Structure
Ways & Means

Methionine Scanning as an NMR Tool for Detecting and Analyzing Biomolecular Interaction Surfaces

Mira C. Stoffregen,[1] Matthias M. Schwer,[1] Fabian A. Renschler,[1] and Silke Wiesner[1,*]
[1]Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany
*Correspondence: silke.wiesner@tuebingen.mpg.de
DOI 10.1016/j.str.2012.02.012

SUMMARY

Methyl NMR spectroscopy is a powerful tool for studying protein structure, dynamics, and interactions. Yet difficulties with resonance assignment and the low abundance of methyl groups can preclude detailed NMR studies, particularly the determination of continuous interaction surfaces. Here we present a straightforward strategy that overcomes these problems. We systematically substituted solvent-exposed residues with reporter methionines in the expected binding site and performed chemical shift perturbation (CSP) experiments using methyl-TROSY spectra. We demonstrate the utility of this approach for the interaction between the HECT domain of the Rsp5p ubiquitin ligase and its cognate E2, Ubc4. Using these mutants, we could instantaneously assign all newly arising reporter methyl signals, determine the Ubc4 interaction surface on a per-residue basis, and investigate the importance of each individual mutation for ligand binding. Our data show that methionine scanning significantly extends the applicability, information content, and spatial resolution of methyl CSP experiments.

INTRODUCTION

All biological processes are founded on the proper organization of proteins in specific interaction networks. To understand cellular function and behavior on a mechanistic level, it is thus essential to understand how proteins recognize their binding partners, how binding specificity is conferred, and how protein function is controlled and modified through biomolecular interactions. NMR spectroscopy has proven to be a particularly powerful method for studying biomolecular interactions, because the measured chemical shifts are highly sensitive to changes in the local chemical environment of the observed atomic nuclei (Gao et al., 2004). Furthermore, unlike other biophysical techniques, chemical shift perturbation studies also capture low-affinity or transient interactions and yield site-specific information about residues located in binding sites. NMR binding studies, however, rely on the knowledge of the protein structure and the resonance assignment of the chemical shifts that are affected by ligand binding. Most commonly, chemical shift perturbation studies are performed on ^{15}N-labeled protein with an NMR-inactive binding partner by recording 2D ^{1}H,^{15}N-correlation spectra. In this case, the resonance assignments of the amide groups necessary for a per-residue mapping of the binding site can be obtained from multidimensional ^{1}H,^{15}N-based experiments. This typically requires that the protein is smaller than 50 kD and stable at concentrations in the hundreds of μM for several days. These restrictions have therefore considerably limited studies of large or unstable proteins by NMR spectroscopy.

More recently, methods have been developed for the ^{1}H,^{13}C-labeling of Ile, Leu, Val, Met, and Ala residues exclusively at their methyl positions in an otherwise highly deuterated background. Along with spectroscopic advances aimed at preserving only slowly relaxing magnetization, methyl labeling has enabled NMR studies of protein structure, dynamics, and interactions in high-molecular-weight systems (>>50 kD) (Ruschak and Kay, 2010). However, although binding interfaces are in general enriched in hydrophobic residues, the overall frequency of solvent-exposed aliphatic amino acids is too low to map continuous interaction surfaces using methyl NMR spectroscopy alone (Gelis et al., 2007; Hamel and Dahlquist, 2005; Wiesner et al., 2007). Another major bottleneck in the use of methyl groups as NMR probes is their laborious resonance assignment. Methyl groups are most commonly assigned from a large number of single-amino acid substitutions (Amero et al., 2011; Ogunjimi et al., 2010; Sprangers and Kay, 2007), NOESY and paramagnetic relaxation enhancement data (Gelis et al., 2007; Sprangers and Kay, 2007; Venditti et al., 2011), and "divide-and-conquer" strategies where the individual components of oligomeric or multidomain proteins are separately expressed and assigned by traditional ^{1}H,^{15}N-based experiments (Gelis et al., 2007; Ogunjimi et al., 2010; Sprangers and Kay, 2007).

Here we present an approach to using methyl NMR spectroscopy as a tool for determining and analyzing protein interaction surfaces with per-residue resolution. Instead of exclusively relying on naturally occurring methyl groups, we strategically introduced methionine substitutions of solvent-exposed residues as NMR reporters of ligand binding. This approach is generally applicable and has several advantages over previously described methods. First, the low natural occurrence of methionines in proteins (only ca. 2%) (Brooks et al., 2002) in combination with the unique chemical shifts of their methyl groups results in well-resolved and highly sensitive NMR spectra that can be acquired at low protein concentrations (in the tens of μM range). Second, the favorable spectroscopic properties of methionine methyl groups enable interaction studies of high-molecular-weight systems (>>100 kD) (Religa et al., 2010). Third, the resonance assignment of the introduced methionine methyl group is instantaneous, because the mutation results in an additional,

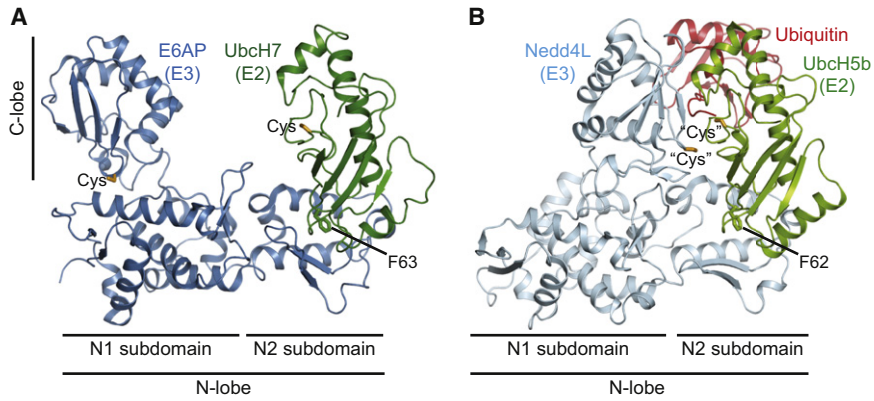

Figure 1. Structural Comparison of E2-HECT Domain Complexes

(A) Ribbon representation of the E6AP HECT domain (E3) (blue) in complex with its cognate E2 (green), UbcH7 (PDB code 1C4Z).

(B) As in (A) but for the Nedd4-like HECT domain in complex with a UbcH5b-ubiquitin oxyester (PDB ID code 3JW0). In the Nedd4L HECT-UbcH5b-ubiquitin complex, the catalytic cysteines were replaced by serines and are thus indicated in quotation marks. The catalytic cysteines, the HECT specificity-determining E2 residue F63 (F62 in UbcH5b), and the HECT subdomains are labeled.

easily identifiable peak in the spectrum. Fourth, continuous binding surfaces can be mapped on a per-residue basis through a series of substitutions of solvent-exposed residues. Fifth, systematic binding-site analyses are performed in a manner similar to alanine-scanning mutagenesis through the identification of functional epitopes as those methionine substitutions that severely interfere with or even abrogate ligand binding (Clackson and Wells, 1995). Last, in contrast to other methods used in combination with alanine scanning, naturally occurring methyl groups are independent internal NMR probes for the structural integrity of the mutated proteins.

In order to explore the utility of methionine-scanning mutagenesis as an NMR tool to define and analyze binding surfaces in a protein that is not amenable to $^{1}H,^{15}N$-based resonance assignment, we studied the 62 kD complex between the catalytic HECT domain of the ubiquitin ligase (E3) Rsp5p and its cognate E2, Ubc4. Ubiquitination is one of the most abundant posttranslational modifications in eukaryotes, where the 76 amino acid protein ubiquitin is attached to substrates by the sequential action of an activating (E1), a conjugating (E2), and a ligating (E3) enzyme (Hershko and Ciechanover, 1998). In contrast to other catalytic domains of E3s, HECT domains both bind the E2 enzyme and accept ubiquitin from E2 to form a HECT-ubiquitin thioester intermediate via a conserved Cys residue prior to transferring ubiquitin to a substrate lysine residue (Huibregtse et al., 1995). In general, a given E2 can function with various E3s, whereas each E3 only interacts with a distinct subset of E2s (Glickman and Ciechanover, 2002; Nuber and Scheffner, 1999). The large number of resulting possible E2-E3 combinations is thought to provide a mechanism for these enzymes to modulate their activities, recognize their bona fide substrates with high specificity, and conjugate different types of ubiquitin chains to ultimately generate a remarkably diverse array of signaling outcomes. Therefore, it is of considerable interest to elucidate the sequence determinants of E2-E3 interactions, such as the Rsp5p HECT domain-Ubc4 interaction studied here.

RESULTS

The Ubc4-Rsp5p HECT Domain Complex as a Test Case

Crystallographic studies have revealed that HECT domains adopt a bilobal fold, with the E2 binding site located on the N2 subdomain of the N lobe (Figure 1) (Huang et al., 1999; Kamadurai et al., 2009). Interestingly, loading of the E2 with ubiquitin induces a reorientation of the HECT subdomains, resulting in a limited number of additional contacts between the E2 and the region of the catalytic loop in the HECT domain C lobe (Figure 1B) (Kamadurai et al., 2009). This supports the notion that conformational flexibility underlies the catalytic activity of HECT domains (Verdecia et al., 2003). Because no structural information is available for the yeast Rsp5p-Ubc4 complex, we wished to characterize this interaction in detail. Moreover, we sought to address whether the E2-HECT C lobe interaction may be preformed in solution even in the absence of ubiquitin. Due to its tendency to aggregate at high μM concentrations, the Rsp5p HECT domain was not amenable to $^{1}H,^{15}N$-based backbone resonance assignment in our hands. To structurally characterize the interaction between the 45 kD Rsp5p HECT domain and the 17 kD Ubc4 E2 enzyme, we therefore performed chemical shift perturbation studies by recording 2D methyl-TROSY (HMQC) spectra. To this end, we used a 40 μM Rsp5p HECT domain sample that was $^{1}H,^{13}C$-labeled at the Ile δ_1- and Met ε-methyl positions but otherwise U-$^{2}H,^{12}C$-labeled. As shown in Figure 2A, two of the ten methionine resonances in the WT Rsp5p HECT domain exhibited significant chemical shift changes upon stepwise addition of unlabeled (NMR-inactive) Ubc4. In contrast, only 1 out of 22 resonances in the isoleucine region of the HMQC spectrum showed a small chemical shift change (Figure 2A). Taken together, these results demonstrate that the Rsp5p HECT domain and Ubc4 form a complex in solution that can be detected by methyl NMR spectroscopy.

Based on the crystal structures of other E2-HECT domain complexes that have been solved previously (Huang et al., 1999; Kamadurai et al., 2009), we expected that the main interactions with Ubc4 are mediated by the Rsp5p N2 subdomain (Figure 1), which contains two Met and three Ile residues (Figure 2B). To identify which of these methyl resonances are affected by the presence of the Ubc4 ligand, we introduced single-amino acid substitutions of the two methionine (M584I and M598V) and the three isoleucine (I603V, I633V, and I652L) residues in the N2 subdomain and compared their HMQC spectra to that of the WT Rsp5p HECT domain (see Figure S1 available online). Using this mutagenesis approach, all the methionine ε- and isoleucine δ_1-methyl groups in the N2 subdomain could be readily assigned by the absence of the peak of interest. With these assignments in hand, we can show that the methyl

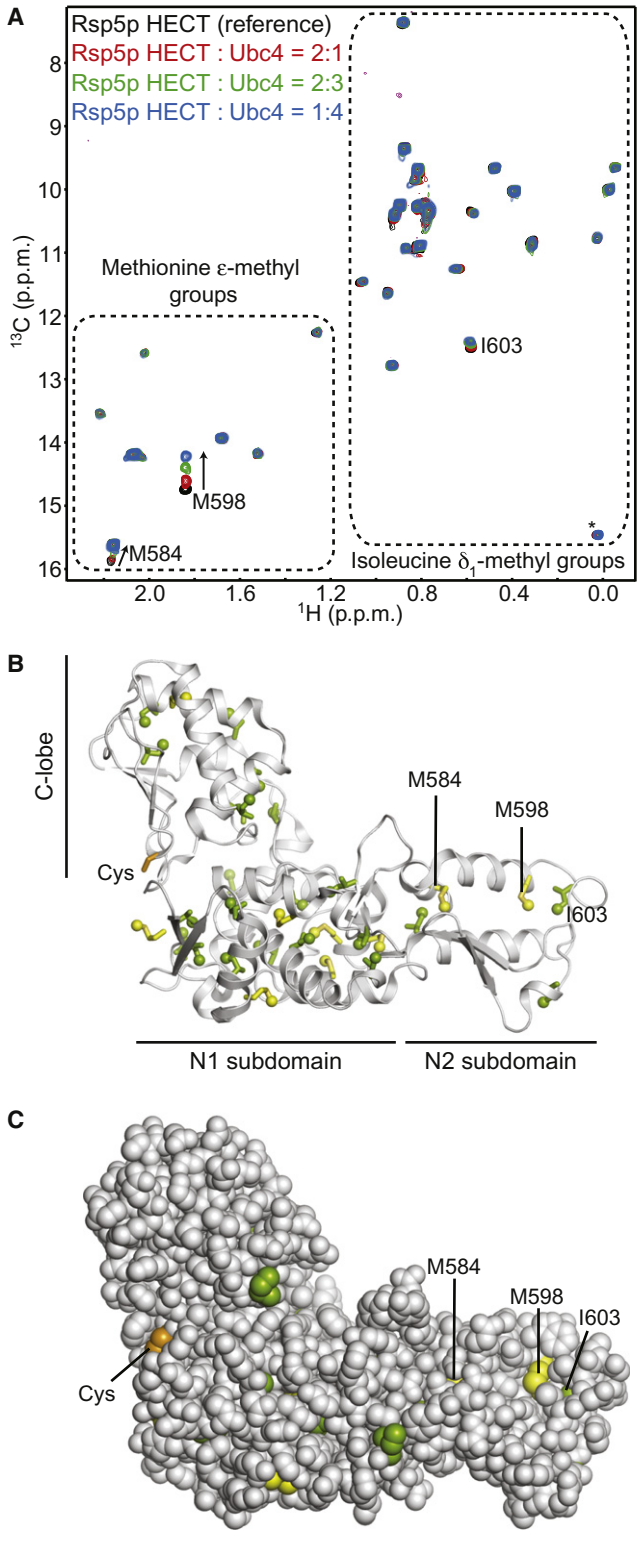

Figure 2. The Rsp5p HECT Domain Interacts with Ubc4

(A) Methyl-TROSY spectra of a uniformly Met-[ε $^{13}CH_3$], Ile-[δ_1 $^{13}CH_3$]-labeled WT Rsp5p HECT domain in the absence (black; reference spectrum) and presence of increasing amounts of unlabeled (and hence NMR-inactive) Ubc4.

(B) Ribbon representation of the Rsp5p HECT domain (PDB code 3OLM) highlighting the naturally occurring methionine (yellow) and isoleucine (green) residues. Met and Ile residues exhibiting chemical shift changes upon Ubc4 binding (M584, M598, and I603) are labeled.

(C) As in (B) but as a sphere representation demonstrating that M584 and I603 are buried in the core of the N2 subdomain.

See also Figure S1.

chemical shifts of M584, M598, and I603 are perturbed upon Ubc4 binding. Hence, the Ubc4 binding region is indeed located on the N2 subdomain, as has been observed in other E2-HECT domain complexes (Huang et al., 1999; Kamadurai et al., 2009) (Figure 2B). However, when we mapped these residues onto the Rsp5p HECT domain structure, we observed that these residues do not define a continuous E2 binding pocket (Figure 2C). Moreover, the side chains of M584 and I603 are buried in the core of the N2 subdomain (the relative solvent-accessible surface area equals 0% for both M584 and I603) (Figure 2C). This suggests that, in contrast to M598, the M584 and I603 side chains do not directly contribute to Ubc4 binding but rather are reporters of Ubc4 binding due to indirect effects (secondary chemical shift changes). It should also be noted that the Rsp5p C lobe lacks methionine and isoleucine residues at or in the vicinity of the positions equivalent to those that form the additional contacts of the C lobe with UbcH5b in the Nedd4L complex. Due to this lack of NMR probes in this region of the Rsp5p HECT domain, this experiment can thus not rule out the existence of additional contacts between Ubc4 and the C lobe.

Methionine Substitutions as NMR Reporters for Ligand Binding

Although the naturally occurring methyl groups in the Rsp5p HECT domain are sufficient as NMR reporters to detect the interaction with the Ubc4 enzyme, this approach fails to provide detailed information about the amino acids constituting the binding pocket and their individual importance for Ubc4 binding. To better define the sequence determinants of the Rsp5p HECT domain-Ubc4 interaction and to test a potential involvement of the C lobe, we systematically introduced Met residues in the Rsp5p HECT domain as NMR reporters of ligand binding. To this end, we mutated a series of solvent-exposed positions in the area laid out by the M584, M598, and I603 residues in the N2 subdomain that we have identified above as being affected by Ubc4 binding (Figure 2). In addition, we substituted two amino acids in the C lobe (A765 and F778) at positions equivalent to those that interacted with UbcH5b in the Nedd4L complex. In total, we individually substituted 19 amino acids to Met (E585, E590, N593, S594, W597, N601, V606, L607, D608, T610, S612, D614, E616, V621, T623, Y643, Y647, A765, and F778) (Figure 3A; Figure S2) and examined the Ubc4 binding properties of each of these mutants by chemical shift perturbation experiments.

In order to have a sufficient number of evenly distributed methyl groups as NMR reporters not only for ligand binding but also for the structural integrity of the designed mutants, we used a U-^{2}H, $^{13}CH_3$-labeling scheme for all mutants where both the ε-methyl groups of methionines and the δ_1-methyls of isoleucines were labeled. Comparison of the HMQC spectra of the individual mutant Rsp5p HECT domains to the WT spectrum showed that the methyl resonances of the newly introduced

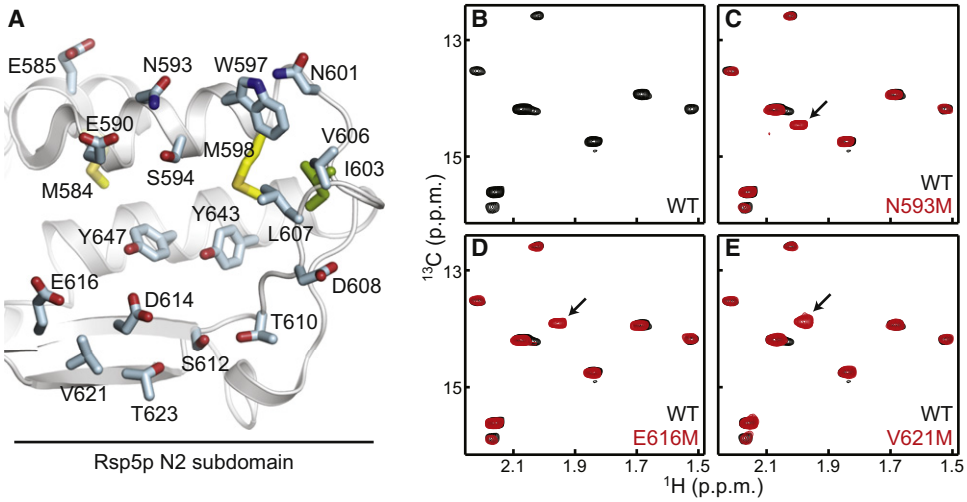

Figure 3. Strategic Placement of Reporter Methionine Residues for Binding Site Mapping by NMR

(A) Ribbon representation of the N2 subdomain of the Rsp5p HECT domain highlighting mutated residues.

(B) Representative region of the methyl-TROSY spectrum of the 45 kD WT Rsp5p HECT domain.

(C–E) Overlay of methyl-TROSY spectra of the WT and reporter methionine mutant Rsp5p HECT domain showing the additional peak resulting from Met substitution. The fact that apart from the additional peak the spectra (also in the isoleucine region; not shown) are virtually identical demonstrates that these mutants are properly folded.

See also Figure S2.

reporter methionines (M_{mut}) could be easily identified in all mutants (Figures 3B–3E). Apart from the presence of the single additional methionine peak, all mutant spectra are very similar to the WT spectrum, demonstrating that all mutants were properly folded. Having established that all reporter methionines could be instantaneously assigned and that all 19 mutants are properly folded, we probed the Ubc4-Rsp5p HECT domain interaction with each of these mutants. This approach thus significantly increases the number of NMR probes in the region of interest and thereby the information content and spatial resolution of this chemical shift perturbation study.

Characterization of the Binding Properties of the Methionine Mutants

In order to investigate the ligand binding capabilities of the individual methionine substitutions, we performed single-point chemical shift perturbation experiments. To this end, we recorded HMQC spectra of all Met-substituted U-^{2}H, Met-[ε ^{13}CH$_3$]-, Ile-[δ_1 ^{13}CH$_3$]-labeled Rsp5p HECT domains in the absence and presence of a 2-fold stoichiometric excess of unlabeled Ubc4 (Figure 4). In all experiments, the chemical shift perturbations of the naturally occurring M584 and M598 served as independent internal references for the Ubc4 binding properties of the individual mutants. For comparison, the methyl-HMQC spectra of the WT Rsp5p HECT domain in the absence and presence of a 2-fold excess of unlabeled Ubc4 are shown in Figure S3.

Overall, we anticipated three different scenarios of ligand binding. First, the mutated residue lies outside the binding pocket (Figure 4A). In this case, the chemical shift of the reporter Met (M_{mut}) should remain largely unchanged upon addition of the ligand, and the chemical shifts of the naturally occurring Met residues in the binding pocket (M584 and M598) should be as perturbed as in the WT protein. Second, the mutated residue is part of the binding pocket but not key to binding (Figure 4B). Then, the binding affinity of the ligand should be comparable to that of the WT protein and, hence, the chemical shift changes of the naturally occurring Met residues in the binding pocket. At the same time, however, the reporter methionine should exhibit significant chemical shift perturbations, as it experiences the presence of the ligand. Finally, the mutated residue may be crucial for the interaction (a binding hot spot) (Figure 4C). Then, the mutation would significantly interfere with or completely abrogate ligand binding. This should result in only minor or no chemical shift perturbations of both the reporter Met and the naturally occurring Met methyl groups.

Indeed, among the 19 mutant HECT domains that we have studied, we observed all three types of binding behavior, indicating that we sampled a large portion of the binding surface (Figure 4). To analyze our chemical shift perturbation experiments in a more quantitative fashion, we calculated the average chemical shift perturbations ($\Delta\delta_{Av}$) of M_{Mut}, M584, and M598 for the WT and mutant Rsp5p HECT domains (Figure 5A) in order to characterize their Ubc4 binding capabilities. We found that 6 (E585M, N601M, V606M, T610M, A765, and F778) out of the 19 mutants exhibited no or only very small chemical shift changes (less than a peak width) for M_{Mut}, whereas the resonances of M584 and M598 shifted significantly (more than one peak width) as observed for the WT protein. We thus conclude that these residues are not part of the Ubc4 binding pocket. In contrast, for eight mutants (E590M, N593M, D608M, S612M, E616M, V621M, T623M, and Y643M), addition of Ubc4 to the mutant proteins resulted in significant chemical shift changes of M_{mut}, M584, and M598 reporting on the fact that these mutant residues are located within the Ubc4 binding pocket but are not key to Ubc4 binding. Last, five mutants (S594M, W597M,

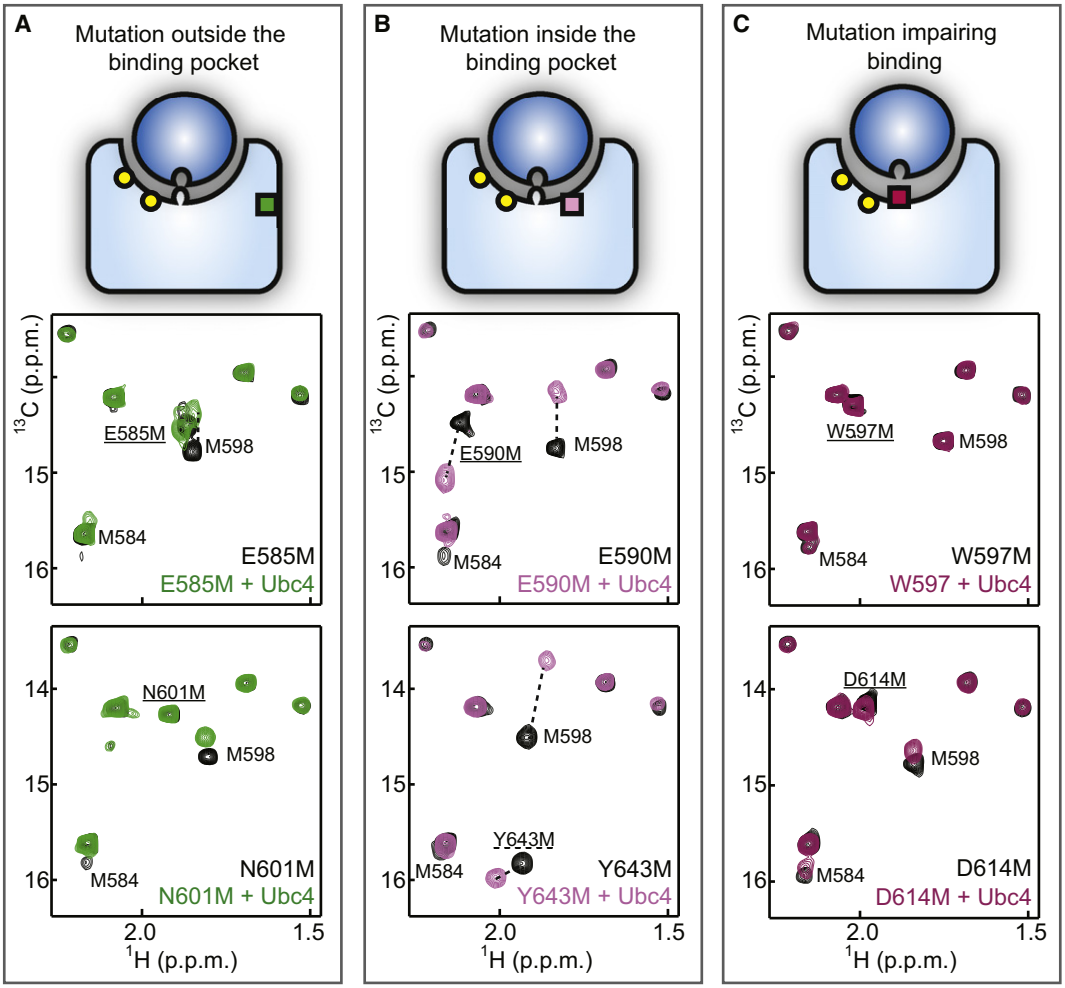

Figure 4. Effects of Methionine Mutations on Ligand Binding

(A) The residue mutated to methionine (green square) is located outside the Ubc4 binding pocket. Top: schematic of the interaction, with yellow circles representing naturally occurring Met residues located in the E2 binding site that are used as internal references for the ligand binding properties of the mutants. The square represents the introduced reporter methionine. Middle and bottom: overlay of a representative region of mutant ^{1}H,^{13}C-HMQC spectra in the presence and absence (black) of a 2-fold stoichiometric excess of unlabeled Ubc4.

(B) The methionine mutant (pink square) is located in the binding site but does not significantly reduce ligand binding affinity. Panels are otherwise as in (A).

(C) The methionine mutation (magenta square) strongly impairs or completely abolishes ligand binding. In this case, no significant chemical shift changes are observed for the mutated Met or the two naturally occurring methionines in the Ubc4 binding site (M584 and M598). Panels are otherwise as in (A).

See also Figure S3.

L607M, D614M, and Y647M) showed only very small or virtually no chemical shift changes upon Ubc4 addition for all Met and Ile residues, indicating that these mutations severely impair binding and hence that the mutated residues can be considered crucial determinants of Ubc4 binding to the Rsp5p HECT domain. It is important to note that all of these loss-of-function mutants are structurally intact, because their HMQC spectra are essentially identical to that of the WT protein except for the presence of the additional methionine reporter peak (Figure 4C).

Determination and Analysis of the Ubc4-Rsp5p HECT Domain Interaction Surface

Previous crystal structures have revealed that the primary E2-HECT domain contact surface is created by a large hydrophobic groove on the N2 subdomain that tightly surrounds the E2 residues F63 and P96 (Huang et al., 1999; Kamadurai et al., 2009). Whereas P96 is common to all E2 enzymes, F63 is conserved only in E2 enzymes that preferentially interact with HECT-type E3s (Eletr and Kuhlman, 2007; Nuber and Scheffner, 1999). In fact, in previous studies, F63 has been identified as an E2 residue that is essential for E2 binding to Rsp5p and E6AP (Eletr and Kuhlman, 2007; Nuber and Scheffner, 1999). Because no structural information is available for the Rsp5p-Ubc4 complex, we generated a structural model of the Ubc4-Rsp5p HECT domain complex to evaluate our results (Figures 5B and 5C). In agreement with our data, the naturally occurring M598, which showed the most significant chemical shift perturbation in the WT Rsp5p HECT domain, is situated close to both Ubc4 residues F63 and

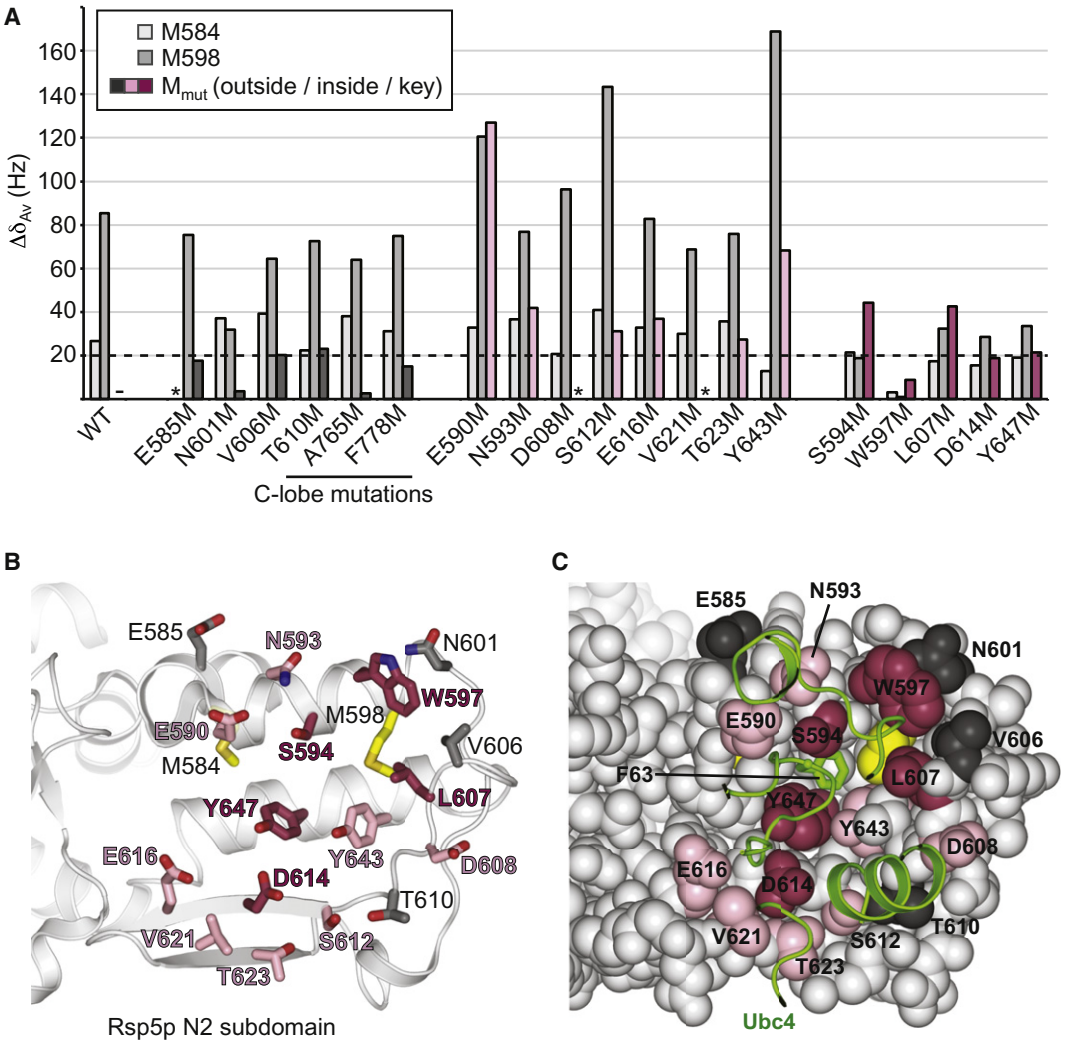

Figure 5. Classification of Ubc4 Binding Capabilities of the Rsp5p HECT Domain Mutants and Mapping of the Binding Site

(A) Average chemical shift perturbations ($\Delta\delta_{Av} = ((\Delta\delta(^1H))^2 + (\Delta\delta(^{13}C))^2)^{1/2}$) of M584 (light gray), M598 (gray), and M_{mut} (dark gray, pink, and magenta) observed in the WT Rsp5p HECT domain and individual methionine-substituted proteins upon addition of a 2-fold stoichiometric excess of Ubc4.

(B and C) Ribbon (B) and sphere (C) representation of the Rsp5 HECT domain color coded corresponding to the Ubc4 binding properties of the Rsp5p HECT domain mutants.

See also Figure S4.

P96. In contrast to M598, the naturally occurring M584 and I603 are buried directly underneath the edge of the Ubc4 binding pocket.

When we mapped the methionine substitutions according to their Ubc4 binding capabilities onto the Rsp5p HECT domain structure (Figure 5C; Figure S4A), we found that four (S594M, W597M, L607M, and Y647M) out of the five residues that severely interfere with the Ubc4 interaction are located in the hydrophobic groove that is centered around the Ubc4 residues F63 and P96. Whereas Y647 is involved in stacking interactions with the aromatic ring of F63, S594 forms a hydrogen bond with the F63 carbonyl. In agreement with our findings, in a previous alanine-scanning study (Eletr and Kuhlman, 2007), mutation of the equivalent positions in E6AP led to a significant destabilization of the E6AP-UbcH7 complex (Figure S4). The Rsp5p residues W597 and L607 form hydrophobic interactions with the conserved P96 residue in Ubc4. An aliphatic amino acid at the L607 position is highly conserved in all HECT domains, and the equivalent position (M653) is also a binding hot spot in the E6AP-UbcH7 complex (Figure S4) (Eletr and Kuhlman, 2007). Of all the mutants that we examined, the W597M mutation in Rsp5p showed the strongest effect on Ubc4 binding and completely abolished the interaction with Ubc4. This may be explained by its involvement in hydrogen bonding with the Ubc4 residue S95 and in van der Waals contacts with the Ubc4 residues P96 and A97. The tryptophan in position 597 in Rsp5p is strictly conserved only in C2-WW-HECT (Nedd4-type) E3s, such as Rsp5p and Nedd4L, as are S95 and P97 in all Ubc4/5 E2s that have been shown to preferentially interact with Nedd4-like ligases (Figure S4). In other HECT-type E3s such as

E6AP, HERC, and HECTD proteins, W597 is replaced by a glutamine or asparagine that in the case of E6AP forms a salt bridge with K96 in UbcH7 (Huang et al., 1999). As noticed previously, the sequence differences at these positions and their importance for the interaction likely reflect the role of these interaction pairs in determining E2-HECT domain specificity (Kamadurai et al., 2009). The only residue that we identified as crucial for the Ubc4 interaction that is not located in the hydrophobic groove is D614. This residue makes contacts to the N terminus, the L2 loop, and the specificity-determining L4 loop of Ubc4. However, the fact that the polar character of this position is well conserved among all HECT domains suggests that it is not specificity determining.

All of the eight mutations that exhibited significant chemical shift perturbations of M_{mut} but did not impair the interaction (E590M, N593M, D608M, S612M, E616M, V621M, T623M, and Y643M) are located close to Ubc4 and surround the HECT-domain residues identified as crucial binding determinants. As shown in Figure 4, we are thus able to map a continuous binding surface with a per-residue resolution and assess the functional importance of individual amino acids within the binding pocket. In contrast, the mutations that were classified as being situated outside the Ubc4 binding pocket (E585M, N601M, V606M, T610M, A765, and F778) all lie on the outer edge of the Ubc4 binding pocket or, in the case of A765 and F778, on the C lobe (Figure S2A). This suggests that despite the conformational flexibility of the HECT domain, the additional binding pocket on the C lobe is not preformed in solution in the absence of the E2-ubiquitin thioester.

Overall, we note that our NMR-based classification of the Met mutants is in excellent agreement with the structural model of the Ubc4-Rsp5p HECT domain complex. Most of the residues that we identified as crucial Rsp5p-Ubc4 interaction determinants also contribute substantially to the binding energy of the UbcH7-E6AP complex (Eletr and Kuhlman, 2007). Although many of these crucial residues are conserved among all types of HECT domains, our data strengthen the suggestion that the position equivalent to W597 in Rsp5p may play an essential role in conferring specificity to the interactions between Nedd4-family E3s and their cognate E2s (Kamadurai et al., 2009).

Taken together, our results show that by introducing methionines as NMR reporters, we can obtain an instantaneous resonance assignment of the newly arising methyl peaks and that these mutants can be used to map continuous binding surfaces and identify crucial interaction determinants. This approach thus significantly extends the scope of biomolecular interaction studies using methyl chemical shift perturbation experiments.

DISCUSSION

One fundamental concept in NMR spectroscopy is to obtain resonance assignments for as many probes as possible in order to maximize site-specific information content. However, this approach rapidly becomes a veritable challenge in high-molecular-weight systems due to line broadening and spectral overlap or, in the case of methyl NMR spectroscopy, due to a lack of straightforward assignment strategies (Ruschak and Kay, 2010). On the other hand, reducing spectral complexity, for example by using methyl groups, bears the risk that the region of interest may contain no NMR probes (and hence be NMR invisible) or too few NMR probes to achieve the desired level of detail. As such, the future trend for the study of protein systems that are not amenable to H,N-based spectroscopy may rather be to introduce as many methyl NMR reporters at particular sites of interest as necessary to address the questions at hand. A first step in this direction has been undertaken recently with the introduction of a small number of cysteines at sites of interest into two supramolecular systems (Religa et al., 2011). These cysteines were then chemically modified to form S-methylthiocysteine as a unique methyl NMR probe to report on the conformational dynamics of these high-molecular-weight complexes.

To overcome the two major drawbacks of methyl NMR spectroscopy, namely the difficulty of resonance assignment and the potentially inadequate number of methyl probes, and yet obtain highly detailed structural information, we introduced a series of reporter methionines to map the Ubc4 binding pocket on the Rsp5p HECT domain on a per-residue basis (Figure 3A). The choice of using methionine substitutions, as opposed to other methyl-containing or chemically modified amino acids, is based on a number of reasons. First and foremost, the average frequency of naturally occurring methionines in proteins is much lower (~2%) than for other methyl-containing amino acids (Ala: ~8%; Leu: ~9%; Ile: ~7%; Val: ~8%) (Brooks et al., 2002). Therefore, the methionine region is the most resolved in a methyl-HMQC spectrum and, thus, the most suitable for a straightforward identification of the additional peak arising from the introduced NMR reporter. As we have shown, all methionine-substitution mutants yielded well-resolved, high-sensitivity methyl NMR spectra that enabled an immediate resonance assignment of the reporter methionine (Figure 3B). Second, the incorporation of methionines into proteins is highly efficient and can be achieved simply by addition of the comparatively inexpensive $^{1}H,^{13}C$-methyl-labeled amino acid to the growth medium (Gelis et al., 2007). Moreover, diversion of the methyl isotope label into other amino acids (scrambling) does not occur for methionines. In contrast, methyl labeling of Ala side chains necessitates the use of various perdeuterated additives to prevent scrambling (Ayala et al., 2009). Last, whereas the line widths of branched amino acid methyl signals may be adversely affected by rotameric jumps, the high degree of rotational freedom of the long, unbranched methionine side chain gives rise to very favorable NMR relaxation properties. Therefore, our approach can be expected to be highly effective for studies of supramolecular-weight complexes (>500 kD) (Religa et al., 2010, 2011) and systems of limited stability or solubility, as demonstrated here for the Rsp5p HECT domain. In this regard, it is worth noting that the applicability of the presented method could be extended even further by using sparse-sampling and fast-pulsing NMR schemes such as SOFAST-methyl-TROSY experiments (Amero et al., 2009).

We have shown here that a per-residue mapping of the interaction surface (as is routinely achieved in H,N-based studies) can be accomplished using methyl chemical shift perturbation experiments through the introduction of a series of methionine substitutions (Figures 4 and 5). This significantly increases the spatial resolution of methyl NMR binding studies. In general, neither a resonance assignment of the naturally occurring methyl

groups nor a prior knowledge of the binding site is required for the utility of our approach. A careful analysis of sequence conservation and surface properties of the proteins of interest in combination with in silico predictions of interaction hot spots should permit the design of a limited number of mutations for an initial identification of binding sites. The identified binding site can then be scanned with higher resolution in a second step through the generation of additional methionine mutants. However, if complex formation induces substantial, previously unknown conformational changes or if the interaction surface is discontinuous, it may be required to perform a more extensive methionine screening and/or an initial NMR titration experiment with I-, L-, V-, A-, M-methyl-labeled WT protein and a subsequent assignment of at least those methyl resonances that experience chemical shift changes upon binding.

In general, protein interactions are not random, but highly specific. Although three-dimensional structures of protein complexes or binding site mapping by NMR spectroscopy provide spatial details about residues located at the interaction surface, these studies are insufficient to determine which contacts are primarily responsible for the strength and specificity of the interaction. Moreover, although biomolecular interfaces often bury large surfaces, it has been recognized that only a subset of the residues at the contact area contributes the majority of the binding energy (Clackson and Wells, 1995). The identification of these functional epitopes is fundamental to understanding protein function, albeit experimentally difficult to achieve. Most studies in this area are based on alanine-scanning mutagenesis, where functional epitopes are identified as those residues that, when mutated to alanine, considerably impair ligand binding (Clackson and Wells, 1995; DeLano, 2002). In addition, computational methods have been developed to identify hot spots in biomolecular interfaces. However, these methods are faced with the challenging complexity of binding events and rely on the availability of experimental data and the knowledge of high-resolution structural data of protein complexes.

Our method combines the advantages of alanine scanning (to identify residues crucial for binding) and the power of methyl-TROSY NMR spectroscopy (that is applicable to high-molecular-weight systems and can detect even weak or transient interactions). By examining the ligand binding capabilities of the individual methionine-substituted proteins based on their chemical shift perturbations, the mutations can be classified as being (1) located outside the binding interface (no chemical shift changes of the reporter methionine), (2) located inside the binding pocket (significant chemical shift perturbation of the reporter methionine), or (3) key to binding (complete loss of the interaction) (Figure 5). To distinguish between the first and last scenarios, the method exploits the presence of naturally occurring methyl groups (or additional NMR reporters) as internal, but independent, indicators of binding. In addition, the naturally occurring methyl groups report on the structural integrity of the methionine-substituted proteins (Figure 3B). As we have shown here, this allows a clear assessment as to whether a mutation interferes with ligand binding or whether it leads to an altered or unfolded conformation of the mutant protein and therefore abolishes the interaction. Methionine-scanning mutagenesis thus permits the study of protein interactions at the resolution of individual amino acid residues and enables a classification of the substitution mutants with respect to their importance for ligand binding.

In summary, we present methionine-scanning mutagenesis in combination with methyl chemical shift perturbation experiments as a straightforward and generally applicable approach to identifying sequence and structural requirements for ligand recognition on a per-residue basis. The knowledge gained from this type of study will further our understanding of the crucial determinants of protein interactions and thereby also help to design complexes of high affinity and specificity as well as small-molecule compounds for pharmaceutical purposes. However, this methodology will not be limited to binding studies, but will be equally valuable to probing protein structure and dynamics. Overall, methionine scanning thus significantly extends the utility, spatial resolution, and information content of methyl NMR studies.

EXPERIMENTAL PROCEDURES

Constructs, Mutagenesis, and Protein Purification

Saccharomyces cerevisiae Ubc4 was cloned by PCR from plasmid DNA (provided by D. Rotin, Hospital for Sick Children, Toronto, Canada), whereas the *S. cerevisiae* Rsp5p HECT domain (amino acids 425–809) was cloned from an Rsp5p full-length construct (Wiesner et al., 2007). Both DNA fragments were ligated into pProEx HTb vectors (Invitrogen) for recombinant expression of His_6-tagged proteins containing a tobacco etch virus (TEV) protease cleavage site directly C-terminal to the His_6 tag. Single-point mutations were introduced into the Rsp5p HECT domain DNA sequence using the QuikChange Site-Directed Mutagenesis protocol (Stratagene). All clones were verified by DNA sequencing. For resonance assignment, all Ile and Met residues in the N2 subdomain were mutated individually (M584I, M598V, I603V, I633V, and I652L), whereas for Ubc4 binding site mapping, in total 19 methionine point mutations were introduced into the Rsp5p HECT domain (E585M, E590M, N593M, S594M, W597M, N601M, V606M, L607M, D608M, T610M, S612M, D614M, E616M, V621M, T623M, Y643M, Y647M, A765M, and F778M).

Unlabeled Ubc4 was expressed in Luria Broth medium using *Escherichia coli* BL21-CodonPlus (DE3) RIL cells (Stratagene), whereas U-^{2}H, Met-[ε $^{13}CH_3$]-, Ile-[δ_1 $^{13}CH_3$]-labeled Rsp5p HECT domain was obtained by overexpression in *E. coli* BL21-CodonPlus (DE3) RIL cells in 150 ml 100% D_2O M9 minimal medium containing 0.3 g $^{2}H,^{12}C$-labeled glucose as the main source of carbon and 0.0375 g $^{14}NH_4Cl$ as the sole source of nitrogen. Approximately 1 hr before induction, 9 mg $^{13}CH_3$-labeled α-ketobutyrate (Sigma) after proton/deuterium exchange at the C3 position (Goto et al., 1999) and 15 mg $^{1}H,^{13}C$-ε-labeled methionine (Sigma) (Gelis et al., 2007) were added to the medium. Protein expression was induced at 25°C with 1 mM IPTG at an OD_{600} of 0.8–1.0. The cells were harvested ca. 14 hr after induction and then lysed by sonication in 1× PBS containing 15 mM imidazole and 1 mM DTT. Both proteins (Ubc4 and the Rsp5p HECT domain) were purified by Ni-affinity chromatography followed by TEV protease cleavage. After cleavage, Ubc4 was separated from His_6-tagged cleavage products and His_6-tagged TEV protease by a second Ni-affinity chromatography step followed by size-exclusion gel filtration. Both the Rsp5p HECT domain and Ubc4 were exchanged into NMR buffer (99% D_2O, 20 mM sodium phosphate [pD 6.5], 150 mM NaCl, 1 mM DTT, and 0.03% NaN_3) for chemical shift-mapping experiments. It should be noted that the Rsp5p HECT domain construct contains one additional methionine resulting from the NcoI restriction site.

NMR Spectroscopy

For chemical shift perturbation experiments, 2D $^{1}H,^{13}C$-methyl-TROSY (HMQC) spectra of 40 μM U-^{2}H, Met-[ε $^{13}CH_3$]-, Ile-[δ_1 $^{13}CH_3$]-labeled WT and mutant Rsp5p HECT domain samples were recorded before and after addition of a 2-fold stoichiometric excess of unlabeled Ubc4. The NMR data were collected at 25°C in 1 hr with an acquisition time of 28 ms and a spectral width of 10.5 ppm in the ^{13}C dimension on an 800 MHz Bruker Avance-III

spectrometer equipped with a room temperature probe head. All NMR data were processed and analyzed using the NMRPipe/NMRDraw program suite (Delaglio et al., 1995) and visualized with NMRView (OneMoonScientific).

Other Methods

Solvent-accessible surface areas were calculated using NACCESS version 2.1 (Hubbard and Thornton, 1993). A homology model of Ubc4 was generated with MODELLER (Eswar et al., 2008). The Ubc4-Rsp5p HECT domain complex was modeled by superposition of the N2 subdomain of the Rsp5p HECT domain (Protein Data Bank [PDB] code 3OLM) with those of the Nedd4L and E6AP HECT domains that are in complex with their cognate E2 enzymes (PDB codes 3JW0 and 1C4Z). As a next step, the Ubc4 homology model was structurally aligned with the E2 enzymes (UbcH5b and UbcH7) in the Nedd4L and E6AP complexes to get a structural model of the Ubc4-Rsp5p HECT domain complex (Huang et al., 1999; Kamadurai et al., 2009; Kim et al., 2011). All structure representations were generated with PyMOL (Schrödinger, LLC).

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and can be found with this article online at doi:10.1016/j.str.2012.02.012.

ACKNOWLEDGMENTS

We are grateful to Remco Sprangers and Andrei Lupas for critical remarks and discussions. This work was supported by the Max Planck Society and a Marie Curie Reintegration grant to S.W. (PIRG04-GA-2008-239418 UBLIGATION).

Received: December 30, 2011
Revised: January 27, 2012
Accepted: February 3, 2012
Published: April 3, 2012

REFERENCES

Amero, C., Schanda, P., Durá, M.A., Ayala, I., Marion, D., Franzetti, B., Brutscher, B., and Boisbouvier, J. (2009). Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies. J. Am. Chem. Soc. *131*, 3448–3449.

Amero, C., Asunción Durá, M., Noirclerc-Savoye, M., Perollier, A., Gallet, B., Plevin, M.J., Vernet, T., Franzetti, B., and Boisbouvier, J. (2011). A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. J. Biomol. NMR *50*, 229–236.

Ayala, I., Sounier, R., Usé, N., Gans, P., and Boisbouvier, J. (2009). An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. J. Biomol. NMR *43*, 111–119.

Brooks, D.J., Fresco, J.R., Lesk, A.M., and Singh, M. (2002). Evolution of amino acid frequencies in proteins over deep time: inferred order of introduction of amino acids into the genetic code. Mol. Biol. Evol. *19*, 1645–1655.

Clackson, T., and Wells, J.A. (1995). A hot spot of binding energy in a hormone-receptor interface. Science *267*, 383–386.

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

DeLano, W.L. (2002). Unraveling hot spots in binding interfaces: progress and challenges. Curr. Opin. Struct. Biol. *12*, 14–20.

Eletr, Z.M., and Kuhlman, B. (2007). Sequence determinants of E2-E6AP binding affinity and specificity. J. Mol. Biol. *369*, 419–428.

Eswar, N., Eramian, D., Webb, B., Shen, M.Y., and Sali, A. (2008). Protein structure modeling with MODELLER. Methods Mol. Biol. *426*, 145–159.

Gao, G., Williams, J.G., and Campbell, S.L. (2004). Protein-protein interaction analysis by nuclear magnetic resonance spectroscopy. Methods Mol. Biol. *261*, 79–92.

Gelis, I., Bonvin, A.M., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A., and Kalodimos, C.G. (2007). Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell *131*, 756–769.

Glickman, M.H., and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol. Rev. *82*, 373–428.

Goto, N.K., Gardner, K.H., Mueller, G.A., Willis, R.C., and Kay, L.E. (1999). A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated ^{15}N-, ^{13}C-, ^{2}H-labeled proteins. J. Biomol. NMR *13*, 369–374.

Hamel, D.J., and Dahlquist, F.W. (2005). The contact interface of a 120 kD CheA-CheW complex by methyl TROSY interaction spectroscopy. J. Am. Chem. Soc. *127*, 9676–9677.

Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. Annu. Rev. Biochem. *67*, 425–479.

Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P.M., Huibregtse, J.M., and Pavletich, N.P. (1999). Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. Science *286*, 1321–1326.

Hubbard, S.J., and Thornton, J.M. (1993). NACCESS computer program. Department of Biochemistry and Molecular Biology, University College London.

Huibregtse, J.M., Scheffner, M., Beaudenon, S., and Howley, P.M. (1995). A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. Proc. Natl. Acad. Sci. USA *92*, 2563–2567.

Kamadurai, H.B., Souphron, J., Scott, D.C., Duda, D.M., Miller, D.J., Stringer, D., Piper, R.C., and Schulman, B.A. (2009). Insights into ubiquitin transfer cascades from a structure of a UbcH5B∼ubiquitin-HECT(NEDD4L) complex. Mol. Cell *36*, 1095–1102.

Kim, H.C., Steffen, A.M., Oldham, M.L., Chen, J., and Huibregtse, J.M. (2011). Structure and function of a HECT domain ubiquitin-binding site. EMBO Rep. *12*, 334–341.

Nuber, U., and Scheffner, M. (1999). Identification of determinants in E2 ubiquitin-conjugating enzymes required for hect E3 ubiquitin-protein ligase interaction. J. Biol. Chem. *274*, 7576–7582.

Ogunjimi, A.A., Wiesner, S., Briant, D.J., Varelas, X., Sicheri, F., Forman-Kay, J., and Wrana, J.L. (2010). The ubiquitin binding region of the Smurf HECT domain facilitates polyubiquitylation and binding of ubiquitylated substrates. J. Biol. Chem. *285*, 6308–6315.

Religa, T.L., Sprangers, R., and Kay, L.E. (2010). Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. Science *328*, 98–102.

Religa, T.L., Ruschak, A.M., Rosenzweig, R., and Kay, L.E. (2011). Site-directed methyl group labeling as an NMR probe of structure and dynamics in supramolecular protein systems: applications to the proteasome and to the ClpP protease. J. Am. Chem. Soc. *133*, 9063–9068.

Ruschak, A.M., and Kay, L.E. (2010). Methyl groups as probes of supra-molecular structure, dynamics and function. J. Biomol. NMR *46*, 75–87.

Sprangers, R., and Kay, L.E. (2007). Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature *445*, 618–622.

Venditti, V., Fawzi, N.L., and Clore, G.M. (2011). Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear Overhauser enhancement spectroscopy. J. Biomol. NMR *51*, 319–328.

Verdecia, M.A., Joazeiro, C.A.P., Wells, N.J., Ferrer, J.-L., Bowman, M.E., Hunter, T., and Noel, J.P. (2003). Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. Mol. Cell *11*, 249–259.

Wiesner, S., Ogunjimi, A.A., Wang, H.R., Rotin, D., Sicheri, F., Wrana, J.L., and Forman-Kay, J.D. (2007). Autoinhibition of the HECT-type ubiquitin ligase Smurf2 through its C2 domain. Cell *130*, 651–662.

