Research Article

Journal of PeptideScience

Received: 11 January 2010 Revised: 27 April 2010 Accepted: 30 April 2010 Published online in Wiley Interscience: 22 June 2010

(www.interscience.com) DOI 10.1002/psc.1257

NMR-based detection of acetylation sites in peptides

Caroline Smet-Nocca,[a*‡] Jean-Michel Wieruszeski,[b] Oleg Melnyk[c] and Arndt Benecke[a,d]

Acetylation of histone tails as well as non-histone proteins was found to be a major component of the 'chromatin code' that regulates transcription through the recruitment of transcription factors, co-regulators and DNA-binding proteins. Acetylation can have several effects modifying protein–protein interactions, protein activity, localization and stability. Using NMR spectroscopy, we provide a simple way to detect acetyl moieties at the ε-amino function of lysine residues based on peptides derived from Histone H4 and TDG amino-terminal domains. Significant changes of acetyl-lysine resonances as compared to non-acetylated residues allow a direct identification of specific acetylated lysine. We also show that, in unfolded peptides, acetylation of lysine side chains leads to characteristic NMR signals that vary only weakly depending on the primary sequence or the total number of acetylated sites, indicating that the acetamide group does not establish any interactions with other residues. Furthermore, resonance changes upon acetylation are restricted to residues nearby the acetylation site, indicating that acetylation does not modify the overall peptide conformation. Copyright © 2010 European Peptide Society and John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article

Keywords: acetylation; histone; thymine-DNA glycosylase; protein structure; NMR spectroscopy

Introduction

Posttranslational modifications of proteins are crucial events in the regulation of protein functions or protein–protein interactions. Acetylation of lysine residues is an abundant posttranslational modification within the eukaryotic cells. Acetylation ranks in the same order as phosphorylation [1] in terms of abundance and exhibits the same features as phospho-dependent signaling pathways because addition and removal of acetyl moities are highly dynamic and are supported by several (de)acetylase enzymes [2]. A large number of non-histone acetylated proteins, the so-called acetylome [1], has been identified in eukaryotic cells. Reversible protein acetylation is involved in the regulation of signaling pathways, in the modification of DNA-binding properties or protein stability and localization, and rivals with other posttranslational modifications targeting lysine residues, such as methylation, ubiquitination and sumoylation. These effects result from neutralizing the positively charged lysine residues and generating new binding sites. Bromodomains were described as specialized domains involved in specific binding of acetylation sites and are found in a number of transcription factors or co-activators like HATs [3,4]. In malignant cells, aberrant acetylation patterns were found as well as modifications of HAT and HDAC activities [5]. The role of HATs and HDACs in development and their deregulation in human diseases like leukemia, breast and colorectal cancers or human developmental disorders emphasizes their functional prevalence. Furthermore, inhibition of HDAC activities has been found to be of promising therapeutic use in cancer treatment because HDAC inhibitors are able to modify the acetylation patterns of histone and non-histone proteins as well, implying a change in their protein–protein interactions and in their subsequent cellular localization and lifetime [6–9].

Many transcription-related factors and DNA-binding proteins were found to be acetylated. A common feature of these proteins is that they often contain unfolded regions which (i) are specifically involved in protein–protein or protein–DNA interactions, (ii) display structural changes upon substrate binding and

* *Correspondence to: Caroline Smet-Nocca, Unité de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, Université des Sciences et Technologies de Lille, Bâtiment C9, 59655 Villeneuve d'Ascq Cedex, France. E-mail: caroline.smet@univ-lille1.fr*

a *Institut de Recherche Interdisciplinaire, CNRS USR3078, Université de Lille1, Parc de la Haute Borne, 50 Avenue de Halley, 59658 Villeneuve d'Ascq Cedex, France*

b *Unité de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, Université de Lille1, 59655 Villeneuve d'Ascq Cedex, France*

c *Institut de Biologie de Lille, CNRS UMR 8161, Université de Lille 1 et 2, 1 rue Professeur Calmette, 59021 Lille Cedex, France*

d *Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440 Bures-sur-Yvette, France*

‡ *Present address:Université des Sciences et Technologies de Lille, Unité de Glycobiologie Structurale et Fonctionnelle, CNRS UMR8576, Bâtiment C9, 59655 Villeneuve d'Ascq Cedex.*

Abbreviations used: *CBP, Creb (cAMP-responsive element binding protein)-binding protein; HAT, histone acetyltransferase; HDAC, histone deacetylase; HSQC, heteronuclear single quantum coherence spectroscopy; SSP, secondary structure propensity; TDG, Thymine-DNA Glycosylase; TOCSY, total correlation spectroscopy.*

Journal of PeptideScience

(iii) are often preferential targets for posttranslational modifications [10–17]. Histones, the protein units at the basis of the nucleosome particle and chromatin architecture, are among the best studied examples of unfolded targets of the posttranslational modification machinery [18–20]. Multiple and combinatorial post-translational modifications were characterized for histone proteins and more precisely the histone amino-terminal tails [21]. Acetylation is a key modification of histones as their acetylation levels are well correlated with transcriptionally active chromatin domains [22]. Histone acetylation, as it has been well described for the histone H4-K16 site [23,24], is involved in chromatin relaxation [25–27] helping to increase DNA accessibility for transcription-related factors. Acetylation also provides novel binding interfaces for proteins targeting acetylated residues and therefore is an integral component of the chromatin code hypothesis [28]. However, a lack of structural details at the protein level hampers the establishment of structure–activity relationship upon acetylation.

Identification of acetylation sites and their binding partners is crucial for a better description of the acetylome and acetylation-dependent signaling pathways. A major issue consists in the determination of acetylation sites when multiple acetylations occur as this is the case of the histone amino-terminal tails. The identification of the acetylated proteins within cells and the determination of the acetylation sites have been made possible through the use of sensitive radiolabeled 14C- or 3H-acetyl-CoA substrates [29–31]. Determination of specific acetylated lysine residues is typically realized by Edman degradation of peptide substrates [32] or acetylation of chemically synthesized peptides *in vitro* [30], which are either destructive or indirect methods. Finally, mass spectrometry is the method of choice to detect acetylation within purified proteins isolated from cells. However, multiple acetylations render difficult the identification of acetylation sites and necessitate a tandem mass spectrometry strategy [33,34]. Quantification of acetylation levels is also a major issue that requires the isotopic labeling of acetyl moieties to ensure a normalization of signal intensities [33,34].

We describe here the impact of lysine side chain acetylation in peptide substrates using NMR spectroscopy. We have studied the structural effects of acetylation on unfolded peptides derived from two CBP/p300 substrates: the histone H4 amino-terminal tail [31,35] and a non-histone protein, the human TDG [32], which is involved in both transcriptional regulation and DNA repair processes [36–41]. They both contain at the molecular level an extended amino-terminal region that is targeted by the CBP/p300 acetyltransferase activity.

Materials and Methods

Peptide synthesis and acetylation

Peptides used in this study are those of the H4 histone *N*-terminal tail and human TDG, whose sequences are Ac-SGRGKGGK$_8$GLGKGGAK$_{16}$RHRKVLR-NH$_2$ and Ac-QPVEPKKPVESK$_{83}$K$_{84}$SGK$_{87}$SAKSKEKQ-NH$_2$, respectively (Ac and NH$_2$ indicate that the *N*- and the *C*-terminus are acetylated and amidated, respectively, and the acetylated lysine residues are numbered). Peptides were synthesized automatically by solid-phase synthesis on a continuous flow synthesizer (Perseptive Biosystems, Framingham, MA, USA) using the Fmoc group as temporary protection for the alpha-amine function of each amino acid. Peptides were assembled at a 0.1-mmol scale on a Rink amide resin (purchased from Novabiochem, Merck Biosciences AG, Läufelfingen, Switzerland), leading after the final TFA cleavage step to an amide function at the carboxy-terminus. Synthesis was carried out using ten equivalents of *N*-α-Fmoc-protected amino acids and coupling was performed with TBTU/HOBt and DIPEA in DMF. Acetylated lysine residues were introduced by coupling of the commercially available *N*-α-Fmoc-*N*-ε-acetyl-lysine synthon (Novabiochem). Fmoc cleavage was performed using 20% piperidine in DMF. Capping steps were performed after each coupling step using acetic anhydride (3%) and DIPEA (0.3%) in DMF. Amino-terminal functions of peptides were acetylated by a final acetylation step that immediately succeeds to the amino-terminal Fmoc deprotection. Peptides were released from the solid support by resin incubation with 5 ml of a TFA solution containing 2.5% triisopropylsilane and 2.5% water for 3 h at room temperature. Crude peptides were precipitated in 250 ml of a cold ether : heptane mixture (v/v) and recovered by centrifugation. Crude peptides were purified by reverse-phase chromatography in 0.05% TFA using a linear gradient of 0–40% acetonitrile in 60 min. Homogeneous fractions, as checked by analytical reverse-phase HPLC and MALDI-TOF mass spectrometry, were pooled and lyophilized.

NMR spectroscopy

NMR experiments were performed at 293 K on a Bruker DMX 600-MHz spectrometer (Bruker, Karlsruhe, Germany) equipped with a cryogenic triple-resonance probe head. For NMR experiments, peptides were dissolved at 2 mM in a buffer containing 50 mM NaH_2PO_4/Na_2HPO_4, pH 6.8, 100 mM NaCl and 5% D_2O. All ^{1}H spectra were calibrated with 1 mM sodium 3-trimethylsilyl-3,3′,2,2′-d4-propionate as a reference. For peptide assignment, standard NOESY and TOCSY experiments were recorded with 400 and 69 ms mixing times, respectively, with 2048 and 512 points and using a DIPSI2 sequence for mixing. Both are recorded with a Water suppression by GrAdient-Tailored Excitation (WATERGATE) pulse sequence for water suppression [42]. ^{1}H-^{15}N HSQC spectra were recorded at nitrogen-15 natural abundance with 64 scans per increment, with 2048 and 256 points in the proton and nitrogen dimensions, respectively, and with a window of 24.1802 ppm centered on 118.006 ppm for the nitrogen-15 dimension. ^{1}H-^{13}C HSQC spectra were recorded with 128 scans per increment, with 1024 and 512 points in the proton and carbon dimensions, respectively, and with a window of 70.4950 ppm centered on 37.362 ppm for the carbon-13 dimension. Heteronuclear experiments were recorded with a WATERGATE sequence for water suppression, a Carr–Purcell–Meiboom–Gill sequence and a double-INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) for sensitivity improvement. All experiments were acquired with a recycle delay of 1 s.

The chemical shift perturbations of individual resonances in peptides were calculated with Eqn (1) taking into account the relative dispersion of the proton and nitrogen chemical shifts (1 and 20 ppm, respectively).

$$\Delta\delta(\text{ppm}) = \sqrt{[\Delta\delta(^1\text{H})]^2 + 0.05[\Delta\delta(^{15}\text{N})]^2} \quad (1)$$

Integration of lysine NHε-acetyl signals was performed either on the proton spectrum for the methyl signals or on the ^{1}H traces extracted from the ^{15}N-^{1}H HSQC experiment for the NHε signal.

Reverse-phase chromatography and MALDI-TOF mass spectrometry analyses of peptide mixtures

Concentrated solutions of purified peptides at 10 mM in water were mixed up to equimolar ratio at final concentrations of

0.4 mM. The peptide mixture was analyzed by reverse-phase chromatography on a C2/C18 column (GE Healthcare Life Sciences, Uppsala, Sweden) equilibrated in 0.1% TFA buffer containing 2% acetonitrile; 20 nmol of each peptide was separated by a linear gradient of acetonitrile (from 2 to 40% acetonitrile in 16.67 ml, i.e. ten-column volume) at room temperature. The same peptide mixture was analyzed by MALDI-TOF mass spectrometry (Voyager DE-PRO 2000, Applied Biosystems, Foster City, California, USA) on α-cyano-4-hydroxycinnamic acid matrix.

Results

Detection of lysine side chain acetylation in peptides

Histone H4 and TDG peptides analyzed in this study were chosen among the well-characterized acetylation sites in the literature. Peptides were synthesized by Fmoc solid-phase chemistry and acetyl moieties were introduced by the use of the *N*-α-Fmoc-(L)-Lys(*N*-ε-acetyl)-OH derivative. The NMR ^{15}N-^{1}H HSQC experiments were acquired at ^{15}N natural abundance in peptides with the use of a cryoprobe head at 600 MHz. Resonance assignments were performed with the acquisition for each peptide of homonuclear TOCSY and NOESY experiments. A remarkable consequence of lysine acetylation on NMR resonances is that it converts an amine function which is in fast exchange with water leading to undetectable NMR signals into an acetamide group which adopts the same behavior as any backbone amide. In consequence, the acetamide moiety at the NHε function of lysine side chain exhibits two characteristic signals: the NHε amide at 8.01 ppm (Figures 1 and 2) and the methyl group at 1.98 ppm (Figure 1). Moreover, the relative integration of the acetamide methyl signal of lysine side chain at 1.98 ppm on one hand and those of the *N*-terminal acetyl moiety at 2.09 ppm on the other indicates the number of acetylation sites (Figure 1). Here, the integration values obtained with a recycle delay of 1 s and the number of acetylations introduced by selective NHε-acetyl-lysine incorporation during solid-phase synthesis are in good agreement and illustrates the applicability of the integration of NHε-acetyl signals to quantify the acetylation level. However, one should take care about the recycle delays in the case of unknown acetylation pattern for which a fine measurement is required. Longer recycle delays are recommended to get accurate integration values at the expense of the acquisition time. The NHε amide group of acetyl-lysine exhibits multiple scalar contacts with all side chain protons and a weak contact with Hα in the ^{1}H-^{1}H TOCSY experiment allowing the identification of the corresponding acetylated lysine residue in the absence of any NOE contacts between the NHε and other residues (Figure 2). Identification of specific acetylated lysine in unfolded peptides is hampered by the fact that ^{1}H chemical shifts differ only weakly from random coil values, hence giving a lot of resonance redundancies. However, as observed in our peptide models from amino-terminal tail of histone H4 and human TDG, a significant change in resonances comprised between +0.13 and +0.20 ppm for the Hε signals and between −0.19 and −0.22 ppm for the Hδ signals as compared to non-acetylated peptides (Tables S1 and S2, Supporting Information) facilitates the identification of acetylation sites (Figure 2). In contrast to the Hε resonances that shift in the lower field, the Hδ resonances shift in the higher field as do the Hγ signals, however, with a lower amplitude. The backbone amide signal of the acetylated residue is also significantly shifted as compared to the corresponding non-acetylated lysine falling within a range of −0.08 to −0.14 ppm (Tables S1 and S2). In contrast, the acetyl effect is less pronounced on Hα and Hβ resonances.

Additionally, shifts of peptide resonances have been mapped on a ^{1}H-^{13}C HSQC spectrum that can detect variations of ^{13}C resonances upon lysine side chain acetylation. Resonances of the modified lysines are the most affected with a marked effect for the Cδ (Table S3), whereas the remaining resonances are weakly or not perturbed (Figure 3(A)). However, given their high degree of degeneracy, ^{13}C values of either non-acetylated or acetylated lysine residues are not useful for the identification of the acetylation sites. As ^{13}C resonances are sensitive to changes in backbone structure, perturbations of ^{13}C resonances could be used to detect a propensity to adopt α-helical or β-sheet structures upon lysine acetylation. In the TDG [72–95] peptide, ^{13}Cα and ^{13}Cβ resonances are within a range of 1 ppm with respect to a purely random coil structure indicating that the peptide adopts mainly an extended structure [43,44]. We have measured, by calculating the SSP with ^{13}Cα and ^{13}Cβ values [45], an overall content of 10.7% β-structure localized within the *N*- and *C*-termini (residues 72–79 and 93–95, Figure 3(B)). Small variations of ^{13}Cα and/or ^{13}Cβ resonances of residues flanking the acetylation sites (V80, E81, S82, S85, G86, S88 and A89) are detected upon multiple lysine acetylation (Figure 3(A)) inducing only slight modifications of the overall β-structure content (Figure 3(B)).

Acetylation of lysine side chains has also an impact on the local peptide conformation and affects only the resonances of neighboring residues. We have examined, using ^{15}N-^{1}H HSQC at natural abundance combined with TOCSY/NOESY experiments, the effect of lysine acetylation (i) in the case of a single acetylation at different positions within the same peptide sequence and (ii) in the case of multiple close or distant acetylation sites.

Effect of different single acetylations on TDG peptide resonances

Multiple acetylation sites has been identified *in vitro* and found to be in close vicinity within the amino-terminal domain of TDG [32], which is involved in the regulation of its DNA repair activity [37,46]. This functional region encompassing the acetylation sites, the so-called regulatory domain, interacts in an intramolecular manner with the catalytic domain and undergoes a conformational equilibrium between this 'closed' state and an 'open' state that adopts an extended conformation with little residual structure [46]. We have synthesized peptides encompassing the three identified acetylation sites at K83, K84 and K87 positions (human TDG numbering) located within the TDG regulatory domain [32]. In the non-acetylated peptide, the detection of the sole sequential contacts between NH and Hα of residues *i* and *i* − 1, respectively, is indicative of a random coil structure. This observation is supported by ^{13}Cα and ^{13}Cβ resonances coinciding with random coil values. We have found that a single-lysine acetylation at each position leads to local perturbations spanning over the two or three neighboring residues in addition to the acetylation site (Figure 4). The resonance of the NHε acetamide moiety is not sensitive to the chemical environment in these unfolded peptides as it does not vary with the acetylated lysine residue indicating no interaction between the acetamide group and other residues.

Furthermore, no NOE contacts were detected between the NHε acetamide and other peptide resonances, indicating an absence of interactions mediated by the acetamide moiety. We have not observed a significant difference of dipolar couplings between residues upon acetylation suggesting that lysine acetylation does

Journal of PeptideScience

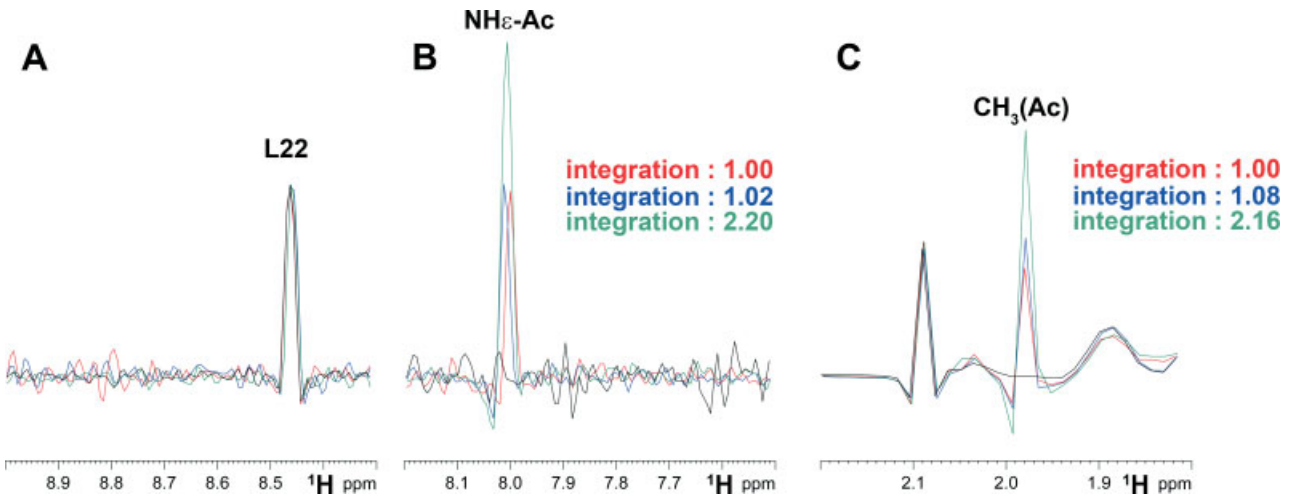

Figure 1. One-dimensional projections extracted from ^{15}N-^{1}H spectra in the ^{1}H dimension (A and B) for the H4 [1–23] peptides non-acetylated (black), mono-acetylated on K8 (red) or on K16 (blue), di-acetylated at both K8 and K16 (green) showing the row corresponding to L22 residue at 127.2531 ppm (A) as a reference and the row extracted from NHε acetamide signal at 127.3816 ppm (B). ^{1}H wet spectra of the same peptides centered on the methyl signal of the acetamide moiety (C). Relative peak intensities are given (B and C): the NH acetamide signals in (B) are normalized on the L22 signal (A) and the methyl signals at 1.98 ppm in (C) are normalized on the signal at 2.09 ppm corresponding to the methyl group of the amino-terminal acetamide moiety.

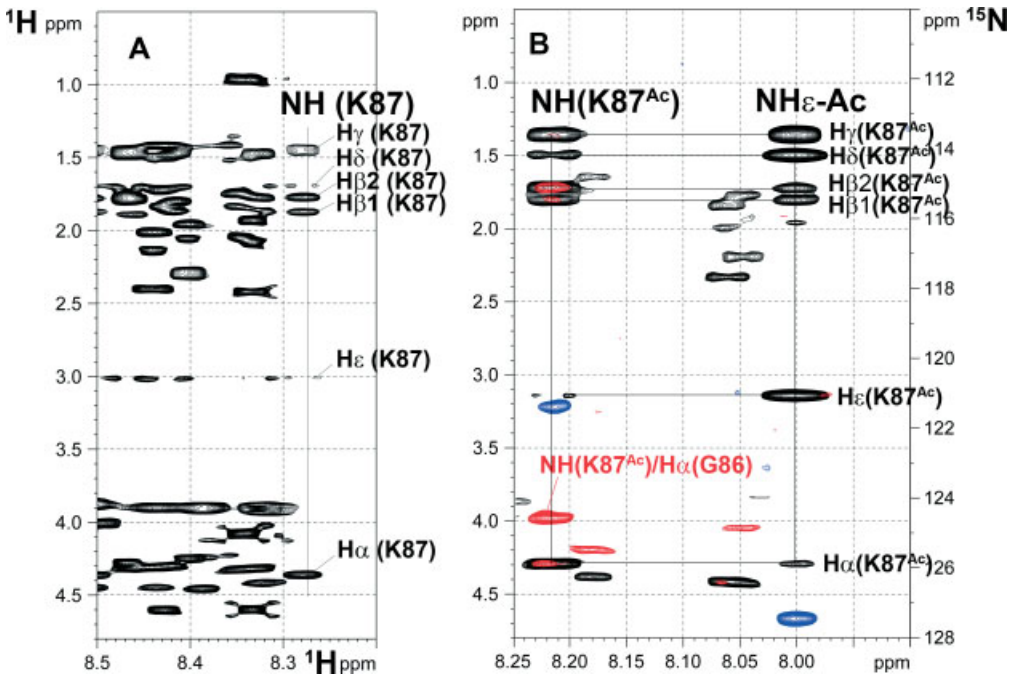

Figure 2. Scalar couplings of the K87 residue in its non-acetylated (A) or acetylated (B) forms. Detail of the ^{1}H-^{1}H TOCSY spectra (black) of the TDG [72–95] K84-Ac (A) or K87-Ac (B) peptides showing the ^{1}H-^{1}H scalar couplings between Hα, Hβ, Hγ, Hδ, Hε and the backbone NH for the K87 (A) and K87-Ac (B) residues or the NHε acetamide of the K87-Ac residue (B). ^{1}H-^{1}H NOESY spectrum is depicted in red indicating the NH/Hα dipolar coupling between K87-Ac and G86, and the ^{1}H-^{15}N HSQC spectrum is shown in blue (B). In the panel (B), the *x*-axis is common for the three TOCSY, NOESY and ^{15}N-^{1}H HSQC spectra, whereas the *y*-axis corresponds on the left to the ^{1}H dimension of the TOCSY and NOESY spectra and on the right to the ^{15}N dimension of the ^{15}N-^{1}H HSQC spectrum.

not change the overall extended structure of the peptides. The region encompassed in the TDG [72–95] peptide has a marked propensity to adopt a residual structure within the TDG protein [46]. An analysis of the primary sequence reveals a high content of charged residues that could be responsible for transient long-range interactions (hydrogen bonds, salt bridges, etc.) in the polypeptide chain. In this context, neutralization of positively charged lysine residues by acetylation could contribute to modify these non-covalent interactions. One can further notice that acetylation of K83 induces larger averaged variations of chemical shifts than acetylation of K84 or K87 suggesting more pronounced conformational changes in the TDG [72–95] peptide when K83 is acetylated (Figure 4). In its non-acetylated form, the NH_3 group of lysine is hardly detected due to fast proton exchange with water that causes severe line broadening. However, at low pH and depending on the surrounding environment, exchange rates with

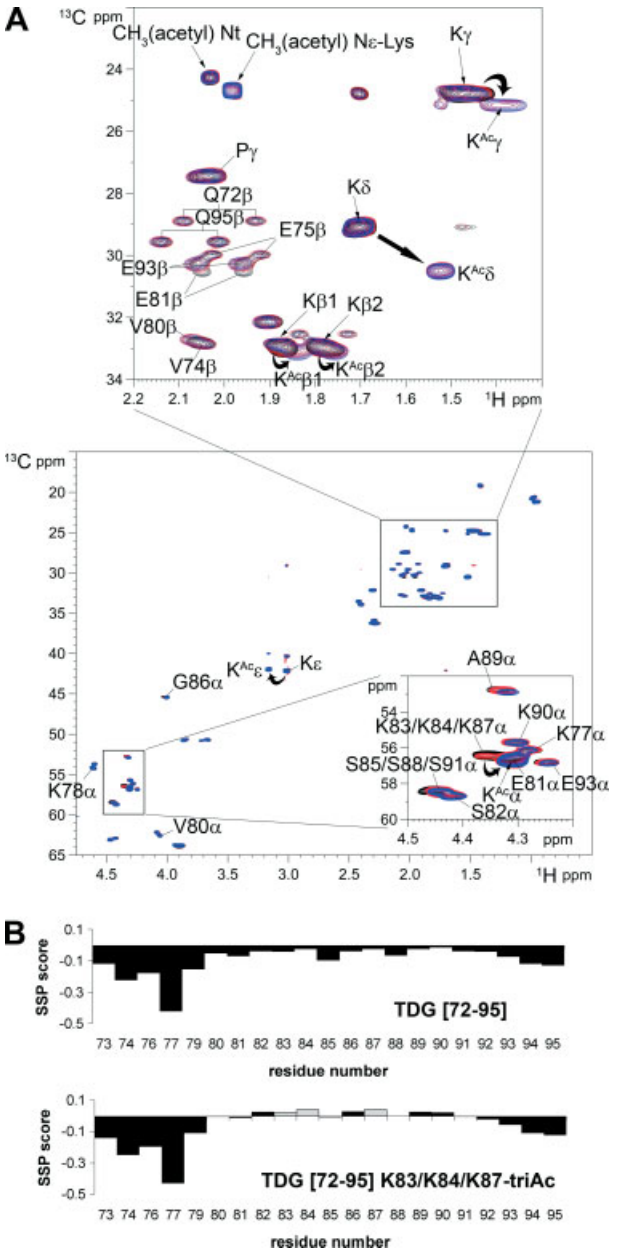

Figure 3. (A) ^{1}H-^{13}C HSQC spectra of TDG [72–95] non-acetylated (black), K83-Ac/K84-Ac di-acetylated (red) and K83-Ac/K84-Ac/K87-Ac tri-acetylated (blue) peptides (QPVEPKKPVESK$_{83}$K$_{84}$SGK$_{87}$SAKSKEKQ). Relevant resonances are annotated. Acetylated residues are indicated by an Ac superscript and the *N*-terminus by an Nt. (B) SSP score along the TDG [72–95] primary sequence for the non-acetylated (upper panel) and the K83-Ac/K84-Ac/K87-Ac tri-acetylated (lower panel) peptides. The SSP scores of acetylated lysines are indicated by grey bars.

water are slow enough to allow the observation of NH_3 resonances. In the TDG [72–95] peptide, the signals of lysine NH_3 groups were detected only at lower pH values (pH 4.01 and 2.99, Figure S1(B)). The corresponding ^{15}N resonances are found typically around 30–50 ppm [47]. Here, given the spectral width in the nitrogen dimension, the signal was four-fold aliased and observed at 129.5 ppm. Furthermore, decoupling of nitrogen was poor at this frequency, leading to the observation of a $^1J_{HN}$ coupling of 70.44 Hz (Figure S1(B)), which is significantly smaller than $^1J_{HN}$ couplings observed for backbone amide (around 90 Hz) [47]. In the TOCSY experiment acquired at pH 2.99, scalar couplings between NH_3 and Hα or side chain protons do not allow for an identification of the corresponding lysines, all of them being potentially detectable. The important degeneracy of lysine signals (except for K78 Hα at 4.585 ppm due to the presence of a proline at the position 79) hampers the assignment of the NH_3 lysine resonances (Figure S1(C)). Salt bridges between lysine and nearby glutamate residues (E75, E81 and/or E93) are quite unfavorable at this low pH and the low content of hydrophobic residues argue rather for an extended structure, so that such slow water-exchange rates can merely be explained by a pH effect. Furthermore, a pH titration has been performed to evaluate the impact of a modification of the glutamate charge on the resonances of lysine backbone amides. Because the observation of backbone amides is also related to water-exchange rates, the pH range that can be experienced is restricted to pH values lower than 7. We have acquired TOCSY and ^{15}N-^{1}H HSQC spectra on the TDG [72–95] peptide at pH 2.99, 4.01, 5.10 and 6.65. Marked perturbations of chemical shifts upon pH variations were detected for all glutamate residues with a larger shift between pH 4 and 5 (Figure S1(A)) which is in good agreement with the standard pK_R value of glutamate (around 4.25). The same transition is found to a lesser extent for surrounding residues in the primary sequence (for residues 72–82 in the *N*-terminus and 90–95 in the *C*-terminus). In contrast, in the central region of the peptide (for residues 83–89) that encompasses the three acetylation sites, a transition between pH 5.1 and 6.65 was observed together with smaller shift amplitudes (Figure S1(A)). These results indicate that a modification of the glutamate charge has little impact on the environment of backbone amide of K83, K84 or K87 residues in this unfolded peptide. Hence, the chemical shift perturbations detected upon lysine acetylation stem from the acetyl group itself rather than from disruption of long-range interactions.

Multiple acetylations: the case of distant acetylation sites in the histone H4 amino-terminus

We have evaluated the reciprocal effect of two distant acetylation sites in the histone H4 peptide (residues 1–23) located within the unfolded amino-terminal tail. The mono-acetylated peptides at K8 and K16 positions, and the di-acetylated peptide at both positions were analyzed by NMR spectroscopy (Figures 1 and 5). Single acetylation of either K8 or K16 residue leads to a minor modification of the peptide conformation restricted to the neighboring residues in the primary sequence (Figure 5) as seen for TDG peptides. Furthermore, no intramolecular interaction involving the acetamide moiety was observed because the NHε signal has nearly the same ^{1}H and ^{15}N values in every peptide (Figure 5). Remarkably, the pattern of chemical shift modifications in the di-acetylated H4 peptide corresponds to a simple additional effect of both K8 and K16 acetyl moieties indicating that remote acetylation sites have no cross-effect on peptide resonances. A detailed analysis of this latter peptide also allows the quantification of its acetylation level (i) through the comparison of the methyl signal integration of *N*-ε-acetyl-lysine with those of the amino-terminal acetamide (Figure 1(C)) or (ii) through the comparison of NHε signal integration of the acetyl-lysine residues with those of other NH signals (Figure 1(A) and (B)).

In contrast to mass spectrometry for which equimolar amounts of peptides at different acetylation states lead to unequal signal intensities (Figures S2 and S3), NMR analyses allow the relative quantification of acetylation levels at each position. For the histone

Journal of PeptideScience

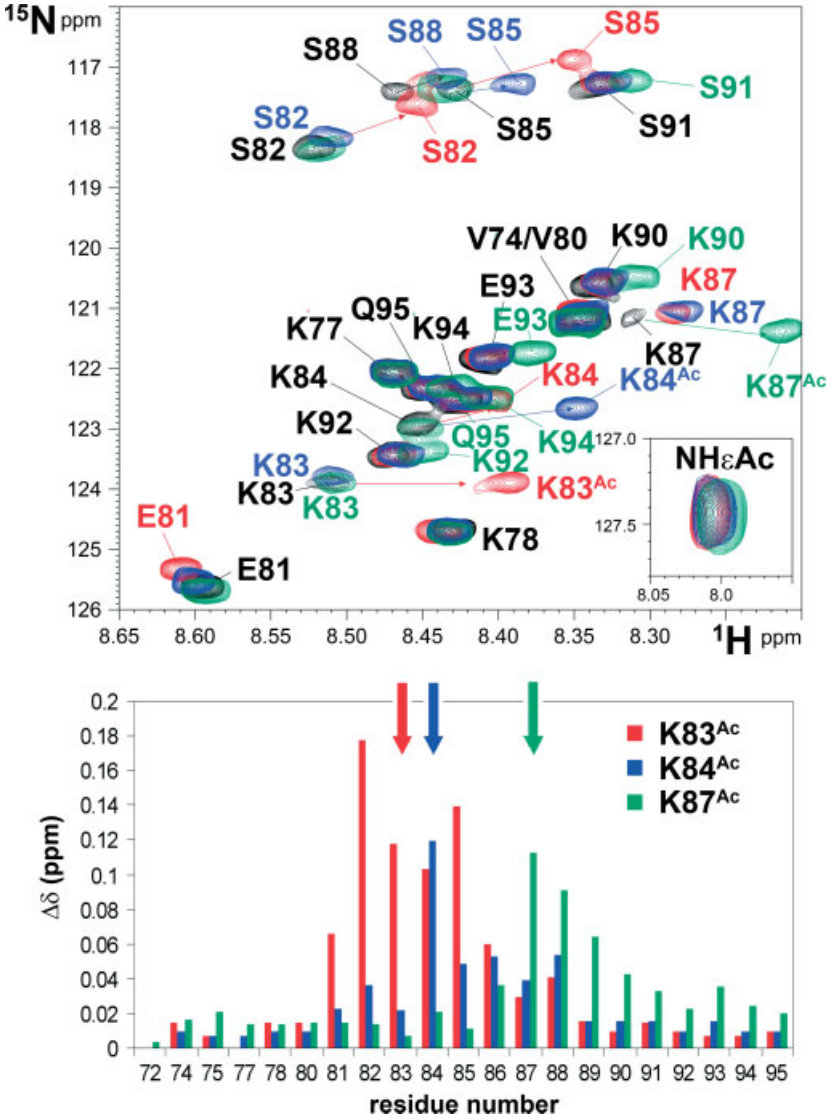

Figure 4. Effect of individual acetylations on the amide backbone resonances. ^{15}N-^{1}H HSQC spectra of TDG [72–95] peptides (QPVEPKKPVESK$_{83}$K$_{84}$SGK$_{87}$SAKSKEKQ) at natural abundance (upper panel). The non-acetylated peptide is depicted in black, the mono-acetylated peptide on K83 residue in red, on K84 residue in blue and on K87 residue in green. Residues are annotated and acetylated residues are indicated by an Ac superscript. The chemical shift variations between each of the three mono-acetylated peptides and the non-acetylated one along the peptide primary sequence using the same color code as for the NMR spectra (lower panel) are represented graphically. $\Delta\delta$ values are the averaged ^{1}H and ^{15}N chemical shift variations calculated according to Eqn (1). Acetylation sites are indicated by arrows.

H4 di-acetylated peptide, we have found identical acetylation levels at both K8 and K16 positions with the integration of NHε/Hα and NH/Hα contacts in the ^{1}H-^{1}H TOCSY experiment. We obtained for these couple of signals the same values of peak intensities indicating comparable acetylation levels at both positions (Figure 6).

Multiple acetylations: the case of close acetylation sites in the TDG amino-terminus

Again, we have used the sensitivity of the amide backbone resonance to probe modifications of the peptide conformation upon multiple nearby acetylations in the TDG amino-terminal region (residues 72–95) at positions K83, K84 and K87. We have compared the effect of either two vicinal acetylations at positions K83 and K84 or three acetylations (Figure S4). NOE contacts restricted to H$_N$/Hα between residues i and $i-1$, respectively, as well as ^{13}C resonances that fall within a range of 1 ppm from random coil values still indicate that both peptides adopt an extended conformation with a slight propensity (8.5%) to form β-sheet structures in their N- and C-terminal parts as indicated by the SSP score [45] (Figure 3(B)). This propensity ranges, however, in the same order than for the non-acetylated peptide suggesting no marked modification of the peptide structure upon acetylation. In the case of K83/K84 acetylations, a simple additional effect of lysine acetylation was observed (Figure S4(A) and (B)) and no significant shift was detected out of the 81–88 region. In contrast, an additional acetylation on K87 partially counteracts the effect of K83/K84 acetylation on K84 and S85 resonances and induces larger shifts spanning over the entire C-terminus with a more pronounced effect than the sum of K83^{Ac}/K84^{Ac} and K87^{Ac} individual contributions (Figure S4(C) and (D)). Such long-range perturbations of amide backbone resonances indicate

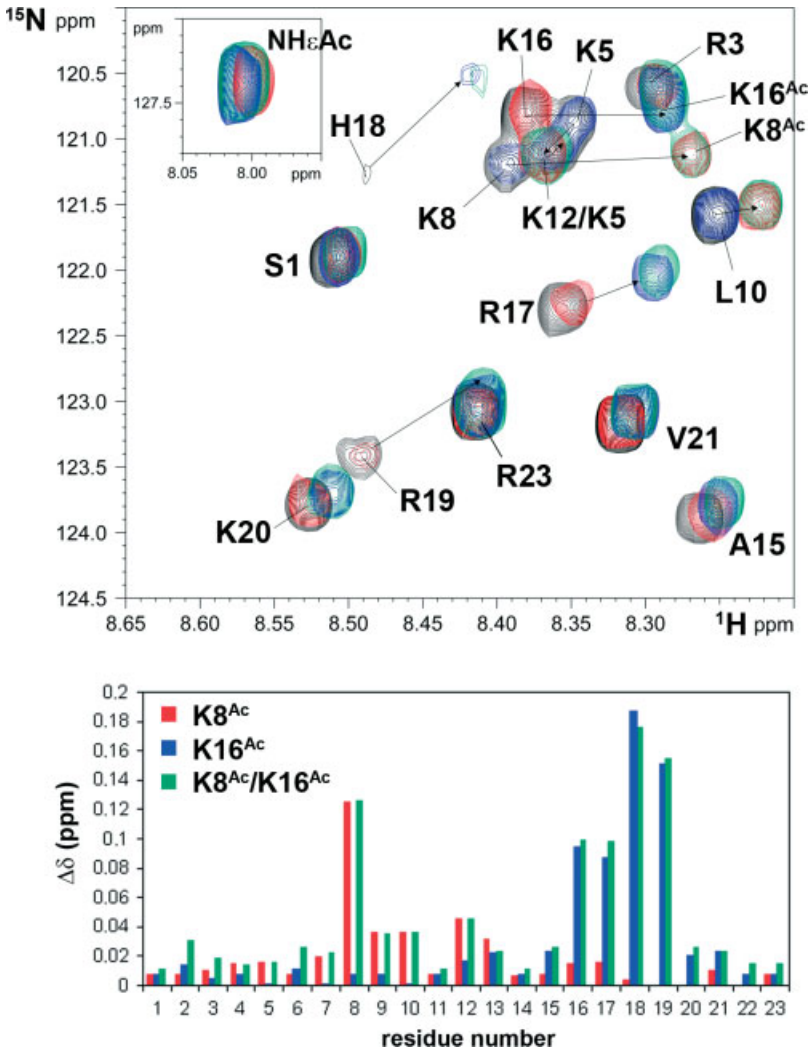

Figure 5. Effect of two remote acetylations on the amide backbone resonances. ^{15}N-^{1}H HSQC spectra of the histone H4 [1–23] peptides (SGRGKGGK$_8$GLGKGGAK$_{16}$RHRKVLR) at natural abundance (upper panel). The non-acetylated peptide is depicted in black, the mono-acetylated peptide on K8 residue in red or on K16 residue in blue, and the di-acetylated peptide at both K8 and K16 residues in green. Residues are annotated and acetylated residues are indicated by an Ac superscript. The chemical shift variations between each of the acetylated peptides and the non-acetylated one along the peptide primary sequence with the same color code as for the NMR spectra (lower panel) are represented graphically. $\Delta\delta$ values are the averaged ^{1}H and ^{15}N chemical shift variations calculated according to Eqn (1).

that K87 acetylation could have a broader effect on the peptide conformation.

Discussion

We have investigated by NMR spectroscopy the effect of lysine side chain acetylation in peptide substrates using two peptide models, one from the histone H4 and the other from a non-histone protein, the human TDG, which are both substrates of the acetyltransferase CBP/p300. We have shown that acetylation at distinct positions within both substrates has a similar impact on peptide resonances. This methodology allows a direct identification of acetylation sites based on a classical strategy of peptide assignment using TOCSY and NOESY experiments. Scalar contacts between lysine side chain protons and the characteristic NHε acetamide signal at 8.01 ppm enable the identification/quantification of acetylation sites even in the context of a multiple acetylated substrate. The acquisition of ^{15}N-^{1}H and ^{13}C-^{1}H HSQC at natural abundance has allowed for the investigation of the effect of an acetyl moiety on the ^{15}N and ^{13}C chemical shifts and on the peptide structure. These experiments can be used to detect lysine acetylation through characteristic resonances and allow for an overall quantification of acetylation levels. However, they cannot be used for the identification/quantification of each acetylation site due to a high degree of resonance degeneracy concerning the (non)acetylated lysines in the case of random coil peptides. Given the high sensitivity of $^{13}C\alpha$ and $^{13}C\beta$ to backbone structure, these chemical shifts can be used to map a perturbation of the peptide structure upon acetylation. Based on the sensitivity of backbone amide to their chemical environment, a mapping of NH variations upon acetylation of one or more lysine residues can be achieved using the ^{15}N-HSQC experiment. If used in the context of a folded protein, long-range perturbations can be detected for residues near the acetylation site. Here, we have shown in the context of unfolded peptides that acetylation of individual lysines has only a restricted effect on peptide conformation. No overall changes in the extended structure was observed but only significant shifts

Journal of PeptideScience

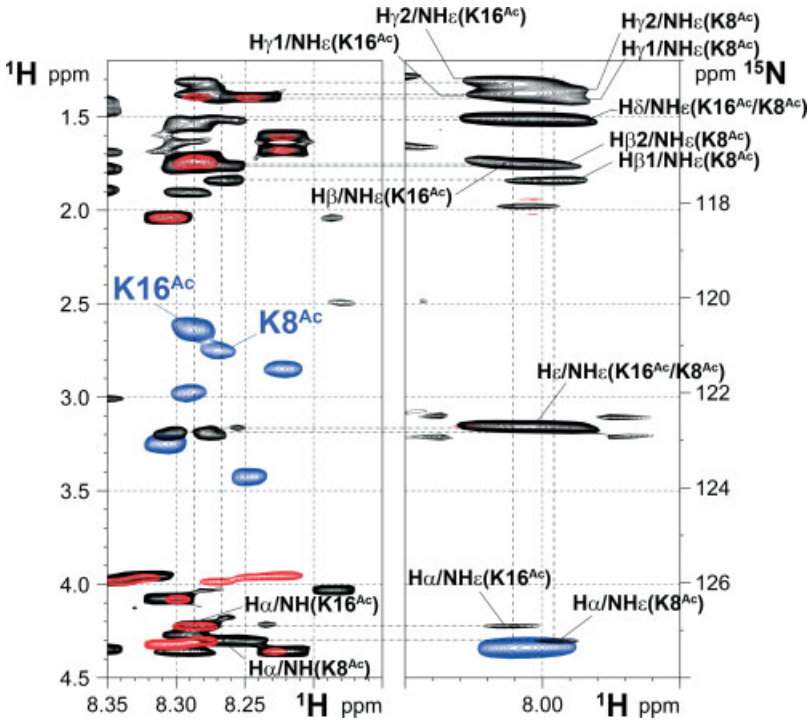

Figure 6. Detail of the ^{1}H-^{1}H TOCSY spectrum (black) of the histone H4 [1–23] di-acetylated peptide (SGRGKGGK$_8^{Ac}$GLGKGGAK$_{16}^{Ac}$RHRKVLR) showing the ^{1}H-^{1}H scalar couplings between Hα, Hβ, Hγ, Hδ, Hε and the backbone NH (left panel) or the NHε acetamide (right panel) for both the K8-Ac and K16-Ac residues. The ^{1}H-^{1}H NOESY spectrum is depicted in red and the ^{1}H-^{15}N HSQC spectrum in blue. Note that the *x*-axis is common for the three TOCSY, NOESY and ^{15}N-^{1}H HSQC spectra, whereas the *y*-axis corresponds on the left to the ^{1}H dimension of the TOCSY and NOESY spectra and on the right to the ^{15}N dimension of the ^{15}N-^{1}H HSQC spectrum.

of the backbone NH signals of the two or three residues in the immediate vicinity of the acetylation site. In these unfolded peptides, multiple acetylations being close or far from each other have also little influence on the peptide structure. No cross-effect was detected in the case of two remote acetylation sites as studied using a histone H4 K8/K16 di-acetylated peptide. The effect of two or three close acetylation sites exemplified in the case of the TDG peptide has been analyzed and it shows an additional contribution of each acetylation on the overall variations of peptide resonances with a more pronounced effect of K87 acetylation in the triply acetylated peptide. A detailed study of the latter peptide in its non-acetylated state indicates that despite a high charge density spanning over the entire sequence, no long-range interaction is observed between glutamate and lysine residues that could be perturbed by charge neutralization upon lysine acetylation. Moreover, the NHε or methyl resonances of the acetamide moiety are almost the same for all the acetylation sites, be they in the histone H4 or TDG amino-terminal domains. These features suggest that the acetyl group does not establish any interactions with other residues and always experienced the same chemical environment.

A number of works has highlighted the use of NMR spectroscopy to detect residual structures in intrinsically unfolded proteins or the presence of disordered domains in globular proteins [46,48–51]. The use of glutamine as acetylated lysine mimics has shown that a single-charge neutralization can modify the chromatin organization suggesting that acetylation of various histone tails has distinct roles in nucleosome self-assembly and in the regulation of the higher order chromatin structure [52]. Our methodology, when combined with a native peptide ligation strategy [53,54], could determine the structural impact of lysine acetylation itself in a full protein context. This strategy can take advantage from chemical peptide synthesis to generate peptide fragments of up to 50-amino acid length while selectively and quantitatively introducing one or several posttranslational modifications. The other part of the protein of interest can be synthesized as a recombinant protein with the advantage to be isotopically labeled (with ^{15}N and/or ^{13}C and/or ^{2}H) when expressed in *Escherichia coli* (for review, see Refs 55–57). The production of high amounts of natively folded proteins can be achieved at relatively low costs as compared to the introduction of isotopically labeled amino acids during solid-phase peptide synthesis. Then, the native chemical ligation provides segmental isotopically labeled proteins with a fully controlled incorporation of posttranslational modifications. Their structural effect can be investigated by NMR spectroscopy through the assignment of a sub-spectrum corresponding to the ^{15}N or ^{13}C-labeled protein fragment. In other words, for histone proteins, the structural effect of homogeneous covalent, and even combinatorial, histone tail acetylations introduced within the unlabeled, and therefore invisible, part of the protein can be studied through the spectrum analysis of the labeled part (the histone core) to evaluate their impact on histone and nucleosome structures. Isotopic labels can also be introduced at strategic points during the chemical synthesis of acetylated peptides to investigate the direct effect of lysine acetylation on the *N*-terminal tail structure and function such as in the establishment of intramolecular interactions or in the binding of protein or DNA molecules. The detection and characterization of *in vitro* acetylation after treatment of recombinant proteins or protein fragments with purified acetyltransferases or cell extracts could take advantage from the per-residue resolution of NMR spectroscopy as similarly described for phosphorylation [58,59].

Such a methodology applied to the detailed investigation of chromatin structure and function would have a significant impact on the comprehension of the chromatin code hypothesis [28], a central dogma in systems biology.

Acknowledgements

This work has been funded by the Centre National de la Recherche Scientifique (CNRS), the Région Nord-Pas de Calais, the Institut de Recherche Interdisciplinaire, the Institut des Hautes Études Scientifiques and the Agence Nationale de la Recherche. The 600-MHz facility used in this study was funded by the Région Nord-Pas de Calais (France), the CNRS, the University of Lille 1 and the Institut Pasteur de Lille. The authors thank Drs Guy Lippens and Isabelle Landrieu for insightful discussions.

Supporting information

Supporting information may be found in the online version of this article.

References

1 Spange S, Wagner T, Heinzel T, Kramer OH. Acetylation of non-histone proteins modulates cellular signalling at multiple levels. *Int. J. Biochem. Cell Biol.* 2009; **41**: 185–198.
2 Kouzarides T. Acetylation: a regulatory modification to rival phosphorylation? *EMBO J.* 2000; **19**: 1176–1179.
3 Mujtaba S, Zeng L, Zhou MM. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* 2007; **26**: 5521–5527.
4 Zeng L, Zhou MM.Bromodomain: an acetyl-lysine binding domain. *FEBS Lett.* 2002; **513**: 124–128.
5 Timmermann S, Lehrmann H, Polesskaya A, Harel-Bellan A. Histone acetylation and disease. *Cell. Mol. Life Sci.* 2001; **58**: 728–736.
6 Mottet D, Castronovo V. Histone deacetylases: target enzymes for cancer therapy. *Clin. Exp. Metastasis* 2008; **25**: 183–189.
7 Selvi RB, Kundu TK. Reversible acetylation of chromatin: implication in regulation of gene expression, disease and therapeutics. *Biotechnol. J.* 2009; **4**: 375–390.
8 Ocker M, Schneider-Stock R. Histone deacetylase inhibitors: signalling towards p21cip1/waf1. *Int. J. Biochem. Cell Biol.* 2007; **39**: 1367–1374.
9 Carew JS, Giles FJ, Nawrocki ST. Histone deacetylase inhibitors: mechanisms of cell death and promise in combination cancer therapy. *Cancer Lett.* 2008; **269**: 7–17.
10 Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 2004; **32**: 1037–1049.
11 Obradovic Z, Peng K, Vucetic S, Radivojac P, Brown CJ, Dunker AK. Predicting intrinsic disorder from amino acid sequence. *Proteins* 2003; **53**(Suppl 6): 566–572.
12 Tompa P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* 2002; **27**: 527–533.
13 Radivojac P, Obradovic Z, Smith DK, Zhu G, Vucetic S, Brown CJ, Lawson JD, Dunker AK. Protein flexibility and intrinsic disorder. *Protein Sci.* 2004; **13**: 71–80.
14 Fink AL. Natively unfolded proteins. *Curr. Opin. Struct. Biol.* 2005; **15**: 35–41.
15 Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, Uversky VN, Vidal M, Iakoucheva LM. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol.* 2006; **2**: e100.
16 Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK. Intrinsic disorder and functional proteomics. *Biophys. J.* 2007; **92**: 1439–1456.
17 Uversky VN, Radivojac P, Iakoucheva LM, Obradovic Z, Dunker AK. Prediction of intrinsic disorder and its use in functional proteomics. *Methods Mol. Biol.* 2007; **408**: 69–92.
18 Hansen JC, Lu X, Ross ED, Woody RW. Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J. Biol. Chem.* 2006; **281**: 1853–1856.
19 Bang E, Lee CH, Yoon JB, Lee DW, Lee W. Solution structures of the N-terminal domain of histone H4. *J. Pept. Res.* 2001; **58**: 389–398.
20 Cary PD, Crane-Robinson C, Bradbury EM, Dixon GH. Effect of acetylation on the binding of N-terminal peptides of histone H4 to DNA. *Eur. J. Biochem.* 1982; **127**: 137–143.
21 Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000; **403**: 41–45.
22 Eberharter A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep.* 2002; **3**: 224–229.
23 Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 2006; **311**: 844–847.
24 Shogren-Knaak M, Peterson CL. Switching on chromatin: mechanistic role of histone H4-K16 acetylation. *Cell Cycle* 2006; **5**: 1361–1365.
25 Kurdistani SK, Tavazoie S, Grunstein M. Mapping global histone acetylation patterns to gene expression. *Cell* 2004; **117**: 721–733.
26 Schubeler D, MacAlpine DM, Scalzo D, Wirbelauer C, Kooperberg C, van Leeuwen F, Gottschling DE, O'Neill LP, Turner BM, Delrow J, Bell SP, Groudine M. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* 2004; **18**: 1263–1271.
27 Ura K, Kurumizaka H, Dimitrov S, Almouzni G, Wolffe AP. Histone acetylation: influence on transcription, nucleosome mobility and positioning, and linker histone-dependent transcriptional repression. *EMBO J.* 1997; **16**: 2096–2107.
28 Benecke A. Chromatin code, local non-equilibrium dynamics, and the emergence of transcription regulatory programs. *Eur. Phys. J. E Soft matter* 2006; **19**: 353–366.
29 Lau OD, Kundu TK, Soccio RE, Ait-Si-Ali S, Khalil EM, Vassilev A, Wolffe AP, Nakatani Y, Roeder RG, Cole PA. HATs off: selective synthetic inhibitors of the histone acetyltransferases p300 and PCAF. *Mol. Cell* 2000; **5**: 589–595.
30 Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 1996; **87**: 953–959.
31 Schiltz RL, Mizzen CA, Vassilev A, Cook RG, Allis CD, Nakatani Y. Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. *J. Biol. Chem.* 1999; **274**: 1189–1192.
32 Tini M, Benecke A, Um SJ, Torchia J, Evans RM, Chambon P. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. *Mol. Cell* 2002; **9**: 265–277.
33 Smith CM. Quantification of acetylation at proximal lysine residues using isotopic labeling and tandem mass spectrometry. *Methods* 2005; **36**: 395–403.
34 Smith CM, Gafken PR, Zhang Z, Gottschling DE, Smith JB, Smith DL. Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Anal. Biochem.* 2003; **316**: 23–33.
35 Brownell JE, Allis CD. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr. Opin. Genet. Dev.* 1996; **6**: 176–184.
36 Chen D, Lucey MJ, Phoenix F, Lopez-Garcia J, Hart SM, Losson R, Buluwela L, Coombes RC, Chambon P, Schar P, Ali S. T:G mismatch-specific thymine-DNA glycosylase potentiates transcription of estrogen-regulated genes through direct interaction with estrogen receptor alpha. *J. Biol. Chem.* 2003; **278**: 38586–38592.
37 Gallinari P, Jiricny J. A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. *Nature* 1996; **383**: 735–738.
38 Hardeland U, Bentele M, Jiricny J, Schar P. Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. *J. Biol. Chem.* 2000; **275**: 33449–33456.
39 Hardeland U, Bentele M, Jiricny J, Schar P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, Drosophila and fission yeast orthologs. *Nucleic Acids Res.* 2003; **31**: 2261–2271.
40 Lucey MJ, Chen D, Lopez-Garcia J, Hart SM, Phoenix F, Al-Jehani R, Alao JP, White R, Kindle KB, Losson R, Chambon P, Parker MG, Schar P, Heery DM, Buluwela L, Ali S. T:G mismatch-specific thymine-DNA glycosylase (TDG) as a coregulator of transcription interacts with SRC1 family members through a novel tyrosine repeat motif. *Nucleic Acids Res.* 2005; **33**: 6393–6404.

41 Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, Wiebauer K, Jiricny J. Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J. Biol. Chem.* 1996; **271**: 12767–12774.
42 Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* 1992; **2**: 661–665.
43 Wishart DS, Sykes BD, Richards FM. The chemical shift index : a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* 1992; **31**: 1647–1651.
44 Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* 1995; **5**: 67–81.
45 Marsh JA, Singh VK, Jia Z, Forman-Kay JD. Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein : implications for fibrillation. *Prot. Sci.* 2006; **15**: 2795–2804.
46 Smet-Nocca C, Wieruszeski J-M, Chaar V, Leroy A, Benecke A. The Thymine-DNA Glycosylase regulatory domain: residual structure and DNA binding. *Biochemistry* 2008; **47**: 6519–6530.
47 Iwahara J, Jung YS, Clore GM. Heteronuclear NMR spectroscopy for lysine NH(3) groups in proteins: unique effect of water exchange on (15)N transverse relaxation. *J. Am. Chem. Soc.* 2007; **129**: 2971–2980.
48 Bussell R, Jr, Eliezer D. Residual structure and dynamics in Parkinson's disease-associated mutants of alpha-synuclein. *J. Biol. Chem.* 2001; **276**: 45996–46003.
49 Lippens G, Sillen A, Smet C, Wieruszeski JM, Leroy A, Buee L, Landrieu I. Studying the natively unfolded neuronal Tau protein by solution NMR spectroscopy. *Protein Pept. Lett.* 2006; **13**: 235–246.
50 Smet C, Leroy A, Sillen A, Wieruszeski JM, Landrieu I, Lippens G. Accepting its random coil nature allows a partial NMR assignment of the neuronal Tau protein. *ChemBioChem* 2004; **5**: 1639–1646.
51 Sung YH, Eliezer D. Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Mol. Biol.* 2007; **372**: 689–707.
52 Wang X, Hayes JJ. Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating the stability of higher-order chromatin structure. *Mol. Cell. Biol.* 2008; **28**: 227–236.
53 Shogren-Knaak MA, Fry CJ, Peterson CL. A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J. Biol. Chem.* 2003; **278**: 15744–15748.
54 Shogren-Knaak MA, Peterson CL. Creating designer histones by native chemical ligation. *Methods Enzymol.* 2004; **375**: 62–76.
55 Muir TW. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* 2003; **72**: 249–289.
56 Muralidharan V, Muir TW. Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat. Methods* 2006; **3**: 429–438.
57 Otomo T, Ito N, Kyogoku Y, Yamazaki T. NMR observation of selected segments in a larger protein: central-segment isotope labeling through intein-mediated ligation. *Biochemistry* 1999; **38**: 16040–16044.
58 Landrieu I, Lacosse L, Leroy A, Wieruszeski JM, Trivelli X, Sillen A, Sibille N, Schwalbe H, Saxena K, Langer T, Lippens G. NMR analysis of a Tau phosphorylation pattern. *J. Am. Chem. Soc.* 2006; **128**: 3575–3583.
59 Landrieu I, Leroy A, Smet-Nocca C, Huvent I, Amniai L, Hamdane M, Sibille N, Buée L, Wieruszeski J-M, Lippens G. Nuclear Magnetic Resonance Spectroscopy of the neuronal Tau protein: normal function and implication in Alzheimer's disease. *Biochem. Soc. Trans.* 2010; 38 (in press) .

