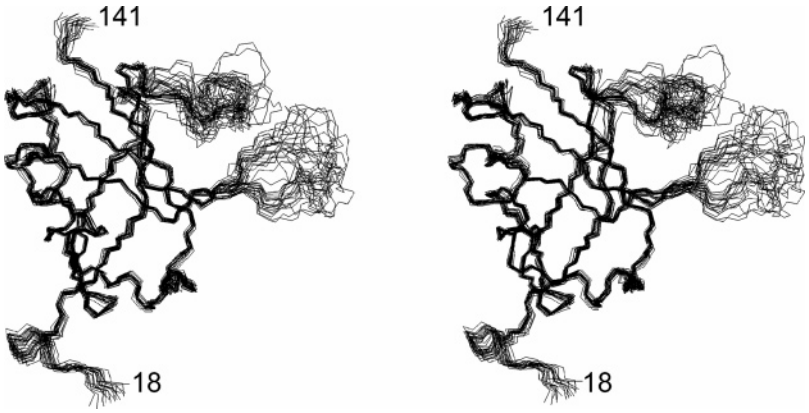

FIGURE 2: Stereoviews of the final 20 SAP18 structures are shown superposed. The superimposition was optimized for backbone atoms (N, C^{α}, and C) of ordered residues 20−36, 54−91, and 102−139 to the mean coordinates. Residues from the nonstructured N (6−17) and C (142−149) termini are not shown.

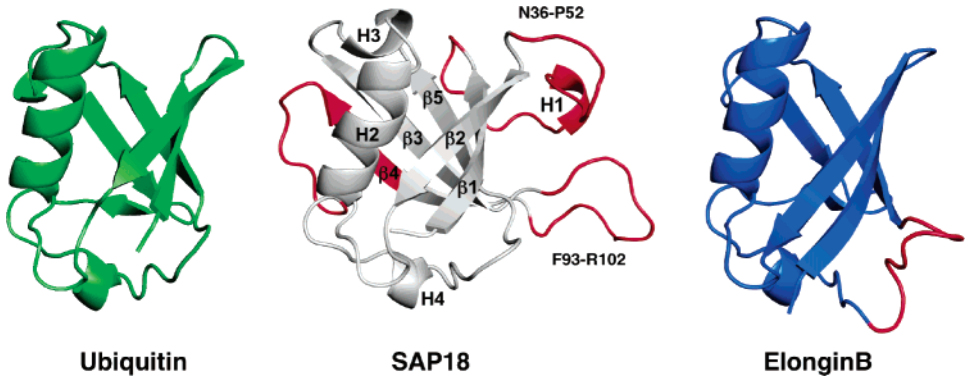

FIGURE 3: Ribbon representation for residues 25−139 of SAP18 is shown with structurally related proteins ubiquitin and Elongin B. Inserts within the ubiquitin-fold architecture are shown in red for SAP18 and Elongin B. This figure was made using PyMOL (*46*).

(PDB file 1UBI) with approximately a 1.8 Å rmsd, a surprising level of structural conservation given the 11% sequence identity between SAP18 and ubiquitin.

Notable unique features in the SAP18 architecture relative to the compact ubiquitin fold include 3 large insertions comprising approximately 40 residues; in addition, 25 and 14 extra residues at the N and C termini, respectively, form flexible chain extensions to the core fold. The position and size of the inserts are highlighted in Figure 3 in red relative to ubiquitin [PDB ID 1D3Z (*28*)] and the ubiquitin-like protein Elongin B [PDB ID 1VCB (*29*)]. Two loop insertions form spatially close extended fingers between β strands β1 and β2, and β3 and β4. As indicated previously, these loops are poorly defined in the calculated structures and are likely dynamically disordered in solution. In contrast, the third insert is structurally well-defined in SAP18 and extends the β4 strand by several residues relative to ubiquitin. Interestingly, a β bulge (residue I106) is located at the residue position at which the β4 strand typically terminates in other ubiquitin-like proteins.

Additional structural differences between SAP18 and the archetype ubiquitin fold arise from the extension of the β4 strand. In SAP18, the β3 strand is extended by several residues, forcing a significant displacement of the H3 helical turn within the connecting loop between helix H2 and strand β3, relative to ubiquitin. In the ubiquitin fold, a helical turn caps the β3 strand in a conformation that clashes with the closely apposed termini of β3 and β4 strands in SAP18. As a consequence, the helical H3 turn in SAP18 is nestled in a cleft formed on one side by the C end of the major H2 helix and on the other side by residues from the extended ends of the β3 and β4 staves. The 25- and 15-residue chain extensions beyond the ubiquitin fold at the N and C termini are not unique to SAP18; for example, the small ubiquitin-related modifier (SUMO) has a 20-residue extension and Elongin B has a 39-residue extension at the N and C termini, respectively, relative to the ubiquitin fold.

While ubiquitin and its close relatives [together, the ubiquitin-like superfamily in the SCOP-fold hierarchy (*30*, *31*)] form the largest cohort of resolved β-grasp structures, the remainder are a remarkably diverse lot of molecules with neither sequence similarity or functional overlap. This extraordinarily broad subscription to a common core fold is characteristic of what have been termed superfolds, select protein architectures with perhaps an ancient origin, very stable chain topologies, and plasticity to great sequence and functional divergence (*32*). To better classify SAP18 within many branches of the β-grasp superfold class, targeted, domain-level structural comparisons against SCOP (β-grasp) fold family d.15 were conducted with the SSM program (*33*). As judged by several measures, *Q*, *P*, or *Z* scores, rmsd, and/or the length of Cα alignment, this survey consistently grouped SAP18 first with members of the parent ubiquitin superfamily (SCOP index d.15.1.1) and then progressively superposed the FERM module family (d.15.1.4), Ras-association (RA) domains (d.15.1.5), UBX domains (d.15.1.2), the PB1 family (d.15.2.2), Ig-binding domains (d.15.7.1), GABA(A) receptor-associated protein (GABARAP) modules (d.15.1.3), and doublecortin domains (d.15.11.1) (see Figure S2 in the Supporting Information). With few exceptions, the

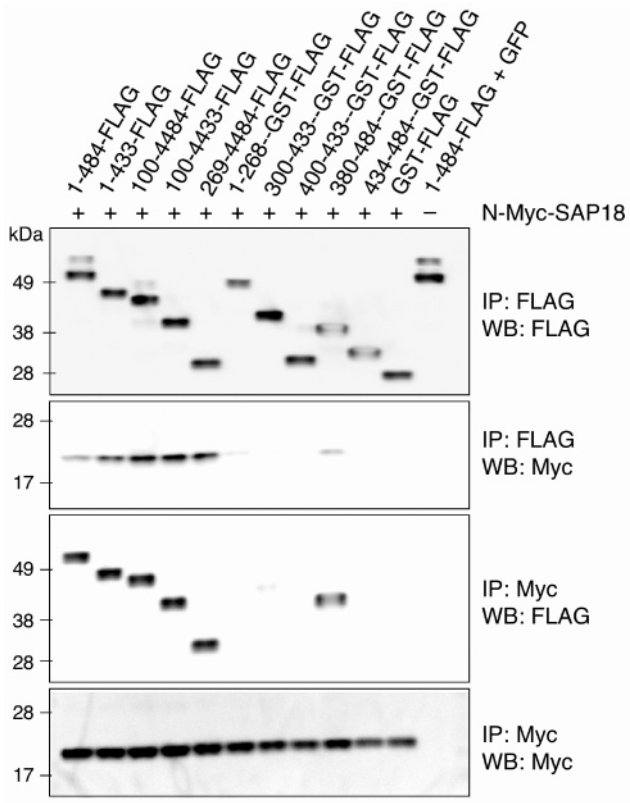

FIGURE 4: Identification of a SAP18-binding region in Su(fu). Results are summarized from co-immunoprecipitation (IP) and Western blot (WB) analysis between FLAG epitope-tagged versions of C-terminal fusion constructs of Su(fu) and a N-terminal Myc-SAP18 fusion construct. Size in kilodaltons is denoted on the side of the blots.

sequence identity over aligned residues between SAP18 and the β-grasp domains in SCOP remains in the 10−15% range, where it is difficult to gauge if the structural resemblance is due to sequence divergence. However, the closer structural matches between SAP18 and the ubiquitin-like superfamily, despite the loop embellishments of SAP18, invite a SCOP-level classification in a new family within the d.15.1 ubiquitin-like fold group, at the level of the ubiquitin-related (grouping ubiquitin with molecules like SUMO, NEDD8, and Elongin-B) RA, UBX, FERM, or GABARAP domain families.

Identification of a Minimal SAP18-Binding Fragment of Su(fu). The identification and characterization of SAP18-interacting domains will be critical to developing an understanding of the functional role of SAP18 as an adapter protein and the structural organization of the HDAC complex. Because a direct interaction between Su(fu) and SAP18 has been reported to be integral to Gli recruitment of the HDAC complex, identification of the SAP18-binding domain of Su(fu) will provide insight into the structural basis for how the HDAC activity is recruited by the transcription factor. To address what portion of Su(fu) interacts with SAP18, an amino-terminal Myc-epitope-tagged SAP18 construct (N-Myc-SAP18) was overexpressed in cells expressing various FLAG-epitope-tagged truncation variants of Su(fu) and tested for binding by co-immunoprecipitation. Full-length Su(fu) [Su(fu)1−484-FLAG] was capable of pulling down N-Myc-SAP18 (see Figure 4). Likewise, Su(fu) regions 1−433, 100−484, 100−433, and 269−484 were all capable of interacting very strongly with N-Myc-SAP18. The N-terminal domain of Su(fu) encoded by Su(fu)1−268-GST-FLAG was unable to interact with SAP18 (Figure 4). This segment of Su(fu) contains a discrete globular domain whose function is to engage the C-terminal end of Gli and whose $\alpha + \beta$ fold has been earlier resolved by X-ray crystallography (*26*). GST-FLAG constructs expressing C-terminal Su(fu) residues 300−433, 400−433, and 433−484 were all unable to bind to SAP18. However, a Su(fu)-380−484-GST-FLAG construct was able to pull down SAP18, albeit less robustly than Su(fu) regions 1−433, 100−484, 100−433, and 269−484 (see Figure 4). These results indicate that Su(fu) binds to SAP18 through a minimal region encoded between residues 380 and 484. This C-terminal segment of Su(fu) is predicted to encase a conserved globular domain whose boundaries are well-predicted by disorder analysis using the Disprot program [VL3H method (*34*)] and is of mixed $\alpha + \beta$ secondary-structure composition by PsiPRED analysis (*35*) (see Figure S1 in the Supporting Information).

DISCUSSION

As the fold analysis of SAP18 recapitulates, protein structure (in the form of a cryptic ubiquitin-like β-grasp domain) is far better preserved than sequence. The manner in which domain folds interact appears to be similarly well-conserved (*36*), suggesting that perhaps the biochemical function of SAP18 is to serve as a ubiquitin-like adapter or interaction module, akin to ubiquitin-related molecules that serve as covalent protein tags, like ubiquitin (*37*), SUMO (*38*), Fat10 (*39*), and Nedd8 (*40*), and recruit similar molecular machinery in the cell (*41*, *42*). The ubiquitin-like fold is also found in a N-terminal FERM domain, where it functions as an adapter module within key proteins that form cytoskeleton−membrane junctions (*43*, *44*) and in the GABA(A)-receptor-associated protein that links GABA(A) receptors to the cytoskeleton (*45*). In addition, the UBX, Elongin B, and the family of RA domain proteins all utilize the ubiquitin-like fold as an adapter module in the recognition or assembly of larger protein complexes (*30*). Is there accordingly a common structural theme in the interaction modalities of these ubiquitin-like folds with other molecules? True to the highly gregarious nature of superfolds (*32*), ubiquitin-like proteins have been captured in a variety of different complexes that, in sum, sample nearly every face of the core fold surface (*30*). Still, some preferred sites for protein−protein interactions do emerge within the circle of ubiquitin-like families. For instance, ubiquitins famously utilize a hydrophobic patch centered on β3 to engage a wide variety of helical ubiquitin-interaction motifs and zinc-finger modules (*41*). In contrast, the related SUMO proteins capture linear SUMO-binding motifs by bidirectional β-strand addition in the cleft between β2 and helix H2 (*46*). Elongin B contacts the BTB fold Elongin C protein by pairing of the selfsame β2 strand in parallel (*29*), while RA domains engage their prey in a related but antiparallel fashion (*30*). Continuing this focus on the β2 edge of the ubiquitin-like fold, PB1 domains form asymmetric heterodimers with a type-II module using its 2 side to dock the β4 end of a type-I module (*47*, *48*).

Conservation analysis has proven quite accurate in mapping the functional epitopes of protein structures that are tied to evolutionarily diverse sequence families (*49*, *50*). Drawing from a wide range of SAP18 homologues collected by

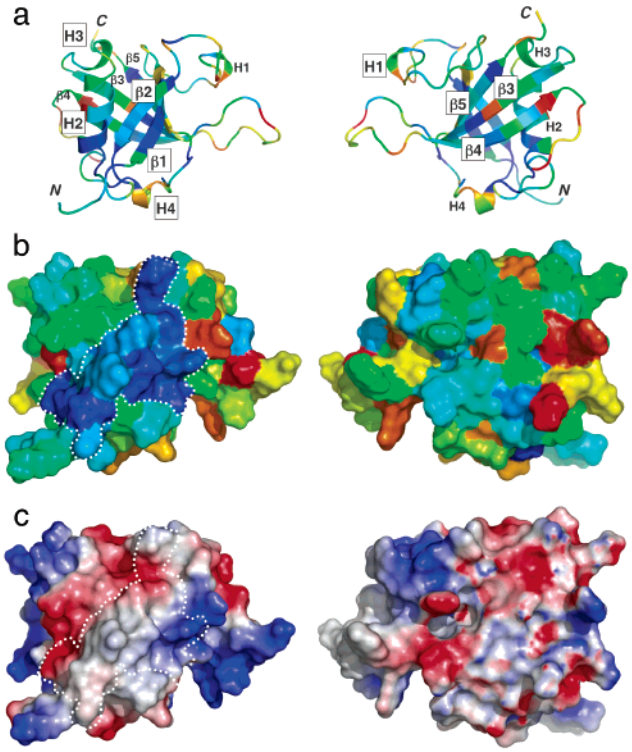

FIGURE 5: Conservation analysis of SAP18 shows a likely functional epitope (a) Consurf (*66*) was used to derive a surface-mapped conservation profile for SAP18, derived from a phylogenetically diverse alignment of SAP18 homologues. The conservation pattern is color-coded onto the ribbon diagram of SAP18 (encompassing amino acids 23−140) using PyMOL (*65*) spanning unconserved (red color) through moderately conserved (yellow-green) to very well-conserved (dark blue) positions (*66*, *67*). The two views are related by a 180° rotation about the vertical axis. (b) Same conservation profile is mapped to the solvent-accessible surface of SAP18, showing a clearly visible epitope (colored light to dark blue, toward maximal conservation, further highlighted by a white dotted outline) on one side of the structure but not the other, which runs along strand β2 and incorporates additional residues from helix H2 and strand β1. (c) Electrostatic potential surfaces of SAP18 (calculated by PyMOL using the vacuum electrostatics option) show that the conserved epitope of b overlaps with a largely uncharged (white) area that might be well-suited for protein−protein interactions. Conversely, the long, flexible loops of SAP18 show a high degree of positive charge (blue).

iterative PsiBLAST searches of a nonredundant sequence database (*51*), the T-coffee program (*52*) was used to generate a structurally faithful SAP18 multiple-sequence alignment to drape over the SAP18 structure and feed the ConSurf algorithm (*49*) (see Figure S3 in the Supporting Information). Surface patches of residues that retain identity or close chemical similarity across the phylogenetic gap between mammalian and fission yeast SAP18 molecules clearly point to a more conserved face of the SAP18 structure (Figure 5). This potential functional site in SAP18 encompasses the helix H2/strand β2 interface (as well as the neighboring N-terminal end of strand β1) in a manner that is reminiscent of the ConSurf-predicted and now structurally validated helix H2/strand β2 interaction site in SUMO for SUMO-binding motifs (*53*, *54*). This site is uncharged, largely hydrophobic, and surrounded by charged residues (Figure 5c) and resembles known protein−protein interaction epitopes. We therefore propose that SAP18 utilizes a SUMO-like β-strand addition mechanism (*55*) for the recognition and specific binding of short peptide stretches by strand pairing with SAP18 strand β2 and contact with nearby areas of helix H2, a type of interaction that is also similar to that used by Elongin B and RA domains (*30*). Interestingly, alanine-scanning analysis of ubiquitin surface residues has provided a hint of an analogous, secondary (to the primary β3 patch), but still silent β2-centered interaction site in ubiquitin (*54*, *56*). Conversely, SAP18 is unlikely to use a ubiquitin β3 patch-interaction mechanism because the corresponding β3 site is not well-conserved and is also partly occluded by the long, variable loop between β1 and β2 (Figure 5).

The structural and evolutionary delineation of a preferred interaction site in SAP18 invites functional parallels to molecular adapters like SUMO, Elongin B, and RA domains, all of which have converged on a similar interaction zone and strategy. Still, as pointed out in the case of the RA domain family, significant fold and sequence similarity do not ensure a similar binding profile (*57*). In the case of SAP18, there are now seven reported molecular interactors that include Sin3 in the HDAC complex (*7*), *Drosophila* GAGA (*10*), bicoid (*9*, *11*), Enhancer of zeste (*58*), the DNA polymerase-ϵ subunit (*59*), the ASAP and EJC protein complexes involved in pre-mRNA splicing [apoptosis- and splicing-associated protein (ASAP) (*60*) and exon junction complex (EJC) (*61*)], and Su(fu). Aside from some functional parallels [e.g., GAGA, bicoid, and Enhancer of zeste are DNA-binding proteins (*9*−*11*, *58*)], the SAP18 binders appear to be structurally unrelated proteins that do not share any common globular domains by sensitive SMART module analysis (*62*). The present work maps the earlier observed interaction between SAP18 and Su(fu) (*6*, *12*) to a C-terminal globular domain in the latter molecule; this region has an $\alpha + \beta$ secondary-structure signature of $\beta_2\alpha\beta_4\alpha\beta$ that may itself have a distant resemblance to ubiquitin-like folds with a $\beta_2\alpha\beta_2\alpha\beta$ signature (see Figure S1 in the Supporting Information). Two other SAP18-binding proteins with recognizable domains in mapped interaction regions include GAGA and the DNA polymerase-ϵ subunit, bearing predicted BTB and OB modules, respectively (*10*, *59*); there is a dearth of globular domains in other defined SAP18-interacting regions [e.g., bicoid and Enhancer of zeste (*9*−*11*, *58*)]. Together, this suggests that SAP18 is capable of flexible interactions with a structurally diverse set of ligands that may take the form of peptide motifs embedded in globular domains [Su(fu) or GAGA] or sitting in unstructured chain (bicoid or Sin3). The predicted helix H2/β2 strand interaction region of SAP18 could accommodate an Elongin B/Elongin C-type interaction (*29*) with GAGA, where the β3 strand of the GAGA BTB domain fold could slot alongside the SAP18 β2 strand. Analogously, an edge β strand from the Su(fu) ubiquitin-like fold could form a PB1-like heterodimer interaction (*47*, *48*) with the same SAP18 β2 strand. For the remaining molecules, short peptide stretches functionally analogous to subtle SUMO-binding motifs (*54*) could dock alongside the SAP18 β2 strand, following the β-strand addition mechanism (*55*). An unbiased effort to locate subtle yet conserved linear peptide motifs shared among the SAP18-binding proteins was conducted with the computational method of Neduva and Russell (*63*) and revealed a few candidate sequences that are being further evaluated (not shown). The fact that these types of potential β2-strand-centered interactions can bridge the gap from globular

domain to linear peptide recognition is suggested by the ability of the Elongin C−BTB domain β3-strand-derived synthetic peptides to effectively bind to the Elongin B ubiquitin-like fold (*64*). The binding promiscuity of SAP18 may have structural roots in a common binding mode that is convergently used by several ubiquitin-like families.

The diverse binding partners of SAP18 tie this ubiquitin-like molecular adapter to a range of larger, multisubunit assemblies like HDAC or the ASAP complex and are predicted to make use of a recognition mechanism, β-strand addition (*55*), that is particularly well-used by members of ubiquitin-like families that are more closely structurally similar to SAP18. The broad nature of SAP18-nucleated complexes also argues for a cast of supporting, unique interactions that make use of other, more variable surfaces of SAP18, including the long, variable loops and the C- and N-terminal chain extensions, to contact proteins specific to one molecular ensemble. For example, the Elongin B/Elongin C interface is centered on the β2 strand of Elongin B, but additional contacts are made by the inserted chain (relative to ubiquitin) between strands β4 and β5 (colored red in Figure 3) and by residues within a 39-residue C-terminal extension (*29*). Outside of convergent β-strand addition interactions between the ubiquitin-like core of SAP18 with key recruitment proteins [like Su(fu)], we do not need to invoke a greater homology between the types of larger complexes that have evolved to accommodate or make use of SAP18.

In the specific case of the transcription factor Gli in the Hh pathway, a heterodimeric assembly involving SAP18 and Su(fu) is speculated to bridge Gli and the Sin3 scaffold protein (*6*, *12*). A high-affinity binding site for the Gli transcription factors has been shown to be centered around highly conserved Su(fu) residues 390−400 (*26*) shown here to bind SAP18 (Figure 4) and likely fold into a discrete globular domain. Current efforts are focused on determining the structure of the SAP18-binding domain of Su(fu) in complex with SAP18 to start to decipher the binding specificity and plasticity of the predicted SAP18-binding groove. To date, however, a Su(fu) construct that is amenable to structural analysis in an unbound state and/or in complex with SAP18 has not been identified.

ACKNOWLEDGMENT

We thank the members of the DNA sequencing, protein sequencing, and mass spectrometry groups at Genentech for their analytical support.

SUPPORTING INFORMATION AVAILABLE

Details about the preparation of partially aligned SAP18 samples for RDC analysis and the construction of Su(fu) expression vectors; in addition, results of the directed structural comparison between SAP18 and ubiquitin-like folds, as well as an alignment of diverse SAP18 homologues that forms the basis of the conservation analysis; and a domain-based diagram of Su(fu) to highlight the SAP18-binding region. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Dominguez, M., Brunner, M., Hafen, E., and Basler, K. (1996) Sending and receiving the hedgehog signal: Control by the *Drosophila* Gli protein cubitus interruptus, *Science 272*, 1621−1625.
2. Cohen, M. M., Jr. (2003) The hedgehog signaling network, *Am. J. Med. Genet. A 123*, 5−28.
3. Aza-Blanc, P., Ramirez-Weber, F. A., Laget, M. P., Schwartz, C., and Kornberg, T. B. (1997) Proteolysis that is inhibited by hedgehog targets cubitus interruptus protein to the nucleus and converts it to a repressor, *Cell 89*, 1043−1053.
4. Kingston, R. E., Bunker, C. A., and Imbalzano, A. N. (1996) Repression and activation by multiprotein complexes that alter chromatin structure, *Genes Dev. 10*, 905−920.
5. de Ruijter, A. J., van Gennip, A. H., Caron, H. N., Kemp, S., and van Kuilenburg, A. B. (2003) Histone deacetylases (HDACs): Characterization of the classical HDAC family, *Biochem. J. 370*, 737−749.
6. Cheng, S. Y., and Bishop, J. M. (2002) Suppressor of fused represses Gli-mediated transcription by recruiting the SAP18−mSin3 corepressor complex, *Proc. Natl. Acad. Sci. U.S.A. 99*, 5442−5447.
7. Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (1997) Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex, *Cell 89*, 357−364.
8. Boehmelt, G., Antonio, L., and Iscove, N. N. (1998) Cloning of the murine transcriptional corepressor component SAP18 and differential expression of its mRNA in the hematopoietic hierarchy, *Gene 207*, 267−275.
9. Zhu, W., and Hanes, S. D. (2000) Identification of *Drosophila* bicoid-interacting proteins using a custom two-hybrid selection, *Gene 245*, 329−339.
10. Espinas, M. L., Canudas, S., Fanti, L., Pimpinelli, S., Casanova, J., and Azorin, F. (2000) The GAGA factor of *Drosophila* interacts with SAP18, a Sin3-associated polypeptide, *EMBO Rep. 1*, 253−259.
11. Singh, N., Zhu, W., and Hanes, S. D. (2005) SAP18 is required for the maternal gene bicoid to direct anterior patterning in *Drosophila melanogaster*, *Dev. Biol. 278*, 242−254.
12. Paces-Fessy, M., Boucher, D., Petit, E., Paute-Briand, S., and Blanchet-Tournier, M. F. (2004) The negative regulator of Gli, suppressor of fused (Sufu), interacts with SAP18, Galectin3 and other nuclear proteins, *Biochem. J. 378*, 353−362.
13. McCallum, S. A., Yin, J., and Fairbrother, W. J. (2005) ^{1}H, ^{13}C, and ^{15}N resonance assignments of SAP18, *J. Biomol. NMR 31*, 259.
14. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, *J. Biomol. NMR 6*, 135−140.
15. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1995) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press: San Diego, CA.
16. Stone, M. J., Fairbrother, W. J., Palmer, A. G., III, Reizer, J., Saier, M. H., Jr., and Wright, P. E. (1992) Backbone dynamics of the *Bacillus subtilis* glucose permease IIA domain determined from ^{15}N NMR relaxation measurements, *Biochemistry 31*, 4394−4406.
17. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation, *Biochemistry 33*, 5984−6003.
18. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, *J. Mol. Biol. 319*, 209−227.
19. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol. Biol. 278*, 353−378.
20. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR 13*, 289−302.
21. Chou, J. J., Gaemers, S., Howder, B., Louis, J. M., and Bax, A. (2001) A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles, *J. Biomol. NMR 21*, 377−382.
22. Ishii, Y., Markus, M. A., and Tycko, R. (2001) Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel, *J. Biomol. NMR 21*, 141−151.
23. McCallum, S. A., and Pardi, A. (2003) Refined solution structure of the iron-responsive element RNA using residual dipolar couplings, *J. Mol. Biol. 326*, 1037−1050.

24. Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude, *J. Magn. Reson. 131*, 159−162.
25. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
26. Merchant, M., Vajdos, F. F., Ultsch, M., Maun, H. R., Wendt, U., Cannon, J., Desmarais, W., Lazarus, R. A., de Vos, A. M., and de Sauvage, F. J. (2004) Suppressor of fused regulates Gli activity through a dual binding mechanism, *Mol. Cell Biol. 24*, 8627−8641.
27. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices, *J. Mol. Biol. 233*, 123−138.
28. Bax, A., and Tjandra, N. (1997) High-resolution heteronuclear NMR of human ubiquitin in an aqueous liquid crystalline medium, *J. Biomol. NMR 10*, 289−292.
29. Stebbins, C. E., Kaelin, W. G., Jr., and Pavletich, N. P. (1999) Structure of the VHL−ElonginC−ElonginB complex: Implications for VHL tumor suppressor function, *Science 284*, 455−461.
30. Kiel, C., and Serrano, L. (2006) The ubiquitin domain superfold: Structure-based sequence alignments and characterization of binding epitopes, *J. Mol. Biol. 355*, 821−844.
31. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) SCOP: A structural classification of proteins database for the investigation of sequences and structures, *J. Mol. Biol. 247*, 536−540.
32. Orengo, C. A., Jones, D. T., and Thornton, J. M. (1994) Protein superfamilies and domain superfolds, *Nature 372*, 631−634.
33. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2256−2268.
34. Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., and Dunker, A. K. (2005) Exploiting heterogeneous sequence properties improves prediction of protein disorder, *Proteins 61 (Supplement 7)*, 176−182.
35. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server, *Bioinformatics 16*, 404−405.
36. Aloy, P., Ceulemans, H., Stark, A., and Russell, R. B. (2003) The relationship between sequence and interaction divergence in proteins, *J. Mol. Biol. 332*, 989−998.
37. Hershko, A., and Ciechanover, A. (1998) The ubiquitin system, *Annu. Rev. Biochem. 67*, 425−479.
38. Matunis, M. J., Coutavas, E., and Blobel, G. (1996) A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex, *J. Cell Biol. 135*, 1457−1470.
39. Hipp, M. S., Kalveram, B., Raasi, S., Groettrup, M., and Schmidtke, G. (2005) FAT10, a ubiquitin-independent signal for proteasomal degradation, *Mol. Cell Biol. 25*, 3483−3491.
40. Freed, E., Lacey, K. R., Huie, P., Lyapina, S. A., Deshaies, R. J., Stearns, T., and Jackson, P. K. (1999) Components of an SCF ubiquitin ligase localize to the centrosome and regulate the centrosome duplication cycle, *Genes Dev. 13*, 2242−2257.
41. Harper, J. W., and Schulman, B. A. (2006) Structural complexity in ubiquitin recognition, *Cell 124*, 1133−1136.
42. Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins, *Annu. Rev. Cell Dev. Biol. 22*, in press.
43. Tsukita, S., and Yonemura, S. (1999) Cortical actin organization: Lessons from ERM (ezrin/radixin/moesin) proteins, *J. Biol. Chem. 274*, 34507−34510.
44. Bretscher, A., Edwards, K., and Fehon, R. G. (2002) ERM proteins and merlin: Integrators at the cell cortex, *Nat. Rev. Mol. Cell Biol. 3*, 586−599.
45. Knight, D., Harris, R., McAlister, M. S., Phelan, J. P., Geddes, S., Moss, S. J., Driscoll, P. C., and Keep, N. H. (2002) The X-ray crystal structure and putative ligand-derived peptide binding properties of γ-aminobutyric acid receptor type A receptor-associated protein, *J. Biol. Chem. 277*, 5556−5561.
46. Song, J., Zhang, Z., Hu, W., and Chen, Y. (2005) Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: A reversal of the bound orientation, *J. Biol. Chem. 280*, 40122−40129.
47. Wilson, M. I., Gill, D. J., Perisic, O., Quinn, M. T., and Williams, R. L. (2003) PB1 domain-mediated heterodimerization in NADPH oxidase and signaling complexes of atypical protein kinase C with Par6 and p62, *Mol. Cell. 12*, 39−50.
48. Yoshinaga, S., Kohjima, M., Ogura, K., Yokochi, M., Takeya, R., Ito, T., Sumimoto, H., and Inagaki, F. (2003) The PB1 domain and the PC motif-containing region are structurally similar protein binding modules, *EMBO J. 22*, 4888−4897.
49. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005) ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures, *Nucleic Acids Res. 33*, W299−W302.
50. Lichtarge, O., and Sowa, M. E. (2002) Evolutionary predictions of binding surfaces and interactions, *Curr. Opin. Struct. Biol. 12*, 21−27.
51. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
52. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-coffee: A novel method for fast and accurate multiple sequence alignment, *J. Mol. Biol. 302*, 205−217.
53. Chupreta, S., Holmstrom, S., Subramanian, L., and Iniguez-Lluhi, J. A. (2005) A small conserved surface in SUMO is the critical structural determinant of its transcriptional inhibitory properties, *Mol. Cell Biol. 25*, 4272−4282.
54. Duda, D. M., and Schulman, B. A. (2005) Tag-team SUMO wrestling, *Mol. Cell. 18*, 612−614.
55. Remaut, H., and Waksman, G. (2006) Protein−protein interaction through β-strand addition, *Trends Biochem. Sci. 31*, 436−444.
56. Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M., and Hicke, L. (2001) Distinct functional surface regions on ubiquitin, *J. Biol. Chem. 276*, 30483−30489.
57. Wohlgemuth, S., Kiel, C., Kramer, A., Serrano, L., Wittinghofer, F., and Herrmann, C. (2005) Recognizing and defining true Ras binding domains I: Biochemical analysis, *J. Mol. Biol. 348*, 741−758.
58. Wang, L., Ding, L., Jones, C. A., and Jones, R. S. (2002) *Drosophila* Enhancer of zeste protein interacts with dSAP18, *Gene 285*, 119−125.
59. Wada, M., Miyazawa, H., Wang, R. S., Mizuno, T., Sato, A., Asashima, M., and Hanaoka, F. (2002) The second largest subunit of mouse DNA polymerase ϵ, DPE2, interacts with SAP18 and recruits the Sin3 co-repressor protein to DNA, *J. Biochem. 131*, 307−311.
60. Schwerk, C., Prasad, J., Degenhardt, K., Erdjument-Bromage, H., White, E., Tempst, P., Kidd, V. J., Manley, J. L., Lahti, J. M., and Reinberg, D. (2003) ASAP, a novel protein complex involved in RNA processing and apoptosis, *Mol. Cell Biol. 23*, 2981−2990.
61. Tange, T. O., Shibuya, T., Jurica, M. S., and Moore, M. J. (2005) Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core, *RNA 11*, 1869−1883.
62. Letunic, I., Copley, R. R., Pils, B., Pinkert, S., Schultz, J., and Bork, P. (2006) SMART 5: Domains in the context of genomes and networks, *Nucleic Acids Res. 34*, D257−D260.
63. Neduva, V., and Russell, R. B. (2006) DILIMOT: Discovery of linear motifs in proteins, *Nucleic Acids Res. 34*, W350−W355.
64. Ohh, M., Takagi, Y., Aso, T., Stebbins, C. E., Pavletich, N. P., Zbar, B., Conaway, R. C., Conaway, J. W., and Kaelin, W. G., Jr. (1999) Synthetic peptides define critical contacts between elongin C, elongin B, and the von Hippel−Lindau protein, *J. Clin. Invest. 104*, 1583−1591.
65. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.
66. Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003) ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information, *Bioinformatics 19*, 163−164.
67. Pupko, T., Bell, R. E., Mayrose, I., Glaser, F., and Ben-Tal, N. (2002) Rate4Site: An algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues, *Bioinformatics 18 (Supplement 1)*, S71−S77.
68. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.

BI060687L