



NMR characterization of solvent accessibility and transient structure in intrinsically disordered proteins

Christoph Hartlmüller¹ · Emil Spreitzer² · Christoph Göbl³ · Fabio Falsone⁴ · Tobias Madl^{2,5}

Received: 3 February 2019 / Accepted: 11 April 2019 / Published online: 11 July 2019
© The Author(s) 2019

Abstract

In order to understand the conformational behavior of intrinsically disordered proteins (IDPs) and their biological interaction networks, the detection of residual structure and long-range interactions is required. However, the large number of degrees of conformational freedom of disordered proteins require the integration of extensive sets of experimental data, which are difficult to obtain. Here, we provide a straightforward approach for the detection of residual structure and long-range interactions in IDPs under near-native conditions using solvent paramagnetic relaxation enhancement (sPRE). Our data indicate that for the general case of an unfolded chain, with a local flexibility described by the overwhelming majority of available combinations, sPREs of non-exchangeable protons can be accurately predicted through an ensemble-based fragment approach. We show for the disordered protein α -synuclein and disordered regions of the proteins FOXO4 and p53 that deviation from random coil behavior can be interpreted in terms of intrinsic propensity to populate local structure in interaction sites of these proteins and to adopt transient long-range structure. The presented modification-free approach promises to be applicable to study conformational dynamics of IDPs and other dynamic biomolecules in an integrative approach.

Keywords Solvent paramagnetic relaxation enhancement · Intrinsically disordered proteins · Residual structure · p53 · FOXO4 · α -Synuclein

Christoph Hartlmüller and Emil Spreitzer are the shared first authors.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10858-019-00248-2>) contains supplementary material, which is available to authorized users.

✉ Tobias Madl
tobias.madl@medunigraz.at

- ¹ Center for Integrated Protein Science Munich (CIPSM) at the Department of Chemistry, Technische Universität München, Lichtenbergstrasse 4, 87548 Garching, Germany
- ² Gottfried Schatz Research Center for Cell Signaling, Metabolism and Aging, Institute of Molecular Biology & Biochemistry, Medical University of Graz, Neue Stiftingtalstrasse 6, 8010 Graz, Austria
- ³ The Campbell Family Institute for Breast Cancer Research at Princess Margaret Cancer Centre, 610 University Avenue, Toronto, ON M5G 2M9, Canada
- ⁴ Institute of Pharmaceutical Sciences, University of Graz, Schubertstrasse 1, 8010 Graz, Austria
- ⁵ BioTechMed-Graz, Graz, Austria

Introduction

The well-established structure–function paradigm has been challenged by the discovery of intrinsically disordered proteins (IDPs) (Dyson and Wright 2005). It is suggested that about 40% of all proteins have disordered regions of 40 or more residues, with many proteins existing solely in the unfolded state (Tompa 2012; Romero et al. 1998). Although they lack stable secondary or tertiary structure elements, this large class of proteins plays a crucial role in various cellular processes (Theillet et al. 2014; Wright and Dyson 2015; van der Lee et al. 2014; Uversky et al. 2014). Disorder serves a biological role, where conformational heterogeneity granted by disordered regions enables proteins to exert diverse functions in response to various stimuli. Unlike structured proteins, which are essential for catalysis and transport, disordered proteins are crucial for regulation and signaling. Due to their intrinsic flexibility they can act as network hubs interacting with a wide range of biomolecules forming dynamic regulatory networks (Dyson and Wright 2005; Tompa 2012; Babu et al. 2011; Flock et al. 2014; Wright and Dyson 1999; Uversky 2011; Habchi et al.

2014). Given the plethora of potential interaction partners, it is not surprising that the interaction of IDPs with binding partners are often tightly regulated via an intricate ‘code’ of post-translational modifications, including phosphorylation, methylation, acetylation, and various others (Wright and Dyson 2015; Bah and Forman-Kay 2016). These proteins, and distortions in their interaction networks, for example by mutations and aberrant post-translational modifications (PTMs), are closely linked to a range of human diseases, including cancers, neurodegeneration, cardiovascular disorders and diabetes, they are currently considered difficult to study (Dyson and Wright 2005; Tompa 2012; Babu et al. 2011; Habchi et al. 2014; Metallo 2010; Uversky et al. 2008; Dyson and Wright 2004). Complications arise from the following factors: these proteins lack well-defined stable structure, they exist in a dynamic equilibrium of distinct conformational states, and the number of experimental techniques and observables renders IDP conformational characterization underdetermined (Mittag and Forman-Kay 2007; Eliezer 2009). Thus, an integration of new sets of experimental and analytical techniques are required to characterize the conformational behavior of IDPs.

Although IDPs are highly dynamic, they often contain transiently-folded regions, such as transiently populated secondary or tertiary structure, transient long-range interactions or transient aggregation (Marsh et al. 2007; Shortle and Ackerman 2001; Bernado et al. 2005; Mukrasch et al. 2007; Wells et al. 2008). These transiently-structured regions are of particular interest to study the biological function of IDPs as they can report on biologically-relevant interactions and encode biological function. Examples are aggregation, liquid–liquid phase separation, binding to folded co-factors, or modifying enzymes (Yuwen et al. 2018; Brady et al. 2017; Choy et al. 2012; Maji et al. 2009; Putker et al. 2013).

NMR spectroscopy is exceptionally well-suited to study IDPs, and in particular to detect transiently folded regions (Meier et al. 2008; Wright and Dyson 2009; Jensen et al. 2009). Several NMR observables provide atomic resolution, and ensemble-averaged information reporting on the conformational energy landscape sampled by each amino acid, including chemical shifts, residual dipolar couplings (RDCs), and paramagnetic relaxation enhancement (PRE) (Dyson and Wright 2004; Eliezer 2009; Marsh et al. 2007; Shortle and Ackerman 2001; Meier et al. 2008; Gobl et al. 2014; Gillespie and Shortle 1997; Clore et al. 2007; Huang et al. 2014; Ozenne et al. 2012; Clore and Iwahara 2009; Otting 2010; Hass and Ubbink 2014; Gobl et al. 2016). RDCs, and PREs, either alone or in combination have been used successfully in recent years to characterize the conformations and long-range interactions of IDPs (Bernado et al. 2005; Ozenne et al. 2012; Dedmon et al. 2005; Bertocini et al. 2005; Parigi et al. 2014; Rezaei-Ghaleh et al. 2018). However, both techniques rely on a modification of the IDP

of interest, either by external alignment media in case of RDCs or the covalent incorporation of paramagnetic tags in the case of PREs.

We and others have proposed applications of soluble paramagnetic agents to obtain structural information by NMR without any modifications of the molecules of interest (Gobl et al. 2014; Guttler et al. 2010; Hartlmuller et al. 2016; Hocking et al. 2013; Madl et al. 2009, 2011; Respondek et al. 2007; Zangger et al. 2009; Pintacuda and Otting 2002; Bernini et al. 2009; Wang et al. 2012; Sun et al. 2011; Gong et al. 2017; Gu et al. 2014; Hartlmuller et al. 2017). The addition of soluble paramagnetic compounds leads to a concentration-dependent and therefore tunable increase of relaxation rates, the so-called paramagnetic relaxation enhancement (here denoted as solvent PRE, sPRE; also known as co-solute PRE, Fig. 1a). This effect depends on the distance of the spins of interest (e.g. ^1H , ^{13}C) to the biomolecular surface. The nuclei on the surface are affected the strongest by the sPRE effect, and this approach has been shown to correlate well with biomolecular structure in the case of proteins and RNA (Madl et al. 2009; Pintacuda and Otting 2002; Bernini et al. 2009; Hartlmuller et al. 2017). sPREs have gained popularity for structural studies of biomolecules, including in the structure determination of proteins (Madl et al. 2009; Wang et al. 2012), docking of protein complexes (Madl et al. 2011), and qualitative detection of dynamics (Hocking et al. 2013; Sun et al. 2011; Gong et al. 2017; Gu et al. 2014).

The most commonly used paramagnetic agent for measuring sPRE data is the inert complex Gd(DTPA-BMA) (gadolinium diethylenetriaminepenta-acetic acid bismethylamide, commercially available as ‘Omniscan’), that is known to not specifically interact on protein surfaces (Guttler et al. 2010; Madl et al. 2009, 2011; Pintacuda and Otting 2002; Wang et al. 2012; Respondek et al. 2007; Zangger et al. 2009; Göbl et al. 2010). Previously, we and others could show that sPRE data provide in-depth structural and dynamic data for IDP analysis (Madl et al. 2009; Sun et al. 2011; Gong et al. 2017; Emmanouilidis et al. 2017; Johansson et al. 2014). For example, sPRE data helped to characterize α -helical propensity in a previously postulated flexible region in the folded 42 kDa maltodextrin binding protein (Madl et al. 2009), and dynamic ligand binding to the human “survival of motor neuron” protein (Emmanouilidis et al. 2017). While writing this manuscript, and based on sPRE data for exchangeable amide protons, the Tjandra lab has shown that sPREs can detect native-like structure in denatured ubiquitin (Kooshapur et al. 2018).

Here, we present an integrative ensemble approach to predict the sPREs of IDPs. This ensemble representation is used to calculate conformationally averaged sPREs, which fit remarkably well to the experimentally-measured sPREs. We show for the disordered protein α -synuclein,

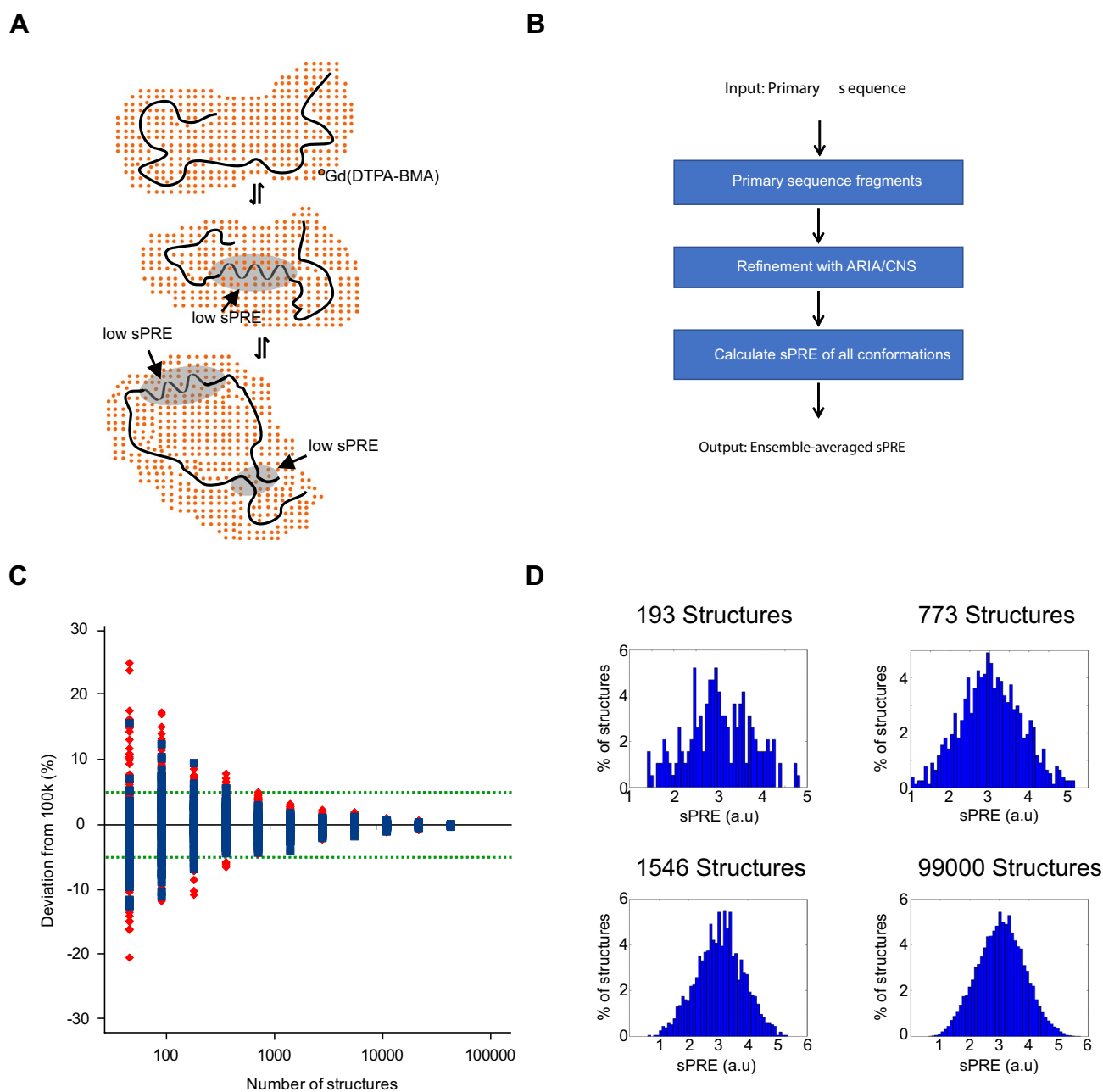


Fig. 1 Principle and workflow for solvent PRE. **a** Transient secondary structures of IDPs are characteristic for protein–protein interaction sites and are therefore crucial for various cellular functions. NMR sPRE data provide quantitative and residue specific information on the solvent accessibility as the effect of paramagnetic probes such as Gd(DTPA-BMA) is distance dependent, which can be used to detect secondary structures within otherwise unfolded regions and long-range contacts within a protein. **b** Prediction of sPRE is based on an ensemble approach of a library of peptides. Each peptide has a length of 5 residues, and is flanked by triple-Ala on both termini (e.g.

AAAXXXXAAA, where XXXXX is a 5-mer fragment of the target primary sequence). Following water refinement using ARIA/CNS, sPRE values of all conformations are calculated and the average solvent PRE value of the ensemble is returned. **c** Predicted C α sPRE (blue) and standard deviation (red) of AAHVAVVAAA ensembles consisting of 99,000 down to 48 structural conformations. The green-dotted line indicates 5% deviation from the ensemble with 99,000 conformations. **d** Histograms of different ensemble sizes showing the distribution of predicted sPRE values

and disordered regions of the proteins FOXO4 and p53 that deviation from random coil behavior can indicate intrinsic propensity to populate transient local structures

and long-range interactions. In summary, this method provides a unique modification-free approach for studying IDPs, that is compatible with a wide range of NMR pulse sequences and biomolecules.

Materials and methods

Protein expression and purification

For expression of human FOXO4^{TAD} (residues 198–505), p53^{TAD} (residues 1–94), pETM11-His₆-ZZ cDNA and including an N-terminal TEV protease cleavage site coding for the respective proteins were transformed into *E. coli* BL21-DE3. To obtain ¹³C/¹⁵N isotope labeled protein, cells were grown for 1 day at 37 °C in minimal medium (100 mM KH₂PO₄, 50 mM K₂HPO₄, 60 mM Na₂HPO₄, 14 mM K₂SO₄, 5 mM MgCl₂; pH 7.2 adjusted with HCl and NaOH with 0.1 dilution of trace element solution (41 mM CaCl₂, 22 mM FeSO₄, 6 mM MnCl₂, 3 mM CoCl₁, 2 mM ZnSO₄, 0.1 mM CuCl₂, 0.2 mM (NH₄)₆Mo₇O₁₇, 24 mM EDTA) supplemented with 2 g of ¹³C₆H₁₂O₆ (Cambridge isotope) and 1 g of ¹⁵NH₄Cl (Sigma). At an OD (600 nm) of 0.8, cells were induced with 0.5 mM isopropyl-β-D-thiogalactopyranosid (IPTG) for 16 h at 20 °C. Cell pellets were harvested and sonicated in denaturing buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 20 mM imidazole, 2 mM tris(2-carboxyethyl)phosphine (TCEP), 20% glycerol and 6 M urea. His₆-ZZ proteins were purified using Ni-NTA agarose (QIAGEN) and eluted in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 200 mM imidazole, 2 mM TCEP and subjected to TEV protease cleavage. Untagged proteins were then isolated by performing a second affinity purification step using Ni-NTA beads for removal of TEV and uncleaved substrate. A final exclusion chromatography purification step was performed in the buffer of interest on a gel filtration column (Superdex peptide (10/300) for p53 and Superdex 75 (16/600) for FOXO4, GE Healthcare).

α-Synuclein was expressed and purified as described (Falsone et al. 2011). Briefly, pRSETB vector containing the human AS gene was transformed into BL21 (DE3) Star Cells. ¹³C/¹⁵N-labeled α-synuclein was expressed in minimal medium (6.8 g/l Na₂HPO₃, 4 g/l KH₂PO₄, 0.5 g/l NaCl, 1.5 g/l (15NH₄)₂SO₂, 4 g/l ¹³C glucose, 1 μg/l biotin, 1 μg/l thiamin, 100 μg/ml ampicillin, and 1 ml 1000× microsalts). Cells were grown to an OD (600 nm) of 0.7. Protein was expressed by addition of 1 mM IPTG for 4 h. After harvesting cells were resuspended in 20 mM Tris-HCl, 50 mM NaCl, pH 7.4, supplemented with a Complete[®] protease inhibitor mix (Roche, Basel, Switzerland). Protein purification was then achieved using a Resource Q column and gel filtration using a Superdex 75 gel filtration column (GE Healthcare, Uppsala, Sweden).

Generation of conformational ensembles

Conformational ensembles were generated using the ARIA/CNS software-package, comprising of 1500 random backbone conformations of all possible 5-mer peptides of the protein of interest, and flanked by triple-alanine. Every backbone conformation served as starting structure in a full-atom water refinement using ARIA (Bardiaux et al. 2012). For every refined structure the solvent PRE is calculated and the averaged solvent PRE of the central residue is stored in the database. To predict sPRE data, a previously published grid-based approach was used (Hartmuller et al. 2016; Pintacuda and Otting 2002). Briefly, the structural model was placed in a regularly-spaced grid representing the uniformly distributed paramagnetic compound and the grid was built with a point-to-point distance of 0.5 Å and a minimum distance of 10 Å between the protein model and the outer border of the grid. Next, grid points that overlap with the protein model were removed assuming a molecular radius of 3.5 Å for the paramagnetic compound. To compute the sPRE for a given protein proton sPREⁱ_{predicted}, the distance-dependent paramagnetic effect (Hartmuller et al. 2016; Hocking et al. 2013; Pintacuda and Otting 2002) was numerically integrated over all remaining grid points according to Eq. (1):

$$\text{sPRE}_{\text{predicted}}^i = c \cdot \sum_{d_{i,j} < 10\text{\AA}} \frac{1}{d_{i,j}^6} \quad (1)$$

where *i* is the index of a proton of the protein, *j* is the index of the grid point, *d_{i,j}* is the distance between the *i*th proton and the *j*th grid point and *c* is an arbitrary constant to scale the sPRE values (1000). Theoretical sPRE values were normalized by calculating the linear fit of experimental and predicted sPRE followed by shifting and scaling of the theoretical sPRE. To predict the solvent PRE of the entire IDP sequence, each peptide with the five matching amino acids is searched and the corresponding solvent PRE values are combined. sPRE data of the two N- and C-terminal residues were not predicted in this setup. All scripts and sample runs can be obtained downloaded from the homepage of the authors (<https://mbbc.medunigraz.at/forschung/forschungseinheiten-und-gruppen/forschungsgruppe-tobias-madl/software/>).

NMR experiments

The setup of sPRE measurements using NMR spectroscopy was performed as described previously (Hartmuller et al. 2016, 2017). To obtain sPRE data, a saturation-based approach was used. The ¹H-R₁ relaxation rates were determined by a saturation-recovery scheme followed by a

read-out experiment such as a ^1H , ^{15}N HSQC, ^1H , ^{13}C HSQC or a 3D CBCA(CO)NH experiment. The read-out experiments were combined with the saturation-recovery scheme in a Pseudo-3D (HSQCs) or Pseudo-4D [CBCA(CO)NH] experiment, with the recovery time as an additional dimension. The CBCA(CO)NH was recorded using non-uniform sampling. Alternatively, ^1H - R_2 relaxation rates can be as described (Clare and Iwahara 2009).

A 7.5 ms ^1H trim pulse followed by a gradient was applied for proton saturation. During the recovery, ranging from several milliseconds up to several seconds, z -magnetization is built up. The individual recovery delays are applied in an interleaved manner, with short and long delays occurring in alternating fashion. For every ^1H - R_1 measurement 10 delay times were recorded and for error estimation, at 1 delay time was recorded as a duplicate.

Measurement of ^1H - R_1 rates were repeated for increasing concentrations of the relaxation-enhancing Gd(DTPA-BMA)/Omniscan (GE Healthcare, Vienna, Austria) and the solvent PRE was obtained as the average change of the proton R_1 rate per concentration of the paramagnetic agent. After each addition of Gd(DTPA-BMA), the recovery delays were shortened such that for the longest delay all NMR signals were sufficiently recovered. The interscan delay was set to 50 ms, as the saturation-recovery scheme does not rely on an equilibrium z -magnetization at the start of each scan. All NMR samples contained 10% $^2\text{H}_2\text{O}$. Spectra were processed using NMRPipe and analyzed with the NMRView and CcpNmr Analysis software packages (Johnson 2004; Delaglio et al. 1995; Skinner et al. 2016).

Measurement of sPRE data used in this study

Assignment of p53^{TAD} was achieved using HNCACB, CBCA(CO)NH and HCAN spectra and analyzed using ccpNMR (Skinner et al. 2016). sPRE data of 300 μM samples of uniformly $^{13}\text{C}/^{15}\text{N}$ labeled p53^{TAD} was measured on a 600 MHz Bruker Avance Neo NMR spectrometer equipped with a TXI probehead at 298 K in a buffer containing 50 mM sodium phosphate buffer, 0.04% sodium azide, pH 7.5. ^1H - R_1 rates of $^1\text{H}^{\text{N}}$, H^{α} and H^{β} were determined using ^1H , ^{13}C HSQC and ^1H , ^{15}N HSQC as read-out spectra (4/4 scans, 200/128 complex points in F2). For assignment of α -synuclein, previously reported chemical shifts were obtained from the BMRB (accession code 6968) and the assignment was confirmed using HNCACB and CBCA(CO)NH spectra. ^1H - R_1 rates of aliphatic protons and amide protons of a 100 μM sample (50 mM bis(2-hydroxyethyl)amino-tris(hydroxymethyl) methane (bis-Tris), 20 mM NaCl, 3 mM sodium azide, pH 6.8) were determined using ^1H , ^{13}C HSQC and ^1H , ^{15}N HSQC read-out spectra, respectively, at 282 K in the presence of 0, 1, 2, 3, 4 and 5 mM Gd(DTPA-BMA). For assignment of FOXO4^{TAD} HNCACB, CBCA(CO)NH and HCAN

spectra were recorded and assigned using ccpNMR (Skinner et al. 2016). Measurements of ^{13}C , ^{15}N labeled FOXO4^{TAD} at 390 μM in 20 mM sodium phosphate buffer, pH 6.8, 50 mM NaCl, 1 mM DTT were performed on a 600 MHz magnet (Oxford Instruments) equipped with an AV III console and cryo TCI probe head (Bruker Biospin). Pseudo-4D CBCA(CO)NH spectra served as read-out for ^1H - R_1 rates and were recorded on a 250 μM sample on a 900 MHz Bruker Avance III spectrometer equipped with a TCI cryoprobe using non-uniform sampling (4 scans, 168/104 complex points in F1 (^{13}C)/F2 (^{15}N) sampled with 13.7% resulting a total number of 600 complex points). Spectra were processed using hmsIST/NMRPipe (Hyberts et al. 2014).

Analysis of NMR data

sPRE data of the model proteins was analyzed as described previously. Briefly, peak intensities were extracted using nmr-glue python package and fitted to a mono-exponential build up curve the SciPy python package and Eq. (2).

$$I(t) = -A \cdot e^{-R_1 \cdot t} + C \quad (2)$$

where $I(t)$ is the peak intensity of the saturation-recovery experiment, t is the recovery delay, A is the amplitude of the z -magnetization build-up, C is the plateau of the curve and R_1 is the longitudinal relaxation rate. Duplicate recovery delays were used to determine the error for the fitted rates R_1 .

$$\epsilon_{\text{exp}} = \sqrt{\frac{1}{2N} \cdot \sum_{i=1}^N \delta_i^2} \quad (3)$$

where N is the number of peaks in the spectrum, i is the index of the peak, and δ_i is the difference of the duplicates for the i th peak. The error of the rates R_1 was then obtained using a Monte Carlo-type resampling strategy. The solvent PRE is obtained by performing a weighted linear regression using the equation

$$R_1(c) = \text{sPRE} \cdot c + R_1^0 \quad (4)$$

where c is the concentration of Gd(DTPA-BMA), $R_1(c)$ is the fitted R_1 rate at the present of Gd(DTPA-BMA) with a concentration c , R_1^0 is the R_1 in the absence of Gd(DTPA-BMA) and sPRE is the slope and the desired sPRE value. For the weighted linear regression, the previously determined errors ΔR_1 for R_1 was used, and the error on the concentration c was neglected.

Results and discussion

To detect transient structural elements in IDPs, an efficient back-calculation of sPREs of IDPs is essential. Whereas back-calculation of sPREs is relatively straightforward for folded rigid structures and can be carried out efficiently using a grid-based approach by integration of the solvent environment (Hartlmuller et al. 2016, 2017), this approach fails in the case of highly conformationally heterogeneous IDPs. In our approach, sPREs are best represented as an average sPRE of an ensemble. NMR observables and nuclear spin relaxation phenomena, including sPREs, directly sense chemical exchange through the distinct magnetic environments that nuclear spins experience while undergoing those exchange processes. The effects of the dynamic exchange on the NMR signals can be described by the McConnell Equations (McConnell 1958) In the case of a two-site exchange process, and assuming that the exchange rate is faster than the difference in the sPREs observed in both states, the observed sPRE is a linear, population-weighted average of the sPRE observed in both states, as seen for covalent paramagnetic labels (Clare and Iwahara 2009). Moreover, the correlation time for relaxation is assumed to be faster than the exchange time among different conformations within the IDP (Jensen et al. 2014; Iwahara and Clare 2010). The effective correlation time for longitudinal relaxation depends on the rotational correlation time of the biomolecule, the electron relaxation time and the lifetime of the rotationally correlated complex of the biomolecule and the paramagnetic agent (Madl et al. 2009; Eletsky et al. 2003). For ubiquitin, the effective correlation time for longitudinal relaxation was found to be in the sub-ns time scale (Pintacuda and Otting 2002), whereas that conformational exchange in IDPs typically appears at slower timescales (Jensen et al. 2014).

Calculating the average of sPREs over an ensemble of protein conformations presents serious practical difficulties that affect both the accuracy and the portability of the calculation. For RDCs it has been shown that convergence to the average requires an unmanageably large number of structures (e.g. 100,000 models for a protein with 100 amino acids), and that the convergence strictly depends on the length of the protein (Bernado et al. 2005; Nodet et al. 2009). To simplify the back-calculation of sPREs we use a strategy proposed for RDCs by the Forman-Kay and Blackledge groups (Marsh et al. 2008; Huang et al. 2013).

To back-calculate the sPRE from a given primary sequence of an IDP we generated fragments of five amino acids of the sequence of interest and flanked them with triple-alanine sequences at the N- and C-termini to simulate the presence of upstream/downstream amino acids (Fig. 1b). An ensemble of structures for these sequences

is then generated using ARIA/CNS including water refinement (Bardiaux et al. 2012). To predict the solvent PRE of the entire IDP, the peptide with the five matching residues is searched and the corresponding solvent PREs averaged for the entire conformational ensemble are returned. This approach is highly parallelizable and dramatically reduces the computational effort compared to simulating the conformations of the full-length IDP.

To determine the number of conformers necessary to converge the back-calculated sPRE of the defined 11-mers, we generated an ensemble of 100,000 structures for a 11-mer AAADVAVVAAA peptide using ARIA/CNS (Bardiaux et al. 2012) and back-calculated the sPRE for subsets with decreasing number of structures. We find that 1500 conformers are sufficient to reproduce the sPRE with a deviation compared to the maximum ensemble below 5% (Fig. 1c, d).

Back-calculation of the sPRE by fast grid-based integration has some advantages compared to alternative approaches relying on surface accessibility (Kooshapur et al. 2018). First, sPREs can be obtained for atoms without any surface accessibility in grid-based integration approaches as they still take into account the distance-dependent paramagnetic effect. This is expected to provide more accurate predictions for regions with a high degree of bulky side chains or transient folding.

To validate our computational approach, we recorded several sets of experimental ¹H-sPREs for the disordered regions of the human proteins FOXO4, p53, and α -synuclein. Similar to many other transcription factors, p53 and FOXO4 are largely disordered outside their DNA binding domains.

In order to demonstrate that surface accessibility data can be obtained for a challenging IDP, we recorded sPRE data for the 307 residue transactivation domain of FOXO4. The FOXO4 transcription factor is a member of the forkhead box O family of proteins that share a highly conserved DNA-binding motif, the forkhead box domain (FH). The FH domain is surrounded by large N- and C-terminal intrinsically disordered regions which are essential for the regulation of FOXO function (Weigel and Jackle 1990). FOXOs control a plethora of cellular functions, such as cell growth, survival, metabolism and oxidative stress, by regulating the expression of hundreds of target genes (Burgering and Medema 2003; Hornsveld et al. 2018). Expression and activity of FOXOs are tightly controlled by PTMs such as phosphorylation, acetylation, methylation and ubiquitination, and these modifications impact on FOXO stability, sub-cellular localization and transcriptional activity (Essers et al. 2004; de Keizer et al. 2010; van den Berg et al. 2013). Because of their anti-proliferative and pro-apoptotic functions, FOXOs have been considered as bona fide tumor suppressors. However, FOXOs can also support tumor development and progression by maintaining cellular homeostasis, facilitating metastasis and inducing therapeutic resistance (Hornsveld

et al. 2018). Thus, targeting FOXO activity might hold promise in cancer therapy.

The C-terminal FOXO4 transactivation domain has been suggested to be largely disordered and to be the binding site for many cofactors. Because it also harbors most of the post-translational modifications (Putker et al. 2013; Burgering and Medema 2003; Hornsveld et al. 2018; Bourgeois and Madl 2018), we set off to study this biologically important domain using our sPRE approach. ^1H , ^{15}N and ^1H , ^{13}C HSQC NMR spectra of FOXO4^{TAD} are of high quality and showed no detectable ^1H , ^{13}C , or ^{15}N chemical shift changes between the spectra recorded in the absence or presence of Gd(DTPA-BMA) (Fig. 2a). sPRE data of FOXO4 had to be recorded in pseudo-4D saturation-recovery CBCA(CO)NH spectra due to the severe signal overlap observed in the 2D HSQC spectra. It should be noted that any kind of NMR

experiment could be combined in principle with a sPRE saturation recovery measurement block in order to obtain ^1H - or ^{13}C sPRE data. The sPRE data of FOXO4^{TAD} yield differential solvent accessibilities in a residue-specific manner (Fig. 2b, c). H^α atoms located in regions rich in bulky residues are showing lower sPREs and H^α atoms located in more exposed glycine-rich regions display higher sPREs. H^β sPRE data was obtained for a limited number of residues and shows overall elevated sPREs due to the higher degree of exposure and a reasonable agreement of predicted and experimental data (Supporting Fig. 1). A comparison of the predicted sPRE data with a bioinformatics bulkiness prediction shows that some features are reproduced by the bioinformatics prediction (Supporting Fig. 2A). However, the experimental sPRE is better described by our approach. Strikingly, the predicted sPRE pattern reproduces the

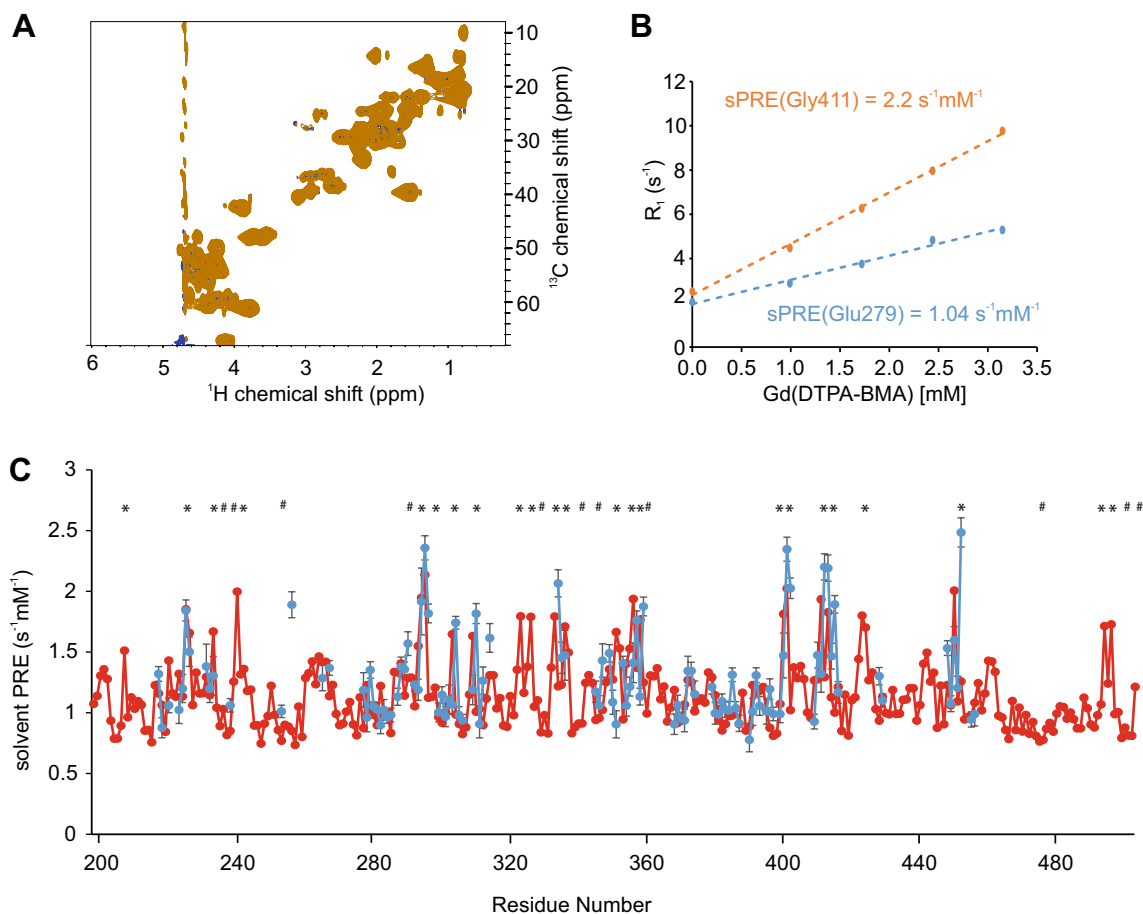


Fig. 2 Comparison of predicted and measured solvent PRE of FOXO4^{TAD}. **a** Overlay of ^1H , ^{13}C HSQC spectra, with full recovery time of a 390 μM ^{13}C , ^{15}N labeled FOXO4^{TAD} sample in the absence (blue) and presence of 3.25 mM Gd(DTPA-BMA) (orange). **b** ^1H - R_1 rates of two selected residues of FOXO4^{TAD} at different Gd(DTPA-BMA) concentrations. **c** Predicted (red) and experimentally-determined (blue) solvent PRE values using CBCA(CO)NH as readout spectrum, of assigned H^α peaks of FOXO4^{TAD}. Experimental sPRE

values are calculated by fitting the data with a linear regression equation. Predicted sPRE values are based on the previously described ensemble approach. Residues with bulky side chains (Phe, Trp, Tyr) are labeled with #, and exposed glycine residues are labeled with * (see Supporting Fig. 2A for a bulkiness profile). Errors of the measured ^1H - R_1 rates were calculated using a Monte Carlo-type resampling strategy and are shown in the diagram as error bars

experimental sPRE pattern exceptionally well, indicating that the FOXO4^{TAD} is largely disordered and does not adopt any stable or transient tertiary structure in the regions for which sPRE data could be obtained.

In order to demonstrate that surface accessibility data can be obtained for a IDP with potential formation of transient local secondary structure we recorded sPRE data for the 94-residue transactivation domain of p53. p53 is a homotetrameric transcription factor composed of an N-terminal trans-activation domain, a proline-rich domain, a central DNA-binding domain followed by a tetramerization domain and the C-terminal negative regulatory domain. p53 is involved in the regulation of more than 500 target genes and thereby controls a broad range of cellular processes, including apoptosis, metabolic adaptation, DNA repair, cell cycle arrest, and senescence (Vousden and Prives 2009). The disordered N-terminal p53 transactivation domain (p53^{TAD}) is a key interaction motif for regulatory protein–protein interactions (Fernandez-Fernandez and Sot 2011): it possesses two binding motifs with α -helical propensity, named p53^{TAD1} (residues 17–29) and p53^{TAD2} (residues 40–57). These two motifs act independently or in combination in order to allow p53 to bind to several proteins regulating either p53 stability or transcriptional activity (Shan et al. 2012; Jenkins et al. 2009; Rowell et al. 2012). Because of its pro-apoptotic function, p53 is recognized as tumor suppressor, and is found mutated in more than half of all human cancers affecting a wide variety of tissues (Olivier et al. 2010). Within this biological and disease context the N-terminal p53-TAD plays a key role: it mediates the interaction with folded co-factors, and comprises most of the regulatory phosphorylation sites.

¹H, ¹⁵N and ¹H, ¹³C HSQC NMR spectra recorded of p53^{TAD} are of high quality and showed no detectable ¹H, ¹³C, or ¹⁵N chemical shift changes between the spectra recorded in the absence or presence of Gd(DTPA-BMA) (Fig. 3a, Supporting Fig. 3A). The sPRE data of p53^{TAD} display differential solvent accessibilities in a residue-specific manner: due to different excluded volumes for the paramagnetic agent H α atoms located in regions rich in bulky residues show lower sPREs and H α atoms located in more exposed regions show higher sPREs (Fig. 3b, c, Supporting Fig. 2B).

sPRE data of structured proteins are often recorded for amide protons. However, chemical exchange of the amide proton with fast-relaxing water solvent protons might lead to an increase of the experimental sPRE, as has been observed for the disordered linker regions in folded proteins and in RNA (Hartlmüller et al. 2017; Gobl et al. 2017). For imino and amino protons of the UUCG tetraloop RNA and a GTP class II aptamer, for example, the increase of ¹H-R₁ rates is larger at small concentrations of the paramagnetic compound, and becomes linear at higher concentrations. Thus, we decided to focus here on experimental and back-calculated sPRE data of H α protons. Nevertheless, ¹H^N-sPREs are

shown for comparison in the supporting information (Supporting Fig. 4A).

Comparison of the back-calculated and experimental p53^{TAD}-sPREs shows that several regions within p53^{TAD} yield lower sPREs than predicted, indicating that p53^{TAD} populates residual local structure or shows long-range tertiary interactions. In line with this, ¹⁵N NMR relaxation data and ¹³C secondary chemical shift data display reduced flexibility of p53^{TAD} and transient α -helical structure (Supporting Fig. 5). This is in line with previous studies which found that the p53^{TAD1} domain adopts a transiently populated α -helical structure formed by residues Phe19-Leu26 and that the p53^{TAD2} domain adopts a transiently populated turn-like structure formed by residues Met40-Met44 and Asp48-Trp53 (Lee et al. 2000). Given that p53^{TAD} has been reported to interact with several co-factors, our data indicate that sPRE data can indeed provide important insight into the residual structure of this key interaction motif (Bourgeois and Madl 2018; Raj and Attardi 2017).

In order to address the question of whether sPREs can be used to detect transient long-range interactions in disordered proteins we recorded ¹H sPRE data for the 141-residue IDP α -synuclein using ¹H, ¹³C and ¹H, ¹⁵N, HSQC-based saturation recovery experiments at increasing concentrations of Gd(DTPA-BMA). α -Synuclein controls the assembly of pre-synaptic vesicles in neurons and is required for the release of the neurotransmitter dopamine (Burre et al. 2010). The aggregation of α -synuclein into intracellular amyloid inclusions coincides with the death of dopaminergic neurons, and therefore constitutes a pathologic signature of synucleinopathies such as Parkinson's disease, dementia with Lewy bodies, and multiple system atrophy (Alafuzoff and Hartikainen 2017). Formation of transient long-range interactions has been proposed to protect α -synuclein from aggregation.

¹H, ¹⁵N and ¹H, ¹³C HSQC NMR spectra of α -synuclein are of high quality and showed no detectable ¹H, ¹³C, or ¹⁵N chemical shift changes between the spectra recorded in the absence or presence of 5 mM Gd(DTPA-BMA) (Fig. 4a). The sPRE data of α -synuclein display variable solvent accessibilities in a residue-specific manner (Fig. 4b), with H α atoms located in regions rich in bulky residues showing lower sPREs and H α atoms located in more exposed regions showing higher sPREs (see also Supporting Fig. 2C for a comparison with the bioinformatics bulkiness profile and Supporting Fig. 4B for the ¹H^N sPRE data). Thus, the sPRE value provides local structural information about the disordered ensemble. Strikingly, we observed decreased sPREs, and therefore lower surface accessibility, in several regions, such as between residues 15–20, 26–30, 52–57, 74–79, 87–92, 102–110, and 112–121, respectively (Fig. 4c). Comparison of these regions with recently-published ensemble modeling using extensive sets of RDC and PRE data (Salmon et al. 2010) shows that the previously-observed

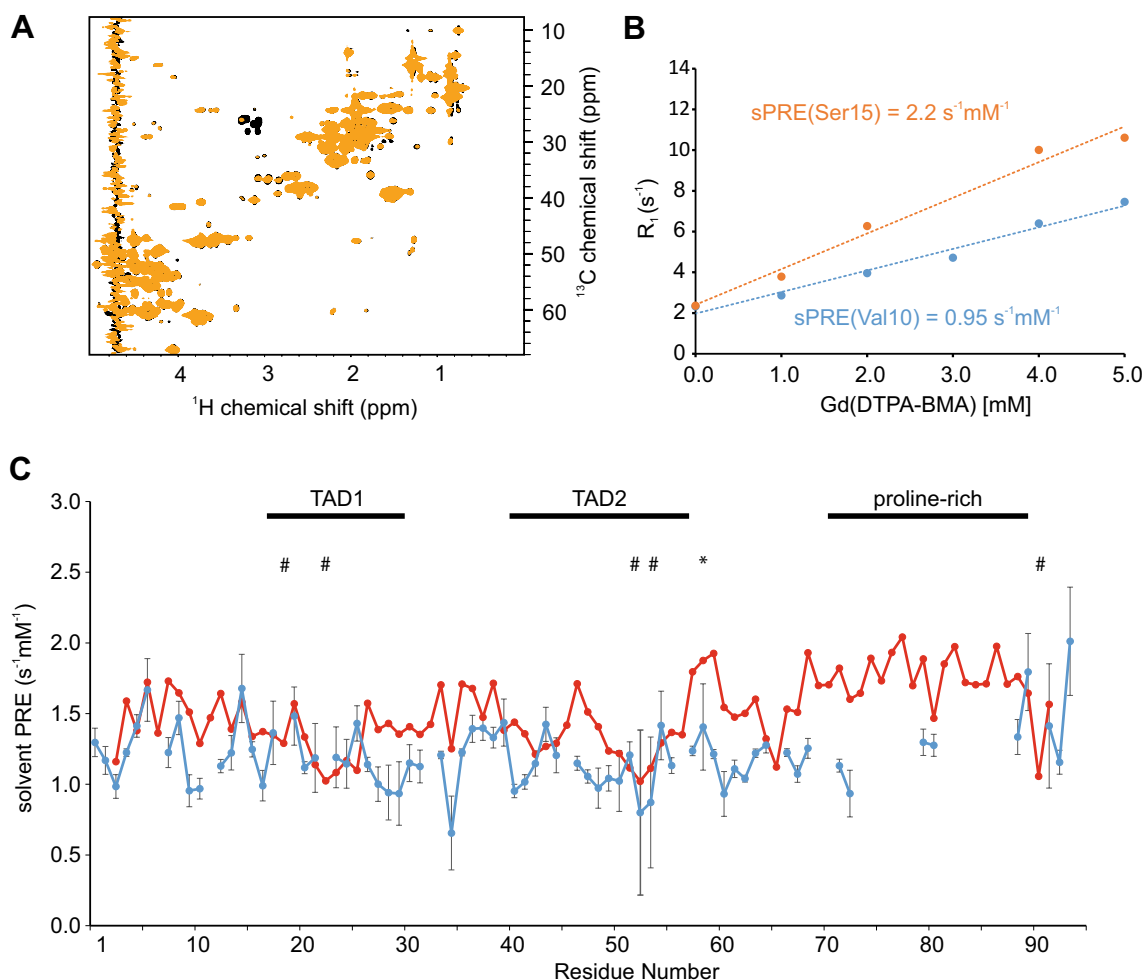


Fig. 3 Comparison of predicted and measured solvent PRE of p53^{TAD} . **a** Overlay of ^1H , ^{13}C HSQC read-out spectra, with full recovery time of a $300\ \mu\text{M}$ ^{13}C , ^{15}N labeled p53^{TAD} in absence (black) and presence of $5\ \text{mM}$ $\text{Gd}(\text{DTPA-BMA})$ (orange). **b** $\text{Gd}(\text{DTPA-BMA})$ -concentration-dependent R_1 rates of two selected residues. **c** Diagram showing predicted (red) and measured (blue) solvent PRE values of each H^α atom of p53^{TAD} . Experimental sPRE values are calculated by fitting the data with a linear regression equation. Predicted sPRE

values are based on the previously described ensemble approach. Regions binding to co-factors (TAD1, TAD2) and the proline rich region are labeled. Residues with bulky side chains (Phe, Trp, Tyr) are labeled with #, and exposed glycine residues are labeled with * (see Supporting Fig. 2B for a bulkiness profile). Errors of the measured ^1H - R_1 rates were calculated using a Monte Carlo-type resampling strategy and are shown in the diagram as error bars

transient intra-molecular long-range contacts involving mainly the regions 1–40, 70–90, and 120–140 within α -synuclein are reproduced by the sPRE data. Thus, sPRE data are highly sensitive to low populations of residual structure in disordered proteins.

Conclusions

In order to understand the conformational behavior of IDPs and their biological interaction networks, the detection of residual structure and long-range interactions is required. The large number of degrees of conformational freedom of IDPs require extensive sets of experimental

data. Here, we provide a straightforward approach for the detection of residual structure and long-range interactions in IDPs and show that sPRE data contribute important and easily-accessible restraints for the investigation of IDPs. Our data indicate that for the general case of an unfolded chain with a local flexibility described by the overwhelming majority of available combinations, sPREs can be accurately predicted through our approach. It can be envisaged that a database of all potential combinations of the 20 amino acids within the central 5-mer peptide can be generated in the future. Nevertheless, generation of sPRE datasets for the entire 3.2 million possible combinations is beyond the current computing capabilities.

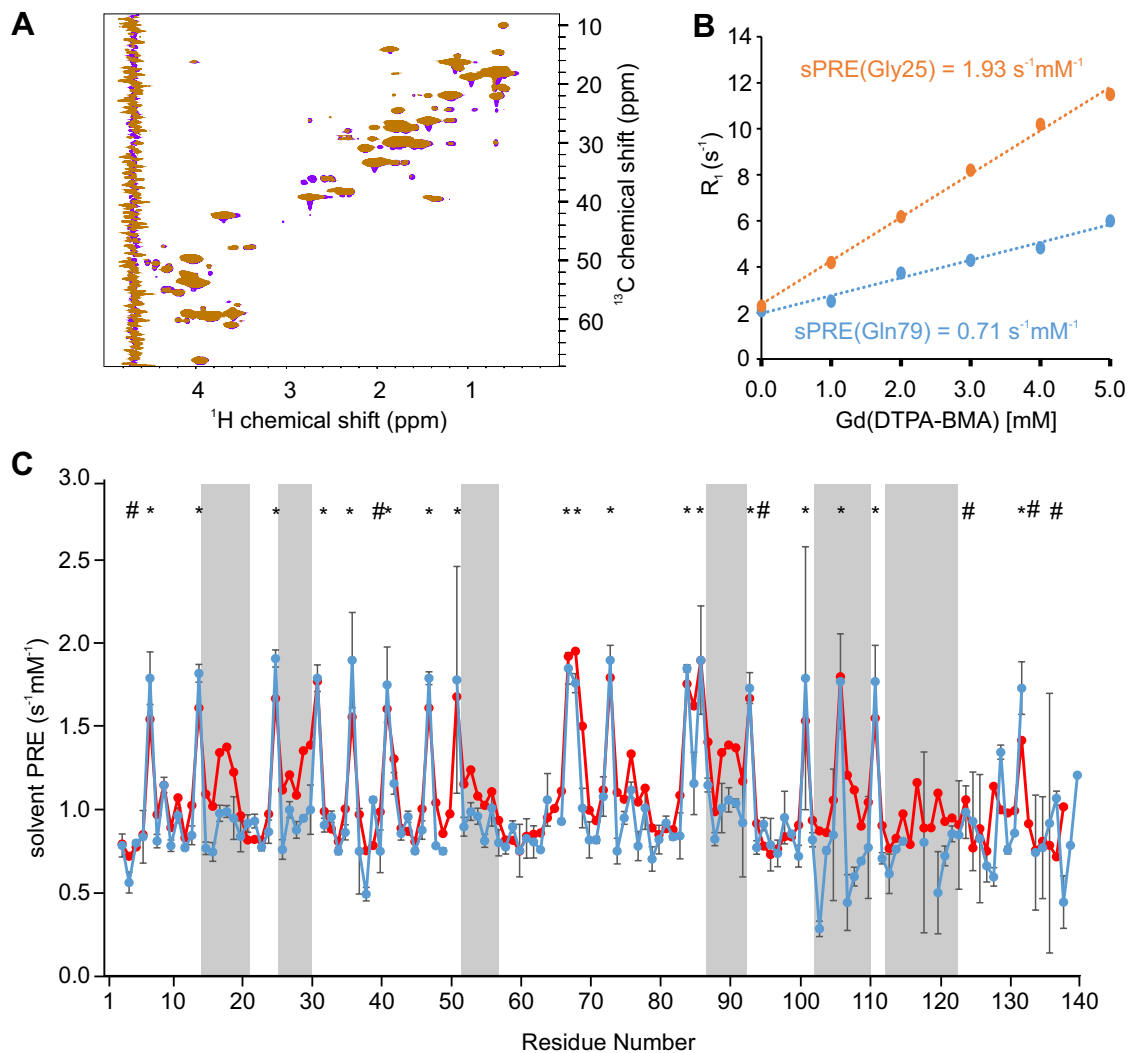


Fig. 4 Comparison of predicted and measured solvent PRE of α -synuclein. **a** Overlay of ^1H , ^{13}C HSQC Read-out spectra, with full recovery time of $100\ \mu\text{M}$ ^{13}C , ^{15}N labeled α -synuclein in absence (violet) and presence of $5\ \text{mM}$ Gd(DTPA-BMA) (orange). **b** Linear fit of relaxation rate ^1H - R_1 and Gd(DTPA-BMA) concentration of two selected residues of α -synuclein. **c** Predicted (red) and experimentally determined (blue) sPRE values from ^1H , ^{13}C HSQC read-out spectra. Regions of strong variations between predicted and measured sPRE

values are highlighted by grey boxes. Experimental sPRE values are calculated by fitting the data with a linear regression equation. Predicted sPRE values are based on the previously described ensemble approach. Residues with bulky side chains (Phe, Trp, Tyr) are labeled with #, and exposed glycine residues are labeled with * (see Supporting Fig. 2C for a bulkiness profile). Errors of the measured ^1H - R_1 rates were calculated using a Monte Carlo-type resampling strategy and are shown in the diagram as error bars

Our approach promises to be a straightforward screening tool to exclude potential specific interactions of the soluble paramagnetic agent with IDPs and to guide positioning of covalent paramagnetic spin labels which are often used to detect long-range interactions within IDPs (Gobl et al. 2014; Clore and Iwahara 2009; Otting 2010; Jensen et al. 2014). Paramagnetic spin labels are preferable placed close to, but not within regions involved in transient interactions in order to avoid potential interference of the spin label with weak and dynamic interactions.

In summary, we used three highly disease-relevant biological model systems for determining the solvent accessibility

information provided by sPREs. This information can be easily determined experimentally and agrees well with the sPREs predicted for non-exchangeable protons using our grid-based approach. Our method proves to be highly sensitive to low populations of residual structure and long-range contacts in disordered proteins. This approach can be easily combined with ensemble-based calculations such as implemented in flexible-meccano/ASTEROIDS (Mukrasch et al. 2007; Nodet et al. 2009), Xplor-NIH (Kooshapur et al. 2018), or other programs (Estana et al. 2019) to interpret residual structure of IDPs quantitatively and in combination with complementary restraints obtained from RDCs and

PREs. In particular for IDP ensemble calculations relying on sPRE data it is essential to exclude specific interactions of the paramagnetic agent with the IDP of interest which would lead to an enhanced experimental sPRE compared to the predicted sPRE.

Acknowledgements Open access funding provided by Austrian Science Fund (FWF). This research was supported by the Austrian Science Foundation (P28854, I3792, DK-MCD W1226 to TM), the President's International Fellowship Initiative of CAS (No. 2015VBB045, to TM), the National Natural Science Foundation of China (No. 31450110423, to TM), the Austrian Research Promotion Agency (FFG: 864690, 870454), the Integrative Metabolism Research Center Graz, the Austrian infrastructure program 2016/2017, the Styrian government (Zukunftsfonds) and BioTechMed/Graz. E.S. was trained within frame of the PhD program Molecular Medicine. We thank Dr. Vanessa Morris for carefully reading the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alafuzoff I, Hartikainen P (2017) Alpha-synucleinopathies. *Handb Clin Neurol* 145:339–353
- Babu MM, van der Lee R, de Groot NS, Gsponer J (2011) Intrinsically disordered proteins: regulation and disease. *Curr Opin Struct Biol* 21:432–440
- Bah A, Forman-Kay JD (2016) Modulation of intrinsically disordered protein function by post-translational modifications. *J Biol Chem* 291:6696–6705
- Bardiaux B, Malliavin T, Nilges M (2012) ARIA for solution and solid-state NMR. *Methods Mol Biol* 831:453–483
- Bernado P, Bertoncini CW, Griesinger C, Zweckstetter M, Blackledge M (2005) Defining long-range order and local disorder in native alpha-synuclein using residual dipolar couplings. *J Am Chem Soc* 127:17968–17969
- Bernini A, Venditti V, Spiga O, Niccolai N (2009) Probing protein surface accessibility with solvent and paramagnetic molecules. *Prog Nucl Magn Reson Spectrosc* 54:278–289
- Bertoncini CW et al (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc Natl Acad Sci USA* 102:1430–1435
- Bourgeois B, Madl T (2018) Regulation of cellular senescence via the FOXO4-p53 axis. *FEBS Lett* 592:2083–2097
- Brady JP et al (2017) Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc Natl Acad Sci USA* 114:E8194–E8203
- Burgering BM, Medema RH (2003) Decisions on life and death: FOXO Forkhead transcription factors are in command when PKB/Akt is off duty. *J Leukoc Biol* 73:689–701
- Burre J et al (2010) Alpha-synuclein promotes SNARE-complex assembly in vivo and in vitro. *Science* 329:1663–1667
- Choy MS, Page R, Peti W (2012) Regulation of protein phosphatase 1 by intrinsically disordered proteins. *Biochem Soc Trans* 40:969–974
- Clore GM, Iwahara J (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem Rev* 109:4108–4139
- Clore GM, Tang C, Iwahara J (2007) Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr Opin Struct Biol* 17:603–616
- de Keizer PL et al (2010) Activation of forkhead box O transcription factors by oncogenic BRAF promotes p21cip1-dependent senescence. *Cancer Res* 70:8526–8536
- Dedmon MM, Lindorff-Larsen K, Christodoulou J, Vendruscolo M, Dobson CM (2005) Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J Am Chem Soc* 127:476–477
- Delaglio F et al (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293
- Dyson HJ, Wright PE (2004) Unfolded proteins and protein folding studied by NMR. *Chem Rev* 104:3607–3622
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6:197–208
- Eletsky A, Moreira O, Kovacs H, Pervushin K (2003) A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using ¹³C spectroscopy. *J Biomol NMR* 26:167–179
- Eliezer D (2009) Biophysical characterization of intrinsically disordered proteins. *Curr Opin Struct Biol* 19:23–30
- Emmanouilidis L et al (2017) Allosteric modulation of peroxisomal membrane protein recognition by farnesylation of the peroxisomal import receptor PEX19. *Nat Commun* 8:14635
- Essers MA et al (2004) FOXO transcription factor activation by oxidative stress mediated by the small GTPase Ral and JNK. *EMBO J* 23:4802–4812
- Estana A et al (2019) Realistic ensemble models of intrinsically disordered proteins using a structure-encoding coil database. *Structure* 27:381–391e2
- Falsone SF et al (2011) The neurotransmitter serotonin interrupts alpha-synuclein amyloid maturation. *Biochim Biophys Acta* 1814:553–561
- Fernandez-Fernandez MR, Sot B (2011) The relevance of protein-protein interactions for p53 function: the CPE contribution. *Protein Eng Des Sel* 24:41–51
- Flock T, Weatheritt RJ, Latysheva NS, Babu MM (2014) Controlling entropy to tune the functions of intrinsically disordered regions. *Curr Opin Struct Biol* 26:62–72
- Gillespie JR, Shortle D (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J Mol Biol* 268:158–169
- Gobl C, Madl T, Simon B, Sattler M (2014) NMR approaches for structural analysis of multidomain proteins and complexes in solution. *Prog Nucl Magn Reson Spectrosc* 80:26–63
- Gobl C et al (2016) Increasing the chemical-shift dispersion of unstructured proteins with a covalent lanthanide shift reagent. *Angew Chem Int Ed Engl* 55:14847–14851
- Gobl C et al (2017) Flexible IgE epitope-containing domains of Phl p 5 cause high allergenic activity. *J Allergy Clin Immunol* 140:1187–1191
- Göbl C, Kosol S, Stockner T, Rückert HM, Zangger K (2010) Solution structure and membrane binding of the toxin fst of the par addiction module. *Biochemistry* 49:6567–6575
- Gong Z, Gu XH, Guo DC, Wang J, Tang C (2017) Protein structural ensembles visualized by solvent paramagnetic relaxation enhancement. *Angew Chem Int Ed Engl* 56:1002–1006
- Gu XH, Gong Z, Guo DC, Zhang WP, Tang C (2014) A decadentate Gd(III)-coordinating paramagnetic cosolvent for protein relaxation enhancement measurement. *J Biomol NMR* 58:149–154

- Guttler T et al (2010) NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol* 17:1367–1376
- Habchi J, Tompa P, Longhi S, Uversky VN (2014) Introducing protein intrinsic disorder. *Chem Rev* 114:6561–6588
- Hartlmüller C, Gobl C, Madl T (2016) Prediction of protein structure using surface accessibility data. *Angew Chem Int Ed Engl* 55:11970–11974
- Hartlmüller C et al (2017) RNA structure refinement using NMR solvent accessibility data. *Sci Rep* 7:5393
- Hass MA, Ubbink M (2014) Structure determination of protein-protein complexes with long-range anisotropic paramagnetic NMR restraints. *Curr Opin Struct Biol* 24:45–53
- Hocking HG, Zangger K, Madl T (2013) Studying the structure and dynamics of biomolecules by using soluble paramagnetic probes. *ChemPhysChem* 14:3082–3094
- Hornsveld M, Dansen TB, Derksen PW, Burgering BMT (2018) Re-evaluating the role of FOXOs in cancer. *Semin Cancer Biol* 50:90–100
- Huang JR, Ozenne V, Jensen MR, Blackledge M (2013) Direct prediction of NMR residual dipolar couplings from the primary sequence of unfolded proteins. *Angew Chem Int Ed Engl* 52:687–690
- Huang JR et al (2014) Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: a combined NMR and SAXS study. *J Am Chem Soc* 136:7068–7076
- Hyberts SG, Arthanari H, Robson SA, Wagner G (2014) Perspectives in magnetic resonance: NMR in the post-FFT era. *J Magn Reson* 241:60–73
- Iwahara J, Clore GM (2010) Structure-independent analysis of the breadth of the positional distribution of disordered groups in macromolecules from order parameters for long, variable-length vectors using NMR paramagnetic relaxation enhancement. *J Am Chem Soc* 132:13346–13356
- Jenkins LM et al (2009) Two distinct motifs within the p53 transactivation domain bind to the Taz2 domain of p300 and are differentially affected by phosphorylation. *Biochemistry* 48:1244–1255
- Jensen MR et al (2009) Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. *Structure* 17:1169–1185
- Jensen MR, Zweckstetter M, Huang JR, Blackledge M (2014) Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chem Rev* 114:6632–6660
- Johansson H et al (2014) Specific and nonspecific interactions in ultraweak protein-protein associations revealed by solvent paramagnetic relaxation enhancements. *J Am Chem Soc* 136:10277–10286
- Johnson BA (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* 278:313–352
- Kooshapur H, Schwieters CD, Tjandra N (2018) Conformational ensemble of disordered proteins probed by solvent paramagnetic relaxation enhancement (sPRE). *Angew Chem Int Ed Engl* 57:13519–13522
- Lee H et al (2000) Local structural elements in the mostly unstructured transcriptional activation domain of human p53. *J Biol Chem* 275:29426–29432
- Madl T, Bermel W, Zangger K (2009a) Use of relaxation enhancements in a paramagnetic environment for the structure determination of proteins using NMR spectroscopy. *Angew Chem Int Ed Engl* 48:8259–8262
- Madl T, Bermel W, Zangger K (2009b) Use of relaxation enhancements in a paramagnetic environment for the structure determination of proteins using NMR spectroscopy. *Angew Chem Int Ed Engl* 48:8259–8262
- Madl T, Gortlich D, Sattler M (2011) Structural analysis of large protein complexes using solvent paramagnetic relaxation enhancements. *Angew Chem Int Ed Engl* 50:3993–3997
- Maji SK et al (2009) Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325:328–332
- Marsh JA et al (2007) Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. *J Mol Biol* 367:1494–1510
- Marsh JA, Baker JM, Tollinger M, Forman-Kay JD (2008) Calculation of residual dipolar couplings from disordered state ensembles using local alignment. *J Am Chem Soc* 130:7804–7805
- McConnell HM (1958) Reaction rates by nuclear magnetic resonance. *J Chem Phys* 28:430–431
- Meier S, Blackledge M, Grzesiek S (2008) Conformational distributions of unfolded polypeptides from novel NMR techniques. *J Chem Phys* 128:052204
- Metallo SJ (2010) Intrinsically disordered proteins are potential drug targets. *Curr Opin Chem Biol* 14:481–488
- Mittag T, Forman-Kay JD (2007) Atomic-level characterization of disordered protein ensembles. *Curr Opin Struct Biol* 17:3–14
- Mukrasch MD et al (2007) Highly populated turn conformations in natively unfolded tau protein identified from residual dipolar couplings and molecular simulation. *J Am Chem Soc* 129:5235–5243
- Nodet G et al (2009) Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings. *J Am Chem Soc* 131:17908–17918
- Olivier M, Hollstein M, Hainaut P (2010) TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol* 2:a001008
- Otting G (2010) Protein NMR using paramagnetic ions. *Annu Rev Biophys* 39:387–405
- Ozenne V et al (2012) Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics* 28:1463–1470
- Parigi G et al (2014) Long-range correlated dynamics in intrinsically disordered proteins. *J Am Chem Soc* 136:16201–16209
- Pintacuda G, Otting G (2002a) Identification of protein surfaces by NMR measurements with a paramagnetic Gd(III) chelate. *J Am Chem Soc* 124:372–373
- Pintacuda G, Otting G (2002b) Identification of protein surfaces by NMR measurements with a paramagnetic Gd(III) chelate. *J Am Chem Soc* 124:372–373
- Putker M et al (2013) Redox-dependent control of FOXO/DAF-16 by transportin-1. *Mol Cell* 49:730–742
- Raj N, Attardi LD (2017) The transactivation domains of the p53 protein. *Cold Spring Harb Perspect Med* 7:a026047
- Respondek M, Madl T, Gobl C, Golser R, Zangger K (2007a) Mapping the orientation of helices in micelle-bound peptides by paramagnetic relaxation waves. *J Am Chem Soc* 129:5228–5234
- Respondek M, Madl T, Gobl C, Golser R, Zangger K (2007b) Mapping the orientation of helices in micelle-bound peptides by paramagnetic relaxation waves. *J Am Chem Soc* 129:5228–5234
- Rezaei-Ghaleh N et al (2018) Local and global dynamics in intrinsically disordered synuclein. *Angew Chem Int Ed Engl* 57:15262–15266
- Romero P et al (1998) Thousands of proteins likely to have long disordered regions. *Pac Symp Biocomput* 3:437–448
- Rowell JP, Simpson KL, Stott K, Watson M, Thomas JO (2012) HMGB1-facilitated p53 DNA binding occurs via HMG-Box/p53 transactivation domain interaction, regulated by the acidic tail. *Structure* 20:2014–2024

- Salmon L et al (2010) NMR characterization of long-range order in intrinsically disordered proteins. *J Am Chem Soc* 132:8407–8418
- Shan B, Li DW, Bruschweiler-Li L, Bruschweiler R (2012) Competitive binding between dynamic p53 transactivation subdomains to human MDM2 protein: implications for regulating the p53. MDM2/MDMX interaction. *J Biol Chem* 287:30376–30384
- Shortle D, Ackerman MS (2001) Persistence of native-like topology in a denatured protein in 8 M urea. *Science* 293:487–489
- Skinner SP et al (2016) CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J Biomol NMR* 66:111–124
- Sun Y, Friedman JI, Stivers JT (2011) Cosolute paramagnetic relaxation enhancements detect transient conformations of human uracil DNA glycosylase (hUNG). *Biochemistry* 50:10724–10731
- Theillet FX et al (2014) Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem Rev* 114:6661–6714
- Tompa P (2012) Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci* 37:509–516
- Uversky VN (2011) Intrinsically disordered proteins from A to Z. *Int J Biochem Cell Biol* 43:1090–1103
- Uversky VN, Oldfield CJ, Dunker AK (2008) Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys* 37:215–246
- Uversky VN et al (2014) Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chem Rev* 114:6844–6879
- van den Berg MC et al (2013) The small GTPase RALA controls c-Jun N-terminal kinase-mediated FOXO activation by regulation of a JIP1 scaffold complex. *J Biol Chem* 288:21729–21741
- van der Lee R et al (2014) Classification of intrinsically disordered regions and proteins. *Chem Rev* 114:6589–6631
- Vousden KH, Prives C (2009) Blinded by the Light: The Growing Complexity of p53. *Cell* 137:413–431
- Wang Y, Schwieters CD, Tjandra N (2012a) Parameterization of solvent-protein interaction and its use on NMR protein structure determination. *J Magn Reson* 221:76–84
- Wang Y, Schwieters CD, Tjandra N (2012b) Parameterization of solvent-protein interaction and its use on NMR protein structure determination. *J Magn Reson* 221:76–84
- Weigel D, Jackle H (1990) The fork head domain: a novel DNA binding motif of eukaryotic transcription factors? *Cell* 63:455–456
- Wells M et al (2008) Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc Natl Acad Sci USA* 105:5762–5767
- Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: reassessing the protein structure-function paradigm. *J Mol Biol* 293:321–331
- Wright PE, Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* 19:31–38
- Wright PE, Dyson HJ (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* 16:18–29
- Yuwen T et al (2018) Measuring solvent hydrogen exchange rates by multifrequency excitation (15)N CEST: application to protein phase separation. *J Phys Chem B* 122:11206–11217
- Zangger K et al (2009a) Positioning of micelle-bound peptides by paramagnetic relaxation enhancements. *J Phys Chem B* 113:4400–4406
- Zangger K et al (2009b) Positioning of micelle-bound peptides by paramagnetic relaxation enhancements. *J Phys Chem B* 113:4400–4406

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.