Published in final edited form as:
Biochemistry. 2014 April 8; 53(13): 2091–2100. doi:10.1021/bi401436h.

The *Arabidopsis* Histone Methyltransferase SUVR4 Binds Ubiquitin via a Domain with a Four-Helix Bundle Structure

Mohummad Aminur Rahman[†,‖], **Per E. Kristiansen**[†], **Silje V. Veiseth**[†], **Jan Terje Andersen**[†,‡], **Kyoko L. Yap**[§], **Ming-Ming Zhou**[§], **Inger Sandlie**[†,‡], **Tage Thorstensen**[†,⊥], and **Reidunn B. Aalen**[*,†]

[†]Department of Biosciences, University of Oslo, N-0316 Oslo, Norway

[‡]Centre for Immune Regulation (CIR), Department of Immunology, Oslo University Hospital Rikshospitalet, and Department of Biosciences, University of Oslo, N-0316 Oslo, Norway

[§]Department of Structural and Chemical Biology, Icahn School of Medicine at Mount Sinai, New York, New York 10029, United States

Abstract

In eukaryotes, different chromatin states facilitate or repress gene expression and restrict the activity of transposable elements. Post-translational modifications (PTMs) of amino acid residues on the N-terminal tails of histones are suggested to define such states. The histone lysine methyltransferase (HKMTase) SU(VAR)3-9 RELATED4 (SUVR4) of *Arabidopsis thaliana* functions as a repressor of transposon activity. Binding of ubiquitin by the WIYLD domain facilitates the addition of two methyl groups to monomethylated lysine 9 of histone H3. By using nuclear magnetic resonance (NMR) spectroscopy, we identified SUVR4 WIYLD (S4WIYLD) as a domain with a four-helix bundle structure, in contrast to three-helix bundles of other ubiquitin binding domains. NMR titration analyses showed that residues of helix α1 (Q38, L39, and D40) and helix α4 (N68, T70, A71, V73, D74, I76, S78, and E82) of S4WIYLD and residues between the first and second β-strands (T9 and G10) and on β-strands 3 (R42, G47, K48, and Q49) and 4 (H68, R72, and L73) undergo significant chemical shift changes when the two proteins interact. A model of the complex, generated using HADDOCK, suggests that the N-terminal and C-terminal parts of S4WIYLD constitute a surface that interacts with charged residues close to the hydrophobic patch of ubiquitin. The WIYLD domains of the closely related SUVR1 and SUVR2 *Arabidopsis* proteins also bind ubiquitin, indicating that this is a general feature of this domain.

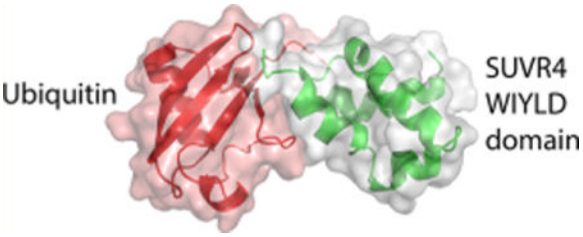
[*]**Corresponding Author:** Telephone: +47 22854437. Fax: +47 22856041. reidunn.aalen@ibv.uio.no.
[‖]**Present Addresses:** MAR.: Oncomatrix Research Laboratory, Department of Biomedicine, University of Bergen, N-5009 Bergen, Norway.
[⊥]**Present Addresses:** T.T.: Bioforsk, Norwegian Institute for Agricultural and Environmental Research, N-1430 Ås, Norway.

Supporting Information
CD spectrum of S4WIYLD (Figure S1), ^{1}H–^{15}N NOE effects (Figure S2), and HSQC spectrum of A32E-WIYLD (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

Author Contributions
M.A.R and P.E.K. contributed equally to this work.

Notes
The authors declare no competing financial interest.

The question of whether SUVR proteins act as both readers of monoubiquitinated H2B and writers of histone PTMs is discussed.



Transcription of eukaryotic genes is dependent on not only transcription factors and RNA polymerase but also the state of the chromatin. The basic unit of chromatin is the nucleosome where the DNA double helix is wound twice around an octamer of histone proteins with two copies each of histones H2A, H2B, H3, and H4. Amino acid (aa) residues on the N-terminal histone tails, which protrude from the nucleosomes, can be modified in a number of ways, including acetylation, phosphorylation, ubiquitination, and methylation. Different combinations of such modifications are found in the loosely structured gene-rich euchromatin compared to the tightly packed gene-poor but repeat-rich heterochromatin.[1,2] Of particular interest are mono-, di-, and trimethylation of different lysine residues that contribute to heterochromatinization and to open or compact states that either facilitate or repress transcriptional activity in euchromatin. Aside from DOT1 that modifies histone H3 lysine 79 (K79),[3] histone lysine methyltransferases (HKMTases) are enzymes with a catalytic SET domain. The genome of the model plant *Arabidopsis thaliana* encodes nearly 40 different SET domain proteins that can be assigned to evolutionarily conserved classes found in both animals and plants.[4,5] These proteins serve different functions depending on the position of the lysine residue they modify (their substrate specificity) and the number of methyl groups they add (their product specificity).

To elucidate how post-translational modifications (PTMs) are established and function in the epigenetic shaping of chromatin states, it is crucial to understand the substrate and product specificities of HKMTases at an atomic level, their crosstalk with other PTMs, and their targeting mechanisms. In animals, proteins similar to SU(VAR)3–9 confer trimethylation of histone H3 lysine 9 (H3K9me3) and are involved in heterochromatinization and suppression of the activity of transposable elements (TEs).[1,6] In *Arabidopsis*, heterochromatin is associated with H3K9me1/me2 and is controlled by SU(VAR)3–9 HOMOLOG (SUVH) proteins.[7–9]

In addition to 10 *SUVH* genes, *Arabidopsis* has five *SU(VAR)3–9 RELATED* (*SUVR1–SUVR5*) genes.[4] SUVR proteins differ from SUVH proteins in their domain structure, and SUVR1, SUVR2, and SUVR4 contain the plant-specific WIYLD domain of ~60 aa, named after its most conserved residues.[10] This domain was recently shown to interact with both free ubiquitin and monoubiquitinated histone H2B (H2Bub1) *in vitro*.[11] Interestingly, the presence of free ubiquitin enhanced the enzymatic HKMTase activity of SUVR4 in a WIYLD-dependent manner and shifted the SUVR4 product specificity from di- to

trimethylation *in vitro*. *In vivo*, SUVR4 was shown to direct H3K9me3 to transposable elements marked with H3K9me1 and to suppress their activity.[11]

SUVR4 is the first HKMTase identified in plants conferring H3K9me3 activity.[11] Aside from transposons, H3K9me3 is prevalent in euchromatin in *Arabidopsis*.[2] Together with H3K4me3, H3K36me3, and H2Bub1, this PTM constitutes chromatin state 1 (CS1) in *Arabidopsis*, representative of transcribed genes. The latter three marks are also associated with actively transcribed genes in animals.[12,13] The establishment of combinations of histone marks designating specific chromatin states requires crosstalk between the histone tail-modifying enzymes. Ubiquitin is a small protein of 76 aa, but large compared to other histone PTMs, and no reader of H2Bub1 has been identified to date.[12,14] Ubiquitin has a diversity of functions; for instance, proteins covalently connected to chains of ubiquitin linked via lysine 48 (Lys48) can be targeted for proteasomal degradation, while proteins with Lys63-linked chains are involved in cell signaling, and monoubiquitinated proteins may be involved in regulation of chromatin remodeling and transcription.[15] Both free ubiquitin and ubiquitinated proteins can be recognized by ubiquitin binding domains (UBDs) found in a number of proteins.[16–19] In general, ubiquitination facilitates conformational changes and creates modified surfaces that can be recognized and promote noncovalent protein–protein interactions with UBDs.[17] A hydrophobic patch formed by the β-sheet of ubiquitin is usually involved in such interactions, while it is the C-terminus of ubiquitin that is linked to other proteins in ubiquitin conjugates.[17,18,20,21] UBDs that recognize conjugated ubiquitin normally interact with the C-terminus.[18]

To improve our understanding of how ubiquitin and H2Bub1 influence SUVR4 function, we have determined the three-dimensional (3D) structure of the SUVR4 WIYLD (S4WIYLD) domain by using NMR spectroscopy. We present a model for the S4WIYLD–ubiquitin interaction based on the solution structure, mutational analysis, and chemical shift perturbations (CSPs) upon binding.

EXPERIMENTAL PROCEDURES

Yeast Two-Hybrid Experiments

Yeast two-hybrid (Y2H) interaction studies were performed as described in ref 11. The gene fragments encoding SUVR1, SUVR2, and SUVR4 WIYLD domains (S1WIYLD, S2WIYLD, and S4WIYLD, respectively) in pDONR/Zeo were recombined into pGBKT7-GW,[22] while UBIQUITIN EXTENSION PROTEIN1 (UBQ1), ubiquitin, and L40 sequences were recombined into pGADT7 (Clontech) using Gateway technology. Direct interactions were tested by mating pGADT7-UBQ1, pGADT7-ubiquitin, and pGADT7-L40 with pGBKT7-S1WIYLD and pGBKT7-S2WIYLD using pGBKT7-S4WIYLD and the empty pGBKT7 vector as positive and negative controls. Diploid colonies were selected on SD −L/−T and then streaked out on SD −L/−T/−H +3 AT medium selective for protein–protein interactions.

Generation of Constructs for Expression of Recombinant Proteins

S4WIYLD with mutations in the codons for aa 32, 37, 61, and 74 was generated with the QuikChange II Site-Directed Mutagenesis Kit (Stratagene).[11] To create glutathione *S*-transferase (GST) fusion proteins for an enzyme-linked immunosorbent assay (ELISA), wild-type and mutant S4WIYLD cDNA fragments were recombined into pGEX-AB-GW.[10] Ubiquitin cDNA was recombined into pHMGWA[23] to generate maltose binding protein (MBP)-tagged ubiquitin. The protein used for structural analysis of S4WIYLD contained an N-terminal His tag followed by an enzyme cleavage site to make it suitable for expression and purification. The numbering of aa residues in the S4WIYLD domain refers to their position in the full-length SUVR4 protein (Figure 1A).

Circular Dichroism (CD)

CD spectra were recorded using a Jasco J-810 spectropolarimeter (Jasco International Co.). All measurements were taken using a quartz cuvette (Starna) with a 0.1 cm path length. Samples were scanned five times at a rate of 50 nm/min with a bandwidth of 0.5 nm and a response time of 2 s over the wavelength range of 190–260 nm. The data were averaged, and the spectrum of a sample-free control was subtracted. The α-helical content of the peptides was determined by application of the spectral fitting methods CONTIN/LL,[24] CDSSTR,[25] and SELCON3,[26] in the CDpro package,[27] and the single-point method using the mean residual ellipticity at 222 nm ($[\theta]_{222}$) and the equation

$$f_{\mathrm{H}}=100\% \times [\theta]_{222}/[-40000(1-2.5n)] \quad (1)$$

where f_{H} is the α-helical content in percent and n the number of residues.[28] Two or more repetitions of each measurement were done.

Enzyme-Linked Immunosorbent Assay (ELISA)

Wells were coated with 100 μL of MBP-ubiquitin in 1× PBS (10 μg/mL) overnight at 4 °C and then blocked with 4% skim milk (skm; Acumedia) for 1 h at room temperature. The wells were then washed four times with a PBS/0.005% Tween 20 mixture (PBS/T) (pH 7.4) before 100 μL of titrated amounts of GST-fused SUVR proteins was added to the wells (50.0–0.3 μg/mL) in PBS/T. GST alone was used as a negative control. After incubation for 2 h at room temperature, wells were washed as described above. Bound SUVR proteins were detected by adding 100 μL of an HRP-conjugated polyclonal anti-GST antibody from goat (1:5000; GE Healthcare), diluted in a 4% skm/PBS/T mixture and incubated for 1 h at room temperature. Following the washes described above, 100 μL of the substrate TMB (Calbiochem) was added to each well and incubated for 45–60 min before 100 μL of 0.25 M HCl was added. The absorbance was measured at 450 nm using a Sunrise TECAN spectrophotometer (TECAN, Maennedorf, Switzerland). Binding of titrated amounts (12.5–0.03 μg/mL) of a His-tagged variant of S4WIYLD protein was measured essentially as described above except that binding was detected using an HRP-conjugated anti-His antibody (Serotec).

NMR Sample Preparation

^{13}C- and ^{15}N-labeled S4WIYLD protein was expressed in *Escherichia coli* BL21-Star DE3 cells and purified to at least 90% purity, estimated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis, and prepared for NMR essentially as previously described.[29] The cells were grown to an OD_{600} of 0.7 in unlabeled rich LB medium at 37 °C, harvested by centrifugation, washed, and finally transferred into ^{15}N- and ^{13}C-labeled M9 defined medium [using [^{15}N]NH_4Cl and [^{13}C]glucose (Sigma)]. Cells from 1 L of LB medium were transferred into 500 mL of M9 medium and incubated for 1 h to adapt the cells to the new medium and deplete the unlabeled metabolites. Protein expression was then induced by addition of 1 mM IPTG. After overnight incubation, the cells were harvested by centrifugation, resuspended in Tris buffer (lysis buffer) [50 mM Tris-HCl (pH 7.5), 400 mM NaCl, 100 mM KCl, 1 mM DTT, 1 mM EDTA, 0.5% Triton X-100, 20 mM imidazole, and protease inhibitor], and lysed using Ex-press. After centrifugation (14000 rpm), the supernatant containing recombinant protein was filtered through 0.45 μm filters and prepared for affinity purification. His-S4WIYLD fusion protein was affinity-purified by Ni-NTA affinity chromatography using a HisTrap FF 5 mL (GE Healthcare) column in the ÄKTA purifier essentially as described in ref 11, yielding ~2–5 mg (GST-tagged proteins) and ~8–10 mg (His- and MBP-tagged proteins) per liter of bacterial culture. The proteins were concentrated to 10.4 mg/mL using Amicon 10000 NMWL centrifugal concentrators. NMR experiments were performed with samples containing 0.6–0.8 mM S4WIYLD in 20 mM sodium phosphate (pH 6.5), 50 mM NaCl, 1 mM DTT, and 0.2 mM DSS in a 95:5 H_2O/D_2O mixture. Labeled and nonlabeled ubiquitin was purchased from ASLA biotech.

NMR Spectroscopy

The following NMR experiments were performed to assign the backbone chemical shifts and determine structural restraints: ^{15}N HSQC, ^{13}C HSQC, ^{15}N NOESY-HSQC, ^{13}C NOESY-HSQC, HNHA, HNCA, CBCAcoNH, CBCANH, HNCO, HNcaCO, HNcoCA, (H)-CCcoNH, HcccoNH, HBHAcoNH, HBHANH, and HCCH-TOCSY.[30,31] ^{1}H–^{15}N NOESY with a mixing time of 3 s was performed to investigate the relaxation properties of the protein.[32] All experiments were conducted on a 600 MHz Bruker Avance II spectrometer with four channels and a 5 mm TCI cryo probe at 25 °C. ^{1}H–^{15}N HSQC experiments were also conducted at 10, 15, 20, 25, and 30 °C.

Data were processed using Topspin version 2.1 (Bruker Biospin). DSS was used as a chemical shift standard, and ^{13}C and ^{15}N data were referenced using frequency ratios as described previously.[33] For visualization and assignment, CARA[34] was used.

The ^{15}N and ^{13}C NOESY-HSQC spectra were manually peak picked using SPARKY.[35] NOESY upper distance constraints were generated by the CANDID routine in CYANA version 3.0.[36] Torsion angle constraints were determined from the chemical shifts by the application of TALOS+;[37] 100 structures were calculated using CYANA version 3.0, and the 20 structures with the lowest energy were kept in the structural ensemble. The structures have been deposited in the Protein Data Bank (PDB) as entry 2lxe and the NMR data in the Biomagnetic Resonance Bank (BMRB) as entry 18672.

Unlabeled ubiquitin was added to ^{15}N-labeled S4WIYLD domain protein stepwise from a 1:1 ratio to a final molar ratio of 5:1. Similarly, unlabeled S4WIYLD domain protein was added stepwise to ^{15}N-labeled ubiquitin from a ½:1 ratio to a 5:1 ratio. The ^{1}H and ^{15}N chemical shifts changes in the S4WIYLD domain were monitored via the two-dimensional ^{1}H–^{15}N HSQC spectrum. The changes of amide ^{1}H and ^{15}N chemical shifts were averaged by using the formula[38]

$$\Delta\delta_{\mathrm{av}}=\left\{\frac{1}{2}\left[(\Delta\delta_{\mathrm{H}})^2+(0.2\Delta\delta_{\mathrm{N}})^2\right]\right\}^{1/2} \quad (2)$$

The dissociation constant (K_d) was determined from the changes in chemical shifts upon addition of S4WIYLD to ubiquitin by applying the formula[39]

$$\Delta\delta_{\mathrm{av}}=\Delta\delta_{\mathrm{av,\ max}}\left[K_{\mathrm{d}}+[\mathrm{L}]_0+[\mathrm{P}]_0-\sqrt{(K_{\mathrm{d}}+[\mathrm{L}]_0+[\mathrm{P}]_0)^2-4[\mathrm{L}]_0[\mathrm{P}]_0}\right]/2[\mathrm{P}] \quad (3)$$

where $[P]_0$ and $[L]_0$ are the concentrations of the protein and ligand, respectively, and $\Delta\delta_{\mathrm{av,max}}$ is the maximal change in chemical shift that can be realized upon addition of ligand to protein. K_d and $\Delta\delta_{\mathrm{av,max}}$ were free parameters during the fit to the experimental data.

Docking for the S4WIYLD–Ubiquitin Complex

The NMR structure of the S4WIYLD domain and the crystal structure of ubiquitin (PDB entry 1UBQ[40]) were used to model the S4WIYLD–ubiquitin complex employing the HADDOCK server.[41] From the PDB files, the solvent-exposed residues were identified by PYMOL. The residues that had CSP changes above the threshold were filtered according to their water exposure and chosen as active residues. The following active residues were used: Q38, D40, N68, T70, D74, Y77, S78, D81, E82, and N83 for S4WIYLD and T9, G10, R42, G47, K48, Q49, H68, R72, and L73 for ubiquitin. In addition, the following solvent-exposed residues, R37, L66, D67, V79, and E80 of S4WIYLD and K6, L8, K11, T12, A46, L71, R74, G75, and G76 for ubiquitin, close to the active residues, were used as passive residues for docking. The interface residues that were allowed to move during the docking and water refinement for S4WIYLD were R37–K45 and D67–N83, with full flexibility for the C-terminus (E80–N83), and for ubiquitin F44–D52 and S65–G76, with full flexibility for H68–G76, similar to the docking of p47 to ubiquitin.[42]

RESULTS

WIYLD Domains Bind Ubiquitin

The ubiquitin binding property of S4WIYLD was discovered through a yeast two-hybrid (Y2H) screen where it was found to interact with the N-terminal ubiquitin moiety of UBIQUITIN EXTENSION PROTEIN1 (UBQ1), but not the C-terminal ribosomal protein L40 part of UBQ1.[11] Different plant species have two or three closely related genes encoding SUVR proteins with a WIYLD domain[5] (Figure 1A,B). To confirm that ubiquitin binding is a common feature of WIYLDs, the corresponding domains of the *Arabidopsis* SUVR1 and SUVR2 proteins (S1WIYLD and S2WIYLD, respectively) (Figure 1B) were

tested and shown to interact with ubiquitin in the Y2H assay previously used for S4WIYLD (Figure 1C).

To compare the ubiquitin binding capacity of the SUVR variants, a sensitive ELISA was established, which demonstrated that S1WIYLD (0.65 ± 0.12) and S2WIYLD (1.23 ± 0.13) bound slightly weaker and stronger, respectively, than S4WIYLD (Figure 1D,E). Thus, the Y2H and ELISA experiments confirm that ubiquitin binds specifically and selectively to the SUVR variants.

The SUVR4 WIYLD Domain Consists of a Four-Helix Bundle That Is Structurally Similar to Three-Helix Bundle UBDs

In general, UBD structures responsible for noncovalent ubiquitin binding can be divided into four categories: (1) α-helices, (2) zinc fingers, (3) plekstrin homology (PH) domains, and (4) ubiquitin-conjugating (Ubc)-like domains (reviewed in ref 43). PhDsec and JPRED secondary structure predictions suggest that the WIYLD domain consists of three α-helices.[10] To test this, we obtained and analyzed the CD spectrum of S4WIYLD (Figure S1 of the Supporting Information), which demonstrated that the domain is mainly α-helical with CDSSTR giving the best fit to the experimental data, predicting an α-helical content of 53%, in good agreement with the secondary structure predictions mentioned above.

To determine the 3D solution structure of the S4WIYLD domain, NMR spectroscopy was employed using a His-tagged ^{13}C- and ^{15}N-labeled protein. The protein was assigned using standard methods, and structural restraints were obtained and used to determine the structure as described in Experimental Procedures. The assigned ^{15}N HSQC spectrum is shown in Figure 2, and the short-range interactions typically found in α-helices are shown in Figure 3. The number and type of structural restraints and the structural statistics are listed in Table 1. In agreement with the CD experiments, NMR analysis revealed a predominantly helical structure. The N-terminal part of the protein before residue 21 was found to be unstructured, in agreement with the results of ^{1}H–^{15}N NOE experiments (Figure S2 of the Supporting Information). The structured part of the molecule (residues 21–83) consists of a four-helix bundle (Figure 4A,B), with highly conserved hydrophobic residues in helices α1, α2, and α4 (Figure 1A).

To relate the structure of S4WIYLD to those of other UBDs, we identified UBD proteins with sequences similar to those of WIYLD domains (Figures 1A and 4C–F). These UBDs, e.g., p47, BMSC-UbP, and hHR23A, all have a predicted three-helix structure.[42,44,45] Comparison of 3D structures demonstrates that the overall fold of three helix-bundle UBDs and the S4WIYLD domain is similar, although instead of the varying length of helix α2 and a loop prior to the last helix, there is an extra turn and helix in the S4WIYLD structure (Figure 4). The interhelical angles fall in a relatively small range for the S4WIYLD domain and other UBDs between helices α1 and α2 (127°, 110°, 125°, and 125°) and between helix α2 and the last helix (113°, 113°, 102°, and 107°) in S4WIYLD, p47, BMSC, and hHR23, respectively, whereas the differences in angles between the first and last helices are more diverse (60°, 40°, 43°, and 72° in S4WIYLD, BMSC, p47, and hHR23, respectively).

Several UBDs contain a Met/Leu-Gly-Tyr/Phe motif in turn 1 and an Asn-X-polar motif between helices α2 and α3 that form a conserved solvent-accessible surface with hydrophobic patches involved in ubiquitin binding.[44,45] A conserved Leu-Gly-Ile motif (Leu39-Asp40-Ile41 in S4WIYLD) is similarly present in turn 1 of plant WIYLD domains, and an Asn-Tyr-Polar motif (Asn68-Tyr69-Thr70 in S4WIYLD) is found to the N-terminal side of helix α4 (Figure 1A). The presence of these conserved motifs in corresponding positions in WIYLD and UBDs suggests similar functional significance for the interaction of WIYLD with ubiquitin.

Titration Analysis Implicates Helix α1, Turn 1, Turn 3, and Helix α4 of S4WIYLD in the Binding of Ubiquitin

To substantiate whether the mode of binding to ubiquitin was similar to that of three-helix bundle UBDs, ^{15}N-labeled S4WIYLD was titrated with unlabeled ubiquitin, and likewise, ^{15}N-labeled ubiquitin was titrated with unlabeled S4WIYLD. The changes in chemical shifts upon binding, which signify the local chemical and conformational changes at the backbone amide nitrogen and proton atoms, were monitored using HSQC spectra (Figure 5A–D). Such chemical shift perturbations (CSPs) point to the amino acid residues that are potentially important for the protein–protein interaction.

The combined CSPs for the backbone ^{15}N and ^{1}H resonances of both labeled S4WIYLD and labeled ubiquitin were mapped onto the respective sequences (Figure 5B,C). Residues of S4WIYLD with CSPs greater than or equal to a cutoff value of 0.10 included the conserved motifs mentioned above, i.e., Leu39–Ile41 of turn 1 and C-terminal residues Lys68–Glu82 (Figure 5B). The S4WIYLD solution structure showed that the side chains of Leu39–Ile41 are in the proximity of helix α4 (Figure 5E), suggesting its involvement in ubiquitin binding in cooperation with turn 1, as hypothesized previously.[11]

Analyses of the CSPs of ubiquitin (BMRB entry 5387) indicated that the residues between Thr7 and Ile13, Leu42 and Leu50, and His68 and Gly75, positioned between the first and second β-strands and on β-strands 3 and 4, respectively, are important for protein–protein interaction (Figure 5C,F). These residues are close to but not the same as those observed for other ubiquitin binding domains, for which binding involves the hydrophobic patch that consists of residues Leu8, Ile44, and Val70 of the ubiquitin β-sheet and the surrounding region.[44,47,48] For the binding to S4WIYLD, we have observed more pronounced changes in the chemical shifts of charged and polar residues than for the hydrophobic residues, indicating that the interaction may be more electrostatic in nature than hydrophobic.

The change in chemical shifts upon addition of S4WIYLD to ubiquitin (Figure 5D) was used to determine the dissociation constant (K_d) for the interaction between the two proteins (Experimental Procedures, eq 3). The K_d was calculated to be virtually the same as that reported for CUE2-1 (157 ± 12 and 155 ± 9 μM,[46] respectively) and ~10-fold weaker than those for BMSC-UbP (17 μM) and DSK2 (14.8 μM).[44]

Docking of S4WIYLD to Ubiquitin Together with Mutational Analyses Suggests Binding via Electrostatic Interactions

To obtain a model of the WIYLD–ubiquitin protein complex, the NMR structure of S4WIYLD obtained here and the structure of ubiquitin (PDB entry 1UBQ) were docked using HADDOCK The results from the CSP titration analyses for S4WIYLD and ubiquitin (Figure 5B,C) were employed as described in Experimental Procedures. The top 10 docking results are superimposed in Figure 5G, and the best docking model is shown in Figure 5H.

We have previously generated mutations in four residues highly conserved in all WIYLD domains [Ala32, Arg37, Trp61, and Asp74 (Figure 1A)] and shown that R37A abolished and W61A and D74A inhibited the binding to H2Bub1 in a pulldown assay.[11] We screened the mutants for binding to MBP-tagged ubiquitin in an ELISA. The A32E mutation reduced the binding affinity by 37%, and the W61R mutation resulted in a 75% reduction (Figure 6); however, neither the CSP analyses nor the docking model indicated that Ala32 or Trp61 interacts directly with ubiquitin (Figure 5C,F,H). They may instead be important for the maintenance of the tertiary structure of S4WIYLD. According to our solution structure, Ala32 is located in helix α1 with the side chain pointing into the hydrophobic core of the protein, and Trp61 at the N-terminus of S4WIYLD-unique helix α3, with its prominent aromatic side chain positioned between helices α1 and α2 (Figure 5E). Exchange of either the hydrophobic side chain of Ala32 with a charged residue or the long aromatic side chain of Trp61 with the long positively charged side chain of Arg may change the spatial relation of helices α1–α3. We therefore evaluated the structural integrity of S4WIYLD containing the A32E mutation by NMR. The peaks of the HSQC spectrum of A32E-WIYLD were relatively undispersed, suggesting that the mutation disrupts the entire domain structure and providing an explanation for its reduced affinity for ubiquitin (Figure S3 of the Supporting Information).

In agreement with the CSP analyses, the docking model indicated that S4WIYLD residues at the end of helix α1 (Arg37–Asp40) and most of the C-terminal end (especially Asp74, Tyr77, Glu80, and Glu82 of helix α4) are important for the interaction with ubiquitin (Figure 5H). Mutation of Arg37 to Ala has previously been shown to have a strong negative effect on the binding to H2Bub1;[11] however, Arg37 showed low CSP values (Figure 5B), and the mutation resulted in an only 29% reduction in the capacity of binding to ubiquitin (Figure 6). During binding, Asp40 and Asp74 of S4WIYLD move to interact with Arg72 and Lys48 of ubiquitin, respectively. Mutation of Asp74 of S4WIYLD to Ala weakens the binding ability to ubiquitin by 56% according to an ELISA (Figure 6), consistent with the weakened binding to histone H2Bub1 shown earlier by an *in vitro* pull-down assay.[11] Furthermore, several residues of S4WIYLD (Ser78, Asp81, and Glu82) may interact with His68 of ubiquitin. In addition, Glu82 may bind Lys6, Glu80 ends up close to Arg42, and Tyr77 of S4WIYLD inserts into a pocket on the ubiquitin surface close to Ile44 and may possibly interact with Gln49 of ubiquitin. Taken together, our model indicates that S4WIYLD has the center of interaction close to Arg42 of ubiquitin and suggests that the closely positioned N- and C-terminal parts of S4WIYLD, including helix α1, turn 1, and helix α4, constitute an interacting surface that makes electrostatic interactions with charged residues on the surface of ubiquitin.

DISCUSSION

This work provides a description of the solution structure of a WIYLD domain, which reveals a novel four-helix bundle configuration. To the best of our knowledge, the S4WIYLD domain is the first UBD found to consist of four helices. The extra helix is a hallmark for the WIYLD domains that distinguish them from other UBDs. The ubiquitin binding capacity of this domain, demonstrated for three different SUVR proteins, and the strong sequence conservation of the domain in different plant species argue that all WIYLD domains bind ubiquitin. The likely interaction surface between WIYLD and ubiquitin was identified by CSP, mutagenesis, and docking model analyses. The β-sheet of ubiquitin is facing the S4WIYLD domain, and binding involves the movement of residues surrounding Leu8, Gly47, and Val70 of ubiquitin. The major sites of contact with ubiquitin for three-helix bundle UBDs, like CUE (PDB entry 1OTR) and UBA (PDB entries 2DEN, 2JY6, and 1WR1), are confined to the hydrophobic surface centered at Ile44.[15] In comparison, our results taken together suggest that the binding site of S4WIYLD is shifted slightly away from the hydrophobic patch of the β-sheet. This shift in binding site may result from the large fraction of charged residues involved in the S4WIYLD–ubiquitin interaction.

The C-terminal Gly76 of ubiquitin provides the covalent bridge to a lysine residue of proteins conjugated to ubiquitin,[47] and recognition of the C-terminus of ubiquitin is common for proteins regulating ubiquitin conjugation and deubiquitination.[48–50] SUVR4 binds H2Bub1 *in vitro*, and intriguingly, mutation in Arg37, which according to our docking model is situated close to Gly76, has a major negative effect on H2Bub1–S4WIYLD binding. One may therefore speculate that Arg37 is important for the recognition of ubiquitinated H2B, but less important for the binding of free ubiquitin.

SUVR4 targets transposon chromatin, which has a low level of H2Bub1 and a high level of the preferred HKMTase substrate, H3K9me1. SUVR2 has, like SUVR4, been implicated as a regulator of transposon activity and DNA methylation.[51] In *Arabidopsis*, a high level of H2Bub1 and a moderate level of H3K9me3 are associated with transcriptionally active euchromatic genes.[11,52] Ubiquitination and deubiquitination of histone H2B are implicated in transcriptional elongation as well as repression of transposons in yeast, animals, and plants.[1,2,14] It is not known whether SUVR4 or the other SUVR proteins with a WIYLD domain *in vivo* prefer binding to free ubiquitin or also act as readers of H2Bub1 or interact with ubiquitin conjugated to other proteins, three scenarios that might have different implications for the chromatin-modulating role of SUVR proteins (Figure 7). It remains to be determined whether the euchromatic PTM context is suitable for any of the SUVR proteins and what relevance WIYLD domain binding of H2Bub1 has for the *in vivo* function of SUVR4 or other SUVR proteins. As the SUVR1, SUVR2, and SUVR4 proteins are very similar, double or triple *suvr* mutants may be needed to address this topic.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Roy Falleth, Solveig H. Engebretsen for technical assistance, and Cathrine Jacobsen for initial investigation of the binding of S1WIYLD and S2WIYLD to ubiquitin. We acknowledge the Research Council of Norway for their investment in the 600 MHz NMR spectrometer at the University of Oslo.

Funding

This work was supported by the Research Council of Norway (Grants 146652/431 and 183609/S10 to S.V.V. and R.B.A. and 179573/V40 to J.TA.) and the South-Eastern Norway Regional Health Authority (Grant 39375 to J.T.A).

ABBREVIATIONS

NMR	nuclear magnetic resonance
PTM	post-translational modification
H2Bub1	monoubiquitinated histone H2B
HKMTase	histone lysine methyltransferase
UBQ1	UBIQ-UITIN EXTENSION PROTEIN1
HRP	horseradish peroxidase

References

1. Kouzarides T. Chromatin modifications and their function. Cell. 2007; 128:693–705. [PubMed: 17320507]
2. Roudier F, Ahmed I, Berard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L, Giraut L, Despres B, Drevensek S, Barneche F, Derozier S, Brunaud V, Aubourg S, Schnittger A, Bowler C, Martin-Magniette ML, Robin S, Caboche M, Colot V. Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*. EMBO J. 2011; 30:1928–1938. [PubMed: 21487388]
3. Jones B, Su H, Bhat A, Lei H, Bajko J, Hevi S, Baltus GA, Kadam S, Zhai H, Valdez R, Gonzalo S, Zhang Y, Li E, Chen T. The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. PLoS Genet. 2008; 4:e1000190. [PubMed: 18787701]
4. Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumann K, Assalkhou R, Schulz I, Reuter G, Aalen RB. The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucleic Acids Res. 2001; 29:4319–4333. [PubMed: 11691919]
5. Thorstensen T, Grini PE, Aalen RB. SET domain proteins in plant development. Biochim Biophys Acta. 2011; 1809:407–420. [PubMed: 21664308]
6. Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, Reuter G. Central role of *Drosophila* SU(VAR)3–9 in histone H3–K9 methylation and heterochromatic gene silencing. EMBO J. 2002; 21:1121–1131. [PubMed: 11867540]
7. Ebbs ML, Bender J. Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. Plant Cell. 2006; 18:1166–1176. [PubMed: 16582009]
8. Naumann K, Fischer A, Hofmann I, Krauss V, Phalke S, Irmler K, Hause G, Aurich AC, Dorn R, Jenuwein T, Reuter G. Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. EMBO J. 2005; 24:1418–1429. [PubMed: 15775980]
9. Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature. 2002; 416:556–560. [PubMed: 11898023]
10. Thorstensen T, Fischer A, Grini PE, Sandvik SV, Johnsen SS, Reuter G, Aalen RB. The *Arabidopsis* SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. Nucleic Acids Res. 2006; 34:5461–5470. [PubMed: 17020925]

11. Veiseth SV, Rahman MA, Yap KL, Fischer A, Egge-Jacobsen W, Reuter G, Zhou MM, Aalen RB, Thorstensen T. The SUVR4 histone lysine methyltransferase binds ubiquitin and converts H3K9me1 to H3K9me3 on transposon chromatin in *Arabidopsis*. PLoS Genet. 2011; 7:e1001325. [PubMed: 21423664]
12. Morishita M, di Luccio E. Structural insights into the regulation and the recognition of histone marks by the SET domain of NSD1. Biochem Biophys Res Commun. 2011; 412:214–219. [PubMed: 21806967]
13. Foreman KW, Brown M, Park F, Emtage S, Harriss J, Das C, Zhu L, Crew A, Arnold L, Shaaban S, Tucker P. Structural and functional profiling of the human histone methyltransferase SMYD3. PLoS One. 2011; 6:e22290. [PubMed: 21779408]
14. Lee JS, Smith E, Shilatifard A. The language of histone crosstalk. Cell. 2010; 142:682–685. [PubMed: 20813257]
15. Kulathu Y, Komander D. Atypical ubiquitylation: The unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. Nat Rev Mol Cell Biol. 2012; 13:508–523. [PubMed: 22820888]
16. Hicke L, Schubert HL, Hill CP. Ubiquitin-binding domains. Nat Rev Mol Cell Biol. 2005; 6:610–621. [PubMed: 16064137]
17. Garner TP, Strachan J, Shedden EC, Long JE, Cavey JR, Shaw B, Layfield R, Searle MS. Independent interactions of ubiquitin-binding domains in a ubiquitin-mediated ternary complex. Biochemistry. 2011; 50:9076–9087. [PubMed: 21923101]
18. Dikic I, Wakatsuki S, Walters KJ. Ubiquitin-binding domains: From structures to functions. Nat Rev Mol Cell Biol. 2009; 10:659–671. [PubMed: 19773779]
19. Hammond-Martel I, Yu H, Affar EB. Roles of ubiquitin signaling in transcription regulation. Cell Signalling. 2011; 24:410–421. [PubMed: 22033037]
20. Komander D. The emerging complexity of protein ubiquitination. Biochem Soc Trans. 2009; 37:937–953. [PubMed: 19754430]
21. Seet BT, Dikic I, Zhou MM, Pawson T. Reading protein modifications with interaction domains. Nat Rev Mol Cell Biol. 2006; 7:473–483. [PubMed: 16829979]
22. Thorstensen T, Grini PE, Mercy IS, Alm V, Erdahl S, Aasland R, Aalen RB. The *Arabidopsis* SET domain protein ASHR3 is involved in stamen development and interacts with the bHLH transcription factor ABORTED MICROSPORES (AMS). Plant Mol Biol. 2008; 66:47–59. [PubMed: 17978851]
23. Busso D, Delagoutte-Busso B, Moras D. Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. Anal Biochem. 2005; 343:313–321. [PubMed: 15993367]
24. Sreerama N, Woody RW. Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods. J Mol Biol. 1994; 242:497–507. [PubMed: 7932706]
25. Johnson WC. Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins. 1999; 35:307–312. [PubMed: 10328265]
26. Sreerama N, Woody RW. A self-consistent method for the analysis of protein secondary structure from circular dichroism. Anal Biochem. 1993; 209:32–44. [PubMed: 8465960]
27. Sreerama N, Woody RW. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem. 2000; 287:252–260. [PubMed: 11112271]
28. Scholtz JM, Qian H, York EJ, Stewart JM, Baldwin RL. Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water. Biopolymers. 1991; 31:1463–1470. [PubMed: 1814498]
29. Rogne P, Fimland G, Nissen-Meyer J, Kristiansen PE. Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G. Biochim Biophys Acta. 2008; 1784:543–554. [PubMed: 18187052]
30. Davis AL, Keeler J, Laue ED, Moskau D. Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. J Magn Reson. 1992; 98:207–216.

31. Sattlera M, Schleucherb J, Griesingerc C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog Nucl Magn Reson Spectrosc. 1999:93–158.
32. Renner C, Schleicher M, Moroder L, Holak TA. Practical aspects of the 2D ^{15}N-[^{1}H]-NOE experiment. J Biomol NMR. 2002; 23:23–33. [PubMed: 12061715]
33. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. J Biomol NMR. 1995; 6:135–140. [PubMed: 8589602]
34. Keller, RLJ. Thèse de doctorat thesis. ETH Zürich; Zürich: 2004. Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment.
35. Goddard, TD.; Kneller, DG. SPARKY 3. University of California; San Francisco: 2008.
36. Herrmann T, Guntert P, Wuthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol. 2002; 319:209–227. [PubMed: 12051947]
37. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR. 1999; 13:289–302. [PubMed: 10212987]
38. Balasubramanian S, Sureshkumar S, Lempe J, Weigel D. Potent induction of *Arabidopsis thaliana* flowering by elevated growth temperature. PLoS Genet. 2006; 2:e106. [PubMed: 16839183]
39. Fielding L. NMR methods for the determination of protein–ligand dissociation constants. Prog Nucl Magn Reson Spectrosc. 2007; 51:219–242.
40. Vijay-Kumar S, Bugg CE, Cook WJ. Structure of ubiquitin refined at 1.8 Å resolution. J Mol Biol. 1987; 194:531–544. [PubMed: 3041007]
41. de Vries SJ, van Dijk M, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. Nat Protoc. 2010; 5:883–897. [PubMed: 20431534]
42. Yuan X, Simpson P, McKeown C, Kondo H, Uchiyama K, Wallis R, Dreveny I, Keetch C, Zhang X, Robinson C, Freemont P, Matthews S. Structure, dynamics and interactions of p47, a major adaptor of the AAA ATPase, p97. EMBO J. 2004; 23:1463–1473. [PubMed: 15029246]
43. Hurley JH, Lee S, Prag G. Ubiquitin-binding domains. Biochem J. 2006; 399:361–372. [PubMed: 17034365]
44. Chang YG, Song AX, Gao YG, Shi YH, Lin XJ, Cao XT, Lin DH, Hu HY. Solution structure of the ubiquitin-associated domain of human BMSC-UbP and its complex with ubiquitin. Protein Sci. 2006; 15:1248–1259. [PubMed: 16731964]
45. Mueller TD, Feigon J. Solution structures of UBA domains reveal a conserved hydrophobic surface for proteinprotein interactions. J Mol Biol. 2002; 319:1243–1255. [PubMed: 12079361]
46. Kang RS, Daniels CM, Francis SA, Shih SC, Salerno WJ, Hicke L, Radhakrishnan I. Solution structure of a CUE-ubiquitin complex reveals a conserved mode of ubiquitin binding. Cell. 2003; 113:621–630. [PubMed: 12787503]
47. Thorne AW, Sautiere P, Briand G, Crane-Robinson C. The structure of ubiquitinated histone H2B. EMBO J. 1987; 6:1005–1010. [PubMed: 3036486]
48. Winget JM, Mayor T. The diversity of ubiquitin recognition: Hot spots and varied specificity. Mol Cell. 2010; 38:627–635. [PubMed: 20541996]
49. Reyes-Turcu FE, Horton JR, Mullally JE, Heroux A, Cheng X, Wilkinson KD. The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. Cell. 2006; 124:1197–1208. [PubMed: 16564012]
50. Drag M, Mikolajczyk J, Bekes M, Reyes-turcu FE, Ellman JA, Wilkinson KD, Salvesen GS. Positional-scanning fluorigenic substrate libraries reveal unexpected specificity determinants of DUBs (deubiquitinating enzymes). Biochem J. 2008; 415:367–375. [PubMed: 18601651]
51. Stroud H, Greenberg MVC, Feng S, Bernatavichute YV, Jacobsen SE. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. Cell. 2013; 152:352–364. [PubMed: 23313553]
52. Sridhar VV, Kapoor A, Zhang KL, Zhu JJ, Zhou T, Hasegawa PM, Bressan RA, Zhu JK. Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination. Nature. 2007; 447:735–738. [PubMed: 17554311]

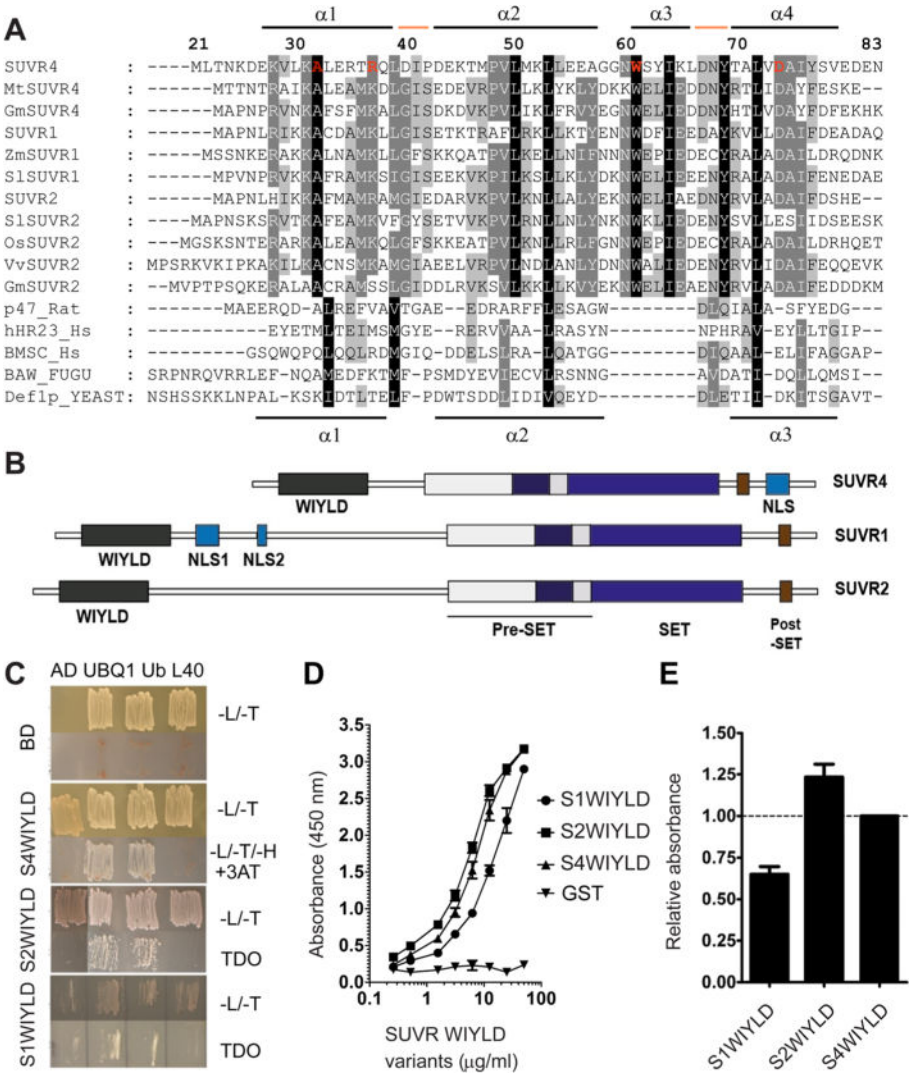


Figure 1.

WIYLD domains bind ubiquitin. (A) Multiple-sequence alignment of plant WIYLD domains and three-helix bundle ubiquitin binding domains from yeast, human (Hs), rat, and fish (Fugu) with conserved amino acids. Numbering refers to residues 21–83 of S4WIYLD. Highly conserved residues are highlighted in black and gray. The secondary structure is shown for WIYLD above and for the other UBDs below the alignment. Mutated amino acids are colored red. The positions of two short conserved motifs found in many UBDs are indicated by red lines. Abbreviations: Mt, *Medicago truncatula; Gm, Glycine max;* Zm, *Zea mais;* Sl, *Solanum lycopersicum;* Os, *Oryza sativa indica;* Vv, *Vitis vinifera*. (B) The domains of the *Arabidopsis* SUVR1, SUVR2, and SUVR4 proteins include the conserved SET domain with pre-SET and post-SET motifs and a conserved WIYLD domain at the N-terminus. Nuclear localization signals (NLSs) have been identified in SUVR1 and SUVR4. (C) Yeast two-hybrid interaction test between WIYLD domains from SUVR1, SUVR2, and SUVR4 (S1WIYLD, S2WIYLD, and S4WIYLD, respectively) and full-length UBQ1, as well as the N-terminal ubiquitin (Ub) and C-terminal ribosomal L40 moieties of UBQ1. −L/−T, medium selective for diploid colonies. −L/−T/−H +3 AT and TDO, medium selective

for protein–protein interactions. AD, control mating with an empty prey vector. BD, control mating with an empty bait vector. (D) Titrated amounts of GST-tagged WIYLD domains were added to ubiquitin coated in ELISA wells. GST alone was included as a negative control. Bound GST and GST–WIYLD fusion proteins were detected using an HRP-conjugated anti-GST antibody. The data are presented on a logarithmic scale. The results are representative of three independent experiments ($N = 3$). (E) Quantification of the ELISA results for S1WIYLD and S2WIYLD relative to S4WIYLD, which was set to 1.

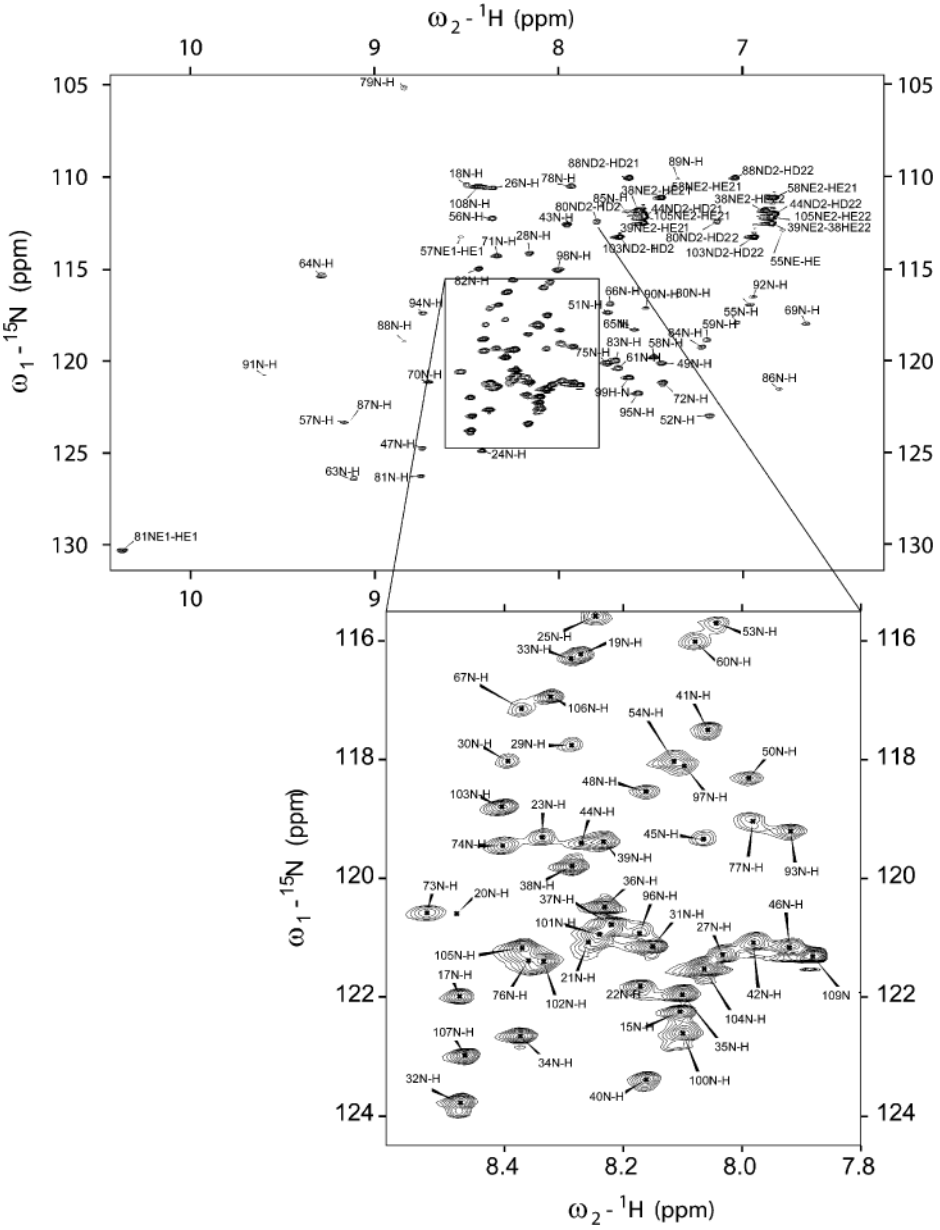


Figure 2.

^{15}N HSQC spectrum of S4WIYLD obtained by NMR spectroscopy: (top) assignment of the full spectrum and (bottom) expansion of the central region. Assignment labels are with regard to the full sequence, including the N-terminal tag of 20 amino acids.

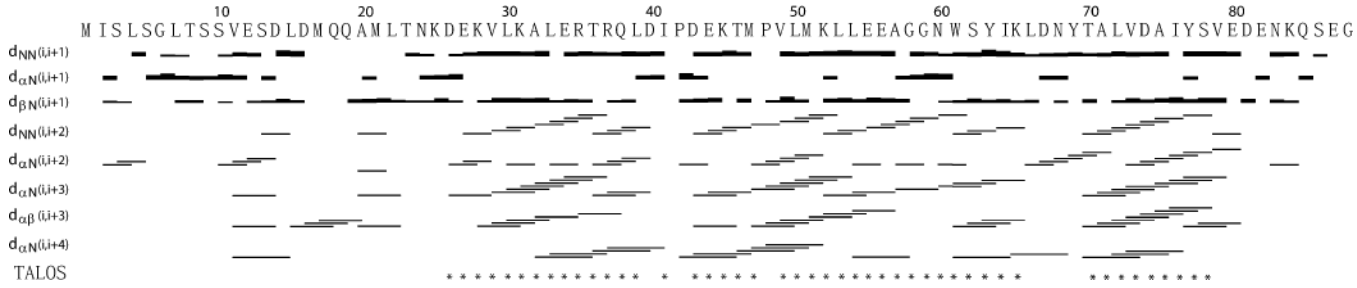


Figure 3.

Short-range interactions suggest a mainly α-helical structure of S4WIYLD. Sequence plot of S4WIYLD with short-range correlations used in the structure calculation illustrated with lines and the existence of torsion angle restraints from TALOS symbolized with stars. The thicknesses of the lines are related to NOE cross-peak intensities.

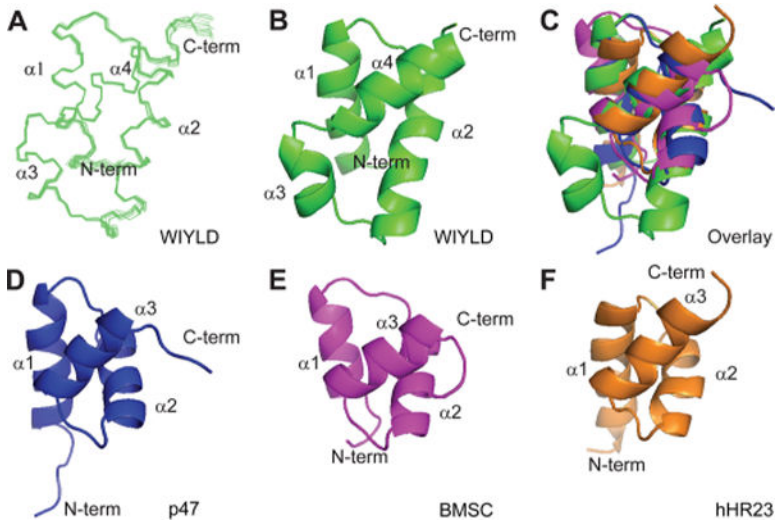


Figure 4.

Solution structure of the S4WIYLD domain that is a four-helix bundle distinct from three-helix bundle UBDs. (A) Cα trace representing the superimposition of 20 S4WIYLD domain structures of aa residues 25–81. (B) Ribbon representation of the lowest-energy structure of the S4WIYLD domain. (C) Structure of selected three-helix bundle UBDs for comparison to S4WIYLD superimposed on one another (S4WIYLD, green). (D–F) Ribbon representation of the UBDs of human proteins (D) p47 (PDB entry 1V92, blue), (E) BMSC (PDB entry 2CWB, magenta), and (F) hHR23 (PDB entry 1IFY, orange).

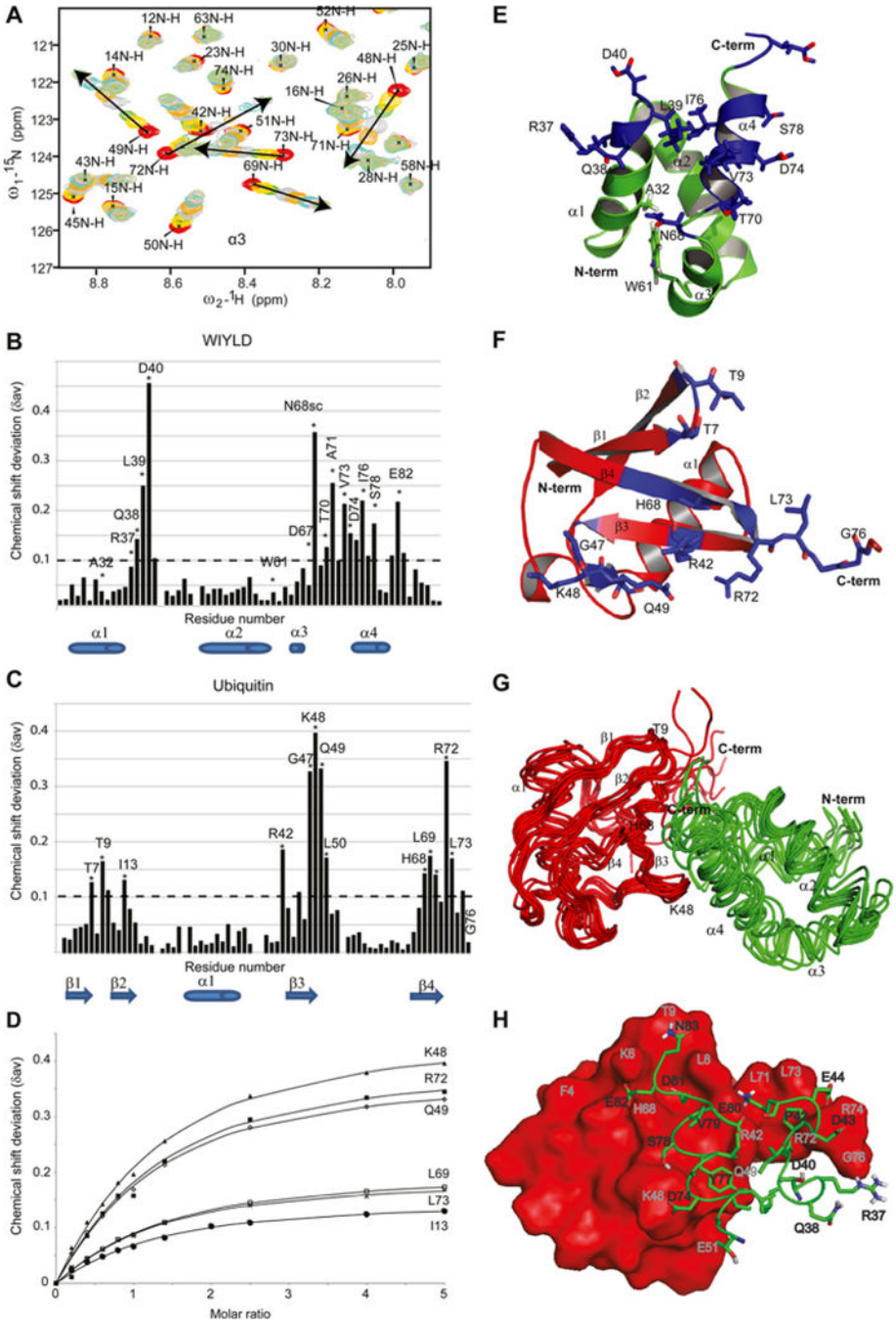


Figure 5.

CSP analyses and docking suggest binding between the S4WIYLD domain and ubiquitin via electrostatic interactions. (A) Expansion of parts of the ^{15}N HSQC spectrum of S4WIYLD (red), with the S4WIYLD:ubiquitin ratio increasing to 1:5. The labels indicate the assignments of peaks with regard to the S4WIYLD sequence in the ^{15}N HSQC spectrum of S4WIYLD. The arrows indicate the direction of the peak motion. (B and C) Sequence dependence of the combined chemical shift deviations for the backbone ^{15}N and ^{1}H resonances of the (B) labeled S4WIYLD in the presence of ubiquitin to a molar ratio 1:5 and

(C) labeled ubiquitin in the presence of S4WIYLD to a molar ratio 1:5. sc represents side chain. The dashed lines indicate the chemical shift change that was used as a threshold value (0.10 ppm). The positions of α-helices and β-strands are indicated below the *X*-axes. (D) Changes in chemical shift for ubiquitin upon titration with S4WIYLD for selected amino acids. The lines indicate the best fit to the K_d value. (E and F) Residues with backbone combined $^{1}H–^{15}N$ chemical shift deviation values greater than or equal to the threshold value (blue), mapped onto a ribbon representation of (E) the S4WIYLD structure and (F) the ubiquitin structure. Selected side chains of interacting ubiquitin and S4WIYLD residues, as well as S4WIYLD aa subjected to mutational analyses, are shown. The α-helices, β-strands, and N- and C-termini are indicated. (G) Superimposition of the 10 best docking models obtained for S4WIYLD and ubiquitin using HADDOCK. α-Helices, β-strands, and N- and C-termini are indicated, as well as selected amino acids. (H) Docking model of S4WIYLD with ubiquitin. Representation of the interaction surface between ubiquitin (surface in firebrick and red; gray residue numbers) and S4WIYLD domain residues interacting with ubiquitin (green ribbon with sticks for side chains; black residue numbers).

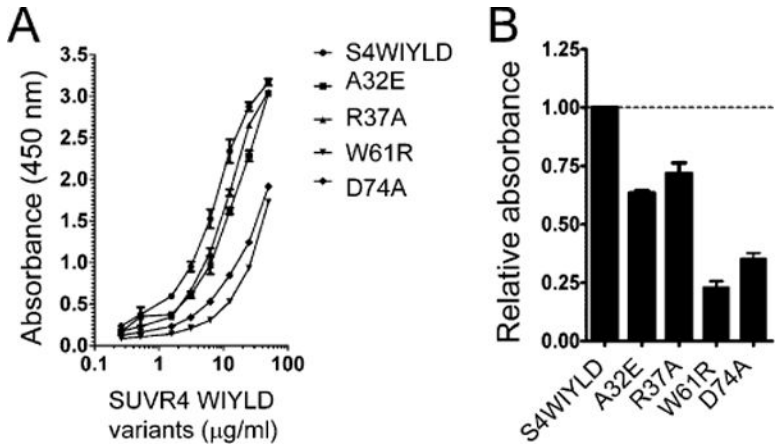


Figure 6.

Identification of S4WIYLD residues crucial for binding to ubiquitin. (A) Titrated amounts of GST-tagged wild-type and mutant variants of S4WIYLD were added to ubiquitin coated in ELISA wells. Bound GST-WIYLD fusion proteins were detected using an HRP-conjugated anti-GST antibody. The data are presented on a logarithmic scale. The results are representative for three independent experiments ($N = 3$). (B) Quantification of the ELISA results relative to that of wild-type S4WIYLD, which was set to 1.

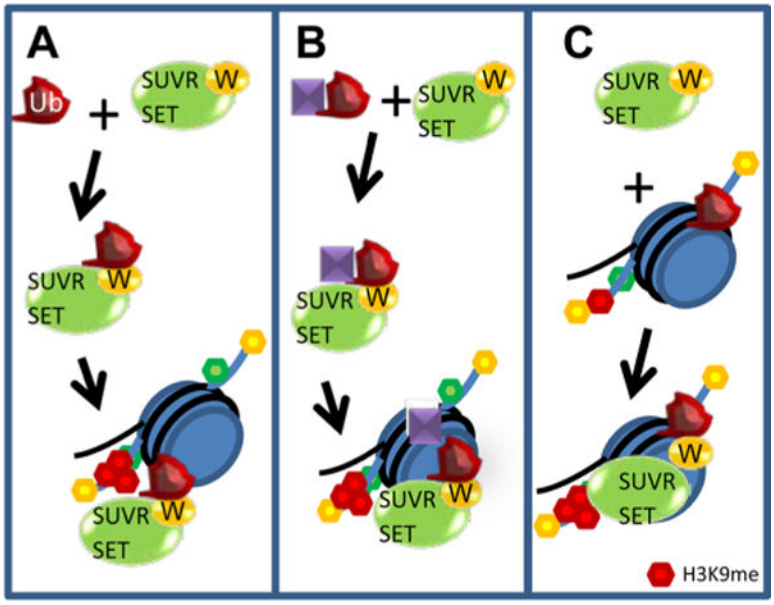


Figure 7.

Scenarios for WIYLD binding of ubiquitin and SUVR HMTase function. (A) Binding of WIYLD to free ubiquitin shifting the SUVR HKMTase from adding one methyl group to adding two methyl groups to monomethylated H3K9. (B) Binding of WIYLD to a ubiquitin-conjugated protein that targets SUVR to specific genes or regions in the genome where it can methylate H3K9me1. (C) WIYLD functioning as a reader of monomethylated H2B facilitating methylation of H3K9me1 on the same or neighboring nucleosomes.

Table 1

NMR Structure Determination Statistics for the SUVR4 WIYLD Domain

no. of experimental restraints	
total distance	1641
intraresidual	429
sequential ($\|i - j\| = 1$)	470
medium-range ($2 \leq \|i - j\| \leq 4$)	496
long-range ($\|i - j\| \geq 5$)	246
dihedral angle	110
total	1751
Structure Statistics	
energy (Å)	1.85 ± 0.24
constraint violations	
distance constraint of >0.1 Å	6 ± 2
angle constraint of >5°	0
distance constraints (Å)	0.0081 ± 0.0015
angle constraints (deg)	0.7122 ± 0.0427
maximal distance violation (Å)	<0.35
maximal angle violation (deg)	<5
coordinate precision (Å),[a] residues 25–81	
backbone (N, Cα, CO)	0.39 ± 0.10
heavy atoms	0.94 ± 0.16
Ramachandran analysis (%)	
residues in most favored regions	80.6
residues in additionally allowed regions	19.4
residues in generously allowed regions	0.0
residues in disallowed regions	0.1

[a]The coordinate precision is defined as the average root-mean-square deviation among the final 20 structures.