Structure and Dynamics of an Intrinsically Disordered Protein Region That Partially Folds upon Binding by Chemical-Exchange NMR

Cyril Charlier, Guillaume Bouvignies, Philippe Pelupessy, Astrid Walrant, Rodrigue Marquant, Mikhail Kozlov, Pablo de Ioannes, Nicolas Bolik-coulon, Sandrine Sagan, Patricia Cortes, et al.

Structure and dynamics of an intrinsically disordered protein region that partially folds upon binding by chemical-exchange NMR

Cyril Charlier,[a,b,†] Guillaume Bouvignies,[a,b] Philippe Pelupessy,[a,b] Astrid Walrant,[a,b] Rodrigue Marquant,[a,b] Mikhail Kozlov,[c] Pablo De Ioannes,[d] Nicolas Bolik-Coulon,[a,b] Sandrine Sagan,[a,b] Patricia Cortes,[c,e] Aneel K. Aggarwal,[d] Ludovic Carlier,[a,b,*] Fabien Ferrage,[a,b,*]

[a] Laboratoire des Biomolécules, Département de chimie, École normale supérieure, UPMC Univ. Paris 06, CNRS, PSL Research University, 24 rue Lhomond, 75005 Paris, France.
[b] Sorbonne Universités, UPMC Univ. Paris 06, École normale supérieure, CNRS, Laboratoire des Biomolécules (LBM), 75005 Paris, France.
[c] Department of Medicine, Immunology Institute, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA.
[d] Department of Structural and Chemical Biology, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA.
[e] Department of Molecular, Cellular and Biomedical Science, CUNY School of Medicine, City College of New York, 160 Convent Avenue, New York, NY 10031, USA.

ABSTRACT: Many intrinsically disordered proteins (IDPs) and protein regions (IDRs) engage in transient, yet specific, interactions with a variety of protein partners. Often, if not always, interactions with a protein partner lead to partial folding of the IDR. Characterizing the conformational space of such complexes is challenging: in solution-state NMR, signals of the IDR in the interacting region become broad, weak and often invisible; while X-ray crystallography only provides information on fully ordered regions. There is thus a need for a simple method to characterize both fully and partially ordered regions in the bound state of IDPs. Here, we introduce an approach based on monitoring chemical exchange by NMR to investigate the state of an IDR that folds upon binding through the observation of the free state of the protein. Structural constraints for the bound state are obtained from chemical shifts and site-specific dynamics of the bound state are characterized by relaxation rates. The conformation of the interacting part of the IDR was determined and subsequently docked onto the structure of the folded partner. We apply the method to investigate the interaction between the disordered C-terminal region of Artemis and the DNA binding domain of Ligase IV. We show that we can accurately reproduce the structure of the core of the complex determined by X-ray crystallography and identify a broader interface. The method is widely applicable to the biophysical investigation of complexes of disordered proteins and folded proteins.

Introduction

Intrinsically disordered proteins (IDPs) and regions (IDRs) are ubiquitous in the proteome.[1,2] The flexibility of IDPs allows them to engage in many interactions with other biomolecules.[2] These interactions often lead to folding of a local sequence motif[3] but most parts of a disordered protein usually retain some disorder. In extreme cases, an entire IDP assembles in a fuzzy complex.[4,5] The presence of both ordered and disordered segments makes the characterization of conformational ensembles of the bound states of IDPs challenging.[6] A widely applicable method to determine the conformational properties of the interacting regions of complexes involving long IDRs in solution is still missing.

A conventional approach to the study of IDPs has been to identify a minimal peptide that interacts with a folded domain of partner protein and to determine the structure of the "rigid" complex. This reductionist approach has led to numerous structures of complexes between globular domains and peptides, solved by X-ray crystallography or nuclear magnetic resonance (NMR). The analysis of protein-peptide interactions has produced an impressive body of knowledge on the interactions of classes of protein domains with tiny regions of intrinsically disordered proteins, for instance, Src homology 3 (SH3) domains[7,8] or calmodulin.[9,10]

Recently, protein structures of folded domains that form a complex with long disordered regions have been explored by X-ray crystallography. Such structures provide little, if any, information on those parts of the IDR that do not become fully ordered in the complex[6,11] while they potentially participate in the interaction. NMR has also been used, in particular when most of the disordered region

binds to a small domain.[12,13] The characterization of multiple binding sites involving different fragments of the disordered protein is possible.[11,14,15] In rare studies, NMR and crystallography data have been combined to determine the conformational ensemble of a disordered protein bound to a domain.[16,17] This allows one to identify how binding modulates the conformational space of the disordered protein and may reveal secondary transient interactions.

The determination of high-resolution protein structures by NMR remains challenging for large systems (MW > 30 kDa) and this includes the region of a disordered protein that folds upon binding to a large domain. This part of the IDR has comparatively unfavorable relaxation properties, behaving as part of a large domain with slow tumbling due to the drag of the disordered region.[18,19] Although perdeuteration and transverse relaxation-optimized methods may attenuate these drawbacks, fast relaxation leads to weak and broad signals that are difficult to identify amongst the many intense and sharp signals from the regions that retain disorder. At worst, conformational dynamics on μs-ms timescales in the complex may lead to dramatic line broadening.[20-22]

Would it be possible to design a general method to determine the structure of a complex between an otherwise disordered protein region and a folded protein domain by NMR? This method should be robust enough to be efficient even in the absence of exploitable spectra of the complex. Here, we introduce such an approach based on chemical-exchange NMR. Recent progress in the characterization of chemical exchange processes by NMR.[23] using Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion (RD) [24,25] and chemical-exchange saturation transfer (CEST)[26,27] provide information on weakly populated states that exchange with a ground state easily observed by NMR. CPMG and CEST experiments have been used to determine the kinetics of folding upon binding for some IDRs.[22,28,29] The chemical shifts of a weakly populated state can be obtained from CEST and CPMG and employed as constraints to determine the structure of this NMR-invisible state, whether it is an excited state[30,31] or a complex in exchange with a free state.[32] Weakly populated states with lifetimes between 1 ms and 100 ms are ideally suited for these approaches. Interestingly, a significant share of IDRs form transient but specific complexes with similar lifetimes.[33]

We employ CPMG and CEST experiments to characterize the bound form of an intrinsically disordered protein region: a 96-residue long construct from the disordered C-terminal region of the protein Artemis bound to the DNA-binding domain (DBD) of Ligase IV. Artemis is a nuclease, which plays a key role in adaptive immunity and DNA repair through its participation in the non-homologous end-joining (NHEJ) pathway.[34] The structure of a complex between an 11-residue peptide from Artemis and the DBD of Ligase IV has been solved by X-ray crystallography.[35,36] The chemical shifts obtained for the bound form of Artemis were used to derive dihedral angles for the backbone and a few side chains. We performed NMR titrations of the DBD of Ligase IV with Artemis to obtain additional structural constraints. We used these data alongside the structure of free Ligase IV to calculate sets of structures with the program HADDOCK.[37] We show that the interaction between Artemis and Ligase IV extends beyond the interface defined by X-ray crystallography through additional transient contacts. Overall, we propose an integrative approach based on NMR to characterize both fully and partially ordered regions in the bound state of IDPs.

Results

As a test case, we have investigated the interaction between the disordered C-terminal region of the protein Artemis and the DBD of Ligase IV. Artemis contains a catalytic core comprised of β-lactamase and β-CASP domains, as well as a long C-terminal region that is predicted to be disordered[38] and proposed to be involved in the regulation of the catalytic activity (see Figure S2.a). We have examined the region of Artemis that encompasses residues 480-575 ($Art^{480-575}$), which includes residues 485-495, shown previously to interact with the DBD of Ligase IV.[39] The crystal structures of the core of the complex[35,36] (PDB codes 4HTP and 3W1G) are remarkably similar and show the Artemis peptide 485-495 nested in a hydrophobic pocket between helices α1 and α2 of Ligase IV (Figure S2.b).

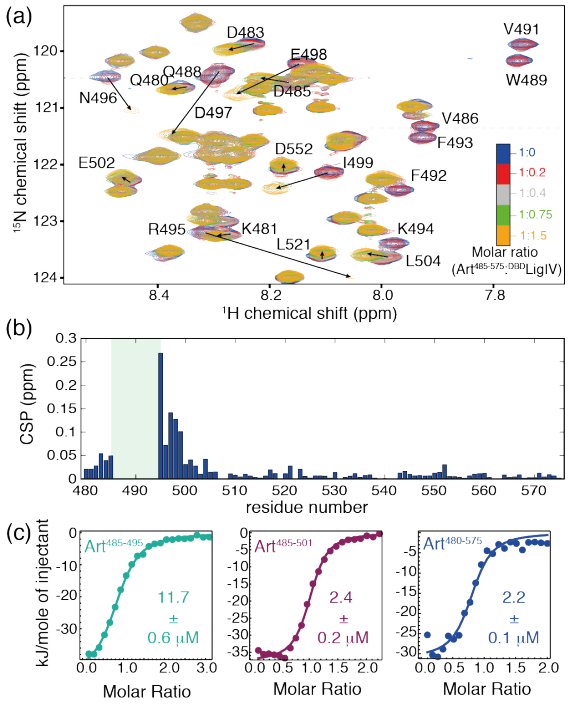

Figure 1. Interaction between Artemis and Ligase IV. (a) Titration of ^{15}N labeled $Art^{480-575}$ by unlabeled $^{DBD}LigIV$ followed by two-dimensional [^{1}H-^{15}N] heteronuclear single-quantum correlation (HSQC) spectroscopy. (b) Chemical shift perturbations (see methods) between the free and bound forms of $Art^{480-575}$ for residues with observable signals in the bound state. (c) Isothermal titration calorimetry of $^{DBD}LigIV$ and three constructs of Artemis: $Art^{485-495}$ (green), $Art^{485-501}$ (purple), and $Art^{480-575}$ (blue).

The ligase IV-binding site of Artemis extends beyond residues 485-495. We obtained a nearly complete NMR assignment of the backbone resonances of Art$^{480\text{-}575}$ (Fig. S3). The narrow dispersion of the amide ^{1}H chemical shifts in the ^{1}H-^{15}N HSQC spectrum shows that Art$^{480\text{-}575}$ is highly disordered in solution, as confirmed by secondary structure propensities based on chemical shifts[40] (Fig. S13).

An NMR titration shows the extent of the interaction of Art$^{480\text{-}575}$ with the DBD of Ligase IV (^{DBD}LigIV). The intensities of ^{1}H-^{15}N cross-peaks for Artemis residues 485-499 decrease substantially upon the addition of substoichiometric amounts of unlabeled ^{DBD}LigIV (Fig. 1.a), which is a marker of slow to intermediate chemical exchange on the chemical shift timescale. This region includes the 11-residue motif 485-495 that has been originally proposed as the Ligase IV binding site[39] as well as the four following residues at the C-terminus. At stoichiometric amounts of ^{DBD}LigIV, additional signals can be observed in the [^{1}H-^{15}N] HSQC spectra of Artemis, which could be assigned to residues 485 and 495-499 in the bound form based on triple resonance experiments. The signals of Artemis residues 486-494 in the bound form remained undetected in both 2D and 3D spectra. The chemical shifts of residues 495-499 drastically change upon Artemis binding to ^{DBD}LigIV (Figures 1.a,b). By contrast, the signals of residues flanking the segment 485-499 exhibit rather small chemical shift perturbations (CSP) and are in the fast exchange regime.

We used isothermal titration calorimetry (ITC) to investigate whether the Ligase IV-binding site of Artemis may extend beyond the C-terminus of the previously studied region 485-495.[39] We synthesized and titrated two Artemis peptides: Art$^{485\text{-}495}$ and a longer peptide, Art$^{485\text{-}501}$. The dissociation constant for the binding of Art$^{485\text{-}495}$ to ^{DBD}LigIV (K_d = 11.7 ± 0.6 μM) is similar to the previously reported value obtained under slightly different conditions[35] (Fig. 1.c and Table S1). The addition of six residues at the C-terminus to give Art$^{485\text{-}501}$ leads to a five-fold reduction in the dissociation constant (K_d = 2.3 ± 0.2 μM), confirming the role of these additional residues in the interaction with ^{DBD}LigIV. Since Art$^{480\text{-}575}$ binds ^{DBD}LigIV with identical affinity, residues 502-575 do not contribute significantly to the thermodynamics of the interaction, as suggested by NMR chemical shift perturbations. Under a higher salt concentration (500 mM NaCl) the dissociation constant for the complex between Art$^{485\text{-}495}$ and ^{DBD}LigIV barely changes while the K_d for the Art$^{485\text{-}501}$:^{DBD}LigIV complex increases 2.6-fold (Table S1, Figure S4). Although all these observations might point to a role of electrostatic interactions involving residues 496-501 of Artemis, the differences in the free energy of binding between Art$^{485\text{-}495}$ and Art$^{485\text{-}501}$ arise mostly from a reduction in the entropic penalty upon binding (see discussion in Supporting Information).

Characterization of the bound state of Art$^{480\text{-}575}$ by chemical-exchange NMR. The observation of NMR signals for the core of Art$^{480\text{-}575}$ in complex with ^{DBD}LigIV is challenging at best. CPMG relaxation dispersion and CEST experiments allow the characterization of states that are not visible in NMR spectra, provided that they interconvert with a major, observable state on millisecond time-scales. Analysis of the "invisible" bound form is achieved by analyzing CPMG[28] or CEST[22,41] profiles of resonances from the free protein measured for a sample containing the protein of interest and substoichiometric amounts of the target protein, so that the bound state is lowly populated.

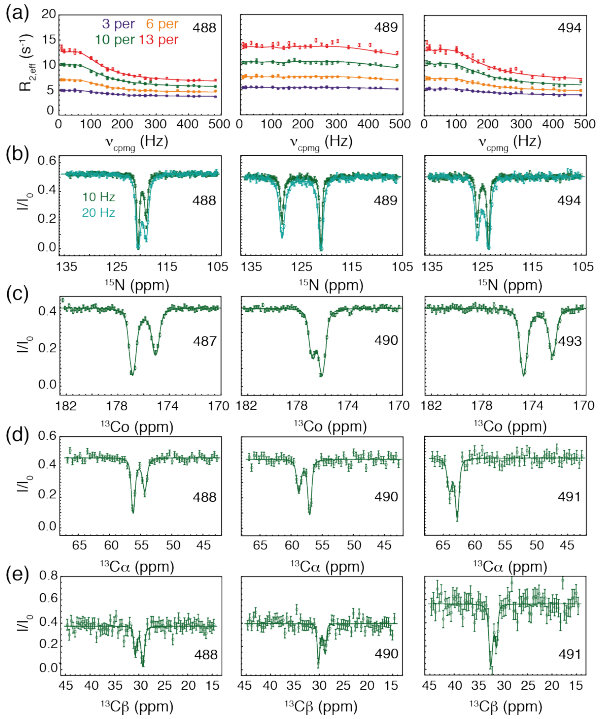

Figure 2. Chemical-exchange NMR investigation of the binding of Art$^{480\text{-}575}$ to ^{DBD}LigIV. (a) Nitrogen-15 CPMG dispersion profiles for representative residues of Art$^{480\text{-}575}$ at the interface with ^{DBD}LigIV. The concentration of ^{DBD}LigIV was varied from 0.03 (purple) to 0.06 (orange), 0.10 (green), and 0.13 equivalent (red). (b) Nitrogen-15 CEST profiles obtained for three representative residues of Art$^{480\text{-}575}$. Two different *rf* field amplitudes were used 10.4 Hz (green) and 20.8 Hz (light green). Representative ^{13}C' (c), $^{13}C^{\alpha}$ (d) and $^{13}C^{\beta}$ (e) CEST profiles of residues of Art$^{480\text{-}575}$ at the interface with ^{DBD}LigIV. Experimental data are shown as small circles. Fitted curves are shown as solid lines in all panels. All CEST experiments were recorded in the presence of 0.1 molar equivalent of unlabeled Ligase IV.

We investigated the bound form of Art$^{480\text{-}575}$ with nitrogen-15 CPMG relaxation dispersion experiments[42] on samples of ^{15}N-labeled Art$^{480\text{-}575}$ with increasing amounts of unlabeled ^{DBD}LigIV (molar ratios ranging from 0 to 13 %) (Fig. 2.a and S5). In the absence of ^{DBD}LigIV, only flat profiles are observed and no exchange process can be detected (data not shown). However, addition of merely 3% of ^{DBD}LigIV leads to RD profiles showing clear signs of exchange contributions to $R_{2,eff}$ for ^{15}N resonances belonging to residues 486-499 (Fig. 2.a and S5). Larger RD profiles are obtained with higher concentrations of ^{DBD}LigIV,

confirming that the exchange process corresponds to the binding of Art$^{480\text{-}575}$ to ^{DBD}LigIV.

To characterize the complex formation in greater detail, CPMG RD experiments were complemented with ^{15}N CEST experiments[26,27,43] performed on nitrogen-15 labeled Art$^{480\text{-}575}$ with 10 % molar ratio of unlabeled ^{DBD}LigIV (Fig. 2.b and S6-7). Most sites from the region 486-499 exhibit profiles with two well-separated intensity dips, with the smaller dip corresponding to the bound state resonances, which confirms that, for these sites, ^{15}N chemical shifts are significantly different between the free and the bound states (Fig. 2.b).

All ^{15}N CPMG RD and CEST data were simultaneously analyzed using the program ChemEx[44] (see Material and Methods) and could be well fit to a global two-site exchange model:

$$\text{Art}^{480\text{-}575} + {}^{DBD}\text{LigIV} \rightleftharpoons \text{Art}^{480\text{-}575} - {}^{DBD}\text{LigIV}$$

with an association rate constant k_{on} = 3.00·10^7 ± 0.01·10^7 s^{-1}.M^{-1} and a dissociation rate constant $k_{off} = k_{on} \cdot K_d$, = 66.1 ± 0.2 s^{-1}, where K_d = 2.2 µM is the dissociation constant determined by ITC. The analysis also provided nitrogen-15 chemical shifts and effective transverse relaxation rates R_2(^{15}N) for the bound state of Art$^{480\text{-}575}$. Chemical shift differences as large as 7.7 ppm for residue 489 were measured. The nitrogen-15 chemical shifts of residues 486-499 in the bound form are shifted away from random coil values (Fig. S12), clearly indicating that this region becomes more structured when bound to ^{DBD}LigIV.

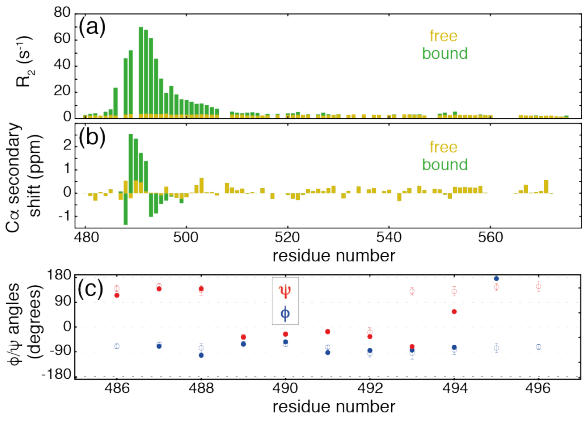

Figure 3. Structural and dynamic information on the bound form of Art$^{480\text{-}575}$ obtained by chemical-exchange NMR. (a) Transverse relaxation rates R_2(^{15}N) and (b) secondary ^{13}C^a chemical shifts for the free (orange) and bound (green) states of Art$^{480\text{-}575}$. (c) Backbone dihedral angles ϕ (blue) and ψ (red) for the bound state of Art$^{480\text{-}575}$. TALOS-N predictions are shown as open symbols, while values found in the crystal structure (PDB 3W1G) are shown as full symbols. Predictions for residue F492 are shown with light colors, as they were not scored the highest confidence level by TALOS-N.

R_2(^{15}N) relaxation rates within the interaction region vary drastically upon binding (Figs. 3.a and S14). Higher rates at the core of the interaction site are expected from folding and slow molecular tumbling of ^{DBD}LigIV within the complex. Indeed, the average R_2(^{15}N) rate, over secondary structure elements, measured on a sample of nitrogen-15 labeled ^{DBD}LigIV in complex with unlabeled Art$^{480\text{-}575}$ is $\langle R_2^{DBD\text{-}bound} \rangle$ = 44 ± 6 s^{-1} (Fig. S16). R_2(^{15}N) rates for Artemis residues 488-494 range between 45 and 75 s^{-1}, defining the core of the interface. The highest rates, obtained for residues 491-493, suggest the presence of an additional fast (µs) exchange process in the bound state. By contrast, R_2(^{15}N) rates confirm the highly disordered nature of regions 480-485 and 510-575 in the complex (Fig. 3.a). Interestingly, in the bound state, ^{15}N transverse relaxation R_2 rates for the 495-510 region are lower than in the core of the complex but larger than in the free state. This region spans about twice the persistence length in disordered or unfolded proteins[45] suggesting a partly restricted conformational space explored on ps-ns timescales, possibly due to transient interactions with Ligase IV.

Additional structural restraints are necessary to determine the conformation of Art$^{480\text{-}575}$ in complex with ^{DBD}LigIV. Carbon-13 chemical shifts of Art$^{480\text{-}575}$ in the bound form were obtained from CEST experiments recorded on carbonyl,[46] carbon-α,[47] and carbon–β,[47] using a uniformly ^{15}N,^{13}C-labeled sample of Art$^{480\text{-}575}$ with 10 % molar ratio of unlabeled ^{DBD}LigIV (Figs. 2.c-e and S8-10). The large, positive ^{13}C$^{\alpha}$ secondary shift values[48,49] observed for residues 489-492 in the bound state (Fig. 3.b) strongly suggest these residues form an α-helix in the complex.[50] All the chemical shifts of the bound form of Artemis obtained from the combined analysis of CEST profiles were used to derive dihedral backbone angles using the program Talos-N.[51] These dihedral angles are in excellent agreement with the crystal structures for the core of the complex, except for residues 493 and 494, which are in an extended conformation in solution, but adopt a helical conformation in the crystals (Fig. 3.c). In addition, Talos-N derived dihedral angles χ1 for residues 486 and 491-493. Notably, our approach provides additional structural restraints for residues 495 and 496.

Mapping the binding to Artemis on the surface of Ligase IV. Relevant structural information on the binding surface of Ligase IV may be provided through the use of chemical shifts perturbations upon binding to Artemis. For that purpose, we obtained the site-specific assignment of backbone resonances for most of the 240 residues of ^{DBD}LigIV (Fig. S15).

The binding surfaces of ^{DBD}LigIV to all Artemis constructs were investigated by NMR titrations. Art$^{480\text{-}575}$ induces selective chemical shift perturbations in the ^{1}H-^{15}N TROSY[52] spectrum of ^{DBD}LigIV (Fig. S17), showing that binding does not lead to a major conformational rearrangement in the large globular domain. Residues displaying the highest CSP values are mostly localized in the N-terminal region encompassing helices α1 and α2 (Fig. 4.a,c), consistent with the binding interface determined by X-ray crystallography using the Art$^{485\text{-}495}$ peptide. Significant perturbations (*i.e.,* one standard deviation beyond the mean CSP) are also observed for residues clustered near the loop connecting helices α6 to α7 (138-144),

the first half of helix α9 (181-184), and helix α10 (198-206). Apart from residues in helix α10, these regions were substantially less affected when ^{DBD}LigIV was titrated by the Art$^{485-495}$ peptide, with changes in chemical shifts mostly limited to the N-terminal region (Fig. S18.c). A detailed comparison between the CSP values of Art$^{480-575}$ and Art$^{485-495}$ (Fig. 4.b,d) suggests a continuous interaction surface that expands beyond the groove between helices α1 and α2, where Art$^{485-495}$ lies in the crystal structure. The extension of the surface includes ^{DBD}LigIV residues F131, L138, G141, S179, and L181 as well as residues localized at the tip of the hairpin α1-α2 (T22, K28, K30, and G31). Importantly, the role of Artemis residues 496-501 in making interactions with these additional docking sites can be inferred from the highly similar CSP profiles obtained with Art$^{480-575}$ and the Art$^{485-501}$ peptide (Fig. S18.a,b).

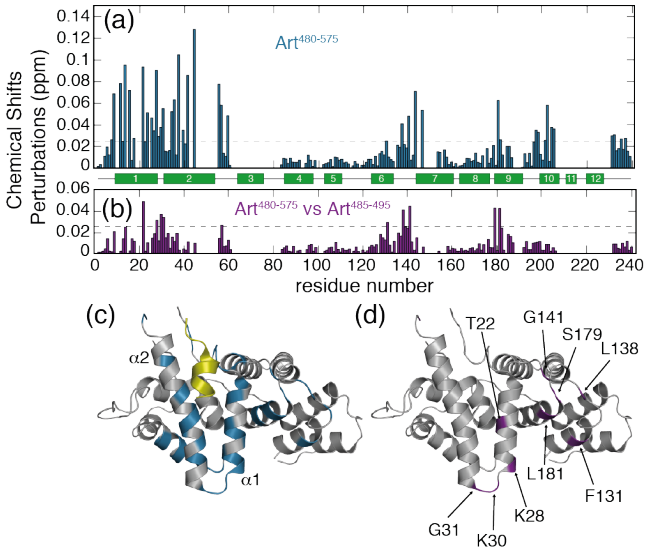

Figure 4. Binding surfaces of ^{DBD}LigIV to Artemis with chemical shift perturbation (CSP). (a) Amide CSP of ^{DBD}LigIV induced by the addition of 1 molar equivalent of Art$^{480-575}$. The average CSP value (0.025 ppm) is indicated by a dotted line. (c) Ligase IV residues with a CSP value larger than 0.025 ppm are colored in blue on the X-ray structure of the complex ^{DBD}LigIV-Art$^{485-495}$ (PDB 3W1G) with Artemis in yellow and ^{DBD}LigIV in grey. (b) Differences in amide CSP of ^{DBD}LigIV between the complexes with Art$^{480-575}$ and Art$^{485-495}$. (d) Residues with a CSP differences larger than 0.025 ppm between the two complexes are shown on the structure of ^{DBD}LigIV (purple).

NMR data-driven model of the Artemis-Ligase IV complex. A structural model of the Artemis-Ligase IV complex in solution was determined by docking computation using experimental constraints derived from NMR chemical-exchange and CSPs data. We first determined the structure of Artemis in the bound state by incorporating in CNS [53] the ϕ, ψ, and χ1 torsion angle restraints derived from the chemical shifts of the bound form. To simplify and speed up the analysis, we only considered the binding site 485-501 in the sequence of Artemis (Art$^{485-501}$) and did not include the region 502-575, which is barely altered by binding. Ten Artemis conformers that are in good agreement with the torsion angle restraints were selected as the starting structural ensemble for the docking process (Table S2). These conformers are remarkably similar to the X-ray structure of Art$^{485-495}$ in complex with Ligase IV in the N-terminal part with an average backbone RMSD of 0.62 Å over residues 486-493 (Fig. 5.a). The C-terminal residues 497-501 are highly disordered in the starting NMR ensemble, as expected from the absence of angular constraints for these residues.

HADDOCK is one of the most popular docking programs for the structure determination of biomolecular complexes using experimental data, in particular NMR CSPs.[37] HADDOCK has been used to determine the structures of homodimers and homo-oligomers of folded proteins from chemical shift perturbations obtained by chemical-exchange NMR[54] as well as the structure of the complex of a short peptide bound to a folded domain[55] based on transferred nuclear Overhauser effects. However, to our knowledge, HADDOCK has never been used in combination with unambiguous structural constraints from chemical-exchange NMR to obtain the structure of a complex involving an IDP that folds upon binding. We docked the structures of Artemis in its bound form onto the structure of the free catalytic core of Ligase IV (residues 1-611, PDB 3W5O)[36] using the HADDOCK web server. The "active" and "passive" residues, as defined by HADDOCK to derive ambiguous intermolecular restraints (AIRs), were selected for Ligase IV on the basis of CSPs. The torsion angle restraints used to calculate the bound structure of Artemis were also used in the docking process. HADDOCK provided 92 final models, sorted in 12 clusters (Fig. S19), each containing 4 to 16 structures with HADDOCK- and *Z*-scores ranging from –100.7 to –51.4 and from –1.3 to +1.4, respectively. The positions of the N-terminal residues 485-496 of Artemis with respect to Ligase IV are remarkably similar in the top 3 clusters (Fig. S20). By contrast, the conformations of residues 497-501 and their contacts with Ligase IV differ markedly between the models and even within clusters (Fig. 5.c).

In the top cluster, the average ligand-RMSD upon superimposition onto ^{DBD}LigIV is 2.1 Å for the structured region encompassing Artemis residues 485-496. The structure of the complex with lowest energy is remarkably similar to the high-resolution crystal structure of the Art$^{485-495}$-LigIV complex (Fig. 5.b). The average backbone interface-RMSD with respect to the crystal structure (PDB 3W1G) is 0.9 ± 0.2 Å for the top cluster. The key hydrophobic interactions involving P487 and aromatic residues W489, F492, and F493 of Artemis are virtually identical to those obtained in the crystal structure. Major differences are only observed for the backbone of residues F493 and K494 that adopt an extended conformation in our NMR structures in solution while they adopt a α-helix conformation in the crystal structures with a short Art$^{485-495}$ peptide.

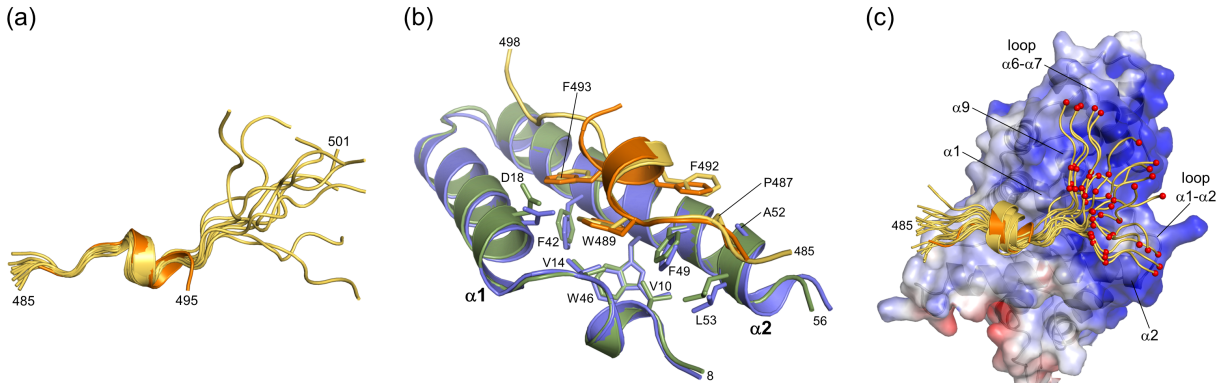

Figure 5. NMR-based structural models of the complex between Artemis and Ligase IV. (a) Solution structure of $Art^{485-501}$ bound to Ligase IV derived from $^{15}N^{H}$, ^{13}CO, $^{13}C\alpha$, and $^{13}C\beta$ chemical shifts. The 10 NMR conformers of $Art^{485-501}$ (yellow) are superimposed onto residues 486-493 from the crystal structure of $Art^{485-495}$ in complex with Ligase IV (orange, PDB 3W1G). (b) Comparison of the contacts at the interface of the complex in the crystal structure (Artemis in orange and Ligase IV in blue, PDB 3W1G) and in the lowest-energy HADDOCK model of the top cluster (Artemis in yellow and ^{DBD}LigIV in green). (c) The 16 models of HADDOCK top cluster are superimposed onto the crystal structure of the complex as in panel (b). The electrostatic surface potential map of ^{DBD}LigIV is shown in blue, white, and red to indicate positive, neutral, and negative electrostatic surface potentials, respectively. Residues of Artemis with acidic side-chains (D497, E498, and D501) are represented by small red spheres.

The consistency of our approach was evaluated by the comparison of experimental chemical shifts with those predicted with Sparta+[56] from the 16 structures of the HADDOCK top cluster. Predicted chemical shifts agree very well with experimental shifts (Figure S21) while small differences are observed with the chemical shifts predicted from the crystal structure (pdb code 3W1G), particularly for residues F493 and K494.

Interestingly, the extended conformations of these residues allow the acidic patch that follows towards the C-terminus (including D497, E498, and D501) to face adjacent regions at the surface of Ligase IV (Fig. 5.c). These additional regions, encompassing the tip of the hairpin α1-α2, the loop α6-α7, and the N-terminal portion of α9, form a continuous positively charged surface. Accordingly, the negatively charged residues D497, E498, and D501 of Artemis contact various arginine and lysine residues of Ligase IV in most docking models through electrostatic interactions. The transient nature of these additional contacts is supported by $R_2(^{15}N)$ rates and chemical shifts which show that Artemis residues 497-501 retain some disorder in the complex. We are well aware that the docking-based structures depicted in Fig. 5.c do not represent *a priori* the ensemble of conformations explored by the Artemis 496-501 region when bound to Ligase IV. However, this ensemble of docking solutions is consistent with the main features identified by NMR data. Predicted chemical shifts from the docking models are in good agreement with experimental values (Fig. S21) and indicate that residues 497-501 are mostly disordered. Weak interactions with the positively charged residues at the surface of Ligase IV may explain the slow decrease of nitrogen-15 R_2 rates in the 497-501 region (Fig. 3.a).

Discussion

The agreement of the NMR-driven docking model with the crystal structure demonstrates that our approach can be applied to obtain structural models of complexes in which the interacting region of an IDP contains both fully ordered and partially disordered regions. The structure of the well-folded parts of the complex comprising Artemis and Ligase IV only differs from the crystal structure, for the backbone dihedral angles of F493 and K494. The proximity of the C-terminus of the short peptide in the crystal structure likely favors the helical conformation. This discrepancy reveals a possible alternative interpretation of this region in the model of the crystal structure. The C-terminus of the short peptide in the crystal structure is built in a helical conformation. However, high B-factors and a weaker electron density in this region are not inconsistent with a more extended path of this segment as obtained by NMR (Fig. S22). This highlights the importance of using long disordered protein regions to grasp accurately the relevant features of such a complex.

The methodology introduced in this work is particularly well suited to the study of IDPs. Indeed, the bound structures of IDPs are usually limited to isolated secondary structure elements (*i.e.,* extended motifs and alpha helices), which lie against the surface of the folded partner protein. Such structure elements can be (i) well defined with the use of backbone dihedral angle constraints, and (ii) accurately docked onto the protein partner from the knowledge of its interacting surface. Here, we show that (i) backbone dihedral angles are obtained in a reliable manner through chemical shifts derived from an analysis of CEST profiles and (ii) the interface is sufficiently well characterized by NMR chemical shift perturbations.

A potential limitation of our approach lies in the ability to define the interface of the complex for large folded protein partners. Structural models of complexes between large proteins and peptides or unfolded proteins have been characterized[15] with the use of methyl transverse relaxation optimized spectroscopy (methyl

TROSY).[57] Protein surfaces contain statistically more methyl-bearing residues at protein-protein interfaces.[58,59] The characterization of protein interactions with methyl probes is thus favorable, as was successfully demonstrated in several studies.[60] CEST profiles, used to obtain the chemical shifts of the bound form, have been recorded on systems with very high transverse relaxation rates, with $R_2(^{15}N) > 1000\ s^{-1}$.[22] This guarantees successful application of our methodology for complexes up to the MDa range.

Our approach is compatible and should now be complemented with methods that allow the characterization of the conformational space of IDRs that remain significantly disordered in the bound state, with the use of programs such as Asteroids[61] or Ensemble.[62] Fragments of the IDR may retain some disorder in the complex and yet be challenging to observe directly by multidimensional NMR. In this case, CEST and CPMG experiments can be used to obtain additional information for the determination of conformational ensembles. In particular, CEST and CPMG experiments have been used to obtain a collection of conformational constraints, such as paramagnetic relaxation enhancements,[63] residual dipolar couplings,[64,65] and pseudo-contact shifts[66] of states with a small population in exchange with a major, observable state. Such data would be highly complementary to chemical shifts measured here. Combining our approach for the folded regions of the IDR and these measurements analyzed with Asteroids[61] or Ensemble,[62] will allow the determination of conformational ensembles for the entire complex including fully ordered, partially ordered, and mostly disordered regions.

Conclusions

Here, we introduce a method for the structural determination of complexes consisting of disordered protein regions that fold in part upon binding to a partner protein. This approach combines chemical-exchange NMR, structural determination of the interacting region of the disordered protein, and docking. Importantly, the method does not require the observation of the bound form of the IDP and provides structural insight on both fully and partially ordered regions in the bound IDP. We apply this method to a complex between a 96-residue disordered region of Artemis and the DBD of Ligase IV. We show that our approach correctly reproduces the crystal structure of the folded core of the complex and provides additional information on regions within the IDP that contribute to the thermodynamic affinity but retain some disorder. The method is not affected by edge effects when short peptides are used in a more traditional approach. From our NMR and ITC-based investigations, we identify additional residues at the C-terminal end of Artemis (496-501) that contribute significantly to the binding to Ligase IV. This NMR-based approach to the structural characterization of complexes of IDPs is robust and is expected to be applicable to a wide range of systems.

Experimental Section

Protein and peptide production. The expression vector for Art$^{480-575}$ encodes a 38 kDa fusion protein with an amino-terminal Glutathion S-transferase (GST) domain, a PreScission protease cleavage site, and the Art$^{480-575}$ region. Art$^{480-575}$ was expressed in *E. coli* either unlabeled in LB rich medium or uniformly labeled with ^{15}N or $^{15}N/^{13}C$ using the procedure developed by Marley *et al.*.[67] Art$^{480-575}$ was purified by glutathione-affinity chromatography (Genscript) and cleaved on a resin at 4°C with recombinant PreScission protease, followed by size-exclusion chromatography on a Superdex 75 16/60 column (GE Healthcare).

The DNA-binding domain of human Ligase IV (^{DBD}LigIV; region 1-240) was expressed as a $(His)_6$-tagged fusion protein in a pET-15b vector and purified as previously described.[35] Expression of uniformly labeled $^{15}N/^{13}C/^{2}D$ ^{DBD}LigIV was carried out from cells cultured in M9 minimum media prepared in 100% D_2O. The two Artemis peptides were produced by solid-phase peptide synthesis and purified by RP-HPLC.

Isothermal Titration Calorimetry. ITC experiments were performed on a Nano ITC calorimeter (TA Instruments) at 23 °C. Titrations were carried out with a fixed concentration of ^{DBD}LigIV (70-100 μM) into which aliquots of each of the Artemis fragments were injected. Thermodynamic parameters were determined by non-linear least-square fitting of the buffer-corrected data using the software NanoAnalyze (version 3.1.2) provided by TA Instruments.

NMR resonance assignments and Chemical Shift Perturbation (CSP) experiments. NMR spectra were acquired at 296.5 K on Bruker Avance-III HD 600 and 800 MHz spectrometers equipped with room-temperature triple resonance (^{1}H, ^{15}N, ^{13}C) probes and on a Bruker Avance-III HD 950 MHz spectrometer equipped with a triple resonance cryoprobe. All NMR data were processed with NMRPipe[68] and analyzed with NMRFAM-Sparky.[69] Backbone resonance assignments of Art$^{480-575}$ were obtained from the analysis of a series of 3D BEST triple resonances experiments[70] as well as an (H)N(COCA)NH experiment,[71,72] all recorded at 800 MHz. The BEST-TROSY versions of these 3D experiments[73] and a ^{15}N-NOESY-HSQC were recorded at 950 MHz to assign the backbone resonances of ^{DBD}LigIV. The interaction between Artemis and Ligase IV was probed by chemical shift perturbations (CSP). Uniformly labeled ^{15}N-Art$^{480-575}$ was titrated at 600 MHz with unlabeled ^{DBD}LigIV and followed with 2D ^{1}H-^{15}N HSQC spectra. Triple-labeled $^{2}D/^{15}N/^{13}C$-^{DBD}LigIV was titrated with each of the unlabeled Artemis fragments and followed with BEST-TROSY 2D ^{15}N-^{1}H correlations at 950 MHz. CSP values were calculated for each residue from the differences in $^{1}H^{N}$ ($\Delta\delta_{H^N}$) and ^{15}N ($\Delta\delta_{N}$) chemical shifts between the free and the bound states using the definition:

$$CSP = \sqrt{(\Delta\delta_{H^N})^2 + (0.1\Delta\delta_{N})^2}$$

CPMG and CEST experiments. All ^{15}N CPMG experiments[42] were carried out at a ^{1}H frequency of 800 MHz on samples of 550 μM ^{15}N-labeled Art$^{480-575}$ with 3%, 6%, 10% and 13% molar ratios of unlabeled ^{DBD}LigIV. All CEST experiments were acquired at 800 MHz using previously published pulse sequences[26,46,47] on uniformly ^{15}N- (for ^{15}N-CEST) and $^{15}N,^{13}C$- (for ^{13}C-CEST) labeled samples of Art$^{480-575}$ (550 μM) with 10% molar ratio of unlabeled ^{DBD}LigIV. Nitrogen-15 transverse relaxation rates (R_2) were measured at 800 MHz on a sample containing 300 μM ^{DBD}LigIV uniformly labeled with nitrogen-15 in complex with 350 μM unlabeled Art$^{480-575}$.

Analysis of CPMG relaxation dispersion and CEST experiments. Peak intensities in CPMG and CEST spectra, I, were quantified by fitting the lineshape of the peaks using an in-house python script (available upon request). A flowchart in

Fig. S1 presents the way the analysis was performed. ^{15}N CPMG and ^{15}N CEST data sets were simultaneously fit to a global two-site exchange model between free (F) and bound (B) states of Art$^{480-575}$ using the program ChemEx,[44] which simulates the evolution of magnetization during the CPMG and CEST periods by numerically integrating of the Bloch-McConnell equation and minimizes a standard χ^2 target function, as described previously.[26] Fitting parameters include the per residue values $\{\varpi_B, R_{2,F}, R_{2,B}, R_1\}$ and the global parameter $\{k_{on}\}$, where ϖ_B is the chemical shift of the bound state, $R_{2,F}$ ($R_{2,B}$) is the transverse spin relaxation rate of the free (bound) state, R_1 is the longitudinal spin relaxation rate that is assumed to be the same for both states and k_{on} is the association rate constant. The dissociation rate constant, k_{off}, was calculated using the relationship $k_{off} = K_d k_{on}$ assuming a dissociation constant K_d = 2.2 μM, as determined by ITC experiments. Note that the value of K_d has no influence on the analysis, provided that the binding partner (here ^{DBD}LigIV) is saturated (Fig. S11). Indeed, under these conditions the pseudo first-order rate constant is independent of K_d: $k_{on}[^{DBD}\text{LigIV}]_{free} = k_{off}\,\alpha/(1-\alpha)$, with α the proportion of ^{DBD}LigIV added to the solution (in the CEST experiment α = 0.1). ^{13}C', $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ CEST profiles were subsequently analyzed using ChemEx employing previously developed protocols.[46,74]

Structure calculations and HADDOCK docking. The structure of Art$^{485-501}$ bound to Ligase IV was calculated by molecular dynamics with CNS 1.2 in ARIA 2.2,[75] using 26 ϕ, ψ, and χ$_1$ angular restraints predicted by Talos-N[51] from the backbone chemical shifts obtained from the analysis of CEST experiments. From 30 initial structures refined in explicit water, the 10 lowest energy conformers of Art$^{485-501}$ were selected and docked onto the free structure of full-length Ligase IV (PDB 3W5O) on the HADDOCK 2.2 web-server[37] using default parameters. The X-ray structure of full-length Ligase IV in the apo form (PDB 3W5O) and the NMR conformers of Ligase IV-bound Art$^{485-501}$ were used as the starting structures. Artemis residues 485-497, for which CEST experiments revealed differences in carbon-13 chemical shifts between the free and the bound states, were considered as directly participating in the binding to Ligase IV and classified as "active" residues. The 26 ϕ, ψ, and χ1 angular restraints used to calculate the bound structure of Art$^{485-501}$ were incorporated into the docking process as unambiguous constraints. The average ligand-RMSD, as defined following CAPRI standards,[76] was calculated in the top cluster using the program PROFIT.[77] The average backbone interface-RMSD of the docking models to the X-ray structure of the Art$^{485-495}$-Ligase IV complex (PDB 3W1G) was calculated in the top cluster based on interface residues that have a heavy atom within 10 Å of any other interacting partner in the crystal structure.

ASSOCIATED CONTENT

Supporting Information

Extended experimental section; prediction of disorder; ITC parameters; assigned 2D correlations of Art$^{480-575}$ and ^{DBD}LigIV; all CPMG RD and CEST profiles for residues 486-499; secondary chemical shifts and secondary structure propensities for Art$^{480-575}$ in both free and bound states; ^{15}N transverse relaxation rates for ^{DBD}LigIV bound to Art$^{480-575}$; chemical shift changes for ^{DBD}LigIV upon Artemis binding; structure statistics for the bound form of Art$^{485-501}$; full docking results.

The Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Authors

* Ludovic.Carlier@upmc.fr; Fabien.Ferrage@ens.fr

Present Addresses

† Laboratory of Chemical Physics, NIDDK, NIH, Bethesda, MD, 20892, USA

Author Contributions

The manuscript was written through contributions of all authors.

Funding Sources

This research has received funding from the European Research Council (ERC) under the European Community's Seventh Framework Programme (FP7/2007-2013), ERC Grant agreement 279519 (2F4BIODYN) to F.F. as well as the Equipex contract ANR-10-EQPX-09. Financial support from the TGIR-RMN-THC FR 3050 CNRS is gratefully acknowledged. A. K. A. is a member of the New York Structural Biology Center, which is a STAR center supported by the New York State Office of Science, Technology, and Academic Research.

ACKNOWLEDGMENT

The authors would like to thank Patrick Fuchs (Université Paris Diderot), Damien Laage (Ecole normale superieure) and Guillaume Stirnemann (Institut de Biologie Physico-Chimique) for fruitful discussions as well as Kaushik Dutta (New York Structural Biology Center), Adrien Favier (Institut de Biologie Structurale, Grenoble) for assistance with NMR experiments, and Geoffrey Bodenhausen (Ecole Normale Supérieure) for his careful reading of the manuscript.

ABBREVIATIONS

CEST, chemical-exchange saturation transfer; CPMG, Carr-Purcell-Meiboom-Gill; CSP, chemical shift perturbations; DBD, DNA-binding domain; GST, Glutathion S-transferase; IDP, intrinsically disordered protein; IDR, intrinsically disordered region; ITC, isothermal titration calorimetry; NHEJ, non-homologous end-joining; NMR, nuclear magnetic resonance; RD, relaxation dispersion; SH3, Src homology 3.

REFERENCES

(1) van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. *Chem. Rev.* **2014**, *114*, 6589.

(2) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197.

(3) Tompa, P.; Davey, N. E.; Gibson, T. J.; Babu, M. M. *Mol. Cell* **2014**, 55, 161.

(4) Mittag, T.; Marsh, J.; Grishaev, A.; Orlicky, S.; Lin, H.; Sicheri, F.; Tyers, M.; Forman-Kay, J. D. *Structure* **2010**, *18*, 494.

(5) Tompa, P.; Fuxreiter, M. *Trends Biochem. Sci* **2008**, *33*, 2.

(6) Ragusa, M. J.; Dancheck, B.; Critton, D. A.; Nairn, A. C.; Page, R.; Peti, W. *Nat. Struct. Mol. Biol.* **2010**, *17*, 459.

(7) Ghose, R.; Shekhtman, A.; Goger, M. J.; Ji, H.; Cowburn, D. *Nat. Struct. Biol.* **2001**, *8*, 998.

(8) Stollar, E. J.; Garcia, B.; Chong, P. A.; Rath, A.; Lin, H.; Forman-Kay, J. D.; Davidson, A. R. *J. Biol. Chem.* **2009**, *284*, 26918.

(9) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632.

(10) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007**, *448*, 325.

(11) O'Connell, N.; Nichols, S. R.; Heroes, E.; Beullens, M.; Bollen, M.; Peti, W.; Page, R. *Structure* **2012**, *20*, 1746.

(12) Radhakrishnan, I.; PerezAlvarado, G. C.; Parker, D.; Dyson, H. J.; Montminy, M. R.; Wright, P. E. *Cell* **1997**, *91*, 741.

(13) Berlow, R. B.; Dyson, H. J.; Wright, P. E. *Nature* **2017**, *543*, 447.

(14) Hurley, T. D.; Yang, J.; Zhang, L.; Goodwin, K. D.; Zou, Q.; Cortese, M.; Dunker, A. K.; DePaoli-Roach, A. A. *J. Biol. Chem.* **2007**, *282*, 28874.

(15) Huang, C. D.; Rossi, P.; Saio, T.; Kalodimos, C. G. *Nature* **2016**, *537*, 202.

(16) Krishnan, N.; Koveal, D.; Miller, D. H.; Xue, B.; Akshinthala, S. D.; Kragelj, J.; Jensen, M. R.; Gauss, C. M.; Page, R.; Blackledge, M.; Muthuswamy, S. K.; Peti, W.; Tonks, N. K. *Nat. Chem. Biol.* **2014**, *10*, 558.

(17) Marsh, J. A.; Dancheck, B.; Ragusa, M. J.; Allaire, M.; Forman-Kay, J. D.; Peti, W. *Structure* **2010**, *18*, 1094.

(18) Khan, S. N.; Charlier, C.; Augustyniak, R.; Salvi, N.; Déjean, V.; Bodenhausen, G.; Lequin, O.; Pelupessy, P.; Ferrage, F. *Biophys. J.* **2015**, *109*, 988.

(19) Abyzov, A.; Salvi, N.; Schneider, R.; Maurin, D.; Ruigrok, R. W. H.; Jensen, M. R.; Blackledge, M. *J. Am. Chem. Soc.* **2016**, *138*, 6240.

(20) Ferreon, J. C.; Martinez-Yamout, M. A.; Dyson, H. J.; Wright, P. E. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 13260.

(21) Martinez, A. I. C.; Weinhaupl, K.; Lee, W. K.; Wolff, N. A.; Storch, B.; Zerko, S.; Konrat, R.; Kozminski, W.; Breuker, K.; Thevenod, F.; Coudevylle, N. *J. Biol. Chem.* **2016**, *291*, 2917.

(22) Kragelj, J.; Palencia, A.; Nanao, M. H.; Maurin, D.; Bouvignies, G.; Blackledge, M.; Jensen, M. R. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 3409.

(23) Charlier, C.; Cousin, S. F.; Ferrage, F. *Chem. Soc. Rev.* **2016**, *45*, 2410.

(24) Loria, J. P.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1999**, *121*, 2331.

(25) Hansen, D. F.; Vallurupalli, P.; Lundstrom, P.; Neudecker, P.; Kay, L. E. *J. Am. Chem. Soc.* **2008**, *130*, 2667.

(26) Vallurupalli, P.; Bouvignies, G.; Kay, L. E. *J. Am. Chem. Soc.* **2012**, *134*, 8148.

(27) Fawzi, N. L.; Ying, J.; Ghirlando, R.; Torchia, D. A.; Clore, G. M. *Nature* **2011**, *480*, 268.

(28) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021.

(29) Schneider, R.; Maurin, D.; Communie, G.; Kragelj, J.; Hansen, D. F.; Ruigrok, R. W. H.; Jensen, M. R.; Blackledge, M. *J. Am. Chem. Soc.* **2015**, *137*, 1220.

(30) Bouvignies, G.; Vallurupalli, P.; Hansen, D. F.; Correia, B. E.; Lange, O.; Bah, A.; Vernon, R. M.; Dahlquist, F. W.; Baker, D.; Kay, L. E. *Nature* **2011**, *477*, 111.

(31) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. *Science* **2010**, *329*, 1312.

(32) Vallurupalli, P.; Hansen, D. F.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 11766.

(33) Huang, Y. Q.; Liu, Z. R. *J. Mol. Biol.* **2009**, *393*, 1143.

(34) Moshous, D.; Callebaut, I.; de Chasseval, R.; Corneo, B.; Cavazzana-Calvo, M.; Le Deist, F.; Tezcan, I.; Sanal, O.; Bertrand, Y.; Philippe, N.; Fischer, A.; de Villartay, J. P. *Cell* **2001**, *105*, 177.

(35) De Ioannes, P.; Malu, S.; Cortes, P.; Aggarwal, A. K. *Cell Reports* **2012**, *2*, 1505.

(36) Ochi, T.; Gu, X.; Blundell, Tom L. *Structure* **2013**, *21*, 672.

(37) van Zundert, G. C. P.; Rodrigues, J.; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A. *J. Mol. Biol.* **2016**, *428*, 720.

(38) Deng, X.; Eickholt, J.; Cheng, J. *BMC Bioinform.* **2009**, *10*, 436.

(39) Malu, S.; De Ioannes, P.; Kozlov, M.; Greene, M.; Francis, D.; Hanna, M.; Pena, J.; Escalante, C. R.; Kurosawa, A.; Erdjument-Bromage, H.; Tempst, P.; Adachi, N.; Vezzoni, P.; Villa, A.; Aggarwal, A. K.; Cortes, P. *J. Exp. Med.* **2012**, *209*, 955.

(40) Marsh, J. A.; Singh, V. K.; Jia, Z. C.; Forman-Kay, J. D. *Protein Sci.* **2006**, *15*, 2795.

(41) Sekhar, A.; Rosenzweig, R.; Bouvignies, G.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 10395.

(42) Hansen, D. F.; Vallurupalli, P.; Kay, L. E. *J. Phys. Chem. B* **2008**, *112*, 5898.

(43) Forsen, S.; Hoffman, R. A. *J. Chem. Phys.* **1963**, *39*, 2892.

(44) Bouvignies, G.

(45) Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. *Science* **2002**, *295*, 1719.

(46) Vallurupalli, P.; Kay, L. E. *Angew. Chem.-Int. Edit.* **2013**, *52*, 4156.

(47) Long, D.; Sekhar, A.; Kay, L. E. *J. Biomol. NMR* **2014**, *60*, 203.

(48) Tamiola, K.; Acar, B.; Mulder, F. A. A. *J. Am. Chem. Soc.* **2010**, *132*, 18000.

(49) Wishart, D. S.; Bigam, C. G.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135.

(50) Wishart, D. S.; Sykes, B. D. *J. Biomol. NMR* **1994**, *4*, 171.

(51) Shen, Y.; Bax, A. *J. Biomol. NMR* **2013**, *56*, 227.

(52) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12366.

(53) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **1998**, *54*, 905.

(54) Sekhar, A.; Rumfeldt, J. A. O.; Broom, H. R.; Doyle, C. M.; Bouvignies, G.; Meiering, E. M.; Kay, L. E. *Elife* **2015**, *4*.

(55) Piserchio, A.; Ramakrishan, V.; Wang, H.; Kaoud, T. S.; Arshava, B.; Dutta, K.; Dalby, K. N.; Ghose, R. *Scientific Reports* **2015**, *5*.

(56) Shen, Y.; Bax, A. *J. Biomol. NMR* **2010**, *48*, 13.

(57) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 10420.

(58) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155.

(59) Tsai, C. J.; Lin, S. L.; Wolfson, H. J.; Nussinov, R. *Protein Sci.* **1997**, *6*, 53.

(60) Wiesnerl, S.; Sprangers, R. *Curr. Opin. Struct. Biol.* **2015**, *35*, 60.

(61) Salmon, L.; Nodet, G.; Ozenne, V.; Yin, G. W.; Jensen, M. R.; Zweckstetter, M.; Blackledge, M. *J. Am. Chem. Soc.* **2010**, *132*, 8407.

(62) Krzeminski, M.; Marsh, J. A.; Neale, C.; Choy, W. Y.; Forman-Kay, J. D. *Bioinformatics* **2013**, *29*, 398.

(63) Sekhar, A.; Rosenzweig, R.; Bouvignies, G.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, E2794.

(64) Vallurupalli, P.; Hansen, D. F.; Stollar, E.; Meirovitch, E.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 18473.

(65) Zhao, B.; Zhang, Q. *J. Am. Chem. Soc.* **2015**, *137*, 13480.

(66) Chen, J. L.; Wang, X.; Yang, F.; Cao, C.; Otting, G.; Su, X. C. *Angew. Chem.-Int. Edit.* **2016**, *55*, 13744.

(67) Marley, J.; Lu, M.; Bracken, C. *J. Biomol. NMR* **2001**, *20*, 71.

(68) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(69) Lee, W.; Tonelli, M.; Markley, J. L. *Bioinformatics* **2015**, *31*, 1325.
(70) Schanda, P.; Van Melckebeke, H.; Brutscher, B. *J. Am. Chem. Soc.* **2006**, *128*, 9042.
(71) Bracken, C.; Palmer, A. G.; Cavanagh, J. *J. Biomol. NMR* **1997**, *9*, 94.
(72) Sun, Z. Y. J.; Frueh, D. P.; Selenko, P.; Hoch, J. C.; Wagner, G. *J. Biomol. NMR* **2005**, *33*, 43.
(73) Lescop, E.; Kern, T.; Brutscher, B. *J. Magn. Reson.* **2010**, *203*, 190.
(74) Bouvignies, G.; Vallurupalli, P.; Kay, L. E. *J. Mol. Biol.* **2014**, *426*, 763.
(75) Rieping, W.; Habeck, M.; Bardiaux, B.; Bernard, A.; Malliavin, T. E.; Nilges, M. *Bioinformatics* **2007**, *23*, 381.
(76) Janin, J. *Mol. Biosyst.* **2010**, *6*, 2351.
(77) Martin, A. C. R., PROFIT, http://www.bioinf.org.uk/software/profit/

Table of Contents graphic.