Sensitivity-Enhanced 13C-NMR for Monitoring Multisite Phosphorylation at Physiological Temperature and pH

Ania Alik, Chafiaa Bouguechtouli, Manon Julien, Wolfgang Bermel, Rania Ghouil, Sophie Zinn-Justin, Francois-Xavier Theillet

COMMUNICATION

Sensitivity-Enhanced ^{13}C-NMR for Monitoring Multisite Phosphorylation at Physiological Temperature and pH.

Ania Alik[a]#, Chafiaa Bouguechtouli[a]#, Manon Julien[a]#, Wolfgang Bermel[b], Rania Ghouil[a], Sophie Zinn-Justin[a], Francois-Xavier Theillet*[a]

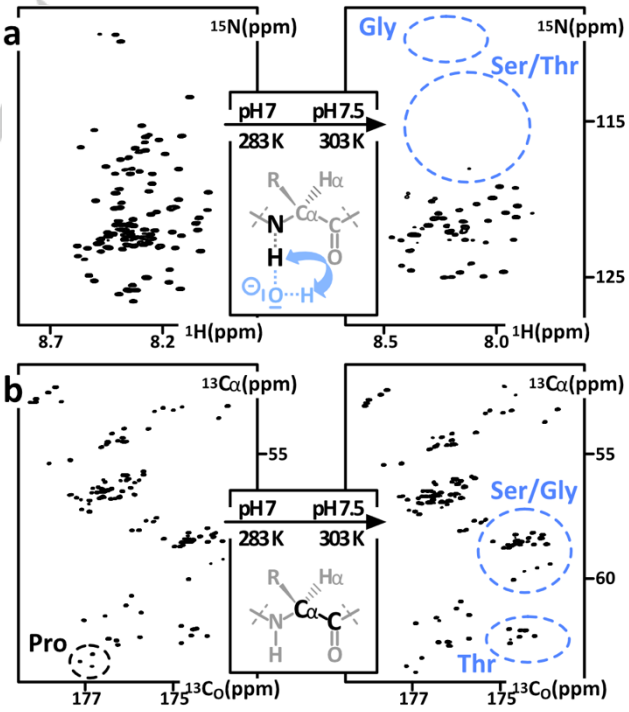
[#] These authors contributed equally.
[*] Corresponding author.
[a] A. Alik, C. Bouguechtouli, M. Julien, R. Ghouil, Dr. S. Zinn-Justin, Dr. F.X. Theillet, Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.
E-mail: francois-xavier.theillet@cnrs.fr
[b] Dr. W. Bermel, Bruker BioSpin GmbH, Silberstreifen, 76287 Rheinstetten, Germany.

Supporting information for this article is given via a link at the end of the document.

Abstract: Abundant phosphorylation events rule the activity of nuclear proteins involved in gene regulation and DNA repair. These occur mostly on disordered regions of proteins, which often contain multiple phosphosites. Comprehensive and quantitative monitoring of phosphorylation reactions is theoretically achievable at a residue-specific level using ^{1}H-^{15}N NMR spectroscopy, but is often limited by low signal-to-noise at pH>7 and T>293K. We developed an improved ^{13}Cα-^{13}CO correlation NMR experiment that works equally at any pH or temperature, i.e. also at those where genuine kinase activities occur. This permits to obtain atomic-resolution information in physiological conditions down to 25 μM. We exemplified the interest of this approach by monitoring phosphorylation reactions, in presence of purified kinases or in cell extracts, on a range of previously problematic targets, namely Mdm2, BRCA2 and Oct4.

In eukaryotes, intracellular signal transduction largely relies on phosphorylation of intrinsically disordered regions of proteins (IDRs)[1]. More than 200,000 phosphosites have been detected on the ~13,000 human phosphoproteins[2,3]. Many of these phosphosites are clustered in unfolded regions of proteins, which conveniently permits to establish crosstalks or robustness in signaling[4,5]. Most of them have been detected by mass-spectrometry (MS) or western-blotting (WB). These popular methods are sensitive, but hardly provide a comprehensive characterization in common cases of multiple neighboring or degenerate phosphosites. High-resolution NMR spectroscopy has emerged as a complementary approach to decipher complex phosphorylation schemes: using the classical detection of ^{1}H-^{15}N amide resonances, residue-specific information is obtained in 2D ^{1}H-^{15}N correlation spectra, even on multiple phosphorylation sites that were not solved by MS/WB[6-8]. In frequent cases where 2D NMR peaks of IDRs overlap, 3D BT-HNCO combined with non-uniform sampling (NUS) shows superior capacities[9,10]. However, these approaches rely on amide protons detection, which is efficient for unfolded peptides at pH<7 and T<293K: because water-amide proton exchange rate is proportional to [OH$^-$] at pH~7 and gets faster at high temperature[11], ^{1}H-^{15}N signals become weaker in physiological conditions where phosphorylation reactions are carried out, especially for fast-exchanging residues like Ser/Thr (Figure1a; Figure S1). Moreover, pH changes of ~0.1 can generate ~25% variations in the H-N crosspeak intensities, making them delicate to use when a quantitative analysis of phosphorylation reaction advancement is needed.

Indeed, to obtain individual phosphorylation rates, time-series of ^{1}H-^{15}N NMR spectra are recorded during the reaction in conditions where kinases are active, i.e. at pH=7 and 298K. Residue-specific peak intensities are then quantified in every spectrum to monitor disappearance or appearance of unphospho- or phospho-species, respectively[6-8]. The non-phosphopeaks (either from the phosphosites themselves or from neighboring residues) provide the most reliable information: reference intensities at 0% phosphorylation are known and permit a straightforward normalization. However, at physiological pH and temperature, disordered proteins generate overlapping and weak ^{1}H-^{15}N signals, as mentioned earlier. The counterpart phosphopeaks display generally more favorable intensities because phosphate groups on the side chain slow down water-



Scheme 1. Figure 1. (a) ^{1}H-^{15}N 2D HSQC and **(b)** ^{13}Cα-^{13}CO 2D (^{1}H-flip*)^{13}Cα^{13}CO-LB spectra of Mdm2(aa284-434) at pH7/283K and pH7.5/303K. Spectral areas of Gly, Ser, Thr resonances are highlighted in blue, those of Pro in black. The same contour levels are applied in all conditions.

amide proton exchange, but exploiting them in a quantitative manner is often problematic: it requires normalization by peak intensities at 100% phosphorylation, a stoichiometry that is often impossible to reach using *in vitro* phosphorylation by purified kinases, especially when these are purchased commercially – commonly at high cost. (Schemes illustrating these normalization issues are shown in Figure S2).

To circumvent these drawbacks, we thought to switch to the observation of proton-less, ^{13}C-detected NMR correlations. This approach became appealing over the last years to characterize unfolded regions of proteins[12-14]. Still, notably because of the lower ^{13}C gyromagnetic ratio, it suffers from lower S/N than conventional ^{1}H-NMR. The ^{13}CO-^{15}N correlation experiments are used to study IDRs because they offer the best crosspeaks dispersion, but necessitate high protein quantities to provide quality spectra (~500 μM in ~300 μL), which is not very adapted to phosphorylation studies: a whole batch of commercial kinase is usually necessary to phosphorylate ~100-1000 nmol of high-preference phosphosites. To reach a more favorable NMR sensitivity, we decided to build on the (^{1}H-flip)^{13}Cα-^{13}CO correlation experiment, the most sensitive 2D ^{13}C-detected pulse sequence[12]. This type of experiment is not spoiled by water-amide proton exchange, at the opposite of the ^{13}CO-^{15}N experiments (Figure S3). ^{13}Cα-^{13}CO experiments are usually thought to generate unresolved spectra showing abundant crosspeaks overlaps for unfolded peptides. However, using prolonged constant-time ^{13}Cα-evolutions, ^{13}Cα-^{13}CO spectra can provide residue-specific information with high coverage along the peptide sequence, as we will see below.

Hence, we thought to improve the (^{1}H-flip)^{13}Cα-^{13}CO sensitivity and adapt the sequence to our application. We first tuned the magnetization transfer delays (we named this first step ^{1}H-flip*, see Figure S4), sacrificing glycine resonances but generating a 15% increase for the rest of the amino acids. On average, this pays off in terms of phospho- and unphosphopeak intensities monitoring. Then, we replaced the final IPAP scheme by a homonuclear decoupling scheme LOW-BASHD applied during the acquisition (Figure S4)[15]. We also implemented a 120 μs triply compensated pulse G5 [16] to invert both carbonyl and alkyl ^{13}C nuclei in the final INEPT block. This yielded on average a 40% increase in S/N without any side-effects on peak line-width (Figure 2a, 2b), as demonstrated by Bax and colleagues[15]. It left the door open for implementing the preservation of equivalent pathways (PEP) routinely applied to ^{1}H-^{15}N HSQC[17]. Despite multiple trials using various pulses and transfer delays (Figure S3), the best combination provided only a ~40% enhancement in signal, while also increasing the noise by √2 because the PEP protocol requires the acquisition of two independent datasets later recombined before processing.

Finally, we thought to use chelated transition metals like Ni-DO_2A for paramagnetic T1 relaxation enhancement, which is a potent way for NMR sensitivity improvement[18,19]. However, to preserve kinase activity, phosphorylation reactions require i) the presence of Mg^{2+}, which can compete with DO_2A-chelated divalent metal ions, and ii) reducing agents, which react with Ni-DO_2A and most of the transition metals complexes that we tried. Among the commercial DOTA derivates available, NOTA is attractive because it is tailored to chelate trivalent transition metals, and thus not Mg^{2+}. We found that the vanadium ion complex [V^{3+}($NOTA^{3-}$)] (Figure 2c) was the only one to remain stable in presence of Mg^{2+} and DTT, as well as at high pH and temperature.

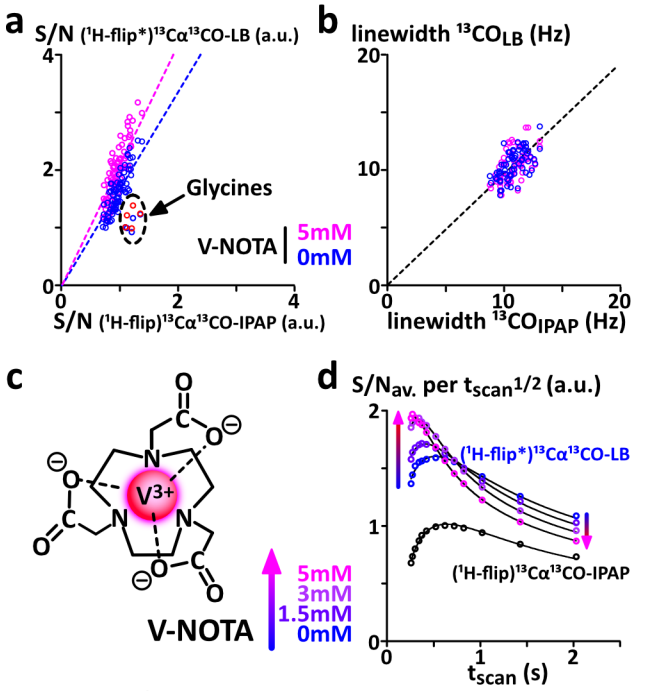


Figure 2. Performances of the (^{1}H-flip*)^{13}Cα^{13}CO-LB sequence, as evaluated on the model protein α-synuclein. **(a)** S/N of well-separated peaks in 2D (^{1}H-flip*)^{13}Cα^{13}CO-LB with 0mM (blue) or 5mM (red) V-NOTA compared to (^{1}H-flip)^{13}Cα^{13}CO-IPAP (0mM V-NOTA); **(b)** Linewidths of isolated peaks in 2D (^{1}H-flip*)^{13}Cα^{13}CO-LB with 0mM (blue) or 5mM (red) V-NOTA compared to (^{1}H-flip)^{13}Cα^{13}CO-IPAP (0mM V-NOTA). **(c)** Chemical structure of V^{3+} chelated by $NOTA^{3-}$; **(d)** Sensitivity and time efficiency: average S/N per $\sqrt{t_{scan}}$ as measured in 1D (^{1}H-flip)^{13}Cα^{13}CO-IPAP (black) and (^{1}H-flip*)^{13}Cα^{13}CO-LB with 0mM (blue), 1.5mM (violet), 3mM (purple) and 5mM (ping) V-NOTA.

It provided up to a 25% increase in S/N for the (^{1}H-flip*)^{13}Cα-linewidths (Figure 2). Fast pulsing with low interscan delays (~0.1s) provides the best S/N per experimental time (Figure 2d). Altogether, using a classical cryoprobe optimized for "indirect" ^{1}H-detection, a ^{13}Cα constant-time evolution of 27 ms and 50% NUS [20,21], we can record in 30 minutes high-resolution 2D ^{13}Cα-^{13}CO correlation spectra providing residue-specific information on IDRs at 50 μM and at any pH or temperature. This can be translated into 25 μM for ^{13}C-direct detection probes. The resulting resolution is often sufficient to get residue-specific information along most of the disordered peptide sequences. To cancel CαCβ coupling in ^{13}Cα dimension, a ^{13}Cα constant-time evolution of 27, 54 or 81 ms must be used. In case of important peak overlaps, an increased resolution is achievable using a ^{13}Cα evolution time of 54 ms, which provides an average 12 Hz linewidth in ^{13}Cα-dimension but provokes about 30% losses in S/N (8 Hz linewidth using an evolution time of 81 ms, Figure S5).

With this dedicated analysis tool in hands, we then attempted to characterize challenging phosphorylation reactions that we failed to monitor accurately using only ^{1}H-^{15}N experiments. This was the case for human Mdm2, an E3-ubiquitin ligase controlling the abundance of the tumor suppressor p53[22,23]. Mdm2 activity is regulated by DNA-damage response (DDR) kinases DNA-PK, ATM or ATR [24,25]. According to the literature, they phosphorylate Mdm2 on 6 sites of its C-terminal region (aa334-434)[24-26], among which mouse S394 (human S395) phosphorylation has consequences on p53 stabilization and DNA-damage signaling.[26-28] To investigate the structural mechanism behind

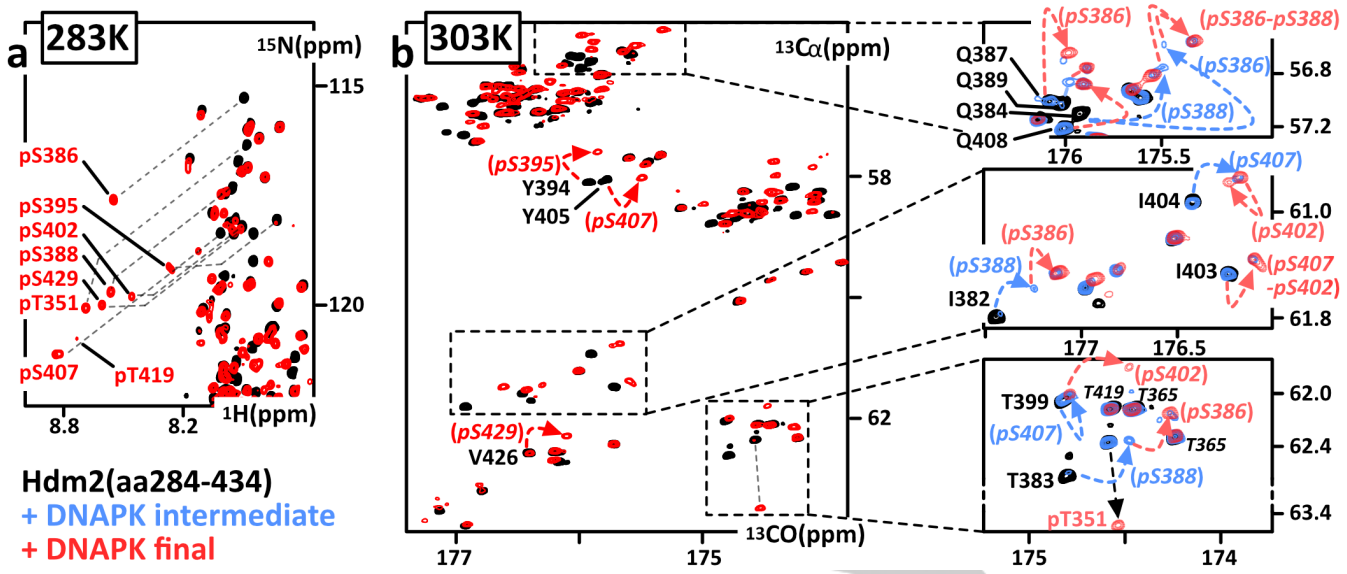


Figure 3. Phosphorylation of human Mdm2(aa284-434) by DNA-PK observed by NMR. (a) Overlay of 2D ^{1}H-^{15}N HSQC spectra at 283K before (black) and after (red) the phosphorylation reaction. **(b)** Overlay of 2D (^{1}H-flip*)^{13}Cα^{13}CO-LB at 303K before (black), in the course (light blue) and at the end (red) of the phosphorylation reaction. At the right: Close-up views of Gln, Ile, and Thr regions. Arrows show peak displacements upon 2 neighbor phosphorylation events: certain sets of peaks report for the 4 phosphostates, e.g. Q384(S386-S388)/Q384(pS386-S388)/Q384(S386-pS388)/Q384(pS386-pS388); in contrast I382(pS388) and T383(pS388) report the appearance of Mdm2(S386-pS388), while I382(pS386) and T383(pS386) both report Mdm2(pS386-S388)+Mdm2(pS386-pS388) (see Experimental Procedures in Supp. Info. for the assignment strategy).

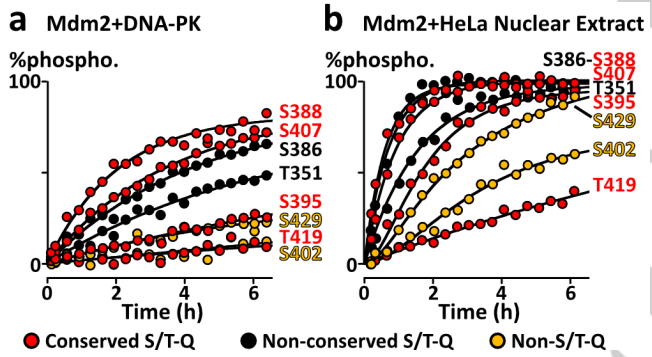


Figure 4. Site-specific phosphorylation kinetics of Mdm2(aa284-434). Build-up curves of phosphorylation reactions as executed **(a)** by purified DNA-PK or **(b)** in HeLa cells Nuclear Extract; The phosphosite labels are in line with their corresponding curves. Non-conserved sites are in black, conserved S/T-Q sites in red, non-S/T-Q sites in orange.

this functional switch, we had first to map comprehensively the phosphorylation scheme established by the DDR kinases. We used the segment Mdm2(aa284-434) comprising the region Mdm2(aa334-434) that we characterized as intrinsically disordered (Figure S6). After a prolonged reaction at 298K in the presence of purified DNA-PK, 2D ^{1}H-^{15}N spectra of Mdm2(aa284-434) recorded at 283K revealed the phosphorylation endpoint (Figure 3): we identified 8 phosphosites T351, S386, S388, S395, S402, S407, T419 and S429. Initially, we could determine T351, S395 and S429 phosphorylation rates by recording time series of ^{1}H-^{15}N SOFAST-HMQC NMR experiments during the reaction at 298K [29]. In the cases of S386, S388, S402, S407 and T419, we were unable to quantify accurately their kinetics because of ^{1}H-^{15}N non-phosphopeaks overlapping or being even non-detectable for S402 and S407. Together with their uncomplete phosphorylation, these prevented the determination of reference ^{1}H-detected peak intensities at 0% and 100% phosphorylation. Here, previously published ^{1}H-detection methods [6-10] found their limits. Then, we recorded times series of alternated ^{1}H-^{15}N SOFAST-HMQC and (^{1}H-flip*)^{13}Cα^{13}CO-LB experiments during the reaction at 298K. These provided complementary information: ^{1}H-^{15}N SOFAST-HMQC spectra reported phosphopeaks build-up curves at high S/N, which we normalized using (^{1}H-flip*)^{13}Cα^{13}CO-LB non-phosphopeaks intensities and their decay curves. We found exploitable non-phosphopeak reporters for S386, S388, S402, S407 and T419 in ^{13}Cα-^{13}CO spectra at 298K (Figure 3, see the Material & Methods for a detailed list, a scheme of the rationale is presented in Figure S2). Using these peaks, we quantified and ranked DNA-PK's site preferences on Mdm2 (Figure 4a). Then, we tested our ability to monitor phosphorylation in cell extracts. Our approach mixing ^{1}H-^{15}N SOFAST-HMQC and (^{1}H-flip*)^{13}Cα^{13}CO-LB spectra worked in cell extracts containing from 0.2 to 5 mg/mL of cellular proteins: the phosphorylation order observed with purified DNA-PK was very similar in HeLa cells Nuclear Extract that were DNA-damaged by short sonication pulses. Thus, our site-specific absolute quantification revealed i) that S388 is a high-preference novel phosphosite, and T351 and S402 are two new lower preference sites, ii) that S429 and S402, which do not correspond to the DDR kinases S-T/Q consensus motif, were steadily modified, iii) that T383 and T419, two S-T/Q consensus sites, were not or slowly phosphorylated, respectively. Similarly, we monitored site-specific phosphorylation on two other disordered segments of nuclear proteins involved in genome maintenance and expression, namely the Breast Cancer type 2 susceptibility protein (BRCA2 aa190-284, see [30]) and the pluripotency transcription factor Oct4 (aa1-145) (Figure S7). Our characterization of BRCA2 phosphorylation by Plk1 (pH7.8,

T=303K), an important mitotic event, permitted to quantify properly two novel phosphosites T219 and T226, which were modified with rates comparable to the already known S193 and T207 (Figure 5a,c). We also wanted to examine Oct4 phosphorylation by MAP kinases, which regulates Oct4 transcriptional activity[31]. In presence of p38α (pH7.4, T298K), we noticed that the classical MAPK target motifs Ser/Thr-Pro S12, S111 and T118 were modified slower than the non-Ser/Thr-Pro T92 (Figure 5b,d). Here, we bring attention to the fact that $^{13}C\alpha^{13}CO$ correlation are very adapted to the study of proline-rich peptides, the residues in N-terminal of Pro showing strong upfield shifts and interesting chemical shift dispersion (see also Figure S8). This is of interest, since the abundant Ser/Thr-Pro motifs are targets of the essential activities of MAP-Kinases and Cyclin-dependent kinases[32]. Because this method is not limited by temperature conditions, it may also help in studies focusing on temperature-sensitive activities of kinases involved in circadian rhythm or hyperthermia consequences[33,34]. Hence, our improved ^{13}C-detected NMR approach expands the range of peptides and conditions accessible to NMR phospho-mapping and -monitoring. NMR methods previously published for PTM characterization are based on amide ^{1}H-detection, which often provides the best sensitivity and possibly higher dimensionality/resolution, using fast acquisition and time-resolved NUS [6-10]. The proposed ^{13}C-detection scheme makes it possible to obtain residue-specific information with high sequence coverage on disordered proteins by NMR in all conditions of pH and temperature.

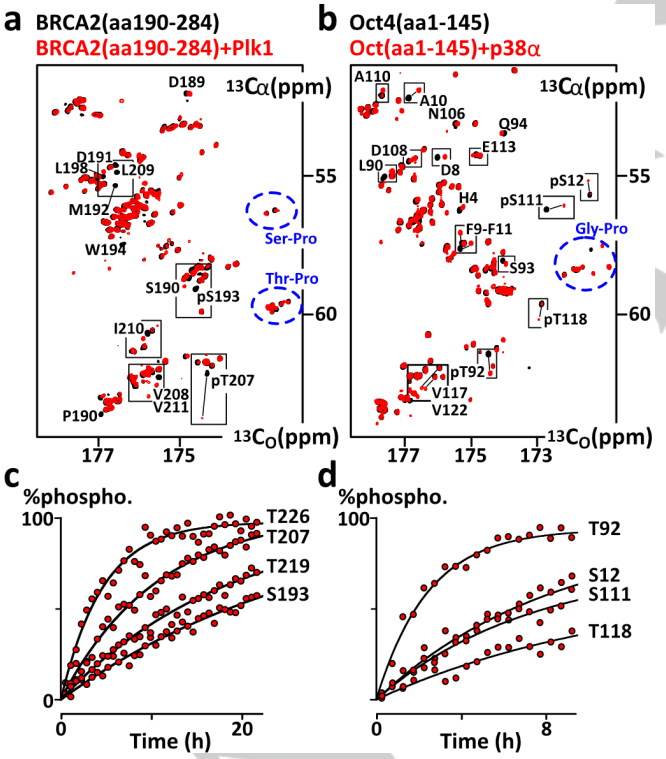


Figure 5. Phosphorylation as observed by 2D (^{1}H-flip*)$^{13}C\alpha^{13}CO$-LB NMR. Overlays of (^{1}H-flip*)$^{13}C\alpha^{13}CO$-LB spectra of **(a)** BRCA2(aa190-284) and **(b)** Oct4(aa1-145) before (black) and after (red) phosphorylation by Plk1 and p38α, respectively. Spectral area corresponding to crosspeaks of Ser/Thr/Gly-Pro motifs are in blue. Kinetics were derived from time-series of (^{1}H-flip*)$^{13}C\alpha^{13}CO$-LB spectra of **(c)** BRCA2(aa190-284) at 100 μM, pH7.8, T=303K **(d)** Oct4(aa1-145) at 100 μM, pH7.4, T=298K at 700MHz using a Bruker-TCI probe.

Experimental section

The Supporting Information contains detailed Material and Methods, HN/CαCO/CON NMR spectra of α-Synuclein, Elk1 and BRCA2 fragments at various pH, temperature and resolution, pulse sequence schemes, proteins' primary sequences presentations.

Acknowledgements

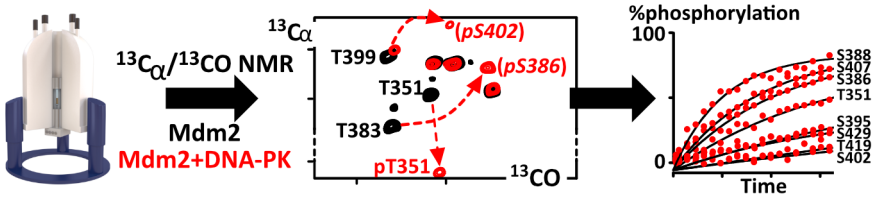
This work was supported by CNRS and CEA-Saclay, by the French Infrastructure for Integrated Structural Biology (https://www.structuralbiology.eu/networks/FRISBI, grant number ANR-10-INSB-05-01, Acronym FRISBI) and by the French National Research Agency (ANR; research grant ANR-14-ACHN-0015). We thank the protein production facility of the Institut Curie for providing the purified kinase Plk1.

Keywords: Cell signaling • Phosphorylation • NMR spectroscopy • Mdm2 • DNA-damage response

[1] P. E. Wright, H. J. Dyson, *Nat Rev Mol Cell Biol* **2015**, *16*, 18–29.
[2] P. Vlastaridis, P. Kyriakidou, A. Chaliotis, Y. Van de Peer, S. G. Oliver, G. D. Amoutzias, *Gigascience* **2017**, *6*, 1–11.
[3] K. Yu, Q. Zhang, Z. Liu, Q. Zhao, X. Zhang, Y. Wang, Z.-X. Wang, Y. Jin, X. Li, Z.-X. Liu, et al., *Nucleic Acids Res.* **2019**, *47*, D451–D458.
[4] F.-X. Theillet, A. Binolfi, T. Frembgen-Kesner, K. Hingorani, M. Sarkar, C. Kyne, C. Li, P. B. Crowley, L. Gierasch, G. J. Pielak, et al., *Chem. Rev.* **2014**, *114*, 6661–6714.
[5] H. Nishi, A. Shaytan, A. R. Panchenko, *Front Genet* **2014**, *5*, 270.
[6] F.-X. Theillet, H. M. Rose, S. Liokatis, A. Binolfi, R. Thongwichian, M. Stuiver, P. Selenko, *Nat. Protoc.* **2013**, *8*, 1416–1432.
[7] M. J. Smith, C. B. Marshall, F.-X. Theillet, A. Binolfi, P. Selenko, M. Ikura, *Curr. Opin. Struct. Biol.* **2015**, *32*, 39–47.
[8] A. Mylona, F.-X. Theillet, C. Foster, T. M. Cheng, F. Miralles, P. A. Bates, P. Selenko, R. Treisman, *Science* **2016**, *354*, 233–237.
[9] M. Mayzel, J. Rosenlöw, L. Isaksson, V. Y. Orekhov, *J. Biomol. NMR* **2014**, *58*, 129–139.
[10] D. Gołowicz, P. Kasprzak, V. Orekhov, K. Kazimierczuk, *Prog. Nucl. Magn. Reson. Spectrosc.* **2020**, *116*, 40–55.
[11] Y. Bai, J. S. Milne, L. Mayne, S. W. Englander, *Proteins* **1993**, *17*, 75–86.
[12] W. Bermel, I. Bertini, I. C. Felli, R. Pierattelli, *J. Am. Chem. Soc.* **2009**, *131*, 15339–15345.
[13] I. C. Felli, L. Gonnelli, R. Pierattelli, *Nat. Protoc.* **2014**, *9*, 2005–2016.
[14] J. Lopez, R. Schneider, F.-X. Cantrelle, I. Huvent, G. Lippens, *Angew. Chem. Int. Ed.* **2016**, *128*, 7544–7548.
[15] J. Ying, F. Li, J. H. Lee, A. Bax, *J. Biomol. NMR* **2014**, *60*, 15–21.
[16] Y. Xia, P. Rossi, M. V. Subrahmanian, C. Huang, T. Saleh, C. Olivieri, C. G. Kalodimos, G. Veglia, *J. Biomol. NMR* **2017**, *69*, 237–243.
[17] P. K. Mandal, A. Majumdar, *Concepts Magn. Reson.* **2004**, *20A*, 1–23.
[18] F.-X. Theillet, A. Binolfi, S. Liokatis, S. Verzini, P. Selenko, *J. Biomol. NMR* **2011**, *51*, 487–495.
[19] N. A. Oktaviani, M. W. Risør, Y.-H. Lee, R. P. Megens, D. H. de Jong, R. Otten, R. M. Scheek, J. J. Enghild, N. C. Nielsen, T. Ikegami, et al., *J. Biomol. NMR* **2015**, *62*, 129–142.
[20] K. Kazimierczuk, V. Y. Orekhov, *Angew. Chem. Int. Ed.* **2011**, *50*, 5556–5559.
[21] M. A. Zambrello, A. D. Schuyler, M. W. Maciejewski, F. Delaglio, I. Bezsonova, J. C. Hoch, *Methods* **2018**, *138-139*, 62–68.
[22] A. C. Joerger, A. R. Fersht, *Annu. Rev. Biochem.* **2016**, *85*, 375–404.
[23] O. Karni-Schmidt, M. Lokshin, C. Prives, *Annu. Rev. Pathol. Mech. Dis.* **2016**, *11*, 617–644.
[24] A. Hafner, M. L. Bulyk, A. Jambhekar, G. Lahav, *Nat Rev Mol Cell Biol* **2019**, *20*, 199–210.
[25] A. N. Blackford, S. P. Jackson, *Mol. Cell.* **2017**, *66*, 801–817.
[26] D. W. Meek, *Biochem. J.* **2015**, *469*, 325–346.
[27] T.-A. Nguyen, D. Menendez, M. A. Resnick, C. W. Anderson, *Hum Mutat* **2014**, *35*, 738–755.
[28] H. S. Gannon, B. A. Woda, S. N. Jones, *Cancer Cell* **2012**, *21*, 668–679.
[29] P. Schanda, Ē. Kupče, B. Brutscher, *J. Biomol. NMR* **2005**, *33*, 199–211.

[30] M. Julien, S. Miron, A. Carreira, F.-X. Theillet, S. Zinn-Justin, *Biomolecular NMR Assignments* **2020**.
[31] 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. U.S.A.* **2012**, *109*, 7162–7168.
[32] C. J. Miller, B. E. Turk, *Trends Biochem. Sci.* **2018**, *43*, 380–394.
[33] Y. Shinohara, Y. M. Koyama, M. Ukai-Tadenuma, T. Hirokawa, M. Kikuchi, R. G. Yamada, H. Ukai, H. Fujishima, T. Umehara, K. Tainaka, et al., *Mol. Cell.* **2017**, *67*, 783–798.e20.
[34] D. Deredge, P. L. Wintrode, M. E. Tulapurkar, A. Nagarsekar, Y. Zhang, D. J. Weber, P. Shapiro, J. D. Hasday, *J. Biol. Chem.* **2019**, *294*, 12624–12637.

Entry for the Table of Contents



We present an NMR method to monitor site-specific phosphorylation at high pH/T, which are necessary for kinases activity but prevented thorough characterization of disordered proteins. Our approach uses sensitivity-enhanced $^{13}C\alpha$-^{13}CO experiments. We detail their improvements and efficiency on Mdm2, BRCA2, Oct4. Much more phosphorylation cases are now investigable by NMR.

Supporting Information

69451 Weinheim, Germany

Sensitivity-Enhanced ^{13}C-NMR for Monitoring Multisite Phosphorylation at Physiological Temperature and pH.

Ania Alik[a]#, Chafiaa Bouguechtouli[a]#, Manon Julien[a]#, Wolfgang Bermel[b], Rania Ghouil[a], Sophie Zinn-Justin[a], Francois-Xavier Theillet*[a]

Abstract: Abundant phosphorylation events rule the activity of nuclear proteins involved in gene regulation and DNA repair. These occur mostly on disordered regions of proteins, which often contain multiple phosphosites. Comprehensive and quantitative monitoring of phosphorylation reactions is theoretically achievable at a residue-specific level using ^{1}H-^{15}N NMR spectroscopy, but is often limited by low signal-to-noise at pH>7 and T>293K. We developed an improved ^{13}Cα-^{13}CO correlation NMR experiment that works equally at any pH or temperature, i.e. also at those where genuine kinase activities occur. This permits to obtain atomic-resolution information in physiological conditions down to 25 μM. We exemplified the interest of this approach by monitoring phosphorylation reactions, in presence of purified kinases or in cell extracts, on a range of previously problematic targets, namely Mdm2, BRCA2 and Oct4.

DOI: 10.1002/anie.2016XXXXX

SUPPORTING INFORMATION

Table of Contents

Experimental Procedures

Peptide constructs

Codon-optimized genes coding for human Mdm2(aa284-434), Mdm2(aa284-333), Mdm2(aa284-380), Mdm2(aa284-393), Mdm2(aa284-415), and mouse Mdm2(aa282-432) were synthesized by Genscript and cloned into pET41(a+) vector between SacII and BamHI restriction sites. Cys362 and Cys374 (mouse Cys372) were Ala-mutated. Codon-optimized Oct4(aa1-145) was synthesized by Genscript and cloned into pET41(a+) vector between SacII and HindIII restriction sites. Codon optimized gene coding for human alpha-synuclein was synthesized by Genscript and cloned into pET22(a+) vector between NdeI and NotI restriction sites. The plasmid encoding the NatB acetyltransferase was obtained from Addgene. Codon-optimized gene coding for mouse Elk1(aa309-429) was synthesized by Genscript, Ala-mutated by Genscript on positions 325, 337, 418 and 423 and cloned into pET45(b+) vector between KpnI and PacI restriction sites. Codon-optimized gene coding for human BRCA2(aa48-218) Ala-mutated on positions Cys132, Cys138, Cys148 and Cys161 was synthesized by Genscript and cloned in pETM-13 between NcoI and BamHI restriction sites. Codon-optimized human p38α (aa1-360, full-length) was synthesized by Genscript and cloned into pET41(a+) between NdeI and AvrII restriction sites. Codon-optimized human Erk2(aa8-360) was synthesized by Genscript and cloned into pET45(b+) vector between KpnI and HindIII restriction sites. The plasmid coding for constitutively active MKK6 (GST-MEKR4V) was a generous gift from Isabelle Landrieu. It was used to activate both Erk2 and p38α. All genes contained a coding sequence for the Tobacco Etch Virus Protease (Tev) cleavage site ENLYFQG on the N-terminal side of the peptide of interest. Oct4(aa1-145) was also followed on the C-terminal side by a Tev cleavage site and a GB1 tag to stabilize the construct during recombinant expression. The Tev construct contained a hexahistidine tag (His6), and was produced in-house recombinantly in *Escherichia coli*.

Recombinant protein production

All peptides were produced in BL21(DE3)Star *E. coli* transformed with the plasmids presented above. BRCA2(aa190-284) was produced as previously described [1]. Alpha-synuclein was produced in cells co-transformed with the plasmid encoding NatB following protocols published earlier.[2] Cells were grown in M9 medium containing ^{13}C-glucose (2 g/L) and $^{15}NH_4^+$ (0.5 g/L) as sole sources of carbon and nitrogen, or in LB for kinases. Media were supplemented with 100 μg/mL ampicillin or 50 μg/mL kanamycine depending on the plasmids, and in 30 μg/mL ampicillin and 15 μg/mL chloramphenicol for alpha-synuclein. Overexpression was induced at an optical density OD_{600}=0.8 by supplementing the medium with 1 mM ITPG at 37°C for all peptides except kinases, which were expressed at 30°C. Cells were harvested by centrifugation 4 hours later and cell pellets stored at -80°C. Cell lysis was performed using sonication in Tris 20 mM, NaCl 150 mM, pH 7.5 (TBS), in presence of benzonase (Sigma-Aldrich), lysozyme, protease inhibitors 1x (EDTA-free cOmplete, Roche) and 10 mM DTT. Soluble and insoluble fractions were separated by 15 min centrifugation at 15,000g. All peptides were purified from the soluble fraction except Elk1(aa309-429), which was resolubilized in TBS supplemented with 8 M urea. Alpha-synuclein was purified using protocols published earlier. [2] For the other peptides, the cell extracts were loaded on a His-Trap FF column (5 mL, GE Healthcare) and eluted using a gradient of imidazole. The eluted fractions were later concentrated, submitted to Tev treatment for 1 hour (after transfer in TBS for Elk1) in TBS supplemented with 10 mM DTT, and then re-loaded on His-Trap column. Fractions containing the peptide of interest were finally submitted to a final size-exclusion chromatography using a Superdex 16/60 75 pg (GE healthcare) previously equilibrated with Hepes 40 mM, NaCl 75 mM or 150 mM, pH 7.0 or 7.4 or 7.8.

Vanadium-NOTA production

VCl_3 was purchased from AlfaAesar–Fisher Scientific. NOTA was purchased from Chematech (Dijon, France). NOTA and VCl_3 were solubilized in water at 0.1 M and 1 M, respectively. They were mixed in a 1:1.1 ratio, and pH was progressively shifted from acidic values (about 2-3) to 7. When isolated, NOTA saturation concentration is lower than 0.1 M; it becomes soluble upon V^{3+} complexation, which requires to elevate pH and reach the anion form $NOTA^{3-}$. $V(OH)_3$ precipitate was sedimented using 14,000 g centrifugation during 10 minutes. 2 mL of the supernatant was applied on a Source 15 RPC column (GE-Healthcare) previously equilibrated with Hepes 1 mM and NaCl 5 mM, pH 7. Salts are eluted first, before the purple-brownish V-NOTA fractions (about 5 mL). These were lyophilized over night before being re-solubilized in water and re-injected on the same Source 15 RPC column and re-lyophilized. The obtained powder was solubilized in ~100-200 μL water. The most accurate method for weighting the final amount of V-NOTA is the following one: we weight the final microtube filled with the chosen volume of water before and after V-NOTA solubilization. Hepes and NaCl weights from the lyophilized 5 mL must not be forbidden. About 20-30 mg V-NOTA are typically obtained from a starting amount of 30 mg NOTA.

SUPPORTING INFORMATION

HeLa Nuclear Extract production

HeLa S3 cells were grown in DMEM (4.5 g/L glucose, Sigma) supplemented with 10 % FBS. 100 million cells were harvested by trypsinization and lysed using the Qproteome Mitochondria isolation kit (Qiagen). We kept the nuclear fraction and resuspended it in 500 μL of Hepes 50 mM, NaCl 400 mM, glycerol 10%, protease inhibitors 1x (EDTA-free cOmplete, Roche), 2 mM DTT and phosphatase inhibitors 1x (PhoSTOP, Roche), pH7. We submitted it to three 5s pulses of sonication using a microtip and a 30% output power (VibraCell, Sonics & Material In. Danbury, Connecticut, USA). The extract was aliquoted and stored at -80°C. This protocol yields extracts at 5 to 10 mg/mL protein concentration. The final working concentration used for the presented figures is 0.5 mg/mL.

NMR spectroscopy

All NMR spectra were recorded on a 700 MHz Bruker Avance Neo spectrometer, equipped with a cryogenically cooled triple resonance $^{1}H[^{13}C/^{15}N]$ probe optimized for ^{1}H detection (TCI Bruker from 2006: ^{13}C ASTM S/N=850, ^{1}H 0.1% ethylbenzene S/N=7260 when delivered; actual ^{13}C-optimized probes ASTM S/N specifications are about 2500). All spectra were processed in Topspin 4.0.4. 3D spectra analysis was carried out using CccpNmr 2.4.2, and 2D spectra were analyzed using NMRFAM-SPARKY.[3] DSS at 100 μM and 3% D_2O were added in all samples.

NMR assignments of backbone amide resonances of uniformly-labeled peptides ($^{13}C/^{15}N$) was achieved via BEST-HNCO, -HN(CA)CO, -HNCACB,[4] (H)N(CA)NH and HNCA 3D experiments, at pH 6.8 and 283K, and at peptide concentrations ranging from 200 to 500 μM in 5 mm diameter Shigemi tubes. HNCO and HN(CA)CO were carried out with 1536 (^{1}H) x 92 (^{13}C) x 80 (^{15}N) complex points and sweep widths of 16.23 ppm (^{1}H), 10 ppm (^{13}C) and 30 ppm (^{15}N), HNCACB 1536 (^{1}H) x 92 (^{13}C) x 80 (^{15}N) complex points and sweep widths of 16.23 ppm (^{1}H), 70 ppm (^{13}C) and 30 ppm (^{15}N), (H)N(CA)NH with 1536 (^{1}H) x 64 (^{13}C) x 64 (^{15}N) complex points and sweep widths of 16.23 ppm (^{1}H), and 30 ppm (^{15}N), and HNCA 1536 (^{1}H) x 192 (^{13}C) x 80 (^{15}N) complex points and sweep widths of 16.23 ppm (^{1}H), 26 ppm (^{13}C) and 30 ppm (^{15}N). Non-uniform sampling was applied (35 to 50%) and processing was achieved using RMDD algorithm implemented in Topspin4.0.4. [5] [6] Spectra were processed with linear prediction of 16 complex points in both ^{13}C, and ^{15}N dimensions, cosine apodization in every dimensions, and zero filling to 2048, 512 and 256 complex points in ^{1}H, ^{13}C, and ^{15}N dimensions, respectively. Assignment 2D ^{1}H-^{15}N HSQC spectra were recorded using 1536 (^{1}H) x 256 (^{15}N) complex points and sweep widths of 16.23 ppm (^{1}H) and 30 ppm (^{15}N), and processed with zero filling to 4K and 1K in the proton and nitrogen dimensions, respectively.

We used this assignment strategy for backbone amide resonances (BEST-HNCO, -HN(CA)CO, -HNCACB, (H)N(CA)NH and HNCA 3D experiments, at 283K) both for non-modified and phosphorylated species. The assignments at 303K in 2D spectra were transferred from those at 283K by recording spectra at 283K, 293K and 303K: this is sufficient to characterize peak-specific temperature shifts. The spectra and the assignment of “DNA-PK intermediate” phosphorylated Mdm2(aa284-434) were obtained using samples whose phosphorylation was stopped by heating at 323K for 10 minutes (provoking DNA-PK inactivation).

Assignments of unmodified and phosphorylated human Mdm2(aa284-434, C362A-C374A), as well as details of NMR experiments used to derive them have been deposited in the Biological Magnetic Resonance Data Bank (BMRB, accession numbers 28011, 28019, respectively). As reported in the literature by Bycroft et al. [7], we observed two stable conformations of the Zn-finger that correspond to trans- and cis-prolines (PDB: 2C6A and 2C6B). The relative NMR peak intensities of the two species were the same in the isolated Zn-finger and in Mdm2(aa284-434) (trans:cis~2:1). Secondary structure propensities of unmodified and phosphorylated Mdm2(aa284-434, C362A-C374A) were obtained with the neighbor-corrected structural propensity calculator ncSCP [8,9] (http://www.protein-nmr.org/ , https://st-protein02.chem.au.dk/ncSPC/) using experimentally determined, DSS referenced Cα and Cβ chemical shifts as input.

We set up 2D spectra during phosphorylation kinetics as follows: 2D ^{1}H-^{15}N SOFAST-HMQC experiments were recorded using 1536 (^{1}H) x 92 (^{15}N) complex points and sweep widths of 16.23 ppm (^{1}H) and 30 ppm (^{15}N), 16 scans (128 dummy scans) and interscan delays of 0.12 s, resulting in 7 minutes long spectra; 2D (^{1}H-flip*)$^{13}C\alpha^{13}CO$-LB experiments were recorded using 1024 (^{13}CO) x 128 ($^{13}C\alpha$) complex points and sweep widths of 20.28 ppm (^{13}CO) and 14 ppm ($^{13}C\alpha$), 80 scans (96 dummy scans) and interscan delays of 0.15 s, and 50% non-uniform sampling optimized for a 35 Hz J-coupling (i.e. low sampling around 14 ms Cα evolution time) and 60 ms relaxation, resulting in 31 minutes long spectra. All spectra were processed zero filling to 4K and 1K in the direct and indirect dimensions, respectively. Cosine apodization was applied in ^{13}CO dimension for (^{1}H-flip*)$^{13}C\alpha^{13}CO$-LB spectra, no apodization was applied for ^{1}H-^{15}N SOFAST-HMQC spectra. 2D (^{1}H-flip*)$^{13}Ca^{13}CO$-LB experiments processing was achieved using the CS algorithm implemented in Topspin4.0.4 [10], parameters were set as follows: Mdd_mode=cs, MddCEXP=false, MddCT_SP=true, Mdd_CsALG=IST, Mdd_CsVE=false, Mdd_CsNITER=0. Two 2D ^{1}H-^{15}N SOFAST-HMQC spectra were initially acquired, before recording alternately 2D (^{1}H-flip*)$^{13}C\alpha^{13}CO$-LB and 2D ^{1}H-^{15}N SOFAST-HMQC spectra. We used 5 mm diameter Shigemi advanced tubes for phosphorylation kinetics, which increases NMR active volume to 330 μL, and yields ~13 % higher ^{13}C-direct detection sensitivity, and no ^{1}H-dectection gains or losses.

Phosphorylation reactions

Phosphorylation reactions were carried out using ^{15}N-labeled target peptides concentrations between 12.5 and 100 μM, in Hepes 40 mM, NaCl 75 mM, DTT 2 mM, ATP 5 mM, $MgCl_2$ 20 mM, protease inhibitors (Roche), 3% D_2O, pH 7.2 to pH 7.8 at 25°C or 30°C in 330 μL. The spectra and kinetics shown in all figures are from kinetics run with 100 μM Mdm2, Oct4, Elk1 or BRCA2 fragments. No changes in phosphorylation preferences were observed with concentration changes. Mdm2 phosphorylation by purified DNA-PK was achieved using commercial DNA-PK (ThermoFisher Scientific / Invitrogen): 2.5 μg of DNA-PK and 35 μg of DNA-PK activator (ThermoFisher Scientific / Invitrogen), i.e. calf thymus DNA, were spiked in the sample immediately before inserting it in the NMR

spectrometer. Regular calf thymus DNA concentrations of 2.5 μg/mL were not reliable enough: initial phosphorylation rates were decreasing with Mdm2 concentration increasing, which goes against kinetics theory and probably reveals an interaction between DNA and Mdm2. We saturated the system with 100 μg/mL of Calf thymus DNA to have constant DNA-PK activation. Moreover, we observed that satisfying reproducibility of DNA-PK phosphokinetics is achievable at 298K, but that this kinase complex appears less stable at 303K, resulting in variable phosphorylation results. BRCA2 kinetics were run at 50 and 100 μM, in hepes 50 mM, NaCl 50mM, pH 7.8 in presence of 250 nM of recombinant Plk1, 2 mM ATP, 1 mM EDTA, 20 mM MgCl2, 2 mM DTT, 1x protease inhibitors cocktail cOmplete (Roche), and 3 mM V-NOTA (Plk1 aliquots gently provided by the protein production platform of the Institut Curie, Paris, France). Oct4 kinetics were run at 50 and 100 μM, in hepes 20 mM, NaCl 50mM, pH 7.4 in presence of 500 nM of recombinant p38α, 2 mM ATP, 5 mM MgCl2, 2 mM DTT 1x protease inhibitors cocktail cOmplete (Roche) using recombinant p38α produced and activated in-house (see the paragraph "Peptide constructs"). Elk1 phosphorylation was performed using recombinant human Erk2 produced and activated in-house (see the paragraph "Peptide constructs").

Phosphorylation kinetics analysis

After processing spectra as detailed in the paragraph "NMR spectroscopy", we measured peak intensities in NMRFAM-SPARKY.[3] Peaks were centered in every spectra to follow peak shifting because of pH drifts. Progress curves were plotted and fitted in Kaleidagraph 4.5. We used only decay curves to normalize phosphorylation build-up curves. We describe the set of disappearing peaks that we used in the following paragraph.

About Mdm2(aa284-434), we could determine pT351, pS395 and pS429 phosphorylation rates using peaks from 2D ^{1}H-^{15}N SOFAST-HMQC NMR spectra. Indeed, we monitored peak intensities of neighboring residues A345+L347 for pT351, Y394+S395 for pS395, L430 for pS429, which all disappear proportionally to the phosphorylation advancement of their closest phosphosite. The other phosphosites (S386, S388, S402, S407, T419) were monitored using 2D (^{1}H-flip*)^{13}Cα^{13}CO-LB spectra as follows. D380, K381, I382 and T383 peaks are good reporters of the concomitant disappearance of non-modified S386 and S388. Even more valuable, a second set of intermediate I382 and T383 peaks appear corresponding to S386/pS388, while a third one report the progressive buildup of pS386/S388+pS386/pS388. In the case of I382 and T383, pS388 does not provoke any further peaks shifting after S386 phosphorylation. At the opposite, (^{1}H-flip*)^{13}Cα^{13}CO-LB peaks of Q384, Q387 and Q387 split in 4 different forms corresponding to the 4 combinations of S386-pS386-S388-pS388, but their interpretation is more tedious because of peak overlaps (Figure 3). Peaks of S402, I403, I404, Y405, S406 and Q408 are good reporters of S407 disappearance, i.e. pS407 appearance, and S398, T399, S402 and I404 of the later appearance of pS402. Finally, E415, R416 and T419 are good reporters of the pT419 reaction advancement. Intensities of peaks that are reporter of a phosphorylation reaction were added to increase the experimental S/N. Decay curves were fitted according to the following equations: $y(t)=I_0*\exp(-k*t)+I_{final}$ with I_0 and k the fitted values. Normalization was performed using (I_0+ I_{final}), and phosphorylation build-up curves are thus $\%phospho(t)=1-y(t)/(I_0+ I_{final})$. We also used phosphopeaks of pS386, pS388, pS402, pS407 and pT419 observed in 2D ^{1}H-^{15}N SOFAST-HMQC NMR spectra, which are well-separated and identifiable in 2D ^{1}H-^{15}N spectra. We normalized them according to the final levels reached at the end of the kinetics, as determined from (^{1}H-flip*)^{13}Cα^{13}CO-LB decay curves. Proton-detected SOFAST-HMQC experiments still provide higher S/N and permit to generate higher quality build-up curves.

BRCA2(aa190-284) phosphokinetics were derived from (^{1}H-flip*)^{13}Cα^{13}CO-LB peak intensities of M192, M192(pS193), S193, W194 and S196 for pS193, T207, pT207, V208, V208(pT207), I210, I210(pT207) and V211 for pT207, A216, T219, pT219 and V220 for pT219, and from H223, A226 and V229 for pT226. Oct4(aa1-145) phosphokinetics were derived from (^{1}H-flip*)^{13}Cα^{13}CO-LB peak intensities of D8, D8(pS12), F9, F9(pS12), A10, A10(pS12), and S12 for pS12, L90, T92 and S93 for pS93, D108, A110, S111 and E113 for pS111, T116, T116(pT118), V117, T118, A121 and V122 for pT118.

Pulse program

We provide two versions of the (^{1}H-flip*)^{13}Cα^{13}CO-LB Bruker pulse program: one for Neo consoles and Topspin4 (c_hcaco_ctre.lowbash.top4), one for earlier systems and Topspin 3 (c_hcaco_ctre.lowbash.top3). They are straightforward to parameterized: only the proton pulse length (p3) has to be determined for each sample, while carbon pulse lengths can be set up only once. The carbon homo-decoupling LOW-BASH pulse is the only pulse that is delicate to parameterize: one should follow carefully the indications provided by Bax and coworkers [11]. The Shapetool in Topspin4 permits to generate the necessary shape pulse: a center lobe of a sinc function of duration 22,600/ν_c in μs (ν_c is the 13C frequency in MHz) should be amplitude modulated by a cosine function at a frequency of 118 x ν_c. Hence, it applies pulses of equivalent power on both sides of the ^{13}CO carrier frequency, centered at 55 and 191 ppm, which efficiently vanishes Bloch-Siegert shift effects[11]. This decoupling pulse is automatically applied every 5 ms by the pulse program, as well as the MLEV16 RF phases. We provide the shape pulse Bruker file for a 700 MHz spectrometer as an example in Supplementary data.

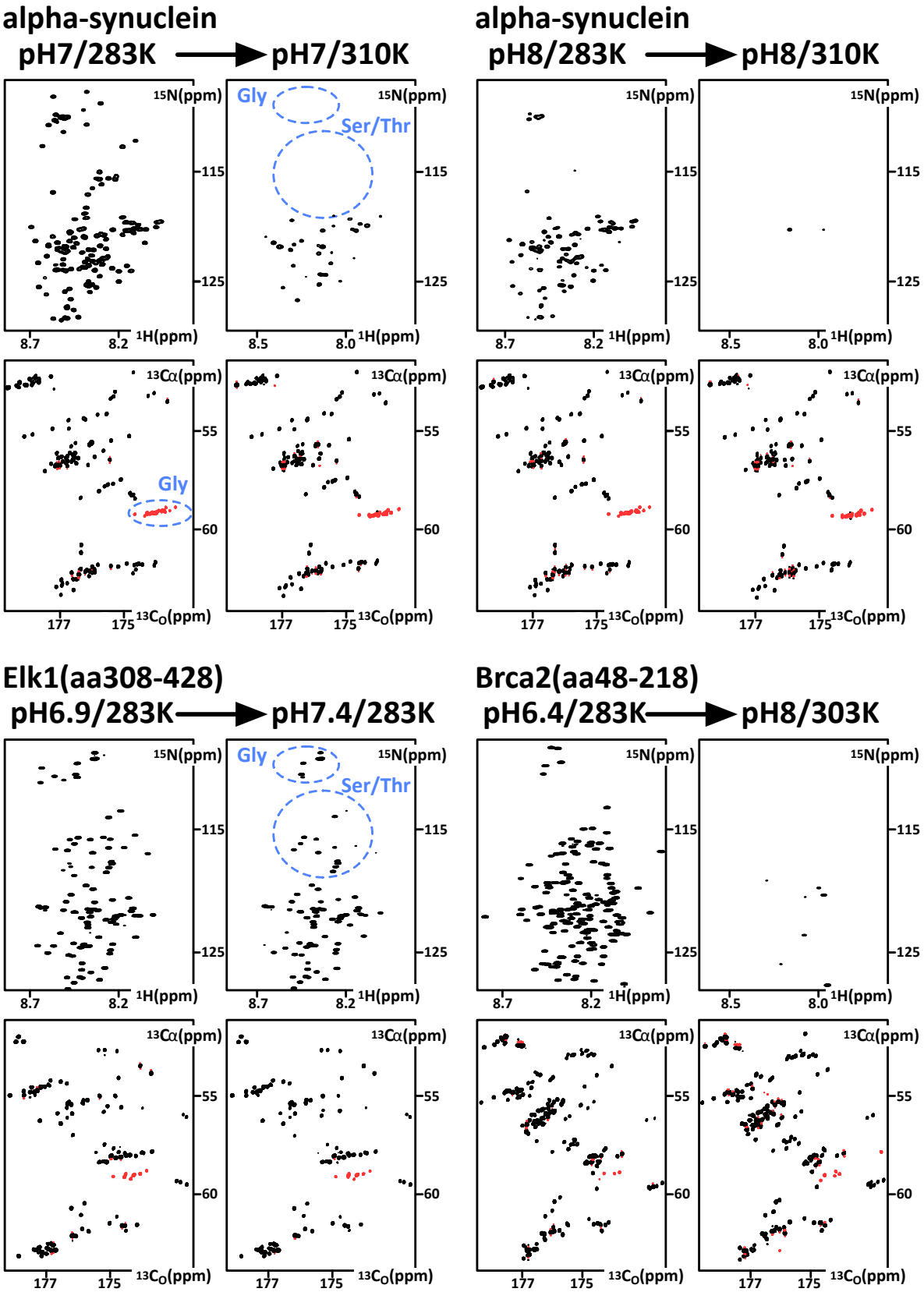
Phosphorylated residues generally provide lower peak intensities than their non-phospho counterparts, due to the difficulties to excite and decouple selectively Cα and Cβ of pSer and pThr.

It is important to use a low D_2O concentration to avoid high amounts of deuterated amides, which generate significant changes in chemical environment for the neighboring Cα and CO, hence provoking peak splitting or measurable exchange relaxation. We use 3% D_2O.

SUPPORTING INFORMATION

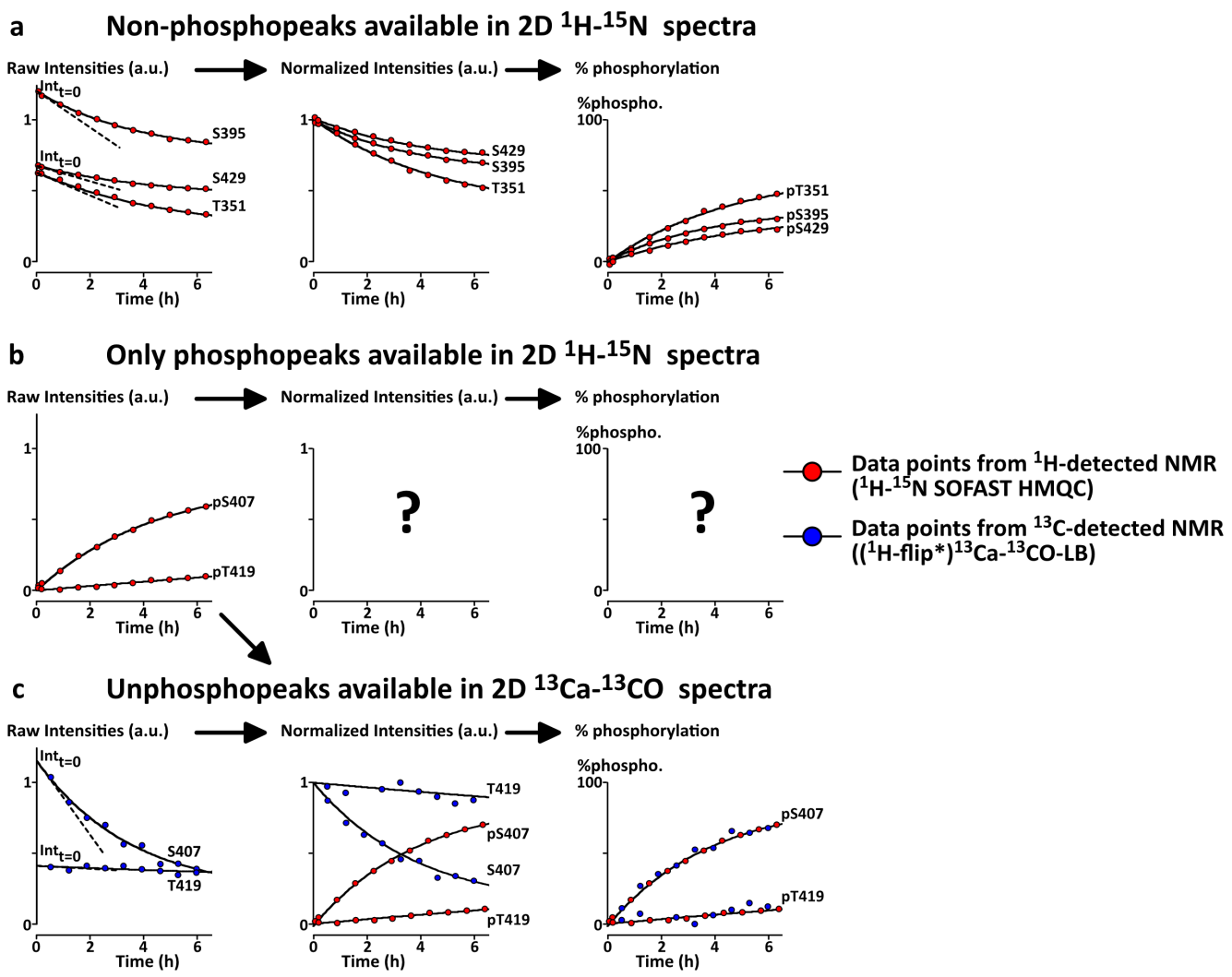
Results and Discussion

Figure S1. ^{1}H-^{15}N 2D HSQC and ^{13}Cα-^{13}CO 2D (^{1}H-flip*)^{13}Cα^{13}CO-LB spectra of alpha-synuclein, Elk1(aa308-428) and Brca2(aa48-218) in low pH and low temperature or high pH and high temperature conditions. The same acquisition parameters and contour levels were applied in both conditions. Positive and negative crosspeaks are in black and red, respectively. Negative peaks in red correspond to Gly residues.



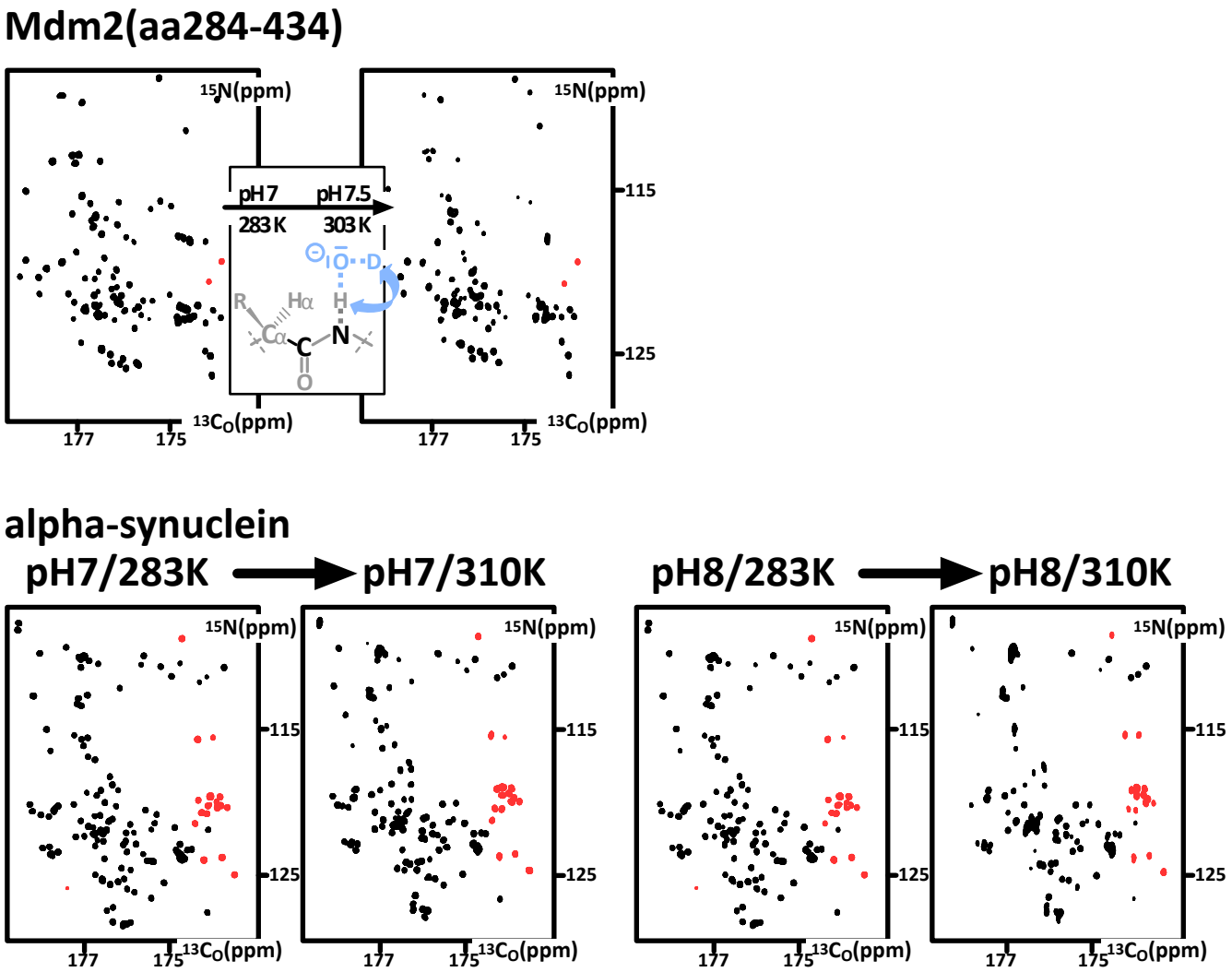
SUPPORTING INFORMATION

Figure S2. Analysis pipeline for data fitting using non-phosphopeaks or phosphopeaks from NMR spectra of Mdm2(aa284-434). **a.** In the case where non-phosphopeaks (from a phosphosite or its neighboring residues) are well-resolved and provide sufficient S/N, intensities measured in every spectrum of the time series can be plotted. They are then fitted using an exponential function $I(t)=I(t=0)*\exp(-t/k)+I(t_{final})$, the fitted parameters being $I(t=0)$, k and $I(t_{final})$. Curves are then normalized by $I(t=0)+I(t_{final})$ and $\%phosphorylation(t)=1-I_{normalized}(t)$. **b.** In the case where only phosphopeaks are well-resolved and provide sufficient S/N, intensities from phosphopeaks can be plotted, but not normalized, unless 100% phosphorylation is reached. **c.** Intensities from non-phosphopeaks can be measured in (^{1}H-flip*)^{13}Cα^{13}CO-LB spectra and the resulting curves can be fitted and used for normalization of both phosphopeaks and non-phosphopeaks build-up curves.



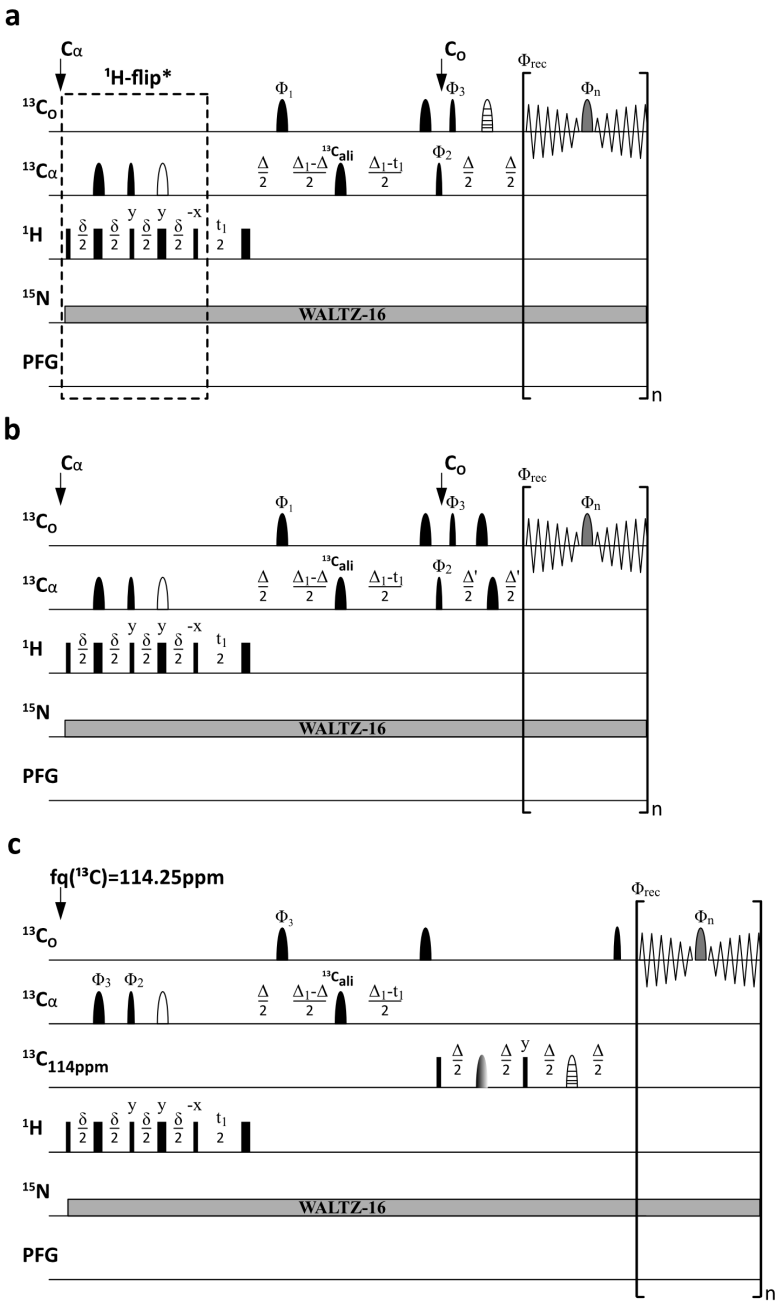
SUPPORTING INFORMATION

Figure S3. ^{15}N-^{13}CO 2D (^{1}H-flip*)CON-IPAP [12] spectra of Mdm2(aa284-434) and alpha-synuclein at low pH and low temperature or high pH and high temperature conditions. The same acquisition parameters and contour levels were applied in both conditions. Negative peaks in red correspond to Gly residues.



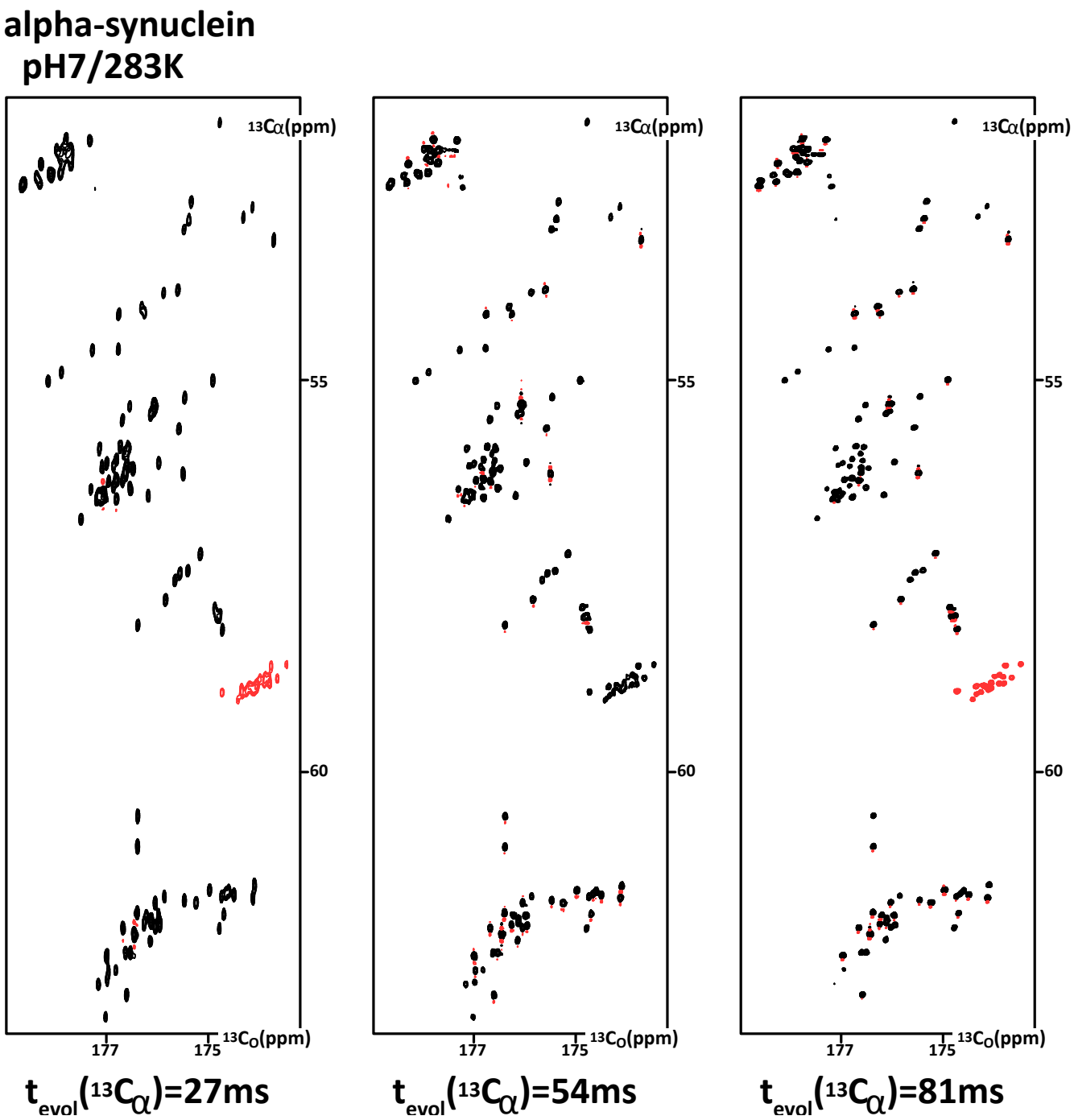
SUPPORTING INFORMATION

Figure S4. Schemes of the **(a)** (^{1}H-flip*)^{13}Cα^{13}CO-LB_1, **(b)** (^{1}H-flip*)^{13}Cα^{13}CO-LB_2 and **(c)** (^{1}H-flip*)^{13}Cα^{13}CO-LB_3 pulse sequences. Pulse lengths are adapted to a spectrometer field of 700 MHz. Delays are: δ=2.9 ms, Δ=9 ms, Δ'= 8.4 ms, Δ_1=27*n ms with n=1, 2 or 3. Phase cycle: ϕ_1=x,x,x,x,-x,-x,-x,-x; ϕ_2=x, -x; ϕ_3=x, x, -x, -x; ϕ_{rec}=x,-x,-x,x; ϕ_n=x,x,-x,-x,-x,x,x,-x,-x,-x,x,x,-x,-x,x. Filled and open rectangles represent hard 90° and 180° pulses, respectively. Narrow and wide black semi-ellipses are Q5-sebop/Q5tr-sebop 90° (231 μs) and Q3-surbop 180° (300 μs) pulses, respectively. ^{13}CO and ^{13}Cα/β and ^{13}C$_{ali}$ carrier frequencies are 173, 55.5 and 41 ppm, respectively. White semi-ellipses are ^{13}Cα centered Q3-surbop 180° (634 μs) pulses applied at 41 ppm. The striped semi-ellipse is a 120 μs triply compensated pulse G5 [13,14] applied at 114.25 ppm. The shaded semi-ellipse is a Q3-surbop 180° (900 μs) pulse applied at 114.25 ppm, amplitude modulated by a cosine function at a frequency of 58.75*ν_c (ν_c is the ^{13}C frequency in MHz), which inverts selectively both ^{13}CO and ^{13}Cα/β at the same time. The grey semi-ellipse is a center lobe of a sinc function of duration 22,600/ν_c in μs (ν_c is the ^{13}C frequency in MHz), amplitude modulated by a cosine function at a frequency of 118*ν_c applied n times every 5 ms, with n=t_{acq}/5ms. Quadrature detection is obtained by incrementing phase ϕ_2 in a States-TPPI manner. The original (^{1}H-flip*)^{13}Cα^{13}CO-IPAP sequence uses a pulse field gradient just before the carrier frequency change to ^{13}CO before the final INEPT transfer. We chose not to apply any pulsed field gradients (G_z), because we did not observe any improvement related to it.



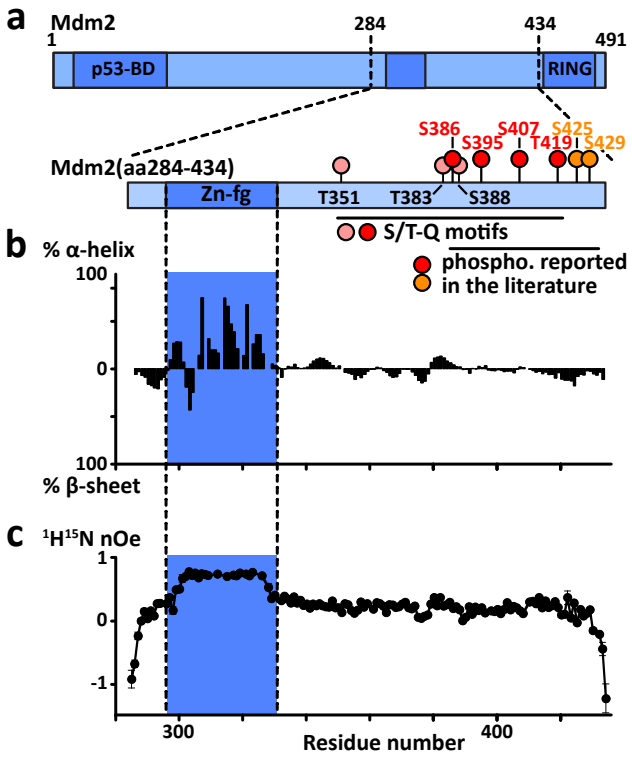
SUPPORTING INFORMATION

Figure S5. (^{1}H-flip*)^{13}Cα^{13}CO-LB spectra of alpha-synuclein using growing constant-time (CT) evolution periods of 27, 54 or 81 ms. These correspond to n/^{1}J(CαCβ) with n=1, 2 or 3. This allows the removal of ^{1}J(CαCβ) =35 Hz coupling effects during Cα chemical shift evolution. Increasing the constant time by 27 ms provokes on average a 30% loss in S/N for IDPs. It generates at the same time an improved resolution: we measured an average 12 Hz linewidth in Cα-dimension at 54 ms and 8 Hz at 81 ms. In these spectra, the signs of NMR signals oscillate according to J-coupling evolution: they are negative for cos(πJ*CT)=-1, i.e. CT=27 or 81 ms, and positive for cos(πJ*CT)=1, i.e. CT=54 ms. Gly residues are an exception, because they do not contain any Cα and their signal is always positive. We have changed the phase by 180° so that most of the peaks appear to be positive (black), and those of Gly residues appear negative at CT=27 or 81 ms.



SUPPORTING INFORMATION

Figure S6. (a) Primary structure of human Mdm2 and Mdm2(aa284-434). Folded domains (p53-BindingDomain: p53-BD; Zinc-finger: Zn-fg; RING-ubiquitin ligase domain: RING) are in deep blue and domains predicted to be disordered are in light-blue. DNA-PK/ATM/ATR (DDR kinases) were reported to phosphorylate sites depicted in solid red or orange circles, representing consensus S/T-Q motifs or non-S/T-Q motifs, respectively. S/T-Q motifs not yet detected to be DDR kinases targets are in light red circles.[15] **(b)** Secondary structure propensity as calculated from $^{13}C\alpha/^{13}C\beta$ chemical shifts using ncSCP.[8,9] **(c)** Low ^{1}H-^{15}N heteronuclear nOes (700 MHz) reveal ps-ns timescale motion typical of unfolded peptides outside of the Zn-fg domain.



SUPPORTING INFORMATION

Figure S7. (a) Primary structure of human Oct4 and Oct4(aa1-145). Folded domains (DNA-BindingDomain: DNA-BD or DBD) are in deep blue and domains predicted to be disordered are in light-blue. **(b)** 2D ^{1}H-^{15}N HSQC spectrum of Oct4(aa1-145) showing that most crosspeaks are in the region (in grey) where resonances of unfolded amino acids are found. **(c)** Secondary structure propensity as calculated from $^{13}C\alpha/^{13}C\beta$ chemical shifts using ncSCP.[8,9] Gly and Pro positions along the primary sequence are indicated by red and blue sticks, respectively.

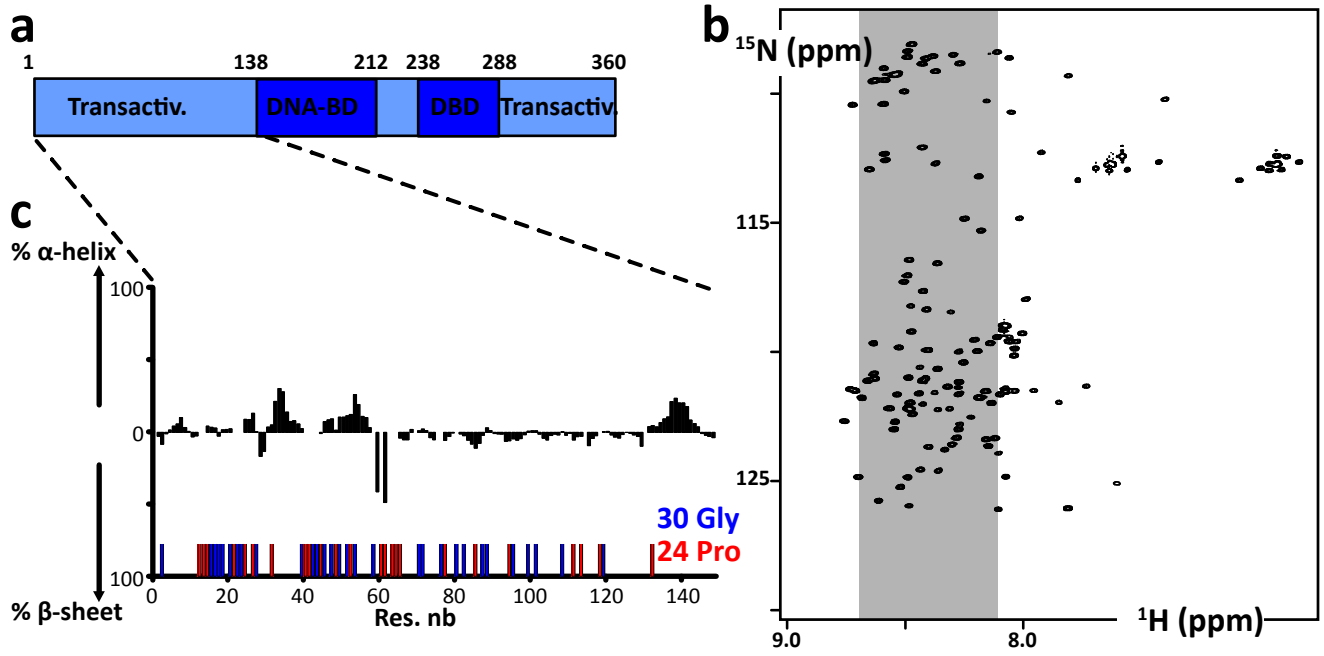
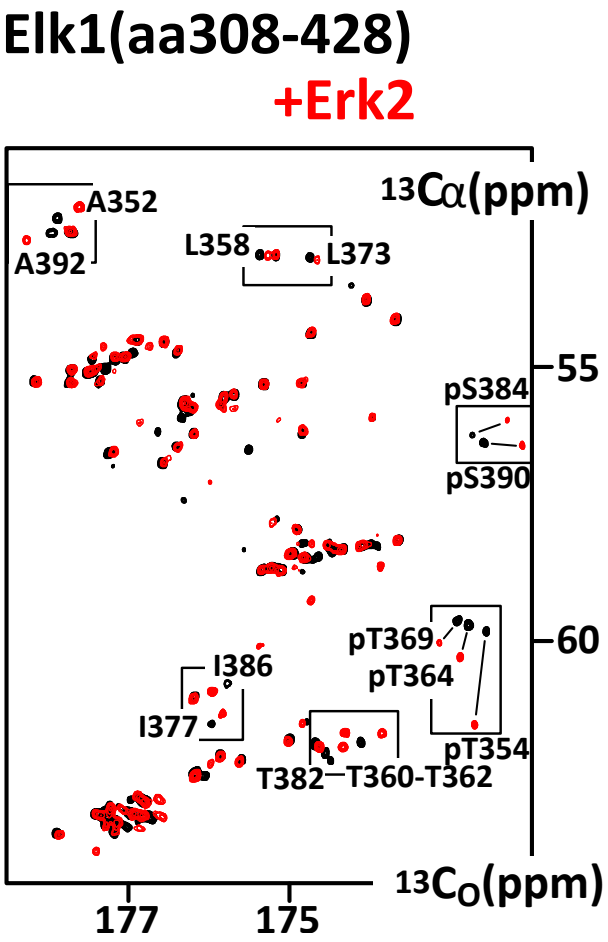


Figure S8. Overlay of (^{1}H-flip*)^{13}Cα^{13}CO-LB spectra of Elk1(aa308-428) before (black) and after (red) phosphorylation by the MAPKinase Erk2.



References

[1] M. Julien, S. Miron, A. Carreira, F.-X. Theillet, S. Zinn-Justin, *Biomolecular NMR Assignments* **2020**, *22*, 295–7.
[2] F.-X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb, et al., *Nature* **2016**, *530*, 45–50.
[3] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325–1327.
[4] E. Lescop, P. Schanda, B. Brutscher, *Journal of Magnetic Resonance* **2007**, *187*, 163–169.
[5] V. Jaravine, I. Ibraghimov, V. Y. Orekhov, *Nat. Meth.* **2006**, *3*, 605–607.
[6] V. Y. Orekhov, V. A. Jaravine, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 271–292.
[7] G. W. Yu, *Protein Science* **2006**, *15*, 384–389.
[8] K. Tamiola, F. A. A. Mulder, *Biochem. Soc. Trans.* **2012**, *40*, 1014–1020.
[9] J. T. Nielsen, F. A. A. Mulder, *J. Biomol. NMR* **2018**, *70*, 141–165.
[10] K. Kazimierczuk, V. Y. Orekhov, *Angew. Chem. Int. Ed.* **2011**, *50*, 5556–5559.
[11] J. Ying, F. Li, J. H. Lee, A. Bax, *J. Biomol. NMR* **2014**, *60*, 15–21.
[12] W. Bermel, I. Bertini, I. C. Felli, R. Pierattelli, *J. Am. Chem. Soc.* **2009**, *131*, 15339–15345.
[13] V. S. Manu, G. Veglia, *J. Magn. Reson.* **2015**, *260*, 136–143.
[14] Y. Xia, P. Rossi, M. V. Subrahmanian, C. Huang, T. Saleh, C. Olivieri, C. G. Kalodimos, G. Veglia, *J. Biomol. NMR* **2017**, *69*, 237–243.
[15] Q. Cheng, B. Cross, B. Li, L. Chen, Z. Li, J. Chen, *Mol. Cell. Biol.* **2011**, *31*, 4951–4963.

Author Contributions

AA, CB, MJ contributed equally. AA, CB, MJ and FXT conducted the research, i.e. sample preparation, data acquisition and analysis. SZJ helped with NMR measurements. FXT achieved resonance assignment, kinetics analysis. WB and FXT optimized the pulse sequences. SZJ and FXT obtained fundings. FXT designed the research and wrote the manuscript. Correspondence and requests for materials should be addressed to FXT.