INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus
Chimie

Chafiaa Bouguechtouli, Rania Ghouil, Ania Alik, Florent Dingli, Damarys Loew and Francois-Xavier Theillet

Structural characterization of stem cell factors Oct4, Sox2, Nanog and Esrrb disordered domains, and a method to detect phospho-dependent binding partners

Published online: 19 April 2024

Part of Special Issue: Breaking Barriers in Chemical Biology – Toulouse 2022

Guest editors: Marie Lopez (CNRS-Univ. Montpellier-ENSCM, IBMM, Montpellier, France), Elisabetta Mileo (Aix-Marseille Univ, CNRS, BIP, IMM, Marseille, France), Eric Defrancq (Univ. Grenoble-Alpes-CNRS, DCM, Grenoble, France), Agnes Delmas (CNRS, CBM, Orléans, France), Boris Vauzeilles (CNRS-Univ. Paris-Saclay, ICSN, Gif-sur-Yvette, France), Dominique Guianvarch (CNRS-Univ. Paris-Saclay, ICMMO, Orsay, France) and Christophe Biot (CNRS-Univ. Lille, UGSF, Lille, France)

https://doi.org/10.5802/crchim.272

The Comptes Rendus. Chimie *are a member of the Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1878-1543

Comptes Rendus
Chimie
Published online: 19 April 2024
https://doi.org/10.5802/crchim.272

Breaking Barriers in Chemical Biology – Toulouse 2022

Structural characterization of stem cell factors Oct4, Sox2, Nanog and Esrrb disordered domains, and a method to detect phospho-dependent binding partners

Chafiaa Bouguechtouli [a], Rania Ghouil [a], Ania Alik [a], Florent Dingli [b], Damarys Loew [b] and Francois-Xavier Theillet [*, a]

[a] Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

[b] Institut Curie, PSL Research University, Centre de Recherche, CurieCoreTech Spectrométrie de Masse Protéomique, Paris cedex 05, France

Current addresses: Structural Motility, Institut Curie, Paris Université Sciences et Lettres, Sorbonne Université, CNRS UMR144, 75005 Paris, France (C. Bouguechtouli), Université de Paris, Institut Cochin, CNRS UMR8104, INSERM U1016, Paris, France (A. Alik)

E-mail: francois-xavier.theillet@cnrs.fr (F.-X. Theillet)

Abstract. The combined expression of a handful of pluripotency transcription factors (PluriTFs) in somatic cells can generate induced pluripotent stem cells (iPSCs). Here, we report the structural characterization of disordered regions contained in four important PluriTFs, namely Oct4, Sox2, Nanog and Esrrb. Moreover, many post-translational modifications (PTMs) have been detected on PluriTFs, whose roles are not yet characterized. To help in their study, we also present a method (i) to produce well-characterized phosphorylation states of PluriTFs, using NMR analysis, and (ii) to use them for pull-downs in stem cell extracts analyzed by quantitative proteomics to detect potential Sox2 binders.

Keywords. Pluripotency transcription factors, Intrinsically disordered proteins, Post-translational modifications, Kinases, NMR, Proteomics, Quantitative mass spectrometry.

Funding. CNRS and CEA-Saclay, French Infrastructure for Integrated Structural Biology (FRISBI, https://frisbi.eu/, grant number ANR-10-INSB-05-01) and French National Research Agency (ANR; research grants ANR-14-ACHN-0015 and ANR-20-CE92-0013), IR INFRANALYTICS FR2054, "Région Ile-de-France" and "Fondation pour la Recherche Médicale".

Manuscript received 3 March 2023, revised 9 November 2023, accepted 10 November 2023.

1. Introduction

The possibility of reprogramming somatic cells to an induced pluripotency state was revealed in the 2000s, giving great hopes in the fields of Biology and Medicine [1–3]. Induced pluripotent stem cells (iPSCs) and embryonic stem cells (ESCs) are characterized by the active state of a pluripotency network, whose core comprises the pluripotency transcription factors (PluriTFs) Oct4, Sox2, Nanog and Esrrb (OSNE). These bind to enhancer sequences and thus activate or repress, or even "bookmark" during mitosis, a wealth of genes related to pluripotency or

* Corresponding author.

cell differentiation [4–9]. Consistently, their misregulation correlates with cancer malignancy and stemness [10–14].

Comprehensive structural descriptions of OSNE are still missing to the best of our knowledge. The folded DNA-binding domains (DBDs) of OSNE have been structurally characterized in complex with their DNA target sequences [15–19] together with the ligand-binding domain (LBD) of Esrrb [20]. Recent studies have depicted even splendid structures of Oct4's and Sox2's DBDs bound to nucleosomes, hence deciphering their "pioneer factor" abilities [21–29]. Another structure of Sox2 bound to the importin Impα3 has also been published, showing how its two Nuclear Localization Sequences (NLSs) flanking the DBD are involved in Sox2 nuclear import [30]. The other segments of OSNE have been predicted to be intrinsically disordered regions of proteins (IDRs) [31], i.e., they should have no stable tertiary fold when isolated [32–36].

These IDRs appear to have important roles in binding partners involved in epigenetic reprogramming, chromatin reorganization, and in recruiting transcription or repression machineries [5,37–41]. These functions are poorly understood, and, to the best of our knowledge, no experimental characterization of the structural behavior of these regions in N- and C-terminal of DBDs has been released yet. Recent studies have shown that C-terminal regions of Oct4 and Sox2 are important for their reprogramming capacities [42,43], notably by contributing to the engagement in molecular phase-separated condensates with the Mediator complex [44]. More generally, the activating or repressive activities of IDRs of transcription factors (TFs) have been scarcely studied at the structural level: these segments are thought to contain hydrophobic patches flanked by acidic amino acids, which favors DNA-binding specificity, phase separation, and low-specificity interactions, notably with the Mediator subunit Med15 [44–53]; more specific interactions have been described in some cases [45,54,55].

Post-translational modifications (PTMs) add a layer of complexity by often regulating IDRs' interactions [32,33,56,57] and notably TFs' activity [58–61]. PTMs are classical carriers of cell signaling by regulating the stability and the interactions of proteins. An increasing number of PTMs have been described on OSNE's IDRs in the recent years [5,37,38,62–77], notably phosphorylation by cyclin-dependent kinases (CDKs) [66,78–87] or by mitogen-activated protein kinases (MAPKs) [83,88,89], or their complementary Ser/Thr O-GlcNAcylation by OGT [65,90–96].

In order to prompt future studies on this topic, we describe here a feasibility study (i) for producing well-characterized samples made of post-translationally modified IDRs of PluriTFs and (ii) to use these as baits in pull-down assays for detecting PTMs' related binding partners. Hence, we characterized some of the phosphorylation reactions of Esrrb and Sox2 by p38α/β, Erk2 and Cdk1/2. Then, we showed that biotinylated chimera of Sox2 and Esrrb coupled to an AviTag peptide could be attached to streptavidin-coated beads. Finally, we loaded truncated segments of the C-terminal IDR of Sox2 (phosphorylated or not) on these beads, and exposed them to extracts of mouse ESCs (mESCs) in pull-down assays, which we analyzed using quantitative mass spectrometry-based proteomics. Among the quantified (phospho-)Sox2 binders, we verified the phospho-dependent interaction between the proline cis-trans isomerase Pin1 and Sox2 using NMR spectroscopy.

2. Material and methods

2.1. *Production of recombinant fragments of Oct4, Sox2, Nanog and Esrrb*

We used human protein sequences, unless specified. Codon-optimized (for expression in *Escherichia coli*) genes coding for human Oct4(aa1–145) and Oct4(aa286–380) were synthesized in the context of larger genes coding for Tev–Oct4(aa1–145)–Tev–GB1 and Tev–Oct4(aa286–380)–Tev–GB1 by Genscript and cloned into pET-41a(+) vector between SacII and HindIII restriction sites, hence permitting the expression of GST–His6–Tev1–Oct4(aa1–145)–Tev2–GB1 and GST–His6–Tev1–Oct4(aa286–380)–Tev2–GB1; Tev1 and Tev2 are the heptapeptide ENLYFQG cleavage site of the TEV protease, Tev2 is separated by GAGGAGG from GB1 (T2Q variant of the immunoglobulin binding domain B1 of the protein G from group G Streptococcus [97,98]). The C-terminal GB1 tag was added to avoid any C-terminal proteolysis of the IDR of interest during the expression and the first purification steps; we did not test constructs without this supplementary folded domain,

whose necessity for the stability of the IDR is thus not proven.

The same rationale (cDNA synthesis, cloning, vectors, chimera constructs) was used for producing Nanog(aa154–305), Nanog(aa154–215), Nanog(aa154–272), Nanog(aa154–305_C185A-C227A-C243A-C251A), Nanog(aa154–272_C185A-C227A-C243A-C251A), and a very similar rationale (chimera constructs missing the C-terminal Tev2–GB1) for Sox2(aa1–42), Sox2(aa115–317_C265A), Sox2(aa115–187), Sox2(aa115–236), Sox2(aa115–282_C265A), Esrrb(aa1–102_C12A-C72A-C91A), Esrrb(aa1–102_C12A-C91A), Nanog(aa1–85) (this latter was cloned in the MfeI/HindIII restriction sites from pET-41a(+)).

The recombinant production and the purification of the protein constructs followed the procedures described previously [99], using the soluble fraction of bacterial lysates, except for the constructs containing the Sox2 C-terminal fragments. These latter constructs were recovered from the insoluble fractions of the lysates and resolubilized in 8 M urea; these were submitted to a His-tag purification in urea, and the last size-exclusion chromatography (SEC) had to be carried out in 2 M urea, which avoided clogging of the column and permitted obtaining regular elution peak widths (these were otherwise extremely broad, up to 100 mL for the longest Sox2(aa115–317_C265A) construct). The samples were concentrated and stored at −20 °C, and thawed just before the NMR experiments. The Sox2 samples containing 2 M urea were submitted to 2–3 cycles of concentration/dilution in Hepes at 20 mM, NaCl at 75 mM to generate samples in urea at 0.25 or 0.125 M.

All purification steps were carefully carried out at 4 °C; protein eluates from every purification step were immediately supplemented with protease inhibitors (EDTA-free cOmplete, Roche) (together with DTT at 10 mM for cysteine-containing protein constructs), before being submitted to a concentration preparing the next purification step.

Chimera constructs of Sox2's and Esrrb's IDR fragments containing a 15-mer peptide AviTag GLNDIFEAQKIEWHE were produced using procedures similar to those described earlier for OSNE constructs. The construct Sox2(aa234–317)–AviTag–His6 was soluble and did not require to be purified in urea.

More details about the production of OSNE peptides are given in the Supplementary Material.

2.2. *Production of the biotin ligase BirA and specific biotinylation of the AviTag–peptide chimera*

The biotin ligase BirA was produced using recombinant production in *E. coli* BL21(DE3)Star transformed with a pET21-a(+) plasmid containing a gene coding for BirA cloned at EcoRI and HindIII restriction sites. pET21a-BirA was a gift from Alice Ting (Addgene plasmid #20857) [100]. The expression was carried out overnight at 20 °C in a Luria–Bertani culture medium. The construct contained a His6 tag in C-terminal and was purified using a two-step purification procedure including a His-trap followed by a SEC. Details about the production of BirA are given in the Supplementary Material.

The biotinylation was executed using a rationale inspired by a published protocol [101], at room temperature during 90 min, in samples containing the AviTag–chimera of interest at 100 μM and BirA at 0.7 μM in a buffer containing ATP at 2 mM, biotin at 600 μM, $MgCl_2$ at 5 mM, DTT at 1 mM, HEPES at 50 mM, NaCl at 150 mM, protease inhibitors (final concentration 1×, EDTA-free cOmplete, Roche), at pH 7.0. To remove some possible proteolyzed peptides and BirA, the biotinylated constructs were purified using a SEC in a column (Superdex 16/60 75 pg, Cytiva) preequilibrated with a buffer containing phosphate at 20 mM, NaCl at 150 mM at pH 7.4 (buffer called thereafter Phosphate Buffer Saline, PBS). The eluted fractions of interest were concentrated and stored at −20 °C.

2.3. *Assignment of NMR signals from OSNE fragments and structural propensities*

The assignment strategy was the same as in previous reports from our laboratory [99]. The ^{15}N relaxation data were recorded and analyzed according to the methods described in previous reports [102]. Details are given in the Supplementary Material.

Disorder prediction was calculated using the ODINPred website (https://st-protein.chem.au.dk/odinpred) [103]. Experimental secondary structure propensities of unmodified OSNE peptides were obtained using the neighbor-corrected structural propensity calculator ncSCP [104,105] (http://www.protein-nmr.org/, https://st-protein02.chem.au.dk/ncSPC/) from the experimentally determined, DSS-

referenced Cα and Cβ chemical shifts as input, with a correction for Gly–Pro motifs (−0.77 ppm instead of −2.0 ppm) [106]. We also used the δ2D method to get requested verifications of the experimental secondary structure propensities [107]. Some signals were too weak in 3D spectra from Sox2(aa115–317_C265A) recorded at 950 MHz, and their chemical shifts were not defined. In these cases, chemical shifts from 3D spectra of Sox2(115-236) or His6–AviTag–Sox2(aa234–317_C265A) were used to complete the lists of chemical shifts used to calculate the chemical shift propensities shown in Figure 2.

$^{1}H_N/^{15}N/^{13}Ca/^{13}Cb/^{13}CO$ NMR assignments of OSNE peptides, together with the corresponding experimental details, have been deposited in the Biological Magnetic Resonance Data Bank (BMRB) with accession numbers 51534 (Sox2_aa1–42), 51717 (Esrrb_aa1–102), 51756 (Oct4_aa1–145), 51758 (Oct4_aa286–360), 51782 (His6–AviTag–Sox2_aa234–317_C265A), and 51780 (Nanog_aa1–85).

2.4. *NMR monitoring of phosphorylation reactions and production of phosphorylated peptides*

We performed the phosphorylation kinetics presented in Figure 4a using commercial recombinant kinases GST–p38β at 10 μg/mL (Sigma-Aldrich, ref. B4437), GST–Erk2 at 20 μg/mL (Sigma-Aldrich, ref. E1283), GST–Cdk1/CyclinA2 at 20 μg/mL (Sigma-Aldrich, stock ref. C0244), and GST–Cdk2/CyclinA2 at 20 μg/mL (Sigma-Aldrich, ref. C0495). Then, we used kinases produced in-house in *E. coli*, using plasmids containing optimized genes coding for p38α(aa1–360, full-length) and Erk2(aa8–360); these were produced, activated, and purified in house as described previously [99]; in-house p38α was used at 40 μg/mL for the experiments shown in Figure 4. Indeed, the limited activities and high costs of commercial kinases motivated us to develop in-house capacities in kinase production. p38α was the most accessible to produce among the MAPKs and CDKs; we produced it and activated it using recombinant MKK6.

Phosphorylation reactions were carried out using ^{15}N-labeled IDRs at 50 μM, in Hepes 20 mM, NaCl 50 mM, DTT or TCEP at 4 mM, ATP 1.5 mM, $MgCl_2$ at 5 mM, protease inhibitors (Roche), 7.5% D_2O, pH 6.8 at 25 °C in 100 μL using 3 mm diameter Shigemi tubes. We monitored the phosphorylation kinetics by recording time series of $^{1}H–^{15}N$ SOFAST-HMQC spectra at 600 or 700 MHz, and by quantifying the NMR signal intensities of the disappearing unphospho- and appearing phospho-residues. We applied the methods that we described in earlier publications [108–111]. More details are given in the Supplementary Material.

2.5. *Pull-down assays*

The mESCs extracts were obtained from mESCs cultured in the conditions previously described [112]. Homogeneous extracts were obtained using DNA shearing by sonication in the presence of benzonase, as described by Gingras and colleagues [113].

The pull-down assays were executed using 25 μL of streptavidin-coated magnetic beads (Magbeads streptavidine, Genscript) loaded with 1 nmol of the biotinylated bait-peptides of interest. These were incubated for one hour at room temperature with mESCs extracts, washed in PBS and eluted using a 2× Laemmli buffer. Details are given in the Supplementary Material.

2.6. *Mass spectrometry-based proteomics analysis of pull-down assays*

The pull-down samples were treated on-beads by trypsin/LysC (Promega). The resulting peptides were loaded and separated on a C18 column for online liquid chromatography performed with an RSLCnano system (Ultimate 3000, Thermo Scientific) coupled to an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). Maximum allowed mass deviation was set to 10 ppm for monoisotopic precursor ions and 0.6 Da for MS/MS peaks. The resulting files were further processed using myProMS v3.9.3 (https://github.com/bioinfo-pf-curie/myproms; Poullet et al. [114]). False-discovery rate (FDR) was calculated using Percolator [115] and was set to 1% at the peptide level for the whole study. Label-free quantification was performed using extracted-ion chromatograms (XICs) of peptides, computed with MassChroQ [116] v.2.2.1. The complete details are given in the Supplementary Material.

2.7. *Recombinant production of Pin1 and NMR analysis of its interaction with Sox2 or phospho-Sox2*

The plasmid containing the gene coding for the Pin1–WW domain was a kind gift from Isabelle Landrieu. The production was executed according to the previously published protocol [117]. NMR analysis of the binding with phospho-Sox2(aa115–240) was performed at 283 K and pH 7.0 with the GST–Pin1–WW construct and ^{15}N-labeled Sox2(aa115–240) mixed in stoichiometric proportions, either at 50 or 10 μM for non-phospho- and phospho-Sox2, respectively. The details on the NMR acquisition, processing, and analysis are given in the Supplementary Information.

3. Results

3.1. *Structural characterization of the N- and C-terminal regions of Oct4*

We produced and purified protein constructs containing the fragments of human Oct4(aa1–145) and Oct4(aa286–360), which were both predicted to be mostly disordered (Supplementary material 3.2.3). The 2D ^{1}H–^{15}N HSQC NMR spectra showed cross-peaks in the region where random coil peptides resonances are usually found (Figure 1). We assigned the backbone NMR signals of $^{1}H_N$, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO for both segments, which permitted to calculate experimentally derived secondary structure propensities (Figures 1c, f, Supplementary Figure S4). We did not find any sign of a stable secondary structure, the highest α-helical propensities reaching about 25% in short stretches of about 5 consecutive amino acids. Hence, we verified experimentally that these N- and C-terminal fragments of human Oct4 are IDRs. We noticed that one motif RTWLSF (aa33–38) generates cross-peaks out of the random coil area in the 2D ^{1}H–^{15}N HSQC, and its chemical shifts reveal the strongest α-helical propensity (about 25% according to two distinct software products, see Figures 1c, f, Supplementary Figure S4). This might typically correspond to an interaction site: IDRs' binding motifs often adopt secondary structure in complexes, whose formation is energetically favored by local conformational preferences for the bound structure in the free state [57,106,118].

We can highlight the fact that Oct4's IDRs contain a high density of prolines, which are not directly detectable in the present $^{1}H_N$-detected experiments, even though most of the $^{13}C\alpha$, $^{13}C\beta$, and ^{13}CO resonances were characterized via HNCAB and HNCO experiments. We have shown previously that the ^{13}C-detected experiments $^{13}C\alpha$ ^{13}CO permitted observing all these Pro residues in Oct4(aa1–145) [99], whose chemical shifts were those of random coil peptides.

3.2. *Structural characterization of the N- and C-terminal regions of Sox2*

We produced and purified peptide fragments of human Sox2, namely Sox2(aa1–42), Sox2(aa115–187), Sox2(aa115–236), Sox2(aa115–282), Sox2(aa234–317_C265A), and Sox2(aa115–317_C265A). We also produced and purified chimera peptides His6–AviTag–Sox2(aa115–240) and His6–AviTag–Sox2(aa234–317_C265A). These were all predicted to be disordered (Supplementary material 3.3.3).

We had solubility issues with all of them but Sox2(aa1–42), Sox2(aa234–317_C265A), and His6–AviTag–Sox2(aa234–317_C265A). We had to recover these troublesome peptides from the insoluble fraction of the bacterial extract after overexpression at 37 °C. We even had to carry out our final SEC purification step in a buffer containing urea at 2 M (at 4 °C) for Sox2(aa115–236), Sox2(aa115–282), His6–AviTag–Sox2(aa115–240), and Sox2(aa115–317_C265A). The assignments of these latter constructs were achieved in 0.25–0.5 M urea, after executing 2 to 3 concentration–dilution steps. Aggregates were forming during the acquisition, which made the assignment rather painful. This behavior correlated with liquid–liquid phase separation (LLPS) propensities (Supplementary Figure S5), which we observed a few months before such a behavior was reported by Young and collaborators [44]. The assignment of Sox2(aa115–317_C265A) was possible only at 950 MHz with the help of the previously assigned smaller fragments Sox2(aa115–236) and His6–AviTag–Sox2(aa234–317_C265A). Some stretches of amino acids were particularly difficult to observe in 3D spectra, e.g., the region aa160–185, because of an apparent fast T2 relaxation. We may investigate these phenomena in later reports.

We observed cross-peaks in the 2D ^{1}H–^{15}N HSQC NMR spectra that were all resonating in the spectral region of random coil peptides' resonances

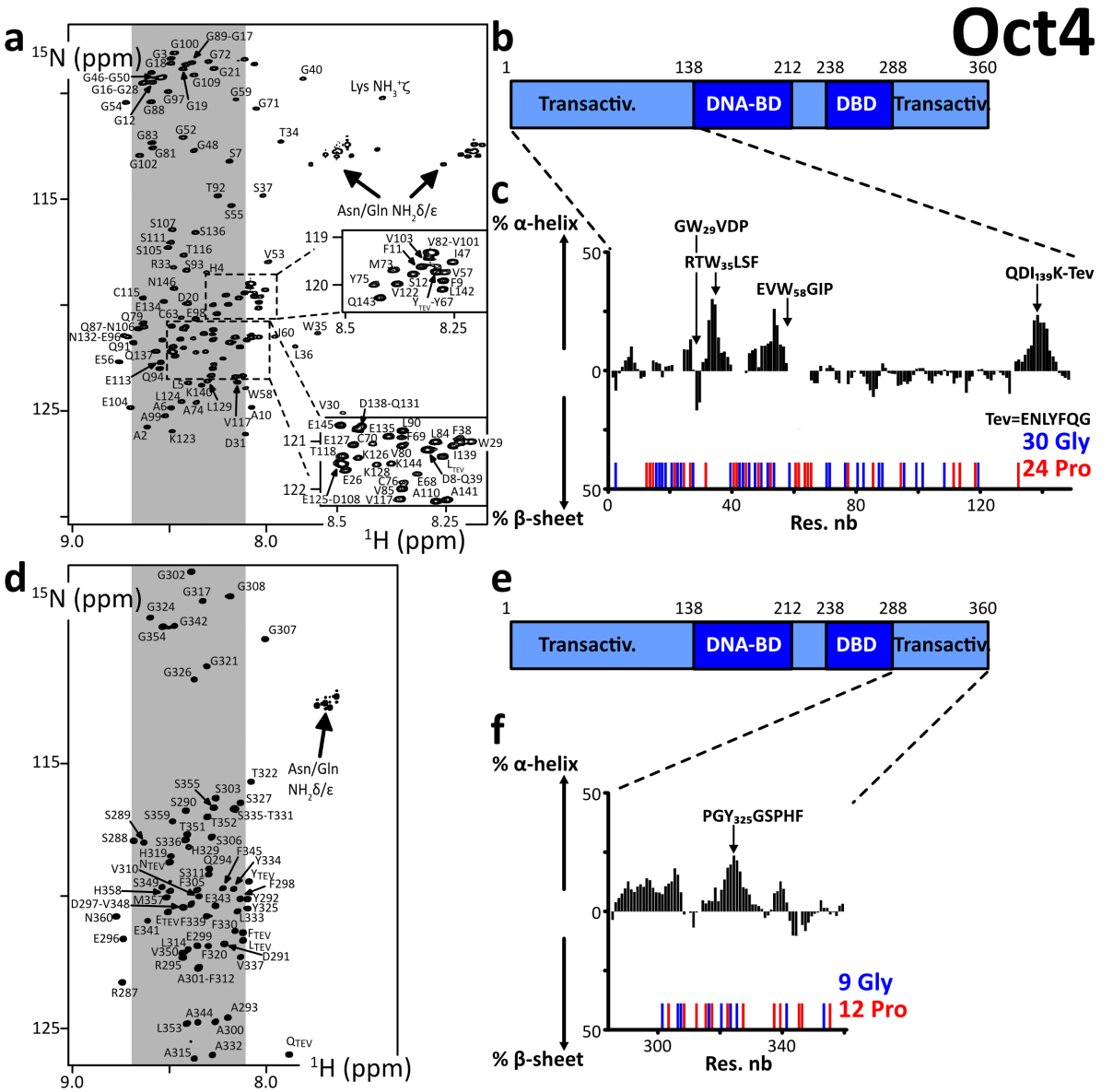

Figure 1. (a,d) 2D ^{1}H–^{15}N HSQC spectra of the N- and C-terminal IDRs of human Oct4, the labels indicating the assignments; the grey areas show the spectral regions where random coil amino acids resonate usually; (b,e) primary structures of Oct4; dark and light colors indicate the folded and disordered domains, respectively; blue and red sticks indicate the positions of Glycines and Prolines, respectively; (c,f) secondary structure propensities calculated from the experimental chemical shifts of the peptide backbone Cα and Cβ, using the ncSPC algorithm [104,105].

(Figure 2). The spectra of the short fragments of the C-terminal region of Sox2 were exactly overlapping with those of Sox2(aa115–317_C265A) (Supplementary Figure S1). This shows that all these fragments have very similar local conformational behaviors, a phenomenon regularly observed with IDRs. We assigned the backbone NMR signals of $^{1}H_N$, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO for Sox2(aa1–42), Sox2(aa115–236), His6–AviTag–Sox2(aa234–317_C265A) and partially for Sox2(aa115–317_C265A). We aggregated the lists

of chemical shifts of the C-terminal fragments and used them to calculate the experimental secondary structure propensities (Figures 2c, f and Supplementary Figure S4). This confirmed the absence of any stable secondary structure elements in Sox2 N- and C-terminal region. The C-terminal region is poorly soluble below 0.25 M urea; this should not affect a stable fold, so we can affirm that these regions of Sox2 are experimentally proven IDRs.

3.3. *Structural characterization of the N-terminal region of Nanog*

We produced and purified the N-terminal peptide fragment of human Nanog(aa1–85). All our attempts to purify C-terminal regions of Nanog failed, even after alanine-mutation of cysteines in Nanog(aa154–305), Nanog(aa154–272), and Nanog(aa154–215). We managed to resolubilize our construct GST–His6–Tev–Nanog(aa154–305) from the insoluble fraction of the bacterial extract, to partially purify it and cleave it using the TEV protease. However, the resulting Nanog(aa154–305) peptide was barely soluble in a detergent (NP-40 at 2% v/v), and not in high-salt buffers, or not even in the presence of urea at 4 M. The 10 tryptophane residues are probably playing a role in this behavior, in the context of a primary structure containing not enough hydrophobic amino acids favoring stable folds.

The cross-peaks of Nanog(aa1–85) in the 2D ^{1}H–^{15}N HSQC NMR spectrum were all in the spectral region of random coil peptides' resonances (Figure 3). The assignment of the backbone NMR signals of $^{1}H_N$, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO permitted calculating experimental secondary structure propensities, which were low through the whole peptide (Figure 3b, Supplementary Figure S4). Thus, we confirmed that this Nanog N-terminal is an IDR.

3.4. *Structural characterization of the N-terminal region of Esrrb*

We produced and purified the N-terminal fragment of human Esrrb(aa1–102), which was predicted to be disordered (Supplementary material 3.5.3), in the alanine-mutated versions Esrrb(aa1–102_C12A-C72A-C92A) and Esrrb(aa1–102_C12A-C92A). This was a strategic choice to attenuate the formation of disulfide bonds; the wild-type N-terminal fragments might however be workable too. Mutating cysteines permitted working in more comfortable conditions and maintaining our construct monomeric for longer periods of time in the next phosphorylation and biotinylation experiments. Cysteines are indeed highly solvent-accessible in IDRs and they are consequently difficult to keep in their thiol, non-disulfide forms, even in the presence of fresh DTT or TCEP at neutral pH. We also produced chimera constructs Esrrb(aa1–102_C12A-C72A-C92A)–AviTag–His6 and Esrrb(aa1–102_C12A-C72A)–AviTag–His6.

All the 2D ^{1}H–^{15}N HSQC NMR spectra revealed cross-peaks in the spectral region of random coil peptides (Figure 3). These spectra are overlapping to a large extent, confirming the weak influence of the mutations of cysteines: the mutation Cys72Ala has almost no consequences on the chemical shifts, below 0.05 ppm even for the neighboring amino acids (Supplementary Figure S2a,b); the mutation Cys91Ala has more impact, with chemical shifts perturbations of about 0.1 ppm for the next 5 amino acids (Supplementary Figure S2c,d), which is at least partially due to the fact that the Ala substitution favors an increase in local α-helicity (about 25%, see Supplementary Figure S2e,f). This N-terminal fragment of human Esrrb is thus an IDR, according to the calculated secondary structure propensities (Figure 3f, Supplementary Figures S2e,f, S4).

3.5. *Phosphorylation of Esrrb and Sox2 by p38α, Erk2, Cdk1/2 as monitored by NMR spectroscopy*

We reported recently the site-specific phosphorylation kinetics of Oct4 by p38α using ^{13}C-direct NMR detection [99]. Here, we used more standard ^{1}H-detected/^{15}N-filtered experiments to rapidly characterize the site preferences of MAPKs and CDKs on Esrrb and Sox2, which we thought to use as baits for performing phospho-dependent pull-down assays (see below).

To start with, we used commercial aliquots of MAPKs, namely p38β and Erk2, and CDKs, namely Cdk1/CyclinA2 and Cdk2/CyclinA2, on Esrrb(aa1–102). We observed the progressive phosphorylation of its three Ser–Pro motifs (at Ser22, Ser34 and Ser58)

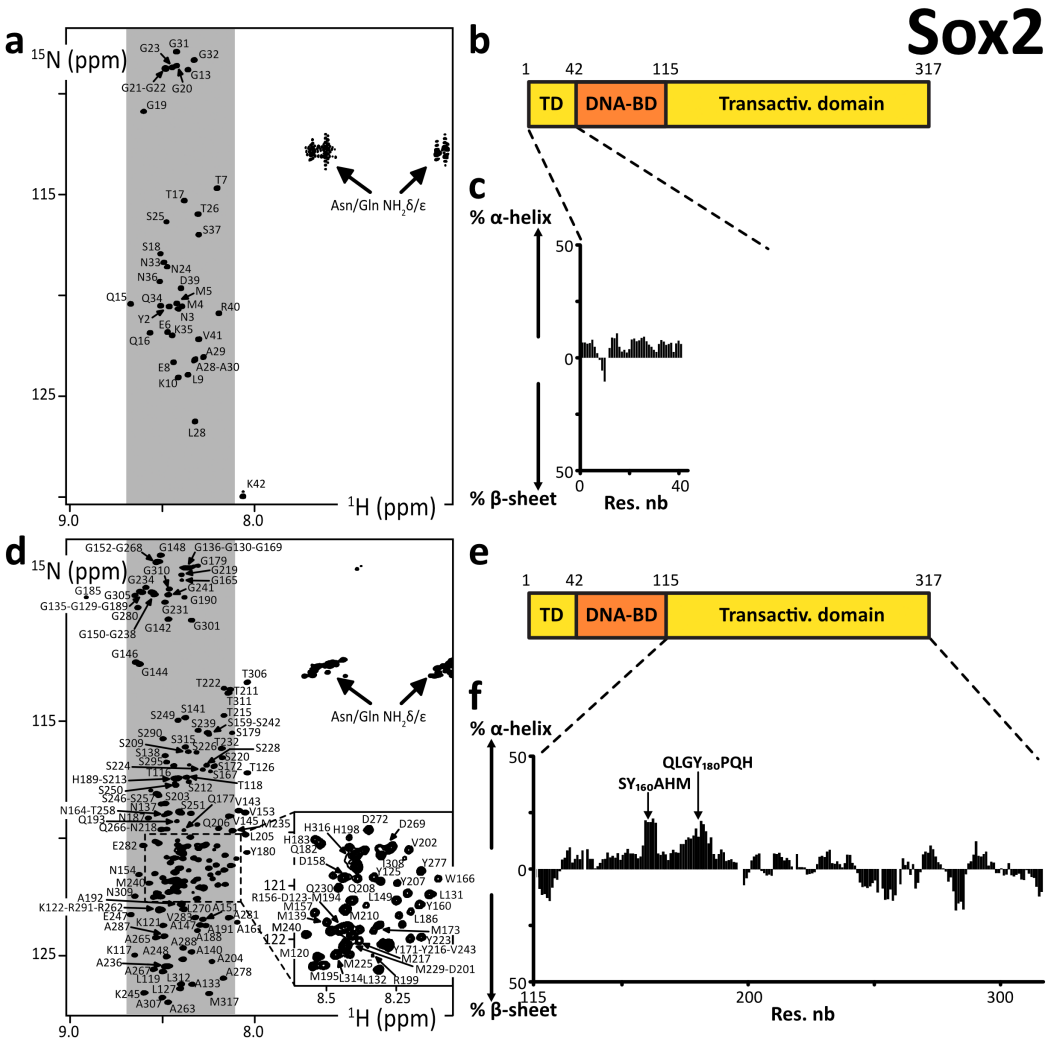

Figure 2. (a,d) 2D ^{1}H–^{15}N HSQC spectra of the N- and C-terminal IDRs of human Sox2, the labels indicating the assignments; the grey areas show the spectral regions where random coil amino acids resonate usually; (b,e) primary structures of human Sox2; dark and light colors indicate the folded and disordered domains, respectively; (c,f) secondary structure propensities calculated from the experimental chemical shifts of the peptide backbone Cα and Cβ, using the ncSPC algorithm [104,105].

in agreement with the consensus motifs of these kinases [119,120]. The kinetics reveal a classical distributive mechanism, where phosphorylation sites are processed independently according to their respective k_{cat} and K_M. Ser22 is the preferred target in all cases, while Ser34 is the least processed by CDKs, if at all: the commercial CDKs are poorly active in our hands, which we have verified with a number of other targets for years in the laboratory; this makes it difficult to distinguish between sites that are only mildly disfavored or those that are more stringently ignored by CDKs in NMR-monitored assays.

Then, we used a potent home-made activated p38α on His6–AviTag–Sox2(aa115–240) and His6–AviTag–Sox2(aa234–317). It phosphorylated all the Ser/Thr–Pro motifs of these two peptides, and also T306 in a $PGT_{306}AI$ context, which shows a favorable proline in position −2 [121], and a less common S212 in a $MTS_{212}SQ$ context.

Hence, we were able to generate AviTag–IDR

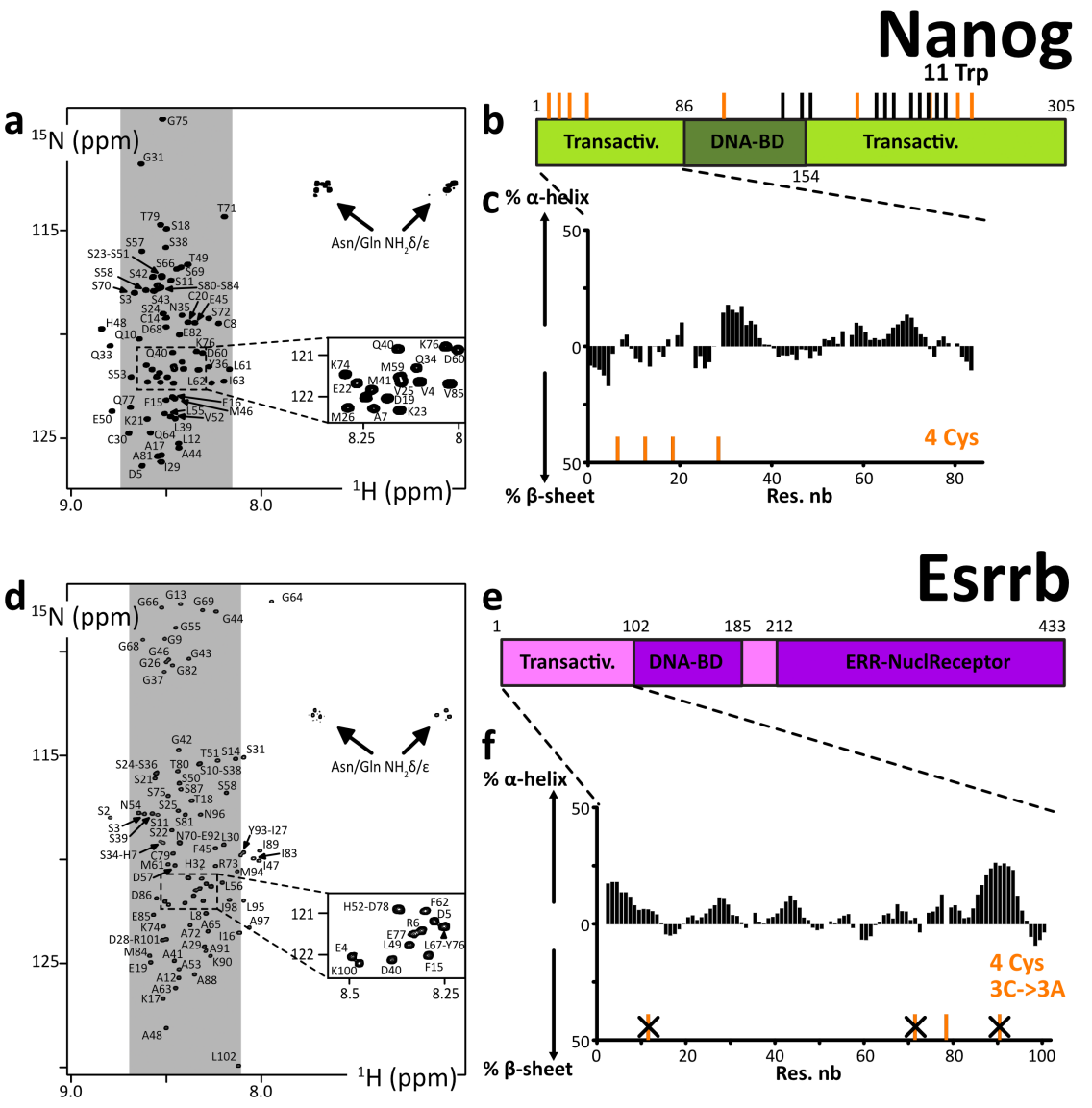

Figure 3. (a,d) 2D ^{1}H–^{15}N HSQC spectra of the N-terminal IDRs of human Nanog and Esrrb, the labels indicating the assignments; the grey areas show the spectral regions where random coil amino acids resonate usually; (b,e) primary structures of human Nanog and Esrrb; dark and light colors indicate the folded and disordered domains, respectively; orange and black sticks indicate the positions of cysteines and tryptophanes, respectively; (c,f) secondary structure propensities calculated from the experimental chemical shifts of the peptide backbone Cα and Cβ, using the ncSPC algorithm [104,105].

chimera in well-defined phosphorylated states. To produce phosphorylated ^{14}N-AviTag–IDR dedicated to pull-down assays, we executed the same protocol on ^{14}N-peptides, while NMR-monitoring in parallel "identical" but ^{15}N-labeled peptides. This allowed us to produce a well-defined phosphorylation state of the ^{14}N-AviTag–IDR for the next experiments.

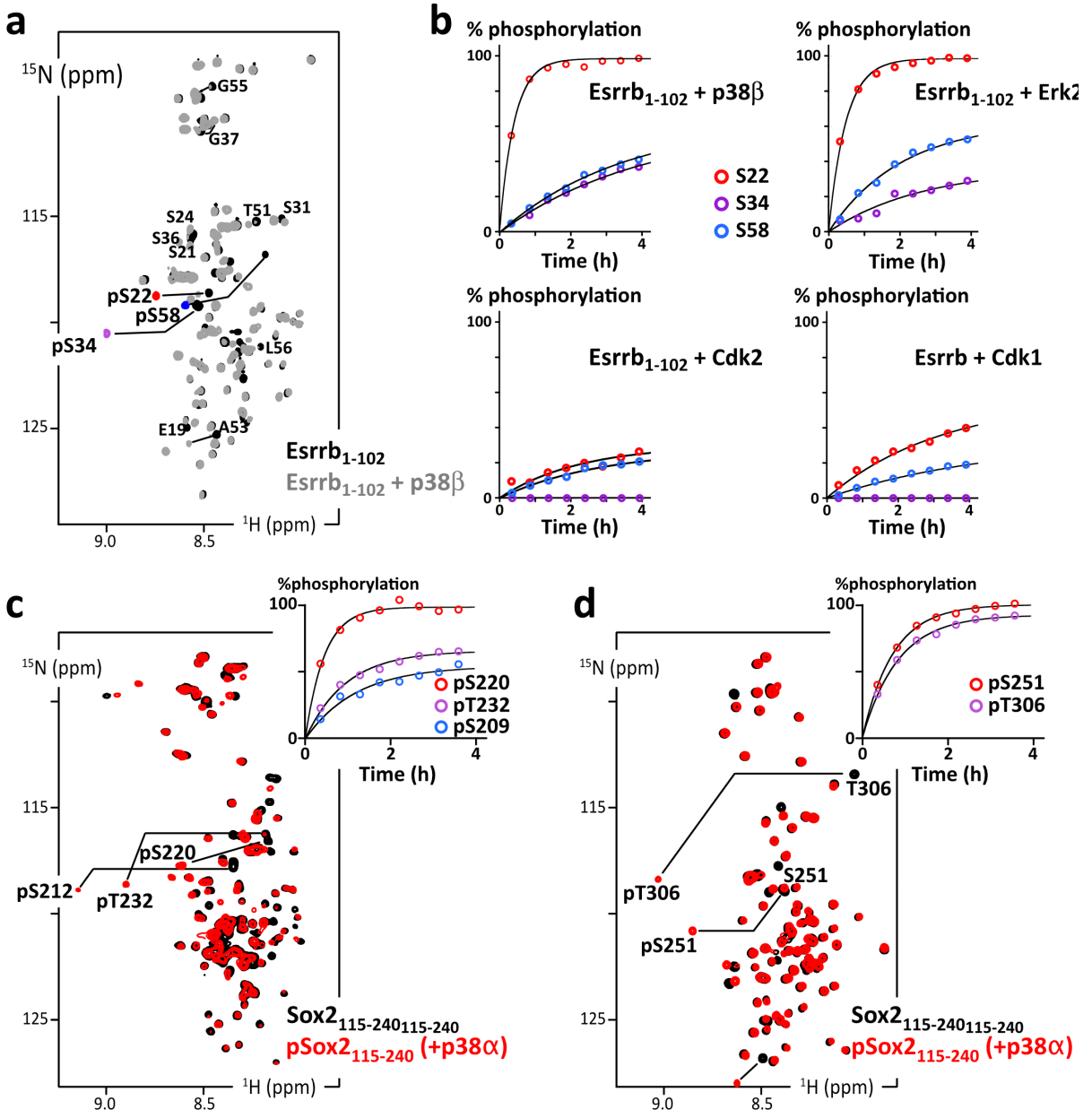

Figure 4. (a) Overlay of 2D ^{1}H–^{15}N HSQC spectra of Esrrb(aa1–102_C12A-C72A-C91A) before (black) and after (grey and phosphosites colored in red/purple/blue) phosphorylation by p38β; (b) residue specific time courses of the phosphorylation of Esrrb(aa1–102) executed by commercial kinases p38β, Erk2, Cdk2 or Cdk1, as measured in time series of 2D ^{1}H–^{15}N SOFAST-HMQC spectra recorded during the reaction; (c) overlay of 2D ^{1}H–^{15}N HSQC spectra of AviTag-Sox2(aa115–240) before (black) and after (red) phosphorylation by p38α; the inset at the top-right shows the residue specific phosphorylation build-up curves, as measured in time series of 2D ^{1}H–^{15}N SOFAST-HMQC spectra; (d) same as (c) with AviTag–Sox2(aa234–317_C265A).

3.6. *Structural characterization of the AviTag–peptide chimera and its biotinylated versions*

We aimed at detecting new partners of OSNE using pull-down assays. We thought to use chimera containing GST at the N-terminus, which appeared as a convenient approach: vectors integrating a GST-coding DNA sequence for overexpression in *E. coli* are available and of common use; glutathione-coated beads are also accessible and permit efficient and specific binding of GST-containing chimera peptides. However, we were unsatisfied by the performances of the method: GST binding to glutathione-coated beads is slow at low temperature (necessary to avoid IDR proteolysis); moreover, it appeared that GST–IDRs chimera are hampered by the IDRs' "molecular cloud" and are even weaker and slower to bind to the beads. Our attempts to bind GST–IDRs to the beads were thus resulting in poor yields, which were not very reproducible. In the context of our aims, i.e., to establish a method allowing quantitative detection of IDRs' binding partners, this unsatisfying lack of reproducibility was only foreboding supplementary variable parameters.

Thus, we decided to switch to another strategy: the use of the specifically biotinylated 15-mer peptide tag called AviTag [101]. This is efficiently and specifically biotinylated by biotin ligase BirA (Figure 5d), which permits high-affinity binding to streptavidin-coated beads. We designed AviTag–IDR chimera, with the AviTag in N-terminal position for Sox2's IDR constructs, and in C-terminal for Esrrb's IDR constructs. We characterized the AviTag and its impact on the IDRs of interest using NMR: the AviTag is unfolded and it does not affect the Sox2 and Esrrb fragments, according to the observed negligible chemical shift perturbations (Figure 5)—the GST Tag provoked also only very weak chemical shift changes on Esrrb(aa1–102) (Supplementary Figure S3). The biotinylated AviTag peptide appears to be slightly less mobile than the common IDPs on the ps–ns timescale, according to the heteronuclear ^{15}N–{^{1}H} NOEs (Figure 5b).

We observed that the biotinylation of the AviTag provokes weak but distinguishable chemical shift perturbations for the close neighbors of the biotinylated lysine, but had no effect on the peptides of interest (Figure 5c). It generated also the appearance of a HN-ester NMR signal, similar to that of acetylated lysine [108,122]. Hence, we could quantify and monitor the reaction advancement using NMR, and determine the incubation time that was necessary and sufficient to obtain a complete biotinylation of our chimera AviTag–IDRs (see Section 2.2). This was one among many optimization steps permitting the production of sufficient quantities of intact IDRs for the pull-down assays.

Next, we tested the binding of the biotinylated AviTag–IDRs on streptavidin-coated beads. This produced very satisfying results, i.e., stoichiometric, specific binding in one hour with no leakage (Supplementary Figure S6). This approach was thus selected for the pull-down assays.

3.7. *Use of AviTag–Sox2 for pull-down assays from mESCs extracts*

We prepared the four peptides AviTag–Sox2(aa115–240) and AviTag–Sox2(aa234–317) in their non-phosphorylated and phosphorylated versions, using p38α to execute the phosphorylation reactions (Figure 6a). These peptides were also biotinylated, and later bound to streptavidin-coated beads, which we used as baits for pull-down assays in extracts from mESCs (Figure 6b). Importantly, an additional SEC was carried out between every step to discard proteolyzed peptides, the enzymes (kinases of BirA) and their contaminants. We performed the pull-down assays with the four samples in parallel with the same cell extract, in duplicate, and then analyzed the bound fractions using quantitative LC-MS/MS analysis (see Supplementary material 1.6 for full description). Hence, we could detect and evaluate the relative quantities of proteins retained by the four AviTag–peptides (Figure 6). On paper, this presents the important advantage of removing false-positive binders, which can interact unspecifically with the streptavidin-coated beads.

We present here results that should be interpreted carefully: we produced only duplicates for every condition, using one single cell extract. To deliver trustful information, the common standards in the field recommend 3 to 5 replicates. Here, we considered only proteins with at least 2 distinct peptides across the 2 replicates (Figure 6c). We observed a two-fold change or more (FC > 2) of some TFs pulled out by AviTag–Sox2(aa115–240), among which the PluriTFs Oct4 and Klf5 are significantly enriched

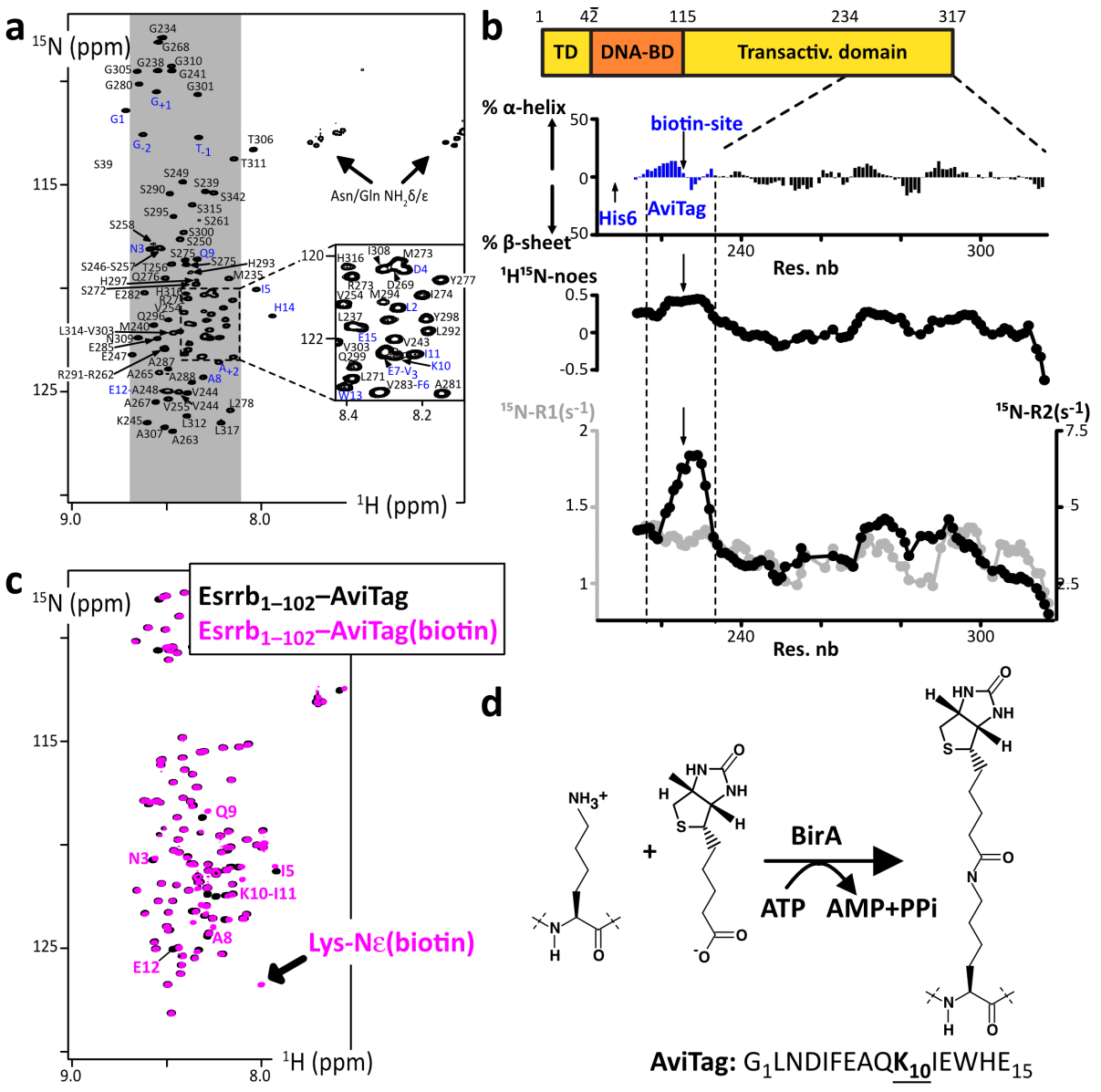

Figure 5. (a) 2D ^{1}H–^{15}N HSQC spectrum of ^{15}N-His6-AviTag-Sox2(aa234–317_C265A), the blue labels indicated the assigned signals from the AviTag residues; (b) secondary structure propensities calculated from the experimental chemical shifts of the peptide backbone Cα and Cβ, using the ncSPC algorithm [104,105]; the residue specific ^{15}N–{^{1}H} NOEs, ^{15}N-R1 (grey) and ^{15}N-R2 (black) measured at 600 MHz are shown below (the profiles show values averaged over three consecutive residues); (c) overlay of 2D ^{1}H–^{15}N HSQC spectra of the Esrrb(aa1–102_C12A-C72A-C91A)–AviTag–His6 before (black) et after (magenta) biotinylation by BirA; the NMR signals from the residues neighboring the biotinylation site are indicated, which permit the quantification of the biotinylated population; (d) scheme of the reaction of AviTag biotinylation executed by the ATP-dependent BirA.

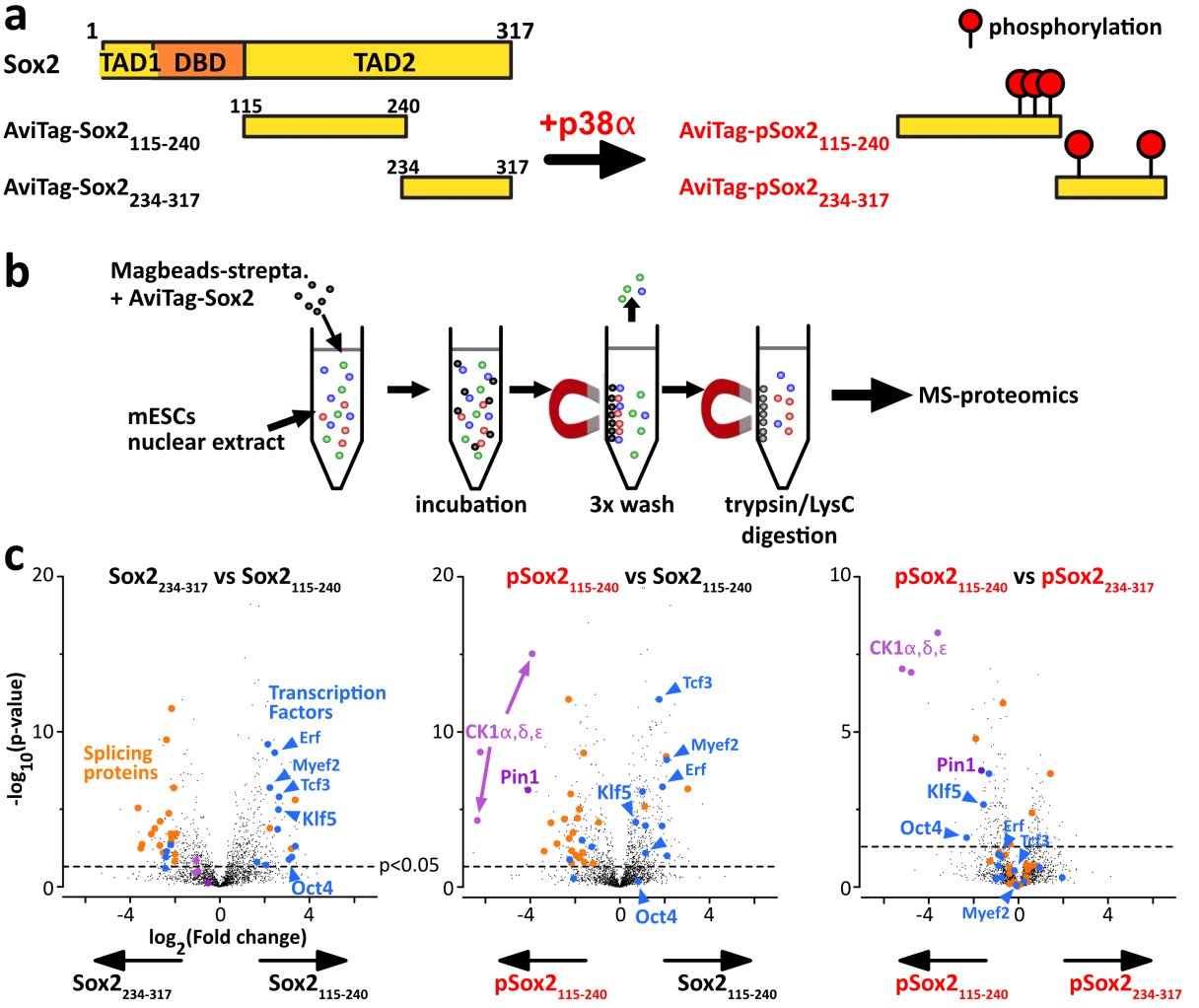

Figure 6. Differential interactomics of Sox2 constructs upon p38α phosphorylation. (a) We have produced unmodified/phospho-Sox2 truncations carrying N-ter biotinylation on the AviTag, and later attached on streptavidin-magnetic beads. (b) We have generated mESCs nuclear extracts for pull-down assays using our Sox2 constructs as baits. (c) The volcano plots of the log2 ratios, showing a quantitative analysis of the proteins present at the end of the pull-down assays; the dashed line indicates the threshold of p-value < 0.05; we highlighted interesting partners in: blue: transcription factors associating with Sox2(aa115–240); magenta: phospho-dependent partners of Sox2(aa115–240) and or Sox2(aa234–317). Pull-downs have been performed in duplicates, using 15 million cells per sample (extract protein conc.: 5 mg/mL), and 1 nmol of bait protein. Experimental conditions may be improved (higher number of replicates, cells, washing conditions, . . .).

(>4 peptides, p-value < 0.02). Also, we found out that pSox2(aa115–240) was pulling out the three isoforms of CK1 (>3 peptides, p-values < 10^{-5}) and the proline isomerase Pin1 (>2 peptides, p-values < 2×10^{-4}). To confirm the value of the method and of the detected interactions, one should at least test them using an orthogonal approach, e.g., NMR spectroscopy of purified proteins.

3.8. *NMR characterization of the interaction between Pin1 and phospho-Sox2*

We decided to test the interaction between pSox2(aa115–240) and Pin1. We recorded ^{1}H–^{15}N NMR spectra of ^{15}N-labeled Sox2(aa115–240) or pSox2(aa115–240) alone or in the presence of the Pin1–WW domain (natural abundance peptide, i.e., 0.6% ^{15}N, 99.4% ^{14}N, hence "NMR-invisible" in ^{15}N-filtered experiments). We observed localized losses in signal intensities for the residues neighboring the three pSox2(aa115–240) phosphosites when mixed with Pin1 (Figure 7b); in contrast, no significant differences showed up in the spectra obtained with non-phosphorylated Sox2(aa115–240) in the absence or presence of Pin1–WW (Figure 7a). Hence, these signal losses are the typical signs of a position-specific interaction between an IDR and a folded protein in the intermediate or slow NMR timescale (μs–s), which corresponds to submicromolar affinities for this type of molecules. Interactions with this range of affinity can therefore be detected using the presented pull-down assay approach.

4. Discussion

The structural biochemistry analysis reported here can be applied to a large list of transcription factors (TFs). These are essential actors of cell signaling: they are key elements for inducing or maintaining pluripotency or differentiation, for cell proliferation or cell cycle arrest, by activating or repressing gene transcription [61,123]. About 90% of the ~1600 TFs are predicted to contain large disordered segments (>30 consecutive amino acids), which is particularly true for PluriTFs [31,124]. A correlation exists actually between TFs and predicted IDRs in all kingdoms of life [45,125,126]. The IDRs of these TFs recruit transcription co-factors or the transcription machinery, which is still not very well characterized in detail [44–55].

Indeed, the fine understanding of TFs' interactions via their IDRs appears to be hampered by the nature of these interactions, which are characterized by weak affinities and multivalency. Moreover, possible redundancy between TFs can emerge from coacervation. Post-translational modifications (PTMs), which can switch on or off IDRs' interactions, are a supplementary source of confusion when searching for binding partners. Among the difficulties, we should also mention the basic biochemistry issues: IDRs are difficult to produce and manipulate in vitro, because they are prone to degradation or aggregation. Here, we have tried to demonstrate the feasibility and interest of some biochemical and spectroscopic approaches to better characterize IDRs of TFs, their phosphorylation and the associated binding partners.

Using a residue specific NMR analysis, we have shown experimentally that the pluripotency TFs OSNE contained IDRs. None of these regions do show any strong secondary structure propensity, which often reveals functional binding sites. Faint ~20% helicities were detected by two independent algorithms (ncIDP and δ2D [104,105,107]) on 5–6 amino acid stretches, which are indicated on Figures 1 and 2 for Oct4 and Sox2. Altogether, the analyzed peptides show local conformational behaviors close to that of a random coil, without any obvious structural elements except the Oct4 stretch between amino acids 33 and 38. Finally, we shall highlight the high propensity for liquid–liquid phase separation of the fragments of Sox2 containing the region aa115–236.

Like other TFs, pluripotency TFs OSNE are post-translationally modified (see Section 1), notably by CDKs and MAPKs [66,78–88]. These two classes of kinases are fundamental actors in all aspects of eukaryotic cellular life, and understanding their activity and regulation in pluripotency or differentiation is of high significance. Interesting questions are still pending: what is the phosphorylation status of OSNE's IDRs in pluripotent cells, what is the impact on their interaction networks, and how does it affect pluripotency or differentiation? The inhibition of MAPK Erk signaling is necessary to maintain pluripotency in the standard culture conditions of ESCs and iPSCs [66,82–84], while these cells show a high CDK activity [8,86,87]; these two kinases family have the same core consensus sites (Ser/Thr–Pro) motifs, which are abundant in OSNE's IDRs and whose phosphorylation has apparently consequences for initiating differentiation [73, 78,80,83,85]. We have shown that we could produce well-defined phosphorylation status of these peptides, using recombinant kinases and NMR analysis, which makes it possible to study their interactions in vitro. We have also demonstrated our capacity to use these peptides as baits in pull-down assays for detecting potential new binding partners.

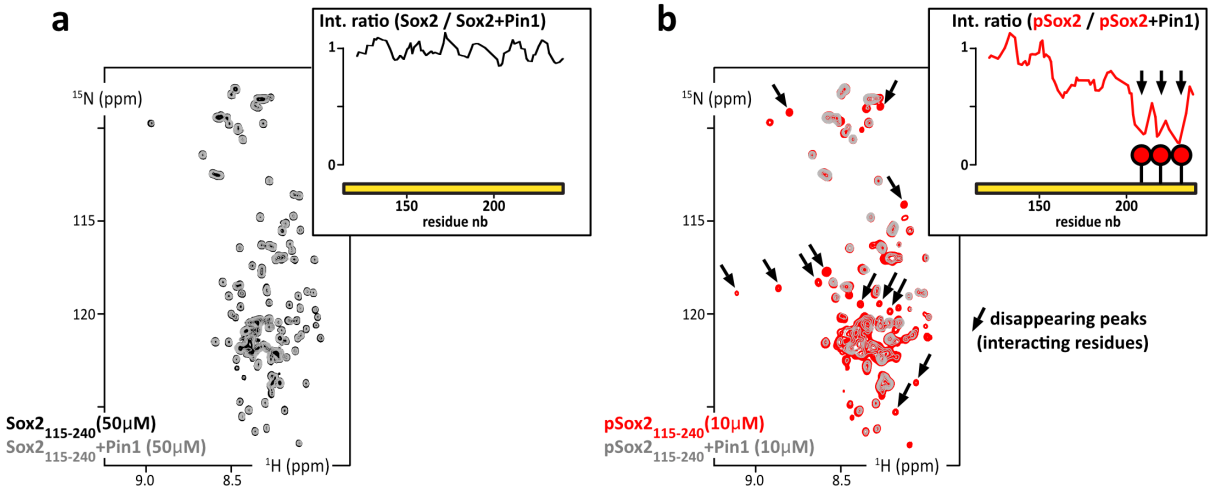

Figure 7. (a) Overlay of 2D ^{1}H–^{15}N HSQC spectra of ^{15}N-Sox2(aa115–240) alone at 50 µM (black) or mixed with Pin1 in isotopic natural abundance and in stoichiometric amounts (grey); inset up-right: residue specific NMR signal intensity ratios as measured in the two HSQC spectra. (b) Overlay of 2D ^{1}H–^{15}N HSQC spectra of ^{15}N-phosphoSox2(aa115–240) alone at 10 µM (red) or mixed with the Pin1–WW domain in stoichiometric amounts (grey); insets, up-right: residue specific NMR signal intensity ratios as measured in the two HSQC spectra (in the absence/presence of Pin1–WW domain).

However, the detected interactions between phospho-Sox2(aa115–240)–pS212–pS220–pT232 and Pin1 or CK1 correspond to widespread, degenerate interactions, whose biological significance may be questionable [127–129]. This is one of the major drawbacks in the field of IDRs' studies: they participate in multiple, degenerate and transient interactions of weak affinities, which can be easily released during the washes of our pull-down assays. In this regard, the "proximity labeling" approaches (BioID, APEX and their derivates) appeared recently to be quite adapted to transient interactions: these methods, developed in the last ten years, use chimera constructs containing enzymes that transfer chemical groups to their intracellular neighbors, which can be later identified by mass spectrometry [130–133]. IDRs are very flexible, solvent-exposed and establish a lot of poorly specific transient interactions. This was raising concerns about the possible production of many false positives if one used proximity labeling methods to detect IDRs' binding partners. This has been partially confirmed by a recent study, but this bias appears to be limited [134]. Interactomes of 109 TFs have actually been described using Bio-ID and affinity purification MS, showing the complementarity between proximity labeling and the pull-down approach proposed here [135]. Yeast two-hybrid, which can detect ~20 µM affinity interactions, and novel phage display approaches will also help in this task [136–139].

Another difficulty in studying IDRs of TFs is their propensity to coacervate [44–55]. Here, we have tried to use Sox2 as a bait in pull-down assays, a protein that has been later recognized to favor liquid–liquid phase separation [44,140]. We met this difficulty during the production steps, which forced us to purify most of the Sox2 constructs in urea at 2 M. We could straightforwardly observe liquid–liquid phase separation of Sox2(aa115–317) at 4 µM using differential interference contrast (DIC) microscopy in the presence of Ficoll-70 (Supplementary Figure S5), but also progressive aggregation and low solubility thresholds while working with our purified samples. These are clear limiting factors for sample production and NMR characterization, which will hamper a number of other studies on IDRs of TFs. This might also affect the results of pull-down assays: we noticed an enrichment in TFs in the samples obtained from pull-downs using Sox2(aa115–240) as a bait, which has a much higher coacervation propensity than Sox2(aa234–317). Is it possible to generate local surface liquid–liquid

phase separation on the surface of streptavidin-coated beads? This might be at the same time a blessing and a curse for future studies, by helping the formation of biologically significant assemblies, or by favoring unspecific, non-native macromolecules interactions.

A final bottleneck in the studies of these IDRs is the capacity to produce post-translationally modified samples. The commercial enzymes are not very well adapted to our NMR studies, because of the required quantities. Here, we have used in-house production of kinase p38α. Since we carried out the present work, we have developed our capacities in producing activated Erk2, Cdk2/CyclinA1 and Cdk1/CyclinB1. These will be part of our future studies. TFs are indeed quite adapted to NMR investigations: they are 300 to 500 residues long and contain large IDRs (~100 amino acids) separating small folded domains (also ~100 amino acids) [31,45]. Their structural characterization would permit understanding a number of cell-signaling mechanisms at the atomic scale, and possibly identifying new therapeutic targets, even though this class of proteins is notoriously difficult to inhibit [60,141,142].

5. Conclusion

We have applied NMR techniques to carry out a primary analysis of the pluripotency transcription factors Oct4, Sox2, Nanog and Esrrb, in particular of their intrinsically disordered regions. We have shown experimentally that they did not adopt a stable fold when isolated, and that we were able to conduct a residue-specific analysis. This relies on the delicate production and purification of these peptides, which are prone to proteolysis and aggregation; producing them in a well-defined PTM status was an even more arduous challenge. We have demonstrated the feasibility of these tasks using recombinant kinases and NMR analysis. We have also evaluated the usefulness of such protein constructs as baits in pull-down assays to detect new binding partners of IDRs. These characterizations and the associated methods provide firm basis for future investigations on transcription factors. The proposed experimental scheme is thus a promising methodology that still needs to be developed and to prove its merits in revealing novel and significant interactions.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Funding

This work was supported by the CNRS and the CEA-Saclay, by the French Infrastructure for Integrated Structural Biology (https://frisbi.eu/, grant number ANR-10-INSB-05-01, Acronym FRISBI) and by the French National Research Agency (ANR; research grants ANR-14-ACHN-0015 and ANR-20-CE92-0013). Financial support from the IR INFRANALYTICS FR2054 for conducting the research is also gratefully acknowledged. This work was also supported by grants from the “Région Ile-de-France” and “Fondation pour la Recherche Médicale” (DL and LSMP).

Acknowledgements

We thank Thaleia Papadopoulou, Amandine Molliex and Navarro Pablo for providing mESCs extracts and for fruitful discussions. We thank Nadia Izadi-Pruneyre for providing the bench space necessary to carry out the pull-down experiments with fresh mESCs extracts. We thank Romain LeBars and the microscopy facility of I2BC Imagerie-Gif for the preliminary tests on liquid–liquid phase separation. We thank the IR-RMN staff (now part of Infranalytics), notably Nelly Morellet, François Giraud and Ewen Lescop, for their reactivity, their support, and their long-standing, efficient care of the 950 MHz spectrometer. We thank Marie Sorin, Baptiste Nguyen and Benjamin Bacri, who contributed to this study during their Master1-Master2 internships.

Supplementary data

Supporting information for this article is available on the journal’s website under https://doi.org/10.5802/crchim.272 or from the author.

References

[1] H. Inoue, N. Nagata, H. Kurokawa, S. Yamanaka, *EMBO J.*, 2014, **33**, 409-417.

[2] Y. Shi, H. Inoue, J. C. Wu, S. Yamanaka, *Nat. Rev. Drug Discov.*, 2017, **16**, 115-130.

[3] R. G. Rowe, G. Q. Daley, *Nat. Rev. Genet.*, 2019, **20**, 377-388.

[4] W. Deng, E. C. Jacobson, A. J. Collier, K. Plath, *Curr. Opin. Genet. Dev.*, 2021, **70**, 89-96.

[5] T. W. Theunissen, R. Jaenisch, *Stem Cell*, 2014, **14**, 720-734.

[6] M. Li, J. C. I. Belmonte, *Nat. Rev. Genet.*, 2017, **18**, 180-191.

[7] C. Chronis, P. Fiziev, B. Papp, S. Butz, G. Bonora, S. Sabri, J. Ernst, K. Plath, *Cell*, 2017, **168**, 442-459, e20.

[8] N. Festuccia, I. Gonzalez, N. Owens, P. Navarro, *Development*, 2017, **144**, 3633-3645.

[9] I. Gonzalez, A. Molliex, P. Navarro, *Curr. Opin. Cell Biol.*, 2021, **69**, 41-47.

[10] P. Mu, Z. Zhang, M. Benelli, W. R. Karthaus, E. Hoover, C.-C. Chen, J. Wongvipat, S.-Y. Ku, D. Gao, Z. Cao, N. Shah, E. J. Adams, W. Abida, P. A. Watson, D. Prandi, C.-H. Huang, E. de Stanchina, S. W. Lowe, L. Ellis, H. Beltran, M. A. Rubin, D. W. Goodrich, F. Demichelis, C. L. Sawyers, *Science*, 2017, **355**, 84-88.

[11] A. C. Hepburn, R. E. Steele, R. Veeratterapillay, L. Wilson, E. E. Kounatidou, A. Barnard, P. Berry, J. R. Cassidy, M. Moad, A. El-Sherif, L. Gaughan, I. G. Mills, C. N. Robson, R. Heer, *Oncogene*, 2019, **38**, 4412-4424.

[12] S. Mirzaei, M. D. A. Paskeh, M. Entezari, S. reza Mirmazloomi, A. Hassanpoor, M. Aboutalebi, S. Rezaei, E. S. Hejazi, A. Kakavand, H. Heidari, S. Salimimoghadam, A. Taheriazam, M. Hashemi, S. Samarghandian, *Biomed. Pharmacother.*, 2022, **156**, article no. 113860.

[13] E.-H. Ervin, R. French, C.-H. Chang, S. Pauklin, *Semin. Cancer Biol.*, 2022, **87**, 48-83.

[14] A. Chaudhary, S. S. Raza, R. Haque, *Semin. Cancer Biol.*, 2023, **88**, 123-137.

[15] D. Esch, J. Vahokoski, M. R. Groves, V. Pogenberg, V. Cojocaru, H. vom Bruch, D. Han, H. C. A. Drexler, M. J. Araúzo-Bravo, C. K. L. Ng, R. Jauch, M. Wilmanns, H. R. Schöler, *Nat. Cell Biol.*, 2013, **15**, 295-301.

[16] A. Reményi, K. Lins, L. J. Nissen, R. Reinbold, H. R. Schöler, M. Wilmanns, *Genes Dev.*, 2003, **17**, 2048-2059.

[17] M. D. Gearhart, S. M. A. Holmbeck, R. M. Evans, H. J. Dyson, P. E. Wright, *J. Mol. Biol.*, 2003, **327**, 819-832.

[18] R. Jauch, C. K. L. Ng, K. S. Saikatendu, R. C. Stevens, P. R. Kolatkar, *J. Mol. Biol.*, 2008, **376**, 758-770.

[19] Y. Hayashi, L. Caboni, D. Das, F. Yumoto, T. Clayton, M. C. Deller, P. Nguyen, C. L. Farr, H.-J. Chiu, M. D. Miller, M.-A. Elsliger, A. M. Deacon, A. Godzik, S. A. Lesley, K. Tomoda, B. R. Conklin, I. A. Wilson, S. Yamanaka, R. J. Fletterick, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, 4666-4671.

[20] B. Yao, S. Zhang, Y. Wei, S. Tian, Z. Lu, L. Jin, Y. He, W. Xie, Y. Li, *J. Mol. Biol.*, 2020, **432**, 5460-5472.

[21] S. O. Dodonova, F. Zhu, C. Dienemann, J. Taipale, P. Cramer, *Nature*, 2020, **580**, 669-672.

[22] A. K. Michael, R. S. Grand, L. Isbel, S. Cavadini, Z. Kozicka, G. Kempf, R. D. Bunker, A. D. Schenk, A. Graff-Meyer, G. R. Pathare, J. Weiss, S. Matsumoto, L. Burger, D. Schübeler, N. H. Thomä, *Science*, 2020, **368**, 1460-1465.

[23] K. Echigoya, M. Koyama, L. Negishi, Y. Takizawa, Y. Mizukami, H. Shimabayashi, A. Kuroda, H. Kurumizaka, *Sci. Rep.*, 2020, **10**, article no. 11832.

[24] G. A. Roberts, B. Ozkan, I. Gachulincová, M. R. O'Dwyer, E. Hall-Ponsele, M. Saxena, P. J. Robinson, A. Soufi, *Nat. Cell Biol.*, 2021, **23**, 834-845.

[25] E. Morgunova, J. Taipale, *Curr. Opin. Struct. Biol.*, 2021, **71**, 171-179.

[26] W. Kagawa, H. Kurumizaka, *Curr. Opin. Struct. Biol.*, 2021, **71**, 59-64.

[27] E. Luzete-Monteiro, K. S. Zaret, *Curr. Opin. Struct. Biol.*, 2022, **75**, article no. 102425.

[28] B. D. Sunkel, B. Z. Stanton, *IScience*, 2021, **24**, article no. 103132.

[29] F. C. M. Gadea, E. N. Nikolova, *J. Mol. Biol.*, 2023, **435**, article no. 167916.

[30] B. Jagga, M. Edwards, M. Pagin, K. M. Wagstaff, D. Aragão, N. Roman, J. D. Nanson, S. R. Raidal, N. Dominado, M. Stewart, D. A. Jans, G. R. Hime, S. K. Nicolis, C. F. Basler, J. K. Forwood, *Nat. Commun.*, 2021, **12**, article no. 28.

[31] B. Xue, C. J. Oldfield, Y.-Y. Van, A. K. Dunker, V. N. Uversky, *Mol. BioSyst.*, 2012, **8**, 134-150.

[32] M. M. Babu, *Biochem. Soc. Trans.*, 2016, **44**, 1185-1200.

[33] P. E. Wright, H. J. Dyson, *Nat. Struct. Mol. Biol.*, 2015, **16**, 18-29.

[34] D. Piovesan, M. Necci, N. Escobedo, A. M. Monzon, A. Hatos, I. Mičetić, F. Quaglia, L. Paladin, P. Ramasamy, Z. Dosztanyi, W. F. Vranken, N. E. Davey, G. Parisi, M. Fuxreiter, S. C. E. Tosatto, *Nucleic Acids Res.*, 2021, **49**, D361-D367.

[35] P. Kulkarni, S. Bhattacharya, S. Achuthan, A. Behal, M. K. Jolly, S. Kotnala, A. Mohanty, G. Rangarajan, R. Salgia, V. Uversky, *Chem. Rev.*, 2022, **122**, 6614-6633.

[36] D. Piovesan, A. Del Conte, D. Clementel, A. M. Monzon, M. Bevilacqua, M. C. Aspromonte, J. A. Iserte, F. E. Orti, C. Marino-Buslje, S. C. E. Tosatto, *Nucleic Acids Res.*, 2023, **51**, D438-D444.

[37] Y. Buganim, D. A. Faddah, R. Jaenisch, *Nat. Struct. Mol. Biol.*, 2013, **14**, 427-439.

[38] A. Rizzino, *Stem Cells*, 2013, **31**, 1033-1039.

[39] S. Jerabek, F. Merino, H. R. Schöler, V. Cojocaru, *Biochim. Biophys. Acta Gene Regul. Mech.*, 2014, **1839**, 138-154.

[40] A. Saunders, F. Faiola, J. Wang, *Stem Cells*, 2013, **31**, 1227-1236.

[41] S. E. Bondos, A. K. Dunker, V. N. Uversky, *Cell Commun. Signal*, 2022, **20**, article no. 20.

[42] K.-P. Kim, Y. Wu, J. Yoon, K. Adachi, G. Wu, S. Velychko, C. M. MacCarthy, B. Shin, A. Röpke, M. J. Arauzo-Bravo, M. Stehling, D. W. Han, Y. Gao, J. Kim, S. Gao, H. R. Schöler, *Sci. Adv.*, 2020, **6**, article no. eaaz7364.

[43] I. Aksoy, R. Jauch, V. Eras, W. A. Chng, J. Chen, U. Divakar, C. K. L. Ng, P. R. Kolatkar, L. W. Stanton, *Stem Cells*, 2013, **31**, 2632-2646.

[44] A. Boija, I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, N. M. Hannett, B. J. Abraham, L. K. Afeyan, Y. E. Guo, J. K. Rimel, C. B. Fant, J. Schuijers, T. I. Lee, D. J. Taatjes, R. A. Young, *Cell*, 2018, **175**, 1842-1855.

[45] L. Staby, C. O'Shea, M. Willemoës, F. Theisen, B. B. Kragelund, K. Skriver, *Biochem. J.*, 2017, **474**, 2509-2532.

[46] C. N. Ravarani, T. Y. Erkina, G. De Baets, D. C. Dudman, A. M. Erkine, M. M. Babu, *Mol. Syst. Biol.*, 2018, **14**, article no. e8190-14.

[47] S. Brodsky, T. Jana, K. Mittelman, M. Chapal, D. K. Kumar, M. Carmi, N. Barkai, *Mol. Cell*, 2020, **79**, 459-471, e4.
[48] A. Erijman, L. Kozlowski, S. Sohrabi-Jahromi, J. Fishburn, L. Warfield, J. Schreiber, W. S. Noble, J. Söding, S. Hahn, *Mol. Cell*, 2020, **78**, 890-902, e6.
[49] G. Næs, J. O. Storesund, P. Udayakumar, M. Ledsaak, O. S. Gabrielsen, *FEBS Open Bio*, 2020, **10**, 2329-2342.
[50] A. L. Sanborn, B. T. Yeh, J. T. Feigerle, C. V. Hao, R. J. Townshend, E. Lieberman Aiden, R. O. Dror, R. D. Kornberg, *ELife*, 2021, **10**, article no. e68068.
[51] L. M. Tuttle, D. Pacheco, L. Warfield, D. B. Wilburn, S. Hahn, R. E. Klevit, *Nat. Commun.*, 2021, **12**, article no. 2220.
[52] L. F. Soto, Z. Li, C. S. Santoso, A. Berenson, I. Ho, V. X. Shen, S. Yuan, J. I. Fuxman Bass, *Mol. Cell*, 2022, **82**, 514-526.
[53] M. V. Staller, E. Ramirez, S. R. Kotha, A. S. Holehouse, R. V. Pappu, B. A. Cohen, *Cell Syst.*, 2022, **13**, 334-345, e5.
[54] B. Bourgeois, T. Gui, D. Hoogeboom, H. G. Hocking, G. Richter, E. Spreitzer, M. Viertler, K. Richter, T. Madl, B. M. T. Burgering, *Cell Rep.*, 2021, **36**, article no. 109446.
[55] K. Teilum, J. G. Olsen, B. B. Kragelund, *Biochem. J.*, 2021, **478**, 2035-2050.
[56] F.-X. Theillet, A. Binolfi, T. Frembgen-Kesner, K. Hingorani, M. Sarkar, C. Kyne, C. Li, P. B. Crowley, L. Gierasch, G. J. Pielak, A. H. Elcock, A. Gershenson, P. Selenko, *Chem. Rev.*, 2014, **114**, 6661-6714.
[57] N. E. Davey, *Curr. Opin. Struct. Biol.*, 2019, **56**, 155-163.
[58] T. M. Filtz, W. K. Vogel, M. Leid, *Trends Pharmacol. Sci.*, 2014, **35**, 76-85.
[59] D. Han, M. Huang, T. Wang, Z. Li, Y. Chen, C. Liu, Z. Lei, X. Chu, *Cell Death Dis.*, 2019, **10**, article no. 290.
[60] M. Qian, F. Yan, T. Yuan, B. Yang, Q. He, H. Zhu, *Drug Discov. Today*, 2020, **25**, 1502-1512.
[61] P. Weidemüller, M. Kholmatov, E. Petsalaki, J. B. Zaugg, *Proteomics*, 2021, **21**, article no. 2000034.
[62] N. Cai, M. Li, J. Qu, G.-H. Liu, J. C. Izpisua Belmonte, *J. Mol. Cell Biol.*, 2012, **4**, 262-265.
[63] L. Fang, L. Zhang, W. Wei, X. Jin, P. Wang, Y. Tong, J. Li, J. X. Du, J. Wong, *Mol. Cell*, 2014, **55**, 537-551.
[64] D. S. Yoon, Y. Choi, Y. Jang, M. Lee, W. J. Choi, S.-H. Kim, J. W. Lee, *Stem Cells*, 2014, **32**, 3219-3231.
[65] H. Jang, T. W. Kim, S. Yoon, S.-Y. Choi, T.-W. Kang, S.-Y. Kim, Y.-W. Kwon, E.-J. Cho, H.-D. Youn, *Stem Cell*, 2012, **11**, 62-74.
[66] J. Brumbaugh, Z. Hou, J. D. Russell, S. E. Howden, P. Yu, A. R. Ledvina, J. J. Coon, J. A. Thomson, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 7162-7168.
[67] S. Dan, B. Kang, X. Duan, Y.-J. Wang, *Biochem. Biophys. Res. Commun.*, 2015, **456**, 714-720.
[68] Y. Cho, H. G. Kang, S.-J. Kim, S. Lee, S. Jee, S. G. Ahn, M. J. Kang, J. S. Song, J.-Y. Chung, E. C. Yi, K.-H. Chun, *Cell Death Differ.*, 2018, **25**, 1781-1795.
[69] C. A. C. Williams, A. Soufi, S. M. Pollard, *Semin. Cancer Biol.*, 2019, **67**, 30-38.
[70] X. Abulaiti, H. Zhang, A. Wang, N. Li, Y. Li, C. Wang, X. Du, L. Li, *Stem Cell Rep.*, 2017, **9**, 1630-1641.
[71] D. K. Kim, B. Song, S. Han, H. Jang, S.-H. Bae, H. Y. Kim, S.-H. Lee, S. Lee, J. K. Kim, H.-S. Kim, K.-M. Hong, B. I. Lee, H.-D. Youn, S.-Y. Kim, S. W. Kang, H. Jang, *Cancers*, 2020, **12**, article no. 2601.
[72] N. P. Mullin, J. Varghese, D. Colby, J. M. Richardson, G. M. Findlay, I. Chambers, *FEBS Lett.*, 2021, **595**, 14-25.
[73] K. T. G. Rigbolt, T. A. Prokhorova, V. Akimov, J. Henningsen, P. T. Johansen, I. Kratchmarova, M. Kassem, M. Mann, J. V. Olsen, B. Blagoev, *Sci. Signal*, 2011, **4**, article no. rs3.
[74] Y. Kamachi, H. Kondoh, *Development*, 2013, **140**, 4129-4144.
[75] J. Shin, T. W. Kim, H. Kim, H. J. Kim, M. Y. Suh, S. Lee, H.-T. Lee, S. Kwak, S.-E. Lee, J.-H. Lee, H. Jang, E.-J. Cho, H.-D. Youn, *Elife*, 2016, **5**, article no. e10877.
[76] P. N. Malak, B. Dannenmann, A. Hirth, O. C. Rothfuss, K. Schulze-Osthoff, *Cell Cycle*, 2015, **14**, 3748-3754.
[77] T. Schaefer, C. Lengerke, *Oncogene*, 2020, **39**, 278-292.
[78] J. Ouyang, W. Yu, J. Liu, N. Zhang, L. Florens, J. Chen, H. Liu, M. Washburn, D. Pei, T. Xie, *J. Biol. Chem.*, 2015, **290**, 22782-22794.
[79] S. Lim, A. Bhinge, S. Bragado Alonso, I. Aksoy, J. Aprea, C. F. Cheok, F. Calegari, L. W. Stanton, P. Kaldis, *Mol. Cell Biol.*, 2017, **37**, article no. e00201-17–24.
[80] H. J. Kim, J. Shin, S. Lee, T. W. Kim, H. Jang, M. Y. Suh, J.-H. Kim, I.-Y. Hwang, D. S. Hwang, E.-J. Cho, H.-D. Youn, *Nucleic Acids Res.*, 2018, **46**, 6544-6560.
[81] M. Moretto-Zita, H. Jin, Z. Shen, T. Zhao, S. P. Briggs, Y. Xu, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 13312-13317.
[82] S.-H. Kim, M.-O. Kim, Y.-Y. Cho, K. Yao, D. J. Kim, C.-H. Jeong, D. H. Yu, K. B. Bae, E.-J. Cho, S. K. Jung, M. H. Lee, H. Chen, J. Y. Kim, A. M. Bode, Z. Dong, *Stem Cell Res.*, 2014, **13**, 1-11.
[83] J. Brumbaugh, J. D. Russell, P. Yu, M. S. Westphall, J. J. Coon, J. A. Thomson, *Stem Cell Rep.*, 2014, **2**, 18-25.
[84] A. Saunders, D. Li, F. Faiola, X. Huang, M. Fidalgo, D. Guallar, J. Ding, F. Yang, Y. Xu, H. Zhou, J. Wang, *Stem Cell Rep.*, 2017, **8**, 1115-1123.
[85] L. Liu, W. Michowski, H. Inuzuka, K. Shimizu, N. T. Nihira, J. M. Chick, N. Li, Y. Geng, A. Y. Meng, A. Ordureau, A. Kołodziejczyk, K. L. Ligon, R. T. Bronson, K. Polyak, J. W. Harper, S. P. Gygi, W. Wei, P. Sicinski, *Nat. Cell Biol.*, 2017, **19**, 177-188.
[86] L. Liu, W. Michowski, A. Kolodziejczyk, P. Sicinski, *Nat. Cell Biol.*, 2019, **21**, 1060-1067.
[87] S. Jirawatnotai, S. Dalton, M. Wattanapanitch, *Semin. Cell Dev. Biol.*, 2020, **107**, 63-71.
[88] R. Spelat, F. Ferro, F. Curcio, *J. Biol. Chem.*, 2012, **287**, 38279-38288.
[89] K. B. Bae, D. H. Yu, K. Y. Lee, K. Yao, J. Ryu, D. Y. Lim, T. A. Zykova, M.-O. Kim, A. M. Bode, Z. Dong, *Stem Cell Rep.*, 2017, **9**, 2050-2064.
[90] Y. Hao, X. Fan, Y. Shi, C. Zhang, D. Sun, K. Qin, W. Qin, W. Zhou, X. Chen, *Nat. Commun.*, 2019, **10**, 1-13.
[91] N. S. Sharma, V. K. Gupta, P. Dauer, K. Kesh, R. Hadad, B. Giri, A. Chandra, V. Dudeja, C. Slawson, S. Banerjee, S. M. Vickers, A. Saluja, S. Banerjee, *Theranostics*, 2019, **9**, 3410-3424.
[92] D. K. Kim, J.-S. Lee, E. Y. Lee, H. Jang, S. Han, H. Y. Kim, I.-Y. Hwang, J.-W. Choi, H. M. Shin, H. J. You, H.-D. Youn, H. Jang, *Exp. Mol. Med.*, 2021, **53**, 1759-1768.
[93] S. Constable, J.-M. Lim, K. Vaidyanathan, L. Wells, *Glycobiology*, 2017, **27**, 927-937.
[94] T. Miura, S. Nishihara, *Trends Glycosci. Glycotechnol.*, 2019, **31**, E69-E75.
[95] L. Ciraku, E. M. Esquea, M. J. Reginato, *Cell. Signal*, 2022, **90**, article no. 110201.

[96] J. Ma, C. Hou, C. Wu, *Chem. Rev.*, 2022, **122**, 15822-15864.

[97] A. M. Gronenborn, D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, G. M. Clore, *Nature*, 1991, **253**, 657-661.

[98] C. K. Smith, J. M. Withka, L. Regan, *Biochemistry*, 1994, **33**, 5510-5517.

[99] A. Alik, C. Bouguechtouli, M. Julien, W. Bermel, R. Ghouil, S. Zinn-Justin, F.-X. Theillet, *Angew. Chem. Int. Ed.*, 2020, **59**, 10411-10415.

[100] M. Howarth, K. Takao, Y. Hayashi, A. Y. Ting, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 7583-7588.

[101] M. Fairhead, M. Howarth, *Methods Mol. Biol.*, 2015, **1266**, 171-184.

[102] F.-X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb, P. Selenko, *Nature*, 2016, **530**, 45-50.

[103] R. Dass, F. A. A. Mulder, J. T. Nielsen, *Sci. Rep.*, 2020, **10**, article no. 14780.

[104] K. Tamiola, F. A. A. Mulder, *Biochem. Soc. Trans.*, 2012, **40**, 1014-1020.

[105] J. T. Nielsen, F. A. A. Mulder, *J. Biomol. NMR*, 2018, **70**, 141-165.

[106] W. Borcherds, F.-X. Theillet, A. Katzer, A. Finzel, K. M. Mishall, A. T. Powell, H. Wu, W. Manieri, C. Dieterich, P. Selenko, A. Loewer, G. W. Daughdrill, *Nat. Chem. Biol.*, 2014, **10**, 1000-1002.

[107] C. Camilloni, A. De Simone, W. F. Vranken, M. Vendruscolo, *Biochemistry*, 2012, **51**, 2224-2231.

[108] F.-X. Theillet, C. Smet-Nocca, S. Liokatis, R. Thongwichian, J. Kosten, M.-K. Yoon, R. W. Kriwacki, I. Landrieu, G. Lippens, P. Selenko, *J. Biomol. NMR*, 2012, **54**, 217-236.

[109] F.-X. Theillet, H. M. Rose, S. Liokatis, A. Binolfi, R. Thongwichian, M. Stuiver, P. Selenko, *Nat. Protoc.*, 2013, **8**, 1416-1432.

[110] A. Mylona, F.-X. Theillet, C. Foster, T. M. Cheng, F. Miralles, P. A. Bates, P. Selenko, R. Treisman, *Science*, 2016, **354**, 233-237.

[111] M. Julien, C. Bouguechtouli, A. Alik, R. Ghouil, S. Zinn-Justin, F.-X. Theillet, in *Intrinsically Disordered Proteins: Methods and Protocols* (B. B. Kragelund, K. Skriver, eds.), Springer US, New York, 2020, 793-817.

[112] N. Festuccia, N. Owens, A. Chervova, A. Dubois, P. Navarro, *Development*, 2021, **148**, article no. dev199604.

[113] J.-P. Lambert, M. Tucholska, T. Pawson, A.-C. Gingras, *J. Proteomics*, 2014, **100**, 55-59.

[114] P. Poullet, S. Carpentier, E. Barillot, *Proteomics*, 2007, **7**, 2553-2556.

[115] M. The, M. J. MacCoss, W. S. Noble, L. Käll, *J. Am. Soc. Mass Spectrom*, 2016, **27**, 1719-1727.

[116] B. Valot, O. Langella, E. Nano, M. Zivy, *Proteomics*, 2011, **11**, 3572-3577.

[117] C. Smet-Nocca, H. Launay, J.-M. Wieruszeski, G. Lippens, I. Landrieu, *J. Biomol. NMR*, 2013, **55**, 323-337.

[118] S. Elkjær, A. D. Due, L. F. Christensen, F. F. Theisen, L. Staby, B. B. Kragelund, K. Skriver, *Commun. Biol.*, 2023, **6**, article no. 63.

[119] W. Peti, R. Page, *Protein Sci.*, 2013, **22**, 1698-1710.

[120] D. J. Wood, J. A. Endicott, *Open Biol.*, 2018, **8**, article no. 180112.

[121] D. L. Sheridan, Y. Kong, S. A. Parker, K. N. Dalby, B. E. Turk, *J. Biol. Chem.*, 2008, **283**, 19511-19520.

[122] S. Liokatis, A. Stützer, S. J. Elsässer, F.-X. Theillet, R. Klingberg, B. van Rossum, D. Schwarzer, C. D. Allis, W. Fischle, P. Selenko, *Nat. Struct. Mol. Biol.*, 2012, **19**, 819-823.

[123] S. A. Lambert, A. Jolma, L. F. Campitelli, P. K. Das, Y. Yin, M. Albu, X. Chen, J. Taipale, T. R. Hughes, M. T. Weirauch, *Cell*, 2018, **172**, 650-665.

[124] J. Liu, N. B. Perumal, C. J. Oldfield, E. W. Su, V. N. Uversky, A. K. Dunker, *Biochemistry*, 2006, **45**, 6873-6888.

[125] I. Yruela, C. J. Oldfield, K. J. Niklas, A. K. Dunker, *Genome Biol. Evol.*, 2017, **9**, 1248-1265.

[126] K. Cermakova, H. C. Hodges, *Trends Biochem. Sci.*, 2023, **48**, 477-490.

[127] X. Z. Zhou, K. P. Lu, *Nat. Struct. Mol. Biol.*, 2016, **16**, 463-478.

[128] Y. Chen, Y. Wu, H. Yang, X. Li, M. Jie, C. Hu, Y. Wu, S. Yang, Y. Yang, *Cell Death Dis.*, 2018, **9**, article no. 883.

[129] J. Gebel, M. Tuppi, A. Chaikuad, K. Hötte, M. Schröder, L. Schulz, F. Lohr, N. Gutfreund, F. Finke, E. Henrich, J. Mezhyrova, R. Lehnert, F. Pampaloni, G. Hummer, E. H. K. Stelzer, S. Knapp, V. Dötsch, *Nat. Chem. Biol.*, 2020, **16**, 1078-1086.

[130] W. Qin, K. F. Cho, P. E. Cavanagh, A. Y. Ting, *Nat. Methods*, 2021, **18**, 133-143.

[131] J.-P. Lambert, M. Tucholska, C. Go, J. D. R. Knight, A.-C. Gingras, *J. Proteomics*, 2015, **118**, 81-94.

[132] A.-C. Gingras, K. T. Abe, B. Raught, *Curr. Opin. Chem. Biol.*, 2019, **48**, 44-54.

[133] X. Liu, K. Salokas, R. G. Weldatsadik, L. Gawriyski, M. Varjosalo, *Nat. Protoc.*, 2020, **15**, 3182-3211.

[134] D.-P. Minde, M. Ramakrishna, K. S. Lilley, *Commun. Biol.*, 2020, **3**, article no. 38.

[135] H. Göös, M. Kinnunen, K. Salokas, Z. Tan, X. Liu, L. Yadav, Q. Zhang, G.-H. Wei, M. Varjosalo, *Nat. Commun.*, 2022, **13**, article no. 766.

[136] C. P. Wigington, J. Roy, N. P. Damle, V. K. Yadav, C. Blikstad, E. Resch, C. J. Wong, D. R. Mackay, J. T. Wang, I. Krystkowiak, D. A. Bradburn, E. Tsekitsidou, S. H. Hong, M. A. Kaderali, S.-L. Xu, T. Stearns, A.-C. Gingras, K. S. Ullman, Y. Ivarsson, N. E. Davey, M. S. Cyert, *Mol. Cell*, 2020, **79**, 342-358, e12.

[137] Y. Ueki, T. Kruse, M. B. Weisser, G. N. Sundell, M. S. Y. Larsen, B. L. Mendez, N. P. Jenkins, D. H. Garvanska, L. Cressey, G. Zhang, N. Davey, G. Montoya, Y. Ivarsson, A. N. Kettenbach, J. Nilsson, *Mol. Cell*, 2019, **76**, 953-964, e6.

[138] C. Benz, M. Ali, I. Krystkowiak, L. Simonetti, A. Sayadi, F. Mihalic, J. Kliche, E. Andersson, P. Jemth, N. E. Davey, Y. Ivarsson, *Mol. Syst. Biol.*, 2022, **18**, article no. e10584.

[139] N. E. Davey, L. Simonetti, Y. Ivarsson, *Trends Biochem. Sci.*, 2022, **47**, 547-548.

[140] G. Krainer, T. J. Welsh, J. A. Joseph, J. R. Espinosa, S. Wittmann, E. de Csilléry, A. Sridhar, Z. Toprakcioglu, G. Gudiškytė, M. A. Czekalska, W. E. Arter, J. Guillén-Boixet, T. M. Franzmann, S. Qamar, P. S. George-Hyslop, A. A. Hyman, R. Collepardo-Guevara, S. Alberti, T. P. J. Knowles, *Nat. Commun.*, 2021, **12**, article no. 1085.

[141] K. Tsafou, P. B. Tiwari, J. D. Forman-Kay, S. J. Metallo, J. A. Toretsky, *J. Mol. Biol.*, 2018, **430**, 2321-2341.

[142] A. Chen, A. N. Koehler, *Trends Mol. Med.*, 2020, **26**, 508-518.