Structural analysis of a complex between small ubiquitin-like modifier 1 (SUMO1) and the ZZ domain of CREB-binding protein (CBP/p300) reveals a new interaction surface on SUMO

Carl Diehl1,2, Mikael Akke3, Simon Bekker-Jensen1, Niels Mailand1, Werner Streicher1,4, and Mats Wikström1,5

1From the Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2200 Copenhagen, Denmark, 2SARomics Biostructures, SE-22363 Lund, Sweden, 3Department of Biophysical Chemistry, Center for Molecular Protein Science, Lund University, SE-22100 Lund, Sweden, 4Novozymes A/S, DK-2880 Bagsvaerd, Denmark, and 5Amgen Inc, One Amgen Center Drive, Thousand Oaks, CA 91320, United States.

Running title: Structural studies of a ZZ domain-SUMO1 complex

To whom correspondence should be addressed: Mats Wikström, Amgen Inc, One Amgen Center Drive, Thousand Oaks, CA 91320, United States. E-mail: mats.wikstrom@amgen.com

Keywords: protein-protein interaction; protein structure; protein dynamics; SIM; NMR

ABSTRACT

We have recently discovered that the ZZ zinc finger domain represents a novel SUMO binding motif. In this study we identify the binding epitopes in the ZZ domain of CBP and SUMO1 using NMR spectroscopy. The binding site on SUMO1 represents a unique epitope for SUMO interaction spatially opposite to that observed for canonical SUMO Interaction Motifs (SIMs). HADDOCK docking simulations using chemical shift perturbations and residual dipolar couplings was employed to obtain a structural model for the ZZ domain-SUMO1 complex. ITC experiments supports this model by showing that the mutation of key residues in the binding site abolishes binding and that SUMO1 can simultaneously and non-cooperatively bind both the ZZ domain and a canonical SIM motif. The binding dynamics of SUMO1 was further characterized using 15N CPMG relaxation dispersions, which define the off rates for the ZZ-domain and SIM motif, and show that the dynamic binding process has different characteristics for the two cases. Furthermore, in the absence of bound ligands SUMO1 transiently samples a high-energy conformation, which might be involved in ligand binding.

SUMO (small ubiquitin-like modifier) is a protein structurally homologous to ubiquitin that functions as a post-translational modifier involved in diverse cellular processes such as chromatin remodeling, ubiquitin E3 ligation, autophagy, cytoskeletal scaffolding, and DNA repair. Higher eukaryotes encode at least three different functional SUMO isoforms, in the human genome designated as SUMO1-3, which can conjugate to a number of target proteins (1-3). SUMO modified target proteins are recognized by SUMO interacting motifs (SIMs), the minimal consensus sequence, which is ψ-K-X-E, where ψ consists of a large hydrophobic amino acid and where X can be any amino acid (4,5). In addition to this canonical type of SIM, closely related variants such as inverted SIMs and phosphorylation-dependent SIMs have also been recently described (6-8). The binding site for the canonical SIM is located in a groove between the α-helix and β-sheet in SUMO, where the SIM motif can bind either in a parallel or anti-parallel fashion (6).

We have recently shown that the ubiquitin ligase HERC2 can bind SUMO1 through its zinc-finger ZZ domain (9), which therefore represents a new class of SUMO binding motifs. Isothermal titration calorimetry (ITC) showed that the ZZ domain binds to SUMO1 with μM affinity in the same range as the canonical SIMs (10). The ZZ domain consists of two anti-parallel β-sheets and a short α-helix
coordinating two zinc ions as revealed by the first published ZZ domain structure originating from CBP (CREB-binding protein/p300) (11). There are approximately 20 ZZ domains (11,12) identified in the human genome, but a general biological function has not been assigned.

Here we present the complex between SUMO1 and the ZZ domain from CBP, where we map the interaction interface using NMR spectroscopy. The structure of this particular ZZ domain has previously been determined in solution by NMR and thus has favorable properties for structural studies (11). Using $^{15}$N HSQC spectra, we could identify the interaction surfaces on the ZZ domain and SUMO1, where the residues affected by the binding were broadened beyond detection. On SUMO1 this constituted a unique binding epitope that subsequently could be used to model the protein-protein complex using HADDOCK docking simulations (13).

Furthermore, in order to address the protein dynamics between different states of SUMO1 we performed CPMG relaxation dispersion experiments to map the micro- to millisecond dynamics of SUMO1, where we discovered that apo-SUMO1 experiences intrinsic conformational exchange. The conformational exchange was quenched in SIM-bound SUMO1, whereas the intrinsic conformational exchange in ZZ domain-bound SUMO1 was largely unperturbed.

**EXPERIMENTAL PROCEDURES**

**Molecular biology** - DNA sequences for the ZZ domain from CBP (UniProtKB = Q92793) and SUMO1 (UniProtKB = P63165) were synthesized by Geneart and sub-cloned into a pNIC28-BsaI vector using ligation independent cloning (14). The resulting expression constructs contained a His-tag capture sequence and a tobacco etch virus (TEV) protease cleavage site at the N terminus preceding the expressed protein of interest. In order to obtain SUMO1 protein that behaves as monomers in solution without the need of the addition of a reducing agent, the single cysteine residue in SUMO1, Cys52, was mutated to serine using Quickchange site-directed mutagenesis (15) and transformed into MACH-1 cells. The SUMO1 mutant was shown to have the same structural and binding properties as the wildtype protein (data not shown). The C52S mutant was used in all subsequent experiments and is hereafter referred to as SUMO1. The sequences for the expression constructs were verified by sequencing. Finally, in order to test the importance of residues in the identified binding epitopes, one mutant was created for the ZZ domain, denoted ZZmut, corresponding to residues N36A, K38A and S39A, while two mutants were designed for SUMO1 corresponding to residues H75A, K78A and E83A denoted SUMO1mut1 and K16A, H43A, H75A, K78A and E83A denoted SUMO1mut2 respectively.

**Protein expression and purification** - Plasmids were transformed into *Escherichia coli* (E. coli) BL21 (DE3) Rosetta cells prior to protein expression. Protein expression was started from overnight cultures grown in Terrific Broth (TB) media supplemented with 50 µg/ml kanamycin and 30 µg/ml chloramphenicol. SUMO1 and the SUMO1 mutants were expressed in TB media with 50 µg/ml kanamycin and 30 µg/ml chloramphenicol, grown to OD$_{600}$ ~ 0.6-0.7 at 37°C, induced with 0.5 mM isopropyl-β-thiogalactopyranoside (IPTG) and grown overnight at 18°C. The ZZ domain and the ZZ mutant were expressed in TB media as above with the addition of 150 µM ZnCl$_2$ according to a previously published protocol (11). For the NMR experiments, SUMO1 and the ZZ domain were labeled with $^{15}$N and $^{13}$C using minimal media (16). SUMO1 was grown in minimal medium to OD$_{600}$~0.6-0.7, induced with 0.5 mM IPTG, and grown overnight at 18°C. The ZZ domain was grown in minimal media with 150 µM ZnCl$_2$ until reaching an OD$_{600}$~1.2-1.5, induced with 1.0 mM IPTG and grown overnight at 18°C. Cells were harvested by centrifugation at 6000 RPM for 10 min, re-suspended in 100 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole with benzamide and Complete-EDTA free tablets (Roche Life Sciences), homogenized using a French press cell disruptor and centrifuged at 11000 g for 30 min. Purification of the proteins was performed using immobilized metal ion affinity chromatography (IMAC) on an ÄKTA express system (GE Healthcare) at 4°C with the HiTrap chelating columns (17) equilibrated with 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole, and eluted using a linear gradient of 10-500 mM imidazole. Protein identity was verified using mass spectrometry (MS). The eluted fractions with protein were collected, dialyzed (Molecular weight
(M_w) cut-off 3500 Da) against 50 mM Tris-HCl pH 7.5, and processed with TEV protease in order to remove the His-tag. The TEV protease treatment retains the amino acids pair Ser-Met at the N-terminal of the expressed proteins. We have therefore designated the first two residues as S_0 and M_1 for both the ZZ domain and SUMO1. The final amino acid sequences for the two proteins are therefore; ZZ domain:

SM_{i}QDRFVYTCECKHHVETRWHCTVCEDY DLCINCYNTKSHAHKMVKWGGLDD_{53}, and SUMO1:

SM_{i}SDQEAKPSTEDLDGKKEGYIKKLKVIQQ DSSEIHFKVKMTTHLKKLESYSQRQGVPMSNLRFLEFGQRIADNHTPKELGMEEDVIEW YQEQTGGHSTV_{101}.

In the final purification step, the proteins were loaded onto an ion exchange MonoQ column (ÄKTA, GE Healthcare) equilibrated with 50 mM Tris pH 7.5 and eluted with a linear gradient of 0-0.4 mM NaCl. The purity was verified by SDS-PAGE and electrospray ionization mass spectrometry (ESI-MS).

**Peptide synthesis** - The SIM-peptide from PIAXS (6) hereinafter denoted SIMPX, corresponding to the following sequence; KVDVILDLIESSDSEEDPPAKROMQ, was synthesized by Biosyntan Gmbh (Berlin, Germany). Purity and molecular mass was confirmed by MALDI-TOF MS.

**Protein characterization** - The secondary structure of the ZZ domains (ZZ wildtype and ZZmut) and the SUMO1 variants (SUMO1, SUMO1mut1, and SUMO1mut2) was verified by circular dichroism (CD) spectroscopy using 0.5-1.0 mg/ml protein in 20 mM MOPS buffer at pH 6.8.

**Isothermal Titration Calorimetry (ITC) experiments** - ITC experiments were performed using a VP-ITC200 instrument (GE Healthcare). The samples were extensively dialyzed against 20 mM MOPS pH 6.8. For the experiments addressing the binding properties of the ZZ domains (ZZ wildtype and ZZmut) and the SUMO variants (SUMO1, SUMO1mut1, and SUMO1mut2), 0.5 mM SUMO1 or SUMO1 mutant (SUMO1mut1 or SUMO1mut2) was titrated into 0.04 mM ZZ wildtype or ZZ mutant. The concentrations were determined from the respective absorbance at 280 nm. In the experiments investigating the SUMO-binding of the ZZ domain and the SIM peptide SIMPX, 0.5 mM SIMPX was titrated into the ITC sample cell containing 0.05 mM SUMO1. Finally, 0.5 mM SIMPX was titrated into a sample containing the pre-formed ZZ domain-SUMO1 complex. All experiments were performed at 10 °C and run until saturation was achieved. The data were fitted using a model describing one binding site using the software provided by the manufacturer (18).

**NMR experiments and resonance assignments** - Protein samples were dialyzed against 20 mM MOPS pH 6.8 and concentrated to 0.3-0.5 mM using Vivaspün concentrators with a M_w cut-off of 3500 Da. D_2O was added to a concentration of 7 % (v/v) for the spectrometer lock and NaN_3 was added to a concentration of 0.02 % (w/v) to prevent bacterial growth in the samples. The backbone resonance assignments for SUMO1 and the ZZ domain were based on published assignments (11,19) in combination with standard 3D triple-resonance experiments for the backbone assignments (HNCA (20), HNCOCA (21), HNCACB (22) and CBCACONH (23)). Spectra were processed with NMRpipe (24), applying zero-filling and linear prediction in the indirect dimensions and solvent filter and polynomial baseline correction in the direct dimension. CCPNMR (25) was used for visualization of spectra and resonance connectivity analysis. Resonance assignment was achieved for 98% of the backbone amides in the ZZ domain and 86% in SUMO1. The SIMPX-bound SUMO1 was assigned by performing a titration of SIMPX into a SUMO1 sample and by following the chemical shift changes upon complex formation. The samples for measuring RDCs were prepared by adding the PF1 phage (ASLA Biotech) to a final concentration of 10 mg/ml in the NMR samples.

**RDC measurements and molecular modeling** - Residual dipolar couplings (RDCs) were measured using interleaved IPAP-experiments (26) at 25°C with spectral widths of 8000 Hz (1H) and 1835 Hz (15N), sampled over 1024 and 256 points respectively. RDCs were measured for each component of the ZZ domain-SUMO1 complex using two separate samples, in which one protein was labeled with 15N/13C and the other unlabeled. In total, 91 RDCs were collected, 44 for SUMO1 and 47 for the ZZ domain.

**Formation of the ZZ-SUMO1 and SIMPX-SUMO1 complexes** - In order to confirm binding and establish complexes with 1:1 stoichiometry, the
NMR samples were prepared by ITC titrations of unlabeled ZZ domain into \(^{15}\text{N}/^{13}\text{C}\) labeled SUMO1, and vice versa. The final NMR samples of the ZZ domain-SUMO1 complex contains 4% unbound SUMO1, as calculated from the Kd and protein concentrations. The complex between \(^{15}\text{N}/^{13}\text{C}\) labeled SUMO1 and SIMPX was formed in a similar way using ITC titrations resulting in NMR samples of the SIMPX-SUMO1 complex that contains 2% unbound SUMO1. \(^{15}\text{N}\) HSQC experiments were subsequently run on the complexes and the resulting spectra were compared with those for unbound proteins.

Structure calculations – Structural refinements of SUMO1 and the ZZ domain was done using Xplor-NIH (27). For the ZZ domain, NOE restraints, dihedral and hydrogen bond restraints were adopted from the published structure (Protein Data Bank code 1TOT) (11) and imported from the Biological Magnetic Resonance Bank (BMRB) and converted from Amber (28) to Xplor-NIH format using in-house scripts. Zinc ions were included into the structure calculations by adding distance and angle restraints in a similar way as for the published structure of the ZZ domain of CBP (11). Structure calculations were initiated from an elongated protein structure using a simulated annealing protocol in which NOE, dihedral, hydrogen bond, and RDC restraints were applied as well as restraints for the zinc coordination. Structure calculations were performed in an iterative fashion, where NOE, dihedral angle, and RDC restraints were iteratively pruned from the calculations in order to obtain a converged molecular structure ensemble. Structure analysis were done using the protein structure validation software suite (PSVS) (29) and PALES (30). Residues in the N- and C-terminal (1-5, 49-53) were considered flexible based on observed peak intensities in the \(^{15}\text{N}\)-HSQC, and were excluded from the structure calculations. In the case of the NMR structure of SUMO1 (Protein Data Bank code 1A5R), the restraints are not available. Synthetic NOE distance restraints were therefore created by collecting all \(^1\text{H}-^1\text{H}\) distances in the first model of the structural ensemble of SUMO1 excluding all restraints corresponding to distances longer than 6.5Å or shorter than 2.8Å. The structure calculation protocol was otherwise the same as for the ZZ domain, where NOE distance restraints and RDCs were iteratively pruned. The weight of RDC restraints in the structure calculations was increased as compared to the calculations of the structure of the ZZ domain due to the use of synthetic NOE restraints. Residues in the N- and C-terminal were considered flexible (1-20, 95-103) based on \(R_2^0\) values from the CPMG relaxation dispersion experiments and were excluded from the structure calculations. The structure analysis were performed using the PSVS suite (29) and PALES (30).

Docking simulations - Docking simulations were performed using HADDOCK (13) and CNS (31) following published protocols (32). Chemical shift changes upon binding were used as ambiguous interaction restraints (AIR) in combination with RDCs measured for the ZZ domain-SUMO1 complex. The docking simulations involved active restraints for those residues experiencing chemical exchange, while neighboring residues were treated as passive. Residues with a solvent-accessible surface-area (SASA) < 20% were filtered out. The active residues in the ZZ domain were N36, K38 and A41, while the passive residues were E11, C12, W20, V24, C31, I32, N33, Y35, T37, S39 and H42. The active residues in SUMO1 were H43, K46, M82, E83 and E85, while the passive residues were K23, K25, T42, K45, G81, E84 and V87. RDCs belonging to residues in the flexible N- and C-terminals were not included as restraints in the docking simulations. The protonation state of the two histidine residues (H40 and H42), coordinating the zinc ion in the metal site of the ZZ domain, was set to match the metal coordination pattern (11). Docking simulations were run using the standard HADDOCK protocol with 1000 initial docking simulations, followed by 200 refinement simulations, and subsequently 200 final refinement simulations including explicit water molecules. RDCs were used as intervector projection angle restraints (33) in the first 2 simulation steps and as direct restraints in the final water refinement procedure.

NMR relaxation dispersion experiments – NMR \(^{15}\text{N}\) CPMG relaxation dispersion were measured using CT-CPMG relaxation dispersion experiments (34,35) at static magnetic field strengths of 14.1 and 18.9 T at 10°C. The temperature was calibrated using a methanol sample before each series of experiments. The constant relaxation time was 40 ms. The experiments utilized the phase cycle proposed by Yip and Zuiderweg to suppress artefacts due to off-

Structural studies of a ZZ domain-SUMO1 complex
resonance effects (36). The effective relaxation rate \( R_{2\text{eff}} \) at a given CPMG refocusing frequency was determined from 2 data points, as described (37). Relaxation dispersions were sampled using 18–24 \( R_{2\text{eff}} \) values. The spectral widths were 8000 Hz (\(^{1}H\)) and 1835 Hz (\(^{15}N\)) at 14.1 T and 10666 Hz (\(^{1}H\)) and 2447 Hz (\(^{15}N\)) at 18.9 T, with 512 and 64 points collected in direct and indirect dimensions, respectively. Spectra were processed using NMRpipe (24). The processing protocol involved a solvent filter, cosine-squared apodization functions, zero filling to twice the number of increments in all dimensions, and a polynomial baseline correction in the \(^{1}H\) dimension. Peak intensities were measured as the integral over 5×3 points (\(^{1}H/^{15}N\)) centered on the peak maximum. The signal-to-noise ratio (S/N) was estimated by calculating the standard deviation of 200 samples of integrated 5×3-point windows in empty regions of each spectrum. Errors in the extracted \( R_{2\text{eff}} \) were estimated from the S/N using error propagation. The relaxation dispersion data were analyzed using CPMGfit v2.3 (38). Relaxation dispersion curves were fitted to the Carver-Richards two-state exchange model (39,40):

\[
R_{2\text{eff}} = R_{20} + R_{cex}(1 / \tau_{cp})
\]

in which:

\[
R_{cex}(1 / \tau_{cp}) = \frac{1}{2} \left( k_{cex} - \frac{1}{\tau_{cp}} \right) \cosh^{-1}[D_{cex} \cosh(\eta_{cex}) - D \cos(\eta_{cex})]
\]

\[
\eta_{cex} = \frac{\tau_{cp}}{\sqrt{2}} \left\{ \psi + \left( \frac{\zeta_{cex}^{2}}{\zeta_{cex}^{2}} \right)^{1/2} \right\}
\]

\[
D_{cex} = \frac{1}{2} \left[ 1 + \frac{\psi + 2\Delta \omega_{\alpha}^{2}}{\left( \frac{\zeta_{cex}^{2}}{\zeta_{cex}^{2}} \right)^{1/2}} \right]
\]

and \( \psi = k_{cex}^{2} - \Delta \omega^{2} \), \( \zeta_{cex}^{2} = -2\Delta \omega k_{cex}(1-2p_{m}) \), \( k_{cex} = k_{1} + k_{-1} \) is the sum of the forward and reverse rate constants, \( \Delta \omega \) is the chemical shift difference between the exchanging conformations, \( R_{20} \) is the average limiting value of the relaxation rate constant for processes other than chemical exchange, \( p_{m} \) is the relative population of the minor (less populated) conformational state, which is related to the major conformational state \( p_{M} \) as \( p_{m} = 1 - p_{M} \), and \( \tau_{cp} = 1/n_{cp} \) is the spacing between refocusing pulses in the CPMG train.

The statistical significance of each fit was assessed by also fitting the data to a constant \( R_{20} \) value (i.e. modeling a flat dispersion profile, indicating the absence of exchange), and the \( F \)-test was used to discriminate between models by rejecting the null hypothesis that the model with more parameters does not provide a significantly better fit than the simpler model at the level \( p < 0.001 \). Errors in the fitted parameters were estimated from 1000 synthetic data sets created using Monte-Carlo simulations (41,42).

RESULTS

To define the molecular details in the interaction between the ZZ domain and SUMO1, we employed a combination of chemical shift mapping, RDCs, docking simulations, and protein engineering. In addition, experiments addressing the dynamic properties of the interactions between SUMO1 and SIMs were performed to advance our understanding of SUMO recognition.

Chemical shift mapping of the binding interface between SUMO1 and the ZZ domain - The binding interface on the ZZ domain and SUMO1 was mapped using binding-induced chemical shift differences measured in \(^{15}N\)-HSQC spectra, where the binding was verified by ITC. A comparison of these \(^{15}N\) HSQC spectra shows that 6 residues in the labeled SUMO1 are broadened beyond detection upon the addition of the ZZ domain due to intermediate chemical exchange between the free and bound states (Fig 1A). In a similar fashion, 5 residues in the ZZ domain are broadened beyond detection upon adding SUMO1 (Fig. 1B). The affected residues in SUMO1 are located in a contiguous region of the structure: L24 is situated in a β-strand, and M82, E83 and E85 are found in an adjacent loop, while H43 and K46 are located in the N-terminal part of the α-helix and preceding loop (Fig. 1C). This binding site is spatially distinct from that observed for classical SUMO interacting motifs (SIMs), which is located on the opposite side of the protein (4-6). The perturbed residues in the ZZ domain are located in the α-helix and the following loop (Fig. 1D). The side chains of C34 and H40 are part of the metal coordination for both of the Zn\(^{2+}\) ions present in the ZZ domain, indicating that the presence of bound Zn\(^{2+}\) ions...
might be important for binding SUMO1. The observed chemical shift changes define a binding site for the interaction between the ZZ domain and SUMO1 not previously described. In order to further define the binding interface, attempts were made to record intermolecular NOEs using 3D \(^{1}N/^{13}C\)-filtered NOESY experiments and 2D \(^{1}N/^{13}C\)-filtered NOESY experiments, where the 3D experiment did not provide reliable intermolecular NOEs due to low signal to noise, while the 2D experiments produced highly ambiguous intermolecular NOEs due to extensive spectral overlap. No intermolecular NOEs were therefore included in the generation of the structure of the ZZ domain-SUMO1 complex.

**RDC measurements and molecular refinements** - Backbone H-N bond vector orientations were determined from 44 RDCs for SUMO1 and 47 RDCs for the ZZ domain. The alignment tensors were calculated for each protein using published NMR structures (SUMO1, Protein Data Bank code 1A5R (19) and the ZZ domain, Protein Data bank code 1TOT (11)) with the structure software PALES (30), with Q-values of 0.93 and 0.90, respectively (43). The initial agreement between the RDCs back-calculated from the published structures and the experimental RDCs is low (Fig. 2A, B), most likely because the structures have not been refined using RDC restraints. In order to improve the correlation between experimental and theoretical RDCs, the published protein structures were therefore first refined versus the experimental RDCs. Given that the changes in chemical shifts are minor upon forming the protein-protein complex, we do not expect any major conformational changes upon formation of the ZZ domain-SUMO1 complex. Hence, as an initial approximation, we assumed that the structure of the complex could be modeled using the published structures and the RDCs measured for the complex. Structural restraints have been published for the ZZ domain (Protein Data Bank code 1TOT) and these were converted to Xplor-NIH format using a structure refinement protocol from the program suite (44). RDCs measured for the protein-protein complex as well as NOE distance restraints, dihedral angle restraints, and hydrogen bond restraints were used for the refinement. Prior to refinement, the protein was modeled to correspond to the human ZZ domain by adding a C-terminal Asp and mutating Thr41 to Ala41, followed by energy minimization using Xplor-NIH. Structural refinement was done using the first model of the published structure ensemble as the input structure. Residues in the N- and C-terminal (1-5, 49-53) were considered flexible based on observed peak intensities in the \(^{1}N\)-HSQC. Refinement was performed in an iterative manner until the resulting structure shows no NOE violations > 0.5 Å. Since there are no published distance restraints available for SUMO1, we employed a different approach for the RDC refinement of this structure. Synthetic NOE distance restraints were created by calculating all \(^1H-^1H\) distances in the first model of the published protein structure (Protein Data Bank code 1A5R), filtering out distances larger than 6.5 Å and shorter than 2.8 Å. In addition, the protein structure was prepared for refinement by changing Cys52 to Ser52. Refinement was done iteratively using the same protocol as for the ZZ domain. Residues in the N- and C-terminal were considered flexible (1-20, 95-103) based on CPMG relaxation dispersion experiments and excluded from the calculations. The refinement was performed iteratively until the resulting structure had no NOE violations > 0.5 Å. For each iteration, NOE violations > 0.5 Å and distances > 6.5 Å were filtered out. Q-values after refinement were 0.30 and 0.26 for the ZZ domain and SUMO1, respectively (Fig 2C,D) (43). The refined protein structures for the ZZ domain and SUMO1 are highly similar to the original structures after RDC refinement with RMSD values for C\(_\alpha\) of 0.29 Å and 0.68 Å for the ZZ domain (residues 6-48) and SUMO1 (residues 21-94), respectively. The main structural differences in terms of C\(_\alpha\) RMSDs for the ZZ domain are observed for residues 10-18 in \(\beta\)-strand 1 and in the SUMO1-binding loop 36-46. The main differences in C\(_\alpha\) RMSDs for SUMO1 are seen for residues 33-43 in \(\beta\)-strand 2 and for residues 70-87 in the ZZ domain binding loop.

**Structural model for the ZZ domain-SUMO1 complex** - After final refinement in water, the docked structures were clustered using the standard HADDOCK algorithm, where 115 structures were clustered into one ensemble (Fig 3A, B). Nine additional clusters of structure ensembles were found, each containing 10 structures or fewer. The RMSD of all residues in the interacting protein interface is 1.6 ± 0.6 Å, whereas the RMSD
between all backbone atoms in the docked structure was 1.6 ± 0.6 Å. The RMSD between the protein interface and backbone atoms was 1.3 ± 0.4 Å and 2.2 ± 0.6 Å for SUMO1 and the ZZ domain, respectively (Fig. 3C, D). The HADDOCK score for the final structure ensemble was -43 ± 18, and the Q-value calculated for the complex was 0.36 showing that the experimental RDCs demonstrate a good agreement with the calculated RDCs for the modeled protein complex (Fig. 3E) (43). The magnitude and orientation of the RDC alignment tensor for the ZZ domain and SUMO1 and the ZZ domain-SUMO1 complex was calculated using Module (45), where the magnitudes and rhombocities are shown in Table 1. Using Module, the RDC alignment tensor was fitted separately for the ZZ domain and SUMO1 in the protein complex indicating highly similar tensor orientations and magnitudes (Fig 3F). The observed structures for the ZZ domain and SUMO1 in the protein complex are nearly identical to the individually refined structures of the uncomplexed proteins with Co RMSDs of 0.001 Å for both the ZZ domain and SUMO1. By extending the comparisons between complexed and uncomplexed proteins to include all atoms in the assessments, also incorporating all surface accessible sidechains, the RMSDs are 2.32 Å and 2.29 Å for the ZZ domain and SUMO1 respectively.

As expected, the binding interface for the docked protein-complex corresponds to the surface residues experiencing chemical shift perturbations upon binding. The binding surface on the ZZ domain consists of the residues spanning from the C-terminal end of the α-helix to the loop connecting the α-helix to β-strand 1. The binding surface on SUMO1 consists of residues belonging to β-strand 1, the N-terminal part of the α-helix and the preceding loop as well as the long loop connecting β-strands 3 and 4. The binding surface on SUMO1 consists of a convex groove on the protein surface. This interaction surface represents a completely new binding epitope for SUMO1, previously not observed in the Protein Data Bank, which is clearly separated in space from known interaction motifs. The previously known interaction motifs for SUMO1 are the SUMO-SIM motif (4-6) the Ubc9 motif and the SENP/RanGAP motif (46,47). None of these interaction motifs are overlapping with the ZZ domain-SUMO1 binding site, suggesting that the ZZ domain can bind to SUMO1 in the presence of other interaction motifs.

In order to verify the binding mode observed in the ZZ domain-SUMO1 complex, we created one mutant of the ZZ domain and two mutants of SUMO1, where we mutated residues in the binding interface to alanines. The residues selected for alanine substitution where either residues with charged side chains or with side chains with the potential to form hydrogen bonds, thus removing important electrostatic interactions in the binding interface. We find that the ZZ domain mutant still has the ability to bind SUMO1, albeit with significantly lower enthalpy of binding as compared to the wildtype (Fig. 4A, B and Table 2). The two mutants of SUMO1, did however, completely abolish binding to the ZZ domains (Fig. 5A-D), clearly indicating that these residues are important for the complex formation and further supporting the molecular model derived from the NMR data.

To compare the SUMO binding characteristics of the ZZ domain with a canonical SIM motif, we performed ITC experiments between SUMO1 and a SIM peptide denoted SIMPX from PIASx (6). As expected, we could show that SIMPX binds to SUMO1 with a similar affinity as the ZZ domain (Fig. 4C and Table 2). To investigate whether SIMPX and the ZZ domain can bind to SUMO simultaneously, ITC experiments between SIMPX and a pre-formed complex between the ZZ domain and SUMO1 were carried out. These experiments revealed that SIMPX binds with a similar affinity and enthalpy of binding to SUMO1 as it does in the absence of the ZZ domain, indicating that SIMPX can bind SUMO1 independent of the ZZ domain in a non-cooperative manner (Fig. 4 C, D and Table 2).

**CPMG relaxations dispersion experiments** - ¹⁵N CPMG relaxation dispersions were measured for apo, SIMPX-bound, and ZZ domain-bound SUMO1 at static magnetic field strengths of 14.1 and 18.9 T (Supplemental Table 1). Apo-SUMO1 showed 52 backbone amides experiencing significant conformational exchange (p<0.01), as exemplified in Fig. 6A. The exchanging residues are highlighted in Fig. 6D, showing that large parts of the core domain (residues 20-95) are undergoing exchange. Each residue was initially fitted to residue-specific exchange parameters, k_ex and p_n. Residues in the core domain with a k_ex < 2000 s⁻¹
were fitted to a global exchange process, resulting in $k_{ex} = 1185 \pm 91 \text{ s}^{-1}$ and $p_m = 0.987 \pm 0.002$.

$^{15}$N CPMG relaxation dispersions for the SIMPX-bound SUMO1 revealed 30 residues experiencing exchange. (Fig. 6B, E). Residues were selected for a global fit based on the same criteria as for apo-SUMO1, and were fitted to a global exchange rate $k_{ex} = 479 \pm 30 \text{ s}^{-1}$ and major population $p_m = 0.942 \pm 0.01$ (Table 3).

$^{15}$N CPMG relaxation dispersion experiments on ZZ domain-bound SUMO1 showed 48 residues with significant exchange (Fig. 6C, F). Again, residues in the core domain of SUMO1 were selected for the global fit, giving a global exchange rate $k_{ex} = 1418 \pm 203 \text{ s}^{-1}$ and major population $p_m = 0.978 \pm 0.019$ (Table 3). In comparison to apo-SUMO1, the ZZ domain bound SUMO1 has a larger number of residues in the flexible N-terminal exhibiting significant dispersions. Residues belonging to the N-terminal part (2-5, 7-8, 10, 13-16 and 18-19) were fitted to a global exchange rate $k_{ex}= 1924 \pm 145 \text{ s}^{-1}$ and a major population $p_m = 0.990 \pm 0.002$.

Comparison of exchanging states in SUMO1
- Upon addition of SIMPX, most of the backbone residues in SUMO1 experiences a significant change in chemical environment as seen from a comparison of the crosspeak positions in the $^{15}$N HSQC for SIMPX-bound and apo-SUMO1 (Fig. 7A). Residues not affected by the binding of SIMPX are located in the N- and C-terminal parts of SUMO1, whereas most backbone residues in the core region of SUMO1 experience a shift in chemical environment upon addition of SIMPX (Fig. 7B, C). However, only a subset of the residues affected by the binding of the SIM-peptide is within 5 Å of SIMPX, as gauged from the SIM-SUMO1 structure (Protein Data Bank code 2ASQ). This result indicates that the binding of SIMPX induces a conformational change in SUMO1. A closer look at the residues exhibiting conformational exchange in apo-SUMO1 reveals that these are primarily located in or close to the SIM-binding site. In summary, all residues within 5 Å of SIMPX, with the exception of residues 39, 43, 48, and 53, exhibit conformational exchange in apo-SUMO1.

Upon addition of SIMPX to SUMO1, the exchange is quenched for many residues, but 20 residues show exchange in both apo- and SIMPX-bound SUMO1. Of these, 6 residues are within 5 Å of SIMPX. A small number of residues in the SIMPX-SUMO1 complex experiences chemical exchange not detected for apo-SUMO1, and of these, only one residue is within 5 Å of the SIM binding site. The remainders of these residues are located in the flexible N-terminal, or in the hydrophobic core of SUMO1, suggesting a change in the hydrophobic packing upon binding.

A comparison of residues that show exchange in the apo- and ZZ domain-bound states of SUMO1 reveals that these groups of residues are similar. In addition, a set of residues shows exchange in the ZZ domain-SUMO1 complex that is not observed in apo-SUMO1. Most of these residues are located in the flexible N-terminal region, which is involved in binding the ZZ domain and therefore flexible in the apo state. Residues that show exchange for apo-SUMO1 but not for the ZZ domain-SUMO1 complex are mainly located in the region indicated to bind the ZZ domain based on the chemical shift perturbations observed for SUMO1.

The $^{15}$N chemical shift changes between the major and minor state was determined from global fits of the CPMG relaxation dispersions for the various states of SUMO1. The chemical shift changes extracted from the global fit of the apo-SUMO1 CPMG relaxation dispersion experiments does not show a direct one-to-one correspondence with the $^{15}$N chemical shift changes observed between the SIMPX-bound and apo-SUMO1 (Fig. 8A) with an RMSD of 0.96 ppm. The data suggest that apo-SUMO1 is exchanging between a major ground state and a high-energy state that involves other conformations than those of the SIMPX-bound state. In contrast, the chemical shift change between the minor and major state in the SIMPX-SUMO1 complex correlates well with the $^{15}$N chemical shift change measured from the $^{15}$N HSQC of the apo- and SIMPX-bound SUMO1 with an RMSD of 0.08 ppm (Fig. 8B). These data imply that SUMO1 is exchanging between free and bound states in the SIMPX-SUMO1 sample. Furthermore, the comparison of the globally fitted chemical shift changes from the CPMG relaxation dispersions between the minor and major state for apo- and ZZ domain-bound SUMO1 yields a good correlation with an RMSD of 0.14 ppm (Fig. 8C). Notably, the chemical exchange rates and the populations of ZZ domain-bound SUMO1 and apo-SUMO1 are the same within errors and with a similar set of residues showing exchange broadening in both states. These observations might suggest that the binding of the
ZZ domain to SUMO1 has no major effect on the intrinsic conformational exchange of SUMO1. A direct comparison of the chemical shift change between the ZZ domain bound SUMO1 and apo-SUMO1 is not feasible due to the lack of 15N cross peaks caused by exchange broadening upon formation of the ZZ domain-SUMO1 complex. An alternative interpretation of the data is that the CPMG dispersions observed for the ZZ domain-bound SUMO1 sample involves exchange between free and bound states; in this scenario, the good match between the chemical shifts of the minor state of apo-SUMO1 and the ZZ domain-bound SUMO1 indicates that SUMO1 binds the ZZ-domain by conformational selection. The latter interpretation is further supported by the expectation, based on the binding affinities from ITC and the NMR sample conditions, that both the SIMPX- and ZZ-domain SUMO1 samples contain a minor population of the peptide-free state, which is in good agreement with the minor state population determined from the CPMG dispersions. The on- and off-rates of the various states was calculated based on a global fit of the populations and exchange rates. For the SIMPX-SUMO1 complex, the \( k_{\text{off}} \) is \( 28 \pm 2 \) s\(^{-1}\) and \( k_{\text{on}} \) is \( 1.6 \pm 0.1 \times 10^6 \) M\(^{-1}\)s\(^{-1}\). The calculated value of \( k_{\text{on}} \) indicates that the binding of SIMPX to SUMO1 is essentially diffusion controlled. In comparison, the off-rate for the ZZ-domain-SUMO1 complex is \( 31 \pm 5 \) s\(^{-1}\), as calculated for the core residues, while the results for the N-terminal residues directly involved in binding of the ZZ-domain peptide yield \( k_{\text{off}} = 19 \pm 2 \) s\(^{-1}\). We speculate that the release of the ZZ-domain peptide might involve a two-step process, with slightly different life times of the interactions with the N-terminal segment and the core residues. The overall slower off-rate for the ZZ-domain is in keeping with the slightly higher affinity of the ZZ domain-SUMO1 complex than SIMPX-SUMO1. The off-rates observed for the different bound states of SUMO1 are in the same regime as the intrinsic conformational exchange of apo-SUMO1, suggesting that these processes might be linked.

**DISCUSSION**

Protein modifications, such as SUMOylation, play important roles in intricate protein-protein interaction networks involving numerous intracellular processes such as the response to DNA-double strand breaks and chromatin remodeling. In the response to DNA damage, the SUMOylated target proteins are recognized by SUMO Interactions Motifs (SIMs) present in various DNA damage signal and repair proteins. Until recently, one type of SIM had been described that is characterized by a signature sequence involving a stretch of charged anionic amino-acid residues (10). The newly discovered SUMO-binding ZZ domain represents a new SUMO recognition motif distinct from the classical SIMs (9). We now show that the ZZ domain interacts with a unique epitope in SUMO1 and that the complex between the ZZ domain and SUMO1 displays a different binding mode from those previously described for SUMO-mediated interactions. Mutation of residues in the interaction surface for the ZZ domain did not significantly decrease the affinity for SUMO1 even though the enthalpy of binding is reduced. From previous studies we have established that the zinc-coordination is crucial for the interaction between the HERC2 ZZ domain and SUMO1 (9). By mutating the zinc-coordinating residues in the ZZ domain, the SUMO-binding capacity was lost. We therefore conclude that the presence of the coordinated zinc in the ZZ domain close to the interaction interface with SUMO1 play an important role in the ZZ domain-SUMO1 complex formation. Importantly, mutation of two epitopes in SUMO1, representing residues in the ZZ domain-SUMO1 interface, completely eliminates binding to the ZZ domain, clearly indicating the importance of these epitopes in ZZ domain recognition.

The CPMG relaxation dispersion data demonstrate that apo-SUMO1 exchanges between the (major) ground state and a (minor) high-energy state, which might be linked to ligand binding. However, it is clear that the minor conformation of apo-SUMO1 does not directly match that of the bound state in the case of SIMPX binding, indicating that this peptide does not bind by strict conformational selection, but induces additional conformational changes upon binding, or that it selects one conformer out of a manifold present in the high-energy state. Also in the case of the ZZ domain, the available data can be interpreted in two ways: either the peptide does not affect the intrinsic dynamics of apo-SUMO1 or it binds by conformational selection. Thus, the available data clearly show that the dynamic process of peptide binding to SUMO1 has markedly
different signatures for the two cases of SIMPX and the ZZ domain. 
Taken together we have shown that the new SUMO-binding ZZ domain binds to SUMO1 through interactions with a unique and spatially distinct site in SUMO1 as compared to the SIMs, and that these two kinds of recognition motifs can bind SUMO1 simultaneously in a non-cooperative fashion. Our data therefore provide new insights into the interactions between SUMO binding motifs and SUMO, and extend the repertoire of proteins recognizing SUMO to also include regulatory proteins harboring ZZ domains.

Acknowledgements: We would like to thank Havva Koç for excellent technical assistance.

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions: CD expressed and purified the proteins, and conducted the NMR experiments including analysis of results. MA contributed with experimental design and interpretation of the CPMG relaxation dispersion experiments. SBJ and NL provided biological data assessments and expertise. WS and MW performed the ITC experiments. MW designed the study, served as the project supervisor, and wrote the paper with input from all other authors.
REFERENCES

1. Geiss-Friedlander, R., and Melchior, F. (2007) Concepts in sumoylation: a decade on. Nature reviews. Molecular cell biology 8, 947-956
2. Gareau, J. R., and Lima, C. D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nature reviews. Molecular cell biology 11, 861-871
3. Becker, J., Barysch, S. V., Karaca, S., Dittner, C., Hsiao, H. H., Berriel Diaz, M., Herzig, S., Urlaub, H., and Melchior, F. (2013) Detecting endogenous SUMO targets in mammalian cells and tissues. Nature structural & molecular biology 20, 525-531
4. Rodriguez, M. S., Dargemont, C., and Hay, R. T. (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. The Journal of biological chemistry 276, 12654-12659
5. Sampson, D. A., Wang, M., and Matunis, M. J. (2001) The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. The Journal of biological chemistry 276, 21664-21669
6. Song, J., Zhang, Z., Hu, W., and Chen, Y. (2005) Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. The Journal of biological chemistry 280, 40122-40129
7. Chang, C. C., Naik, M. T., Huang, Y. S., Jeng, J. C., Liao, P. H., Kuo, H. Y., Ho, C. C., Hsieh, Y. L., Lin, C. H., Huang, N. J., Naik, N. M., Kung, C. C., Lin, S. Y., Chen, R. H., Chang, K. S., Huang, T. H., and Shih, H. M. (2011) Structural and functional roles of Daxx SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. Molecular cell 42, 62-74
8. Cappadocia, L., Mascle, X. H., Bourdeau, V., Tremblay-Belzile, S., Chaker-Margot, M., Lussier-Price, M., Wada, J., Sakaguchi, K., Aubry, M., Ferbeyre, G., and Omichinski, J. G. (2015) Structural and functional characterization of the phosphorylation-dependent interaction between PML and SUMO1. Structure (London, England : 1993) 23, 126-138
9. Danielsen, J. R., Povlsen, L. K., Villumsen, B. H., Streicher, W., Nilsson, J., Wikström, M., Bekker-Jensen, S., and Mailand, N. (2012) DNA damage–inducible SUMOylation of HERC2 promotes RNF8 binding via a novel SUMO-binding Zinc finger. The Journal of cell biology 197, 179-187
10. Hecker, C.-M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006) Specification of SUMO1-and SUMO2-interacting motifs. Journal of Biological Chemistry 281, 16117-16127
11. Legge, G. B., Martinez-Yamout, M. A., Hambly, D. M., Trinh, T., Lee, B. M., Dyson, H. J., and Wright, P. E. (2004) ZZ domain of CBP: an unusual zinc finger fold in a protein interaction module. Journal of molecular biology 343, 1081-1093
12. Ponting, C. P., Blake, D. J., Davies, K. E., Kendrick-Jones, J., and Winder, S. J. (1996) ZZ and TAZ: new putative zinc fingers in dystrophin and other proteins. Trends in biochemical sciences 21, 11-13
13. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. Journal of the American Chemical Society 125, 1731-1737
14. Aslanidis, C., and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). Nucleic acids research 18, 6069-6074
15. Wang, W., and Malcolm, B. (1999) Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. Biotechniques 26, 680-682
16. Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974) Culture medium for enterobacteria. Journal of bacteriology 119, 736-747
17. Gileadi, O., Burgess-Brown, N. A., Colebrook, S. M., Berrett, G., Savitsky, P., Smee, C. E., Loppi, M., Johansson, C., Salah, E., and Pantic, N. H. (2008) High throughput production of recombinant human proteins for crystallography. in Structural Proteomics, Springer. pp 221-246
Wiseman, T., Williston, S., Brandts, J., and Lin, L. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* **179**, 131-137

Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R., and Becker, J. (1998) Structure determination of the small ubiquitin-related modifier SUMO-1. *Journal of molecular biology* **280**, 275-286

Clore, G. M., Bax, A., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) Assignment of the side-chain proton and carbon-13 resonances of interleukin-1 beta. using double-and triple-resonance heteronuclear three-dimensional NMR spectroscopy. *Biochemistry* **29**, 8172-8184

Bax, A., and Ikura, M. (1991) An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins. *Journal of biomolecular NMR* **1**, 99-104

Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha-and beta-carbon resonances in proteins. *Journal of Magnetic Resonance, Series B* **101**, 201-205

Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *Journal of the American Chemical Society* **114**, 6291-6293

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. *Journal of biomolecular NMR* **6**, 277-293

Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696

Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *Journal of magnetic resonance (San Diego, Calif.* : **1997** 131, 373-378

Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *Journal of magnetic resonance (San Diego, Calif.* : **1997** 160, 65-73

Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *Journal of computational chemistry* **26**, 1668-1688

Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins* **66**, 778-795

Zweckstetter, M., and Bax, A. (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. *Journal of the American Chemical Society* **122**, 3791-3792

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., and Pannu, N. S. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallographica Section D: Biological Crystallography* **54**, 905-921

de Vries, S. J., van Dijck, M., and Bonvin, A. M. (2010) The HADDOCK web server for data-driven biomolecular docking. *Nature protocols* **5**, 883-897

Meiler, J., Blomberg, N., Nilges, M., and Griesinger, C. (2000) A new approach for applying residual dipolar couplings as restraints in structure elucidation. *Journal of biomolecular NMR* **16**, 245-252

Loria, J. P., Rance, M., and Palmer, A. G. (1999) A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *Journal of the American Chemical Society* **121**, 2331-2332

Mulder, F. A., Mittermaier, A., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Studying excited states of proteins by NMR spectroscopy. *Nature structural & molecular biology* **8**, 932-935
Structural studies of a ZZ domain-SUMO1 complex

36. Yip, G. N., and Zuiderweg, E. R. (2004) A phase cycle scheme that significantly suppresses offset-dependent artifacts in the $\text{R}^{i} \text{R}^{i+1}$-CPMG experiment. *Journal of Magnetic Resonance* **171**, 25-36

37. Mulder, F. A. A., Mittermaier, A., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Studying excited states of proteins by NMR spectroscopy. *Nature Structural & Molecular Biology* **8**, 932-935

38. Weininger, U., Modig, K., and Akke, M. (2014) Ring Flips Revisited: 13C Relaxation Dispersion Measurements of Aromatic Side Chain Dynamics and Activation Barriers in Basic Pancreatic Trypsin Inhibitor. *Biochemistry* **53**, 4519-4525

39. Carver, J. P., and Richards, R. E. (1972) A general two-site solution for the chemical exchange produced dependence of T2 upon the Carr-Purcell pulse separation. *Journal of Magnetic Resonance (1969)* **6**, 89-105

40. Davis, D. G., Perlman, M. E., and London, R. E. (1994) Direct Measurements of the Dissociation-Rate Constant for Inhibitor-Enzyme Complexes via the T1 [$\rho$] and T2 (CPMG) Methods. *Journal of Magnetic Resonance, Series B* **104**, 266-275

41. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986) *Numerical Recipes. The Art of Scientific Computing*, Cambridge University Press, Cambridge

42. Mandel, A. M., Akke, M., and Palmer, A. G. (1995) Backbone dynamics of *Eschericia coli* Ribonuclease HI: Correlations with structure and function in an active enzyme. *JMB* **246**, 144-163

43. Cornilesu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of biomolecular NMR* **13**, 289-302

44. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Marius Clore, G. (2003) The Xplor-NIH NMR molecular structure determination package. *Journal of Magnetic Resonance* **160**, 65-73

45. Dosset, P., Hus, J. C., Marion, D., and Blackledge, M. (2001) A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *Journal of biomolecular NMR* **20**, 223-231

46. Bernier-Villaamor, V., Sampson, D. A., Matunis, M. J., and Lima, C. D. (2002) Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* **108**, 345-356

47. Reverter, D., and Lima, C. D. (2006) Structural basis for SENP2 protease interactions with SUMO precursors and conjugated substrates. *Nature structural & molecular biology* **13**, 1060-1068

48. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics* **14**, 51-55

49. DeLano, W. L. (2002) The PyMOL molecular graphics system.
FOOTNOTES

This work was supported by the Novo Nordisk Foundation grant NNF14CC0001, the European Commission under the Seventh Framework Program (FP7) contract Bio-NMR 261863 providing access to the Swedish NMR Centre, Gothenburg, Sweden, and the MAX laboratory, Lund, Sweden enabling access to the MAX laboratory (ID 20130289).

The chemical shifts for the ZZ domain and SUMO1 (code 25553) have been deposited in the Biological Magnetic Resonance Data Bank (http://www.bmrb.wisc.edu/). The atomic coordinates for the model of the ZZ domain-SUMO1 complex (Protein Data Bank code 2N1A) have been deposited in the Protein Data Bank (http://wwpdb.org/).

To whom correspondence may be addressed: Mats Wikström, the Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2200 Copenhagen, Denmark. Present address: Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320, United States. E-mail: mats.wikstrom@amgen.com

The abbreviations used are: ZZ domain, CBP/p300 ZZ domain; RDC, residual dipolar coupling; SUMO, Small Ubiquitin-related modifier; SIM, SUMO interacting motif; HSQC, heteronuclear single quantum coherence; ITC, isothermal titration calorimetry; AIR, ambiguous interaction restraints; SASA, solvent-accessible surface-area; HADDUCK, high ambiguity driven protein-protein docking.

FIGURE LEGENDS

FIGURE 1. Chemical shift differences observed when titrating the unlabeled ZZ domain or SUMO1 into $^{15}$N$^{13}$C-labeled SUMO1 or ZZ domain respectively. A. Cutout of $^{15}$N HSQC showing backbone amides in SUMO1 affected by the binding of the ZZ domain. B. Cutout of $^{15}$N HSQC showing backbone amides in the ZZ domain affected by the binding of SUMO1. Contours colored in black correspond to the apo-form, whereas contours colored in red correspond to ZZ domain-bound SUMO1. C. Residues in SUMO1 affected by the binding of ZZ highlighted in red: L24, H43, K46, M82, E83, and E85 (Protein Data Bank code 1A5R). D. Residues in the ZZ domain affected by the binding of SUMO1 highlighted in red: C34, N36, K38, H40, and A41. Zinc ions are depicted as blue spheres. (Protein Data Bank code 1TOT). Protein images were made using Molmol (48).

FIGURE 2. Residual dipolar couplings for SUMO1 (A, C) and the ZZ domain (B, D) displayed before and after structure refinement using residual dipolar couplings as additional restraints. Experimental RDCs are plotted versus back calculated RDCs as calculated from the initial (A, B) and the final structures after refinement (C, D).

FIGURE 3. A. Model of the ZZ domain-SUMO1 complex shown in a surface representation made using Pymol (49). SUMO1 is colored green, the ZZ domain is colored red, and the peptide corresponding to a SIM-motif colored in blue (Protein Data Bank code 2ASQ (6)). B. Model of the ZZ domain-SUMO1 complex shown in a ribbon representation using the same color scheme as in A with the two zinc ions depicted as blue spheres. The residues in SUMO1 and the ZZ domain affected by the interaction in the NMR epitope mapping experiments are indicated in the model, corresponding to the same residues shown for the individual protein models in Fig. 1 C,D. C. Per residue RMSD between interface residues and backbone Ca atoms in the modeled complex for SUMO1 residues 1-103. The location for the secondary structure elements of SUMO1 are indicated by arrows (yellow) for β-strands and a cylinder for the single α-helix (red). D. Per residue RMSD between interface residues and backbone Ca atoms in the modeled complex for the ZZ domain residues 1-53. The location for the secondary structure elements of the ZZ domain are
indicated by arrows (yellow) for β-strands and a cylinder for the helical segment (red). E. Experimental RDCs plotted versus calculated RDCs for the modeled protein complex. F. Orientations of the RDC alignment tensor for the ZZ domain and SUMO1 in the ZZ domain-SUMO1 complex, in which tensor orientations were fitted using Module (45). The ZZ domain (yellow) and SUMO1 (blue) are shown in ribbon representations.

**FIGURE 4.** Isothermal titration (ITC) experiments showing the binding between the ZZ domain and SUMO1, the SIM peptide and SUMO1, and the binding of the SIM peptide to the complex between the ZZ domain and SUMO1. The experiments show the titrations of: A. The binding of the wildtype ZZ domain (ZZwt) to the wildtype SUMO1 (SUMO1wt). B. The binding of the ZZ domain mutant (ZZmut) to the wildtype SUMO1 (SUMO1wt). C. The binding of the SIM peptide to SUMO1. D. The binding of the SIM peptide to the preformed complex between the ZZ domain and SUMO1. The raw data of the experiments are presented on the top panel. The area underneath each injection peak is equal to the total heat released for that injection.

**FIGURE 5.** Isothermal titration (ITC) experiments between ZZ domain variants and SUMO1 variants. The experiments show the titrations of: A. Wildtype ZZ domain (ZZwt) to the SUMO1 mutant 1 (SUMO1mut1). B. ZZwt to SUMO1 mutant 2 (SUMOmut2). C. ZZ domain mutant (ZZmut) to SUMO1mut1. D. ZZmut to SUMO1mut2. The raw data of the experiments are presented on the top panel. The area underneath each injection peak is equal to the total heat released for that injection.

**FIGURE 6.** Example of 15N CPMG relaxation dispersion curves for apo-, SIMPX-bound and ZZ domain-bound states of SUMO1 showing transverse relaxation rates R_2 plotted versus the effective field ν_{cp}. The solid line corresponds to a CPMG relaxation dispersion curve fitted to a 2-state chemical exchange process, whereas the dotted line corresponds to a model with no chemical exchange. A. CPMG relaxation dispersion curve for V38 in apo-SUMO1. B. CPMG relaxation dispersion curve for E20 in SIMPX-bound SUMO1. C. CPMG relaxation dispersion curve for V38 in the ZZ domain-bound SUMO1. Residues exhibiting significant (p<0.01) CPMG relaxation dispersion curves are colored red on the structure of SUMO1. Structures are depicted in cartoon representation. D. Residues with significant CPMG relaxation dispersions in apo-SUMO1: 7, 13, 15-16, 18-26, 28-38, 40, 42, 45-47, 49-50, 54-56, 61, 65, 67, 69-70, 74-76, 78, 81-83, 87, 90, 92, 94, 100-101 (Protein Data Bank code 1ASR). E. Residues with significant CPMG relaxation dispersions in SIMPX-bound SUMO1: 7, 10, 15, 17-18, 20-21, 23, 26-28, 32, 38, 42-43, 48, 55, 57, 61-62, 64-65, 70, 74, 80-81, 83, 87, 100-101 (Protein Data Bank code 2ASQ). SIMPX is colored blue. F. Residues with significant CPMG relaxation dispersions in ZZ domain-bound SUMO1: 2-5, 7-8, 10, 13-16, 18-22, 24-38, 40, 42, 45, 47, 49-50, 60-61, 67, 70, 74, 81-82, 90, 92, 94, 100. The ZZ domain is colored blue and the zinc ions are depicted as green spheres. In panels D, E and F, a selected set of residues from different regions of SUMO1 with significant CPMG relaxation dispersions are indicated with their respective residue number.

**FIGURE 7.** Chemical shift differences observed when titrating SIMPX into 15N13C SUMO1. A. Cutout of 15N HSQC of SUMO1 showing backbone amides in SUMO1 affected by the binding of SIMPX. Black corresponds to apo and red corresponds to the SIMPX-bound SUMO1. The affected residues are indicated by their respective residue number. B. Weighted 1H, 15N chemical shift differences between the SIMPX- and the apo-state of SUMO1 plotted per backbone residue. The location for the secondary structure elements of SUMO1 are indicated by arrows (yellow) for β-strands and a cylinder for the single α-helix (red). C. SUMO1 is colored in green, residues with a significant weighted chemical shift differences (>0.05 PPM) are colored in red, where a subset is indicated by their respective residue number, while SIMPX is shown in blue (Protein Data Bank code 2ASQ).
FIGURE 8. A. \(^{15}\)N Chemical shift difference for backbone amides between major and minor state of apo-SUMO1 from global fits of CPMG relaxation dispersion curves plotted versus \(^{15}\)N chemical shift difference between apo- and SIMPX-bound states. B. \(^{15}\)N Chemical shift difference for backbone amides between the major and minor state of SIMPX-bound SUMO1 from global fits of CPMG relaxation dispersion curves plotted versus \(^{15}\)N chemical shift difference between apo- and SIMPX-bound states. C. \(^{15}\)N chemical shift differences for backbone amides between major and minor state from global fits of CPMG relaxation dispersion curves for apo- and ZZ domain-bound SUMO1 plotted in a covariance graph.
**TABLES**

**TABLE 1.** Summary of structural statistics for the structural refinement using RDCs of the ZZ domain and SUMO1, as well as structural statistics of the docked ZZ domain-SUMO1 complex. Structural analysis was performed using PSVS (29) and PALES (30). The magnitude and orientation of the RDC alignment tensor for the ZZ domain and SUMO1, and the ZZ domain-SUMO1 complex was calculated using Module (45). *AIR restraints indicate the ambiguous restraints used in the HADDOCK docking procedure.

NMR and refinement statistics for the protein structures

|                  | ZZ domain | SUMO1 | ZZ domain-SUMO1 complex |
|------------------|-----------|-------|-------------------------|
| **NMR distance and dihedral constraints** |           |       |                         |
| Distance constraints | ZZ domain | SUMO1 | ZZ domain-SUMO1 complex |
| Total NOE        | 547       | 3304  |                         |
| Intra-residue    | 135       | 476   |                         |
| Inter-residue    |           |       |                         |
| Sequential (|i-j| = 1) | 172      | 1073  |                         |
| Medium-range (|i-j| < 4) | 78       | 644   |                         |
| Long-range (|i-j| > 5) | 162      | 1111  |                         |
| Intermolecular   |           |       |                         |
| Hydrogen bonds   | 5         | 0     |                         |
| Total dihedral angle restraints | 80       | 0     |                         |
| Residual dipolar coupling restraints | 42       | 37    | 79                      |
| AIR restraints*  |           |       |                         |
| Magnitude (D_a)  | 6.03      | 5.95  | 5.89                    |
| Rhombicity       | 0.56      | 0.536 | 0.56                    |
| Q-factor         | 0.30      | 0.26  | 0.36                    |
| **Structure statistics** | ZZ domain | SUMO1 | ZZ domain-SUMO1 complex |
| Violations (RMS) | ZZ domain | SUMO1 | ZZ domain-SUMO1 complex |
| Distance constraints (Å) | 0.08Å | 0.05Å |
| Dihedral angle constraints (°) | 1.73° |
| Max. dihedral angle violation (°) | 8.60° |
Structural studies of a ZZ domain-SUMO1 complex

Max. distance constraint violation (Å) 0.55Å 0.77Å

Structure Quality Factors - overall statistics

| Quality Factor | Value 1 | Value 2 | Value 3 |
|----------------|---------|---------|---------|
| Procheck G-factor (phi / psi only) | -1.18 | -0.65 | -0.83 |
| Procheck G-factor (all dihedral angles) | -1.04 | -0.39 | -0.56 |
| Verify3D | 0.27 | 0.30 | 0.29 |
| ProsaII (-ve) | -0.08 | 0.41 | -0.29 |
| MolProbity clashscore | 64.01 | 43.88 | 54.67 |

Ramachandran Plot Summary from Procheck

| Region type | ZZwt | ZZmut | SIMPX |
|-------------|------|-------|-------|
| Most favoured regions (%) | 63.3 | 70.0 | 77.6 |
| Additionally allowed regions (%) | 30.6 | 22.2 | 15.1 |
| Generously allowed regions (%) | 2.0 | 4.4 | 7.2 |
| Disallowed regions (%) | 4.1 | 3.3 | 0.1 |

TABLE 2. Summary of thermodynamic parameters from ITC experiments of the binding of the ZZ domain wildtype (ZZwt), the ZZ domain mutant (ZZmut), and the SIM peptide SIMPX (6) to SUMO1 respectively. The fourth row shows the data for the binding of SIMPX to the pre-formed complex between the ZZ domain (ZZwt) and SUMO1.

| Protein complex | N (sites) | Kd (µM) | ΔH (kcal/mol) | ΔS (cal/mol/deg) |
|-----------------|----------|---------|---------------|------------------|
| ZZwt/SUMO1      | 1.04±0.06| 5.4±0.2 | -1.7±0.1      | 18.1             |
| ZZmut/SUMO1     | 0.86±0.07| 6.5±0.3 | -0.72±0.08    | 21.2             |
| SIMPX/SUMO1     | 0.87±0.01| 6.2±0.6 | -2.49±0.05    | 15.5             |
| SIMPX/(ZZwt-SUMO1) | 0.88±0.02| 6.1±0.7 | -2.27±0.06    | 16.2             |

TABLE 3. Chemical exchange rates (kex), off-rates (koff) and major populations (pM) for the three different states of SUMO1 extracted from the global fits of significant CPMG relaxation dispersion curves.

| State            | kex (s⁻¹) | koff (s⁻¹) | pM       |
|------------------|-----------|------------|----------|
| Apo-SUMO1        | 1185 ± 91 | 28 ± 2     | 0.987 ± 0.002 |
| SIMPX-SUMO1      | 479 ± 30  | 15 ± 1     | 0.942 ± 0.01  |
| ZZ domain-SUMO1  | 1418 ± 203| 31 ± 5     | 0.978 ± 0.019 |
Structural studies of a ZZ domain-SUMO1 complex

Figure 1
Figure 2

A, B, C, D: Scatter plots showing the relationship between RDC calculated and RDC experimental values. Each plot includes data points and a trend line.
Structural studies of a ZZ domain-SUMO1 complex

Figure 3
Figure 4

A

B

C

D

Structural studies of a ZZ domain-SUMO1 complex
Figure 5

![Graph A](image1)

![Graph B](image2)

![Graph C](image3)

![Graph D](image4)
Figure 6
Figure 7
Figure 8

A

B

C
Structural analysis of a complex between small ubiquitin-like modifier 1 (SUMO1) and the ZZ domain of CREB-binding Protein (CBP/p300) reveals a new interaction surface on SUMO
Carl Diehl, Mikael Akke, Simon Bekker-Jensen, Niels Mailand, Werner Streicher and Mats Wikström

J. Biol. Chem. published online April 25, 2016

Access the most updated version of this article at doi: 10.1074/jbc.M115.711325

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2016/04/25/M115.711325.DC1