The Intrinsically Disordered RNR Inhibitor Sml1 Is a Dynamic Dimer[†]

Jens Danielsson,[‡,§] Leena Liljedahl,[§] Elsa Bárány-Wallje,[‡] Pernille Sønderby,[§] Line Hyltoft Kristensen,[§] Maria A. Martinez-Yamout,[||] H. Jane Dyson,[||] Peter E. Wright,[||] Flemming M. Poulsen,[§] Lena Mäler,[‡] Astrid Gräslund,[‡] and Birthe B. Kragelund*[,§]

Department of Biochemistry and Biophysics, Stockholm University, S-106 91 Stockholm, Sweden, Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Ole Maaloes Vej 5, DK-2200 Copenhagen N, Denmark, and Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received June 3, 2008; Revised Manuscript Received August 26, 2008

ABSTRACT: Sml1 is a small ribonucleotide reductase (RNR) regulatory protein in *Saccharomyces cerevisiae* that binds to and inhibits RNR activation. NMR studies of ^{15}N-labeled Sml1 (104 residues), as well as of a truncated variant (residues 50−104), have allowed characterization of their molecular properties. Sml1 belongs to the class of intrinsically disordered proteins with a high degree of dynamics and very little stable structure. Earlier suggestions for a dimeric structure of Sml1 were confirmed, and from translation diffusion NMR measurements, a dimerization dissociation constant of 0.1 mM at 4 °C could be determined. The hydrodynamic radius for the monomeric form of Sml1 was determined to be 23.4 Å, corresponding to a protein size between those of a globular protein and a coil. Formation of a dimer results in a hydrodynamic radius of 34.4 Å. The observed chemical shifts showed in agreement with previous studies two segments with transient helical structure, residues 4−20 and 60−86, and relaxation studies clearly showed restricted motion in these segments. A spin-label attached to C14 showed long-range interactions with residues 60−70 and 85−95, suggesting that the N-terminal domain folds onto the C-terminal domain. Importantly, protease degradation studies combined with mass spectrometry indicated that the N-terminal domain is degraded before the C-terminal region and thus may serve as a protection against proteolysis of the functionally important C-terminal region. Dimer formation was not associated with significant induction of structure but was found to provide further protection against proteolysis. We propose that this molecular shielding and protection of vital functional structures from degradation by functionally unimportant sites may be a general attribute of other natively disordered proteins.

The suppressor of Mec1 lethality (Sml1)[1] is a small ribonucleotide reductase (RNR) regulatory protein in *Saccharomyces cerevisiae* that binds to and inhibits RNR activation. The highly conserved RNR reduces the ribonucleotides to deoxyribonucleotides necessary for DNA synthesis, and in yeast, it is a heterotetramer of one homodimer, R1, and the heterodimer R2−R4 proteins (*1, 2*). Sml1 binds specifically to the R1 component of RNR in a 1:1 stoichiometry with a dissociation constant in the low micromolar regime (*3, 4*), and at an RNR R1:Sml1 ratio of 1:1, 50% inhibition is achieved (*3*). RNR is particularly active during the S phase of the cell cycle. Its activity is regulated on different levels. In *S. cerevisiae*, one of the regulatory mechanisms is via Sml1. First, the expression level of Sml1 is decreased during the S phase of the cell cycle (*5, 6*). Second, a posttranslational modification of Sml1, i.e., phosphorylation of one or more of the serines (S56, S58, and S60), leads to degradation of Sml1, and this constitutes a second level of regulation (*5, 7*). S60 seems to be the most potent regulation site, since it is phosphorylated to the greatest extent of the three. Sml1 regulation of RNR activity is seen not only during normal DNA synthesis but also as a response to DNA damage (*5, 8, 9*).

Sml1 is a small protein of 104 residues, and its sequence (shown in Figure 1A) suggests that it is mostly unstructured and belongs to a group of intrinsically disordered proteins (IDPs) (*10, 11*). Sml1 has, however, been suggested to adopt transient structures in three regions along the chain, two in the N-terminal part of the protein, involving residues 4−14 and 20−35, and one in the C-terminal part, involving residues 61−80. Between these more structured regions, the polypeptide chain is essentially random in structure (*12*). Regions 4−14 and 61−80 have high propensities to adopt transient α-helical conformation, although no well-defined helical structures have been determined. Mutation and deletion studies of Sml1 suggest that the C-terminal helical region is important for Sml1 inhibition of RNR (*12*). Helix breaking

[†] This work was supported in part by the Carlsberg Foundation (F.M.P.), Grant GM7374 from the National Institutes of Health (H.J.D.), and the Danish Natural Research Councils (21040604, B.B.K.).

* To whom correspondence should be addressed. E-mail: bbk@bio.ku.dk. Phone: +45 35 32 20 81. Fax: +45 35 32 21 28.

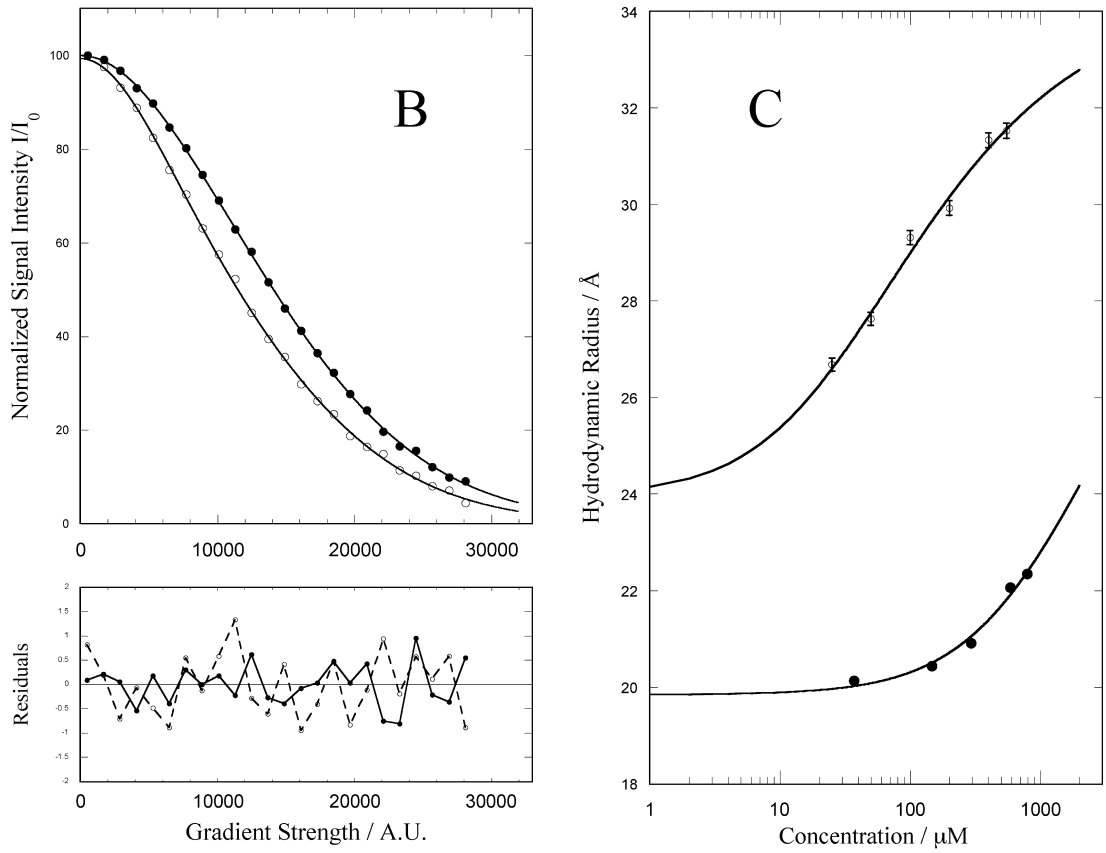
[‡] Stockholm University.

[§] University of Copenhagen.

[||] The Scripps Research Institute.

[1] Abbreviations: Sml1, suppressor of Mec1 lethality; RNR, ribonucleotide reductase; IDP, intrinsically disordered protein; NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum correlation; PFG, pulsed field gradient; NOE, nuclear Overhauser effect; SSP, secondary structure propensity; PRE, paramagnetic relaxation enhancement; MTSL, (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)methanethiosulfonate; MS, mass spectrometry.

10.1021/bi801040b CCC: $40.75
Published on Web 11/17/2008



A

MQNSQDYFYA$_{10}$QNRCQQQQAP$_{20}$STLRTVTMAE$_{30}$FRRVPLPPMA$_{40}$
EVPMLSTQNS$_{50}$MGSSASASAS$_{60}$SLEMWEKDLE$_{70}$ERLNSIDHDM$_{80}$
NNNKFGSGEL$_{90}$KSMFNQGKVE$_{100}$EMDF

FIGURE 1: Hydrodynamic radii of Sml. (A) Amino acid sequence of Sml1. (B) Signal attenuation curves from ^{1}H PFG−NMR translational diffusion measurements of Sml1 at two different concentrations [(●) 550 μM and (○) 25 μM] at 278 K. The lower panel shows the residuals after attenuation data had been fit to the modified Stejskal−Tanner equation described in Experimental Procedures. (C) Concentration dependence of the hydrodynamic radius, calculated from the diffusion coefficients, of full-length wild-type Sml1 (○) and truncated Sml1(50−104) (●). The curves are generated from a simple two-state model in which three parameters are fitted, R_H(monomer), R_H(dimer), and K_D.

mutations inside this region, such as S75P, reduce the inhibitory effect of Sml1 significantly. However, mutations outside the helical region but still in the C-terminal region also decrease Sml1 inhibitory efficiency significantly. Together with the observation that replacement of the C-terminal Phe104 with Leu leads to a considerable reduction in the level of inhibition, this suggests that a large part of the C-terminus of Sml1 is important for its inhibitory function. Even a short peptide with the same sequence as the nine C-terminal residues of Sml1 inhibits RNR activity to some extent. In contrast, deletion of several N-terminal parts of Sml1 had no effect on RNR R1 inhibition (*12*), indicating that the N-terminal region is probably not important for Sml1 function.

Sml1 has been suggested to form a stable dimer, and initially, a disulfide bridge was assumed to form between the cysteine residues of the monomers (*7*). However, mutation studies showed that a dimer was formed even in the absence of the cysteine, suggesting the presence of important intermolecular noncovalent interactions. Since truncation of the 20 outermost N-terminal residues reduced and even eliminated the presence of dimers (*13*), the N-terminal part of Sml1 may control this process.

IDPs make up an important group of proteins, typically involved in regulation of transcription, translation, and cell signaling (*10*). Upon interaction with functional partners, which themselves may be folded or unfolded, IDP domains frequently undergo binding-coupled folding (*10*). The induced structure is often directly related to structural propensities already present in the unbound disordered state. Nuclear magnetic resonance spectroscopy (NMR) is the method of choice for structural and dynamical characterization of such disordered peptide chains (*14*−*17*), and the results can mostly be substantiated by information obtained with other biophysical methods such as circular dichroism, fluorescence, and mass spectrometry (*18, 19*).

Previous NMR studies have shown that the peptide backbone of Sml1 has two regions (4−14 and 61−80) with high α-helix propensities, and relaxation studies have shown that the backbone structure is highly flexible at pH 7.0 with

all S^2 order parameters being <0.6 (*12*). The two helices were suggested to interact with each other and to run antiparallel. Via mapping of solvent-protected side chains of Sml1, it was also suggested to adopt a solvent-excluding three-dimensional structure (*20*). However, no detailed analysis of the structural and dynamical characteristics of Sml1 in solution was reported, and no information regarding details of the dimerization of Sml1 has been reported either. In this work, we have characterized in detail the structure and dynamics of Sml1 using a variety of biophysical methods. We have also specifically studied and characterized the dimerization in terms of the dissociation constant and have identified residues important for dimerization. Furthermore, we have resolved a likely biological role of the N-terminal region.

It is a challenge to couple the structural susceptibilities and dynamic properties of a protein to its functional state, particularly when the functional bound state is not yet known, as for Sml1. Here we have used truncation as a strategy to correlate the structural and dynamical properties of Sml1 to its function. This is important in the ongoing work in understanding Sml1 binding and inhibition of RNR and for the subsequent manipulation of the potential regulation of eukaryotic RNR.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. The *S. cerevisiae* WT *SML1* gene was cloned into the pET-3a vector between the BamHI and NdeI restriction sites. To ensure that the cloning was correct, the plasmid was cut by BamHI and NdeI restriction enzymes and PCR amplification of the insert was performed. BL21(DE3) and BL21(DE3) pLysS competent *Escherichia coli* cells were used for the IPTG-induced expression of Sml1 from vector pET-3a. Recombinant wild-type Sml1 and the studied variants of Sml1 were expressed as described by Chabes et al. (*3*). Cells were harvested after 4 h by centrifugation, resuspended in 50 mM Tris-HCl (pH 7.4) and 1 mM EDTA, and lysed by sonication for 3 × 30 s on ice. The cells were centrifuged at 20000*g* for 25 min at 4 °C. Proteins in the supernatant were precipitated by adding solid ammonium sulfate to 25% saturation at 0 °C (136 g/L) and centrifuged at 12000*g* for 15 min at 4 °C. The supernatant was carefully removed, and the pellet was dissolved in 10 mM Tris-HCl (pH 8.0) (RPC buffer A). Further purification was done using reverse phase HPLC using a SOURCE 15 RPC ST 4.6/100 column. Aliquots of protein were injected onto the column, and Sml1 eluted in a linear gradient in RPC buffer B [10 mM Tris-HCl (pH 8.0) and 60% MeCN], increasing from 25 to 60% B (v/v) over 10 column volumes, with a flow rate of 1 mL/min.

Spectroscopic Methods. All experiments were performed at 4 °C and pH 7 (6.9−7.4) in 10 mM phosphate buffer. C^{α}, C, N, and HN chemical shifts of ^{13}C- and ^{15}N-labeled Sml1 were recorded at pH 7, 4 °C, and 0.6 mM using standard ^{15}N HSQC, HNCA, HN(CO)CA, HNCACB, CBCACONH, HN(CA)CO, and HNCO experiments on a Varian Inova 800 MHz spectrometer equipped with a cryogenically cooled triple-resonance probe. Spectra were processed using nmrPipe and analyzed using Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) and Pronto (*21*).

Diffusion experiments were performed using a PFG-LED sequence with a gradient prepulse (*22, 23*) on a Varian Inova 750 MHz spectrometer. The ^1H signal intensity was determined at 24 linearly spaced gradient strengths, and the attenuating intensity was fitted to a modified Stejskal−Tanner equation that accounts for nonlinearities in the gradient field profile (*24, 25*). The signal intensity was measured on methyl groups and aromatics separately. The diffusion delay was 150 ms, and the gradient pulse was 5 ms. The hydrodynamic radius was calculated using the Stokes−Einstein expression and was determined three times at each concentration. The dilution series was performed using a high-concentration sample that was diluted in several steps, and the diffusion coefficient was determined after a 1−2 h delay, necessary to achieve an equilibrium state. The fitting routine was performed using MATLAB (Mathworks, Natick, MA). The gradient was calibrated using the HDO and α-cyclodextrin diffusion at 4 °C (10.5×10^{-10} and 1.46×10^{-10} m^2/s, respectively). The concentration dependence of the hydrodynamic radii was fitted to a model with an equilibrium between monomer, M, and dimer, D, assuming that the exchange between these states is fast on the diffusion time scale

$$M + M \leftrightarrow D \quad (1)$$

The dissociation constant, K_D, is the ratio of the equilibrium concentrations of monomers and dimers.

$$K_D = [M]^2/[D] = (C_0 p_{monomer})^2/[(1/2)C_0(1 - p_{monomer})] \quad (2)$$

where C_0 is the total protein concentration and $p_{monomer}$ is the monomer population. At each concentration, the population of monomer can be determined as the weighted mean of the observed diffusion coefficient.

$$D_{obs} = p_{monomer} D_{monomer} + (1 - p_{monomer}) D_{dimer} \quad (3)$$

By combining eq 1−3 and determining D_{obs} at several concentrations (C_0), we can determine K_D, $D_{monomer}$, and D_{dimer} from a nonlinear fit. This was performed using MATLAB.

The relaxation experiments were performed using standard HSQC sequences and at three different fields, 500 MHz (Bruker Avance, Karlruhe, Germany) and 600 and 800 MHz (Varian Inova, Palo Alto, CA) proton frequency fields. The signal attenuation, from 10 different relaxation delays, was fitted to a single-exponential decay, and the relaxation rates were determined. The relaxation rates were determined at a high protein concentration, 0.7 mM.

Paramagnetic relaxation enhancement experiments were performed as described previously, and Sml1 was spin-labeled with MTSL as described previously (*26*). The labeled and unlabeled fractions were separated using reverse phase high-pressure liquid chromatography. The MTSL spin-label was reduced using a 4-fold molar excess of ascorbic acid and by readjusting the pH.

Determination of Hydrodynamic Radii by Gel Filtration. Samples of WT-Sml1 were analyzed at different concentrations (9.8 and 0.8 mM) using an ÄKTA HPLC purifier on a prepacked Superdex G75 10/300 Tricorn high-performance column with a V_t of 24 mL with a 20 mM Tris-HCl (pH 7.5) and 150 mM NaCl, at a flow rate of 0.5 mL/min. The column was calibrated using ovalbumin (43 kDa, 30 Å), myoglobin (17 kDa, 17 Å), BSA (66 kDa, 35.5 Å) α-chymotrypsinogen A (25 kDa, 20.9 Å), and ribonuclease A (13.7

kDa, 16.4 Å), as well as blue dextran (V_0) and acetone (V_t). K_{AV} was calculated with the equation $K_{AV} = (V_e - V_0)/(V_t - V_0)$, where V_e is the elution volume of each protein.

Protease Resistance. Protease K and trypsin proteomics grade were purchased from Sigma Aldrich (St. Louis, MO) and used without further purification. The protease K assays were conducted in 50 mM Tris-HCl (pH 8) and 5 mM $CaCl_2$ at room temperature. Protease K was added to protein samples in a molar ratio of 1:10000. Aliquots of the protein/protease K mixture were withdrawn after varying delays, and the reaction was stopped via addition of PMSF to a final concentration of 5 mM. Aliquots were analyzed using SDS−PAGE. Limited trypsin digestion was carried out in 100 mM NH_4HCO_3 (pH 7.4) at room temperature. Trypsin was added to the protein samples in a molar ratio of 1:2000. Aliquots of the protein/trypsin mixture were withdrawn after 0, 5, 10, 20, or 40 min, and the reactions were stopped by adding TFA to a final concentration of 1% (v/v). Samples were analyzed via SDS−PAGE or MALDI-TOF MS.

MALDI-TOF MS. One microliter of 20 μM digest was mixed with 1 μL of HCCA [10 mg/mL in 50% (v/v) MeCN and 0.1% (v/v) TFA] and spotted on a MTP 384 ground steel TF target plate by the dried droplet method. Spectra were recorded on a Bruker Daltonics Autoflex II TOF mass spectrometer in linear positive mode. An average of 100 shots was recorded. Calibration was performed with quadric calibration of either protein calibration standard 1 or peptide calibration standard (both purchased from Bruker Daltonics). Spectra were analyzed with MoverZ (Proteometrics, Inc.). Primary sites were identified from two corresponding fragments giving the full sequence.

RESULTS

Structural Characterization. A number of NMR experiments were performed to elucidate the residual secondary structure of Sml1 and the oligomeric state of the protein. The $^1H-^{15}N$ heteronuclear single-quantum correlation (HSQC) spectra of Sml1 at a wide range of concentrations were well-resolved; cross-peaks corresponding to 96 of 99 non-proline residues have been assigned. The assignment is in general agreement with earlier findings (*12*) with a few exceptions. The small spectral width of the amide signals of Sml1 clearly indicates that the protein is mainly unfolded, as predicted also from the sequence properties.

Hydrodynamic Dimensions of Sml1. From the hydrodynamic radius, it is possible to distinguish between the presence of folded and unfolded protein in a sample (*16, 27, 28*) and between monomeric and dimeric protein. It has been suggested that Sml1 exists as a dimer (*13*). Assuming that the protein is in equilibrium between monomeric and dimeric states and that the exchange between these states is fast on the diffusion time scale, NMR pulsed field gradient (PFG) diffusion experiments in a dilution series should give direct evidence of such an equilibrium (*24, 27, 28*). Figure 1B shows the attenuation curves from Sml1 at 4 °C in D_2O at two different concentrations, 550 and 25 μM. It is immediately clear that Sml1 diffuses significantly faster at the low concentration. We performed a dilution series and determined the hydrodynamic radius at several concentrations. The results are presented in Figure 1C. These data fit well to a model with a monomer−dimer equilibrium using three fitted parameters, the hydrodynamic radius for the monomer (23.4 ± 1 Å) and for the dimer (34.4 ± 0.5 Å) and a dissociation constant K_d of 93 ± 25 μM (Table 1).

Table 1: Measured and Theoretically Estimated (*21*) Hydrodynamic Radii of Monomeric and Dimeric Sml1

	R_H(monomer) (Å)	R_H(dimer) (Å)
experimental	23.4	34.4
theoretical unfolded	29.4	41
theoretical folded	18.8	23.2

The hydrodynamic radius obtained for the monomer of Sml1 is significantly smaller than the R_H of ≈29.4 Å expected for an unstructured protein the same size as Sml1 (*27*). A folded monomer with a mass of 11.8 kDa is expected to have an R_H of ≈18.8 Å. This suggests that the monomer of Sml1 exhibits some globular structure, but with few contributions from well-defined secondary structure. This globularity is presumably due to intermittent tertiary interactions in the monomeric state. The expected hydrodynamic radius of a totally unfolded dimer with a mass of 2 × 11834 Da is 41 Å, and for a compactly folded dimer, we expect an R_H of ≈23.2 Å. The obtained value for the Sml1 dimer falls between these calculated values. Assuming that the dimer is formed by two relatively unstructured monomers, the expected increase in hydrodynamic radius is approximately $2^{1/2}$. The experimentally obtained increase is 1.47 and is in good agreement with the expected increase factor (*27*).

Size exclusion chromatography was used to also estimate the hydrodynamic dimensions of Sml1 (Figure 2S). Loading the column with a high concentration leads to dilution on the column into both monomer and dimer, and wild-type (WT) Sml1 eluted as two peaks, corresponding to 33 and 23 Å, in good agreement with the results from the diffusion experiments.

Chemical Shift Analysis. Chemical shifts are very sensitive probes of secondary structure (*29−32*). One strategy in the chemical shift analysis is to compare the obtained chemical shifts with a reference value for a random coil. The chemical shift assignments were reperformed at a high protein concentration (>0.6 mM) with a dimer fraction of approximately 80%.

The secondary chemical shift analysis (Figure 2) suggests the presence of two helical regions in Sml1. One is located in the central part of the protein and involves residues 60−86. A second region is located in the N-terminal part and includes residues 4−20, in agreement with previous observations (*12*). The maximum ^{15}N secondary chemical shift in each helical region shows a systematic four-residue shift toward the C-terminus compared to the carbon shifts. This may indeed reflect hydrogen bonding between CO_i and H^N_{i+4} as present in an α-helix. This has also been observed in the unfolded state of the all-helical protein ACBP (*33*). The effect is most prominent in the helical region of residues 60−86, but it is also observed in the N-terminal helix and may reflect transient hydrogen bond formation. A third region of Sml1 involving residues 30−45 shows by both scales chemical shifts distinctly deviating from random coil shifts. This region exhibits a different behavior than the other two transient helical regions, and it is not straightforward to conclude the precise nature of the transiently formed structure in this region.

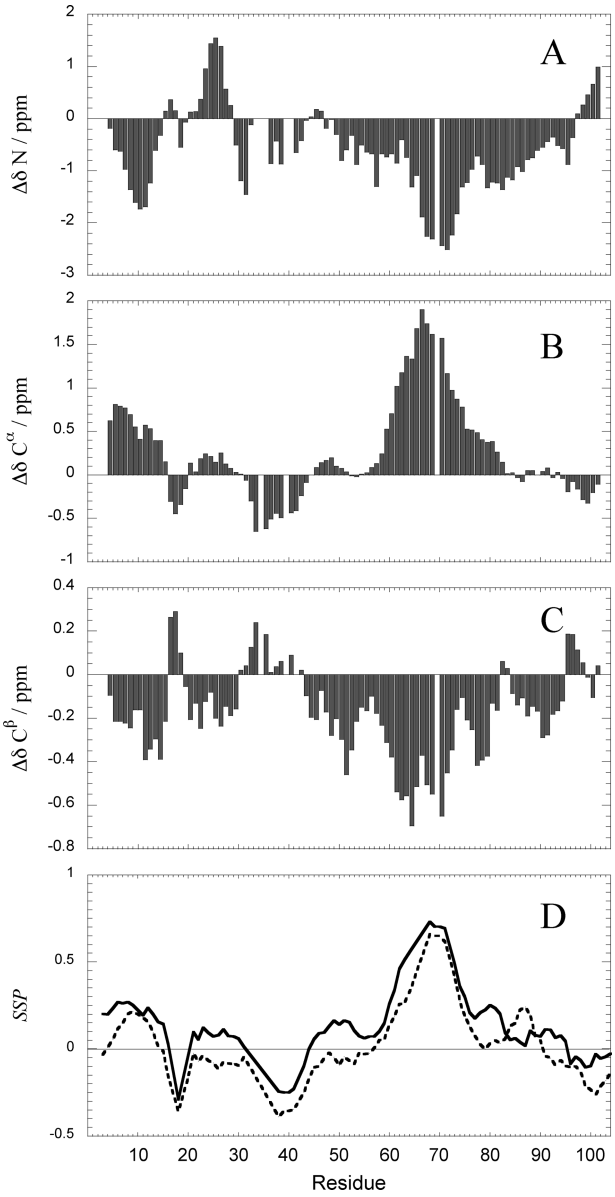


FIGURE 2: Secondary chemical shifts. (A−C) Secondary chemical shifts of Sml1 relative to random coil values (*41*). (D) SSP analysis performed as described by Marsh et al. (*34*). Values close to 1 correspond to fully formed α-helix, and values close to −1 correspond to fully formed β-sheet.

To quantify the secondary structure propensity, a secondary structure propensity (SSP) analysis was performed, following the procedure described by Marsh et al. (*34*). The SSP analysis was performed using both the carbon shifts alone and all available secondary chemical shifts, and the results are shown in the bottom panel of Figure 2. The SSP analysis shows that the helix propensity in the region of residues 60−80 is very high and almost corresponds to a fully formed helix. Interestingly, the region of residues 33−45 has SSP values corresponding to β-sheet propensity. From the SSP analysis, the helix and β-population may be estimated, and we found that Sml1 is 14% helical and has 2% β-strand.

The chemical shift analysis of Sml1 does, however, not suggest a well-defined structure but rather shows localized populations of residual secondary structures in distinct parts of Sml1 in solution.

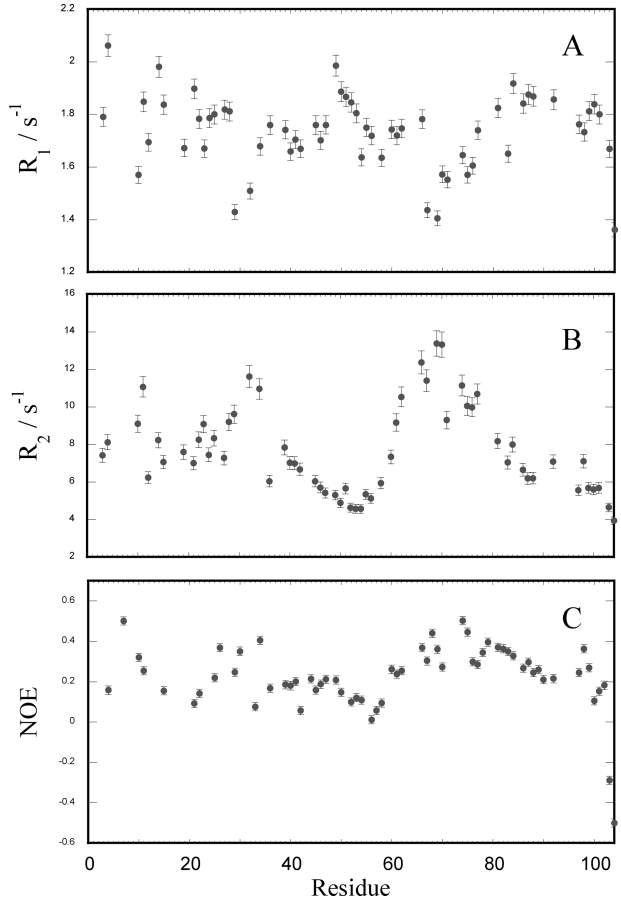


FIGURE 3: Relaxation properties of Sml1. ^{15}N relaxation data of 0.7 mM Sml1 (pH 7.2) in 10 mM phosphate buffer at 278 K. The data shown are R_1, R_2, and NOE values at a proton frequency of 600 MHz. Missing data points are due to poor resolution and/or an unstable data fit and/or proline residues.

Table 2: Averaged Relaxation Values for R_1 and R_2 at Three Magnetic Fields

magnetic field (T)	R_1 (s^{-1})	R_2 (s^{-1})
11.74	1.8 ± 0.1	7.4 ± 2.7
14.09	1.8 ± 0.2	8.4 ± 2.9
18.79	1.7 ± 0.2	9.4 ± 3.4

NMR Relaxation. Regions with experimental evidence of populations adopting secondary structure are expected to exhibit slower dynamics than truly unfolded regions. The NMR relaxation phenomenon is directly linked to both the local and global dynamics of the protein (*35, 36*). R_1, R_2, and ^{15}N NMR relaxation measurements combined with steady state heteronuclear ^{15}N NOEs were used to obtain information about the dynamical properties of Sml1. The relaxation rates and NOEs were determined at three different magnetic fields, 11.74, 14.09, and 18.79 T. The results from 14.09 T are shown in Figure 3. The average relaxation rates at the different fields are listed in Table 2.

The R_2 values reveal three regions with higher R_2 relaxation rates, and two of these regions are evident also in the R_1 data. The region involving residues 60−86 has the highest R_2 values, as high as 13 s^{-1} at 14.09 T, but the two N-terminal regions of residues 4−20 and 25−45 exhibit relaxation faster than average with rates of up to 11 s^{-1}. The NOE values are all above 0, and the pattern is similar to that for the relaxation rates with higher NOE values for these

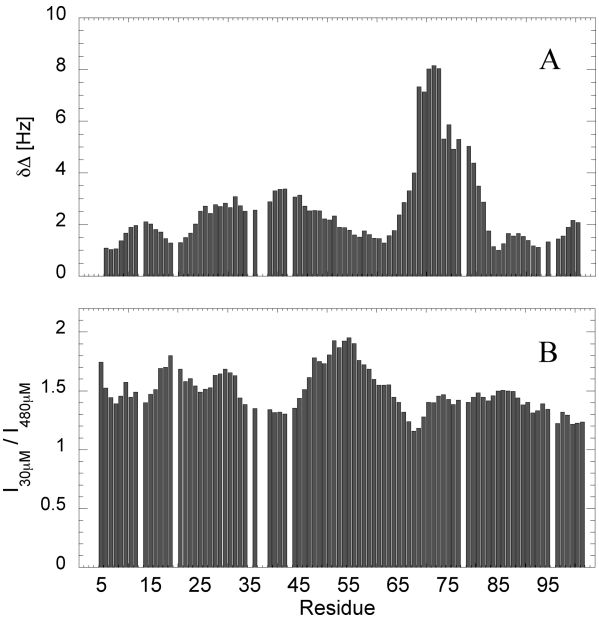


FIGURE 4: Dimerization interface of Sml1. (A) Induced chemical shift changes of $^{1}H^{N}$ and ^{15}N upon dilution of Sml1 from 0.48 mM to 30 μM in 10 mM phosphate buffer (pH 7.2) at 278 K. The values are averaged over five residues. The most prominent changes are seen in the region involving residues 60−80, suggesting involvement in the dimerization, but residues in the region of the N-terminal helix are also affected. (B) Relative change in signal intensity upon dilution of the sample from 0.48 mM to 30 μM. The values are averaged over five residues. Here the highly mobile region of residues 45−60 is most affected, suggesting that this region increases its mobility upon dissociation from the dimeric to monomeric state.

three regions. Heteronuclear NOEs are very sensitive to local mobility, and reduced NOEs, i.e., larger values, typically indicate restricted motion. Taken together, these data indicate that the three regions involving residues 4−20, 25−45, and 60−86 are more ordered and separated by regions with higher mobility. The C-terminal end of the protein has a higher mobility than the N-terminus.

Dimerization Interface. We studied the dimerization interface further using dilution series and induced spectral changes. Lowering the protein concentration drives the monomer−dimer equilibrium toward a higher monomer fraction. The induced chemical shift changes presumably reflect the regions of the protein that are involved in the dimerization. Figure 4 shows the induced chemical shift changes in the $^{1}H^{N}$ and ^{15}N resonances induced upon dilution of the sample from 480 to 30 μM, corresponding to a decrease in the dimer fraction from ~0.7 to 0.2. Here one region shows significant changes, involving residues around residues 60−80, corresponding to the region with the highest helix propensity. There are also some general changes mainly involving the two regions centered on residues 10 and 35, also identified as regions with populations of transient structure. The sign of the chemical shift change suggests an increase in secondary structure content, suggesting that dimer formation stabilizes the transient secondary structure.

The signal intensity change upon dilution of Sml1 was also studied, and the relative changes in signal intensity after dilution of Sml1 from 480 to 30 μM are shown in Figure 4B. The signal intensity is generally increased for the diluted sample, suggesting a lower molecular weight and a faster overall tumbling rate. The region of residues 45−60, however, shows a significantly greater increase in intensity, suggesting that these regions exhibit a greater gain in mobility upon dissociation of the dimer. This region corresponds to the highly mobile region between the more structured parts of Sml1.

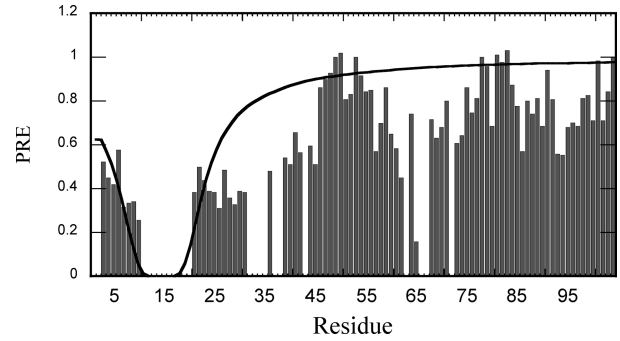


FIGURE 5: Long-range interactions in Sml1. Paramagnetic relaxation enhancement of Sml1 after selective spin labeling of C14 with MTSL. Two regions show long-range interactions with the N-terminus and the regions of residues 60−75 and 85−95. The solid line represents the expected value of the PREs for a Gaussian random coil. This was calculated by creating 10000 Gaussian-excluded volume bead-on-a-string chains, with an excluded volume diameter (D_V) of 5 Å, parametrized to yield hydrodynamic radii corresponding to a classic random coil.

Long-Range Interactions in Sml1. Long-range interactions in Sml1 were examined by the application of spin-labeled Sml1 studied by paramagnetic relaxation enhancement (PRE) measurements. C14 was used to label the protein with MTSL, according to standard methods (*26*). $^{1}H-^{15}N$ HSQC spectra were recorded with the spin-label in a paramagnetic and in a diamagnetic state. The cross-peak intensity changes are related to the relaxation enhancement which depends on the inverse sixth power of the distance between the amide nitrogen−proton vector and the spin-label. Figure 5 shows the PREs calculated from the paramagnetic:diamagnetic ratios of the cross-peak intensity per residue. As expected, residues sequentially closest to the spin-label are broadened beyond detection, but there are (at least) two additional regions that are affected, residues 60−70 and 85−95. For Sml1, the general broadening effect is closely correlated to the amide proton exchange rate and hence the peak intensity in the $^{1}H-^{15}N$ HSQC spectra. The two regions described above, however, fall outside this correlation and seem to originate from long-range interactions within the protein (Figure 1S of the Supporting Information).

To distinguish if the PREs observed are either intra- or intermolecular, spin-labeled [^{14}N]Sml1 was mixed with [^{15}N]Sml1 without a spin-label. No significant line broadening was observed in the [^{15}N]Sml1 HSQC spectrum, suggesting strongly that the observed spin-label-induced PREs are intramolecular. Interestingly, changes in chemical shift on [^{15}N]Sml1 as a result of reduction of the MTSL label on [^{14}N]Sml1 were, however, observed for residues around position 70. This suggests that this region may be involved in both intra- and intermolecular interactions within Sml1 (data not shown).

A rigorous treatment of the PRE data should include several spin-label positions as well as an ensemble simulation (*37*). These data suffer from the fact that the spin-label has only been introduced at one site of the protein, i.e., at the single native cysteine, C14. However, the PRE data are in

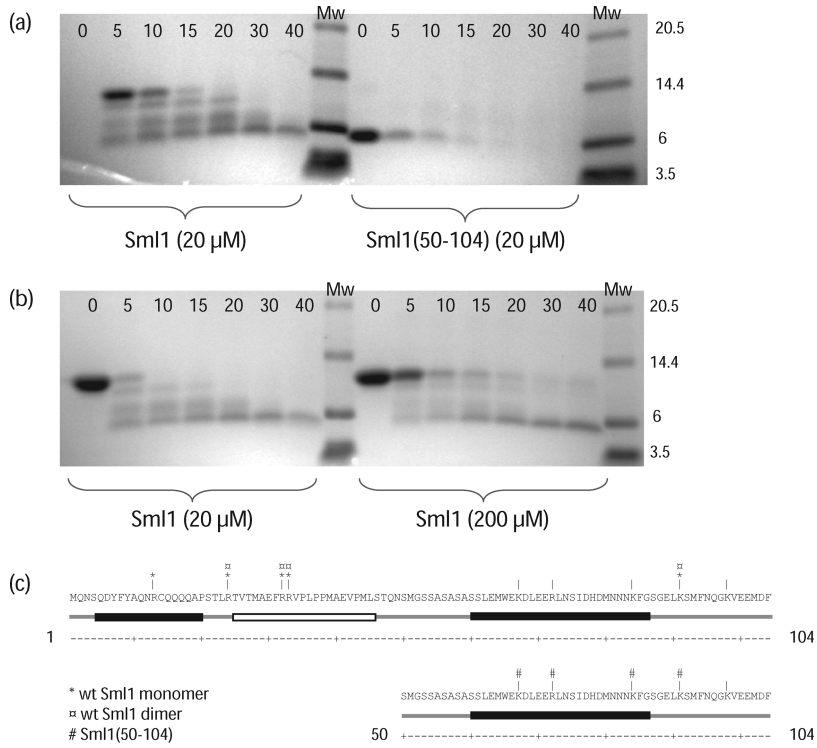


FIGURE 6: Protection of Sml1 from protease degradation. (A) SDS−PAGE following the protease K degradation of wild-type Sml1 and Sml1(50−104) at 20 μM as a function of time (minutes), showing a slower degradation rate for the full-length protein when compared to that of the N-terminally truncated Sml1(50−104). Note that the first time point for wild-type Sml1 is missing due to precipitation. (B) SDS−PAGE following the protease K degradation of Sml1 at two different concentrations corresponding to a high (200 μM, 69% dimer) and a low (20 μM, 31% dimer) population of dimers. The degradation experiment shows that the dimer is significantly more slowly degraded, suggesting that dimerization also protects the protein from degradation. (C) Tryptic cleavage sites in Sml1(monomer), Sml1(50−104), and Sml1(dimer) as observed by MALDI-TOF MS from limited trypsin digestion at 40 min of reaction. Possible trypsin cleavage sites in the sequences are indicated with vertical lines above the residues. Asterisks denote sites observed in monomeric Sml1 at 20 μM, circles denote sites observed in dimeric Sml1 (200 μM), and number signs denote sites observed in Sml1(50−104) at 20 μM. The lines below the sequences show the partial structured region with helix populations colored black and mixed structure colored white.

good agreement with our results obtained from other methods. As seen directly from the PRE data, the N-terminus is close to two regions in the C domain. This indicates that Sml1 has a transient overall folded structure where the N-terminal domain folds onto and caps the C domain. This is in agreement with the relatively small hydrodynamic radius of Sml1 (23.4 Å) measured by diffusion, as described above.

NMR Studies of a Truncated Variant of Sml1. The N-terminal domain may be defined to consist of residues 1−49 and the C-terminal domain to consist of residues 50−104. The C-terminal fragment, Sml1(50−104), was obtained and studied by NMR, both to determine the oligomeric state and to measure the transient secondary structure population of this inhibitory effective fragment.

The nitrogen secondary chemical shifts of Sml1(50−104) exhibited a pattern similar to that of the C-terminal domain of full-length WT Sml1. The C-terminal fragment has a high population of transient helical structure in a region spanning residues 53−70, suggesting that Sml1(50−104) has helical properties similar to those of the wild-type protein. The transient helical region has similar length in the wild type and Sml1(50−104), while the region is shifted toward the N-terminus in Sml1(50−104) (data not shown). To define the oligomeric state of Sml1(50−104), the concentration dependence of the translational diffusion coefficient was studied. The dimerization of WT Sml1 was shown above to involve residues in both the N-terminal and C-terminal parts of Sml1, and thus, fragment 50−104 was not expected to form dimers with the same specificity as the wild type. Figure 1B also shows that the diffusion coefficient of Sml1(50−104) displays only a weak concentration dependence. The diffusion coefficient of Sml1(50−104) corresponds to a hydrodynamic radius (R_H) of 19.9 Å and suggests that Sml1(50−104) has a more extended, random structure with some transient secondary structure regions, as shown with chemical shift analysis.

The central helical region involving residues 60−86 is crucial for Sml1 inhibition of RNR R1 (*12*), and this gives rise to the question of the role of the N-terminal domain, if it is not important for function. The N-terminal domain has two regions with significant populations of secondary structure which suggest also that this domain has a role in Sml1 overall function.

Protease Resistance of Sml1 and Sml1(50−104). To study the importance of the N-terminal domain of wild-type Sml1, we performed protease degradation studies with both wild-type Sml1 and the N-terminally truncated variant, both at monomeric and dimeric concentrations. In separate studies, protease K or trypsin was added to stock solutions of Sml1 or Sml1(50−104) at a low protein concentration. The degradation of the proteins was followed over time and analyzed by SDS−PAGE and MALDI-TOF MS (Figure 6). The degradation produces several distinct fragments, manifested as clear-cut bands in the gel (Figure 6A). It is clear that the degradation resistance of wild-type Sml1 is significantly stronger than for the shorter variant, and we qualita-

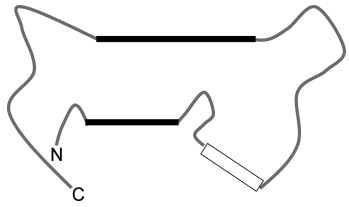


FIGURE 7: Naïve structural model of Sml1. A simple structural sketch of the backbone of Sml1 providing a naïve model which is in agreement with our data. The lines show a partial structured region with helix populations colored black and mixed structure colored white. The structural propensities are not absolute, and thus, we choose to represent the structure using this naïve model.

tively estimate the degradation time to be at least 4 times the degradation time of the truncated variant Sml1(50−104) (Figure 6A). From the limited trypsin digestion of wild-type Sml1, we observe fragments resulting only from the N-terminal domain, whereas fragments resulting from cleavage in the C-terminal domain were not seen. For Sml1(50−104), we distinctly observe degradation fragments resulting from cleavage in the C domain, which were not observed for wild-type Sml1 at the same time point (Figure 6C). Independently, the protease K and trypsin digestion data suggest that the N-terminal domain protects the C domain of Sml1 from degradation and may provide an explanation for why the N-terminal region has been conserved in evolution.

We also performed degradation studies of the full-length protein at two different concentrations to determine the effect of dimerization on the degradation rate. Figure 6B shows the degradation of wild-type Sml1 at 20 and 200 μM, corresponding to approximately 20 and 60% of Sml1 dimers, respectively. The degradation of the sample with a high dimer concentration is significantly slower as seen when comparing qualitatively the intensities of the degradation fragment bands to those of the nondegraded bands. The low-concentration sample of Sml1 (mainly monomeric) degrades approximately 4 times faster than the high-concentration sample of Sml1 (mainly dimeric), suggesting that the formation of dimers significantly protects the protein from degradation. Apart from delayed degradation, the only difference in the cleavage patterns in relation to dimerization as analyzed by limited trypsin digestion and MALDI-TOF MS was seen at residue 13, which could be observed as a cleavage point only at a low protein concentration.

DISCUSSION

We present here a structural characterization of the intrinsically unfolded protein Sml1 and show that even in the absence of well-defined secondary structures, Sml1 adopts a tertiary fold where the N-terminal domain folds onto and caps the C-terminal domain in a protease protective way. A structural sketch is shown in Figure 7. For the sake of clarity, we represent the transient structural regions as pronounced lines and refer to these regions as helical or structured, although these structural elements are highly transient.

The hydrodynamic radius of Sml1 as determined by PFG−NMR and size exclusion chromatography suggests that Sml1 does not behave as a random coil or as an extended unfolded protein. The small hydrodynamic radius suggests that Sml1 adopts a loosely folded structure. This is in agreement with earlier suggestions that Sml1 adopts a loose overall tertiary structure (*12, 20*). This is also supported by the PRE data which indicate long-range contacts between the N-terminal domain and regions 60−80 and 85−95 and is also suggested from the chemical shift analysis of the 50−104 fragment. There is no evidence of a well-defined overall structure, and we suggest transient but specific capping of the C domain by the N domain. The PRE data of disordered proteins shall be interpreted with caution. The distances in a highly mobile system should be considered as distributions of distances rather than fixed distances. Also, the amide proton relaxation rates and the amide proton exchange rates are expected to vary significantly along the polypeptide chain in a highly dynamical protein. The amide proton exchange rate is reflected by the signal intensity in the HSQC spectrum, and we used this effect to be able to separate the significant PREs from those mainly affected by amide proton exchange effects. Despite all these interpretation problems arising from the dynamic features of a disordered protein, the PRE data are in good agreement with the relaxation and diffusion data, and thus, the two long-range contact regions found here are considered significant.

The induced changes in chemical shift upon dimerization as shown in Figure 4 suggest that the dimerization does not induce any significant change in the monomeric structure of Sml1. Our finding that the most affected residues (60−80) are involved in specific dimerization is in apparent disagreement with an earlier finding that the dimerization site is located in region 1−20 (*13*). However, as the N-terminal domain, and especially helical region 4−20, is important for the overall fold of Sml1, as described above, deletion of this region can be expected to affect dimerization indirectly. This suggests that the N-terminal region of Sml1 is important for dimerization. However, the largest impact on chemical shifts is seen in region 60−80, and we therefore propose the following model; for specific dimerization interactions to occur, Sml1 has to fold, i.e., the N-terminus has to cap the C-terminus. The dimerization leads to stabilization of the transient secondary structures of Sml1. The dimerization together with the induced stabilization then restricts the mobility of the flexible region of residues 45−60, leading to a reduction in the intensity of NMR signals from these residues.

We were able to show that both the dimerization and the N domain capping of the C domain protect Sml1 from degradation. The N-terminal domain capping decreases the degradation rate by a factor of 4 and protects specific cleavage sites in the C domain as seen by MALDI-TOF MS. This leads to a model in which the C domain is the inhibition domain and the N domain is a shield, a molecular aegis, for the functional sites, protecting these from degradation. Dimerization serves as an additional and significant protection from protease degradation. This is interesting and to the best of our knowledge a novel finding, since functional dimerization giving rise to protection from degradation has hitherto not been frequently observed in IDPs, although nonfunctional dimerization has been observed (*38*), as well as disordered proteins that fold upon dimerization (*39*). In a recent paper, Simon et al. showed that the IDP UmuD2 forms stable homodimers, however with a significant gain of structure (*40*). This is in contrast to what we observe here. We propose that the molecular shielding and protection of vital functional structures from degradation by functionally unimportant sites may be a general attribute of other IDPs.

The protective role of the N-terminus may be used in the regulation of Sml1 function by the yeast cell. When RNR activity is needed, e.g., as a response to DNA damage, phosphorylation of serines 56, 58, and/or 60 is induced. This may perturb long-range interactions within Sml1 and increase the rate of degradation which results in an increased amount of free, active RNR. This strategy to use phosphorylation as a regulation has been found in other unfolded proteins (*10* and references therein).

Our data lead to a model of Sm1l function and provide an explanation for the observation that Sml1 forms dimers. The dissociation constant for Sml1 dimerization is relatively high, suggesting that the biological relevance may be questioned. However, at translation, the local concentration of Sml1 may be high enough to promote dimerization, and this may then be an effective protection against degradation directly after translation. After translation, Sml1 can be assumed to be more diluted, and this may shift the equilibrium toward a higher population of monomers. To protect against degradation, the N-terminal domain is folded onto the C-terminal domain. However, the C domain may have to be uncapped to be able to bind specifically to subunit R1 of RNR. This may provide an explanation for why the protein does not form a tight and well-defined structure. To enable interaction between the helical region (60−86) and the binding site on R1, the N domain may need to rearrange and expose the helical region.

ACKNOWLEDGMENT

We thank Lars Thelander and Vladimir Domkin for fruitful discussions and Signe Agernæs Sjørup for skilled technical assistance. We acknowledge the points raised by the reviewers concerning the dimerization process.

SUPPORTING INFORMATION AVAILABLE

Normalized peak height of $^{1}H-^{15}N$ HSQC spectra of the Sml1 sample versus I_{OX}:I_{RED} ratios of the [^{15}N]Sml1-MTSL sample with peak heights measured from a $^{1}H-^{15}N$ HSQC spectrum (Figure S1) and size exclusion chromatography profiles of Sml1 (Figure S2). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Chabes, A., Domkin, V., Larsson, G., Liu, A., Gräslund, A., Wijmenga, S., and Thelander, L. (2000) Yeast ribonucleotide reductase has a heterodimeric iron-radical-containing subunit. *Proc. Natl. Acad. Sci. U.S.A. 97*, 2474–2479.
2. Gräslund, A., and Sahlin, M. (1996) Electron paramagnetic resonance and nuclear magnetic resonance studies of class I ribonucleotide reductase. *Annu. Rev. Biophys. Biomol. Struct. 25*, 259–286.
3. Chabes, A., Domkin, V., and Thelander, L. (1999) Yeast Sml1, a protein inhibitor of ribonucleotide reductase. *J. Biol. Chem. 274*, 36679–36683.
4. Zhao, X., Muller, E. G. D., and Rothstein, R. (1998) A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. *Mol. Cell 2*, 329–340.
5. Zhao, X., Chabes, A., Domkin, V., Thelander, L., and Rothstein, R. (2001) The ribonucleotide reductase inhibitor Sml1 is a new target of the Mec1/Rad53 kinase cascade during growth and in response to DNA damage. *EMBO J. 20*, 3544–3553.
6. Zhao, X., and Rothstein, R. (2002) The Dun1 checkpoint kinase phosphorylates and regulates the ribonucleotide reductase inhibitor Sml1. *Proc. Natl. Acad. Sci. U.S.A. 99*, 3746–3751.
7. Uchiki, T., Dice, L. T., Hettich, R. L., and Dealwis, C. (2004) Identification of phosphorylation sites on the yeast ribonucleotide reductase inhibitor Sml1. *J. Biol. Chem. 279*, 11293–11303.
8. Elledge, S. J. (1996) Cell cycle checkpoints: Preventing an identity crisis. *Science 274*, 1664–1671.
9. Taylor, S. D., Zhang, H., Eaton, J. S., Rodeheffer, M. S., Lebedeva, M. A., O'Rourke, T. W., Siede, W., and Shadel, G. S. (2005) The conserved Mec1/Rad53 nuclear checkpoint pathway regulates mitochondrial DNA copy number in *Saccharomyces cerevisiae*. *Mol. Biol. Cell 16*, 3010–3018.
10. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol. 6*, 197–208.
11. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol. 293*, 321–331.
12. Zhao, X., Georgieva, B., Chabes, A., Domkin, V., Ippel, J. H., Schleucher, J., Wijmenga, S., Thelander, L., and Rothstein, R. (2000) Mutational and structural analyses of the ribonucleotide reductase inhibitor Sml1 define its Rnr1 interaction domain whose inactivation allows suppression of mec1 and rad53 lethality. *Mol. Cell. Biol. 20*, 9076–9083.
13. Gupta, V., Peterson, C. B., Dice, L. T., Uchiki, T., Racca, J., Guo, J.-T., Ying, X., Hettich, R., Zhao, X., Rothstein, R., and Dealwis, C. G. (2004) Sml1p is a dimer in solution: Characterization of denaturation and renaturation of recombinant Sml1p. *Biochemistry 43*, 8568–8578.
14. Dyson, H. J., and Wright, P. E. (2002) Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Adv. Protein Chem. 62*, 311–340.
15. Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR. *Chem. Rev. 104*, 3607–3622.
16. Uversky, V. N. (2002) What does it mean to be natively unfolded? *Eur. J. Biochem. 269*, 2–12.
17. Vendruscolo, M., and Dobson, C. M. (2005) A glimpse at the organization of the protein universe. *Proc. Natl. Acad. Sci. U.S.A. 102*, 5641–5642.
18. Frieden, C., Chattopadhyay, K., and Elson, E. L. (2002) What fluorescence correlation spectroscopy can tell us about unfolded proteins. *Adv. Protein Chem. 62*, 91–109.
19. Keiderling, T. A., and Xu, Q. (2002) Unfolded peptides and proteins studied with infrared absorption and vibrational circular dichroism spectra. *Adv. Protein Chem. 62*, 111–161.
20. Sharp, J. S., Guo, J.-T., Uchiki, T., Xu, Y., Dealwis, C., and Hettich, R. (2005) Application of photochemical surface mapping of C16S Sml1p to constrained computational modeling. *Anal. Biochem. 340*, 201–212.
21. Kjær, M., Andersen, K. V., and Poulsen, F. M. (1994) Automated and semiautomated analysis of homo- and heteronuclear multidimensional nuclear magnetic resonance spectra of proteins: The program Pronto. *Methods Enzymol. 239*, 288–307.
22. Gibbs, S. J., and Johnson, C. S. (1991) A PFG NMR experiment for accurate diffusion and flow studies in the presence of eddy currents. *J. Magn. Reson. 93*, 395–402.
23. Yao, S., Howlett, G. J., and Norton, R. S. (2000) Peptide self-association in aqueous trifluoroethanol monitored by pulsed field gradient NMR diffusion measurements. *J. Biomol. NMR 16*, 109–119.
24. Stejskal, E. O., and Tanner, J. E. (1965) Spin diffusion measurements: Spin echoes in the presence of a time dependent field gradient. *J. Chem. Phys. 42*, 288–292.
25. Damberg, P., Jarvet, J., and Gräslund, A. (2001) Accurate measurement of translational diffusion coefficients: A practical method to account for nonlinear gradients. *J. Magn. Reson. 148*, 343–348.
26. Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin sabelling. *J. Mol. Biol. 324*, 349–357.
27. Danielsson, J., Jarvet, J., Damberg, P., and Gräslund, A. (2002) Translational diffusion measured by PFG-NMR on full length and fragments of the Alzheimer Aβ(1−40) peptide. Determination of hydrodynamic radii of random coil peptides of varying length. *Magn. Reson. Chem. 40*, S89−S97.
28. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry 38*, 16424–16431.
29. Berjanskii, M. V., and Wishart, D. S. (2007) The RCI server: Rapid and accurate calculation of protein flexibility using chemical shifts. *Nucleic Acids Res. 35*, W531−W537.

30. Berjanskii, M. V., and Wishart, D. S. (2008) Application of the random coil index to studying protein flexibility. *J. Biomol. NMR 40*, 31–48.
31. Schwarzinger, S., Kroon, G. J. A., Foss, T. R., Chung, J., Wright, P. E., and Dyson, H. J. (2001) Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc. 123*, 2970–2978.
32. Wishart, D. S., and Sykes, B. D. (1994) Chemical shifts as a tool for structure determination. *Methods Enzymol. 239*, 363–392.
33. Modig, K., Juergensen, V. W., Lindorff-Larsen, K., Fieber, W., Bohr, H. G., and Poulsen, F. M. (2007) Detection of initiation sites in protein folding of the four helix bundle ACBP by chemical shift analysis. *FEBS Lett. 581*, 4965–4971.
34. Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. *Protein Sci.,* 2795–2804.
35. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc. 104*, 4546–4559.
36. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc. 104*, 4559–4570.
37. Lindorff-Larsen, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C. M., Poulsen, F. M., and Vendruscolo, M. (2004) Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme A binding protein. *J. Am. Chem. Soc. 126*, 3291–3299.
38. Fieber, W., Kragelund, B. B., Meldal, M., and Poulsen, F. M. (2005) Reversible dimerization of acid-denatured ACBP controlled by helix A4. *Biochemistry 44*, 1375–1384.
39. Song, J., Guo, L.-W., Muradov, H., Artemyev, N. O., Ruoho, A. E., and Markley, J. L. (2008) Intrinsically disordered γ-subunit of cGMP phosphodiesterase encodes functionally relevant transient secondary and tertiary structure. *Proc. Natl. Acad. Sci. U.S.A. 105*, 1505–1510.
40. Simon, S. M., Sousa, F. J. R., Mohana-Borges, R., and Walker, G. C. (2008) Regulation of *Escherichia coli* SOS mutagenesis by dimeric intrinsically disordered *umuD* gene products. *Proc. Natl. Acad. Sci. U.S.A. 105*, 1152–1157.
41. Zhang, H., Neal, S., and Wishart, D. S. (2003) RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR 25*, 173–195.

BI801040B