BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Cooperative Unfolding of Compact Conformations of the Intrinsically Disordered Protein Osteopontin

Dennis Kurzbach,[†] Gerald Platzer,[‡] Thomas C. Schwarz,[‡] Morkos A. Henen,[‡,§] Robert Konrat,*[,‡] and Dariush Hinderberger*[,†]

[†]Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

[‡]Max F. Perutz Laboratories, Dr. Bohr-Gasse 9 (VBC 5), 1030 Vienna, Austria

[§]Faculty of Pharmacy, Mansoura University, Mansoura, Egypt

Supporting Information

ABSTRACT: Intrinsically disordered proteins (IDPs) constitute a class of biologically active proteins that lack defined tertiary and often secondary structure. The IDP Osteopontin (OPN), a cytokine involved in metastasis of several types of cancer, is shown to simultaneously sample extended, random coil-like conformations and stable, cooperatively folded conformations. By a combination of two magnetic resonance methods, electron paramagnetic resonance and nuclear magnetic resonance spectroscopy, we demonstrate that the OPN ensemble exhibits not only characteristics of an extended and flexible polypeptide, as expected for an IDP, but also simultaneously those of globular proteins, in particular sigmoidal structural denaturation profiles. Both types of states, extended and cooperatively folded, are populated simultaneously by OPN in its apo state. The heterogeneity of the structural properties of IDPs is thus shown to even involve cooperative folding and unfolding events.

Intrinsically disordered proteins (IDPs) have revolutionized structural biology in recent years. Despite a lack of well-folded, crystallizable structure in the conventional sense, they fulfill essential functions in eukaryotic life and thus challenge the traditional structure−function paradigm.[1,2] The flexible and dynamic structure of IDPs and their ability to adopt different functional structures (e.g., folding upon binding) yet allow for multiple interactions of a particular protein with several binding partners.[3,4] This makes IDPs intriguing substrates for studies in modern proteomics. Also from a biophysical and structural biology point of view, IDPs remain puzzling in many aspects. Because of the limited number of experimental techniques suited for investigations of IDPs, their solution states, conformational space, and modes of conformational sampling are not well understood.[2,5,6] However, there is a growing body of evidence that these proteins commonly classified as disordered or unstructured should be conceived as "ensembles of a continuum of rapidly interconverting structures"[7] that contain a heterogeneous assembly of conformations, ranging from random coils to compact structures that have regions that have stronger tendencies toward secondary structures.[8−10] Often, preformed local secondary structure elements comprise epitopes for biologically relevant protein interactions.[11−14] Although these partially preformed elements typically undergo folding-upon-binding events resulting in stable structural arrangements of separated interaction elements, no distinct tertiary structure is observed in the apo state.

Motivated by the fact that in the past several years intrinsically local magnetic resonance techniques, especially nuclear magnetic resonance (NMR), have led to intriguing insight into conformational and dynamic properties of IDPs,[15] we here apply solution-state NMR in combination with frozen-state electron paramagnetic resonance (EPR) spectroscopy to an IDP. We aim to combine data gained from a dynamic system state (NMR) with data about a static snapshot of the system (EPR) to gain a detailed picture of structural transition events that are potentially comprised in the conformational space of an IDP. We have chosen Osteopontin (OPN) as a model compound, because earlier studies have shown that this IDP exhibits interesting structural properties like preformed ligand binding sites and a varying compactness profile along the disordered protein backbone. From a biological point of view, OPN is a cytokine involved in metastasis of several kinds of cancer (see ref 16 for a biophysical characterization of OPN).[4,16] Here we show through the combination of EPR and NMR that OPN is also interesting from a biophysical point of view. The compound samples a broad distribution of compact and expanded conformations as expected for an IDP, and the conformational sampling also comprises cooperative folding and unfolding events. Cooperative folding is well-documented in classical proteomics, where transitions between random coil and globular states with distinct long-range interactions are typically described as first-order processes. These transitions are in most cases sigmoidal in nature,[17] and different conformations constitute energetically different

Received: April 22, 2013
Revised: July 10, 2013
Published: July 12, 2013

   dx.doi.org/10.1021/bi400502c | *Biochemistry* 2013, 52, 5167−5175

thermodynamic states.[18] The unexpected finding reported here requires that the classification of IDPs in terms of rapidly interconverting structures has to be augmented by simultaneous conformational sampling of extended as well as cooperatively folded conformations, i.e., with the fact that in IDPs different conformations of a single protein may interconvert via cooperative (phase) transitions.

Initial studies combining NMR and EPR for partially unfolded proteins have already been published.[19] The complementary combination of EPR with NMR spectroscopy applied here leads to coarse-grained information about the conformational states of the disordered OPN (EPR) as well as detailed information about its folded conformations (NMR). This combined magnetic resonance methodology and data interpretation may be applicable to other disordered protein systems.[20−24]

EXPERIMENTAL PROCEDURES

Protein Preparation. The expression and purification of recombinant quail OPN protein (OPN220) mutants were performed as described previously.[16] All details of protein expression, purification, and spin labeling are given in ref 16. Essentially, cysteine mutations were introduced using the QuickChangeII site-directed mutagenesis kit (Stratagene). For NMR and EPR analysis, all protein samples were concentrated to 0.8 mM in phosphate buffer [50 mM sodium phosphate and 50 mM NaCl (pH 6.5)] in a 90% H_2O/D_2O mixture. EPR double mutants were tagged with the nitroxide spin label (1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL), in a process analogous to the labeling procedure described in ref 16. For the purpose of this work, the choice of the label is, however, not crucially important, because the increase in ring rigidity one would gain by changing to PROXYL is negligible. The protein mutants were subject to rigorous purification using PD-10 desalting columns to remove all excess spin labels. The labeling efficiency was determined by means of 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) and UV−vis absorbance. The labeling efficiency was always >95%.

Experimental Double Electron−Electron Resonance (DEER). DEER is applied to glassy solids obtained by freeze quenching the solutions after addition of 30% (v/v) glycerol. [Note that the presence of 20−30% (v/v) glycerol in buffered solutions of proteins is known to not affect protein conformations significantly.[44]] This is achieved by immersing the sample tube in supercooled isopentane. In this way, a snapshot representative of the solution at the glass transition temperature is detected. Prior to being freeze-quenched, the samples were transferred to 3 mm outer diameter quartz tubes. The sample volume was approximately 100 μL and always large enough to fill the complete sensitive volume in the resonator. The four-pulse DEER sequence $\pi/2(\nu_{obs})-\tau_1-\pi(\nu_{obs})-(\tau_1 + t)-\pi(\nu_{pump})-(\tau_2 - t)-\pi(\nu_{obs})-\tau_2-$echo was used to obtain dipolar time evolution data at X-band frequencies (9.2−9.4 GHz) with a Bruker Elexsys 580 spectrometer equipped with a Bruker Flexline split-ring resonator (ER4118X_MS3). The dipolar evolution time t was varied, whereas τ_2 (3 μs) and τ_1 were kept constant. Proton modulation was averaged by the addition of eight time traces of variable τ_1 values, starting with a $\tau_{1,0}$ of 200 ns and using increments ($\Delta\tau_1$) of 8 ns. The resonator was overcoupled to $Q \approx 100$. The pump frequency, ν_{pump}, was set to the maximum of the EPR spectrum. The observer frequency, ν_{obs}, was set to ν_{pump} + 61.6 MHz, coinciding with the low-field local maximum of the nitroxide spectrum. The observer pulse lengths were 32 ns for both $\pi/2$ and π pulses, and the pump pulse length was 12 ns. The temperature was set to 50 K by cooling with a closed cycle cryostat (ARS AF204, customized for pulse EPR, ARS, Macungie, PA). The total measurement time for each sample was ∼12 h. The resulting time traces were normalized to $t = 0$.

DEER Background Correction. Background correction was conducted by dividing by experimental functions gained by DEER on four spin labeled single mutants (C54, C108, C188, C427). All single-mutant data corresponded to a homogeneous three-dimensional distribution of spins (homogeneous exponential decay). Measurements on single mutants were taken at similar protein concentrations as in the case of the double mutants (0.8 mM) to ensure a similar excluded volume. Because the spin concentration was therefore only half the concentration as in the case of double mutants, this adds a factor of 2 to the exponent of the homogeneous background function to compensate for the lower concentration:

$$V_{DEER}^{hom} = \exp\left(-\frac{2\pi\mu_0 g_1 g_2 \beta_e}{9\sqrt{3}\hbar}\lambda c_{pump} t^{d/3}\right) \quad (1)$$

where c_{pump} is the concentration of the spins resonant at frequency ν_{pump} and $d = 3$. g_1 and g_2 are the g values of the observer and the pump spin, respectively, in a two-spin system. μ_0 is the magnetic moment of the electron spin. β_e is the Bohr Magneton. λ is an experimental constant that denotes the fraction of spins flipped by the pump pulse ($\hbar = h/2\pi$). The concentration factor of 2 can, however, be neglected (as we did for background correction) for the relative comparison used for analysis of the data presented here.

Determination of Δ_{eff}. Δ_{eff} was determined by fitting the background-corrected DEER data with a Gaussian distance distribution according to eqs S3−S5 of the Supporting Information. As such, a $V_R(t)$ was determined from the fit and $\Delta_{eff} = 1 - V_R(t = 3\ \mu s)$.

Experimental NMR Spectroscopy. NMR spectra were recorded at 20 °C on Varian spectrometers operating at 500, 600, and 800 MHz. OPN220 protein samples were dissolved in 50 mM sodium phosphate and 50 mM NaCl (pH 6.5) with 10% D_2O as the lock solvent. PRE intensity ratios were derived from pulsed field gradient (PFG) sensitivity-enhanced two-dimensional $^{15}N-^{1}H$ HSQC spectra of ^{15}N-labeled OPN220 mutants C54, C108, C188, and C247.[45] NMR spectra were processed using NMRPipe[46] and analyzed using SPARKY.

Paramagnetic Relaxation Enhancements (PREs). Single-cysteine OPN220 mutants (C54, C108, C188, and C247) were tagged with the nitroxide spin label (1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL). The overall PRE effect on OPN220 was measured as the intensity ratio of cross-peaks in presence (I_{+MTLS}) and absence (I_{-MTSL}) of the cysteine-attached spin labels, as $\Delta_{MTSL} = I_{+MTLS}/I_{-MTSL}$. To account for possible interactions of the spin label with the protein, we added MTSL to untagged, ^{15}N-labeled OPN220 at a final concentration of 1 mM. The intensity ratios (peak intensity) of OPN220 with unbound and free ($I_{+FREEMTSL}$) and without (I_{-MTSL}) MTSL ($\Delta_{UNSPEC} = I_{+FREEMTSL}/I_{-MTSL}$) were combined with the intensity ratios from attached MTSL (Δ_{MTSL}) to calculate the PRE effect on the protein: PRE = $\Delta_{MTSL} + (1 - \Delta_{UNSPEC})$. ΔPRE was calculated as the difference in signal heights under nondenaturing conditions and high-urea and -salt conditions as ΔPRE = $^{15}N-^{1}H$ HSQC intensity (high urea and salt) − $^{15}N-^{1}H$ HSQC intensity (no urea or salt).

dx.doi.org/10.1021/bi400502c | Biochemistry 2013, 52, 5167−5175

The usage of two different samples in the presence and absence of free MTSL is a necessary to ensure that MTSL itself does not have an intrinsic binding affinity for certain preferential protein sites. This cannot be probed by reduction of covalently attached MTSL, because the native protein conformations could not be separated from MTSL-induced folding in this case.

It should be noted that significant spectral overlap at 8 M urea and high NaCl concentrations precluded the extraction of a complete PRE data set. Additionally, PRE measurements under complete denaturation conditions using both high urea and NaCl concentrations were not possible because of substantial viscosity effects (and thus substantially broadened NMR signals).

$^{13}C-^{1}H$ HSQC NMR. Reductive methylation procedures were performed as described by Means and Feeney.[47] The protein was dialyzed against 10 mM HEPES, 150 mM NaCl, and 1 mM DTT buffer (pH 7.4); 0.25 mL of 1.6 mM borane dimethylamine complex $[(CH_3)_2NH{\cdot}BH_3]$ and 0.5 mL of 1.6 mM ^{13}C-labeled formaldehyde were added to 0.5 mL of 0.1 mM protein, and the reaction mixture was incubated while being stirred at 4 °C. Subsequently, the addition of the borane ammonium complex and [^{13}C]formaldehyde was repeated, and the reaction mixture was incubated for an additional 2 h. After 0.12 mL of a 1.6 mM borane ammonium complex solution had been added, the reaction mixture was incubated at 4 °C while being stirred overnight. The reaction was quenched by adding glycine to yield a concentration of 200 mM. Undesired reaction products as well as excess reagents were removed by dialysis against Tris buffer (pH 7.4). The sample was concentrated to a final concentration of approximately 0.1 mM. Two-dimensional $^{13}C-^{1}H$ HSQC NMR experiments were conducted with synthesized ^{13}C-methylated Osteopontin on a Varian Innova 600 MHz spectrometer.

RESULTS AND DISCUSSION

We first probe structural preferences of OPN by applying the EPR-based method double electron−electron resonance (DEER) spectroscopy to six spin labeled Cys double mutants of 220-residue truncated OPN [residues 45−264 of the native protein; we denote the spin labeled OPN double mutants as C*x*−C*y* (*n*/3 E, L, or S) with *x* and *y* being the labeling sites, *n*/3 the fraction spanned by the respective mutant, and E, L, or S the shape (exponential, linear, or sigmoidal, respectively) of the denaturation profile as will be explained below]. Mutants C54−C108 (1/3 L), C108−C188 (1/3 L), and C188−C247 (1/3 E) each span approximately one-third of the whole protein; mutants C54−C188 (2/3 S) and C108−C247 (2/3 L) each span approximately two-thirds of the protein, and mutant C54−C247 (3/3 S) spans nearly the whole truncation mutant (see Figure 1 for a schematic representation of the spanned ranges). Conformational stabilities, understood here as resistance to urea unfolding, of these individual structural segments of OPN are investigated by recording DEER time traces for the different double mutants that are dependent on the urea concentration. In a second step, we investigate cooperatively compacted states by means of paramagnetic relaxation enhancements and rationalize our results on the basis of noncovalent structuring principles in OPN.

Cooperative Transition Events. DEER experiments yield nonaveraged data [i.e., the superposition of data from every single OPN molecule in the shock-frozen solution (see Experimental Double Electron−Electron Resonance (DEER) and section 1 and Figure S1 of the Supporting Information for a detailed description of DEER)] that, after background correction that eliminates intermolecular contributions, display intramolecular dipole−dipole couplings between the two electrons of the spin labels of double mutants as damped cosine modulation of time domain traces. These traces are the evolution of echo intensity with interpulse delay for the four-pulse DEER sequence as described in the Experimental section.[22]

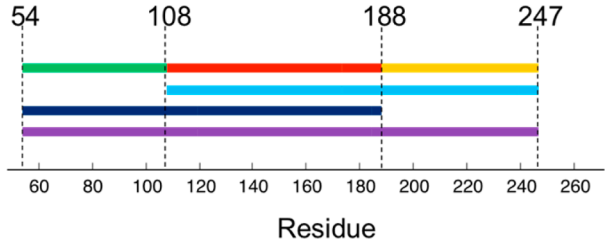

Figure 1. Scheme of spin labeling sites along the protein backbone and sketch of residues spanned by each of the six OPN double mutants.

In particular, the modulation is related to the dipolar coupling frequency that in turn depends on the interspin distance, *R*, as R^{-3}.[22] Typical DEER data for an OPN double mutant at different urea concentrations are shown in Figure 2

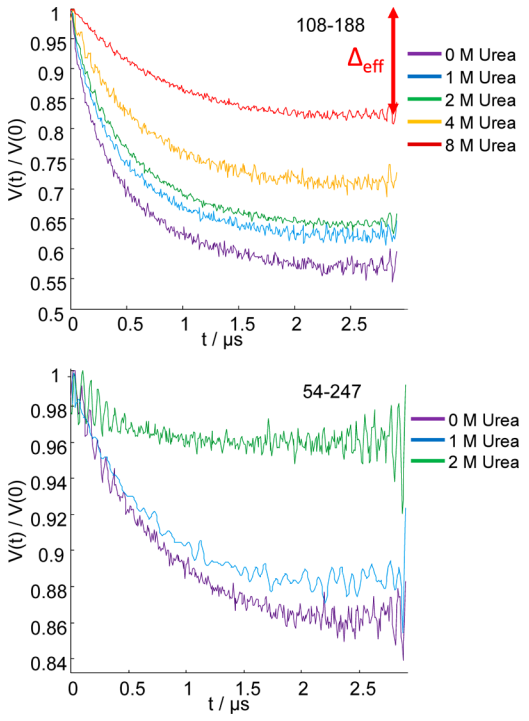

Figure 2. DEER time traces of C108−C188 (1/3 L) and C54−C247 (3/3 S) at different urea concentrations. Δ_{eff} is defined as the signal decay at 3 µs, as indicated by the double-headed arrow [note the different $V(t)/V(0)$ scales].

for two different double mutants [C108−C188 (1/3 L), representative of a double mutant defining a small OPN segment, and C54−C247 (3/3 S), representative of a large segment]. No clear-cut modulations can be observed after experimental background correction. This indicates that the pair distribution functions, $P(R)$, between the two spin labels of the two double mutants (*R* being the distance between the two

dx.doi.org/10.1021/bi400502c | *Biochemistry* 2013, 52, 5167−5175

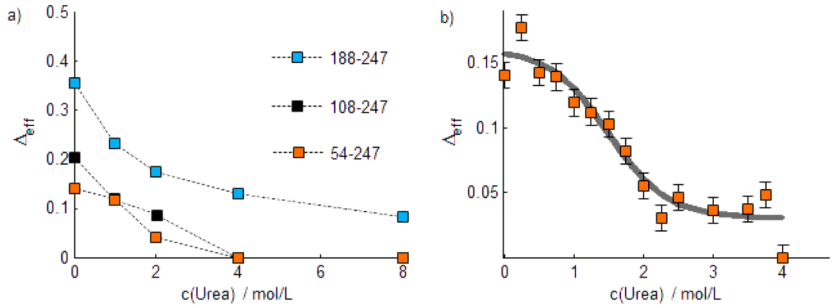

Figure 3. (a) Δ_{eff} for selected double mutants as a function of urea concentration. All the data for all double mutants under investigation can be found in Figures S2–S4 of the Supporting Information. (b) Detailed representation of Δ_{eff} for C54–C247 (3/3 S) as a function of urea concentration. Error bars stem from signal noise. The gray curve is based on a sigmoidal data fit to confirm the visual observation of sigmoidality. The fit is based on the relationship $\Delta_{eff} = a + b/\{1 + \exp[-c(\text{urea}) - m]/s\}$.

spins) are quite broad, because the sum over varying damped cosines converges to exponential decay functions. This is expected for an IDP with very broad conformational ensembles like OPN,[16] because every conformation (corresponding to a certain R) gives rise to a certain modulation function (the complete data set for all double mutants is shown in Figures S2 and S3 of the Supporting Information; all time traces are devoid of modulations).[25] In contrast, globular proteins frequently display DEER time traces with significant and apparent dipolar modulations because their narrow conformational ensembles give rise to a restricted interspin distance range (see Figure S1b of the Supporting Information for time traces calculated from a single discrete interspin distance).[21] Because the established standard analysis methods[26] cannot be applied for the nonmodulated DEER data under investigation, we analyze DEER time traces through an effective modulation depth, Δ_{eff} (as sketched in Figure 2 and Figure S1b of the Supporting Information), which denotes the total signal decay at a t_{max} of 3 μs. As such, $\Delta_{eff} = 1 - V(t = 3\ \mu s)/V(t = 0)$. $V(t)$ is the DEER echo intensity at time t [for details on the determination of Δ_{eff}, see Experimental Double Electron–Electron Resonance (DEER)]. Three microseconds is with our setup the longest achievable experimental DEER evolution time for this study but is generally arbitrary for the proposed analysis of very broad distance distributions reflected in large conformational ensembles. Δ_{eff} is an approximate measure of the average interspin distance for broad $P(R)$ values. For broad distance distributions Δ_{eff} decreases with an increasing interspin distance R. As such, a decrease in Δ_{eff} with an increasing urea concentration is representative for unfolding and expansion of a doubly spin labeled protein of interest. This is shown and explained in detail in section 1 of the Supporting Information and graphically illustrated in Figure S1b,c of the Supporting Information for calculated data.

In Figure 3a, experimental Δ_{eff} values are shown as a function of urea concentration for selected double mutants comprising segments of OPN of different lengths (i.e., starting from the C-terminus approximately one-third, two-thirds, and the entire truncation mutant; for the entire data set, see Figures S2–S4 of the Supporting Information). For the C-terminal part of OPN, comprised by the mutant C188–C247 (1/3 E), an exponential decay of Δ_{eff} (i.e., an increase in interspin distance) can be observed with an increasing urea concentration. This mutant gives rise to the steepest observed slope of any of the Δ_{eff} functions and can hence be regarded as a relative reference for the effect of conformational denaturation on unstably folded protein segments of potentially random coil-like character. Already for low urea concentrations, such segments show significant conformational expansion (i.e., a decrease in Δ_{eff}) in accordance with the idea of very low stability of transient or residual structural elements in IDPs. For mutant C108–C247 (2/3 L) [as well as for mutants C54–C108 (1/3 L) and C108–C188 (1/3 L) (see Figure S4 of the Supporting Information)], one observes an approximately linear decrease in Δ_{eff} with urea concentration, indicating that the OPN segment framed by this mutant (approximately two-thirds of the protein) is on average conformationally more stable than the segment between C188 and C247 [C188–C247 (1/3 E)], although still largely unstructured, random coil-like or (pre)molten globule-like. Strikingly, however, for mutant C54–C247 (3/3 S) of OPN and C54–C188 (2/3 S), we observe a sigmoidal development of the Δ_{eff}-derived denaturation profiles with urea concentration (see Figure 3b). Sigmoidality is a hallmark of cooperative folding of protein conformations and unexpected for an IDP.[17] The sigmoidal development of Δ_{eff} is depicted in more detail for C54–C247 (3/3 S) in Figure 3b. In Figure 4,

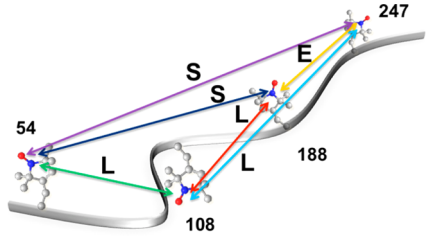

Figure 4. Sketch of OPN double mutants assessed by DEER. Labels E (exponential), L (linear), and S (sigmoidal) denote the profile shapes of the respective Δ_{eff} functions (see Figure 3 and Figures S2 and S4 of the Supporting Information).

the six spin label pairs probed through the six double mutants of OPN are sketched and labeled with the shape of the respective denaturation or Δ_{eff} profile, i.e., linear (L), exponential (E), or sigmoidal (S).

A sigmoidal denaturation profile is indicative of stably and cooperatively folded tertiary structures of OPN, because for low urea concentrations of ≤0.75 M the whole protein does not expand significantly (as seen in a nearly constant Δ_{eff}). This observation of a cooperatively folded conformation is surprising as $P(R)$ values for OPN are generally quite broad, which can be deduced from prior studies concerning OPN's conformational space[16] and is reflected in the nonmodulated DEER time traces (see section 1 and Figures S2–S4 of the Supporting Information and Figure 2). This interesting finding can,

dx.doi.org/10.1021/bi400502c | Biochemistry 2013, 52, 5167–5175

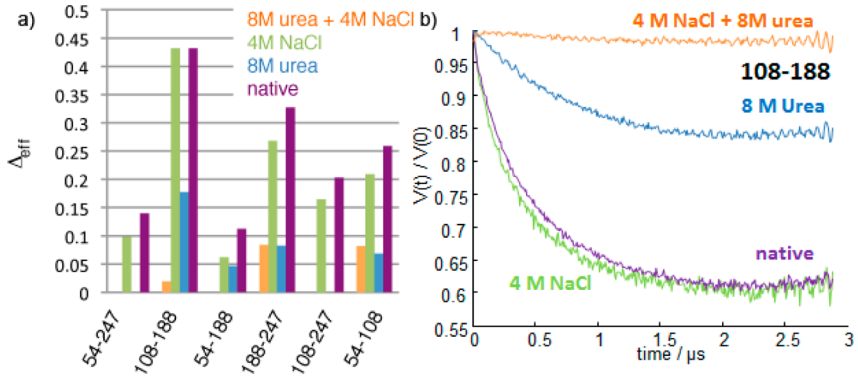

Figure 5. (a) Δ_{eff} values for the different double mutants under different denaturing conditions (4 M NaCl, 8 M urea, or 4 M NaCl with 8 M urea). (b) Exemplary (for C108−C188 (1/3 L)) DEER time traces for different denaturation conditions (4 M NaCl, 8 M urea, or 4 M NaCl with 8 M urea).

however, be understood by concluding that the structural ensemble of OPN contains both cooperatively folded and unfolded conformations and that both contribute to the DEER signals. It should be noted that the interpretation of DEER data is complicated by the fact that for rather small separations of spin label sites both compact and extended (sub)structures contribute significantly to the observed Δ_{eff} values. This means that compact conformations contribute more strongly to DEER time domain data of systems with distant spin labels (e.g., C54−C247 (3/3 S)). In contrast, if the spin labels come closer along the primary sequence and the mean distance becomes shorter (e.g., C188−C247 (1/3 L)), longer distances more significantly dominate the DEER data. This is shown in detail in section 2 of the Supporting Information. Because cooperatively folded conformations are more compact than unfolded ones, sigmoidal Δ_{eff} profiles can therefore only be observed for double mutants with labeling sites that are separated by more than 130 residues, because in these cases the mean distance of the conformational ensemble is large and hence the time domain data are dominated by contributions of folded conformations comprising short (electron−electron) distances. Hence, the corresponding Δ_{eff} profiles are sigmoidal. In contrast, for the three mutants that comprise only one-third of OPN, the DEER signal is dominated by contributions of extended structures. These exhibit linear or exponential denaturation profiles, lacking any sigmoidal contributions to the development of Δ_{eff}. In general, one can thus state for broad $P(R)$ values that with increasing separations between two labeling sites the relative contributions of compact conformations to the Δ_{eff} profiles increase. Compact conformations consequently dominate the urea dependence of Δ_{eff} for C54−C247 (3/3 S), while extended conformations contribute more significantly to Δ_{eff} for C188−C247 (1/3 E), C108−C188 (1/3 L), and C54−C108 (1/3 L). In summary, it is important to note that only the simultaneous sampling of both extended and compact (cooperatively folded) substates in OPN leads to superposition of DEER data with a significantly different urea dependence: a sigmoidal Δ_{eff} profile for mutant C54−C247 (3/3 S) and only gradual changes for double mutants spanning segments smaller than this mutant. Partial structuring as an underlying reason for this observation can be ruled out. For C54−C247 (3/3 S) (nearly the whole length of the truncation mutant), cooperative unfolding can be observed, while this is not the case for the inner segments comprising only approximately one-third of OPN. The latter would, however, necessarily be the case if any segment of OPN would be statically, partially structured. This deduction is possible here only because EPR of freeze-quenched solutions elucidates the whole set of coexisting conformations; ensemble averaged data here would not allow one to discern between partial structuring and sampling of compact conformations. In summary, OPN's structural behavior comprises cooperative phase transition events between compact and expanded conformations.

Noncovalent Structuring Principles of OPN Conformations As Seen in DEER Data. Given the enrichment (compared to the whole proteome) of polar and charged amino acids in IDPs, one can expect the stabilization of cooperative folded structures of OPN (as a necessary consequence of the cooperative phase transition events) likely to be triggered by electrostatic interactions.[3] In Figure 5, the effect of 4 M NaCl on Δ_{eff} is shown for the six double mutants under investigation in the presence and absence of 8 M urea. Figure 5a shows Δ_{eff} values for denaturation conditions with 4 M NaCl, 8 M urea, or 4 M NaCl with 8 M urea. Figure 5b shows exemplary DEER time traces for double mutant C108−C188 under these conditions and its native state. Note that NaCl does not significantly affect the effective modulation depth in the absence of urea but does in its presence. For C108−C188 (1/3 L) and C54−C188 (2/3 S), Δ_{eff} decreases with an increasing NaCl concentration even at 8 M urea. As such, the interspin distance still increases because of the increasing NaCl concentration even if 8 M urea is already present in the solution. Thus, one can state that urea alone does not expand OPN as strongly as urea in combination with NaCl. NaCl alone has only small effects on Δ_{eff}. This can be rationalized as follows. NaCl screens electrostatic interactions, while urea does not.[27−29] Hence, screening of electrostatic interactions seems not to be enough to significantly expand OPN's conformations. Only complementary screening of hydrophobic interactions and hydrogen bonds by urea and screening of electrostatics through NaCl lead to the most effective expansion of OPN's conformations. Hence, one might speculate that urea alone is not sufficient to completely denature OPN and eliminate all residual structural elements from its conformations; only the combination of complementary screening agents might be sufficiently strong. This is remarkable because earlier biophysical characterizations of OPN undoubtedly classify this protein as intrinsically disordered.[16] The significant electrostatic contribution to the energetics of OPN's conformational sampling modes is discussed in more detail below, taking into account paramagnetic relaxation enhancement (PRE) data. Protein dimerization as a possible source of error is ruled out

dx.doi.org/10.1021/bi400502c | *Biochemistry* 2013, 52, 5167−5175

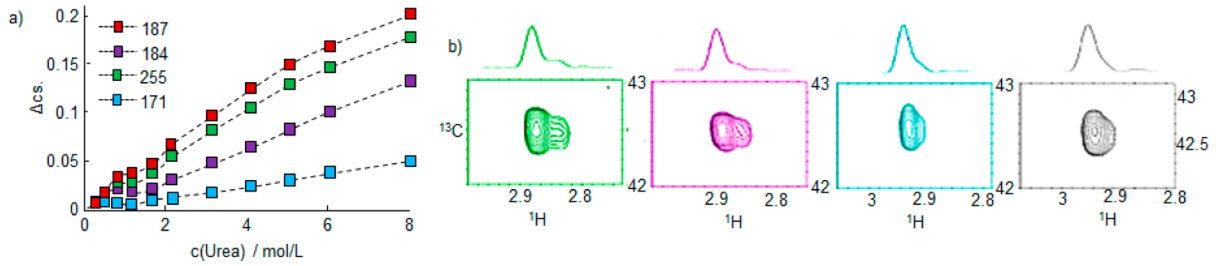

Figure 6. (a) $^{15}N-^{1}H$ NMR chemical shift changes {calculated as cs[c(urea) = 0 M] − cs[c(urea) = x M]} of selected backbone positions as a function of urea concentration. (b) $^{13}C-^{1}H$ HSQC of $^{13}CH_3$-Lys-labeled side chains: green for 0 M urea, pink for 2 M urea, blue for 4 M urea, and black for 6 M urea. Note that the shift in the ^{1}H dimension is merely a consequence of readjusting the transmitter offset in the dependence of the urea concentration to achieve suppression of water signals.

through DEER measurements performed on the four corresponding single mutants (see section 3 and Figure S7 of the Supporting Information). There, completely homogeneous spin distributions are observed, indicating that OPN does not show any form of aggregation at the concentrations (0.8 mM) used for the DEER (and NMR) measurements.

It should be noted that there is a growing body of evidence of the so-called direct mechanism of urea denaturation to describe protein−urea interaction correctly.[30,31] This mechanism states that urea directly binds to the protein backbone likely (primarily) through dispersive interactions and thereby interrupts protein structure-stabilizing interactions.[30,31] For this case of urea denaturation of OPN, this direct mechanism might be important for gaining a full understanding of the DEER data, because the Δ_{eff} values for some cases indicate mean distances between two labeling sites that are longer than the distance one would expect in a random coil polypeptide. For example, for the 54−108 mutants, one would expect distances between the two labels of ∼10 nm from Δ_{eff} (see Figure S1c of the Supporting Information), while for a true random coil with a Flory characteristic ratio of 2, distances of ∼6 nm would be expected.[32] This discrepancy might be traced back to binding of urea to the protein backbone, which leads to longer persistence lengths or rather scaling exponents and thus also to longer inter-residue distances.[33]

Noncovalent Structuring Principles of OPN Conformations As Seen in NMR Data. The urea dependence of NMR backbone chemical shift $^{15}N-^{1}H$ (cs) data was analyzed for residues of the core region and of the terminal region of OPN (see Figure 6a, Experimental NMR Spectroscopy, and Figure S5 of the Supporting Information; the superposition of $^{15}N-^{1}H$ NMR spectra at different urea concentration is shown in Figure S8 of the Supporting Information). The data show only marginal chemical shift changes (Δcs) observed below 1 M for some residues in the compact core region (171 and to some degree 144), and Δcs increases substantially only with ≥2 M urea. For most residues in and outside the core region, a more or less steady increase in Δcs can be observed. The core regions are approximately located between residues 100 and 180 (see Figure S5 of the Supporting Information for more cs data for residues of the core segment of OPN).[15] Overall, larger chemical shift changes were observed for residues located in the compact core of OPN (100−180). Most importantly, slight deviations from the linear Δcs versus urea concentration behavior were observed for residues in the core segment (171 and 144). It should be noted that all conformational substates of OPN contribute to the observed chemical shift changes. Given the small population of the compact structure, only small contributions can be expected. Although the cooperative phase transition observed by means of EPR, indicating the existence of rather stable conformations of OPN that resist denaturation by lower urea concentrations, cannot generally be reproduced through Δcs, the chemical shift changes clearly provide additional evidence of a more compact segment in OPN located between residues 100 and 180.[4,16] As such, the Δcs data are not in conflict with the interpretation derived above from DEER.

An additional indication of the existence of compact structures in the conformational ensemble was provided by NMR observations of side chain positions. In Figure 6b, data from $^{13}C-^{1}H$ HSQC [heteronuclear single-quantum coherence (see experimental $^{13}C-^{1}H$ HSQC NMR)] of $^{13}CH_3$-Lys-labeled OPN are shown. The majority of cross-peaks is overlapped and stems from side chains of residues in random coil-like conformations (signal at a higher number of parts per million of the ^{1}H dimension), which are typically more solvent-exposed and flexible than residues in folded protein segments. However, a fraction of methyl cross-peaks (approximately 20% as determined from fitting signal volumes; signal at a lower number of parts per million of the ^{1}H dimension) is significantly shifted from the bulk signals. This shows that a fraction of the lysine residues in the conformational ensemble are exposed to an environment that is different from that observed for random coil polypeptides. This might indicate that the conformational ensemble partially exists in a compacted form in which the lysine residues are embedded in a more water-depleted core. As such, the $^{13}C-^{1}H$ HSQC is not in conflict with the EPR-derived conclusion that part of the conformational ensemble of OPN cooperatively folds into compact conformations. The shifted lysine peaks even remain unchanged and clearly separated from the bulk of lysine side chains at urea concentrations of ≤1 M. We refrain here from analyzing this observation in the context of cooperativity, yet it further supports the existence of compact structures in the ensemble of OPN. The NMR $^{15}N-^{1}H$ chemical shift changes and $^{13}C-^{1}H$ HSQC data (that is, backbone-based as well as side chain-based data) are in agreement with a compact, presumably cooperatively folded substate in the conformational ensemble of OPN besides large fractions of extended conformations.

Paramagnetic Relaxation Enhancements. Because in a hypothetical random coil polymer paramagnetic relaxation enhancements (PRE; i.e., enhanced relaxation rates of nuclear resonances due to the presence of an electron spin) are limited to residues flanking the spin label sites, the observation of specific and sizable long-range PRE effects provides unambiguous evidence of the existence of compact states.[34,35] PRE

dx.doi.org/10.1021/bi400502c | Biochemistry 2013, 52, 5167−5175

effects were measured for the four single mutants C54, C108, C188, and C247. The different PRE-residue plots (see Figure 7a) show that the conformational ensemble of OPN indeed

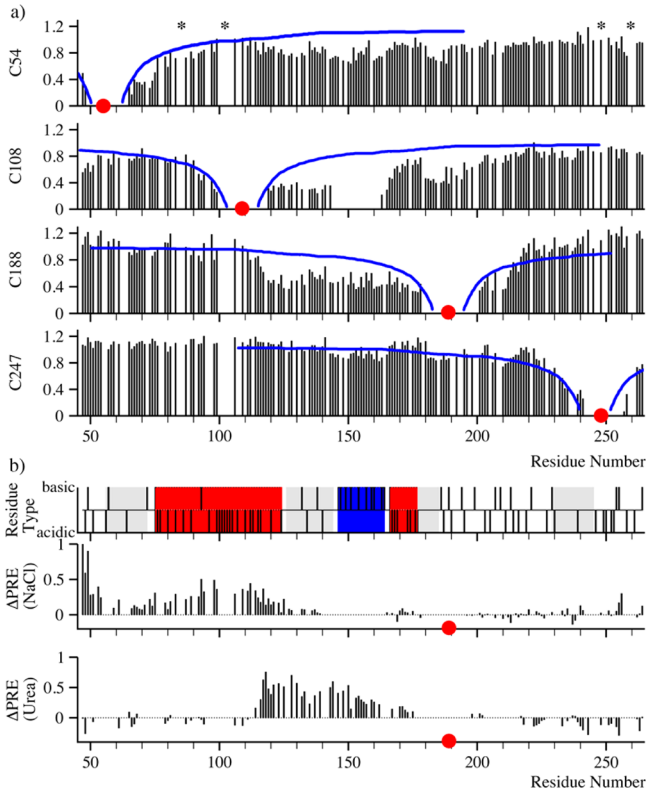

Figure 7. (a) PRE data for the four single mutants C54, C108, C188, and C247. Superimposed in blue are PREs calculated for random coils with a Flory characteristic ratio of 2 by the Solomon–Bloembergen relation.[23] The red dots mark the different labeling sites. The asterisks mark stretches comprising larger numbers of unassigned resonances. (b) Charge map of OPN (top; blue corresponds to patches of primarily basic residues, red to patches of acidic residues, and gray to primarily hydrophobic patches) and PRE changes (ΔPRE) for high-salt (center) and high-urea (bottom) conditions obtained for the C188 mutant [ΔPRE = ^{15}N–^{1}H HSQC intensity (high urea and salt) – ^{15}N–^{1}H HSQC intensity (no urea or salt)].

features distinct long-range interactions. Specifically, the PRE results obtained for mutants C108 and C188 provide clear evidence of the prevalence of a structurally compact region in OPN encompassing residues 100–200 (recall that intermolecular contributions can be ruled out by DEER on the single mutants at the given concentration). The structural stability of this compact conformation as a function of urea and NaCl concentration was monitored further by condition-dependent PRE changes. Figure 7b shows experimental PRE differences (ΔPRE) measured under NaCl and high-urea conditions.

The significantly charged region encompassing residues 75–125 is nearly unaffected by the addition of urea but displays sizable PRE changes under high-NaCl conditions, while residues 125–150 are strongly affected by urea. Hence, from these results and the EPR results (also compare with Figure 5), we can conclude that hydrophobic interactions contribute to the structural stability of OPN and electrostatics play a pivotal role in stabilizing the compact substates of OPN in solution. These findings can be rationalized by a closer inspection of the charge map of OPN (Figure 7b, top). In OPN, negative charges (acidic residues, red) are concentrated in the region between residues 75 and 125, while there is a high density of positive charges (basic residues, blue) in the region between residues 145 and 165. The attraction between these positively and negatively charged regions and the hydrophobic patches around residues 60, 130, and 180 therefore suggest stabilizing interactions and consequently stronger tertiary structure propensity between residues 60 and 180, compared to other regions of OPN.[16] In Figure 8, a sketch of OPN's compact

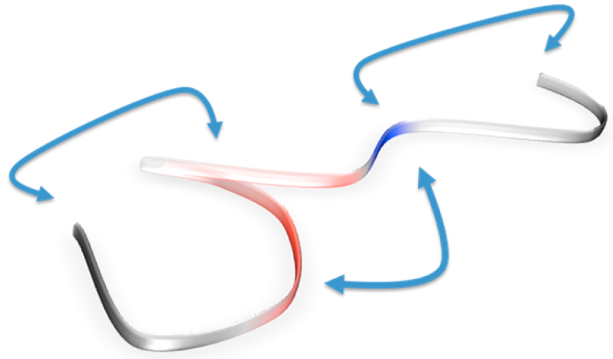

Figure 8. Sketch of the assumed "average" structure of OPN based on the PRE data. The arrows indicate significant PRE effects. As such, OPN can be pictured as having a more compact core and back-folded termini. The colors refer to the charge map in Figure 7b (blue corresponds to patches of primarily basic residues, red to patches of acidic residues, and gray to primarily hydrophobic patches).

comformation is shown, as one would derive it from the PRE data in Figure 7a. Long-range intrachain contacts between stretches between residues 100 and 180 as well as a slight sampling of the more central regions of OPN by the two termini are depicted. In conclusion, the NMR data indicate significantly populated compact structures in OPN that are stabilized (even cooperatively stabilized as evidenced by the DEER data above) by both hydrophobic and electrostatic interactions. This is also in excellent agreement with the NaCl dependence of Δ_{eff} (see Figure 5). The significant resistance to both urea and NaCl unfolding is clearly remarkable for an IDP. It is thus reasonable to conclude that the subtle interplay between conformation-stabilizing enthalpic contributions and destabilizing entropic contributions ultimately account for OPN's conformational flexibility and its ability to cooperatively sample both unfolded and cooperatively folded structures.

■ CONCLUSION

Altogether, we have shown that the IDP OPN cannot be described by polymer physical models such as random coil or molten globule polymers.[18] Instead, OPN simultaneously populates extended as well as cooperatively folded structures and sigmoidal molecular interconversion. This observation for OPN is a convincing experimental demonstration of conformational sampling of different thermodynamic states in an IDP.[36] The fact that OPN samples cooperatively stabilized as well as extended conformations further is particularly intriguing in the context of IDP binding mechanisms. Often protein recognition by IDPs proceeds via folding-upon-binding events accompanied by disorder-to-order transitions,[37] although even in the bound state IDPs (can) retain substantial conformational flexibilities ("fuzziness"),[38] be it static (multiple conformations) or dynamic disorder (fluctuation between different states).

dx.doi.org/10.1021/bi400502c | *Biochemistry* 2013, 52, 5167–5175

Protein−protein interaction is typically described either as an induced fit[39,40] or as a conformational selection fit.[40] While the induced fit model postulates the formation of an encounter complex followed by structural adaptation, conformational selection indicates the existence of a conformational ensemble in which the final bound state is partly present and populated by stabilizing intermolecular interactions, although there is evidence that both mechanisms can be active simultaneously.[41] The existence of a cooperatively folded substate in the structural ensemble of OPN suggests protein−protein interactions occur largely via conformational selection characterized by a significant reduction of the entropic penalty and presumably reduced fuzziness in the bound state.[37,42] The unexpected long-range preformation of the apo state of OPN might thus be of relevance for providing specific interaction interfaces across cellular surfaces and might thus endow OPN with unique abilities to modulate interaction patterns with its several natural ligands.[4]

Furthermore, our results substantiate recent insights that urea-unfolded states of proteins differ significantly from the native state of intrinsically disordered proteins.[43] We show that urea can induce drastic structural rearrangements of IDPs and changes in their conformational space. This is valid in terms of elongation of end-to-end distances of random coils through coordination of urea to the protein backbone, as stated above, and secondary structure propensities become significantly altered and populations of compact substates change when urea interacts with IDP backbones directly.[30,31] Most importantly, the existence of structural cooperative transitions from folded to unfolded states and vice versa in IDPs calls for a novel conceptual view of IDPs that goes beyond the traditional binary scheme of order versus disorder. The subtleties of heterogeneous conformational sampling in IDPs and their putative relevance for biological functions have to be adequately addressed.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary NMR and EPR data, experimental details, and details of sample preparation. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*D.H.: e-mail, dariush.hinderberger@mpip-mainz.mpg.de; phone, +49 6131 379 126. R.K.: e-mail, robert.konrat@univie.ac.at; phone, +43 1 4277 52202.

Author Contributions

D.K. and G.P. contributed equally to this work.

Funding

This work was financially supported by the Deutsche Forschungsgemeinschaft (DFG) via Grant HI 1094/2-1, the Max Planck Graduate Center with the University of Mainz (MPGC), and the Austrian Science Fund (FWF) via Grants P20549-N19, P22125-B12, and W1221. T.C.S. was in part supported by Forschungsstipendium 2011. M.A.H. was supported by a grant from the ÖAD entitled "KU mit Ägypten art 9/2". D.K. acknowledges support by the Gutenberg Academy of the University of Mainz.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Christian Bauer for technical support and Prof. Hans W. Spiess for continuing support. M.A.H. acknowledges the support of Professor Dr. Hassan Eisa, Dr. Alaa El-Din Barghash, and Dr. Laila A. Abouzeid (Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Mansoura University). The expression plasmid for wild-type OPN was provided by the group of K. Bister (University of Innsbruck, Innsbruck, Austria).

■ REFERENCES

(1) Dunker, A. K., and Obradovic, Z. (2001) The protein trinity: Linking function and disorder. *Nat. Biotechnol. 19*, 805−806.

(2) Tompa, P. (2002) Intrinsically unstructured proteins. *Trends Biochem. Sci. 27*, 527−533.

(3) Uversky, V. N. (2011) Intrinsically disordered proteins from A to Z. *Int. J. Biochem. Cell Biol. 43*, 1090−1103.

(4) Bellahcène, A., Castronovo, V., Ogbureke, K. U. E., Fisher, L. W., and Fedarko, N. S. (2008) Small integrin-binding ligand N-linked glycoproteins (SIBLINGs): Multifunctional proteins in cancer. *Nat. Rev. 8*, 212−226.

(5) Konrat, R. (2010) The meandering of disordered proteins in conformational space. *Structure 18*, 416−149.

(6) Uversky, V. N., and Dunker, A. K. (2010) Understanding protein non-folding. *Biochim. Biophys. Acta 1804*, 1231−1264.

(7) Ozenne, V., Schneider, R., Yao, M., Huang, J.-r., Salmon, L., Zweckstetter, M., Jensen, M. R., and Blackledge, M. (2012) Mapping the Potential Energy Landscape of Intrinsically Disordered Proteins at Amino Acid Resolution. *J. Am. Chem. Soc. 134*, 15138−15148.

(8) Huang, J.-r., and Grzesiek, S. (2010) Ensemble Calculations of Unstructured Proteins Constrained by RDC and PRE Data: A Case Study of Urea-Denatured Ubiquitin. *J. Am. Chem. Soc. 132*, 694−705.

(9) Mittag, T., Marsh, J., Grishaev, A., Orlicky, S., Lin, H., Sicheri, F., Tyers, M., and Forman-Kay, J. D. (2010) Structure/Function Implications in a Dynamic Complex of the Intrinsically Disordered Sic1 with Cdc4 Subunit of an SCF Ubiquitin Ligase. *Structure 18*, 494−506.

(10) Marsh, J. A., and Forman-Kay, J. D. (2009) Structure and Disorder in an Unfolded State under Nondenaturating Conditions from Ensemble Models Consistent with a Large Number of Experimental Restraints. *J. Mol. Biol. 391*, 359−374.

(11) Vacic, V., Oldfield, C. J., Mohan, A., Radivojac, P., Cortese, M. S., Uversky, V. N., and Dunker, A. K. (2007) Characterization of molecular recognition features, MoRFs, and their binding partners. *J. Proteome Res. 6*, 2351−2366.

(12) Fuxreiter, M., Simon, I., Friedrich, P., and Tompa, P. (2004) Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J. Mol. Biol. 338*, 1015−1026.

(13) Neduva, V., and Russell, R. B. (2005) Linear motifs: Evolutionary interaction switches. *FEBS Lett. 579*, 3342−3345.

(14) Hsu, W. L., Oldfield, C. J., Xue, B., Meng, J., Huang, F., Romero, P., Uversky, V. N., and Dunker, A. K. (2013) Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. *Protein Sci. 22*, 258−273.

(15) Jensen, M. R., Ruigrok, R. W. H., and Blackledge, M. (2013) Describing intrinsically disordered proteins at atomistic resolution by NMR. *Curr. Opin. Struct. Biol.*, doi: 10.1016/j.sbi.2013.1002.1007.

(16) Platzer, G., Schedbauer, A., Chemeli, A., Ozdowy, P., Coudevylle, N., Auer, R., Kontaxis, G., Hartl, M., Miles, A. J., Wallace, B. A., Glaser, O., Bister, K., and Konrat, R. (2011) The Metastasis-Associated Extracellular Matrix Protein Osteopontin Forms Transient Structure in Ligand Interaction Sites. *Biochemistry 50*, 6113−6124.

(17) Holde, K. E. v., Johnson, W. C., and Ho, P. S. (1998) *Principles of Physical Biochemistry*, Prentice Hall, Upper Saddle River, NJ.

(18) Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci. 11*, 739−756.

(19) Rao, J. N., Jao, C. C., Hegde, B. G., Langen, R., and Ulmer, T. S. (2010) A Combinatorial NMR and EPR Approach for Evaluating the Structural Ensemble of Partially Folded Proteins. *J. Am. Chem. Soc. 132*, 8657−8668.
(20) Schneider, R., Huang, J.-r., Yao, M., Communie, G., Ozenne, V., Mollica, L., Salmon, L., Jensen, M. R., and Blackledge, M. (2012) Towards a robust description of intrinsic protein disorder using nuclear magnetic resonance spectroscopy. *Mol. BioSyst. 8*, 58−68.
(21) Drescher, M. (2012) EPR in Protein Science. *Top. Curr. Chem. 321*, 91−120.
(22) Jeschke, G., and Polyhach, Y. (2007) Distance measurements on spin labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys. 9*, 1895−1910.
(23) Clore, G. M., Tang, C., and Iwahara, J. (2007) Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr. Opin. Struct. Biol. 17*, 603−616.
(24) Morin, B., Bourhis, J. M., Belle, V., Woudstra, M., Carriere, F., Guigliarelli, B., Fournel, A., and Longhi, S. (2006) Assessing induced folding of an intrinsically disordered protein by site-directed spin labeling electron paramagnetic resonance spectroscopy. *J. Phys. Chem. B 110*, 20596−20608.
(25) Schweiger, A., and Jeschke, G. (2001) *Principles of pulse electron paramagnetic resonance*, Oxford University Press, New York.
(26) Jeschke, G., Panek, G., Godt, A., Bender, A., and Paulsen, H. (2004) Data Analysis Procedures for Pulse ELDOR Measurements of Broad Distance Distributions. *Appl. Magn. Reson. 26*, 223−244.
(27) Tanford, C. (1964) Isothermal unfolding of globular proteins in aqueous urea solutions. *J. Am. Chem. Soc. 86*, 2050−2059.
(28) Sinha, N., and Smith-Gill, S. J. (2002) Electrostatics in Protein Binding and Function. *Curr. Protein Pept. Sci. 3*, 601−614.
(29) Avbelj, F., and Fele, L. (1998) Role of Main-chain Electrostatics, Hydrophobic Effect and Side-chain Conformational Entropy in Determining the Secondary Structure of Proteins. *J. Mol. Biol. 279*, 665−684.
(30) Candotti, M., Esteban-Martin, S., Salvatella, X., and Orozco, M. (2013) Toward an atomistic description of the urea-denatured state of proteins. *Proc. Natl. Acad. Sci. U.S.A. 110*, 5933−5938.
(31) Yang, Z. X., Xiu, P., Shi, B. Y., Hua, L., and Zhou, R. H. (2012) Coherent Microscopic Picture for Urea-Induced Denaturation of Proteins. *J. Phys. Chem. B 116*, 8856−8862.
(32) Rubinstein, M., and Colby, R. H. (2003) *Polymer Physics*, Oxford University Press, New York.
(33) Huang, J. R., Gabel, F., Jensen, M. R., Grzesiek, S., and Blackledge, M. (2012) Sequence-Specific Mapping of the Interaction between Urea and Unfolded Ubiquitin from Ensemble Analysis of NMR and Small Angle Scattering Data. *J. Am. Chem. Soc. 134*, 4429−4436.
(34) Yi, Q., Scalley-Kim, M. L., Alm, E. J., and Baker, D. (2000) NMR characterization of residual structure in the denatured state of protein L. *J. Mol. Biol. 299*, 1341−1351.
(35) Mittag, T., and Forman-Kay, J. D. (2007) Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol. 17*, 3−14.
(36) Uversky, V. N., Santambrogio, C., Brocca, S., and Grandori, R. (2012) Length-dependent compaction of intrinsically disordered proteins. *FEBS Lett. 586*, 70−73.
(37) Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol. 293*, 321−331.
(38) Tompa, P., and Fuxreiter, M. (2008) Fuzzy complexes: Polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci. 33*, 2−8.
(39) Koshland, D. E. (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci. U.S.A. 44*, 98−104.
(40) Boehr, D. D., Nussinov, R., and Wright, P. E. (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol. 5*, 789−796.
(41) Grunberg, R., Leckner, J., and Nilges, M. (2004) Complementarity of structure ensembles in protein-protein binding. *Structure 12*, 2125−2136.
(42) Tompa, P., and Fuxreiter, M. (2008) Fuzzy complexes: Polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci. 33*, 2−8.
(43) Hofmann, H., Soranno, A., Borgia, A., Gast, K., Nettels, D., and Schuler, B. (2012) Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U.S.A. 109*, 16155−16160.
(44) Junk, M. J. N., Spiess, H. W., and Hinderberger, D. (2010) The Distribution of Fatty Acids Reveals the Functional Structure of Human Serum Albumin. *Angew. Chem., Int. Ed. 49*, 8755−8759.
(45) Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry 33*, 5984−6003.
(46) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.
(47) Means, G. E., and Feeney, R. E. (1968) Reductive alkylation of amino groups in proteins. *Biochemistry 7*, 2192−2201.