Site-Specific Mapping and Time-Resolved Monitoring of Lysine Methylation by High-Resolution NMR Spectroscopy

François-Xavier Theillet, Stamatios Liokatis, Jan Oliver Jost, Beata Bekei, Honor May Rose, Andres Binolfi, Dirk Schwarzer, Philipp Selenko

Site-specific mapping and time-resolved monitoring of Lysine methylation by high-resolution NMR spectroscopy.

François-Xavier Theillet, Stamatios Liokatis, Jan Oliver Jost†, Beata Bekei, Honor May Rose, Andres Binolfi, Dirk Schwarzer†, Philipp Selenko*

Leibniz Institute of Molecular Pharmacology (FMP Berlin), Robert Roessle Strasse 10, 13125 Berlin, Germany

Supporting Information Placeholder

ABSTRACT: Methylation and acetylation of protein lysine residues constitute abundant post-translational modifications (PTMs) that regulate a plethora of biological processes. In eukaryotic proteins, lysines are often mono-, di-, or tri-methylated, which may signal different biological outcomes. Deconvoluting these different PTM types and PTM states is not easily accomplished with existing analytical tools. Here, we demonstrate the unique ability of NMR spectroscopy to discriminate between lysine acetylation and mono-, di-, tri-methylation in a site-specific and quantitative manner. This enables mapping and monitoring of lysine acetylation and methylation reactions in a non-disruptive and continuous fashion. Time-resolved NMR measurements of different methylation events in complex environments including cell extracts contribute to our understanding of how these PTMs are established *in vitro* and *in vivo*.

Protein lysine residues are subject to two main types of chemically distinct post-translational modifications: lysine acetylation (Figure 1a) and lysine methylation (Figure 1b). Here we describe time-resolved, high-resolution NMR experiments to monitor lysine acetylation and methylation reactions in parallel. We identify the key NMR characteristics of lysine mono-, di- and tri-methylation and outline experimental approaches to identify lysine methylation states and to map specific lysine methylation sites in a residue-resolved manner both *in vitro* and in complex environments such as mammalian cell extracts.

While acetylation neutralizes the positive charge of lysine residues (Figure 1a), lysine methylation does not. In addition, lysines can be methylated in mono-, di-, and tri-methylated forms (Figure 1b). These differences in chemical and physical properties of lysine side-chains have important functional consequences, especially in acting as switches to signal different biological responses[1-4]. In many instances, a given lysine residue can exist in either an acetylated-, or differentially methylated form. Variations in lysine acetylation/methylation are therefore tightly controlled and of fundamental importance in many areas of eukaryotic biology.

In contrast to other analytical methods[5-7], high-resolution NMR spectroscopy can detect differentially modified lysine residues in a straightforward manner. As acetylated and methylated lysines display very different NMR properties, it is easy to discriminate between these two PTMs. We and others have previously delineated the NMR characteristics of lysine acetylation[8-10]. In brief, acetylation results in up-field chemical shift changes ($\Delta\delta(^1H)$~0.1 ppm) of backbone amide resonances of the modified lysine residues that are readily detected in 2D hetero-nuclear (^1H-^{15}N) correlation experiments. In addition, each acetylation event produces one additional amide resonance signal, the generic *acetylation indicator* at ~7.9/127.5 ppm (^1H-$^{15}N\varepsilon$) that corresponds to the newly formed side-chain amide group of modified lysine residue (Figure 1a).

Lysine mono-, di-, or tri-methylation does not lead to detectable

lysine backbone amide chemical shift changes (data not shown). Instead, methylation produces characteristic changes in proton/carbon (^1H-^{13}C) lysine side-chain correlations (Figure 1c). Especially lysine $CH_2\varepsilon$ resonances experience large ^{13}C-carbon down-field chemical shift changes upon methylation. $CH_2\varepsilon$ groups of non-methylated lysines ($CH_2\varepsilon$-Me_0) resonate at ~3.0/42.0 ppm

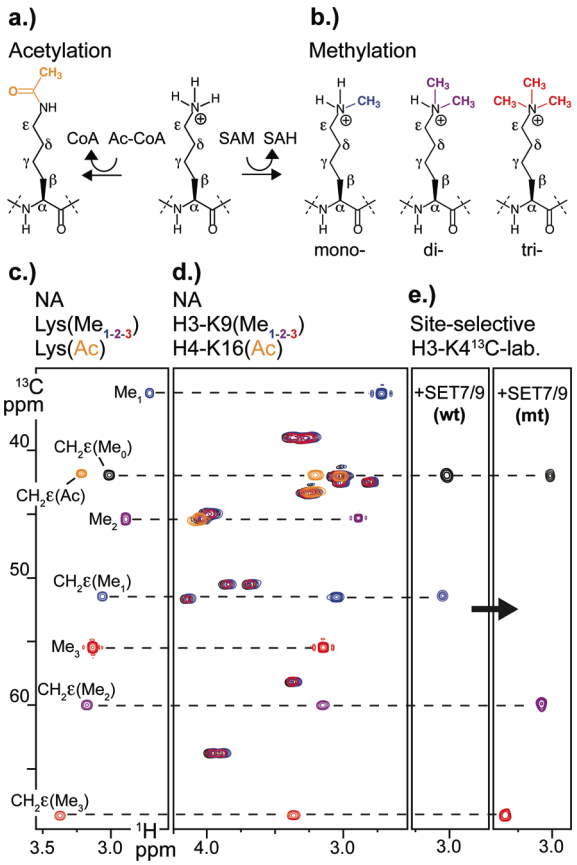


Figure 1. Chemical structures of **a)** acetylated and **b)** mono-, di- and tri-methylated lysines. **c)** Overlay of 2D natural abundance (NA) ^{1}H-^{13}C correlation spectra of non-, mono-, di- and tri-methylated and acetylated lysines (black, blue, purple, red and orange, respectively). **d)** Overlay of 2D NA ^{1}H-^{13}C correlation spectra of histone H3 (aa1-20) non-, mono-, di- and tri-methylated on K9, and of histone H4 (aa1-19) acetylated on K16. **e)** Overlay of 2D ^{1}H-^{13}C correlation spectra of H3 (aa1-15), site-selective K4 ^{13}C-labeled, mono-methylated by wild-type (wt) SET7/9 and successively di- and tri-methylated by mutant (mt) SET7/9 (Y245A).

(^{1}H-^{13}C), while mono- (~3.1/51.0 ppm, $CH_2\varepsilon$-Me_1), di- (~3.2/60.0 ppm, $CH_2\varepsilon$-Me_2), or tri-methylated (~3.4/68.5 ppm, $CH_2\varepsilon$-Me_3) lysines display uniquely different correlation signals. In contrast, $CH_2\varepsilon$ resonances of acetylated lysines display characteristic down-field chemical shift changes in the proton dimension (~3.15/42.0 ppm, $CH_2\varepsilon$-Ac), which enables facile discrimination between non-modified, acetylated, or methylated lysines (Figure 1c). The lysine methyl groups resonate at ~2.7/35.5 ppm (mono-, Me_1), ~2.9/45.5 ppm (di-, Me_2), ~3.1/55.5 ppm (tri-methyl, Me_3), while the methyl groups of acetyl-lysines have a resonance frequency of ~2.0/24.0 ppm (acetyl, Ac) as reported previously[11-13]. Because most modifiable lysine residues in folded and unfolded proteins are solvent exposed and sample similar chemical environments, the aforementioned NMR properties are generally preserved (Figure 1d and unpublished data). Thus, lysine $CH_2\varepsilon$ resonances function as unique identifiers of acetylation and methylation states, and additionally report the respective forms of lysine methylation in an unbiased manner.

We set out to explore the suitability of these NMR properties to report lysine methylation states by probing a prototypic lysine methyltransferase (KMT), recombinant SET7/9, on histone H3. In a reaction with a synthetic H3 peptide (aa1-15), selectively ^{13}C-labeled at Lys4 (K4), wild-type (wt) SET7/9 displayed exclusive mono-methylation (Figure 1e), in line with published data[14]. Furthermore, and in accordance with published data, a SET7/9 mutant (mt) (Y245A) di- and tri-methylated K4 when the substrate was presented in a pre-mono-methylated form (Figure 1e)[14, 16].

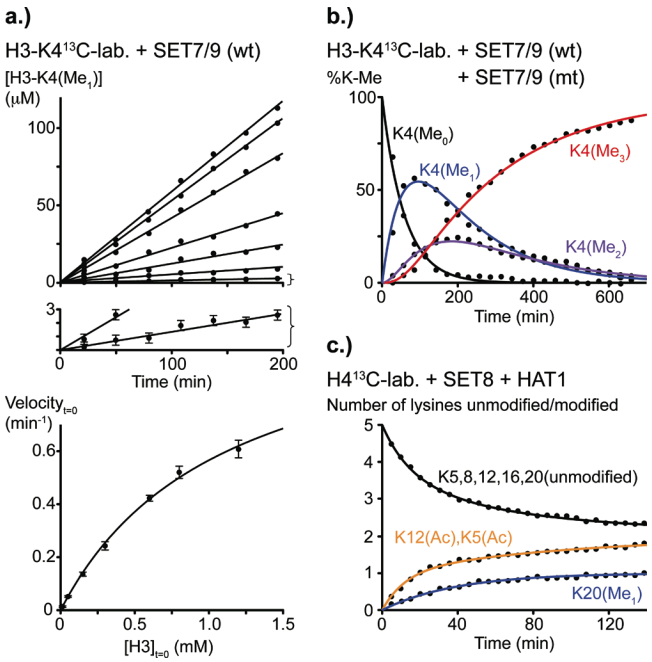


Figure 2. a) *Top:* Histone H3 K4 mono-methylation kinetics by SET7/9 (wt). The expansion below depicts the sizes of measurement errors due to experimental noise, at the lowest enzyme concentration. *Bottom*: Michaelis-Menten plot of initial velocities against H3 substrate concentrations. Error bars were obtained from two independent measurements. **b)** Build-up curves of simultaneous H3 K4 (100 μM in 100 μL) mono-, di- and tri- methylation reactions in a mixture of wt and mt SET7/9. **c)** Build-up curves of simultaneous histone H4 K20 (100 μM in 300 μL) mono-methylation by SET8 and H4 K12 and K5 acetylation by HAT1. Conversion percentages are directly calculated from peak integrals in **b)** and **c)**. In **c)** NMR time points were obtained every 2.5 minutes. For clarity, time intervals of only 5 minutes are shown.

Together, these results indicated that NMR detection of H3 K4 methylation faithfully recapitulated the known modification behaviors of wild-type and mutant SET7/9.

We set up a tailored ^{1}H-^{13}C presat-SOFAST-HMQC pulse sequence[1,5] to specifically detect lysine $CH_2\varepsilon$ resonances with high sensitivity (Figure S1 in the Supporting Information), which, in turn, enabled rapid quantifications of lysine methylation states by simple cross peak integration routines (see also Supporting Information). Following this rationale, methylated lysines in substrates at low concentrations (12.5 μM, ^{13}C-labeled) could accurately be measured within 30 minutes of experimental time (experimental error +/- 2 %, 750MHz, TCI cryo-probe). The same level of precision could be achieved four times faster at twice this concentration and with the same number of scans, as the NMR signal-to-noise ratio increases linearly with concentration, or with the square of the number of scans. Hence, kinetics of enzymatic methylation reactions

that produced at least 5 pmol.min^{-1} of methylated substrate species were reliably reported.

To exemplify this notion, we followed the methylation kinetics of SET7/9 by time-resolved NMR monitoring using the above pulse scheme. Initial velocity measurements of methylation rates at different substrate concentrations (12.5μM to 1.2mM) were performed with an average quantification error of +/- 0.25 μM, or 75 pmol (Figure 2a). Linear fitting to the Michaelis-Menten kinetic model revealed a k_{cat} of 1.1 +/- 0.1 min^{-1} and a K_m of 1.1 +/- 0.2 mM. These values were in excellent agreement with previously reported methylation rates that had been obtained in spectrophotometric assays[14]. We also delineated the respective modification characteristics in a mixed (wt and mt SET7/9) reaction setup (Figure 2b). As expected, H3 K4 mono-methylation was followed by di-, trimethylation, which indicated that mechanistic insights into SET7/9-specific methylation properties could be obtained from the individual modification profiles in a single set of time-resolved NMR experiments.

As lysine acetylation and methylation produce distinctively different ^{1}H-^{13}C correlations, we asked whether we could monitor and individually characterize both types of PTMs in the same NMR experiment. To answer this question, we set up a mixed acetylation/methylation reaction using uniformly ^{13}C-labeled histone H4 (aa1-26) and catalytic amounts of unlabeled, recombinant HAT1 and SET8 (Figure 2c). We had previously shown that the acetyltransferase HAT1 exclusively modified K5 and K12 of histone H4[9]. SET8, in turn, is a well-characterized H4 K20-specific mono-methyltransferase[17]. Following the respective modification profiles by time-resolved 2D ^{1}H-^{13}C correlation experiments, we found that both PTM reactions produced NMR readouts that quantitatively reported the individual modification events. Moreover, both kinetic profiles suggested that acetylation of H4 K5 and K12, and methylation of K20 did not influence each other and proceeded with similar rates as under isolated conditions (data not shown).

While NMR monitoring of lysine methylation based on the aforementioned NMR chemical shift parameters is straightforward, site specific mapping of individual methylation sites is more challenging. Due to the limited chemical shift dispersion of lysine side-chain resonances (^{1}H-^{13}C) in peptides and proteins, all residue-specific information about methylation sites provided thus far was solely afforded by site-selective ^{13}C-lysine labeling, or existing knowledge about KMT target sites. To also enable *de novo* NMR mapping of methylated lysine residues, we took advantage of the large Ce-carbon chemical shift differences between non-, mono-, di-, and trimethylated lysines and the superior chemical shift dispersion of lysine backbone amide resonances in 2D ^{1}H-^{15}N correlation spectra. A 2D (HCe(Me$_X$)-TOCSY-Ca)NH pulse scheme was developed, in which carbon-Ce resonances were selectively excited at either the mono-, di-, or tri-methylated lysine ^{13}Ce resonance frequencies (Figure S2 in the Supporting Information). In combination with lysine-selective ^{13}C/^{15}N labeling schemes, only those lysine residues that existed in matched methylation states gave rise to 2D ^{1}H-^{15}N correlation cross peaks.

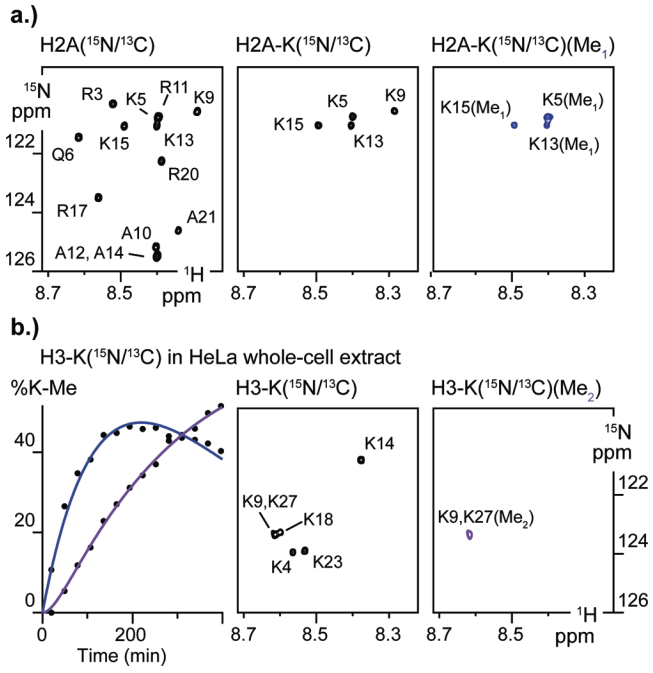


Figure 3. a) *Left:* Selected region of the 2D ^{1}H-^{15}N correlation spectrum of uniformly labeled histone H2A (aa1-23, 150 μM in 100 μL). *Middle:* Same region of the 2D ^{1}H-^{15}N correlation spectrum of ^{15}N/^{13}C/ lysine-selective labeled H2A. *Right:* Identical view of the 2D (HCε(Me1)-TOCSY-Cα) ^{1}H-^{15}N NMR spectrum shows mono-methylated H2A lysines K5, K13 and K15, but not K9, after the reaction with SET7/9. **b)** *Left:* Time course of lysine mono- and di-methylation of H3 (aa1-33)-GB1 (200 μM in 200 μL) in HeLa cell extract. *Middle:* Selected region of the 2D ^{1}H/^{15}N correlation spectrum of lysine-selective labeled H3 (aa1-33)-GB1. *Right:* Identical view of the 2D (HCε(Me1)-TOCSY-Cα) ^{1}H-^{15}N NMR spectrum shows di-methylated H3 lysines K9 and/or K27, after the reaction in HeLa extracts.

Therefore, 'scanning' of the characteristic CH_2ε-Me$_X$ resonance frequencies enabled residue-specific mapping of correspondingly methylated lysine residues. To exemplify this notion, we reacted lysine-specific, ^{13}C/^{15}N-labeled histone H2A (aa1-23) with SET7/9 (wt) (Figure 3a). Multi-site methylation of N- terminal histone H2A by wild-type SET7/9 has recently been reported, although no information about possible methylation sites had been obtained[18]. NMR measurements revealed that SET7/9 mono-methylated K5, K13 and K15 of H2A, but not K9 (Figure 3b). No di- or tri-methylated species were detected (data not shown). 2D (HCε(Me$_1$)-TOCSY-Cα)NH NMR spectra were acquired on 50 μM H2A samples (15 nmol) within 4 hours of acquisition time. Hence, the outlined scheme proved particularly suitable to map methylation sites, as well as to identify the respective methylation states. Alternatively, fast 2D ^{1}H-^{13}C experiments can initially be employed to qualitatively check for the presence of methylated lysine residues, to monitor the progression of lysine methylation in a time-resolved fashion and to identify the types of methylation states that are being established. The TOCSY-type NMR experiment may then be employed to map the respective methylation sites.

To determine whether NMR spectroscopy could also be used to directly monitor lysine methylation by endogeneous KMTs in cellular environments, we prepared lysates from cultured human HeLa cells, to which we added recombinant, ^{13}C/^{15}N-lysine labeled histone H3 (aa1-33) C-terminally pro-

tected by fusion to the *Streptococcal* protein G B1 domain (GB1). The GB1 moiety prevented unspecific C-terminal proteolysis of the free histone H3 peptide, which was observed at prolonged extract incubation times (>6 hrs.). We detected lysine mono- and subsequent di-methylation in a highly reproducible manner in different extract preparations (Figure 3b). Minor tri-methylation (~1%) was only observed at later time points (data not shown). The initial rate of mono-methylation, for 200 μM of substrate, was determined to correspond to 50 +/- 10 pmol.min^{-1} per milligram of total lysate protein. Employing the (HCε(Me$_X$)-TOCSY-Cα)NH NMR experiment, mono/di-methylation was mapped to H3 K9 and/or K27 (Figure 3b). Due to spectral overlap, no unambiguous assignment of the methylation site(s) could be obtained. However, both residues are known target sites for mono- and di-methylation by several cellular KMTs, including G9a and GLP[19-21], which establish transcriptionally silent chromatin structures *in vivo*[22, 23]. Both KMTs also methylate the transcription factor and oncoprotein p53[24] and constitute prominent drug targets for cancer therapy[25].

In summary, NMR monitoring of lysine methylation and acetylation by 2D ^{1}H-^{13}C correlation experiments offers a convenient means to confirm the presence and nature of these PTMs and in the case of lysine methylation, to determine the respective methylation states. In comparison to other analytical methods[5], which are mostly based on indirect readout schemes that require narrowly defined *in vitro* experimental setups and preclude measurements in complex environments such as cell extracts, NMR spectroscopy allows direct quantitative monitoring of acetylation and methylation reactions in a time-resolved manner in a single experiment. While these extract-based NMR measurements may not accurately reflect the spatially regulated, localized enzymatic activities of intact cells, they nevertheless enable comparative assessments of global HAT and KMT activities in different cellular environments and of changes in enzymatic activities upon internal, or external cues. Methyl- and acetyl-transferase enzymes, as well as their de-modifying counterparts i.e. de-methylases and de-acetylases, constitute prominent drug targets[26-29]. The advantages of NMR spectroscopy for the study of these PTMs makes it an ideal tool to analyze the modes of action of the above enzymes and to elucidate the mechanisms by which they regulate the *writing* and *reading* of PTM-based signaling codes *in vitro*, in complex environments such as cell extracts and possibly also in intact cells.

ASSOCIATED CONTENT

Supporting Information. Supplementary figures, Material and Methods and pulse sequence codes.

AUTHOR INFORMATION

Corresponding Author

* E-mail: selenko@fmp-berlin.de

Present Addresses

† Present address: Interfaculty Institute of Biochemistry (FIB), University of Tübingen, Hoppe-Seyler-Str. 4, D-72076 Tübingen (Germany).

ACKNOWLEDGMENT

We thank Pr. Dr. Raymond C. Trievel for providing the plasmids for recombinant SET7/9 and SET8 expression and Dr. Peter Schmieder and Monika Beerbaum for excellent NMR infrastructure maintenance. F.X.T was supported by a post-doctoral grant from the Association pour la Recherche contre le Cancer (ARC). P.S. is supported by an Emmy Noether program grant (SE-1794/1-1) by the Deutsche Forschungs Gemeinschaft (DFG).

ABBREVIATIONS

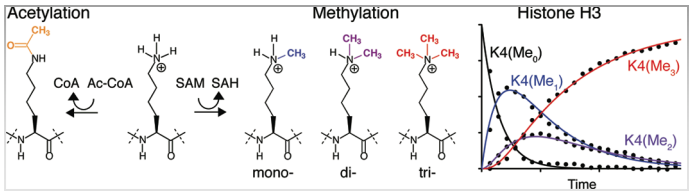
KMT, lysine methyltransferase; mt, mutant; NA, natural abundance; PTM, post-translational modification; SAH, *S*-adenosyl homocysteine; SAM, *S*-adenosyl methionine; wt, wild type; mt, mutant.

REFERENCES

(1) Barth, T. K.; Imhof, A. Trends Biochem. Sci. **2010**, 35, 618-626.
(2) Choudhary, C.; Kumar, C.; Gnad, F.; Nielsen, M. L.; Rehman, M.; Walther, T. C.; Olsen, J. V.; Mann, M. Science **2009**, 325, 834-840.
(3) Egorova, K. S.; Olenkina, O. M.; Olenina, L. V. Biochemistry (Moscow) **2010**, 75, 535-548.
(4) Bannister, A. J.; Kouzarides, T. **Cell Res.** 2011, 21, 381-395.
(5) Luo, M. ACS Chem. Biol. **2011**.
(6) Zhang, K.; Yau, P. M.; Chandrasekhar, B.; New, R.; Kondrat, R.; Imai, B. S.; Bradbury, M. E. Proteomics **2004**, 4, 1-10.
(7) Zee, B. M.; Levin, R. S.; Dimaggio, P. A.; Garcia, B. A. Epigenetics Chromatin **2010**, 3, 22.
(8) Smet-Nocca, C.; Wieruszeski, J. M.; Melnyk, O.; Benecke, A. J. Pept. Sci. **2010**, 16, 414-423.
(9) Liokatis, S.; Dose, A.; Schwarzer, D.; Selenko, P. J. Am. Chem. Soc. **2010**, 132, 14704-14705.
(10) Dose, A.; Liokatis, S.; Theillet, F. X.; Selenko, P.; Schwarzer, D. ACS Chem. Biol. **2010**, 6, 419-424.
(11) Ashfield, J. T.; Meyers, T.; Lowne, D.; Varley, P. G.; Arnold, J. R.; Tan, P.; Yang, J. C.; Czaplewski, L. G.; Dudgeon, T.; Fisher, J. Protein Sci. **2000**, 9, 2047-2053.
(12) Macnaughtan, M. A.; Kane, A. M.; Prestegard, J. H. J. Am. Chem. Soc. **2005**, 127, 17626-17627.
(13) Abraham, S. J.; Kobayashi, T.; Solaro, R. J.; Gaponenko, V. J. Biomol. NMR **2009**, 43, 239-246.
(14) Couture, J. F.; Collazo, E.; Hauk, G.; Trievel, R. C. Nat. Struct. Mol. Biol. **2006**, 13, 140-146.
(15) Schanda, P.; Kupce, E.; Brutscher, B. J. Biomol. NMR **2005**, 33, 199-211.
(16) Del Rizzo, P. A.; Couture, J. F.; Dirk, L. M.; Strunk, B. S.; Roiko, M. S.; Brunzelle, J. S.; Houtz, R. L.; Trievel, R. C. J. Biol. Chem. **2010**, 285, 31849-31858.
(17) Couture, J. F.; Dirk, L. M.; Brunzelle, J. S.; Houtz, R. L.; Trievel, R. C. Proc. Natl. Acad. Sci. U S A **2008**, 105, 20659-20664.
(18) Dhayalan, A.; Kudithipudi, S.; Rathert, P.; Jeltsch, A. Chem. Biol. **2011**, 18, 111-120.
(19) Esteve, P. O.; Patnaik, D.; Chin, H. G.; Benner, J.; Teitell, M. A.; Pradhan, S. Nucleic Acids Res. **2005**, 33, 3211-3223.
(20) Wu, H.; Min, J.; Lunin, V. V.; Antoshenko, T.; Dombrovski, L.; Zeng, H.; Allali-Hassani, A.; Campagna-Slater, V.; Vedadi, M.; Arrowsmith, C. H.; Plotnikov, A. N.; Schapira, M. PLoS One **2010**, 5, e8570.
(21) Wu, H.; Chen, X.; Xiong, J.; Li, Y.; Li, H.; Ding, X.; Liu, S.; Chen, S.; Gao, S.; Zhu, B. Cell Res. **2011**, 21, 365-367.
(22) Kouzarides, T. Cell **2007**, 128, 693-705.
(23) Suganuma, T.; Workman, J. L. Annu. Rev. Biochem. **2011**, 80, 473-499.
(24) Huang, J.; Dorsey, J.; Chuikov, S.; Perez-Burgos, L.; Zhang, X.; Jenuwein, T.; Reinberg, D.; Berger, S. L. J. Biol. Chem. **2010**, 285, 9636-9641.
(25) Chang, Y.; Zhang, X.; Horton, J. R.; Upadhyay, A. K.; Spannhoff, A.; Liu, J.; Snyder, J. P.; Bedford, M. T.; Cheng, X. Nat. Struct. Mol. Biol. **2009**, 16, 312-317.
(26) Haberland, M.; Montgomery, R. L.; Olson, E. N. Nat. Rev. Genet. **2009**, 10, 32-42.
(27) Kelly, T. K.; De Carvalho, D. D.; Jones, P. A. Nat. Biotechnol. **2010**, 28, 1069-1078.

(28) Spannhoff, A.; Hauser, A. T.; Heinke, R.; Sippl, W.; Jung, M. ChemMedChem **2009**, 4, 1568-1582.
(29) Copeland, R. A.; Solomon, M. E.; Richon, V. M. Nat. Rev. Drug Discov. **2009**, 8, 724-732.

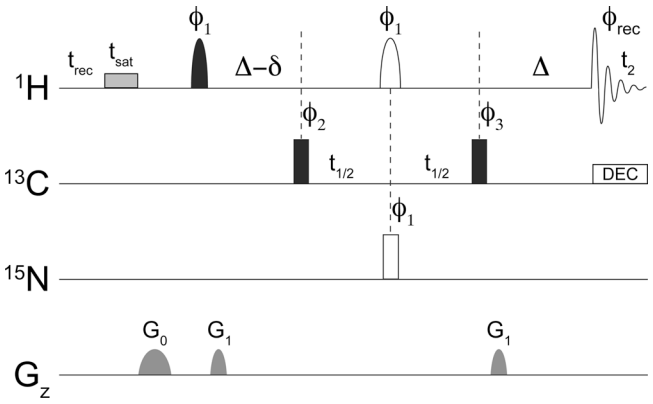
Synopsis TOC



Site-specific mapping and time-resolved monitoring of lysine methylation by high-resolution NMR spectroscopy

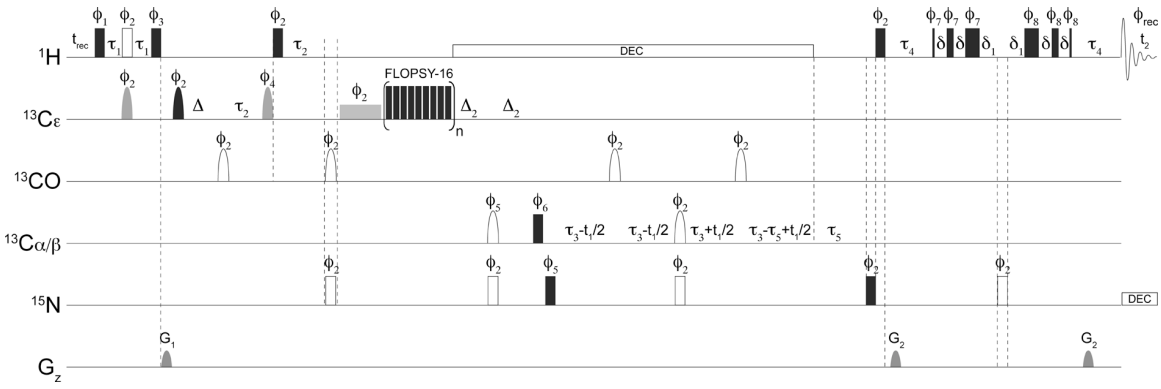
François-Xavier Theillet, Stamatios Liokatis, Jan Oliver Jost, Beata Bekei, Honor May Rose, Andres Binolfi, Dirk Schwarzer, Philipp Selenko

Figure S1



Scheme of the $^{1}H/^{13}C$ presat-SOFAST-HMQC pulse sequence. Delays are: t_{rec}=30 ms, t_{sat}=50 ms, Δ=3.08ms, δ=1.27 ms. Phase cycle: Φ_1=x; Φ_2=x, -x; Φ_3=x, x, -x, -x; Φ_{rec}=x,(-x),(-x),x. Quadrature detection is obtained by incrementing Φ_2 in a States-TPPI manner. Filled and open rectangles represent hard 90° and 180° pulses, respectively. Black, and open semi-ellipses are PC9 90° (3400 μs) and REBURP 180° (2800 μs) pulses, respectively, with a ^{1}H frequency carrier of 3.0 ppm for lysine-Hε selection. Pulsed field gradients (G_z) are shown as grey shapes, G_0 and G_1 are 4 ms and 1 ms, respectively.

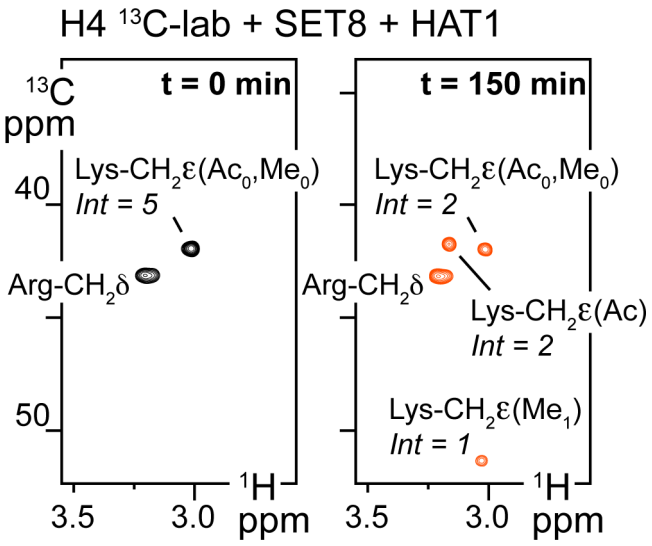
Figure S2



Scheme of the (HCε-TOCSY-Cα)NH pulse sequence. Delays are: t_{rec}=250 ms, τ_1=1.75 ms, Δ=2*pulse$_{1H}$, τ2=1.05 ms, Δ2=11 ms, τ3=5.7 ms, τ5=5.5 ms, τ4=2.25 ms, δ=168 μs, δ_1= (δ –pulse$_{15N}$)/2. Phase cycle: Φ_1=y; Φ_2=x; Φ_3=x, -x; Φ_4=4(x),4(y),4(-x),4(-y); Φ_5=4(x),4(-x); Φ_6=2(x),2(-x), Φ_7=y; Φ_8=-y; Φ_{rec}=x,(-x),(-x),x,(-x),x,x,(-x),(-x),x,x,(-x),x,(-x),(-x),x. 8 cycles of FLOPSY transfer were used (50 ms). Water suppression is achieved via the final 3-9-19 water-gate pulse train. Quadrature detection is obtained by incrementing Φ_5 in a States-TPPI manner. Filled and open rectangles represent hard 90° and 180° pulses, respectively. Black, grey and open semi-ellipses are Q5 90° (400 μs), Q3 180° (860 μs) and Q3 180° (220 μs) pulses[1], respectively, with a ^{13}C frequency carrier of 39 ppm for selection of non-methylated lysine Cε, 54 ppm for mono- and di-methylation, 63 ppm for di- and tri-methylation and 74 for tri-methylation. Longer ^{13}Cε pulses (~1200 μs) can be employed for more stringent selectivity. However, this will negatively impact the overall sensitivity of the experiment. Pulsed field gradients (G_z) are

shown as grey shapes. The grey rectangle before the Cε/CO magnetization transfer via the FLOPSY-16 pulse train denotes a low power trim pulse (1 ms)[2].

Figure S3



2D $^{1}H/^{13}C$ correlation spectra of H4(aa1-26) simultaneously methylated by SET8 and acetylated by HAT1. Values from integration of peaks are shown in italic.

Materials and Methods:

Recombinant protein production and isotope labeling:

Plasmids of histone H3 (aa1-33), histone H4 (aa1-26), histone H2a (aa1-23) and H3 (aa1-33)-GB1 were generated as previously described[3]. Recombinant histone peptides were produced as N-terminal 6xHis/GB1-fusions, followed by a TEV protease cleavage site and the respective peptides of interest. Yeast HAT1 was prepared as a His-tag fusion protein, as previously described[4]. Expression plasmids for the catalytic KMT domains of wild-type and mutant human SET7/9 (aa110-366) and SET8 (aa191-352) were kind gifts from Dr. Raymond C. Trievel. These KMTs were prepared as described[5-7].

Proteins/peptides were expressed in *E.coli* BL21 (DE3)-Express cells (Lucigen). KMTs/HATs were expressed in standard LB medium. Histone substrates were produced

in uniformly-, or lysine-selective isotope labeled forms. Uniform isotope labeling was achieved in minimal M9 medium containing $^{15}NH_4Cl$, or $^{15}NH_4Cl/^{13}C$-glucose for single/doubly labeled samples. Lysine-selective isotope labeling was performed via a two-step protocol: first *E.coli* BL21 (DE3)-Express cells were grown in 800 mL of unlabeled LB until an OD_{600} of 0.8, centrifuged and resuspended in 200 mL of M9 medium containing the following unlabeled amino acids: Ala (76 mg/L), Arg (56 mg/L), Asp (48 mg/L), Asn (48 mg/mL), Cys (140 mg/L), Glu (51 mg/L), Gln (51 mg/L), Gly (64 mg/L), His (23 mg/L), Ile (31 mg/L), Leu (84 mg/L), Met (16 mg/L), Phe (83 mg/L), Pro (56 mg/L), Ser (44 mg/L), Thr (384 mg/L), Tyr (38 mg/L), Trp (40 mg/L), Val (45 mg/L), and isotope-labeled $^{13}C/^{15}N$ Lys (240 mg/L). Recombinant protein/peptide production was induced with 1 mM IPTG for 3 hours at 310 K, except for all KMTs/HATs, which were produced over night at 303 K. Recombinant protein/peptide purifications were performed according to published protocols[3, 5].

Solide-phase peptide synthesis

Non-labeled histone H3-K9(Me_1) (aa1-20), H3-K9(Me_2) (aa1-20), H3-K9(Me_3) (aa1-20), histone H4-K16(Ac) (aa1-19) and histone H3-K4($^{13}C/^{15}N$) (aa1-15) were synthesized using standard Fmoc-based solid-phase chemistry as described[4]. For lysine-selectively labeled peptides, U-13C/U-15N-Fmoc-Lys(Boc)-OH was synthesized from 13C6, 15N2-L-Lysine • HCl (Silantes, Munich, Germany) as described before[4].

KMT/HAT profiling

In vitro KMT assays were performed in 40 mM deuterated Tris, 50 mM NaCl, 2 mM DTT. K_m and k_{cat} values for SET7/9 histone H3-K4($^{13}C/^{15}N$) reactions were determined at pH 8.0 and 303 K based on initial velocity measurements at enzyme concentrations of 1 μM and substrate concentrations of 12.5 μM, 50 μM, 150 μM, 300 μM, 600 μM, 800 μM and 1.2 mM (SAM 500 μM). Experimental noise was determined from root mean square averages of signal intensity integrals of 50 randomly chosen regions of spectral noise in 2D NMR experiments. Initial velocity measurements were performed until 20% of substrate turnover was reached, which enabled linear fitting. Independent duplicate experiments were performed for K_m measurements. Mono-, di- and trimethylation reactions were performed with SET7/9 (wt, 6 μM) and SET7/9 (mt, Y245A, 6 μM) mixed with H3-K4 ($^{13}C/^{15}N$, 100 μM) (400 μM SAM) at 303K and pH 8.8 for a faster reactions. Simultaneous KMT/HAT profiling was performed at pH 8.0 and 303 K with SET8 (0.5 μM) and HAT1 (0.25 μM) and uniformly $^{13}C/^{15}N$ labeled histone H4 (100 μM) (1 mM acetyl-CoA, 400 μM SAM).

NMR spectroscopy

All NMR spectra were recorded on a 750 MHz Bruker Avance spectrometer, equipped with a cryogenically cooled triple resonance $^{1}H[^{13}C/^{15}N]$ probe. All 2D experiments were processed with iNMR4.2 by zero filling (ZF) to 4K and 1K in the proton and nitrogen/carbon dimensions, respectively. Natural abundance (NA) NMR measurements of mono-, di-, tri-methylated and acetylated lysine amino acids (10 mg/mL), and of synthesized peptides H3-K9(Me_1), H3-K9(Me_2), H3-K9(Me_3), and H4-

K16(Ac) (40 mg/mL), were performed at 277 K and at pH 7.0 in 20 mM phosphate buffer, 50 mM NaCl. Natural abundance 2D $^{1}H/^{13}C$ correlation NMR spectra were recorded with 8 scans, 1024 (^{1}H) x 256 (^{13}C) complex points, and sweep widths of 16.66 ppm (^{1}H) and 60 ppm (^{13}C), using standard hetero-nuclear single quantum coherence (HSQC) NMR experiments.

2D $^{1}H/^{15}N$ HSQC correlation NMR spectra of uniformly labeled H2A ($^{13}C/^{15}N$) and lysine-selective labeled H2A ($^{13}C/^{15}N$) were recorded at 250 μM, 277 K and at pH 6.0 in 20 mM phosphate buffer, 50 mM NaCl, with 8 scans, 1024 (^{1}H) x 256 (^{15}N) complex points, and sweep widths of 16.66 ppm (^{1}H) and 24 ppm (^{15}N). 2D $^{1}H/^{15}N$ (HCε(Me_X)-TOCSY-Cα)NH correlation NMR spectra of uniformly labeled H2A ($^{13}C/^{15}N$) and lysine-selective labeled H2A-K ($^{13}C/^{15}N$) were recorded in a 3 mm diameter Shigemi tube at 150 μM, 277 K and at pH 6.0 in 40 mM deuterated Tris, 50 mM NaCl, with 2K scans, 1024 (^{1}H) x 16 (^{15}N) complex points, and sweep widths of 16.66 ppm (^{1}H) and 4 ppm (^{15}N). An interscan delay of 250 ms was used, which provided the best signal-to-noise per unit of experimental time. The total experimental time was 4 hours. Linear prediction was performed for the nitrogen dimension to 24 complex points before Fourier transformation.

NMR assignments of backbone amide resonances of uniformly-labeled H2A ($^{13}C/^{15}N$) was achieved via BEST HNCACB and CBCACONH experiments[8], at pH 6.4 and 277K and at 500 μM of H2A, with 16 scans 1024 (^{1}H) x 128 (^{13}C) x 80 (^{15}N) complex points and sweep widths of 16.66 ppm (^{1}H), 66 ppm (^{13}C) and 24 ppm (^{15}N). 3D NMR spectra were processed with Topspin 2.1 and analyzed with CccpNmr 2.0. 2D $^{1}H/^{13}C$ SOFAST-HMQC NMR spectra for KMT/HAT reactions were performed at the indicated

concentrations, pH and temperatures (see above). For large substrate concentrations (> 50 μM), NMR experiments were run with 1024 (^{1}H) x 64 (^{13}C) complex points and sweep widths of 16.66 ppm (^{1}H) and 12 ppm (^{13}C), or of 16.66 ppm (^{1}H) and 30 ppm (^{13}C). The number of scans was adjusted depending on sample concentrations. For low substrate concentrations 512 scans and 1024 (^{1}H) x 24 (^{13}C) complex points were recorded, resulting in 29 minutes of acquisition time per experiment. Interscan delays of 30 ms were used.

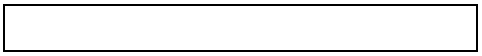
Methylation in HeLa whole-cell extracts

HeLa cells were grown in DMEM low glucose medium, harvested at ~80% confluency, washed with phosphate buffered saline, centrifuged 3 minutes at 500 g in a centrifuge equipped with swing-out buckets. The cell pellet was sonicated 4 times for 5 seconds. Cell lysis, as well as lysis of nuclear organelles was confirmed by light microscopy. The extract was then buffered with 80 mM of deuterated Tris, pH 8.5, and supplemented with 1 mM SAM, 2 mM DTT, protease inhibitors. 200 μM of uniformly ^{15}N/^{13}C-labeled H3-GB1 was added to NMR monitor endogenous KMR activities. The final sample volume was 200 μL and the reaction was followed at 310 K in a 4 mm Shigemi NMR tube. At the chosen reaction end point, the pH was adjusted to 6.2 and methylation site mapping was performed employing the (HCε(Me$_X$)-TOCSY-Cα)NH pulse sequences as outlined above (at 277 K).

Quantitative data analysis

Methylation and acetylation levels were deduced based on $^{1}H/^{13}C\varepsilon$ resonance intensities, as determined by signal integration. Normalization was performed as follows: Sums of all CH2ε peak volumes (i.e. non-modified, mono-, di- and trimethylated, and acetylated) were averaged in each NMR spectrum of the time-series. This average was normalized and used as the reference to determine the respective percentages of individually modified species in every time point. Progress curves were fitted in GraphPad Prism 5.0, based on equations solved with Mathematica and shown below.

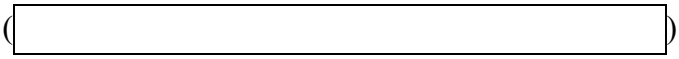
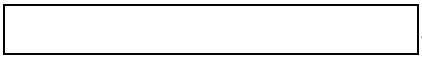
- For data shown in figure 2a:



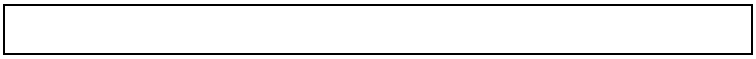
With *k* being a function of substrate concentration.

The differential equation to be solved was:

(1)

Equation (1) describes an exponential rate, whose initial slope () can be fitted by linear evolution:  .

- For data shown in figure 2b:



The differential equations to be solved were:

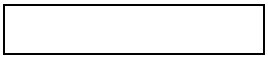
(2)

(3)

(4)

(5)

We assumed constant k_1, k_2 and k_3's at different substrate concentrations, which is a fair approximation given the associated K_m values of ~ 1 mM. We also assumed similar K_{m2} and K_{m3}'s, inferring that mono- and di-methylated lysines are equally good substrates for SET7/9 (mt). The equations used for fitting of the experimental data were:



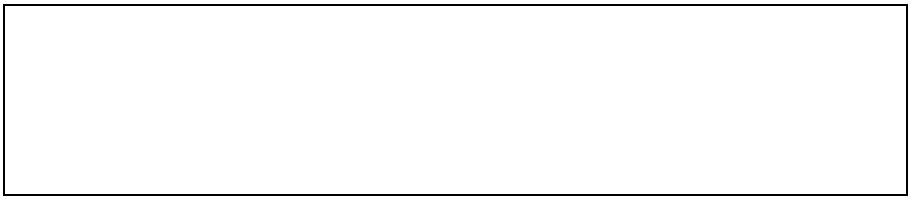
- For data shown in figure 2c:

Based on observations under isolated reaction conditions, we assumed that acetylation of K5 and K12 pursued independent of methylation of K20 and that progressive

accumulation of acetylated and methylated substrate molecules did not influence the respective other modification event. We also considered that the H4 peptide concentration was lower than the previously determined K_m of HAT1 and SET8 at their respective substrate sites. A further approximation was afforded by the assumption that HAT1 had similar K_m's for K5 and K12. Despite these approximations, 12 different rate equations were taken into account.

The differential equations to be solved were:

(6)

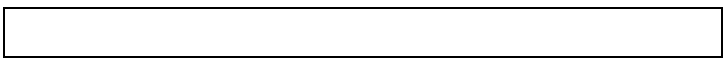


(7)

(8)

H4(Ac) representing the accumulation of H4-K5Ac, H4-K12Ac, H4-K5K12Ac, H4-K5AcK20Me$_1$, H4-K12AcK20Me$_1$ and H4-K5K12AcK20Me$_1$, and H4(Me$_1$) representing the accumulation of H4-K20Me$_1$, H4-K5AcK20Me$_1$, H4-K12AcK20Me$_1$ and H4-K5K12AcK20Me$_1$.

The equations used for fitting of the experimental data were:



Supplementary Bibliography:

(1) Emsley, L.; Bodenhausen, G. *J. Mag. Res.* **1992,** *97*, 135-148.

(2) Kadkhodaie, M.; Rivas, O.; Tan, M.; Mohebbi, A.; Shaka, A. J. *J. Mag. Res.* **1991,** *91*, 437-443.

(3) Liokatis, S.; Dose, A.; Schwarzer, D.; Selenko, P. *J. Am. Chem. Soc.* **2010,** *132*, 14704-14705.

(4) Dose, A.; Liokatis, S.; Theillet, F. X.; Selenko, P.; Schwarzer, D. *ACS Chem. Biol.* **2010,** *6*, 419-424.

(5) Couture, J. F.; Collazo, E.; Hauk, G.; Trievel, R. C. *Nat. Struct. Mol. Biol.* **2006,** *13*, 140-146.

(6) Del Rizzo, P. A.; Couture, J. F.; Dirk, L. M.; Strunk, B. S.; Roiko, M. S.; Brunzelle, J. S.; Houtz, R. L.; Trievel, R. C. *J. Biol. Chem.* **2010,** *285*, 31849-31858.

(7) Couture, J. F.; Dirk, L. M.; Brunzelle, J. S.; Houtz, R. L.; Trievel, R. C. *Proc. Natl. Acad. Sci. U S A* **2008,** *105*, 20659-20664.

(8) Lescop, E.; Schanda, P.; Brutscher, B. *J. Magn. Reson.* **2007,** *187*, 163-169.