8020 *Biochemistry* **2010,** *49,* 8020–8032
DOI: 10.1021/bi101153w

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Structural Analysis of the Smad2−MAN1 Interaction That Regulates Transforming Growth Factor-β Signaling at the Inner Nuclear Membrane[†]

Emilie Kondé,[‡,⊥] Benjamin Bourgeois,[‡,⊥] Carine Tellier-Lebegue,[‡] Wei Wu,[§] Javier Pérez,[∥] Sandrine Caputo,[‡] Wika Attanda,[‡] Sylvaine Gasparini,[‡] Jean-Baptiste Charbonnier,[‡] Bernard Gilquin,[‡] Howard J. Worman,[§] and Sophie Zinn-Justin*[,‡]

[‡]*Laboratoire de Biologie Structurale et Radiobiologie, URA CNRS 2096, CEA Saclay, 91190 Gif-sur-Yvette, France,* [§]*Department of Medicine and Department of Pathology and Cell Biology, College of Physicians and Surgeons, Columbia University, New York, New York 10032, and* [∥]*Synchrotron SOLEIL, BP 48, 91192 Gif-sur-Yvette Cedex, France.* [⊥]*These authors contributed equally to this work.*

Received July 20, 2010; Revised Manuscript Received August 16, 2010

ABSTRACT: MAN1, an integral protein of the inner nuclear membrane, influences transforming growth factor-β (TGF-β) signaling by directly interacting with R-Smads. Heterozygous loss of function mutations in the gene encoding MAN1 cause sclerosing bone dysplasias and an increased level of TGF-β signaling in cells. As a first step in elucidating the mechanism of TGF-β pathway regulation by MAN1, we characterized the structure of the MAN1 C-terminal region that binds Smad2. Using nuclear magnetic resonance spectroscopy, we observed that this region is comprised of a winged helix domain, a structurally heterogeneous linker, a U2AF homology motif (UHM) domain, and a disordered C-terminus. From nuclear magnetic resonance and small-angle X-ray scattering data, we calculated a family of models for this MAN1 region. Our data indicate that the linker plays the role of an intramolecular UHM ligand motif (ULM) interacting with the UHM domain. We mapped the Smad2 binding site onto the MAN1 structure by combining GST pull-down, fluorescence, and yeast two-hybrid approaches. The linker region, the UHM domain, and the C-terminus are necessary for Smad2 binding with a micromolar affinity. Moreover, the intramolecular interaction between the linker and the UHM domain is critical for Smad2 binding. On the basis of the structural heterogeneity and binding properties of the linker, we suggest that it can interact with other UHM domains, thus regulating the MAN1−Smad2 interaction.

The transforming growth factor-β (TGF-β)[1] family of cytokines comprises key regulators of metazoan embryonic development and adult tissue homeostasis. In the canonical pathway, ligands of both the TGF-β and bone morphogenetic protein members of this family bind to heteromeric serine/threonine kinase receptor complexes, which in turn phosphorylate R-Smad transcription factors at their C-terminal tails. This phosphorylation induces Smad1, -5, and -8 in the bone morphogenic protein pathway and Smad2 and -3 in the TGF-β pathway to accumulate in the nucleus and assemble transcriptional complexes that regulate hundreds of target genes (*1*, *2*).

Controlling the duration of TGF-β signaling is critical in development. This can be achieved by restricting the ability of Smad2 and Smad3 to accumulate in the nucleus. Smad2 constitutively shuttles between the cytoplasm and the nucleus, in particular through binding to nucleoporins (*3*). However, these nucleoporins compete with the cytoplasmic retention factor SARA and the nuclear Smad2 partner FAST-1 for binding to a hydrophobic corridor on the MH2 surface of Smad2 (*3*, *4*). SARA overexpression forces Smad2 into clustered locations in the cytoplasm, and FAST-1 overexpression forces Smad2 into the nucleus. On the basis of these insights, it has been proposed that localization of Smad2 is regulated by its binding to different nuclear and cytoplasmic proteins. TGF-β signaling does not influence Smad2 shuttling by changing its affinity for nucleoporins but by controlling its association with retention factors (*3*).

Smad proteins exhibit multiple phosphorylation sites that could participate in the regulation of their binding properties. In particular, phosphorylation of *Xenopus* Smad2 in its linker region restricts the time during which it remains in the nucleus (*5*). Multiple kinases have been identified that mediate the phosphorylation of Smad linkers (*6*−*10*). For example, in response to mitogens, the mitogen-activated protein kinase (MAPK) extracellular signal-regulated kinase catalyzes the phosphorylation of Smad linkers in vivo (*6*, *7*). This phosphorylation results in inhibition of Smad1 activity (*6*, *8*) and a decreased level of nuclear accumulation (*6*). Similarly, Smad2 linker phosphorylation causes MAPK-mediated attenuation of Smad2 nuclear accumulation and activity (*7*). Thus, the MAPK pathway influences the ability of Smads to transduce TGF-β signals.

[†]H.J.W. and W.W. were supported in part on this work by a gift from Ms. Lyn Pickel. E.K. and W.A. were supported by Association Française contre les Myopathies (Ph.D. Grant AFM13401 and ANR-06-MRAR-006-02). S.Z.-J., B.B., C.T.-L., S.G., J.-B.C., and B.G. were supported by Agence Nationale de la Recherche (ANR-06-MRAR-006-02), Commissariat à l'Energie Atomique, and Centre National de la Recherche Scientifique.

*To whom correspondence should be addressed: LBSR Bât. 144, CEA Saclay, 91190 Gif-sur-Yvette, France. Phone: +33 169083026. Fax: +33 169084712. E-mail: szinn@cea.fr.

[1]Abbreviations: TGF-β, transforming growth factor-β; MAPK, mitogen-activated protein kinase; UHM, U2AF homology motif; ULM, UHM ligand motif; NMR, nuclear magnetic resonance; SAXS, small angle X-ray scattering; ITC, isothermal titration calorimetry; CD, circular dichroïsm; HSQC, heteronuclear single-quantum correlation; NOESY, nuclear Overhauser effect spectroscopy; PDB, Protein Data Bank; SIM, Smad interacting motif; TEV, tobacco etch virus.

 Published on Web 08/17/2010

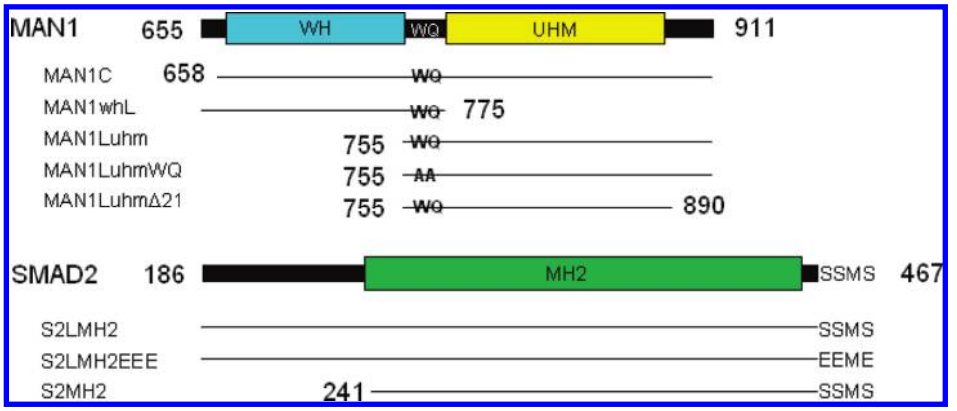

FIGURE 1: Architecture of the main protein fragments cited in this study: MAN1C, the C-terminal nucleoplasmic region of MAN1; MAN1whL consisting of the WH domain and the linker; MAN1Luhm containing the linker, the UHM domain, and the C-terminus; two mutants of MAN1Luhm, namely, MAN1LuhmWQ and MAN1LuhmΔ21; S2MH2 comprising the MH2 domain of Smad2; S2LMH2 consisting of the linker and the MH2 domain, and a mutant of S2LMH2, namely S2LMH2EEE, mutated at the three C-terminal serines of the SSXS motif.

MAN1 is an integral protein of the inner nuclear membrane with an N-terminal nucleoplasmic region, two transmembrane segments, and a C-terminal nucleoplasmic region (*11*). The C-terminal region of MAN1 directly interacts with Smad2 and Smad3 and represses TGF-β signaling (*12*−*16*). Heterozygous loss of function mutations in the gene encoding MAN1 cause the sclerosing bone dysplasias osteopoikilosis, Buschke-Ollendorff syndrome, and nonsporadic melorheostosis, which likely result from enhanced TGF-β signaling in cells (*12*). The MAN1−Smad interaction may lead to the cytoplasmic accumulation of Smad2 and Smad3, thus causing inhibition of TGF-β signaling (*14*). However, the mechanism by which the nucleoplasmic MAN1 C-terminal region would stimulate Smad2 and Smad3 nuclear export remains unclear. One hypothesis, as proposed for the MAN1 paralog NET25 (*17*), is that MAN1 functions as a scaffold protein recruiting Smad2, Smad3, and kinases to the inner nuclear membrane. Phosphorylation of Smad linker residues by kinases such as MAPKs would then lead to enhanced nuclear export. Alternatively, MAN1 may more simply act as a "nuclear envelope sink" for activated nuclear Smad proteins, sequestering them at the inner nuclear membrane and competitively inhibiting interactions with other nuclear factors involved in gene regulation.

To improve our understanding of the function of MAN1, we initiated structural studies of its C-terminal region that binds to Smad2 and Smad3. We previously showed that this region contains a winged helix (WH) DNA binding domain (*18*). In this study, we characterized the three-dimensional (3D) structure of the whole C-terminal region of MAN1 also comprising a structurally heterogeneous linker, a U2AF homology motif (UHM) domain, and a disordered C-terminus. Our data revealed that the linker plays the role of an intramolecular UHM ligand motif (ULM), and that its positioning onto the MAN1 UHM domain is critical for Smad2 binding.

EXPERIMENTAL PROCEDURES

Protein Preparation. We generated various protein fragments of human MAN1 that are shown in Figure 1: MAN1C corresponding to the C-terminal nucleoplasmic region of MAN1, MAN1whL consisting of the WH domain and the linker, MAN1Luhm containing the linker, the UHM domain, and the C-terminus, and several mutants of MAN1Luhm, such as MAN1LuhmWQ and MAN1LuhmΔ21. Similarly, several regions of Smad2 were produced: S2MH2 comprising the MH2 domain of Smad2, S2LMH2 consisting of the linker and the MH2 domain, and a mutant of S2LMH2, namely S2LMH2EEE, mutated at the three C-terminal serines of the SSXS motif.

MAN1C was overexpressed in *Escherichia coli* strain BL21 DE3 Star transformed with a plasmid that encodes a ZZ domain, a cleavage site for tobacco etch virus protease, and MAN1C. MAN1whL was overexpressed in *E. coli* strain DE3 pLys S transformed with a construct generated in pGEX-4T-1 (GE Healthcare) that encodes glutathione *S*-transferase (GST) and a thrombin site fused to MAN1whL. Both proteins were produced and purified as described previously (*18*).

MAN1 constructs MAN1Luhm (755−911) and MAN1LuhmΔ21 (755−890) and the mutants of MAN1Luhm (W765A/Q766A, P778A/P779A, and K864A/R870A) were overexpressed in *E. coli* strain BL21 DE3 Star transformed with a plasmid that encodes a six-histidine tag and the protein of interest. Fusion proteins were purified using a His trap column (GE Healthcare) and a gel filtration column (Superdex 200, Hi-Load, 120 mL, GE Healthcare). Their circular dichroism (CD) spectra were recorded on a Jobin Yvon CD6 instrument between 190 and 250 nm at 5 °C on 15 μM protein samples in a buffer containing 50 mM Tris-H_2SO_4 (pH 8), 75 mM NaF, and 5 mM β-mercaptoethanol. CD spectra obtained for wild-type MAN1Luhm and its variants analyzed in this study are similar, suggesting no overall structural changes due to the deletions or mutations (Figure S0 of the Supporting Information).

S2MH2, S2LMH2, and S2LMH2EEE were produced in *E. coli* strain BL21 DE3 Star transformed with a plasmid that encodes a six-histidine tag and the protein of interest. The fusion proteins were purified using a His trap column (GE Healthcare) and a gel filtration column (Superdex 200, Hi-Load, 120 mL, GE Healthcare). GST-S2MH2 and GST-S2LMH2 were overexpressed in *E. coli* strain BL21 DE3 Star transformed with a plasmid that encodes GST and the protein of interest. They were purified on a glutathione-Sepharose 4B column and by gel filtration (Superdex 200, Hi-Load, 120 mL, GE Healthcare).

Nuclear Magnetic Resonance (*NMR*) *Spectroscopy*. For NMR experiments, proteins were produced in M9 minimal medium containing $^{15}NH_4Cl$ (Cortecnet and Eurisotop) as the only nitrogen source or $^{15}NH_4Cl$ and [^{13}C]glucose (Cortecnet and Eurisotop) to yield ^{15}N-labeled or ^{15}N- and ^{13}C-labeled proteins, respectively. NMR samples were prepared in 50 mM Tris buffer (pH 6.7) containing 150 mM NaCl (90% H_2O/10% D_2O), 1 mM EDTA, 1 mM PMSF, and 1 mM NaN_3. 3-(Trimethylsilyl)[2,2,3,3-$^2H_4$]propionate was added as a chemical

shift reference. The protein concentration was between 0.1 and 0.3 mM.

NMR experiments were performed at 20 and 30 °C on Bruker (Saclay, France) 600 and 700 MHz spectrometers equipped with a triple-resonance cryoprobe. Spectra allowing chemical shift assignment and structural characterization were acquired as described previously (*19*). Two-dimensional (2D) $^{1}H-^{15}N$ heteronuclear single-quantum correlation (HSQC) and $^{1}H-^{15}N$ nuclear Overhauser effect spectroscopy (NOESY) were conducted with MAN1C, MAN1whL, and wild-type and mutated MAN1Luhm fragments. 3D HNCO, HNCA, CBCA(CO)NH, HNCACB, HBHA(CO)NH, HNHA, $^{1}H-^{15}N$ NOESY-HSQC, and $^{1}H-^{13}C$ NOESY-HSQC spectra of MAN1Luhm were recorded. All NMR spectra were analyzed and exploited with Sparky (T. D. Goddard and D. G. Kneller, University of California, San Francisco, CA). Chemical shift assignment was accelerated by the use of MARS (*20*).

Small Angle X-ray Scattering (*SAXS*) *Experiments*. SAXS samples were prepared in 50 mM Tris buffer (pH 8.0) containing 150 mM NaCl and 10 mM DTT. Their initial protein concentrations yielded between 12 and 14 mg/mL. Before data collection, they were injected on a KW402-5-4F gel filtration column (Shodex) to avoid the presence of aggregates. The SAXS measurements were taken just after protein elution at a concentration of ~1 mg/mL.

Synchrotron radiation X-ray scattering data were collected at 17 °C following standard procedures (*21*) on the SWING line at Synchrotron SOLEIL (Gif-sur-Yvette, France). The scattering due to buffer was subtracted by averaging the buffer measurements enclosing the actual protein measurement. All data manipulations were performed with PRIMUS (*22*). GNOM (*23*) provided access to the radius of gyration and the maximal distance of the molecule. It also provided a value of the SAXS intensity at the origin I_0, from which the mass was calculated using water as a reference, following

$$M = I_0/K \times N/(c\Delta\rho_m{}^2)$$

where M is expressed in kilodaltons, K is the ratio between theoretical (0.0164 cm^{-1}) and measured water scattering, N is Avogadro's number, c is the mass concentration in grams per liter, and $\Delta\rho_m$ is the protein electron density (taken to be 2×10^{10} cm/g, which corresponds to a protein partial specific volume of 0.74). CRYSOL (*24*) was used to generate SAXS curves from calculated models.

Molecular Modeling. MetaPrDos was ran on the server http://prdos.hgc.jp/cgi-bin/meta/top.cgi. The solution structure of the MAN1 WH domain (663−753) was previously determined by NMR in our laboratory [PDB entry 2CH0 (*18*)]. The MAN1 UHM domain (785−889) was modeled on the basis of the structure of the U2AF65 UHM domain (PDB entry 10OP; 28% identical to MAN1) and secondary structure constraints given by TALOS (*25*) using Modeler 9v7 (*26*).

To position the linker onto the UHM domain, the globular domain was centered at the origin and the linker (755−784) in a completely extended conformation was located 20 Å from the center of mass of the globular domain parallel to the *x*-axis. To generate 5000 initial structures of the fragment comprising the linker and the UHM domain, the globular domain was randomly rotated around its center of mass and the linker was rotated 10° by 10° around the *x*-axis. The peptide bond between residues 784 and 785 was deleted and replaced by a distance restraint. The C-terminus (890−910) was generated in a completely extended conformation. Ambiguous distance restraints were derived from chemical shift perturbation mapping analysis of mutant MAN1LuhmWQ. The 30 models of MAN1Luhm with the lowest distance restraint energies were selected, and the SAXS curves corresponding to these models were calculated with CRYSOL (*24*). Comparison with the experimental SAXS curves showed that 10 models are characterized by χ values of <3.5. The quality of these models is characterized by a Molprobity score of 2.6 ± 0.1 (*27*).

To model MAN1C, the WH (663−753) and UHM (785−889) domains were centered at the origin and the WH domain was translated 30 Å along the *x*-axis. The linker, in a completely extended conformation, was located 50 Å from the origin and parallel to the *x*-axis. To generate 5000 initial structures of MAN1C, the WH and UHM domains were randomly rotated around their respective center of mass and the linker was rotated 10° by 10° around the *x*-axis. The C-terminus (890−910) was generated in a completely extended conformation. The peptide bonds between the WH domain and the linker (residues 753 and 754) and between the linker and the UHM domain (residues 784 and 785) were deleted and replaced by two distance restraints. Ambiguous distance restraints were derived from chemical shift perturbation mapping analysis of mutant MAN1LuhmWQ. Additional distance restraints were derived from the comparison of the chemical shifts of MAN1whL and MAN1Luhm with those of MAN1C. The 30 models of MAN1C with the lowest distance restraint energies were selected, and SAXS curves corresponding to these models were calculated with CRYSOL (*24*). Comparison with experimental SAXS curves showed that 10 models are characterized by χ values of <1.7. The quality of these models is characterized by a Molprobity score of 2.5 ± 0.1 (*27*).

The same molecular modeling procedure was used for both calculations. It consists, in 14000 steps, of high-temperature molecular dynamics at 500 K under ambiguous distance restraints followed by 40000 steps of cooling to 300 K. The time step was 2 fs. Modeling calculations were performed using CHARMM version c32b1 and the CHARMM27 force field on a local cluster of 22 AMD OPTERON.

Isothermal Titration Calorimetry (*ITC*). ITC measurements were taken at 20 °C with a MicroCal VP-ITC instrument (GE Healthcare). MAN1Luhm and S2LMH2EEE were equilibrated in the same buffer containing 100 mM Na_2HPO_4 (pH 8), 150 mM NaCl, and 5 mM β-mercaptoethanol. S2LMH2EEE (65 μM) in the 1.3 mL calorimeter cell was titrated by 47 automatic injections of 6 μL of MAN1Luhm (855 μM). The first injection of 2 μL was ignored in the final data analysis. Integration of peaks corresponding to each injection, subtraction of the contribution of MAN1Luhm dilution (an endothermic reaction observed when MAN1Luhm is diluted into the buffer and contributing to 10−15% of the ITC signal), and correction for the baseline were performed using the Origin-based 7.0 software provided by the manufacturer. Curve fitting was done with a standard one-site model and gives the stoichiometry (n), equilibrium binding constant (K_a), and enthalpy of the complex formation (ΔH).

Fluorescence Measurements. Native and mutated MAN1-Luhm fragments were labeled with Alexa 488 (Alexa fluor 488 Labeling Kit, Molecular Probes). These fragments (~2 mg/mL) were incubated in 200 μL of 100 mM Na_2HPO_4 (pH 8), 150 mM NaCl, and 5 mM β-mercaptoethanol in the presence of a 2−8-fold molar excess of Alexa 488 for 3 h at room temperature. Free Alexa 488 was then separated from the protein on a gel filtration

column (Superdex 75, 24 mL, GE Healthcare) equilibrated with the same buffer. The labeling yield was evaluated on the basis of optical density measurements at 280 and 495 nm, related to the concentrations of protein and fluorophore, respectively. It was between 40 and 100%.

Measurements were taken at room temperature in black 384-well plates (Corning 3820). Filters were selected as a function of Alexa 488 optical characteristics (λ_{ex} = 495 nm, and λ_{em} = 530 nm). MAN1Luhm (50 nM) labeled with Alexa 488 was incubated with increasing concentrations of S2LMH2EEE in a final volume of 30 μL. The polarization data were fitted to the following equation:

$$P = P_0 + (P_{max}L)/(L + K_d)$$

where P_0 represents the polarization of MAN1Luhm in the absence of S2LMH2EEE, P_{max} the highest polarization of the binding curve, corresponding to the saturation of the interaction involving the labeled MAN1Luhm, L the concentration of S2LMH-2EEE, and K_d the dissociation constant.

GST Pull-Down Experiments. Approximately 1.5 mg of purified GST-S2LMH2 was incubated for 30 min at 4 °C with 100 μL of GS4B beads (GE Healthcare), equilibrated in 50 mM Tris (pH 8.0), 150 mM NaCl, and 1 mM DTT. After centrifugation for 5 s at 6000 rpm and an extensive washing of the beads, 30 μg of MAN1Luhm proteins (wild-type and variants) was added and incubated for 30 min at 4 °C. After similar washing steps, the proteins fixed on the beads were eluted with 50 mM Tris (pH 8.0), 150 mM NaCl, and 15 mM glutathione. The elutions were resolved by SDS−PAGE.

Yeast Two-Hybrid Experiments. DNAs encoding full-length Smad2, the MH2 domain of Smad2, and mutants of the region of residues 730−911 of MAN1 were generated by polymerase chain reactions and cloned into pACT2 and pGBKT7, respectively (*13*). pGBKT7 was used to express proteins fused at their N-terminus to the Gal4-DNA binding domain and a c-Myc epitope tag. pACT2 was used to express proteins fused at their N-terminus to the Gal4 transcription activation domain and an HA epitope tag. Mutants were obtained using the QuickChange kit (Stratagene). All plasmid constructs were confirmed by DNA sequencing, and the protein expression level was controlled by Western blotting. Yeast two-hybrid assays were performed according to the manufacturer's instructions (Clontech).

RESULTS

The MAN1 C-Terminal Nucleoplasmic Region Is Composed of Two Globular Domains. MetaPrDos predicts that MAN1 C-terminal nucleoplasmic region [MAN1C 658−911 (see Figure 1)] contains two globular domains (Figure 2A). Finding results consistent with this prediction, we have previously reported that residues 666−750 form a WH motif (*18*). Residues 775−880 were assigned to the UHM domain family on the basis of sequence alignments (*28*). To further characterize the 3D structure of MAN1C, we have produced ^{15}N-labeled fragments of MAN1C, MAN1whL (for WH and Linker), and MAN1Luhm (for Linker and UHM) containing the predicted UHM domain.

Analysis of the NMR ^{1}H−^{15}N HSQC spectra of MAN1C, MAN1whL, and MAN1Luhm, which represent fingerprints of the fragment 3D structures, shows that the spectrum of the whole C-terminal region, MAN1C, is similar to the sum of the spectra of the fragments MAN1whL and MAN1Luhm (Figure 2B). Thus, MAN1whL and MAN1Luhm adopt similar folds when they are either free in solution or within the whole C-terminal region.

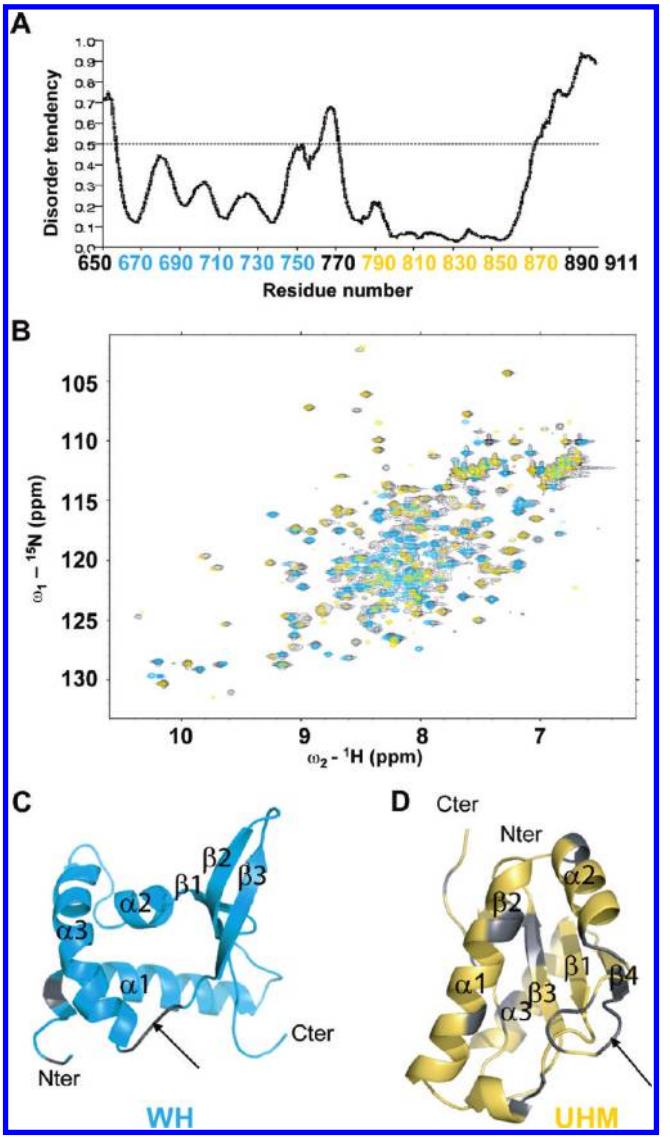

FIGURE 2: MAN1C exhibits two globular domains. (A) Disorder tendency of the MAN1 C-terminal nucleoplasmic region of residues 655−911 (MAN1C) as a function of its sequence. This plot was calculated by MetaPrDos (http://prdos.hgc.jp/cgi-bin/meta/top.cgi). (B) Superposition of the ^{1}H−^{15}N HSQC spectra of MAN1C (black), MAN1wh (cyan), and MAN1Luhm (yellow). These spectra were recorded at 700 MHz, pH 6.7, and 30 °C. (C) Ribbon representation of the 3D structure of the WH domain (PDB entry 1CH0). (D) Ribbon representation of the model calculated for the UHM domain of MAN1 on the basis of the secondary structure deduced from NMR chemical shift analysis and the 3D structure of the UHM domain of U2AF65 (PDB entry 1O0P). In panels C and D, the gray color indicates unassigned residues and the arrow points to the main unassigned region.

As we already determined the 3D structure of the WH domain [Figure 2C (*18*)], we further characterized the 3D structure of MAN1Luhm in solution. We produced ^{15}N- and ^{13}C-labeled MAN1Luhm and assigned the frequencies of the NMR backbone nuclei of 70% of the non-proline residues of this fragment (Figure S1 of the Supporting Information). The unassigned peaks observed on the MAN1Luhm ^{1}H−^{15}N HSQC spectrum are not related to peaks of the 3D NMR spectra recorded for assignment, thus highlighting the dynamics of the corresponding residues. From this assignment and using TALOS (*25*), we deduced the localization of the secondary structure elements of MAN1Luhm. Five β-strands of more than three residues were identified, four of

them corresponding to the predicted β-strands of the UHM fold (*28*): C785−I788 (β1), H822−D826 (β2), C833−C838 (β3), and V863−R867 (β4). The fifth β-strand is located at the C-terminus of the UHM domain and comprises N882−P887. Similarly, three α-helices were identified, two of them corresponding to the predicted helices of the UHM fold (*28*): H803−E812 (α1) and E842−A850 (α2). The third helix is located at the C-terminus of the UHM fold and comprises L868−880. Analysis of the $^{1}H-^{1}H$ nuclear Overhauser effect (NOE) patterns, reflecting $^{1}H-^{1}H$ proximities, confirmed the pairing of predicted β-strands β1−β4. The additional helix α3 could be positioned onto the β-sheet; such a positioning has already been found in the 3D structure of the UHM domain of U2AF65, which is 28% identical to the UHM of MAN1 [PDB entry 1O0P (*29*)]. No $^{1}H-^{1}H$ NOE could be observed that allowed pairing of the additional β-strand β5; this strand is close to β-strand β2 and might be transitorily paired with β2, as found in other domains with a similar fold (*30*, *31*). Incomplete assignment of NMR side chain nuclei frequencies precluded the calculation of a well-resolved solution structure for the UHM domain of MAN1. At this stage, we modeled the structure of the domain of residues 785−889 of MAN1 on the basis of secondary structure constraints corresponding to TALOS results and the 3D structure of the U2AF65 UHM domain (Figure 2D).

The MAN1 C-Terminal Fragment Comprises Several Regions in Conformational Exchange. After several days at 30 °C, MAN1C is degraded. Analysis of the degradation products by N-terminal sequencing and SDS−PAGE shows that the main fragment has residue 658 at its N-terminus and has a size similar to that of MAN1whL. Thus, MAN1whL represents the most stable fragment of the C-terminal region of MAN1. In this fragment, the NMR signals of 20% of the residues are not visible and cannot be assigned. These residues are located at the N-terminus (664, 665, and 668), in the α3−β1 loop (726−728), and in the linker region (753−756 and 762−775) (Figure 2C). In the less stable fragment MAN1Luhm, the NMR signals of 30% of the residues are not visible. Most of these residues are located in the linker region (764−774), in the α2−β4 loop (859−863), and in the C-terminus (890−901) (Figure 2D). Finally, we expected the appearance of additional NMR signals in the $^{1}H-^{15}N$ spectrum of MAN1C compared to the spectra of MAN1whL and MAN1Luhm, because of the potential stabilization of several structural elements in the larger fragment. However, no new peak was observed, and several peaks present in the NMR spectra of the smaller fragments were not observed in the spectrum of MAN1C. These peaks correspond to residues of the linker region (755−763), the N-terminus of the UHM domain (775−780), the β1−α1 loop (792, 795, and 804), and the β2−β3 loop (826, 830, and 831). In total, the NMR signals of 30% of the C-terminal region were not observed, showing that MAN1C exhibits large regions in conformational exchange.

The Linker Region between the Two Globular Domains Is Anchored at the Surface of the UHM Domain. In the linker region of MAN1C, residues 750−754 are poorly structured, as deduced from the analysis of NMR $^{1}H-^{15}N$ NOE experiments (Figure S2 of the Supporting Information), and residues 755−780 are not observed because of conformational exchange. In MAN1Luhm, most of the NMR signals corresponding to the linker region are not observed. On these bases, we fashioned the hypothesis that this linker region exhibits conformational exchange by interacting transitorily with the rest of the MAN1Luhm fragment. UHM domains are described as interacting with peptides centered on a tryptophan residue (*28*). Interestingly, alignment of 22 sequences of MAN1 homologues from insects to mammals revealed that, in the linker region, next to the UHM domain, residues W765, Q766, and A769 are strictly conserved (Figure S3A of the Supporting Information).

Because of their conservation across species, we determined if W765 and Q766 could contribute to a critical transient interaction between the linker and the UHM domain. To do so, we expressed a MAN1Luhm fragment in which these residues are replaced with alanines (MAN1LuhmWQ double mutant). We produced a ^{15}N-labeled sample of this mutant and compared its $^{1}H-^{15}N$ HSQC spectrum to that of MAN1Luhm (Figure S4 of the Supporting Information). Clearly, the spectra are superimposable, which shows that the overall folding of the two fragments is the same. However, 22 HSQC peaks are largely shifted so that they could not be easily assigned in the mutant spectrum, and 16 HSQC peaks are significantly displaced because of the mutations. The chemical environment of the residues corresponding to these peaks is affected by the mutations. We assigned these peaks to several backbone N−H bonds of the linker region, backbone N−H bonds of I788 (strand β1), R789, M791, F792, V795, M796, I798, G799, D800, and Q801 (β1−α1 loop), L804, A805, E808, I810, E812, S815, and D816 (helix α1), D818 and I820 (α1−β2 loop), I823 and D826 (strand β2), K827, N828, R830, and G832 (β2−β3 loop), V834 (strand β3), G853 and G858 (α2−β4 loop), Y871 and L879 (helix α3), and T883 and S888 (strand β5), and indole N−H bonds of W802 (β1−α1 loop) and W855 (α2−β4 loop) in the UHM domain. The mapped residues form a continuous surface located on one side of the protein and centered on a cavity defined by the β-sheet, helix α1, the β1−α1 loop, and the β2−β3 loop (Figure 3A). The largest dimension of the surface is approximately 55 Å, which suggests that residues W765 and Q766 anchor ∼20 residues of the linker to the UHM domain.

We performed SAXS experiments on samples of MAN1Luhm, MAN1LuhmWQ, and MAN1C. SAXS provides accurate information about folding and conformation in solution for both rigid and flexible macromolecules. The indirect Fourier transform of the SAXS intensity gives access to the real space electron pair distance distribution of molecules in solution. However, SAXS data might be distorted by the scattering because of small amounts of aggregates. To eliminate this problem, each sample was eluted from a gel filtration column at ∼1 mg/mL just prior to data collection. Moreover, the monodispersity of the samples was verified by a Guinier analysis (Figure S5 of the Supporting Information). Figure 3B shows that the SAXS curves of MAN1Luhm and MAN1LuhmWQ are similar, reflecting their common global fold already illustrated by NMR. Calculation of the electron pair distance distributions using GNOM gave access to the radius of gyration (R_g), the maximal distance, and the I_0 value that cannot be measured experimentally and is proportional to the mass. From these values, a Kratky representation was displayed (Figure 3B) that highlighted structural differences between MAN1Luhm and MAN1LuhmWQ: the maximum of the curve corresponding to $(SR_g)^2(I/I_0)$ is shifted toward higher SR_g values, reflecting the more elongated and less structured state of the mutant (*32*). Moreover, MAN1Luhm is a monomer of 20.1−20.6 kDa (compared to a theoretical mass of 19.5 kDa), with a radius of gyration between 20 and 21 Å and a maximal distance between 68 and 72 Å. MAN1C is a monomer of 30 kDa (compared to a theoretical mass of 29.2 kDa) with larger dimensions; its radius of gyration is 25.3 Å and its maximal distance 90 Å. Interestingly, MAN1LuhmWQ is a monomer of 20−21 kDa

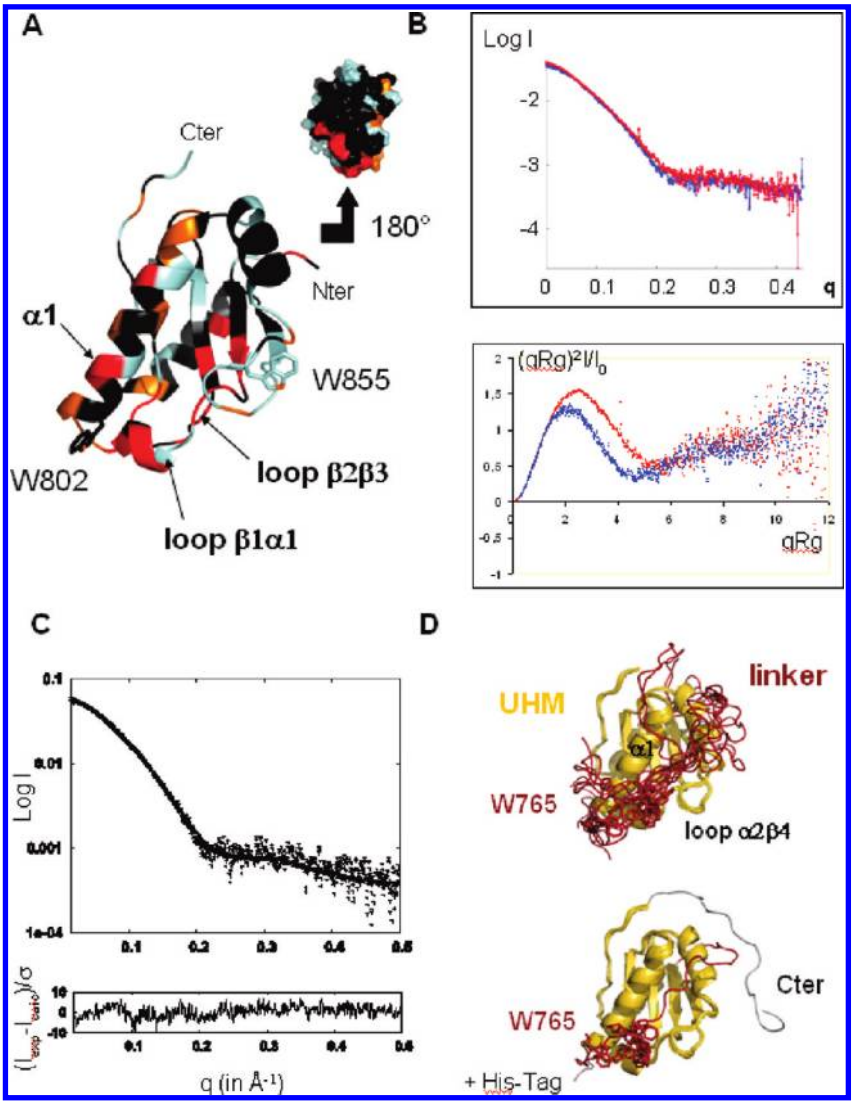

FIGURE 3: Modeling the MAN1Luhm fragment. (A) NMR chemical shift mapping of the UHM region interacting with the linker. On the black ribbon representation of the MAN1 UHM model, residues whose backbone $^{1}H-^{15}N$ HSQC peak is largely shifted and could not be assigned due to the W765A and Q766A mutations are colored red, residues whose backbone $^{1}H-^{15}N$ HSQC peak is shifted by more than 1.5 times the root-mean-square deviation of the chemical shift difference distribution but could still be assigned are colored orange, and residues whose backbone $^{1}H-^{15}N$ HSQC peak could not be assigned in the MAN1Luhm and MAN1LuhmWQ spectra are colored cyan. (B) Comparison of the SAXS data recorded on MAN1Luhm (blue) and MAN1LuhmWQ (red). The top panel displays the superposition of the logarithm of the SAXS intensities as a function of the amplitude of the diffusion vector [$q = 4\pi \sin(\theta)/\lambda$]. The bottom panel corresponds to the superposition of the Kratky plots: the diffusion vector amplitude multiplied by the gyration radius is plotted on the abscissa, and the square of the abscissa multiplied by the SAXS intensity divided by the intensity at 0 is plotted on the ordinate. This representation reveals the type of structure: compact, partially folded, or unfolded. (C) Representation of the fit of the models to the SAXS data. As an example, in the top panel, the SAXS curve averaged on the 10 models is superimposed onto the experimental curve, and in the bottom panel, the difference between the calculated and experimental curves divided by the experimental error is plotted as a function of the diffusion vector amplitude. (D) Models of MAN1Luhm consistent with the NMR and SAXS data. Ribbon representations of 10 MAN1Luhm models are superimposed, with the linker and residue W765 colored red, the UHM domain yellow, and the C-terminus gray. In the top panel, the linker and UHM domain backbones (with the side chain of W765 shown as sticks) of the 10 models are displayed. In the bottom panel, 10 UHM backbones and W765 side chains but only one representative model of the linker and C-terminus are displayed.

(compared to a theoretical mass of 19.3 kDa) whose radius of gyration is intermediate between those of MAN1uhm and MAN1C (23.3−24 Å) and whose maximal distance is close to that of MAN1C. Hence, SAXS data indicate that mutations W765A and Q766A provoked a local unfolding in MAN1Luhm. This result validates our previous NMR-based analysis that strongly suggests a role for W765 and Q766 in anchoring the linker onto the UHM domain.

We modeled MAN1Luhm on the basis of the NMR and SAXS data recorded on MAN1Luhm and MAN1LuhmWQ. First, we calculated 5000 structures of MAN1Luhm consistent with the chemical shift mapping results obtained from the NMR analysis of the MAN1LuhmWQ mutant. Therefore, we applied 28 ambiguous distance constraints between each of the residues whose NMR signals are affected by the mutations (see Figure 3A) and the residues from the fragment of residues I757−R774 of the linker. Thirty models that are consistent with the NMR-derived constraints were selected: at least 22 constraints (> 80%) correspond to distances of < 6 Å. Second, back-calculated SAXS curves were obtained for each of these models using CRYSOL (*24*). They were superimposed onto the experimental SAXS curve recorded on the MAN1Luhm sample, and the χ parameters reflecting the fit of the calculated curves to the experimental data were calculated. Although the χ is dependent in part of the experimental

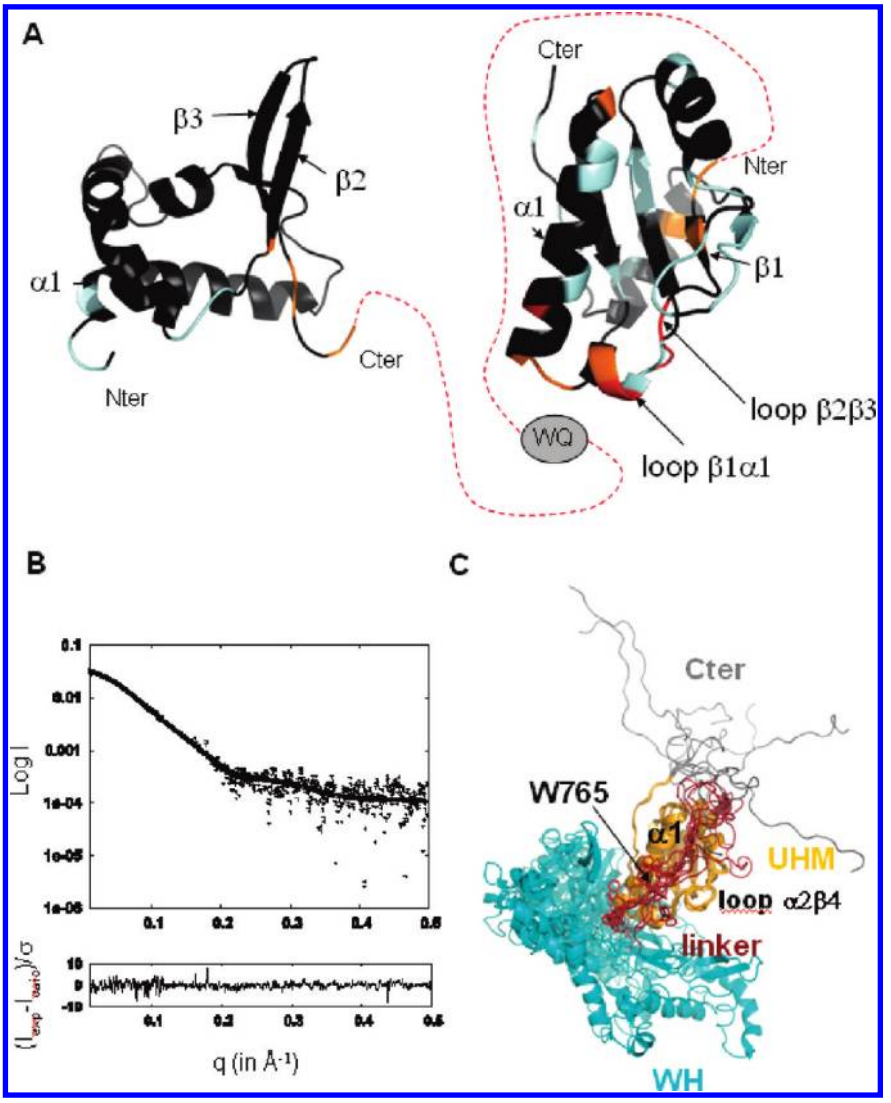

FIGURE 4: Modeling the whole C-terminal nucleoplasmic region of MAN1. (A) NMR chemical shift mappings of the WH region whose chemical environment is affected by the presence of the UHM domain and the C-terminus (right) and of the UHM region whose chemical environment is affected by the presence of the WH domain (left). On the black ribbon representation of the MAN1 WH and UHM domains, residues for which the backbone $^{1}H-^{15}N$ HSQC peak is largely shifted and could not be assigned in the MAN1C spectrum are colored red, residues for which the backbone $^{1}H-^{15}N$ HSQC peak is shifted by more than 1.5 times the root-mean-square deviation of the chemical shift difference distribution but could still be assigned are colored orange, and residues for which the backbone $^{1}H-^{15}N$ HSQC peak could not be assigned in the MAN1whL and MAN1Luhm spectra are colored cyan. (B) Representation of the fit of the models to the SAXS data. As an example, in the top panel, the SAXS curve averaged on the 10 models is superimposed onto the experimental curve, and in the bottom panel, the difference between the calculated and experimental curves divided by the experimental error is plotted as a function of the diffusion vector amplitude. (C) Models of MAN1C consistent with the NMR and SAXS data. Ribbon representations of 10 MAN1Luhm models are superimposed, with the WH domain colored cyan, the linker and its W765 red, UHM domain yellow, and the C-terminus gray.

noise, a lower χ value generally corresponds to a better fit of the model(s) to the experimental data. Ten models were selected that satisfied the distance restraints and were consistent with the SAXS data ($\chi < 3.5$). Subtraction of the average SAXS curve calculated on these 10 models to the experimental curve shows that the differences are randomly distributed along zero for all q values, as predicted for random errors (Figure 3C). Superimposition of the 10 models leads to a first view of the MAN1Luhm structures consistent with our experimental data (Figure 3D). In particular, the position of W765 relative to the UHM domain is displayed for the 10 models. Clearly, despite the structural heterogeneity of the UHM domain, the W765 residue lies in or next to the cavity between helix α1 and the α2−β4 loop.

Modeling the Whole C-Terminal Region of MAN1. Having modeled the MAN1Luhm fragment, we attempted to position this fragment relative to the WH domain in the whole MAN1C region. Comparison of the $^{1}H-^{15}N$ HSQC spectra of MAN1whL alone and within the whole C-terminal region shows that only a small number of residues have their backbone chemical shifts influenced by the presence of the UHM domain and the C-terminus. These residues are essentially found at the N-terminus of β2 (R731), at the C-terminus of β3 (I748 and Q749), and at the C-terminus of the fragment (A752, S753, I757, L758, and V759) (Figure 4A). Comparison of the NMR $^{1}H-^{15}N$ HSQC spectra of MAN1Luhm and MAN1C shows that 16 peaks are present in the spectrum of MAN1Luhm but absent in the spectrum of the whole C-terminal region. These peaks correspond to backbone N−H bonds of residues D755−I760, S762, and K763 in the linker region between the WH and UHM domains and residues R775, S777, and N780 (N-terminus), F792, V795, and L804 (β1−α1 loop), and R830 and E831 (β2−β3 loop) in the UHM domain. Moreover, the NMR nitrogen and amide proton chemical shifts of 11 residues are

significantly affected by the presence of the WH domain. These residues are L782 (N-terminus), C785, L786, and I788 (β1), M796 and G799 (β1−α1 loop), D800 (α1), N817 (α1−β2 loop), N828 and S829 (β2−β3 loop), and Y871 (α3) in the UHM domain. All these residues are found either in the linker between the two domains or within a surface centered on the β1−α1 and β2−β3 loops (Figure 4A).

We calculated the MAN1C structures consistent with our NMR and SAXS data. From our chemical shift mapping results, we deduced five ambiguous distance constraints between each mapped WH domain residue and either the linker residues or the UHM domain residues involved in WH domain binding. We also introduced 10 ambiguous distance constraints between each mapped UHM domain residue and either the linker residues or the WH residues involved in UHM domain binding. Using these 15 constraints together with the 28 constraints previously obtained from the analysis of the mutant MAN1LuhmWQ, we calculated 5,000 models of MAN1C. Thirty models that nicely fit to the NMR-derived constraints were selected: at least 34 constraints (> 80%) correspond to distances of < 6 Å. Back-calculated SAXS curves were obtained for each of these MAN1C models using CRYSOL (*24*). Finally, we filtered 10 models that are consistent with both the NMR and SAXS data (χ < 1.7). The error function measuring the difference between the experimental and calculated SAXS values is randomly distributed along zero for all q values, as predicted for random errors (Figure 4B). Superimposition of the 10 models leads to a description of the MAN1C conformational heterogeneity consistent with our experimental data (Figure 4C). Clearly, in these models, the W765 residue is again in or close to the cavity defined by helix α1 and the α2−β4 loop.

The Flexible Linker and C-Terminal Fragments of MAN1 Are Involved in the MAN1−Smad2 Interaction. To identify the MAN1 surface involved in Smad2 recognition, we tested using GST pull-down and yeast two-hybrid experiments the binding of various MAN1 wild-type and mutated fragments (Figure 5A) to either S2MH2, S2LMH2, or full-length Smad2 (Figure 5B,C). Yeast two-hybrid experiments previously demonstrated that fragments of residues 680−910 and 730−910 of MAN1 bind to the MH2 domains of Smad2 and Smad3, suggesting that the WH domain of residues 663−753 is not required for these interactions (*13*). We demonstrated by GST pull-down that MAN1Luhm binds to S2LMH2 corresponding to the linker and MH2 domains of Smad2 (Figure 5B). We also confirmed by two-hybrid analysis that MAN1 730−910, comprising MAN1Luhm, recognizes both the MH2 domain S2MH2 and full-length Smad2 (Figure 5C).

Then we tested the role of the UHM domain in the MAN1−/Smad2 interaction. We measured the binding to Smad2 of several mutants of MAN1Luhm, in which conserved and solvent-exposed residues are converted into alanines (Figure S3B of the Supporting Information). To prevent interference with the linker positioning, we chose to mutate only residues located on the side of the UHM domain opposite the linker binding cavity. Yeast two-hybrid experiments showed that mutations R789A, K864A, R870A, and R874A do not affect the binding of MAN1 730−911 to either full-length Smad2 or the Smad2 MH2 domain (Figure 5C). These results suggest that the surface of the UHM domain opposite the linker binding cavity is not involved in Smad2 binding.

The linker region of MAN1 is necessary for Smad2 and Smad3 binding in yeast two-hybrid experiments (*13*). We therefore checked the role of the conserved W765 and Q766 residues in the MAN1−Smad2 interaction. GST pull-down experiments revealed no binding of MAN1LuhmWQ to GST-S2LMH2 (Figure 5B). Yeast two-hybrid experiments further showed no binding of MAN1 fragment 730−911 mutated at positions 765 and 766 to either full-length Smad2 or the Smad2 MH2 domain (Figure 5C). We thus confirmed that the MAN1 linker plays an essential role in the MAN1−Smad2 interaction.

P778 and P779, next to the linker region and at the N-terminus of the UHM region, were described as belonging to a putative MAN1 Smad interaction motif (SIM) (*18*, *33*). GST pull-down experiments showed no impact of these mutations on the binding of MAN1Luhm to GST-S2LMH2 (Figure 5B). Yeast two-hybrid experiments revealed that mutations P778A and P779A affect the binding of MAN1 to Smad2 MH2, but not to full-length Smad2 (Figure 5C). Thus, residues P778 and P779 influence the binding of MAN1 to the MH2 domain but not to full-length Smad2.

We finally examined the role of the MAN1 C-terminus by testing the interaction between MAN1LuhmΔ21 and Smad2 using two-hybrid experiments. We showed that the MAN1 fragment 730−890 poorly recognizes the MH2 domain of Smad2 but binds to full-length Smad2 (Figure 5C). Thus, we concluded again that the C-terminus is involved in the MH2 domain recognition but is not critical for full-length Smad2 binding.

Our results show that the linker, the N-terminus of the UHM domain, and the C-terminus are involved in the recognition of the MH2 domain of Smad2 and underline the critical role of the linker region in full-length Smad2 binding.

MAN1Luhm Recognizes both the Wild-Type Smad2MH2 Domain and a Trimeric Mimic of Its Phosphorylated State through Its Linker and C-Terminal Regions. Difficulty in obtaining large quantities of pure S2MH2 and S2LMH2 precluded measurement of a precise K_d value corresponding to the MAN1Luhm−Smad2 interaction. However, fluorescence measurements indicated that the affinity of labeled MAN1Luhm for the mainly monomeric GST-S2MH2 is in the micromolar range (data not shown). It is known that following TGF-β signaling, Smad2 and Smad3 are phosphorylated on their C-terminal serines and their MH2 domains oligomerize. A mimic of the phosphorylated Smad3MH2 domain was designed that exhibits three serine to aspartic acid mutations and trimerizes (*34*). We produced a similar Smad2 fragment that we called S2LMH2EEE (see Figure 1). We confirmed that S2LMH2EEE is trimeric via gel filtration experiments (Figure S6 of the Supporting Information). Fluorescence measurements revealed that the affinity of labeled MAN1Luhm for S2LMH2EEE yields 9.8 ± 3.0 μM (Figure 6A). ITC measurements confirmed that the K_d of this interaction is in the micromolar range (2.7 ± 0.1 μM) and further showed that the ΔH value is −8.9 kcal/mol and the $T\Delta S$ value is −1.4 kcal/mol (Figure 6B). The stoichiometry of the reaction is 1, which indicates that the Smad2 trimer recognizes three MAN1-Luhm molecules. This interaction is driven by a large enthalpic component largely responsible for the observed micromolar affinity.

Fluorescence measurements revealed the impact of various MAN1Luhm mutations on the recognition of trimeric S2LMH-2EEE (Figure 6C). They indicated that mutations P778A and P779A and mutations K864A and R870A in the UHM domain have no impact on the binding of MAN1Luhm to S2LMH-2EEE. MAN1LuhmWQ binds with a 148.8 ± 50.5 μM affinity to S2LMH2EEE (at least 7 times weaker than MAN1Luhm). MAN1LuhmΔ21 binds with a 25.2 ± 4.0 μM affinity to S2LMH2EEE (~2.5 times weaker than MAN1Luhm). These

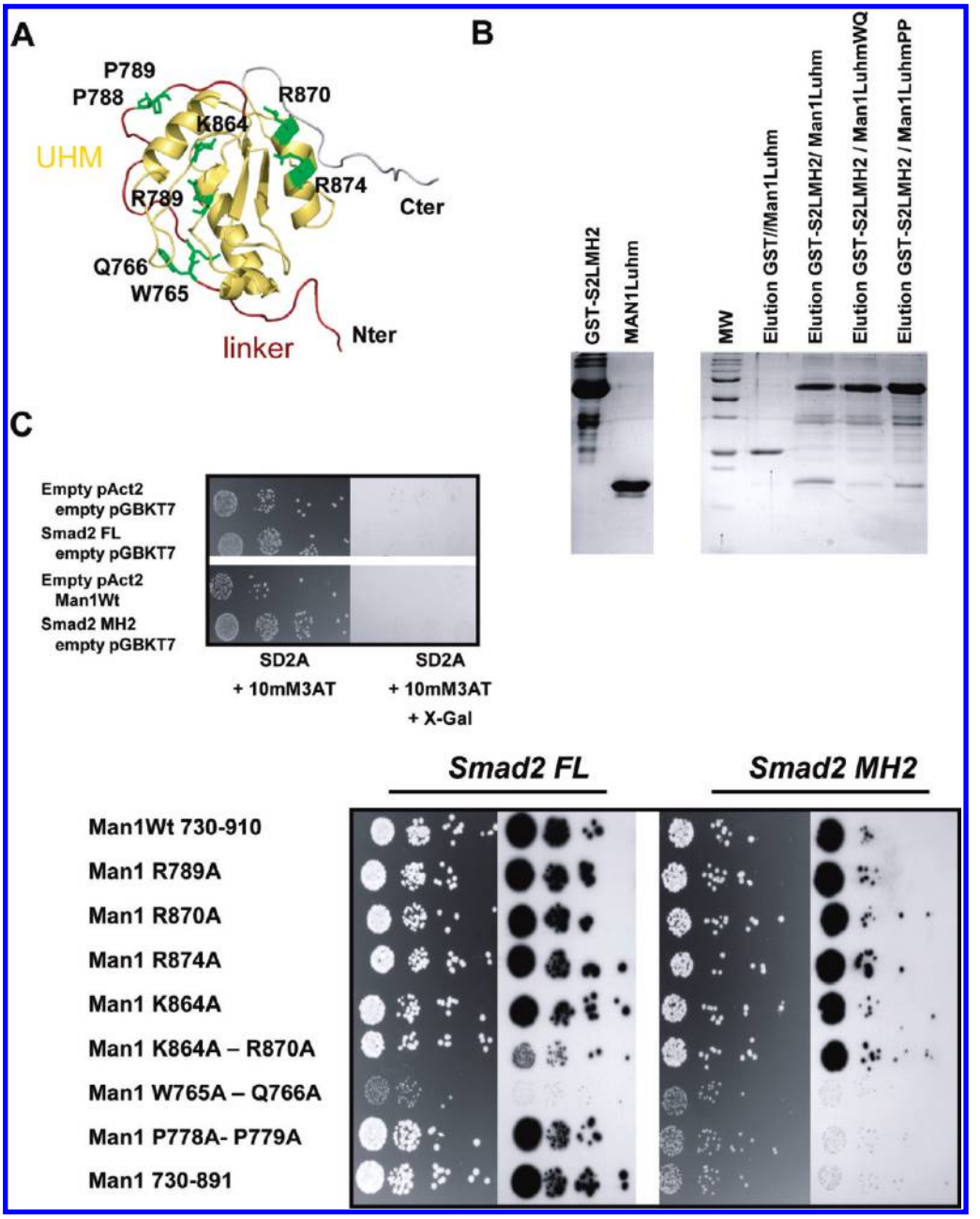

FIGURE 5: Mapping the MAN1Luhm region interacting with wild-type Smad2 fragments. (A) Localization of the mutated residues (green) on a ribbon representation of MAN1Luhm (linker colored red, UHM yellow, and C-terminus gray). (B) GST pull-down experiments reporting the binding of MAN1Luhm and its linker variants to GST-S2LMH2. Input proteins are shown in lanes 1−3: GST-S2LMLH2 in lane 1, Man1Luhm in lane 2, and Man1LuhmΔ21 in lane 3. After incubation of the GS4B beads with either GST (lane 5) or GST-S2LMH2 (lanes 6−8) and different MAN1 fragments, the complexes are eluted using a 10 mM glutathione buffer. The W765A and Q766A mutations clearly hinder Smad2 binding, as judged by the very weak band of Man1Luhm retained by the corresponding variant. (C) Two-hybrid experiments reporting the binding of MAN1 730−910 and its variants to either the Smad2 MH2 domain or full-length Smad2. In the top panel, empty vectors were used to perform control experiments. In the bottom panel, the interactions between native and mutated MAN1 730−911 fragments and either full-length Smad2 (left) or the Smad2 MH2 domain (right) are visualized. In both panels, the cells containing the different plasmids were spotted using 10-fold serial dilutions onto SD2A plates containing 10 mM 3-aminotriazole and incubated at 30 °C for 3 days, to test the expression of the first reporter gene *HIS3* (dark background columns). The cells were also streaked on SD2A plates to test for β-galactosidase expression in an overlay plate assay (white background columns). Physical interaction between Smad2 and Man1 (730−911) activated both expressions.

results show that the linker and C-terminus are involved in the recognition of the trimeric MH2 domain of Smad2 and thus demonstrate that MAN1Luhm is the smallest MAN1 fragment that can bind with a micromolar affinity to this trimeric Smad2.

Taken together, our data revealed that the surface of MAN1 involved in either wild-type or trimeric Smad2 recognition comprises the structurally heterogeneous regions called the linker and C-terminus. Consistently, the melting temperature of the MAN1Luhm−S2LMH2EEE complex is 48 °C, which is lower than the melting temperatures of MAN1Luhm (55 °C) and S2LMH-2EEE (50 °C) (Figure S7 of the Supporting Information). Thus, our results suggest that conformational rearrangements are involved in the interaction between MAN1 and Smad2.

DISCUSSION

Our NMR and SAXS analysis of the C-terminal nucleoplasmic region of MAN1 revealed that it is composed of a WH domain, a structurally heterogeneous linker region, a UHM domain, and

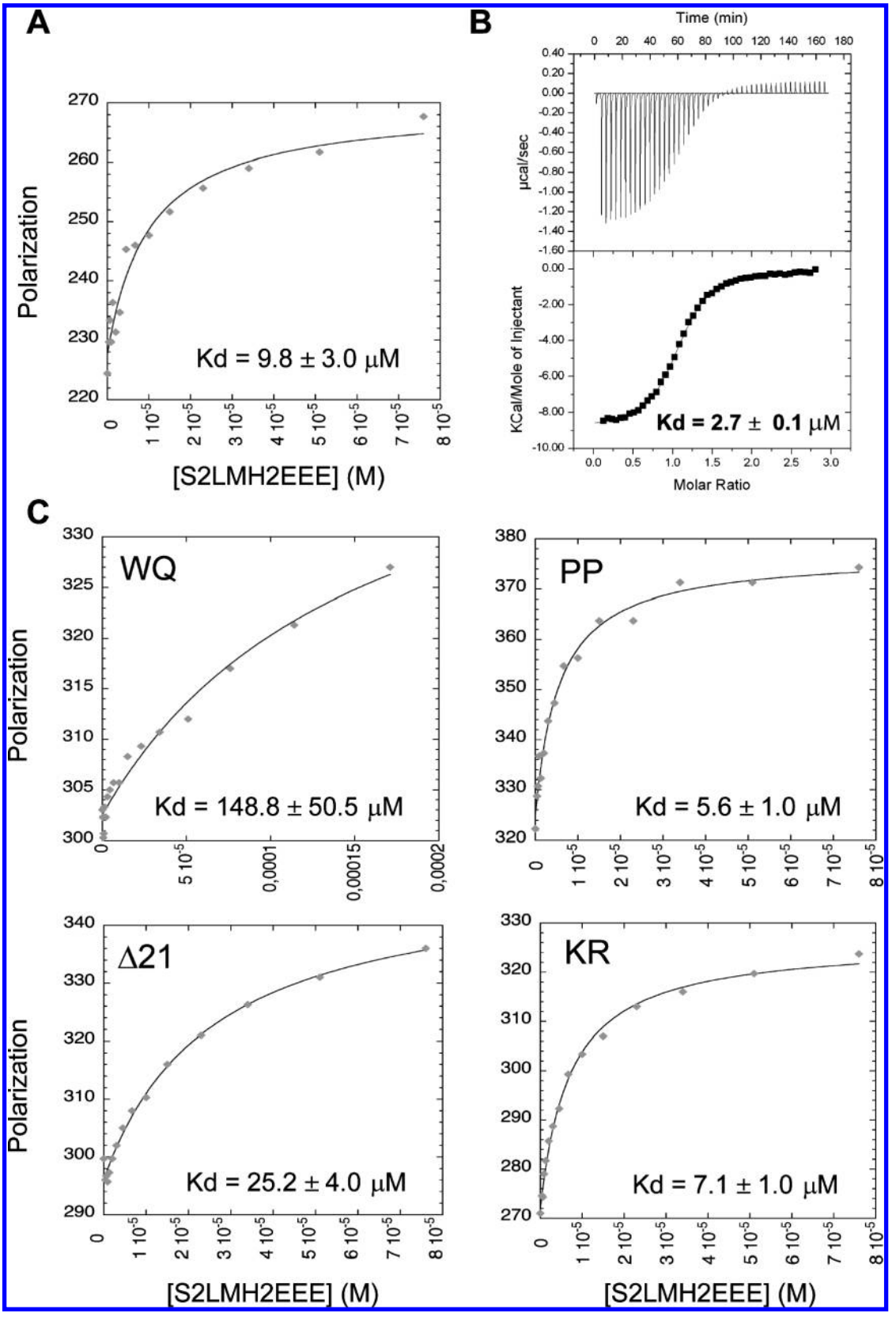

FIGURE 6: Mapping the MAN1Luhm region interacting with the trimeric S2LMH2EEE fragment. (A) Fluorescence polarization of the labeled MAN1Luhm fragment as a function of the concentration in S2LMH2EEE. Fitting the curve yielded a K_d value of $9.8 \pm 3.0\,\mu$M. (B) Binding curve obtained by ITC when MAN1Luhm is added onto S2LMH2EEE. Fitting this curve yielded a K_d value of $2.7 \pm 0.1\,\mu$M, a stoichiometry of 1, a ΔH value of -8.9 kcal/mol, and a $T\Delta S$ value of -1.4 kcal/mol. (C) Effect of different mutations in MAN1Luhm on its binding to S2LMH2EEE, as followed by fluorescence. Analysis of the binding curves was conducted for MAN1LuhmWQ ($K_d = 148.8 \pm 50.5\,\mu$M), MAN1Luhm(P778A/P779A) ($K_d = 5.6 \pm 1.0\,\mu$M), MAN1LuhmΔ21 ($K_d = 25.2 \pm 4.0\,\mu$M), and MAN1Luhm(K864A/R870A) ($K_d = 7.1 \pm 1.0\,\mu$M). The x-axis range is different in the case of the binding to MAN1LuhmWQ because of the lower affinity of this interaction.

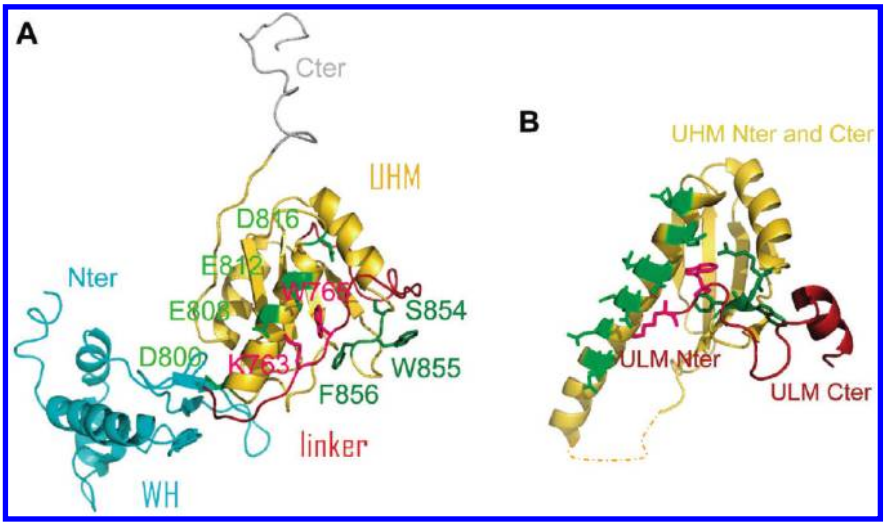

FIGURE 7: MAN1 linker−UHM interface (A) that is reminiscent of a classical ULM−UHM interface (B). Ribbon representations of the UHM domains of MAN1C and U2AF65 are colored yellow, the MAN1 linker and the U2AF35 ULM peptide red, and the MAN1 WH domain is colored cyan and the MAN1 C-terminus gray. The negatively charged residues of helix α1 are depicted as light green sticks; the three residues of the α2−β4 loop motif are depicted as dark green sticks, and the positively charged and tryptophan residues of the MAN1 linker and the U2AF35 ULM are depicted as magenta sticks.

a poorly folded C-terminus. Because of the small number of WH NMR signals affected by the presence of the UHM domain and the C-terminus, the interface between the WH and UHM domains is limited in our MAN1 models. Reciprocally, several chemical shifts of the UHM domain are affected by the presence of the WH domain, but most of them are also affected by the mutations W765A and Q766A. The presence of the WH domain in MAN1C probably influences the linker structure, provoking local rearrangements at the linker−UHM interface. As the WH domain is dispensable for Smad2 recognition, we further focused on the structural characterization of the MAN1Luhm−Smad2 interaction.

Within the C-Terminal Region of MAN1, an Intramolecular Interaction Involves the ULM-like Motif of the Linker and the UHM Domain. The UHM domain and the C-terminus are rapidly degraded in the NMR tube at 30 °C or in trypsin proteolysis assays (data not shown). Several UHM domains have been reported to be poorly stable in the absence of their ULM partner (*28*). X-ray structures have been obtained of UHM domains bound to their partners (*35*). These studies have shown that UHM domains have a cavity located between a negatively charged helix α1 and an R-X-F motif in the α2−β4 loop. This cavity is able to recognize a peptide containing several positively charged residues followed by a tryptophan residue. In the case of MAN1, helix α1 is negatively charged and an S-W-F motif is found in the α2−β4 loop. We tested the binding of MAN1-Luhm to a typical ULM peptide [KKKKVRKYWDVPP (*35*)] but did not observe any modification of the NMR signals of MAN1Luhm upon peptide addition (data not shown). However, we detected a highly conserved K763-X-W765-Q766-$(X)_2$-A769-F770 motif, also containing a positively charged residue and a tryptophan, in the linker region of MAN1. When W765 and Q766 are mutated into alanines, the indole chemical shifts of W802 in helix α1 and W855 in the α2−β4 loop are significantly affected, suggesting that the linker-conserved motif is positioned in the cavity defined by helix α1 and the α2−β4 loop. Further SAXS and NMR analyses revealed that indeed the K763-X-W765-Q766 motif interacts with the ULM-binding cavity. In our MAN1C model, positively charged residue K763 of the linker is located close to the negatively charged residues of helix α1 and W765 of the linker is close to the S-W-F motif of the α2−β4 loop, similar to what was observed in ULM−UHM complexes like the U2AF35−U2AF65 complex displayed in Figure 7 (*35*). Characterization of such an intramolecular interaction involving a short peptide motif and a structural module has already been described for SH2, SH3 (*36*), PDZ (*37*), and FHA domains (*38*), but this is the first example of an intramolecular interaction involving a ULM−UHM interaction. It has been demonstrated that ULMs recognize various UHM domains with distinct preferences (*39*, *40*). Our intramolecular interaction could be displaced if MAN1 recognizes ULM peptides or UHM domains belonging to other proteins (*41*).

The Intramolecular Linker−UHM Interface Is Critical for Smad2 Binding. MAN1Luhm is the smallest MAN1 fragment that can bind with a micromolar affinity to the wild-type and trimeric Smad2MH2 domains. On one hand, we have shown that the structurally heterogeneous linker and C-terminal regions of MAN1 are involved in Smad2 recognition. Flexibility in these regions suggests that the MAN1−Smad2 interaction involves conformational changes. This is further supported by the finding that the melting temperature of the complex is lower than that of the free proteins: conformational rearrangements necessary for binding of MAN1 to Smad2 diminish the stability of both proteins. On the other hand, the linker alone is not capable of Smad2 binding (*13*) and the MAN1 C-terminus is not essential for Smad2 recognition. We showed that the face of the UHM domain opposite the linker binding region is not involved in Smad2 recognition. However, mutating two hydrophobic residues of the UHM domain, one of which is a buried residue of the hydrophobic core, hindered the MAN1−Smad3 interaction in vivo, suggesting that the UHM domain is necessary for Smad3 binding (*14*). Together, these data suggest that Smad2 recognizes both the linker and the UHM cavity and that the positioning of the linker onto this UHM cavity regulates Smad2 binding (Figure 7). Thus, interaction of the MAN1 linker region with a UHM domain found in another protein should influence recognition of Smad2 by MAN1.

Does MAN1 Compete with Transcription Factors for Smad2 Binding? Several transcription factors of the Fast and Mix families have been described as binding to Smad2 and

Smad3 (*42*, *43*). They are composed of a DNA binding domain and a predicted unfolded region containing the SIM. This motif consists of 19 residues and is centered on a P-P-N-K sequence. In the transcription factors Mixer and Milk, mutation of the two prolines of the SIM motif into alanines abrogated Smad2 binding (*43*). The sequence of MAN1 located after the DNA binding WH domain and centered on residues P778 and P779 is similar to the SIM described in transcription factors (*18*). In MAN1, two-hybrid experiments demonstrated a role for P778 and P779 in the interaction with the MH2 domain of Smad2 but not with full-length Smad2. As the expression yields of MAN1 and Smad2 were very different in the two yeast two-hybrid experiments, as observed by Western blotting (data not shown), a similar decrease in MAN1−Smad2 affinity might have different consequences on the concentration of these complexes in yeast. However, fluorescence measurements revealed that, in vitro, the P778A/P779A double mutation does not influence the K_d value corresponding to the MAN1Luhm−S2LMH2EEE interaction. Thus, the prolines might be located in the Smad2MH2 recognition surface but are not critical for the affinity of MAN1 for the trimeric form of Smad2MH2. Finally, mutations of Smad2MH2 hydrophobic residues described as abrogating the interaction of Smad2 with several transcription factors did not affect the binding to MAN1 (730−910), as observed by two-hybrid methods (data not shown). Our data suggest that the binding of MAN1 to Smad2 does not follow the same structural rules as the binding of transcription factors to Smad2.

In conclusion, a variety of biophysical approaches were necessary to characterize the structurally heterogeneous C-terminal region of MAN1 and to identify the residues critical for its transient interaction with Smad2. We show that, in MAN1, the linker region located between the WH and UHM domains plays the role of an intramolecular UHM ligand motif. We propose that MAN1 can recruit both monomeric and trimeric Smad2 to the inner nuclear membrane through a surface involving this linker, the ULM-binding cavity of the UHM domain, and the C-terminus. The MAN1 linker region could recruit other UHM-containing proteins, for example, protein kinases (*39*), to the inner nuclear membrane. Such interactions would clearly influence recognition of Smad2 by MAN1 and could cause the phosphorylation of MAN1 partners. Finally, the MAN1 linker possesses some sequence similarities with the SIM critical for Smad2 recognition by transcription factors, but further studies are needed to confirm a competition between MAN1 and transcription factors for Smad2 binding.

ACKNOWLEDGMENT

We are grateful to Marie-Bénédicte Barrault and Anne Peyroche for their help in setting up the two-hybrid experiments. We also thank Gabriel David for useful advice during the SAXS experiments, Emmanuel Margeat for essential discussions during fluorescence measurement analyses, and Benjamin Chagot for its help during ITC experiments. Finally, we are grateful to Jose A. Marquez of the High Throughput Crystallization Laboratory at Grenoble for the Thermofluor experiments.

SUPPORTING INFORMATION AVAILABLE

CD spectra of the different mutants of MAN1Luhm, assigned NMR $^{1}H-^{15}N$ HSQC spectrum of MAN1Luhm and its comparison with that of MAN1LuhmWQ, the $^{1}H-^{15}N$ nOes of the MAN1 fragments, analysis of their sequence conservation, Guinier representations of their SAXS intensities, gel filtration profile of S2LMH2EEE, and the Thermofluor analysis of the denaturation temperatures of MAN1Luhm, the MAN1Luhm−S2LMH2EEE complex, and S2LMH2EEE. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Massague, J. (1998) TGF-β signal transduction. *Annu. Rev. Biochem. 67*, 753–791.
2. Feng, X. H., and Derynck, R. (2005) Specificity and versatility in TGF-β signaling through Smads. *Annu. Rev. Cell Dev. Biol. 21*, 659–693.
3. Xu, L., Kang, Y., Col, S., and Massague, J. (2002) Smad2 nucleocytoplasmic shuttling by nucleoporins CAN/Nup214 and Nup153 feeds TGFβ signaling complexes in the cytoplasm and nucleus. *Mol. Cell 10*, 271–282.
4. Tsukazaki, T., Chiang, T. A., Davison, A. F., Attisano, L., and Wrana, J. L. (1998) SARA, a FYVE domain protein that recruits Smad2 to the TGFβ receptor. *Cell 95*, 779–791.
5. Grimm, O. H., and Gurdon, J. B. (2002) Nuclear exclusion of Smad2 is a mechanism leading to loss of competence. *Nat. Cell Biol. 4*, 519–522.
6. Kretzschmar, M., Liu, F., Hata, A., Doody, J., and Massague, J. (1997) The TGF-β family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. *Genes Dev. 11*, 984–995.
7. Kretzschmar, M., Doody, J., Timokhina, I., and Massague, J. (1999) A mechanism of repression of TGFβ/Smad signaling by oncogenic Ras. *Genes Dev. 13*, 804–816.
8. Pera, E. M., Ikeda, A., Eivers, E., and De Robertis, E. M. (2003) Integration of IGF, FGF, and anti-BMP signals via Smad1 phosphorylation in neural induction. *Genes Dev. 17*, 3023–3028.
9. Fuentealba, L. C., Eivers, E., Ikeda, A., Hurtado, C., Kuroda, H., Pera, E. M., and De Robertis, E. M. (2007) Integrating patterning signals: Wnt/GSK3 regulates the duration of the BMP/Smad1 signal. *Cell 131*, 980–993.
10. Sapkota, G., Alarcon, C., Spagnoli, F. M., Brivanlou, A. H., and Massague, J. (2007) Balancing BMP signaling through integrated inputs into the Smad1 linker. *Mol. Cell 25*, 441–454.
11. Lin, F., Blake, D. L., Callebaut, I., Skerjanc, I. S., Holmer, L., McBurney, M. W., Paulin-Levasseur, M., and Worman, H. J. (2000) MAN1, an inner nuclear membrane protein that shares the LEM domain with lamina-associated polypeptide 2 and emerin. *J. Biol. Chem. 275*, 4840–4847.
12. Hellemans, J., Preobrazhenska, O., Willaert, A., Debeer, P., Verdonk, P. C., Costa, T., Janssens, K., Menten, B., Van Roy, N., Vermeulen, S. J., Savarirayan, R., Van Hul, W., Vanhoenacker, F., Huylebroeck, D., De Paepe, A., Naeyaert, J. M., Vandesompele, J., Speleman, F., Verschueren, K., Coucke, P. J., and Mortier, G. R. (2004) Loss-of-function mutations in LEMD3 result in osteopoikilosis, Buschke-Ollendorff syndrome and melorheostosis. *Nat. Genet. 36*, 1213–1218.
13. Lin, F., Morrison, J. M., Wu, W., and Worman, H. J. (2005) MAN1, an integral protein of the inner nuclear membrane, binds Smad2 and Smad3 and antagonizes transforming growth factor-β signaling. *Hum. Mol. Genet. 14*, 437–445.
14. Pan, D., Estevez-Salmeron, L. D., Stroschein, S. L., Zhu, X., He, J., Zhou, S., and Luo, K. (2005) The integral inner nuclear membrane protein MAN1 physically interacts with the R-Smad proteins to repress signaling by the transforming growth factor-β superfamily of cytokines. *J. Biol. Chem. 280*, 15992–16001.
15. Ishimura, A., Ng, J. K., Taira, M., Young, S. G., and Osada, S. (2006) Man1, an inner nuclear membrane protein, regulates vascular remodeling by modulating transforming growth factor β signaling. *Development 133*, 3919–3928.
16. Cohen, T. V., Kosti, O., and Stewart, C. L. (2007) The nuclear envelope protein MAN1 regulates TGFβ signaling and vasculogenesis in the embryonic yolk sac. *Development 134*, 1385–1395.
17. Huber, M. D., Guan, T., and Gerace, L. (2009) Overlapping functions of nuclear envelope proteins NET25 (Lem2) and emerin in regulation of extracellular signal-regulated kinase signaling in myoblast differentiation. *Mol. Cell. Biol. 29*, 5718–5728.
18. Caputo, S., Couprie, J., Duband-Goulet, I., Konde, E., Lin, F., Braud, S., Gondry, M., Gilquin, B., Worman, H. J., and Zinn-Justin, S. (2006) The carboxyl-terminal nucleoplasmic region of MAN1 exhibits a DNA binding winged helix domain. *J. Biol. Chem. 281*, 18208–18215.

19. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc. 34*, 93–158.
20. Jung, Y. S., and Zweckstetter, M. (2004) Mars: Robust automatic backbone assignment of proteins. *J. Biomol. NMR 30*, 11–23.
21. David, G., and Pérez, J. (2009) Combined sampler robot and high-performance liquid chromatography: A fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Crystallogr. 42*, 892–900.
22. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr. 36*, 1277–1282.
23. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr. 25*, 495–503.
24. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRYSOL: A Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr. 28*, 768–773.
25. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR 13*, 289–302.
26. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234*, 779–815.
27. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., III, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res. 35*, W375–W383.
28. Kielkopf, C. L., Lucke, S., and Green, M. R. (2004) U2AF homology motifs: Protein recognition in the RRM world. *Genes Dev. 18*, 1513–1526.
29. Selenko, P., Gregorovic, G., Sprangers, R., Stier, G., Rhani, Z., Kramer, A., and Sattler, M. (2003) Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. *Mol. Cell 11*, 965–976.
30. Conte, M. R., Grune, T., Ghuman, J., Kelly, G., Ladas, A., Matthews, S., and Curry, S. (2000) Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold. *EMBO J. 19*, 3132–3141.
31. Simpson, P. J., Monie, T. P., Szendroi, A., Davydova, N., Tyzack, J. K., Conte, M. R., Read, C. M., Cary, P. D., Svergun, D. I., Konarev, P. V., Curry, S., and Matthews, S. (2004) Structure and RNA interactions of the N-terminal RRM domains of PTB. *Structure 12*, 1631–1643.
32. Durand, D., Vives, C., Cannella, D., Perez, J., Pebay-Peyroula, E., Vachette, P., and Fieschi, F. (2010) NADPH oxidase activator p67-(phox) behaves in solution as a multidomain protein with semi-flexible linkers. *J. Struct. Biol. 169*, 45–53.
33. Randall, R. A., Howell, M., Page, C. S., Daly, A., Bates, P. A., and Hill, C. S. (2004) Recognition of phosphorylated-Smad2-containing complexes by a novel Smad interaction motif. *Mol. Cell. Biol. 24*, 1106–1121.
34. Chacko, B. M., Qin, B., Correia, J. J., Lam, S. S., de Caestecker, M. P., and Lin, K. (2001) The L3 loop and C-terminal phosphorylation jointly define Smad protein trimerization. *Nat. Struct. Biol. 8*, 248–253.
35. Kielkopf, C. L., Rodionova, N. A., Green, M. R., and Burley, S. K. (2001) A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. *Cell 106*, 595–605.
36. Kuriyan, J., and Cowburn, D. (1997) Modular peptide recognition domains in eukaryotic signaling. *Annu. Rev. Biophys. Biomol. Struct. 26*, 259–288.
37. Cheng, H., Li, J., Fazlieva, R., Dai, Z., Bu, Z., and Roder, H. (2009) Autoinhibitory interactions between the PDZ2 and C-terminal domains in the scaffolding protein NHERF1. *Structure 17*, 660–669.
38. Barthe, P., Roumestand, C., Canova, M. J., Kremer, L., Hurard, C., Molle, V., and Cohen-Gonsaud, M. (2009) Dynamic and structural characterization of a bacterial FHA protein reveals a new autoinhibition mechanism. *Structure 17*, 568–578.
39. Manceau, V., Kielkopf, C. L., Sobel, A., and Maucuer, A. (2008) Different requirements of the kinase and UHM domains of KIS for its nuclear localization and binding to splicing factors. *J. Mol. Biol. 381*, 748–762.
40. Corsini, L., Hothorn, M., Stier, G., Rybin, V., Scheffzek, K., Gibson, T. J., and Sattler, M. (2009) Dimerization and protein binding specificity of the U2AF homology motif of the splicing factor Puf60. *J. Biol. Chem. 284*, 630–639.
41. Pufall, M. A., and Graves, B. J. (2002) Autoinhibitory domains: modular effectors of cellular regulation. *Annu. Rev. Cell Dev. Biol. 18*, 421–462.
42. Chen, Y., Lebrun, J. J., and Vale, W. (1996) Regulation of transforming growth factor β- and activin-induced transcription by mammalian Mad proteins. *Proc. Natl. Acad. Sci. U.S.A. 93*, 12992–12997.
43. Germain, S., Howell, M., Esslemont, G. M., and Hill, C. S. (2000) Homeodomain and winged-helix transcription factors recruit activated Smads to distinct promoter elements via a common Smad interaction motif. *Genes Dev. 14*, 435–451.