BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 10237–10245 10237
DOI: 10.1021/bi101321y

The Effect of Asp-His-Ser/Thr-Trp Tetrad on the Thermostability of WD40-Repeat Proteins†

Xian-Hui Wu,‡ Rong-Chang Chen,‡ Ying Gao,‡ and Yun-Dong Wu*,‡,§,∥

‡*School of Chemical Biology and Biotechnology, Laboratory of Chemical Genomics, Shenzhen Graduate School of Peking University, Shenzhen 518055, China,* §*Department of Chemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China, and* ∥*College of Chemistry, Peking University, Beijing 100871, China*

Received August 17, 2010; Revised Manuscript Received October 8, 2010

ABSTRACT: We recently found that Asp-His-Ser/Thr-Trp hydrogen-bonded tetrads are widely and uniquely present in the WD40-repeat proteins. WDR5 protein is a seven WD40-repeat propeller with five such tetrads. To explore the effect of the tetrad on the structure and stability of WD40-repeat proteins, the wild-type WDR5 and its seven mutants involving the substitutions of tetrad residues have been isolated. The crystal structures of the wild-type WDR5 and its three WDR5 mutants have been determined by X-ray diffraction method. The mutations of the tetrad residues are found not to change the basic structural features. The denaturing profiles of the wild type and the seven mutants with the use of denaturant guanidine hydrochloride have been studied by circular dichroism spectroscopy to determine the folding free energies of these proteins. The folding free energies of the wild type and the S62A, S146A, S188A, D192E, W330F, W330Y, and D324E mutants are measured to be about −11.6, −2.7, −3.1, −2.9, −3.6, −7.1, −7.0, and −7.5 kcal/mol, respectively. These suggest that (1) the hydrogen bonds in these hydrogen bond networks are unusually strong; (2) each hydrogen-bonded tetrad provides over 12 kcal/mol stability to the protein; thus, the removal of any single tetrad would cause unfolding of the protein; (3) since there are five tetrads, the protein must be in a highly unstable state without the tetrads, which might be related to its biological functions.

WD40-repeat proteins belong to β-propeller proteins (*1*−*5*), and they compose about 1% of total proteins in different species (*6*). These proteins are found to be important building blocks in assembling protein complexes (*6*−*10*). The protein complexes are extensively involved in regulating cell cycle (*11*−*14*), activating or inactivating gene expression by alternating or recognizing the histone modification (*15*−*26*), regulating DNA replication and DNA damage repair (*27*, *28*), protein ubiquitination (*12*, *29*−*36*), apoptosis (*37*−*41*), and signal transduction (*42*, *43*). The structural features and biological functions of the majority of WD40-repeat proteins remain to be explored, which is attracting increasing attention.

About 30 crystal structures of WD40-repeat proteins from various species are available in the Protein Data Bank (*44*), of which about 10 are protein-peptide/protein complexes. Through analyzing these structures, we found that every WD40-repeat protein has at least one Asp-His-Ser/Thr-Trp (DHSW) tetrad (Scheme 1) hydrogen bond network (*45*). Some of the proteins have up to six or seven DHSW tetrads. On average each WD40-repeat protein has about three DHSW tetrads. Very similar situations have been found in predicted WD40-repeat proteins or domains in the human genome, as well as in the genomes of other species such as *Drosophila*, *Arabidopsis*, and yeast (data not shown) by the SUPERFAMILY database (*46*, *47*).

Moreover, such DHSW tetrad appears to be unique for WD40-repeat proteins. We have not found such tetrad structures in other types of sheet proteins in our analysis using an in-house code. An examination of the X-ray crystal structures of 65 non-WD40 β-propeller proteins also reveals no such DHSW tetrad (Supporting Information Table S1), although the two types of proteins have similar β-propeller structures. Therefore, it is meaningful to study the importance of the hydrogen-bonded tetrad to the structural stability and possible biofunctions of the WD40-repeat proteins.

Garcia-Higuera et al. studied the influence of Asp residues in the turns connecting strands b and c in Gβ and Sec13 on the folding of the two WD40-repeat proteins (*48*). They found that the conserved Asp residues do not have equivalent effect on the folding of the proteins. While some are critical to the folding of the proteins, others have much less effect (*48*).

WDR5 is a well-known WD40-repeat protein, which folds into a seven WD40 domain propeller structure (*49*). Due to its involvement in the regulation of H3K4 methylation(*16*) by physical contact with the histone H3 N-terminus and MLL1, the crystal structure and the function of WDR5 have been studied intensively (*19*−*21*, *23*, *24*). The crystal structures show that it has five tetrads, which are illustrated in Scheme 2. It is interesting to ask several questions: (1) How much stabilization do these hydrogen bond networks contribute to the stability of WDR5? (2) Is each of the five tetrads needed for the folding of WDR5? (3) How do mutations of the tetrad residues influence the structure of WDR5? To address these questions, we have measured and compared the basic folds and folding free energies of WDR5 wild type (WT) and its various mutants designed to delete some hydrogen bonds in the hydrogen bond network of DHSW tetrad with the use of mutagenesis, crystallographic studies, and protein

†This work was supported by the Research Grants Council of Hong Kong 663509 (to Y.-D.W.) and the Shenzhen Graduate School of Peking University.

*To whom correspondence should be addressed. Phone: 0086-2603-2700. Fax: 0086-2603-3174. E-mail: wuyd@szpku.edu.cn.

Published on Web 10/12/2010

unfolding induced by guanidine hydrochloride. Here we report our findings.

METHODS

Construction of WDR5 Mutants, Protein Overexpression, and Purification. The sequence of WDR5 cDNA encoding 31−334 aa was cloned into pET28a(+) expression vector. Site-directed mutagenesis (Scheme 2) was prepared using the procedures provided by the QuickChange site-directed mutagenesis kit. All of the constructions were verified by DNA sequencing. In the expression of the WT WDR5 and mutants, IPTG was added to 0.1 mM for induction as the OD_{600nm} of cultures reached 0.6. After an 8−10 h induction at a constant temperature of 25 °C, the cell was collected by centrifugation at 5000*g* for 10 min at 4 °C. Then the centrifuged cell was resuspended and sonicated in a 20 mM Tris buffer and 50 mM imidazole, pH 8.0, containing 500 mM NaCl. The debris was removed by centrifugation at 20000*g* for 3 × 30 min. The resulting supernatant was loaded to a Ni-NTA column after being filtered through a 0.22 μm pore filter from Millipore. The needed fraction was eluted at the concentration of about 300 mM imidazole. The elution was concentrated to no more than 2 mL, which can be loaded to Superdex 75 for the following purification with the use of buffer 20 mM Tris-HCl, pH 8.0, including 200 mM NaCl. The collected fractions were concentrated to about 10−25 mg/mL for crystallization and CD measurements. The final protein concentration has been determined with using the calculated extinction coefficient at the wavelength of 280 nm (*50*).

Determination of the Crystal Structures for WDR5 Mutants. To further understand the influence of the overall fold and the hydrogen bond formed by the DHSW tetrad, we have raised the crystal structures for WDR5 WT and its S62A, W330F, and W330Y mutants. The crystallizations were carried out using the sitting-drop or hanging-drop vapor-diffusion method at 293−298 K. Crystals showed up after 2 days in the condition of 18−25% PEG3350 and 150 mM $(NH_4)_2SO_4$, pH 7.0. The crystals were transferred into the cryosolution, containing 90% of the reservoir and 10% glycerol, for about 20 s, then flash-cooled, and maintained at 100 K (Oxford Instruments) using a cold nitrogen-gas stream during the whole data collection process. The diffraction data were collected by a Rigaku Saturn 944+ CCD detector using Cu Kα radiation from a Rigaku MicroMax002+ generator operated at 45 kV and 88 mA. The CrystalClear 1.4 (Rigaku, USA) software package was used to process the data. The structure was determined by the molecular replacement (MR) method using the program MolRep in the CCP4 package (*51*). Further model building and refinement were performed manually using the graphical program COOT(*52*) and the program Refmac5 (*53*). The model geometry was evaluated by PROCHECK (*54*).

Equilibrium Circular Dichroism (CD) Measurements. CD measurements were carried out on a Chirascan spectrometer with a Peltier temperature-controlled cell based on the protocol (*55*−*58*). All CD measurements were carried out in 1 × PBS buffer at pH 7.5. For the GdmHCl-induced denaturation, samples with 100 μg/mL protein at different GdmHCl concentrations were made up individually and equilibrated at 25 °C for 4 h. Due to the disturbing noise of GdmHCl below 210 nm, the CD signal was measured at 230 nm with a path of 0.4 cm, and data were averaged over six independent measurements.

Two-State Analysis of Denaturation by GdmHCl. The GdmHCl denaturation profile was fit using a two-state assumption. The folding free energy without GdmHCl was estimated by a linear extrapolation method (*58*):

$$\alpha_i = \frac{[\theta_i] - [\theta_U]}{[\theta_F] - [\theta_U]} \quad (1)$$

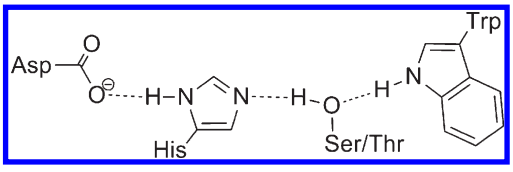
where $[\theta_i]$ is the ellipticity at the *i*th concentration of GdmHCl, $[\theta_F]$ is the ellipticity at which the protein is fully folded, and $[\theta_U]$ is the ellipiticity of protein in the 6 M GdmHCl. The protein in 6 M GdmHCl is assumed to be the fully unfolded.

$$K_i = \frac{\alpha_i}{1 - \alpha_i} \quad (2)$$

where K_i is the folding constant for a monomeric protein at the *i*th concentration of GdmHCl, which can be calculated by the folded fraction α_i. The free energy of folding of a protein can be evaluated from the equation:

$$\Delta G_F = -RT \ln K_F \quad (3)$$

Scheme 1: The Hydrogen Bond Network of Asp-His-Ser/Thr-Trp (DHSW) Tetrad



Scheme 2: Five DHSW Tetrads Are Present in WDR5[a]

[a]The color-marked substitutions are purified and studied by the GdmHCl induced unfolding.

Table 1: Data Collection Statistics of WDR5 S62A, W330F, and W330Y Mutants[a]

	S62A	W330F	W330Y
data collection			
wavelength (Å)	1.5418	1.5418	1.5418
space group	$P12_11$	$P12_11$	$P12_11$
unit cell parameters			
a, b, c (Å)	43.63, 79.65, 45.66	43.72, 79.65, 45.97	43.75, 79.68, 45.79
α, β, γ (deg)	90.00, 90.00, 102.02	90.00, 90.00, 101.73	90.00, 90.00, 101.91
resolutions (Å)	42.67−2.05 (2.12−2.05)	31.85−2.05 (2.12−2.05)	39.05−2.10 (2.18−2.10)
R_{sym}[b] or R_{merge}	0.064 (0.142)	0.079 (0.218)	0.123 (0.242)
$I/\sigma I$	11.8 (5.9)	9.1 (3.5)	9.0 (4.8)
completeness (%)	99.75 (98.9)	99.8 (98.2)	100.0 (99.9)
redundancy	3.49 (2.84)	3.51 (2.82)	6.72 (6.23)
refinement			
resolutions (Å)	42.67−2.05	31.85−2.30	39.05−2.20
no. of unique reflections	19198	19383	17994
R_{work}/R_{free}	0.188/0.239	0.167/0.233	0.220/0.273
no. of atoms			
protein	2424	2433	2434
water	53	172	212
B factors			
protein	16.08	17.35	12.57
water	14.38	22.67	20.07
rmsd			
bond lengths (Å)	0.0221	0.0194	0.0189
bond angles (deg)	1.8642	1.6715	1.7659
Ramachandran (%)			
most favored	85.1	87.2	86.1
additionally allowed	14.9	12.8	13.5
generously allowed	0.0	0.0	0.4
disallowed	0.0	0.0	0.0

[a]Values in parentheses are for the highest resolution shell. [b]$R_{sym} = \sum\sum_i |I(h) - I(h)_i| / \sum\sum_i I(h)$, where $I(h)$ is the mean intensity.

where R is the gas constant, which is equal to 1.98 kcal/mol, T is the absolute temperature, and K_F is the folded constant of monomeric protein, which can be calculated by the function K_F = [F]/[U], where [F] and [U] represent the folded and unfolded fractions, respectively.

$$\Delta G_i = \Delta G_0 + m[\text{GdmHCl}] \quad (4)$$

The free energy of protein folding is a linear function of the amount of GdmHCl added, where ΔG_i is the free energy of the protein in the presence of the ith concentration of GdmHCl and ΔG_0 is the free energy of the folded protein in the absence of GdmHCl.

RESULTS AND DISCUSSION

To investigate the influence of the hydrogen bond network formed by DHSW tetrad on the WDR5 stability, a series of mutations that involve residues in the tetrads, such as D192E/N, D324E/N, W198F/Y, W156F/Y, W330F/Y, S62A, S104A, S146A, S188A, and S318A, were carried out to alter the hydrogen bond networks (Scheme 2). Mutants of H44A, H300A, S62A/S188A, and S62A/S146A were also overexpressed in *Escherichia coli*. However, these mutants could not be purified due to their insolubilities, suggesting that these mutants could not fold properly. As will be shown later, these mutations destroy strong hydrogen bonds in tetrads and cause too much destabilization.

We successfully purified the WDR5 WT protein and its S62A, S146A, S188A, D192E, W324E, W330F, and W330Y mutants with the use of Ni-NTA and Superdex 75 gel filtration tandemly. Although these mutants have remarkably different solubilities, the final samples were all concentrated up to 5−25 mg/mL without precipitation for the further crystallization and CD measurement.

The Crystal Structure of WDR5 S62A and D330F/Y Mutants. In order to confirm the preservation of the WDR5 entire fold and the alternation of the hydrogen bond network as we remove the hydrogen bonds in the DHSW tetrad by substituting Asp, Ser, and Trp residues, high quality crystal structures of WDR5 S62A and W330F/Y mutants were determined (Table 1). The resolutions of three mutants are all better than 2.5 Å, which allow us to further analyze and understand the results on the basis of protein structures.

By the superposition of crystal structures of WDR5 WT and S62A, W330F, and W330Y mutants, the overall folds are almost identical as shown in Figure 1A. Therefore, the WDR5 S62A, W330F, and W330Y mutations that disrupt the hydrogen bond network of tetrad have little effect on the basic fold. The hydrogen bonds in the tetrads are analyzed as the polar protons are assigned. The parameters of hydrogen bonds formed by the WDR5 tetrads suggest that the single substitutions of residues involved in the tetrads only alter the hydrogen bond network locally (Supporting Information Table S2). Comparing the hydrogen bond network of DHSW tetrads in the WDR5 WT, the WDR5 S62A mutation removes two hydrogen bonds between H44 and S62 and between S62 and W72 (Figure 1B, Supporting Information Table S2). According to the crystal structure of WDR5 S62A mutant, no water molecule is found between H44 and W72, which can play the similar role as the S62 side chain. In addition, the W72 indole ring is found to rotate by about 180° so that the N−H of the indole ring exposes to solvent. In the WDR5 W330F and W330Y mutants, panels C and D of Figure 1 show that the hydrogen bond formed between S318 and

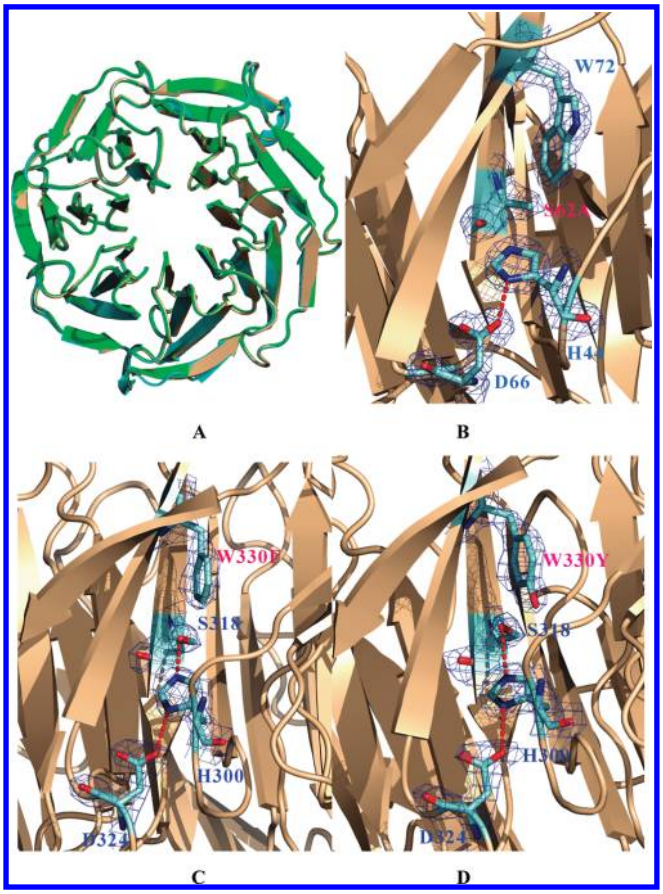


FIGURE 1: (A) The superposition of WDR5 WT and S62A, W330F, and W330Y mutants. (B), (C), and (D) are the electron densities of WDR5 S62A (3MXX), W330F (3N0D), and W330Y (3N0E) mutants. The dashed line is the hydrogen bond. (B) The electron density of WDR5 S62A mutant. (C) The electron density of WDR5 W330F mutant. (D) The electron density of WDR5 W330Y mutant.

W330 is deleted without disturbing the hydrogen bond network formed by the D324-H300-S318 triad. Based on the crystal structure of WDR5 W330Y mutant (Figure 1D), Y330 indeed does not form a hydrogen bond with S318, although the phenol group of the Tyr side chain is a stronger hydrogen bond donor than the N−H of Trp. The orientation of the phenol group does not allow the hydrogen bond formation. Although the crystal structures of other mutants are not obtained, we expect that all of these mutant proteins have similar structures as indicated by the very similar CD spectra as will be discussed later. Since the crystal structures indicate very little structural perturbation by these mutants, the variation of protein stabilities is likely mainly due to the loss of hydrogen bonding.

CD Measurement. As shown in Figure 2, the WDR5 WT and its mutant proteins have quite similar CD profiles in the range of 200−260 nm at 25 °C, pH 7.5, at the 1 × PBS buffer. All of these proteins have a deep minimum near 210 nm and a maximum near 230 nm. These are characteristic for β-sheet proteins. Thus, the CD spectra also indicate that the mutations of the tetrad residues do not affect the overall propeller structure of the protein, in accordance with the crystallographic observations (Figure 1A).

Shown in Figure 3 are the changes of CD ellipticity of WDR5 WT and its S146A mutant upon the addition of different concentrations of denaturant guanidine hydrochloride (GdmHCl). The intensity change near 230 nm is most sensitive to the concentration of GdmHCl. On the other hand, the intensity change at

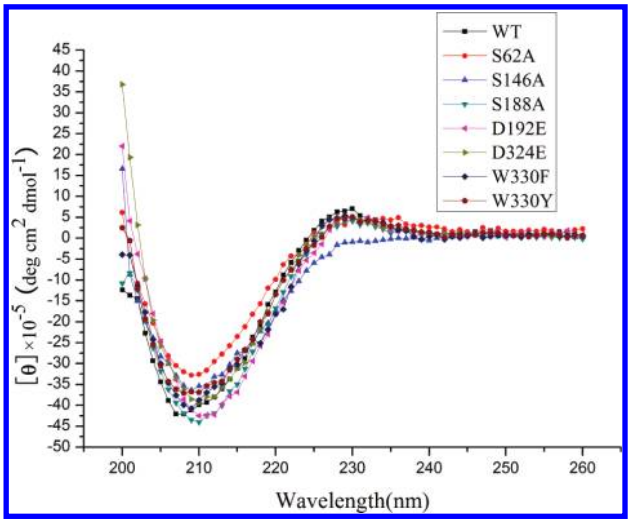


FIGURE 2: Circular dichroism spectrum of WDR5 WT and seven mutants. Spectra were collected at 25 °C and pH 7.5 in 1 × PBS buffer.

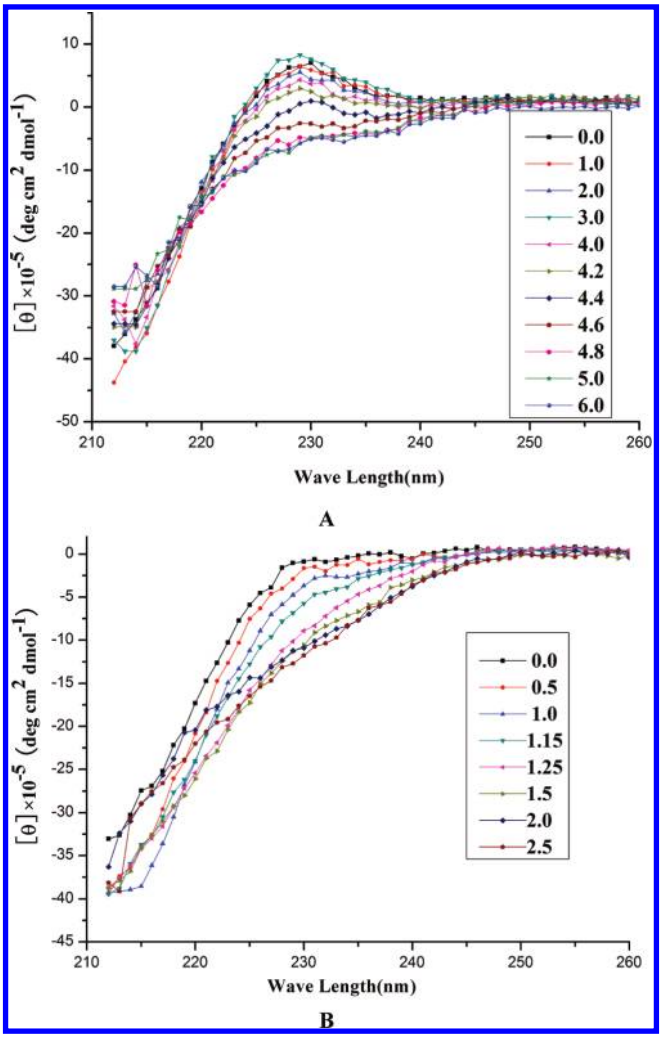


FIGURE 3: Changes of ellipticities of proteins in the range of 210−260 nm as the concentration of GdmHCl is increased at 25 °C and pH 7.5 in the 1 × PBS buffer. (A) WDR5 WT. (B) WDR5 S146A mutant.

the major intensity near 210 nm is not sensitive to the concentration of GdmHCl. This is because GdmHCl itself has absorption below 210 nm, perturbing the CD spectrum at 210 nm. Therefore,

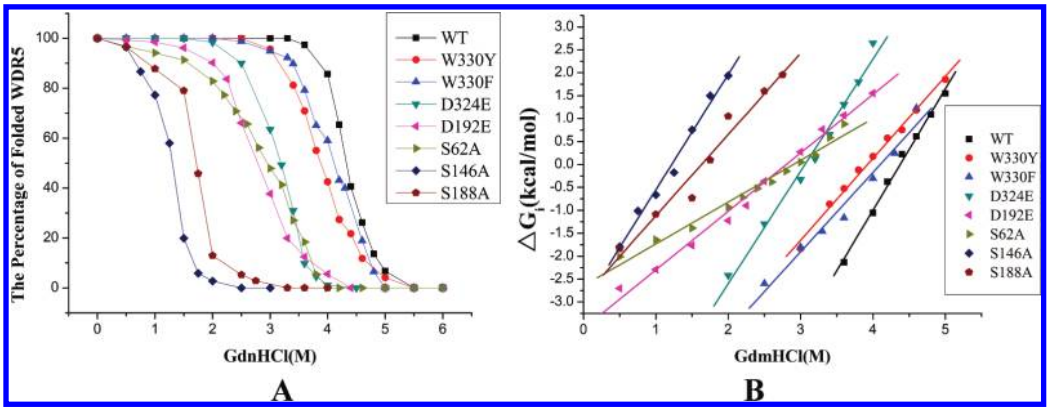


FIGURE 4: (A) The percentage of folded WDR5 WT and mutants vs GdmHCl. (B) The fitted folding free energies of WDR5 WT and mutants using the two-state model by the linear extrapolation method.

Table 2: Measured $C_{1/2[GdmHCl]}$, Folding Free Energy, Standard Error of Folding Free Energy, Change of Folding Free Energy, m Value, and Standard Error of m Value of WDR5 WT and Seven Mutants[a]

	WT	S62A	S146A	S188A	D192E	W330F	W330Y	D324E
$C_{1/2[GdmHCl]}$	4.4	3.0	1.4	1.8	2.8	4.1	4.0	3.2
ΔG_{ND}	−11.6	−2.7	−3.1	−2. 9	−3.6	−7.1	−7.0	−7.5
$SE_{\Delta G}$[b]	0.4	0.14	0.2	0.3	0.1	0.5	0.2	0.4
$\Delta\Delta G_{WT\ mutant}$	0.0	9.0	8.5	8.7	8.0	4.5	4.6	4.14
m value	2.7	0.9	2.5	1.8	1.3	1.7	1.8	2. 5
SE_m[b]	0.1	0.0	0.1	0.2	0.0	0.1	0.1	0.1

[a]The units of $C_{1/2[GdmHCl]}$ and folding free energy are in mol/L and kcal/mol, respectively. [b]The standard errors of folding free energies and m values, respectively.

we chose the intensity change around 230 nm to calculate the folding free energy of WDR5 WT and mutants.

GdmHCl-Induced Denaturation. We first tried to induce denaturation of WDR5 WT and its mutants by increasing temperature. However, it is observed that some proteins undergo aggregation upon unfolding at increased temperatures. This makes the measurement of CD ellipticity unreliable for unfolding transition temperature.

Therefore, we measured the stabilities of WDR5 WT and its mutants by performing protein unfolding with the denaturant guanidine hydrochloride. Circular dichroism spectroscopy was recorded at 25 °C and pH 7.5 in 1 × PBS buffer with different concentrations of GdmHCl. Based on the ellipticity intensity near 230 nm, the percentage of folded protein can be calculated by the two-state analysis (see Methods). Shown in Figure 4A are the plots of the folded fraction of the eight proteins against the concentration of GdmHCl. They indicate that the denaturation induced by GdmHCl is cooperative in an apparently two-state curve. The WDR5 WT is the most stable, and it starts to unfold at the highest GdmHCl concentration. The mutants are more sensitive to the concentration change of GdmHCl. The midpoint transition of WDR5 WT is at 4.4 M, much higher than those of WDR5 mutants (Table 2). The midpoint transitions of WDR5 W330F/Y mutants are about 4.1 and 3.9 M, respectively. WDR5 W330F/Y mutants may have the similar fold and hydrogen bond network in solution due to the similar midpoint transitions and unfolding profiles. Thus, the alternation of midpoint transition mainly corresponds to the deletion of hydrogen bond formed between S318 and W330. The measured midpoint transitions of WDR5 D192E and D324E mutants are 2.8 and 3.2 M, respectively, which are smaller than that of the WDR5 WT. The WDR5 S62A, S146A, and S188A mutants have different midpoint transition values, 3.0, 1.4, and 1.8 M, respectively. The different midpoint transitions indicate different sensitivities of the mutants to the denaturant GdmHCl.

Folding Free Energies. The folding free energies of WDR5 WT and its mutants may provide us the thermodynamic effect of the DHSW tetrad on the WD40-repeat proteins. Shown in Figure 4B are the plots based on eq 4 for the eight proteins. The linear relationship between ΔG_i and concentration of GdmHCl is quite good. The intercept, where $\Delta G_i = \Delta G_0$, gives the folding free energy. The slope, m, indicates the sensitivity of protein stability toward the denaturant. These values are given in Table 2. Each point in Figure 4 is the average of six individual measurements. These detailed data are given in Supporting Information Table S3. Based on these data, the standard errors to ΔG_0 and m for each protein can be derived, which are also given in Table 2.

The fitted folding free energy ΔG_{ND} of WDR5 WT is about −11.6 kcal/mol. The folding free energy for W330F and W330Y is −7.1 and −7.0 kcal/mol, respectively. In these mutations, one hydrogen bond (between S318 and W330 in the WT) is removed while the hydrogen bond network of D324-H300-S318 is intact (Figure 1C,D). Although the mutations may cause some variations in hydrophobic interactions because the residue is in a hydrophobic environment, we may still give a rough estimation of the hydrogen bond strength of about 4 kcal/mol.

The measured folding free energies for WDR5 S62A, S146A, and S188A mutants range from −2.7 to ∼− 3.1 kcal/mol. These mutants are about 8.5−9.0 kcal/mol less stable than the WDR5 WT! The mutations not only remove two hydrogen bonds, leaving only the hydrogen bond between Asp and His, but also disrupt cooperative interaction in the hydrogen bond network of the tetrad, which is substantial (2.0−2.5 kcal/mol) according to our earlier quantum mechanical calculations (*45*).

FIGURE 5: The solvent accessibility of DHSW tetrad. The D66-H44-S62-W72 tetrad and the two WD40 repeats are colored in green, yellow, and blue, respectively. The surrounding hydrophobic residues from two neighboring WD40 repeats shield the hydrogen bonds formed by H44-S62-W72 triad.

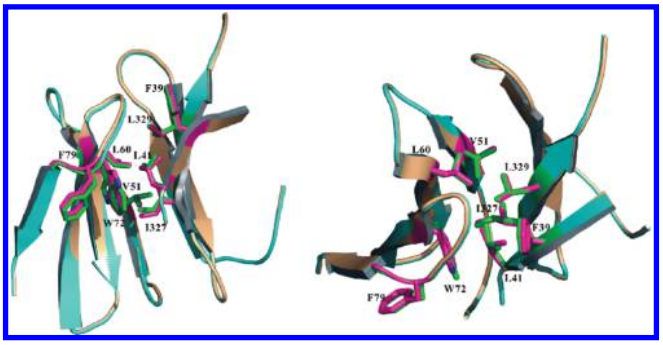


FIGURE 6: The superposition of two WD40 repeats, WDR5 WT (cyan) and its S62A mutant (wheat). The stick represents the hydrophobic residues (WT in green, S62A in red) involved in surrounding the Trp72 indole ring.

Quite different folding free energies were obtained for WDR5 D192E and D324E mutants. The folding free energy of −7.5 kcal/mol for WDR5 D324E mutant appears to be similar to those of WDR5 W330Y and W330F mutants. However, that of WDR5 D192E mutant is only −3.6 kcal/mol, which is about 8 kcal/mol less stable than that of the WDR5 WT. These are in accord with the inequivalent effects of Asp to Gly mutations on the folding of Gβ and Sec13 proteins that were reported by Garcia-Higuera et al. (*48*).

These results raise two questions: (1) Neutral hydrogen bonds in proteins normally are considered to cause less than 3.6 kcal/mol stabilization in terms of free energy (*59*−*62*). Normally, the average magnitudes of most of the neutral hydrogen bonds measured by Pace et al. (*59*, *60*) are less than 1.5 kcal/mol. However, the above hydrogen bonds seem to be over 4 kcal/mol. What causes the unusually large stabilization? (2) What causes the large difference in stabilization between WDR5 D192E and D324E mutants?

Solvent Accessibilities of the Hydrogen Bond Networks. In order to answer the above-mentioned questions, the solvent environment of the tetrads was analyzed because the strength of the hydrogen bond can be remarkably different in different solvent environments (*60*, *63*). Shown in Figure 5 are two views of the hydrogen bond network of the D66-H44-S62-W72 tetrad of WDR5 WT. The hydrogen bond network locates in the cleft of two neighboring WD40 repeats. The hydrogen bonds formed by H44-S62-W72 are surrounded by the hydrophobic residues located on strands a and d of its own WD40 and strand d of neighbored WD40 as well as the loop connecting two WD40 repeats.

To semiquantify the solvent accessibilities of the hydrogen bond networks formed by five DHSW tetrads of WDR5, we have calculated the solvent-accessible surface area (SASA) of their polar atoms using the SURFACE program embedded in the CCP4i package (*51*). According to the calculations, the SASAs of oxygen atoms of solvent-exposed Glu and Asp carboxyl groups are all over 10 Å^2, which is much larger than 0−5 Å^2 SASA of the oxygen atoms belonging to most of Asp carboxyl groups involved in the tetrads (Figure 5 and Supporting Information Tables S4 and S5), suggesting that Asp carboxyl groups are not fully accessible by the solvent. The calculated SASA values of His, Ser, and Trp in the tetrads also indicate that the hydrogen bonds formed between His and Ser and between Ser and Trp are all in the hydrophobic environment. In addition, the solvent accessibilities of polar atoms involved in the formation of hydrogen bonds of WT WDR5 and its mutants have no remarkable change except for the W72 NE1 atom in the WDR5 S62A mutant (Supporting Information Table S4). The SASA of the W72 NE1 atom changes from 0 to 5.8 Å^2 in the WDR5 S62A mutant. This is mainly due to the orientation of the indole ring upon the deletion of the hydrogen bond formed between S62 and W72 (Figure 1B, Supporting Information Table S4).

We also calculated the SASAs of all tetrads in the 30 available crystal structures of WD40-repeat proteins (Supporting Information Table S5). The hydrogen bonds formed by H-S-W in the tetrads are all water inaccessible as indicated by the near zero SASAs of the polar atoms involving the hydrogen bonds. Although the SASAs of most Asp carboxyl groups are not zero, all Asp in the tetrads are mainly in hydrophobic environments, which provide a strong hydrogen bond due to the ionic carboxyl groups. All of the solvent accessibilities and cooperativities of tetrads in 30 available crystal structures of WD40-repeat proteins are similar to those of WDR5, which supports that there is not much difference in the thermostabilities provided by the DHSW tetrads in WD40-repeat proteins.

The m Value. The definition of *m* value can be interpreted by the free energy change of the folded and unfolded protein being transferred from water to 1 M GdmHCl (*64*). It is also well-known that there is strong correlation between *m* value and the SASA upon unfolding (*65*, *66*), because Gdm^+ ions preferentially interact with and stabilize the hydrophobic groups in the unfolded state (*67*, *68*). The reduction of the *m* value suggests two possibilities: (1) The substitutions increase the solvent accessibility of hydrophobic residues in the folded states (*69*). (2) The deletion of some hydrogen bonds by the substitutions weakens the hydrophobic interaction of some particular residues, which increases their accessibilities or sensitivities to GdmHCl. According to our calculations, the marginally different SASAs of WDR5 WT and S62A, W330F, and W330Y mutants in the folded states (Supporting Information Table S6) exclude the possibility that different SASAs account for different reductions of *m* values by the substitutions.

On the other hand, Figure 4B shows that there is little difference among WDR5 S62A, S146A, and S188A mutants at ~0.5 M GdmHCl. As the concentration of GdmHCl increases, the difference becomes larger and larger, suggesting that these three mutants have different sensitivities to GdmHCl in the process of unfolding. The *m* values of the WDR5 S62A, S146A, and S188A are 0.9, 2.5, and 1.8, respectively. We note that seven hydrophobic residues surround W72: F39, L41, I327, and L329 from the neighboring WD40 and V51, L60, and F79 from its own WD40 surround the W72 indole ring as shown in Figure 6. As S62 being replaced by Ala, the W72 indole ring flips about 180° without alternating the hydrophobic interaction significantly. However, the loop between strands c and d is

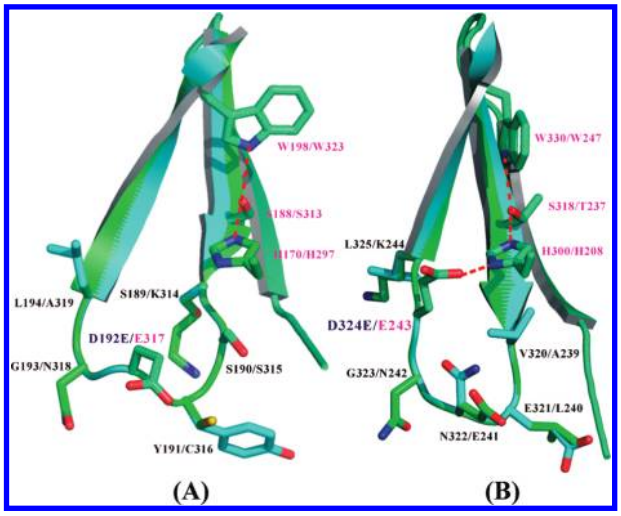


FIGURE 7: The proposed loop structures between strands b and c in the WDR5 D192E and D324E mutants. (A) The superposition between the proposed loop structure of WDR5 D192E mutant (cyan) and EED (2QXV, green). (B) The superposition between the proposed loop structure of WDR5 D324E mutant (cyan) and SIF2 (1R5M, green).

loosened in the WDR5 S62A, indicating that the hydrophobic interaction between W72 and the seven hydrophobic residues may be weakened without the constraint of the hydrogen-bonded tetrad (Figure 6). Moreover, the looseness of hydrophobic stacking may make the seven hydrophobic residues more accessible by GdmHCl and reduce the *m* value. On the other hand, five and six hydrophobic residues surround W156 and W198 indole rings, respectively (Supporting Information Figures S2 and S3). Trp156 indole ring is only buried by V103, L111, I113, L125, and I144, and W198 is surrounded by V145, V153, I155, L167, V177, and I186. The hydrophobic core formed by various numbers of residues may cause the different sensitivities to the denaturant GdmHCl, which is in agreement with the various reductions in *m* values of S62A, S146A, and S188A mutants. This trend is also in correlation with the measured ΔG of these three mutant proteins. WDR5 S62A mutant, which has the smallest *m* value, causes the largest reduction in protein stability ($\Delta\Delta G_{\text{WT mutant}}$ = 9.0 kcal/mol). While WDR5 S146A, which has the smallest reduction in *m* value, also has the smallest reduction in stability ($\Delta\Delta G_{\text{WT mutant}}$ = 8.5 kcal/mol).

The difference between D192E and D324E mutations in causing protein stability reduction of 8.0 and 4.1 kcal/mol, respectively, and variation in *m* values, 1.3 and 2.5, respectively, compared to 2.7 for the WT is most intriguing. Although the crystal structures of the D192E and D324E mutants of WDR5 could not be obtained, we were able to find two similar mutation situations in the available X-ray crystal structures of WD40-repeat proteins. Shown in panels A and B of Figure 7 are the part of the structure between strand b and strand c from EED (2QXV) (*70*) and SIF2 (1R5M) (*71*), respectively. Both structures have a six-residue loop connecting the two strands. In EED (Figure 7A), a Glu residue is at position 317, which is four residues up from S313. The E317 does not form a hydrogen bond with H297, which is in the H297-S313-W323 hydrogen-bonded triad. In SIF2 (Figure 7B), there are two Glu residues in the loop. The first one is E241, which is four residues up from T237. Similar to E317 in EED, it does not form a hydrogen bond with H208. However, the second Glu residue, E243, which is six residues up from T237, forms a hydrogen bond with H208, so that E243-H208-T237-W247 retains a hydrogen-bonded tetrad.

The above two situations are very similar to the D192E and D324E mutants, respectively. As shown by superpositions (Figure 7), The E192 of the D192E mutant is at the position of E317 of EED or E241 of SIF2, while E324 of the D324E mutant is at the position of E243 of SIF2. It is reasonable to hypothesize that, in the D192E mutant, the E192 cannot form a hydrogen bond with H170 because of a ring constraint, and in D324E mutant, the E324 can form a hydrogen bond with H300. Such a hypothesis may allow us to rationalize the fact that the D192E mutation significantly reduces the protein stability by about 8.0 kcal/mol (strong anion−hydrogen bonds have been recognized since 1987; see, for example, ref *72*) while D324 mutation reduces the protein stability by only about 4.1 kcal/mol. The model also accommodates with the observations that the D192E mutant has a considerably reduced *m* value of 1.3 while the D324E mutant has a similar *m* value (2.5) as the WT (2.7) because the latter maintains the DHSW tetrad hydrogen bonding while in the former E192 is exposed.

SUMMARY

We have used X-ray crystallographic method and CD spectroscopy to explore the effects of the hydrogen bond network formed in DHSW tetrad on the structure and stability of WD40-repeat proteins with the use of WDR5 as a model. Based on the crystal structures, WDR5 W330F, W330Y, and S62A mutations have little effect on the basic structure. The CD measurements also suggest that the solution structure is not affected by the mutations. Therefore, the change of folding free energy by mutations mainly corresponds to the deletion of hydrogen bonds. The folding free energies of WDR5 WT and its mutants have been measured with the use of denaturant guanidine hydrochloride. Surprisingly large hydrogen bond energies (over 4 kcal/mol) have been measured for the hydrogen bonds in the tetrads. These are mainly due to highly hydrophobic environment and a cooperative effect of the tetrads. It can be concluded that the complete deletion of a single tetrad will lead to unfolding of WDR5. It can also be expected that the number of tetrads in a WD40-repeat protein may correlate with the instability of the surface of the protein, which may determine the biofunction of the protein. Future research is directed to the structure-based biofunction analysis.

SUPPORTING INFORMATION AVAILABLE

List of PDB_ID of non-WD40 β-propeller, hydrogen bond parameters and solvent-accessible surface area of crystal structures; detailed information of crystal structures and CD experiment of WDR5 and its mutants. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Fülöp, V., and Jones, D. T. (1999) β Propellers: structural rigidity and functional diversity. *Curr. Opin. Struct. Biol. 9*, 715–721.
2. Paoli, M. (2001) Protein folds propelled by diversity. *Prog. Biophys. Mol. Biol. 76*, 103–130.
3. Andrade, M. A., Perez-Iratxeta, C., and Ponting, C. P. (2001) Protein repeats: structures, functions, and evolution. *J. Struct. Biol. 134*, 117–131.
4. Jawad, Z., and Paoli, M. (2002) Novel sequences propel familiar folds. *Structure 10*, 447–454.
5. Pons, T., Gómez, R., Chinea, G., and Valencia, A. (2003) Beta-propellers: associated functions and their role in human diseases. *Curr. Med. Chem. 10*, 505–524.

6. Stirnimann, C. U., Petsalaki, E., Russell, R. B., and Müller, C. W. (2010) WD40 proteins propel cellular networks. *Trends Biochem. Sci.* (doi: 10.1016/j.tibs.2010.04.003).
7. Neer, E. J., Schmidt, C. J., Nambudripad, R., and Smith, T. F. (1994) The ancient regulatory-protein family of WD-repeat proteins. *Nature 371*, 297–300.
8. Smith, T. F., Gaitatzes, C., Saxena, K., and Neer, E. J. (1999) The WD repeat: a common architecture for diverse functions. *Trends Biochem. Sci. 24*, 181–185.
9. Yu, L. H., Gaitatzes, C., Neer, E., and Smith, T. F. (2000) Thirty-plus functional families from a single motif. *Protein Sci. 9*, 2470–2476.
10. Li, D., and Roberts, R. (2001) WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cell. Mol. Life Sci. 58*, 2085–2097.
11. Sansam, C. L., Shepard, J. L., Lai, K., Ianari, A., Danielian, P. S., Amsterdam, A., Hopkins, N., and Lees, J. A. (2006) DTL/CDT2 is essential for both CDT1 regulation and the early G2/M checkpoint. *Genes Dev. 20*, 3117–3129.
12. Jin, J. P., Arias, E. E., Chen, J., Harper, J. W., and Walter, J. C. (2006) A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. *Mol. Cell 23*, 709–721.
13. Yu, H. (2007) Cdc20: a WD40 activator for a cell cycle degradation machine. *Mol. Cell 27*, 3–16.
14. Zeng, C. J. T., Lee, Y. R. J., and Liu, B. (2009) The WD40 repeat protein NEDD1 functions in microtubule organization during cell division in *Arabidopsis thaliana*. *Plant Cell 21*, 1129–1140.
15. Chen, G., and Courey, A. J. (2000) Groucho/TLE family proteins and transcriptional repression. *Gene 249*, 1–16.
16. Wysocka, J., Swigut, T., Milne, T. A., Dou, Y. L., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H., and Allis, C. D. (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell 121*, 859–872.
17. Montgomery, N. D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S. J., Otte, A. P., and Magnuson, T. (2005) The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr. Biol. 15*, 942–947.
18. Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., III, Voigt, P., Martin, S. R., Taylor, W. R., Marco, V. D., Pirrotta, V., Reinberg, D., and Gamblin, S. J. (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature 461*, 762–767.
19. Couture, J. F., Collazo, E., and Trievel, R. C. (2006) Molecular recognition of histone H3 by the WD40 protein WDR5. *Nat. Struct. Mol. Biol. 13*, 698–703.
20. Han, Z. F., Guo, L., Wang, H. Y., Shen, Y., Deng, X. W., and Chai, J. J. (2006) Structural basis for the specific recognition of methylated histone H3 lysine 4 by the WD-40 protein WDR5. *Mol. Cell 22*, 137–144.
21. Schuetz, A., Allali-Hassani, A., Martin, F., Loppnau, P., Vedadi, M., Bochkarev, A., Plotnikov, A. N., Arrowsmith, C. H., and Min, J. R. (2006) Structural basis for molecular recognition and presentation of histone H3 By WDR5. *EMBO J. 25*, 4245–4252.
22. Dou, Y. L., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., Allis, C. D., and Roeder, R. G. (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol. 13*, 713–719.
23. Patel, A., Dharmarajan, V., and Cosgrove, M. S. (2008) Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *J. Biol. Chem. 283*, 32158–32161.
24. Patel, A., Vought, V. E., Dharmarajan, V., and Cosgrove, M. S. (2008) A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J. Biol. Chem. 283*, 32162–32175.
25. Suganuma, T., Pattenden, S. G., and Workman, J. L. (2008) Diverse functions of WD40 repeat proteins in histone recognition. *Genes Dev. 22*, 1265–1268.
26. Song, J. J., Garlick, J. D., and Kingston, R. E. (2008) Structural basis of histone H4 recognition by p55. *Genes Dev. 22*, 1313–1318.
27. Groisman, R., Polanowska, J., Kuraoka, I., Sawada, J., Saijo, M., Drapkin, R., Kisselev, A. F., Tanaka, K., and Nakatani, Y. (2003) The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell 113*, 357–367.
28. Scrima, A., Konickova, R., Czyzewski, B. K., Kawasaki, Y., Jeffrey, P. D., Groisman, R., Nakatani, Y., Iwai, S., Pavletich, N. P., and Thoma, N. H. (2008) Structural basis of UV DNA-damage recognition by the DDB1−DDB2 complex. *Cell 135*, 1213–1223.
29. Orlicky, S., Tang, X., Willems, A., Tyers, M., and Sicheri, F. (2003) Structural basis for phosphodependent substrate selection and orientation by the SCF^{Cdc4} ubiquitin ligase. *Cell 112*, 243–256.
30. Wu, G., Xu, G., Schulman, B. A., Jeffrey, P. D., Harper, J. W., and Pavletich, N. P. (2003) Structure of a β-TrCP1-Skp1-β-catenin complex. *Mol. Cell 11*, 1445–1456.
31. Higa, L. A., Wu, M., Ye, T., Kobayashi, R., Sun, H., and Zhang, H. (2006) CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat. Cell Biol. 8*, 1277–1283.
32. He, Y. J., McCall, C. M., Zeng, Y., and Xiong, Y. (2006) DDB1 functions as a linker to recruit receptor WD40 proteins to CUL4−ROC1 ubiquitin ligases. *Genes Dev. 20*, 2949–2954.
33. Hao, B., Oehlmann, S., Sowa, M. E., Harper, J. W., and Pavletich, N. P. (2007) Structure of a Fbw7-Skp1-cyclin E complex: multisite-phosphorylated substrate recognition by SCF ubiquitin ligases. *Mol. Cell 26*, 131–143.
34. Hu, J., Zacharek, S., He, Y. J., Lee, H., Shumway, S., Duronio, R. J., and Xiong, Y. (2008) WD40 protein FBW5 promotes ubiquitination of tumor suppressor TSC2 by DDB1−CUL4−ROC1 ligase. *Genes Dev. 22*, 866–871.
35. Ghosh, P., Wu, M., Zhang, H., and Sun, H. (2008) mTORC1 signaling requires proteasomal function and the involvement of CUL4-DDB1 ubiquitin E3 ligase. *Cell Cycle 7*, 373–381.
36. Choe, K. P., Przybysz, A. J., and Strange, K. (2009) The WD40 repeat protein WDR-23 functions with the CUL4/DDB1 ubiquitin ligase to regulate nuclear abundance and activity of SKN-1 in *Caenorhabditis elegans*. *Mol. Cell. Biol. 29*, 2704–2715.
37. Zou, H., Henzel, W. J., Liu, X., Lutschg, A., and Wang, X. (1997) Apaf-1, a human protein homologous to *C. elegans* CED-4, participates in cytochrome *c*-dependent activation of Caspase-3. *Cell 90*, 405–413.
38. Acehan, D., Jiang, X., Morgan, D. G., Heuser, J. E., Wang, X., and Akey, C. W. (2002) Three-dimensional structure of the apoptosome. *Mol. Cell 9*, 423–432.
39. Saeki, M., Irie, Y., Ni, L., Yoshida, M., Itsuki, Y., and Kamisaki, Y. (2006) Monad, a WD40 repeat protein, promotes apoptosis induced by TNF-α. *Biochem. Biophys. Res. Commun. 342*, 568–572.
40. Shi, Y. (2006) Mechanical aspects of apoptosome assembly. *Curr. Opin. Cell Biol. 18*, 677–684.
41. Itoh, T., Iwashita, S., Cohen, M. B., Meyerholz, D. K., and Linn, S. (2007) Ddb2 is a haploinsufficient tumor suppressor and controls spontaneous germ cell apoptosis. *Hum. Mol. Genet. 16*, 1578–1586.
42. Wall, M. A., Coleman, D. E., Lee, E., Iñiguez-Lluhi, J. A., Posner, B. A., Gilman, A. G., and Sprang, S. R. (1995) The structure of the G protein heterotrimer $G\alpha 1\beta 1\gamma 2$. *Cell 83*, 1047–1058.
43. Lambright, D. G., Sondek, J., Bohm, A., Skiba, N. P., Hamm, H. E., and Sigler, P. B. (1996) The 2.0 Å crystal structure of a heterotrimeric G protein. *Nature 379*, 311–319.
44. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235–242.
45. Wu, X. H., Zhang, H., and Wu, Y. D. (2010) Is Asp-His-Ser/Thr-Trp tetrad hydrogen-bond network important to WD40-repeat proteins: a statistical and theoretical study. *Proteins: Struct., Funct., Genet. 78*, 1186–1194.
46. Wilson, D., Madera, M., Vogel, C., Chothia, C., and Gough, J. (2007) The SUPERFAMILY database in 2007: families and functions. *Nucleic Acids Res. 35*, D308–D313.
47. Wu, X. H., and Wu, Y. D., unpublished data.
48. Higuera-Garcia, I., Gaitatzes, C., Smith, T. F., and Neer, E. J. (1998) Folding a WD repeat propeller: role of highly conserved aspartic acid residues in the G protein b subunit and Sec13. *J. Biol. Chem. 273*, 9041–9049.
49. Trievel, R. C., and Shilatifard, A. (2009) WDR5, a complexed protein. *Nat. Struct. Mol. Biol. 16*, 678–680.
50. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem. 182*, 319–326.
51. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D 50*, 760–763.
52. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D 60*, 2126–2132.
53. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D 53*, 240–255.
54. Laskowski, R. A. M. M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283–291.

55. Greenfield, N. J. (1999) Applications of circular dichroism in protein and peptide analysis. *Trends Anal. Chem. 18*, 236–244.
56. Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta 1751*, 119–139.
57. Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc. 1*, 2527–2535.
58. Greenfield, N. J. (2006) Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism. *Nat. Protoc. 1*, 2733–2741.
59. Pace, C. N., Horn, G., Hebert, E. J., Bechert, J., Shaw, K., Urbanikova, L., and Scholtz, J. M. (2001) Tyrosine hydrogen bonds make a large contribution to protein stability. *J. Mol. Biol. 312*, 393–404.
60. Takano, K., Scholtz, J. M., Sacchettini, J. C., and Pace, C. N. (2003) The contribution of polar group burial to protein stability is strongly context-dependent. *J. Biol. Chem. 278*, 31790–31795.
61. Pace, C. N., Treviño, S., Prabhakaran, E., and Scholtz, J. M. (2004) Protein structure, stability and solubility in water and other solvents. *Philos. Trans. R. Soc. London, Ser. B 359*, 1225–1235.
62. Pace, C. N. (2009) Energetics of protein hydrogen bonds. *Nat. Struct. Mol. Biol. 16*, 681–682.
63. Gao, J., Bosco, D. A., Powers, E. T., and Kelly, J. W. (2009) Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nat. Struct. Mol. Biol. 16*, 684–691.
64. Auton, M., and Bolen, D. W. (2005) Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc. Natl. Acad. Sci. U.S.A. 102*, 15065–15068.
65. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci. 4*, 2138–2148.
66. Courtenay, E. S., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (2000) Thermodynamic analysis of interactions between denaturants and protein surface exposed on unfolding: interpretation of urea and guanidinium chloride m-values and their correlation with changes in accessible surface area (ASA) using preferential interaction coefficients and the local-bulk domain model. *Proteins* No. Suppl. 4, 7–85.
67. Mason, P. E., Neilson, G. W., Enderby, J. E., Saboungi, M. L., Dempsey, C. E., MacKerell, A. D., Jr., and Brady, J. W. (2004) The structure of aqueous guanidinium chloride solutions. *J. Am. Chem. Soc. 126*, 11462–11470.
68. Godawat, R., Jamadagni, S. N., and Garde, S. (2010) Unfolding of hydrophobic polymers in guanidinium chloride solutions. *J. Phys. Chem. B 114*, 2246–2254.
69. Stagg, L., Samiotakis, A., Homouz, D., Cheung, M. S., and Wittung-Stafshede, P. (2010) Residue-specific analysis of frustration in the folding landscape of repeat β/α protein apoflavodoxin. *J. Mol. Biol. 396*, 75–89.
70. Han, Z., Xing, X., Hu, M., Zhang, Y., Liu, P., and Chai, J. (2007) Structural basis of EZH2 recognition by EED. *Structure 15*, 1306–1315.
71. Cerna, D., and Wilson, D. K. (2005) The structure of Sif2p, a WD repeat protein functioning in the set3 corepressor complex. *J. Mol. Biol. 351*, 923–935.
72. (a) Sprang, S., Standing, T., Fletterick, R. J., Stroud, R. M., Finer-Moore, J., Xuong, N.-H., Rutter, W. J., and Craik, C. S. (1987) The three-dimensional structure of Asn102 mutant of trypsin: role of Asp102 in serine protease catalysis. *Science 237*, 905–909. (b) Craik, C. S., Roczniak, S., Largman, C., and Rutter, W. J. (1987) The catalytic role of the active site aspartic acid in serine proteases. *Science 237*, 909–913.