Published in final edited form as:

Methods Mol Biol. 2009 ; 497: 121–138. doi:10.1007/978-1-59745-566-4_8.

Identification of SUMO Binding Motifs by NMR

Candace S. Seu^{1,2} and Yuan Chen^{1,*}

¹Division of Immunology, Beckman Research Institute of the City of Hope, Duarte, CA 91010.

Abstract

Post-translational modification by the small ubiquitin-like modifier (SUMO) family of proteins is an important cellular regulatory mechanism, and in recent years, has been found to be involved in a large and diverse set of signaling pathways. Most of these SUMO-dependent functions appear to be mediated by the interaction between SUMO on the modified proteins and a “SUMO binding motif” (SBM or SIM) on receptor proteins. Nuclear magnetic resonance (NMR) studies were instrumental in the identification of this SUMO-binding motif, and reveal that, depending on the sequence context, this motif can bind to SUMO in two opposing orientations. In this paper, we provide an overview of how NMR methods can be used to identify such short conserved binding motifs and structurally characterize their interaction with target proteins. These experiments are complimentary to traditional biochemical methods, and are applicable to the identification of other SUMO-binding motifs and to the studies of other ubiquitin-like modification systems.

Keywords

SUMO; SUMO binding motif (SBM); SUMO interacting motif (SIM); NMR; chemical shift perturbation; protein-peptide interaction; protein-protein interaction

1. Introduction

Small ubiquitin-like modifiers (SUMO) are a family of small proteins that are known to be involved in an increasing variety of essential cellular functions, including DNA repair, genetic transcription, and mitotic regulation (1–3). The SUMO family of proteins includes at least three paralogues, SUMO-1, -2, and -3, which have similar structural and chemical characteristics but appear to be involved in different pathways. In a chemical sense, SUMOylation is simply the covalent coupling of the SUMO C-terminal glycine with a target lysine on a substrate. In a biochemical sense, however, the results are anything but simple. Nevertheless, the rapid and reversible nature of the modification enables SUMO to exert flexible and specific control over protein-protein interactions.

SUMOylation, like other post-translational modifications, exerts this control by providing a new binding site for interactions with other proteins. For example, SUMOylation of RanGAP1 is necessary for its interaction with the nuclear pore protein RanBP2/Nup358 (4,5), and the transcription factors p300 and Elk-1 must be SUMOylated to recruit histone deacetylase 6 (HDAC6) (6) and HDAC2 (7), respectively. In principle, SUMO modification could regulate the activity of a protein by altering its conformation. However, this is unlikely to be a general phenomenon, because SUMO modification sites are often located in extended loops, such as in RanGAP1 (8), or in unstructured termini, such as in p53 (9). Additionally, the modification

*To whom correspondence should be addressed at: Yuan Chen, Ph.D., Division of Immunology, Beckman Research Institute of the City of Hope, 1450 East Duarte Road, Duarte, CA 91010. ychen@coh.org, Tel:(626)-930-5408; Fax:(626)-301-8186.

²Current address: Department of Chemistry and Biochemistry, University of California, San Diego.

sites are not required to be in regular secondary structures (10). By studying the interaction between SUMO and several protein fragments known to bind to SUMOylated proteins, our lab has identified and characterized a SUMO binding motif (SBM) that is crucial to the formation of SUMO-dependent protein-protein interactions (11,12). This SBM is different from the previously identified $\Phi K \times E$ substrate motif, which binds the E2 enzyme for covalent modification by SUMO, but does not bind to SUMO non-covalently (10).

The SBM first identified in our NMR experiments has since been confirmed by many other studies. For example, the SBM site on PML has been shown to be critical for binding SUMOylated proteins and for the formation of PML nuclear bodies (13), and the SBM site on PIASX has been shown to be important for its association with SUMO-modified Elk-1 protein in regulation of its transcriptional activities (14). Notably, the importance of the SBM in transcriptional regulation was demonstrated by a study, in which random mutagenesis was used to identify a small, defined surface on SUMO-1 as the only region important for SUMOylation-dependent transcriptional repression, which is a common function of SUMOylation (15). This surface of SUMO-1 binds the SBM determined by our studies, which suggests that the SBM is likely to mediate most transcriptional regulatory activities. The widespread significance of the SBM has been further demonstrated by the publication of a variety of papers over the last two and half years (16–29).

NMR studies have played and will continue to play a critical role in the identification and characterization of other SUMO binding motifs (11,12). There are several advantages to the use of NMR methods for such studies. First, NMR studies are carried out in solution, so sample preparation is simple. Second, NMR chemical shift perturbation is sensitive to the formation of molecular complexes with a wide range of affinities (with K_d from pM to mM). Third, the residues at the binding interface can be efficiently identified when an interaction is observed. NMR studies had two important advantages over other methods in SBM identification. First, SBMs in proteins can be as short as four or five amino acid residues. Unlike long ubiquitin-binding motifs, such as UBA and CUE, which were readily identified by sequence alignment of ubiquitin-binding proteins, the consensus sequence of the short SBM was difficult to identify by alignment alone. Therefore, NMR studies were critical in the identification of the core residues in the SBM. Second, the bound orientation of the SBM is sometimes reversed, depending on the sequence context (12) (Figure 1). Thus, proper alignment of the consensus sequence requires some SUMO-binding protein sequences to be reversed, which is difficult to carry out in the absence of three-dimensional structural information about the bound orientations of different SUMO-binding sequences. The NMR structures of the two SBM peptides in complexes with SUMO-1 allowed a structure-based sequence alignment to define the consensus sequence (12).

SUMO is only one of many ubiquitin-like proteins, and it is likely that similar binding motifs that cannot be identified through purely biochemical means exist in other systems as well. Here, we discuss and attempt to offer a general understanding of the NMR techniques and strategies that have been used to identify and structurally characterize the SBM. These approaches are expected to be applicable to the identification of other SBMs or novel motifs that bind to other ubiquitin-like proteins.

2. Materials

2.1 Protein expression and purification

1. BL21(DE3) *E. coli* cells or appropriate variants.
2. Protein expression plasmids for SUMO-1, -2 or 3.

3. Ampicillin (100 mg/mL), kanamycin (25 mg/mL), or appropriate filter-sterile antibiotic stocks.
4. LB medium: Dissolve 5 g yeast extract, 10 g trypton, and 10 g NaCl in 1 mL H₂O. Autoclave for 20 min, let cool to room temperature. Add 1 mL of the appropriate antibiotic stock solution.
5. M9 medium: Dissolve 6.7 g Na₂HPO₄, 3.0 g KH₂PO₄, 1 g NaCl, and 0.5 g Na₂SO₄ in 1 L H₂O. Autoclave for 20 min, let cool to room temperature. Filter and add 1 mL of 0.1 M CaCl₂, 1 mL of 1.0 M MgCl₂, 5 mL of 0.2 g/mL NH₄Cl, and 5 mL of 0.4 g/mL glucose (all solutions in water). Add 1 mL antibiotic stock solution, 10 mL of 100× Basal Medium Eagle Vitamin Concentrate (FisherSci), and 1 mL of 1000× Trace Mineral Solution (286 mg H₃BO₄, 1.5 g CaCl₂-H₂O, 4 mg CoCl₂-6H₂O, 20 mg CuSO₄-H₂O, 28 mg FeSO₄-H₂O, 20.8 g MgCl₂-6H₂O, 18 mg MnCl₂-4H₂O, 0.2 mg MoO₃, and 20.8 mg ZnCl₂ in 100 mL sterilized H₂O). For ¹⁵N-, ¹³C-, or ²H-labeled protein, replace NH₄Cl, glucose, or H₂O with its isotopically labeled counterpart (¹⁵NH₄Cl, ¹³C₆-glucose, or D₂O, respectively, available from Cambridge Isotope Laboratories or Spetra Stable Isotopes).
6. Isopropyl-B-D-thiogalactopyranoside (IPTG): 1.0 M stock solution in H₂O.
7. Bugbuster 10× (Novagen).
8. Benzonase (Novagen).
9. 14× Protease inhibitor tablet (Roche) or PMSF.
10. Lysis buffer: 5 mM imidazole, 20 mM phosphate, pH 7.5. Filter sterilize.
11. Wash buffer: 20 mM imidazole, 20 mM phosphate, pH 7.5. Filter sterilize.
12. Elution buffer: 200 mM imidazole, 20 mM phosphate, pH 7.5. Filter sterilize.
13. Ni-NTA agarose beads (Qiagen)
14. 5K MWCO dialysis cassette or centrifuge filter

2.2 Peptide expression and purification

1. BLR(DE3)pLysS *E. coli* cells.
2. pET31b+ plasmid with sequence for ketosteroid isomerase-(Met- candidate SBM)_n-His₆ (Note 5).
3. LB medium.
4. ¹⁵N/¹³C labeled M9 medium (as described in the previous section).
5. IPTG: stock solution as above.
6. Denaturing lysis buffer: 100 mM NaH₂PO₄, 10 mM Tris-HCl, 8 M urea, pH 8.0.
7. Wash buffer: 100 mM NaH₂PO₄, 10 mM Tris-HCl, 8 M urea, pH 6.3.
8. Elution buffer: 100 mM NaH₂PO₄, 10 mM Tris-HCl, 8 M urea, pH 5.9. Adjust the pH of buffer prior to use.
9. 70% formic acid in H₂O.

⁵pET31b+ by Novagen is specific for the expression of small peptides. It a ketosteroid isomerase gene in the N-terminus of the construct, which is used to drive the expressed peptide into inclusion bodies in order to protect it from degradation. Methionine residues separate repeats of the peptide, so that the peptides can be cleaved with cyanogen bromide.

10. 0.2 g cyanogen bromide in 6 mL 70% formic acid.
11. Phosphate buffer, 10–100 mM.

2.3 NMR experiments

1. NMR spectrometer, equipped with four channels, pulsed-field gradient, pulse-shaping capabilities, and triple resonance probe.
2. NMR tubes are from Shigemi Inc.: advanced microtube, model BMS-005-TB (<http://www.geocities.com/~shigemi/page7.html>) (Note 1)
3. NMR buffer: 10–100 mM phosphate, 5–10% D₂O, 0.1% NaN₃, pH 6–7.

3. Methods

3.1 General strategy

The usual strategy for studying protein-peptide complexes requires enrichment of one binding partner, but not the other (Figure 2). Most NMR-active nuclei, with the exception of ¹H, exist in low abundance in natural molecules. Therefore, artificially elevated ¹³C and ¹⁵N levels in the protein sample need to be obtained by expressing it in isotopically enriched media (described in detail below). ¹³C or ¹⁵N **edited** NMR experiments, typically ¹H-¹⁵N HSQC or TROSY, are used to selectively observe signals from the ¹³C/¹⁵N-enriched partner. Reciprocally, ¹³C- or ¹⁵N-**filtered** NMR experiments, such as ¹⁵N-filtered TOCSY and NOESY, are used to selectively observe signals from the un-enriched partner in the complex. In these spectra, resonances of the labeled protein are usually well resolved because of large ¹⁵N chemical shift dispersion, while resonances of the peptide are easy to monitor because of the low number of peaks.

3.2 Expression and purification of ¹³C, ¹⁵N-labeled SUMO proteins

1. Transform expression plasmids of SUMO into BL21(DE3) *E. coli* cells using standard molecular biology methods.
2. Streak cells onto an LB plate containing appropriate antibiotic, incubate overnight at 37° C.
3. Select a single colony. Inoculate 50–100 mL of antibiotic supplemented LB media with the colony, and grow overnight at 37° C in a shaker (Note 2).
4. Centrifuge cells (5000 rpm, 5 min), decant media, and resuspend cell pellet in 1 L of ¹⁵N- or ¹⁵N-, ¹³C- enriched M9 media containing appropriate isotopes and antibiotics. ¹⁵N-labeled medium would be sufficient for titration experiments to detect chemical shift perturbation upon the complex formation. ¹⁵N, ¹³C-enriched medium is necessary for structural studies of SUMO-SBM complex.
5. Shake in a 37° C incubator at 220 rpm, 2–3 hours until OD₆₀₀ = 0.6–0.8.

¹Shigemi tubes allow the use of less sample (1/3 of that needed for a regular NMR tube) and are manufactured as an outer tube and an inner plug (Figure 6). To load the sample and inner plug effectively, first pipette the sample into the outer tube and centrifuge to remove any bubbles at the meniscus. Carefully push the inner plug into the tube until its bottom contacts the top of the sample. Hold the tube in one hand at a 45° angle to upright, and tap the top of the inner tube gently but quickly, such that the inner plug travels a short distance downward, hopefully letting any air bubbles slip past the plug. Avoid forcing the inner plug all the way to the bottom of the outer tube; the inner plug can be slowly and gently pulled upward before this happens. Tap and repeat as necessary until all air bubbles are gone and a small amount of sample solution is present above the plug. Air bubbles in the sample volume cause field inhomogeneity, and should be scrupulously eliminated.

²This initial overnight growth in LB is necessary to ensure that the cells grow in a timely manner, as starting with a single colony in M9 media is too slow to be practical.

6. Induce protein expression by adding 0.5 mM IPTG (500 μ L of 1 M stock solution). Incubate 3–4 hours. Do not overgrow the cells; SUMO expression should be induced at $OD_{600}=0.6$ for maximal protein production in M9 media. When using LB media, one can induce SUMO expression at $OD_{600}=0.8 - 1$.
7. Harvest by centrifuging cells (5000 rpm, 5 min). Pellet can be saved at -80° C indefinitely.
8. Resuspend cells in a solution of 20 mL lysis buffer, 2 mL 10 $\times$ Bugbuster, 2 μ L benzonase, and protease inhibitors (Note 3).
9. Centrifuge the suspension (16,000 g, 15 min) and retain the supernatant.
10. Purify the His₆-tagged proteins from the supernatant using standard Ni²⁺ affinity chromatography techniques (please refer to the instructions by the manufacturer of the Ni affinity beads). Buffer exchange into a buffer that does not produce NMR signal by itself, such as phosphate buffer, and concentrate the protein (Note 4).
11. Titration experiments, require a minimum of 0.1 mM protein concentration. For structural studies, it is desirable to have SUMO concentrations over 0.4 mM. Since the sample volume can be as small as 250 μ L, 1 mg of SUMO is sufficient for most experiments.
12. Check protein purity by Coomassie blue staining of a SDS-Page gel. Estimate protein concentrations by amino acid analysis. SUMO can be stored at -80° C for several years. All human SUMO paralogues and yeast homologue are stable in the NMR tube and are soluble to several mM.

3.3 Expression and purification of labeled peptides

1. Clone multiple tandem copies (a larger replication number gives a higher yield) of the peptide encoding sequence, separated by single methionine codons, into the pET31b+ plasmid. The cloning site is upstream of the His₆ tag and downstream of an N-terminal ketosteroid isomerase gene (Note 5).
2. Transform the pET31b+ expression construct into BLR(DE3)pLysS *E. coli* cells using standard molecular biology methods.
3. Follow protocol steps 2–7 in section 3.2.
4. Resuspend cells in 5 mL of denaturing lysis buffer per gram of pellet. Mix gently at room temperature until the solution is translucent (15–60 min), signaling complete lysis.
5. Centrifuge the suspension (10,000 g, 30 min) and retain the supernatant.
6. Purify the His₆-tagged proteins from the supernatant using standard denaturing Ni²⁺ affinity chromatography techniques. Precipitate the expressed protein by dialysis against water.
7. Redissolve precipitate in 70% formic acid. Cleave fusion protein overnight with the cyanogen bromide/formic acid solution in a ventilated hood protected from light. Dry by rotary evaporation.
8. Resuspend gelatinous material in PBS. Adjust pH to 7.4 and stir overnight.

³If you do not have Bugbuster, cell lysis can be achieved through sonication or with a French press.

⁴Buffer exchange filters sometimes contain a wetting agent such as glycerol, which will contaminate samples and produce NMR signals that overlap with the signals of the protein and peptide samples. To avoid this, soak and rinse the filters thoroughly following manufacture's recommendations.

9. Purify peptide-containing supernatant with reverse-phase HPLC, because a small amount of covalent dimer will also be produced. Verify sequence identity with mass spectrometry (Note 6).

3.4 Detection of SUMO-SBM interactions

Nuclear magnetic resonance (NMR) chemical shift perturbation is probably the most sensitive method for the determination of whether your candidate SBM binds to SUMO. The exact frequencies, or “chemical shifts” of the nuclei, are modulated by their chemical and spatial environments. Each peak in an NMR spectrum can be correlated to a source nucleus and interpreted to reveal and monitor the chemical environment and structure around that particular atom. Chemical shifts are extremely sensitive to changes in the local environments of their source nuclei, which can be caused by introduction of aromatic ring currents, peptide bond anisotropy, electrostatic interactions, and hydrogen bonding upon protein complex formation. Therefore, NMR chemical shift perturbation is frequently used for monitoring protein complex formation. To detect whether a potential SBM binds to SUMO, SUMO can be titrated by a peptide harboring the potential SBM. Then ^1H - ^{15}N HSQC or TROSY will be recorded to monitor specific chemical shift changes. Chemical shift perturbation is extremely sensitive for detecting complex formation of a wide range of affinities with dissociation constants ranging from pM to mM.

3.5 Identification of interacting SUMO residues and estimation of K_d

Chemical shift perturbation is highly efficient in identifying the residues involved in the SUMO-SBM interaction. Residues that are located at the binding interface will inevitably experience changes in their chemical shifts, due to changes in the local environments of their nuclei. Residues that have the largest chemical shift changes are usually located within the binding interface, although the surface identified by chemical shift perturbation usually extends slightly beyond the direct contact area, due to small conformational changes induced by complex formation.

Superposition of the HSQC spectra of SUMO, free and in complex with the peptides, will reveal the SBM binding sites on the surface of SUMO. In order to correlate a resonance in the spectra to a specific residue in the protein or peptide, chemical shift assignments must be obtained. The chemical shift assignments for SUMO-1 and SUMO-2 are available (deposited in BioMagResBank with entry numbers 6304 for SUMO-1, and 6801 for SUMO-2). HSQC is a very sensitive 2D NMR experiment which correlates amide hydrogens with their nitrogens, yielding one peak for every residue. Within less than one hour, one can determine whether a putative SBM actually interacts with SUMO, identify the binding site, and obtain information on the K_d of the interaction. Detailed procedures are as follows:

1. Prepare at least 0.2 mL of ^{15}N -labeled SUMO-1 in NMR buffer in a Shigemi tube. A concentration of 0.1 mM is sufficient for a 600 MHz spectrometer with a cryo-probe; 0.3 mM should guarantee signal on a 500 MHz spectrometer with a room temperature probe (Note 7).
2. Take an initial HSQC spectrum of this unbound SUMO sample. This will serve as your baseline and control.
3. Add unlabeled peptide to the sample to specific molar equivalences (e.g. 1:4, 1:2, 3:4, 1:1, 2:1), mixing well and taking a new HSQC spectrum after each addition. The

⁶Isotopically labeled Fmoc amino acid derivatives are extremely expensive. Thus, in vivo production presents an attractive alternative to solid-phase peptide synthesis.

⁷All NMR experiments should be done at the same temperature (recommended: 17° C for SUMO-1) and at the same pH. Chemical shifts are extremely sensitive to both, but especially the latter.

titration should be designed such that the final concentration of SUMO should be at least 0.1 mM at the end of the titration. It is usually not a problem, since SUMO is soluble to at least 5 mM.

4. Process and overlay the spectra using programs such as Felix and NMRView.
5. Observe how each peak changes in response to increasing peptide concentration. A lack of change indicates that the residue does not participate in the interaction. The type of changes (such as gradual shifts, intensity changes, and complete disappearances) provides information on the rate of the complex dissociation (see below: Point 8). Since the association rate is likely to be diffusion-limited, the dissociation rate is therefore a good indication of the affinity constant.
6. Mapping the affected residues onto a 3D model of the protein may aid in confirming the binding site and differentiating between local and long-range structural changes. This requires chemical shift assignments of the protein (Section 3.7).
7. When performing these experiments, it is important to ensure that the both SUMO and the peptides are in an identical buffer to ensure that chemical shift changes are not caused by slight pH changes in the solution when mixing the peptide and the protein together.
8. Note the type of chemical shift changes (Figure 3):
 - Fast exchange: Gradual peak shifting indicates that the exchange between the free and bound molecule is fast relative to the chemical shift difference of the two forms. The NMR spectrum is thus a population-weighted average of the free and bound molecule. As the concentration of the bound form increases, the average is increasingly weighted towards the new position. Fast exchange is usually correlated with weak binding, with dissociation constant ranging on the order of 100 μ M to 1 mM. In this case, a K_d for the interaction can be accurately extracted by fitting the data to a curve using software such as Origin or Sigma Plot. More detailed reviews on this topic can be found in the literature (30).
 - Slow exchange: In some cases, peaks in the spectrum of unbound protein will gradually decrease in intensity until they disappear, while new peaks appear in new positions and gradually increase in intensity at subsequent titration points. This pattern of change occurs when the dissociation rate is slower than the chemical shift difference of the free and bound species. The NMR thus effectively takes a spectrum of a sample containing two distinct compounds (free and bound protein), whose relative intensities are determined by the population of the free and bound portions of the protein. Slow exchange is usually correlated with stronger complex formation, with dissociation constants less than 1 μ M.
 - Intermediate exchange: In some cases, peaks will disappear completely due to extreme line broadening. In this case, the dissociation constant is usually between 1 to 10 μ M.

3.6. Identification of the core SBM residues

Superposition of the TOCSY spectrum of the free peptides and the 15 N-filtered TOCSY spectrum of the peptides in complexes with 15 N-enriched SUMO-1 will reveal the residues in these peptides that have the most significant chemical shift changes upon complexation, identifying the core SBM sequence (11). Identification of the specific SBM residues that participate in the interaction with SUMO can usually be accomplished with the same sample

of ^{13}C , ^{15}N labeled SUMO and unlabeled SBM. $^{15}\text{N}/^{13}\text{C}$ -filtering is used to remove signals from the $^{15}\text{N}/^{13}\text{C}$ -labeled SUMO, leaving only the data from the unlabeled peptide. TOCSY experiments are used to identify the amino acid residue types, while NOESY experiments provide information on their sequential connectivity (31). In particular, the latter are used to establish physical proximity between H_α , H_β , H_N of one residue ($n-1$, Figure 4) and the H_N of the next residue (n , Figure 4). Chemical shift assignments of the peptides, free and in complex with SUMO, can be obtained using a combination of TOCSY and NOESY spectra.

If the SBM contains multiple instances of the same amino acid, the residues may give degenerate shifts and be indistinguishable in the 2D spectrum. In this case, it is necessary to produce ^{13}C , ^{15}N labeled peptide using the methods described in Section 3.3. The amino acid assignment of the peptide's residues can then be solved using the heteronuclear NMR techniques described in Section 3.7. Once this has been achieved, a one-step titration and spectral overlay of labeled peptide with unlabeled SUMO will distinguish the SBM and non-interacting regions of the peptide. A ^{13}C -resolved NOESY spectrum of the sample containing ^{13}C , ^{15}N -labeled peptide and unlabeled SUMO contains mostly intermolecular NOEs, because the small peptide does not produce a large number of intramolecular NOEs. Such NOESY spectrum has significantly higher sensitivity than filtered- and edited NOESY spectra for the identification of intermolecular NOEs for structure determination of the complex.

3.7. Structural Determination of SUMO-SBM complexes

Resonance assignments are the basis for structure determination by NMR methods. If the exchange between free and bound SUMO is fast on the chemical shift timescale, SUMO resonances in the complex with a SBM can be readily assigned by following the shift of the peaks during titration. We have developed a software to automatically accomplish this task (32). If the exchange between the free and bound SUMO is slow on the chemical shift time scale, the resonances of SUMO in the complex will need to be assigned. This can be accomplished by acquiring multiple 3D NMR experiments on 0.3–0.5 mM of a ^{13}C , ^{15}N labeled sample of SUMO in complex with the unlabeled SBM peptide. In these experiments, NMR-active sidechain carbons are correlated with nearby backbone nitrogens and their protons, functionally adding a third, carbon dimension to a HSQC plane. Inter-residue correlations aid in establishing amino acid connectivity, while intra-residue correlations aid in identifying the amino acids and matching them up to a known sequence. A strategy for establishing connectivity is illustrated in Figure 5. Complete resonance assignments of the protein are achieved when the chemical shifts of each residue's H, N, C_α , and C_β are known in the protein. The choice and number of experiments needed to achieve this varies. Data acquisition using current methods generally takes 1–3 days per 3D experiment on a 600 MHz NMR instrument. Data interpretation may take up to an additional week to 6 months, though this depends on the size of the protein, the quality of the NMR data, and the extent of signal overlap. While a complete protocol for solving a protein structure by NMR is beyond the scope of this review, procedures can be found in other publications (33–36). A similar strategy can be used to assign the resonances of the SBM-containing peptide, using a sample containing $^{13}\text{C}/^{15}\text{N}$ -enriched peptide and unlabeled SUMO.

Assignment of chemical shifts to their originating residues does not in itself yield a three-dimensional structure of a protein. Instead, it is necessary to collect “structural restraints”, which identify short inter-proton distances, backbone torsion angles, hydrogen bonds, and bond orientations. These are compiled and incorporated into iterated structural calculations with programs such as HADDOCK (37). Structure determination of a SUMO-SBM complex can provide insights into the recognition mechanism, and the information can be used to design peptidomimetics or small molecular mimics of the SBM.

Nuclear Overhauser effect spectroscopy (NOESY) experiments, which measure NOE between two protons, are the most important way of providing distance restraints. The intensity of an NOE signal is inversely proportional to r^6 (where r is the distance between the two nuclei), and can be used to infer inter-atom distances of up to 5 Å. In a real, non-ideal environment, this effect can be complicated by phenomena such as “spin diffusion”, in which signal is transferred to a third nucleus, giving the erroneous impression that there is an NOE signal between nuclei 1 and 3, and local dynamics, which artificially weaken the NOE. Distances derived from NOEs are thus expressed as one of several ranges in order to reflect this uncertainty.

The second class of restraints used in our SBM studies was generated by inferring backbone torsion angles from chemical shifts. The chemical shifts of the $H\alpha$, $C\alpha$, $C\beta$, and C' residues are affected by local backbone conformation in a highly predictable manner. The chemical shift index (CSI) method is a simple analytical tool that correlates chemical shift patterns with certain secondary structural features (38). More recently, structure prediction programs such as TALOS have been developed to yield more detailed predictions (39).

There are three other common methods of generating restraints:

1. Backbone dihedral angle constraints can be provided by chemical shifts of the backbone atoms, as described above. In addition, J coupling constants between HN and $H\alpha$ atoms are dependent on their dihedral angles, and can be related by the Karplus equation in order to yield Φ backbone angles (40,41).
2. Residual dipole coupling (RDC) constraints measure the relative orientation of dipoles (usually H-N bonds) to an alignment tensor. In normal solution, rapid tumbling of the macromolecule causes the average dipolar coupling constants to be zero. However, bacteriophages, bicelles, or other solid supports can be used as “alignment media” to predispose the molecule towards a preferred alignment in the magnetic field, so that RDCs are not zero and thus can be measured (42). RDCs are most often used to improve the accuracy of structures determined from other constraints.
3. NH hydrogens exchange rapidly with water protons, if not protected by hydrogen bonds. Dissolving the protein in D_2O allows these residues to become deuterated, and hence NMR inactive. An overlay of the HSQC spectra of the protein in H_2O and D_2O should thus identify the amide hydrogen atoms that are protected by hydrogen bonds, by virtue of their persistence in the latter spectrum.

Acknowledgments

This work was supported by NIH grants GM074748 and CA94595.

References

1. Hay RT. Mol Cell 2005;18:1–12. [PubMed: 15808504]
2. Johnson ES. Annu Rev Biochem 2004;73:355–382. [PubMed: 15189146]
3. Seeler JS, Dejean A. Nat Rev Mol Cell Biol 2003;4:690–699. [PubMed: 14506472]
4. Matunis MJ, Wu J, Blobel G. J Cell Biol 1998;140:499–509. [PubMed: 9456312]
5. Mahajan R, Delphin C, Guan T, Gerace L, Melchior F. Cell 1997;88:97–107. [PubMed: 9019411]
6. Girdwood D, Bumpass D, Vaughan OA, Thain A, Anderson LA, Snowden AW, Garcia-Wilson E, Perkins ND, Hay RT. Mol Cell 2003;11:1043–1054. [PubMed: 12718889]
7. Yang SH, Sharrocks AD. Mol Cell 2004;13:611–617. [PubMed: 14992729]
8. Bernier-Villamor V, Sampson DA, Matunis MJ, Lima CD. Cell 2002;108:345–356. [PubMed: 11853669]

9. Rodriguez MS, Desterro JM, Lain S, Midgley CA, Lane DP, Hay RT. *Embo J* 1999;18:6455–6461. [PubMed: 10562557]
10. Lin D, Tatham MH, Yu B, Kim S, Hay RT, Chen Y. *J Biol Chem* 2002;277:21740–21748. [PubMed: 11877416]
11. Song J, Durrin LK, Wilkinson TA, Krontiris TG, Chen Y. *Proc Natl Acad Sci U S A* 2004;101:14373–14378. [PubMed: 15388847]
12. Song J, Zhang Z, Hu W, Chen Y. *J Biol Chem* 2005;280:40122–40129. [PubMed: 16204249]
13. Shen TH, Lin HK, Scaglioni PP, Yung TM, Pandolfi PP. *Mol Cell* 2006;24:331–339. [PubMed: 17081985]
14. Yang SH, Sharrocks AD. *Embo J* 2005;24:2161–2171. [PubMed: 15920481]
15. Chupreta S, Holmstrom S, Subramanian L, Iniguez-Lluhi JA. *Mol Cell Biol* 2005;25:4272–4282. [PubMed: 15870296]
16. Tatham MH, Kim S, Jaffray E, Song J, Chen Y, Hay RT. *Nature Structural & Molecular Biology* 2005;12:67–74.
17. Baba D, Maita N, Jee JG, Uchimura Y, Saitoh H, Sugawara K, Hanaoka F, Tochio H, Hiroaki H, Shirakawa M. *Nature* 2005;435:979–982. [PubMed: 15959518]
18. Izumiya Y, Ellison TJ, Yeh ETH, Jung JU, Luciw PA, Kung HJ. *Journal of Virology* 2005;79:9912–9925. [PubMed: 16014952]
19. Takahashi Y, Kikuchi Y. *Journal of Biological Chemistry* 2005;280:35822–35828. [PubMed: 16109721]
20. Nguyen HV, Chen JL, Zhong J, Kim KJ, Crandall ED, Borok Z, Chen Y, Ann DK. *American Journal of Pathology* 2006;168:1452–1463. [PubMed: 16651613]
21. Cheng CH, Lo YH, Liang SS, Ti SC, Lin FM, Yeh CH, Huang HY, Wang TF. *Genes & Development* 2006;20:2067–2081. [PubMed: 16847351]
22. Uchimura Y, Ichimura T, Uwada J, Tachibana T, Sugahara S, Nakao M, Saitoh H. *Journal of Biological Chemistry* 2006;281:23180–23190. [PubMed: 16757475]
23. Raffa GD, Wohlschlegel J, Yates JR, Boddy MN. *Journal of Biological Chemistry* 2006;281:27973–27981. [PubMed: 16880212]
24. Mukhopadhyay D, Ayaydin F, Kolli N, Tan SH, Anan T, Kametaka A, Azuma Y, Wilkinson KD, Dasso M. *Journal of Cell Biology* 2006;174:939–949. [PubMed: 17000875]
25. Lin DY, Huang YS, Jeng JC, Kuo HY, Chang CC, Chao TT, Ho CC, Chen YC, Lin TP, Fang HI, Hung CC, Suen CS, Hwang MJ, Chang KS, Maul GG, Shih HM. *Molecular Cell* 2006;24:341–354. [PubMed: 17081986]
26. Mohan RD, Rao A, Gagliardi J, Tini M. *Molecular and Cellular Biology* 2007;27:229–243. [PubMed: 17060459]
27. Lee YK, Thomas SN, Yang AJ, Ann DK. *Journal of Biological Chemistry* 2007;282:1595–1606. [PubMed: 17079232]
28. Benson MD, Li QJ, Kieckhafer K, Dudek D, Whorton MR, Sunahara RK, Iniguez-Lluhi JA, Martens JR. *Proceedings of the National Academy of Sciences of the United States of America* 2007;104:1805–1810. [PubMed: 17261810]
29. Kumar A, Srivastava S, Hosur RV. *Journal of Molecular Biology* 2007;367:1480–1493. [PubMed: 17320104]
30. Lian, L-Y. “NMR of Macromolecules: A Practical Approach”. Roberts, G., editor. New York: Oxford University Press; 1993. p. 164–166.
31. Wuthrich, K. John Wiley and Sons. 1986.
32. Fukui L, Chen Y. *Bioinformatics* 2007;23:378–380. [PubMed: 17118956]
33. Clore GM, Gronenborn AM. *Science* 1991;252:1390–1399. [PubMed: 2047852]
34. Driscoll PC, Gronenborn AM, Wingfield PT, Clore GM. *Biochemistry* 1990;29:4668–4682. [PubMed: 2372550]
35. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. *Biochemistry* 1989;28:6150–6156. [PubMed: 2675964]

36. Oshkinat H, Griesinger C, Kraulis PJ, Sorensen OW, Ernst RR, Gronenborn AM, Clore GM. *Nature* 1988;332:374–376. [PubMed: 3352736]
37. Dominguez C, Boelens R, Bonvin AM. *J Am Chem Soc* 2003;125:1731–1737. [PubMed: 12580598]
38. Wishart DS, Sykes BD, Richards FM. *J Mol Biol* 1991;222:311–333. [PubMed: 1960729]
39. Cornilescu G, Delaglio F, Bax A. *J Biomol NMR* 1999;13:289–302. [PubMed: 10212987]
40. Karplus M. *J Chem Phys* 1959;30:11–15.
41. Karplus M. *J Am Chem Soc* 1963;85:2870–2871.
42. Bax A, Kontaxis G, Tjandra N. *Methods Enzymol* 2001;339:127–174. [PubMed: 11462810]

Figure 1.
The SBMs of RanBP2 and PIAS-X-P bind SUMO-1 in opposite orientations.

Figure 2.

Labeling scheme for protein-peptide interaction studies using NMR methods. The left configuration is usually sufficient for most SBMs; preparation of the complex on the right may be necessary when the peptide contains degenerate residues.

Figure 3.
NMR peak behavior at different exchange rates between the free and bound states.

Figure 4.

The NOESY experiment transfers signal between protons that are physically close. TOCSY spectral patterns are used to identify amino acid residues, and inter-residue NOESY connections are then used to match the residues to the peptide sequence.

Figure 5.

The HNCA experiment, true to its name, transfers magnetization in the order H-N-C α and back, and thus correlates the H and N of residue ***n*** with the C α residues of residues ***n*** and ***n-1***. C α_{n-1} is correlated to two H-N pairs; therefore, there should be two peaks with the same C α shift but different H-N values. Unambiguous identification of these two peaks establishes that they belong to contiguous residues. Chains of contiguous residues can then be combined to reveal overall connectivity.

Figure 6.
Proper Shigemi tube assembly.