doi.org/10.26434/chemrxiv.12950942.v1

Site-Specific Modification and Segmental Isotope Labelling of HMGN1 Reveals Long-Range Conformational Perturbations Caused by Posttranslational Modifications

Gerhard Niederacher, Debra Urwin, David J. Tremethick, K. Johan Rosengren, Christian F. W. Becker, Anne Conibear

Submitted date: 14/09/2020 • Posted date: 15/09/2020

Citation information: Niederacher, Gerhard; Urwin, Debra; J. Tremethick, David; Rosengren, K. Johan; Becker, Christian F. W.; Conibear, Anne (2020): Site-Specific Modification and Segmental Isotope Labelling of HMGN1 Reveals Long-Range Conformational Perturbations Caused by Posttranslational Modifications. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.12950942.v1

Interactions between histones, which package DNA in eukaryotes, and nuclear proteins such as the high mobility group nucleosome-binding protein HMGN1 are important for regulating access to DNA. HMGN1 is a highly charged and intrinsically disordered protein (IDP) that is modified at several sites by posttranslational modifications (PTMs) - acetylation, phosphorylation and ADP-ribosylation. These PTMs are thought to affect cellular localisation of HMGN1 and its ability to bind nucleosomes; however, little is known about how these PTMs regulate the structure and function of HMGN1 at a molecular level. Here, we combine the chemical biology tools of protein semi-synthesis and site-specific modification to generate a series of unique HMGN1 variants bearing precise PTMs at their N- and C-termini with segmental isotope labelling for NMR spectroscopy. This study demonstrates the power of combining protein semi-synthesis for introduction of site-specific PTMs with segmental isotope labelling for structural biology, allowing us to understand the roles of PTMs with atomic precision, from both structural and functional perspectives.

File list (2)

Site-specific modification and segmental isotope labelling of HMGN1 reveals long-range conformational perturbations caused by posttranslational modifications

Gerhard Niederacher[1], Debra Urwin[2], David J. Tremethick[2], K. Johan Rosengren[3], Christian F. W. Becker[1], and Anne C. Conibear[3*]

[1] University of Vienna, Faculty of Chemistry, Institute of Biological Chemistry, Währinger Straße 38, 1090 Vienna, Austria.

[2] The Australian National University, John Curtin School of Medical Research, Department of Genome Sciences, ACT 2601, Australia.

[3] The University of Queensland, School of Biomedical Sciences, Brisbane, QLD 4072, Australia.

[*]Corresponding author:

Anne C. Conibear

The University of Queensland, School of Biomedical Sciences, Brisbane, QLD 4072, Australia.

E-mail: a.conibear@uq.edu.au

Phone: + 61-7-3365-1738

Abstract

Interactions between histones, which package DNA in eukaryotes, and nuclear proteins such as the high mobility group nucleosome-binding protein HMGN1 are important for regulating access to DNA. HMGN1 is a highly charged and intrinsically disordered protein (IDP) that is modified at several sites by posttranslational modifications (PTMs) – acetylation, phosphorylation and ADP-ribosylation. These PTMs are thought to affect cellular localisation of HMGN1 and its ability to bind nucleosomes; however, little is known about how these PTMs regulate the structure and function of HMGN1 at a molecular level. Here, we combine the chemical biology tools of protein semi-synthesis and site-specific modification to generate a series of unique HMGN1 variants bearing precise PTMs at their N- or C-termini with segmental isotope labelling for NMR spectroscopy. With access to these precisely-defined variants, we show that PTMs in both the N- and C-termini cause changes in the chemical shifts and conformational populations in regions distant from the PTM sites; up to 50-60 residues upstream of the PTM site. The PTMs investigated did not directly affect binding of HMGN1 to nucleosome core particles, suggesting that they have other regulatory roles. This study demonstrates the power of combining protein semi-synthesis for introduction of site-specific PTMs with segmental isotope labelling for structural biology, allowing us to understand the role of PTMs with atomic precision, from both structural and functional perspectives.

Keywords

Nuclear magnetic resonance spectroscopy, posttranslational modification, protein modification, intrinsically disordered proteins, protein semisynthesis, segmental labelling

Introduction

Eukaryotic cells package DNA into chromatin, a dynamic and complex structure made up of nucleosomes and non-histone proteins.[1-3] Nucleosome cores comprise ~147 base pairs of DNA wrapped around an octamer of two copies of each of the four histones H2A, H2B, H3 and H4 (Figure 1a).[4, 5] Linker DNA joins nucleosome cores into arrays and interacts with linker histone H1, forming higher order chromatin structures.[4] Structural biology techniques, as well as single-molecule and biophysical methods, have given insights into the structural dynamics of nucleosomes, showing that chromatin switches between multiple states, each with specific biological functions.[6, 7] Associated with these chromatin structures are non-histone proteins, such as the high-mobility group proteins, which modulate chromatin architecture and dynamics, thereby influencing a range of biological processes.[8, 9]

High mobility group (HMG) proteins are part of the dynamic network of architectural proteins in the nucleus. They are characterised by a high content of charged residues and are divided into three families – HMGA, HMGB and HMGN – based on their DNA-binding domains.[9, 10] The nucleosome-binding HMGN family comprises five proteins (HMGN1, 2, 3, 4 and 5), all of which have a nuclear localisation signal (NLS) split into two segments, a conserved nucleosome-binding domain (NBD) and a negatively-charged C-terminal chromatin regulatory domain (Figure 1b).[11] HMGN1 and HMGN2 are ubiquitously expressed and are thought to regulate transcription by modulating chromatin remodelling, having an important role in DNA repair, regulation of histone posttranslational modifications (PTMs) and maintenance of cell identity.[12-16] Although the precise mechanisms by which HMGN proteins modulate epigenetic processes are still not clear, several studies indicate that these nucleosomal proteins regulate cell-type-specific gene expression, and therefore have potential implications in developmental processes and disease.[11] For example, HMGN proteins have been reported to function as alarmins,[17-19] to have both pro-[20] and anti-tumour[21] activities and to have roles in DNA repair,[22, 23] immune regulation[17] and Down syndrome.[11, 24] These studies, amongst others, highlight the need to understand the molecular-level interactions of HMGN1.

HMGN proteins are intrinsically disordered and interact dynamically with chromatin.[13-15] Circular dichroism studies indicated that HMGN1 and HMGN2 binding caused changes in nucleosome structure, but that these structural changes differed for the two variants.[25] A detailed

study of the interaction between HMGN2 and nucleosomes was carried out using methyl-TROSY NMR and showed that HMGN2 interacts with the acidic patch (comprising six acidic amino acids of H2A and two acidic amino acids of H2B located on the nucleosome surface) via two conserved arginine residues (Arg22 and Arg26) in the NBD as well as with the nucleosomal DNA via its lysine-rich C-terminal region (Figure 1a).[26] Based on modelling studies, the authors proposed that phosphorylation of Ser24 and Ser28 of HMGN2 during mitosis causes electrostatic repulsion with negatively charged residues in the acidic patch, leading to dissociation of the complex.[26]

Figure 1. HMGN1 interacts with nucleosomes and bears phosphorylation, ADP ribosylation and acetylation PTMs. a) schematic structure of a nucleosome comprising dimers of each of the four histones H2A, H2B, H3 and H4 wrapped by DNA. HMGN1 (dark blue line) is an intrinsically disordered protein that binds to nucleosomes. Orientation and interaction of HMGN1 is based on that reported for HMGN2.[26] b) Protein sequence, domains and PTMs of HMGN1.[27, 28] Amino acids are represented by their one-letter codes and residue numbers are shown in grey. NLS = nuclear localization signal, NBD = nucleosome-binding domain, CHUD/RD = chromatin unfolding domain/regulatory domain. Lysine acetylation is represented by an orange square, serine phosphorylation by a purple hexagon and ADP ribosylation by a green circle.

Posttranslational modifications (PTMs) of HMGN proteins, predominantly serine phosphorylation and lysine acetylation, have been detected at several sites, as shown for HMGN1 in Figure 1b.[27-29] It is proposed that these PTMs regulate the nucleosome binding and subcellular location of HMGN proteins and thereby the accessibility of chromatin and modification of histones.[30-32] However, the PTMs and enzymes responsible appear to differ between the two most widely-studied variants, HMGN1 and HMGN2.[28, 33] The enzymes that

install PTMs on HMGN1 also modify histones. Therefore, access of these enzymes and other histone modifiers might be modulated by HMGN1 binding to the nucleosome.[29, 32] More recently, in a broad search of ADP ribosylation sites, which demonstrated that this PTM is widespread on histones, serine ADP-ribosylation was also reported for HMGN1 at residues Ser6 and Ser24.[34] Although proteomics approaches have identified PTM sites on HMGN proteins, many questions remain about effects of these PTMs, individually and in combination. Access to site-specifically modified HMGN variants will therefore be important to understand how these PTMs and combinations of PTMs influence HMGN interactions with nucleosomes and modulate chromatin dynamics. The value of access to site-specifically modified variants for elucidating chromatin dynamics and transcriptional regulation has been demonstrated in numerous studies on histones,[35-37] however such studies have not yet been widely carried out on non-histone proteins that are associated with chromatin or considered structural effects of their PTMs.

The chemical biology tools of protein semi-synthesis and site-specific modification provide unique access to homogeneous proteins bearing PTMs.[38] Such proteins are difficult to obtain from natural or recombinant sources and are crucial for understanding the structural and functional roles of PTMs. Typically, a segment of the protein is synthesized using solid phase peptide synthesis (SPPS), which allows for introduction of the desired PTM(s) in precise locations. The remaining segment(s) of the protein are produced recombinantly and ligated to the synthetic segment via native chemical ligation, or a derivative thereof.[38, 39] Recent developments and combinations with enzymatic ligations have greatly expanded the scope of protein synthesis and semi-synthesis and allow for incorporation of site-specific PTMs and combinations of PTMs, in principle, in any part of the protein.[38, 40] Furthermore, protein semi-synthesis also allows for introduction of site-specific and segmental isotope labels that are often essential for studying structural and binding properties of proteins by NMR spectroscopy.[41] NMR is a powerful technique for studying dynamic and intrinsically disordered proteins like HMGN1, which are not amenable to X-ray crystallography or cryo-electron microscopy.[26, 42] The potential of combining site-specific protein modification and segmental isotope labelling for NMR spectroscopy has not yet been widely explored but has the opportunity to provide a unique perspective on how PTMs cause structural or conformational changes to proteins and to link these to biological consequences.

Here we combine site-specific modification and segmental isotope labelling to provide unique access to HMGN1 variants, allowing us to explore the structural and functional effects of PTMs on this intrinsically disordered protein. Access to modified and unmodified HMGN1 with various isotope labelling patterns allows for the first atomic-resolution information on PTMs and combinations of PTMs, and comparison of their activities. We report the NMR chemical shift assignments of HMGN1, facilitated by segmental labelling and show how the labelled variants allow conformational perturbations to be observed. PTMs in the N- and C-terminal segments give rise to shifts in signals from domains that are distant in terms of primary sequence, some of which form part of the nucleosome binding domain. The site-specifically modified HMGN1 variants were also compared in functional assays for their nucleosome core-binding abilities. Similar binding of the HMGN1 variants to nucleosome core particles, suggests that the PTMs might have more subtle effects or regulate interactions with other partners, for example linker DNA or act as a code for other nucleosomal binding proteins.

Materials and methods

Solid phase peptide synthesis

Solid phase peptide synthesis (SPPS) of HMGN1 segments was carried out using 9-fluorenylmethoxycarbonyl (Fmoc) chemistry, either manually or on an automated synthesizer. For manual syntheses and automated syntheses on a PTI Tribute synthesizer, Fmoc-protected building blocks were coupled using 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 2.4 equiv., 0.5 M in DMF) as activator in combination with diisopropylethylamine (DIPEA, 5 equiv.) as base at room temperature for 20-30 min. N-terminal deprotection was achieved with piperidine [20% in dimethylformamide (DMF), 2 x 5 min.]. Peptides synthesized on a Liberty Blue microwave peptide synthesizer (CEM) were coupled with diisopropyl carbodiimide (DIC, 0.25 M in DMF) and OxymaPure (0.5 M in DMF), using 5 equiv. of amino acid and microwave irradiation for 4 min. at 90 °C. Deprotection was carried out using piperazine (25% in DMF). C-terminally amidated peptides were synthesized on Rink amide resin and C-terminal acid peptides were synthesized on pre-loaded Wang or Tentagel resins. Short random-coil peptides were N-terminally acetylated with acetic anhydride/DCM/DIPEA 10:85:5 (2 x 5 min.). Peptides with C-terminal hydrazides were synthesized on hydrazide-linked resin using the method of Zheng *et al.*[43] Modified residues were installed using Fmoc-Ser(PO(OBzl)OH)-OH (CEM) or Fmoc-Lys(ac)-OH (Merck) and were coupled without microwave irradiation. After drying, peptides were cleaved from the resin with trifluoroacetic acid (TFA)/dimethylsulfide(DMS)/triisopropylsilane(TIPS)/H_2O 90:5:2.5:2.5 for 2-3 h and then precipitated with diethyl ether and pelleted by centrifugation. Crude peptides were dissolved in acetonitrile (ACN)/water 1:1, lyophilized and then purified by RP-HPLC on either C4 or C18 columns using a gradient of ACN in water with 0.1% TFA. Purity was analysed by electrospray mass spectrometry in positive ion mode and by analytical RP-HPLC.

Construct design and expression of recombinant protein segments

Recombinant Expression and Purification of full-length HMGN1

A DNA construct comprising a hexa-histidine-tag followed by a tobacco etch virus (TEV) cleavage site (ENLYFQ|S) and full-length HMGN1 (His_6-TEV-HMGN1, Eurofins Genomics, full sequence in Supplementary Data S1) was cloned into a pET21a plasmid via NdeI and XhoI

restriction sites. Expression was carried out in *E. coli* BL21(DE3) using 2YT medium (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl) containing 100 µg/mL ampicillin. Overnight cultures were diluted to OD_{600} ~ 0.2, grown at 37 °C until OD_{600} ~ 0.7 and protein overexpression was induced with 1 mM isopropyl thiogalactopyranoside (IPTG). After 4 h, cells were harvested and cell pellets resuspended in TBS (50 mM Tris, 150 mM NaCl, pH 7.5) buffer and lysed by passing twice through a high-pressure cell disruptor (Constant Systems). The lysate was centrifuged and the supernatant loaded on a NiNTA column (GE HisTrap HP 5 mL) equilibrated with TBS. Protein was eluted from the column with a gradient of 0-400 mM imidazole in TBS over 60 min. Fractions containing His_6-TEV-HMGN1 were identified by SDS-PAGE, pooled and dialyzed against TBS. The His_6-tag was removed by TEV protease cleavage at 4 °C overnight with a 1:20 v/v ratio of TEV protease and 1 mM dithiothreitol (DTT). Finally, full length HMGN1 was purified by RP-HPLC on a C4 column with a gradient of 5-40% ACN in water with 0.1% TFA over 35 min at a flow rate of 3 mL/min.

Recombinant Expression and Purification of HMGN1(1-65)-thioester

A synthetic HMGN1(1-65) DNA sequence (Eurofins Genomics, full sequence in Supplementary Data S1) was inserted upstream of the *Mycobacterium xenopi* DNA gyrase A (*Mxe* GyrA) intein, His_7-tag and a chitin-binding domain (CBD) in a pTXB1 (New England Biolabs) vector via NdeI and SpeI restriction sites. Recombinant protein expression was carried out in *E. coli* BL21(DE3) Rosetta2 (Novagen) cells in 2YT medium (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl) containing 100 µg/ml ampicillin and 30 µg/ml chloramphenicol. Overnight cultures were diluted to OD ~0.2, grown at 37 °C until OD ~0.7 and protein overexpression was induced with 1 mM IPTG. After 4 h, cells were harvested by centrifugation and cell pellets were resuspended in TBS buffer and lysed by passing twice through a high-pressure cell disruptor (Constant Systems). Solubilization of inclusion bodies was achieved by treatment with 8 M guanidinium hydrochloride (Gdn.HCl) overnight at room temperature followed by NiNTA affinity purification on an Äkta Prime Plus FPLC system using a GE HisTrap HP 5 mL column equilibrated with 6 M Gdn.HCl/TBS buffer. Bound protein was eluted from the column with a gradient from 0-500 mM imidazole over 1 h at a flow rate of 1 mL/min. The fractions containing the HMGN1(1-65)-Mxe-His_7-CBD fusion protein were identified by SDS-PAGE, pooled and concentrated with centrifugal filters (Merck Amicon, 10 kDa MWCO). Buffer exchange to 8 M urea was performed using

Sephadex G-25 PD10 columns (GE Healthcare). The urea concentration of the obtained sample was reduced from 8 M to 4 M by dilution with TBS. Mxe-intein cleavage was induced by an additional 1:1 dilution with 1 M mercaptoethane sulfonate sodium salt (MesNa) in TBS (final concentration: 500 mM MesNa, 2 M urea in TBS) overnight. HMGN1(1-65)-thioester was purified by RP-HPLC on a C4 column (Kromasil) with a gradient of 5-65% ACN in water with 0.1% TFA over 30 min at a flow rate of 3 mL/min.

Recombinant Expression and Purification of HMGN1(11-99)_A11C

A pET21a plasmid containing the His_6-TEV-HMGN1(11-99)_A11C sequence (BioCat, full sequence in Supplementary Data S1) was transformed into *E. coli* BL21(DE3) Rosetta2 (Novagen). His_6-TEV represents a hexa-histidine-Tag followed by the modified recognition sequence for TEV protease (ENLYFQ|C) to obtain an N-terminal cysteine after TEV protease cleavage. Protein expression was performed in 2YT medium (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl) containing 100 µg/mL ampicillin and 30 µg/mL chloramphenicol. Overnight cultures were diluted to OD ~0.2, grown at 37 °C until OD ~0.7 and protein overexpression was induced with 1 mM IPTG. After 4 h, cells were harvested, and cell pellets were resuspended in TBS buffer and lysed by passing twice through a high-pressure cell disruptor (Constant Systems). The lysate was centrifuged and the supernatant was loaded on a NiNTA column (GE HisTrap HP 5 mL), equilibrated with 10 mM imidazole in TBS. Protein was eluted from the column with a gradient from 10-500 mM imidazole in TBS over 1 h. Fractions containing His_6-TEV-HMGN1(11-99) were identified by SDS-PAGE, pooled and dialyzed against TBS. For the removal of the purification tag and the release of the N-terminal cysteine, 1 mM tris-carboxylethyl phosphine (TCEP) and TEV protease (1:20 ratio) were added and the protein was cleaved at 4 °C overnight. HMGN1(11-99) was purified by RP-HPLC on a C4 column (Kromasil) with a gradient of 5-45% gradient of ACN in water with 0.1% TFA in 60 min at a flow rate of 3 mL/min.

Recombinant Expression of ^{15}N- and $^{13}C/^{15}N$-labeled proteins

For expression of ^{15}N-labeled proteins, cultures were grown to OD 0.7 in 2YT, pelleted by centrifugation and the cell pellet was washed with 1 L M9 buffer (3 g/L KH_2PO_4, 12.8 g/L $Na_2HPO_4*7H_2O$, 0.5 g/L NaCl) and resuspended in a quarter of the initial volume of minimal medium containing 4 g/L glucose [unlabeled or uniformly ^{13}C-labeled (U-^{13}C-6, 99%, Cambridge Isotope Laboratories CLM-1396-10)], 1 g/L $^{15}NH_4Cl$ (^{15}N, 99%, Cambridge Isotope Laboratories

NLM-467-25), 10 mL/L Basal Medium Eagle (Sigma), 2 mM $MgSO_4$, and 0.1 mM $CaCl_2$ in M9 buffer. After a 1 h regeneration period at 37 °C, overexpression was induced with 1 mM IPTG.[44] Cells were harvested, lysed, cleaved and purified as above for the equivalent unlabeled variants.

Native chemical ligation of synthetic and recombinant HMGN1 segments

Ligation of HMGN1_1-65-thioester to HMGN1_66-99_A66C variants (HMGN1_66-99_A66C, HMGN1_66-99_A66C_pS86, HMGN1_66-99_A66C_pS89, or HMGN1_66-99_A66C_pS98)

HMGN1_1-65-MesNa thioester (2 mM) and HMGN1_66-99_A66C variants (4 mM) were dissolved in degassed ligation buffer (6 M GdnHCL, 200 mM NaH_2PO_4/Na_2HPO_4, 100 mM TCEP, 100 mM methyl thioglycolate, MTG). The pH was adjusted to 7.2, the sample was incubated at 40 °C overnight and ligation progress was monitored by LC-MS and SDS-PAGE. On completion, the reaction mixture was purified by RP-HPLC on a C18 column (Kromasil) with a gradient from 5-65% ACN in water with 0.1% TFA in 30 min. Fractions containing the ligation product were identified by ESI-MS, combined and lyophilized. For desulfurization, the ligation product (1 mM) was dissolved in desulfurization buffer (6 M GdnHCL, 200 mM NaH_2PO_4/Na_2HPO_4, 200 mM TCEP, 18 mM VA-44, 18 mM *t*-BuSH, pH 6.6) and incubated for 5 h at 40 °C. The reaction was monitored by LC-MS and the desulfurized protein was purified by RP-HPLC on a C18 column (Kromasil) with a gradient from 5-45% ACN in water with 0.1% TFA over 40 min.

Ligation of HMGN1_1-10-thioester/hydrazide variants (HMGN1_1-10, HMGN1_1-10_acK2, HMGN1_1-10_pS6 or HMGN1_1-10_acK2pS6) to HMGN1_11-99_A11C.

HMGN1_1-10-hydrazide (51 mM) was dissolved in 6 M GdnHCL, 200 mM NaH_2PO_4/Na_2HPO_4, 28 mM $NaNO_2$, pH 3 and stirred at -15°C for 15 min to convert the acyl hydrazide to the corresponding azide. HMGN1_11-99_A11C (5.5 mM) was dissolved in 100 µl ligation buffer (6 M GdnHCL, 200 mM NaH_2PO_4/Na_2HPO_4, 100 mM TCEP, 200 mM MTG), mixed with the hydrazide solution and the ligation mixture was brought to room temperature and adjusted to pH 7.3. Concentrations in the ligation mixture were HMGN1_1-10-thioester 18 mM and HMGN1_11-99_A11C 3.5 mM. The reaction was shaken at room temperature overnight and monitored by LC-MS and SDS-PAGE. The ligation product was purified by RP-HPLC on a C18 column (Kromasil) with a gradient from 5-45% ACN in water over 30 min. Desulfurization was

performed by dissolving the ligation product (2 mM) in desulfurization buffer (6 M GdnHCl, 200 mM NaH_2PO_4/Na_2HPO_4, 200 mM TCEP, 18 mM VA-44, 18 mM *t*-BuSH, pH 6.6) and incubating at 40 °C for 7 h. Desulfurized product was purified by RP-HPLC on a C18 column (Kromasil) with a gradient from 5-65% ACN in water with 0.1% TFA over 30 min.

NMR data collection and processing

Samples for NMR data collection were prepared in 20 mM Na_2HPO_4 buffer with 5% v/v D_2O (Cambridge Isotope Laboratories, DLM-4-99-100), 10 μM 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) as an internal reference,[45] and 0.02% w/v NaN_3 at pH 6.0. NMR spectra were acquired at the University of Vienna NMR Centre on an Avance III HDX 700 MHz NMR spectrometer (Bruker BioSpin, Germany) equipped with an inverse helium cooled quadruple cryoprobe (QCI-F) or at the University of Queensland, Centre for Advanced Imaging on an Avance III HD 700 MHz NMR spectrometer equipped with a cryoprobe or an Avance III HD 900 MHz NMR spectrometer equipped with a cryoprobe and all spectra were acquired at 298 K. Spectra acquired for $^{15}N/^{13}C$-labeled unmodified HMGN1 included ^{1}H-^{15}N HSQC; HNCO; HNCA; HN(CA)CB; HN(CO)CA; HN(CO)(CA)CB; HN(CA)CO; HN(CO)(CA)(N)NH; HN(CA)(N)NH. Spectra acquired for ^{15}N-segmentally labeled modified/unmodified HMGN1 segments included ^{1}H-^{15}N HSQC and ^{1}H-^{15}N heteronuclear NOE experiments. Spectra acquired for unlabeled HMGN1 segments and random coil peptides included ^{1}H-^{15}N HSQC; ^{1}H-^{1}H TOCSY with an isotropic mixing time of 80 ms; ^{1}H-^{1}H NOESY with a mixing time of 300 ms; ^{1}H-^{15}N HSQC; and ^{1}H-^{13}C HSQC. Spectra were Fourier transformed, phased and calibrated on the DSS signal (^{1}H at 0 ppm) in Topspin 4.0.6 (Bruker BioSpin, Germany). ^{15}N and ^{13}C spectra were calibrated on the unified scale according to the IUPAC recommendations,[45, 46] using a ratio of 0.251449530 for ^{13}C and 0.101329118 for ^{15}N. Spectra were assigned in CCP-NMR v2 and v3.[47] Secondary chemical shifts were calculated by subtraction of the respective random coil chemical shift from the observed chemical shift.[48, 49] Chemical shift perturbation values for NH chemical shifts were calculated using the formula:[50-52]

$$\Delta\delta\ (\text{ppm}) = \sqrt{\frac{(wN \times \Delta\delta N)^2 + (wH \times \Delta\delta H)^2}{2}}$$

Where the weighting values are $w_N = 0.158$ and $w_H = 1$, and $\Delta\delta N$ and $\Delta\delta H$ indicate the difference in chemical shift (ppm) between full-length recombinant HMGN1 and the various semi-synthetic and segmentally-labeled HMGN1 variants.

Steady-state ^{1}H-^{15}N nuclear Overhauser effect (NOE) data were acquired for the uniformly- and segmentally-labelled HMGN1 variants at 700 and 900 MHz.[53] ^{1}H-^{15}N NOE ratios (I_{NOE}/I_{st}) were calculated from duplicate pairs of ^{1}H-^{15}N spectra acquired in an interleaved fashion with and without proton saturation. Phasing and scaling parameters were equal for both spectra and peak heights were calculated in CCPNMR v3.

Circular dichroism

Circular dichroism spectra were recorded on a Chirascan Plus CD spectrometer (Applied Photonics, UK) in micro-cuvettes with a 1 mm pathlength at 25 °C. Scans were carried out from 260 to 185 mm with five repetitions. Spectra were examined individually and then averaged and a baseline (H_2O) spectrum was subtracted. The HMGN1 sample was prepared in H_2O at a concentration of 5 μM at pH 5.0.

***In vitro* nucleosome core assembly and electrophoretic mobility shift assays (EMSAs)**

Nucleosome cores were assembled onto a 147 bp AlexaFluor488 labelled DNA fragment containing the 601 nucleosome positioning sequence.[54] The AlexaFluor488-labeled fragment was generated by PCR and purified as described previously.[55] Recombinant *Xenopus laevis* histone octamers were used to assemble nucleosomes, and were produced using standard protocols.[56]

Assembly of nucleosome cores was performed by salt-gradient dialysis as described.[56, 57] HMGN1 variants (2-800 nM) were incubated with AlexaFluor488-labeled nucleosome cores (50 nM) or DNA (50 nM) in gel shift buffer (10 mM Tris pH 7.5, 25 mM NaCl, 25 mM KCl, 0.2 mM DTT, 4% sucrose) on ice for 20-30 min. Binding reactions were loaded onto 5% (w/v) native PAGE gels and electrophoresed in 0.5X TBE at 75 V for 45-90 min at 4 °C. Gels were visualized on a Typhoon™ FLA9000 laser scanner (GE Healthcare).

Results and Discussion

Characterisation and assignment of unmodified HMGN1

HMGN1_1-99, preceded by a hexa-histidine tag and TEV-protease cleavage site was expressed in *E. coli* and purified on a NiNTA (affinity) column from the soluble fraction after cell lysis. Cleavage by TEV protease removed the hexa-histidine tag (Figure 2a), leaving an N-terminal serine preceding the native HMGN1 sequence (Sequence in Table S2). This serine residue is part of the recognition sequence (ENLYFQ|S/G) for TEV protease that remains after cleavage and is designated as Ser0 in this work to preserve consistency of residue numbering with HMGN1 sequences in the literature and UniProt databank. The cleaved protein was then purified by HPLC, giving a final yield of ~ 1.1 mg/L of culture. The desired molecular weight of 10.6 kDa was confirmed by mass spectrometry, however HMGN1 runs abnormally on SDS-PAGE, showing an apparent molecular weight of ~ 22 kDa (Figure 2b, d), likely due to its high content of charged residues and disordered nature.[10, 58] The lack of characteristic alpha-helix or beta-sheet maxima and minima in the CD spectrum (Figure 2c) also gives evidence for the absence of secondary structural features.

Uniform ^{15}N- and ^{15}N/^{13}C-labeling of HMGN1 for NMR spectroscopy was achieved with ~93% labelling efficiency using a two-step expression protocol in which the *E. coli* cell mass was grown up in rich medium, the cells were pelleted by centrifugation and then resuspended in minimal medium containing isotope labelled nutrients for expression. As shown in Figure 2b, expression yield and purity was reduced compared to rich medium, but sufficient quantities of uniformly ^{15}N/^{13}C-labeled HMGN1 (~0.8 mg/L of rich medium, 250 mL minimal medium) were obtained to acquire NMR data that enabled assignment of the backbone resonances (Supplementary Data Table S8). In agreement with the CD data, the limited dispersion of the amide proton chemical shifts confirms the lack of secondary structure; all amide proton shifts lie between 8.0 and 8.6 ppm (Figure 2e). Three-dimensional HNCO, HNCA, HN(CA)CB, HN(CO)CA, HN(CO)(CA)CB, HN(CA)CO, HN(CO)(CA)(N)NH, and HN(CA)(N)NH spectra were acquired on a 700 MHz NMR spectrometer equipped with a cryoprobe and showed sharp signals typical of intrinsically disordered proteins. Some ambiguities caused by overlap of amide resonances were resolved by comparison with spectra of semi-synthetic segmentally-labelled HMGN1 variants generated as described below, illustrating advantages of segmental labelling

techniques. Secondary Cα chemical shifts are shown in Figure 2f and, except for the C-terminal aspartic acid residue, are all < 1 ppm, in agreement with an absence of secondary structural features. Some stretches of positive secondary chemical shifts (e.g. residues 14-23 and residues 41-71), however might indicate an alpha-helical propensity. The overall dispersion of the signals and magnitudes of the secondary chemical shifts of HMGN1 are similar to those of HMGN2,[26] however the latter were assigned for HMGN2 bound to nucleosomes, and show mostly negative secondary Cα chemical shifts.

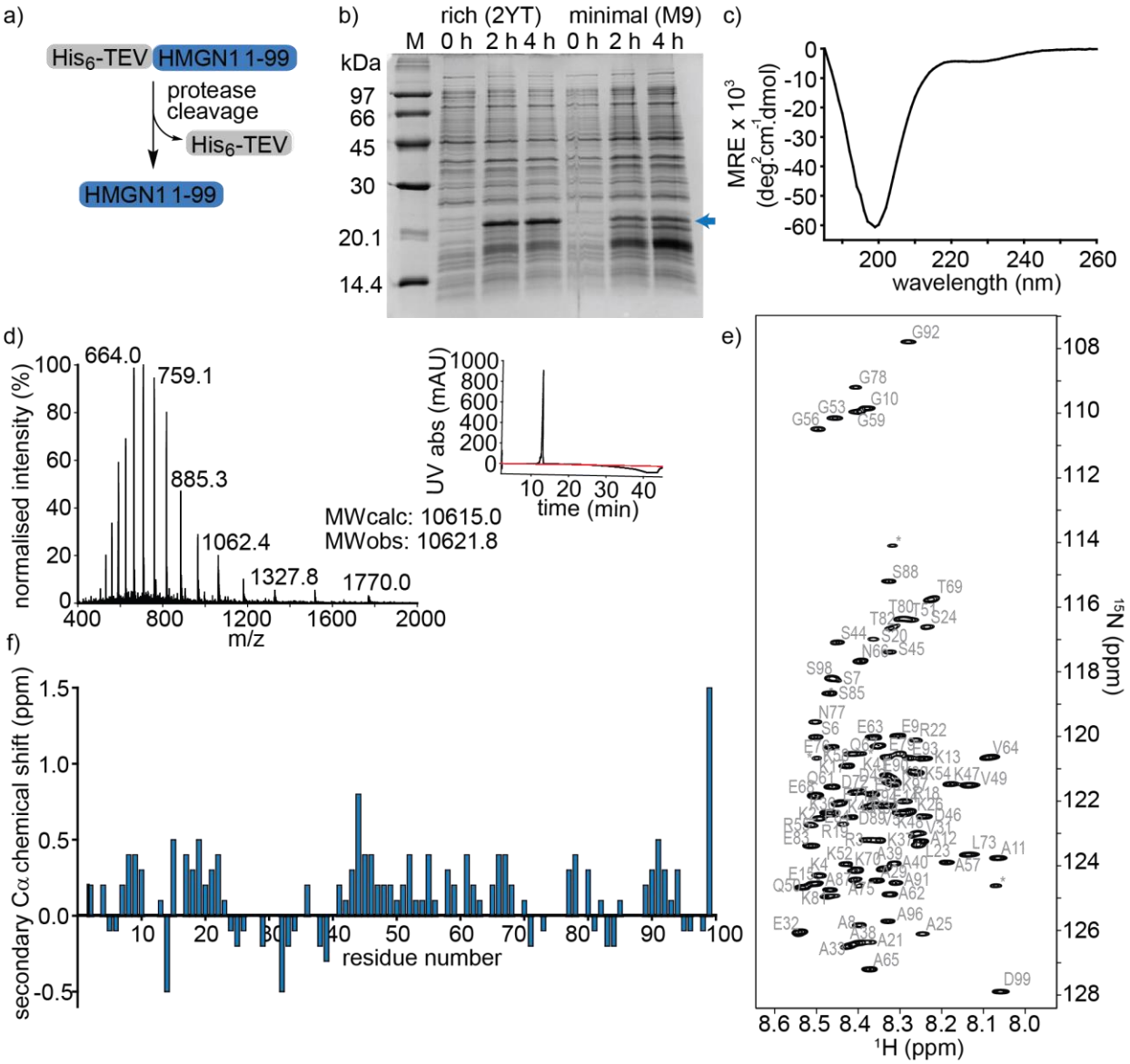


Figure 2. Expression and characterization of full-length unmodified HMGN1. a) HMGN1 fused at the N-terminus to a His_6-tag and TEV protease cleavage site was expressed and purified from *E. coli*. Cleavage with TEV protease and RP-HPLC purification yielded full-length HMGN1. b) SDS-PAGE gel showing expression profile (hours after induction) of His_6-TEV-HMGN1 construct in rich medium (2YT) and minimal (M9) medium used for $^{15}N/^{13}C$ isotope labelling. The blue arrow indicates the position of the His_6-TEV-HMGN1 protein, which runs at an apparent MW of ~22 kDa (Calculated MW of His_6-TEV-HMGN1 construct: 12.4 kDa). c) Circular dichroism (CD) spectrum of purified HMGN1. d) ESI Mass spectrum of purified HMGN1 with analytical HPLC trace (inset) at 214 nm (black) and 280 nm (red). e) ^{1}H-^{15}N HSQC NMR spectrum of uniformly labelled, unmodified HMGN1 with residue assignments. f) Secondary Cα chemical shift plot of unmodified HMGN1.

N-terminal PTMs of HMGN1: semi-synthesis, segmental labelling and NMR spectroscopy

To access HMGN1 variants bearing the N-terminal PTMs acetylation at Lys2 (acK2) and/or phosphorylation at Ser6 (pS6), we designed a protein semi-synthesis strategy involving ligation of a synthetic N-terminal segment (residues 1-10) bearing the desired PTMs to a recombinantly expressed C-terminal segment (residues 11-99), which could be expressed in labelled or unlabelled form (Figure 3a). In the synthetic segment, we can incorporate authentic phosphoserine and acetyl lysine residues using solid phase peptide synthesis (SPPS), in contrast to the alanine, glutamic acid or glutamine mutations widely used for recombinant proteins. As HMGN1 contains no native cysteine residues, we chose a ligation junction for native chemical ligation between residues 10 and 11 to install the N-terminal PTMs. In this strategy, Gly10 at the C-terminus of the synthetic fragment serves as a sterically unhindered acyl donor and Ala11 is mutated to cysteine for ligation. After ligation, the cysteine can be desulfurized using established methods to restore the native HMGN1 sequence.[59] Synthetic HMGN1_1-10 segments were synthesized by SPPS on a hydrazide resin to yield acyl hydrazides as thioester precursors,[43] and posttranslationally modified residues were installed using commercially-available protected building blocks. Cleavage, deprotection and HPLC purification yielded the peptide hydrazides HMGN1_1-10, HMGN1_1-10_acK2, HMGN1_1-10_pS6 and HMGN1_1-10_acK2,pS6 (see Supplementary Data S4 for chemical structures, yields, mass spectra and analytical HPLC data). The C-terminal recombinant segment was expressed in *E. coli* with the A11C mutation, preceded by a His_6 tag and TEV-protease cleavage site. After purification, TEV-protease cleavage proceeded to ~80% completion and purification by HPLC yielded the N-terminal cysteine-

bearing protein segment (Figure 3a, yield ~ 1.2 mg/L of rich medium, 1.5 L minimal medium). For ligation of the two HMGN1 segments, *in situ* conversion of the acyl hydrazides to acyl thioesters was achieved by oxidation of the hydrazide to an azide and thiolysis with methyl thioglycolate (MTG).[43] On addition of the HMGN1_11-99_A11C segment and pH adjustment, the ligation proceeded to completion within 5 h at 40 °C. Although we had selected the alkyl thiol additive MTG for ligation, envisaging a one-pot desulfurization strategy,[38, 60] we found the desulfurization did not go to completion and the ligation product degraded over time. However, HPLC purification of the ligation product prior to desulfurization gave cleaner and more complete semi-synthetic HMGN1 variants with native sequences and the desired PTMs (see Supplementary Data S2 and S3 for sequences, yields, mass spectra and analytical HPLC data). Ligation of uniformly ^{15}N-labelled HMGN1_11-99 segments yielded the corresponding site-specifically modified and segmentally labelled HMGN1 variants.

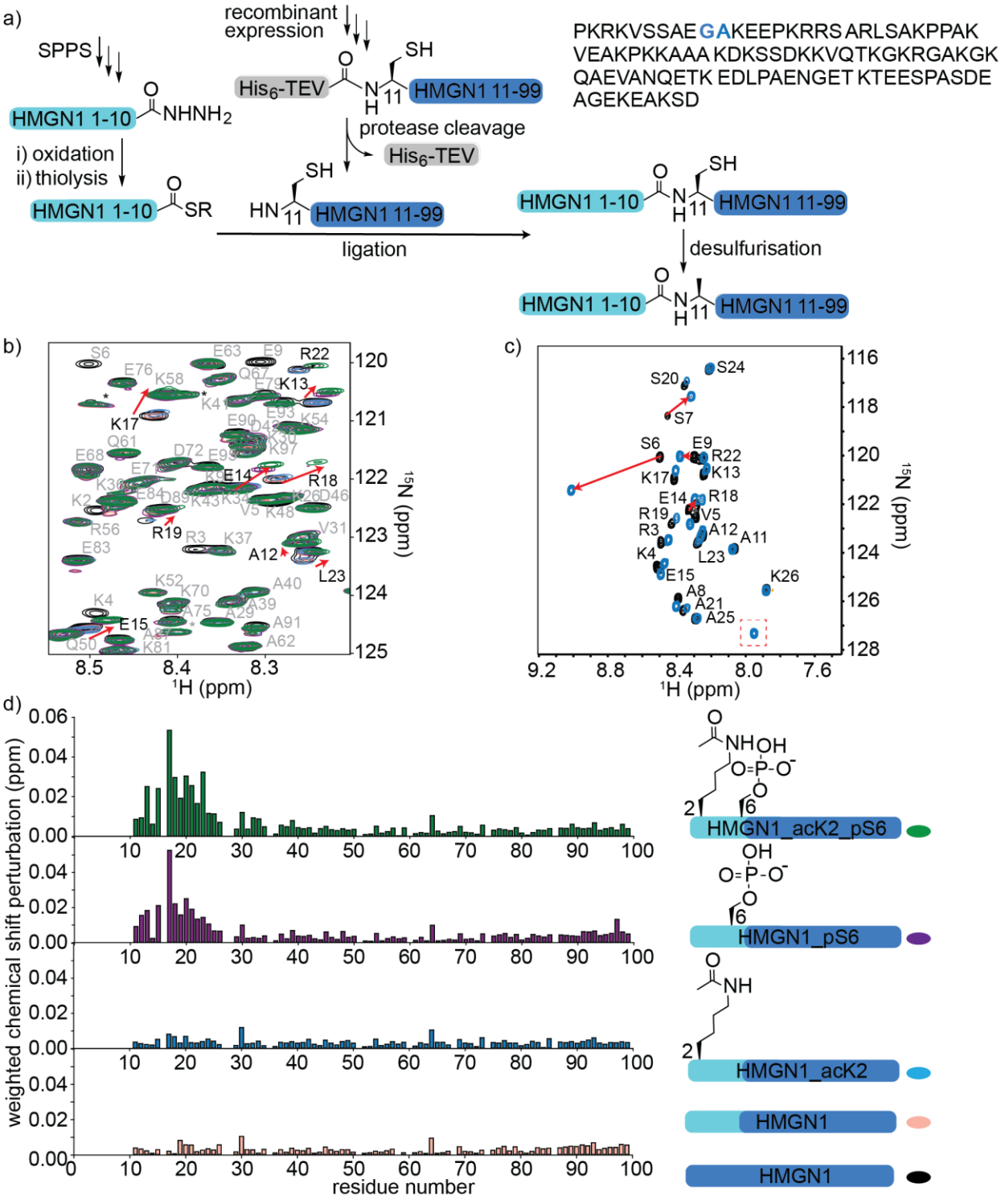


Figure 3. Semi-synthesis and NMR of HMGN1 bearing N-terminal PTMs. a) Semi-synthesis strategy involving synthesis of the N-terminal segment HMGN1_1-10 (and variants bearing PTMs) by solid phase peptide synthesis (SPPS) and recombinant expression of the C-terminal segment HMGN1_11-99 (with or without uniform ^{15}N isotope labeling). Native chemical

ligation followed by desulfurization of cysteine to alanine yielded site-specifically modified and segmentally labelled variants of HMGN1. The sequence of HMGN1 is shown with the selected ligation junction in blue. b) Section of ^{15}N-HSQC NMR spectrum overlay of segmentally labelled HMGN1 variants showing resonances in the labelled section that shift when N-terminal PTMs are present (red arrows and residue labels in black). Spectra are shown for HMGN1_15N13C (black), HMGN1_unmodN_15N (pink), HMGN1_acK2_15N (blue), HMGN1_pS6_15N (purple) and HMGN1_acK2,pS6 (green), corresponding to the schematic structures in (d). Full spectra are shown in Supplementary Data S6. c) Section of ^{15}N-HSQC NMR spectrum overlay of synthetic HMGN1_1-26 (black) and HMGN1_1-26_acK2,pS6 (blue) HMGN1 segments, acquired at natural abundance. Resonances that shift on introduction of PTMs are shown by red arrows and the new amide resonance from acK2 is shown in a red box. Full spectra are shown in Supplementary Data S6. d) Chemical shift perturbation plots for the segmentally-labelled HMGN1 variants bearing N-terminal PTMs, as shown in (b). Light blue segments are synthetic and unlabelled, dark blue segments are ^{15}N isotope-labelled.

Comparison of ^{15}N-HSQC spectra of segmentally labelled unmodified HMGN1 with unmodified uniformly labelled HMGN1 (Figure 3b and S6) shows that they have identical NMR 'fingerprints' and that the ligation-desulfurization protocol does not alter the structural properties of the protein. As the synthetic segment is unlabelled, backbone amide resonances of residues 1-10 are absent in the ^{15}N-HSQC spectrum of segmentally-labelled HMGN1, which was helpful for confirming assignments of these residues. ^{15}N-HSQC spectra of HMGN1 variants bearing PTMs in the synthetic, unlabelled segment show small shifts in several amide resonances within the labelled segment (Figure 3b). This is significant because it shows that even small PTMs, such as addition of a phosphate or acetyl group to a side chain, can cause long-range perturbations in the chemical environment of backbone residues that are distant from the modification site, even in a disordered protein. As shown in the chemical shift perturbation plots in Figure 3d, acK2 alone did not result in any shifts of amide resonances in the labelled segment, however, the HMGN1 variant bearing pS6 displayed shifts in residues 14-24, which falls in the nucleosome-binding domain. These shifts are larger in the HMGN1 variant bearing both acK2 and pS6 PTMs. Although HMGN2 (89 residues) is shorter than HMGN1 (99 residues) and differs in other domains, the nucleosome-binding domains (NBDs) of both proteins are conserved. In a study of HMGN2 binding to nucleosomes, Kato *et al.* proposed that residues 22-28 bind to the acidic patch of nucleosomes and that phosphorylation of Ser24 and Ser28 in HMGN2 (mimicked by mutation to glutamic acid) disrupts electrostatic interactions with the acidic patch.[26] Although

we did not investigate the effects of PTMs in the NBD in this study, our results suggest that in addition to potential direct electrostatic interactions, PTMs in the N-terminus might also influence the conformation and binding of the NBD. This model is consistent with that proposed by Lim *et al.*, who hypothesized that HMGN1 hinders access of mitogen-activated kinases MSK1 and MSK2 to histone H3.[61] Phosphorylation of HMGN1 at Ser6 by these kinases led to reduced binding to nucleosomes, thereby allowing access of MSK1 and MSK2 to histone H3 and phosphorylation of H3 Ser10.[61]

To verify these long-range perturbations and explore any other shorter-range changes, HMGN1_1-26 segments, either unmodified, or bearing both acK2 and pS6 PTMs, were synthesized by SPPS (Supplementary Data S3). These segments are unlabelled because they are obtained by SPPS, but homonuclear TOCSY and NOESY spectra and ^{15}N-HSQC spectra were acquired at natural abundance and were sufficient for backbone assignment for this 26-residue segment. (NMR spectroscopy of full-length HMGN1 (99 residues) at natural ^{15}N/^{13}C abundance would be impractical because of significant signal overlap in the homonuclear spectra and higher concentrations required to compensate for low sensitivity at natural abundance). Hence, segmental labelling strategies as described above were used for the full-length HMGN1 variants. The spectra shown in Figure 3c and S6 show characteristic chemical shift changes of the modified residues as predicted by their random coil shifts,[49] namely a downfield shift (0.51 ppm) of the S6 amide proton chemical shift upon phosphorylation and appearance of a new amide resonance at (7.95 pm/127.2 ppm) from the amide moiety in acetyllysine (Figure 3c), in agreement with literature values.[49] In addition to shifts in the resonances of residues flanking the PTMs, shifts were also seen in the more distant residues 12-20, corresponding to those seen for full-length segmentally-labeled HMGN1_acK2_pS6 (Figure 3b). Although the shifts in residues 12-20 on modification of Lys2 and Ser6 suggest an interaction between these two regions, no new NOESY peaks were observed that would indicate an overall structural change or formation of new secondary structure elements. However, weak transient interactions are not expected to give rise to ^{1}H-^{1}H NOEs between disordered protein regions.

To investigate the possibility of restricted motion caused by interactions of PTMs with regions of HMGN1 that are distant in primary sequence, we measured heteronuclear ^{1}H-^{15}N nuclear Overauser effect (NOE) ratios at 700 MHz and 900 MHz for unmodified HMGN1 and the

HMGN1 variant bearing both acK2 and pS6 PTMs (Supplementary Data S7a).[53, 62] As expected NOE ratios increase with increasing magnetic field strength and the majority of the NOE ratios were close to zero, reflecting the intrinsic disorder of HMGN1. NOE ratios for the N- and C-termini of HMGN1 are more negative than those of the central regions, corresponding to greater flexibility of the termini, particularly residues 1-30 and 90-99. The PTMs acK2 and pS6, however, do not appear to alter the flexibility of HMGN1, which would be indicated by an increase or decrease in the NOE ratio for a region of the protein. Similarly, comparison of the NOE ratios of the other modified variants HMGN1_pS6, HMGN1_pS88, and HMGN1_pS85,88,98 did not reveal any differences in motion compared to unmodified HMGN1 (Supplementary Data S7b and S7c).

C-terminal PTMs of HMGN1: semi-synthesis, segmental labelling and NMR spectroscopy

HMGN1 variants bearing the C-terminal PTMs pS85, pS88, pS98 were accessed via an expressed protein ligation (EPL) semi-synthesis strategy, as shown in Figure 4a.[63] The N-terminal segment HMGN1_1-64 was expressed in *E. coli* fused to the *Mxe* GyrA intein, followed by a His_7 tag and chitin binding domain (Sequence in Supplementary Data S1). The fusion protein was expressed in inclusion bodies and was resolubilized and purified under denaturing conditions. Buffer exchange to urea and dilution of the denaturant in the presence of the thiol sodium mercaptoethane sulfonate (MesNa) led to folding and cleavage of the intein to yield the N-terminal segment of HMGN1 bearing a C-terminal thioester, which was purified by HPLC from the cleaved intein and a small amount of uncleaved fusion protein (Supplementary Data S5). Three variants of the C-terminal segment HMGN1_65-99 were synthesized by SPPS with residue 65 as a cysteine in place of the native alanine (sequences, yields, mass spectra and analytical HPLC are shown in the Supplementary Data S2 and S3). This mutation allows for native chemical ligation and subsequent desulfurization to yield the native sequence, as for the N-terminally modified variants. Despite the choice of β-branched valine as the acyl donor residue in our strategy, ligation proceeded to completion within 5 h at 40 °C and subsequent desulfurization yielded three HMGN1 variants: unmodified, pS88 and pS85,88,98 (Supplementary Data S3). Corresponding segmentally labelled variants were also prepared by expression of the N-terminal intein fusion segment in ^{15}N-labeled medium.

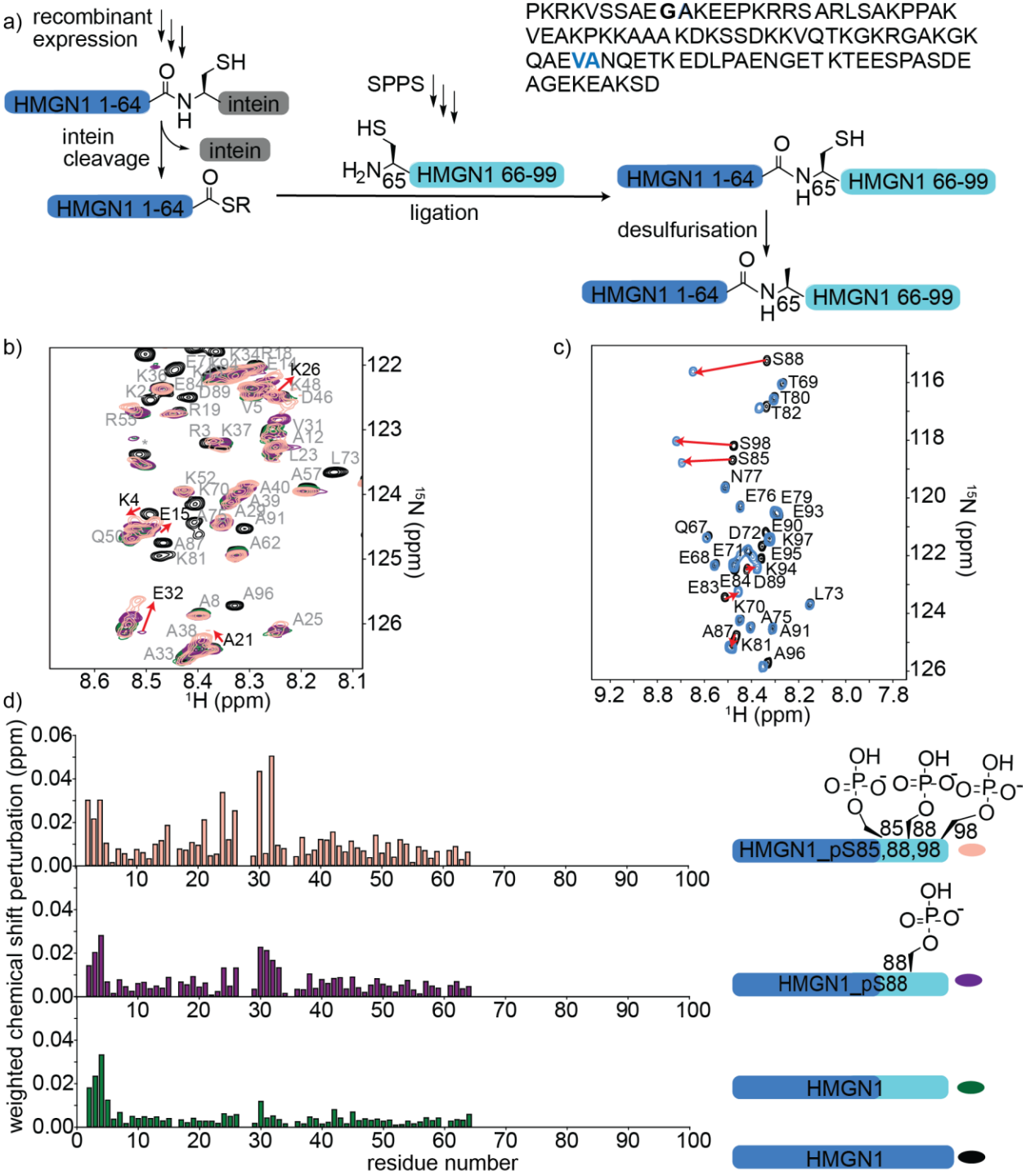


Figure 4. Semi-synthesis and NMR of HMGN1 bearing C-terminal PTMs. a) Semi-synthesis strategy involving recombinant expression of the N-terminal segment HMGN1_1-64 (with or without uniform ^{15}N isotope labelling) and synthesis of C-terminal segment HMGN1_65-99_A65C (and variants bearing PTMs) by solid phase peptide synthesis (SPPS). Native chemical ligation followed by desulfurization of cysteine to alanine yielded site-specifically modified and

segmentally labelled variants of HMGN1. The sequence of HMGN1 is shown with the selected ligation junction in blue. b) Section of ^{15}N-HSQC NMR spectrum overlay of segmentally labelled HMGN1 variants showing resonances in the labelled section that shift when C-terminal PTMs are present (red arrows and residue labels in black). Spectra are shown for HMGN1_15N13C (black), HMGN1_unmodC_15N (green), HMGN1_pS88_15N (purple), HMGN1_pS85,88,98_15N (pink), corresponding to the schematic structures in (d). Full spectra are shown in Supplementary Data S6. c) Section of ^{15}N-HSQC NMR spectrum overlay of synthetic HMGN1_65-99 (black) and HMGN1_65-99_pS85,88,98 (blue), acquired at natural abundance. Resonances that shift on introduction of PTMs are shown by red arrows. Full spectra and spectra of HMGN1_65-99_pS85 and HMGN1_65-99_pS8 are shown in Supplementary Data S6. d) Chemical shift perturbation plots for the segmentally-labelled HMGN1 variants bearing C-terminal PTMs, as shown in (b). Light blue segments are synthetic and unlabelled, dark blue segments are ^{15}N isotope labelled.

As for the N-terminal ligation strategy, comparison of the ^{15}N HSQC spectra of segmentally labelled unmodified HMGN1 with that of fully labelled HMGN1 showed that our semi-synthesis strategy did not change the conformation of unmodified HMGN1 (Supplementary Data S6). Resonances for residues 65-99 absent in the spectra of segmentally labelled HMGN1 aided in assignment of fully labelled HMGN1. ^{15}N-HSQC spectra of the variants bearing pS85 and pS85,88,98 PTMs showed shifts and broadening of several signals, indicating changes in equilibrium conformations, as for the N-terminal PTMs. As shown in Figure 4d, the resonances most affected by the phosphorylation PTMs lie in the region 24-34, which is within the nucleosome binding domain. Larger changes for the triply-modified variant HMGN1_pS85,88,98 compared to the singly-modified variant suggest that the PTMs have a cumulative effect on the conformational preferences. In contrast to the shifts observed on the introduction of N-terminal PTMs, where signals in a region shifted to a new position, the C-terminal phosphorylations caused peaks to broaden, as exemplified by Glu32 (Figure 4b). This indicates that several conformations are present and interchanging on the NMR timescale. Furthermore, the region of residues affected is more dispersed than that observed for the N-terminal PTMs (Figure 3b). Whereas several studies have focused on phosphorylations in the NBD, the role of phosphorylation in the C-terminal regulatory domain is less well characterized. Phosphorylation of Ser6, Ser85, Ser88 and Ser98 was detected in HMGN1 isolated from MCF-7

breast cancer cells and it was noted that phosphorylation would introduce even more negative charges in this region, which is already highly negatively charged.[64]

In the NMR spectra of the synthetic unlabelled HMGN1_65-99 segments (Figure 4c and S6), downfield shifts in the backbone amide resonances of the phosphorylated residues were observed as expected for pSer. However, the changes in chemical shift upon phosphorylation relative to the unmodified residue differ depending on the sequence position; whereas Ser85 shifts by 0.21 ppm in the ^{1}H dimension, Ser88 shifts by 0.32 ppm, Ser98 shifts by 0.24 ppm, Ser6 shifts by 0.51 ppm (Figure 3c) and the shift observed for the random coil peptide is 0.32 ppm.[49] Although HMGN1 is unstructured and most resonances have chemical shifts close to their random coil values (Figure 2e and 2f), this indicates that shifts upon phosphorylation are dependent on the sequence context. We hypothesized that the changes in conformational equilibria observed on phosphorylation might be caused by a shift in proline *cis/trans* isomer proportions, especially as HMGN1 contains several proline residues. In particular, the PTM pS85 precedes a proline residue and we explored whether serine phosphorylation might alter the *cis/trans* proline ratio, as has been reported for some proteins such as those which are substrates of the peptidyl-prolyl *cis/trans* isomerase Pin1.[41, 65] This possibility was also suggested when Ser85 phosphorylation was first detected,[64] but was not further investigated. Using short model peptides derived from those used to determine the random coil shifts of posttranslationally modified residues,[49] we compared the proportion of *cis*- and *trans*-proline conformations when preceded by serine or phosphoserine. As shown in the Supplementary Data (S6), although characteristic shifts are observed for phosphoserine, we did not observe any significant change in the proportion of *cis*-proline when preceded by serine (10%) compared to phosphoserine (11%), as determined by comparison of the relative peak volumes. These ratios are similar to those observed for proline flanked by glycine residues in short random coil peptides.[49] Nevertheless, this result does not preclude the possibility that the phosphoserine could be a recognition motif for a peptidyl-prolyl isomerase.

Binding of modified and unmodified HMGN1 variants to nucleosome core particles and DNA

We compared binding of recombinant and semi-synthetic HMGN1 variants and segments to mononucleosomes to see if the conformational changes observed in the nucleosome binding region correspond to functional changes in nucleosome binding. Figure 5 shows the results of

electrophoretic gel mobility assays in which HMGN1 variants were titrated with reconstituted nucleosomes. Overall, the results show that HMGN1 binds to nucleosomes and the major band corresponds to two HMGN1 molecules per nucleosome, as would be expected from binding of one HMGN1 to each face of the nucleosome. This result agrees with early studies on HMGN1 (previously HMG 14),[66] however in the nucleus, the number of HMGN1 molecules is thought to be limiting, with 0.5 – 1.5 molecules per nucleosome.[15, 27] The semi-synthetic unmodified variants (HMGN1_unmodN and HMGN1_unmodC) showed similar binding to recombinantly produced HMGN1 (Fig. 5a), agreeing with the NMR data in showing that the semi-synthesis protocol does not change the structure or function of the protein. However, the long-range effects of the PTMs observed in the NMR spectra, some of which are in the nucleosome binding region, do not seem to have a significant effect on binding to nucleosome core particles (Fig. 5b and 5c). This suggests that the N- and C-terminal PTMs of HMGN1 affect other functions of HMGN1, such as competition with histones, the formation of high-order chromatin structures, binding to linker DNA, or serving as a docking pad for other chromatin interacting proteins (analogous to the histone code).[1] Future experiments will test these and other possibilities.

Truncated HMGN1 variants in which the N-terminal residues (1-10) or C-terminal residues (65-99) were missing were also compared with full-length HMGN1 for their nucleosome-binding properties, as shown in Supplementary Figure S9a. Considering the lower molecular weights, the truncated variants show similar nucleosome binding to full-length HMGN1, showing that the N- and C-terminal regions are not directly involved in nucleosome core binding. These results support the predicted orientation of HMGN1 shown in Figure 1a, based on that predicted for HMGN2 using molecular modelling techniques.[26] Experiments in which the HMGN1 variants were mixed with DNA (147 bp AlexaFluor488 labelled DNA containing the 601 nucleosome positioning sequence, Supplementary Data Figure S9b) indicated that there is no difference in binding of the HNGN1 variants to naked DNA.

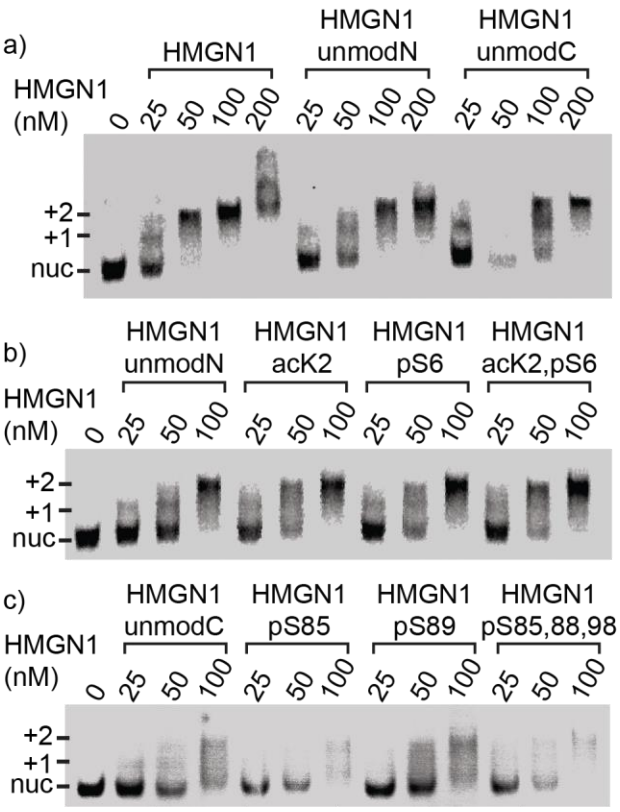


Figure 5. Nucleosome-binding of HMGN1 variants. Electrophoretic mobility shift assays showing mononucleosome binding of HMGN1 variants bearing N- or C-terminal PTMs. Position of nucleosomes bound by one HMGN1 molecules (+1), and two HMGN1 molecules (+2) are shown in relation to the unbound nucleosome (nuc). a) Mononucleosomes (50 nM) were incubated with full length recombinant or semi-synthetic HMGN1 at molar ratios of 0.5:1, 1:1, 2:1 and 4:1 HMGN1 molecules per nucleosome. Binding reactions were separated on 5% TBE-acrylamide gels and then scanned for fluorescence. b) and c) Comparison of semi-synthetic HMGN1 proteins bearing N- or C-terminal PTMs. HMGN1 variants were incubated with mononucleosomes at molar rations of 0.5:1, 1:1, and 2:1 HMGN1 per nucleosome.

In summary, we have generated the first set of HMGN1 variants bearing site-specific PTMs and combinations of PTMs, focusing on the N- and C-terminal PTMs and introducing genuine phosphoserine and acetyllysine residues rather than mimics of these PTMs. We have characterised HMGN1 using NMR spectroscopy and report the chemical shifts of this intrinsically disordered protein. Employing protein semi-synthesis techniques also allowed us to introduce segmental isotope labels and to detect subtle structural changes in regions distant to the modification site; changes that would be difficult to measure without access to specifically-

modified variants or using other structural biology techniques that are more suited to structured proteins. For HMGN1, we show that PTMs cause chemical shift changes in the modified residues, as predicted by the random coil shifts, and shifts in their flanking residues. Moreover, N- and C-terminal PTMs of HMGN1 cause long-range chemical shift perturbations in residues within the nucleosome-binding domain but this did not significantly affect the binding to nucleosome cores. These results suggest that the N- and C-terminal PTMs have other biological functions. Overall, this study shows that our approach of combining protein semi-synthesis techniques with structural biology and functional assays allows us to study the effects of specific posttranslational modifications and combinations of modifications, providing mechanistic insights that link protein structure and dynamics with function. Furthermore, understanding the roles of the abundant non-histone components of chromatin and how they regulate chromatin dynamics will help to gain a more complete picture of how gene expression is regulated in health and disease.

Acknowledgements

We thank Philipp Schilling, Manuel Felkl, Thomas Kremsmayer, Gilyana Kücükkaplan and Oliver Gajsek for assistance with peptide synthesis. We are also grateful to Hanspeter Kählig, Georg Kontaxis (University of Vienna) and Gregory Pierens (The University of Queensland, Centre for Advanced Imaging) for assistance with acquiring NMR data.

ACC is supported by a University of Queensland Development Fellowship (Project 613982). Isotope labelled reagents were provided by Cambridge Isotope Laboratories Inc. through the award of a CIL Research Grant to ACC. CFWB is funded by the Vienna Science and Technology Fund (WWTF) through project LS17-0008.

Supporting Information

DNA and peptide sequences of protein constructs, sequences and labelling patterns of semi-synthetic HMGN1 variants, chemical structures, yields, analytical MS and HPLC data, full gels and NMR spectra, and chemical shift assignments are available in the attached supporting information.

Conflict of Interest

The authors declare no conflict of interest

ORCIDs

Debra Urwin: 0000-0002-9798-1233

David J. Tremethick: 0000-0001-5274-8078

Anne C. Conibear: 0000-0002-5482-6225

Christian F. W. Becker: 0000-0002-8890-7082

K. Johan Rosengren: 0000-0002-5007-8434

References

1. Luger, K.; Dechassa, M. L.; Tremethick, D. J., New insights into nucleosome and chromatin structure: An ordered state or a disordered affair? *Nat. Rev. Mol. Cell. Biol.* **2012,** *13* (7), 436-447.

2. Zhou, K.; Gaullier, G.; Luger, K., Nucleosome structure and dynamics are coming of age. *Nat. Struct. Mol. Biol.* **2019,** *26* (1), 3-13.

3. Baldi, S.; Korber, P.; Becker, P. B., Beads on a string-nucleosome array arrangements and folding of the chromatin fiber. *Nat. Struct. Mol. Biol.* **2020,** *27* (2), 109-118.

4. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J., Crystal structure of the nucleosome core particle at 2.8 a resolution. *Nature* **1997,** *389* (6648), 251-260.

5. McGinty, R. K.; Tan, S., Nucleosome structure and function. *Chem. Rev.* **2015,** *115* (6), 2255-2273.

6. Li, G.; Widom, J., Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol.* **2004,** *11* (8), 763-769.

7. van Emmerik, C. L.; van Ingen, H., Unspinning chromatin: Revealing the dynamic nucleosome landscape by NMR. *Prog. Nucl. Magn. Res. Spectrosc.* **2019,** *110*, 1-19.

8. Bustin, M.; Reeves, R., High-mobility-group chromosomal proteins: Architectural components that facilitate chromatin function. *Prog. Nucl. Acid Res. Mol. Biol.* **1996,** *54*, 35-100.

9. Reeves, R., Nuclear functions of the HMG proteins. *Biochim. Biophys. Acta.* **2010,** *1799* (1-2), 3-14.

10. Goodwin, G. H.; Sanders, C.; Johns, E. W., A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. *Eur. J. Biochem.* **1973,** *38* (1), 14-19.

11. Nanduri, R.; Furusawa, T.; Bustin, M., Biological functions of HMGN chromosomal proteins. *Int. J. Mol. Sci.* **2020,** *21* (2).

12. He, B.; Deng, T.; Zhu, I.; Furusawa, T.; Zhang, S.; Tang, W.; Postnikov, Y.; Ambs, S.; Li, C. C.; Livak, F.; Landsman, D.; Bustin, M., Binding of HMGN proteins to cell specific enhancers stabilizes cell identity. *Nat. Commun.* **2018,** *9* (1), 5240.

13. Murphy, K. J.; Cutter, A. R.; Fang, H.; Postnikov, Y. V.; Bustin, M.; Hayes, J. J., HMGN1 and 2 remodel core and linker histone tail domains within chromatin. *Nucleic Acids Res.* **2017,** *45* (17), 9917-9930.

14. Postnikov, Y. V.; Bustin, M., Functional interplay between histone H1 and HMG proteins in chromatin. *Biochim. Biophys. Acta.* **2016,** *1859* (3), 462-467.

15. Rattner, B. P.; Yusufzai, T.; Kadonaga, J. T., HMGN proteins act in opposition to ATP-dependent chromatin remodeling factors to restrict nucleosome mobility. *Mol. Cell* **2009,** *34* (5), 620-626.

16. Postnikov, Y. V.; Belova, G. I.; Lim, J. H.; Bustin, M., Chromosomal protein HMGN1 modulates the phosphorylation of serine 1 in histone H2A. *Biochemistry* **2006,** *45* (50), 15092-15099.

17. Wei, F.; Yang, D.; Tewary, P.; Li, Y. N.; Li, S.; Chen, X.; Howard, O. M. Z.; Bustin, M.; Oppenheim, J. J., The alarmin HMGN1 contributes to antitumor immunity and is a potent immunoadjuvant. *Cancer Res.* **2014,** *74* (21), 5989-5998.

18. Yang, D.; Postnikov, Y. V.; Li, Y.; Tewary, P.; de la Rosa, G.; Wei, F.; Klinman, D.; Gioannini, T.; Weiss, J. P.; Furusawa, T.; Bustin, M.; Oppenheim, J. J., High-mobility group nucleosome-binding protein 1 acts as an alarmin and is critical for lipopolysaccharide-induced immune responses. *J. Exp. Med.* **2012,** *209* (1), 157-171.

19. Zuo, B.; Qi, H.; Lu, Z.; Chen, L.; Sun, B.; Yang, R.; Zhang, Y.; Liu, Z.; Gao, X.; You, A.; Wu, L.; Jing, R.; Zhou, Q.; Yin, H., Alarmin-painted exosomes elicit persistent antitumor immunity in large established tumors in mice. *Nat. Commun.* **2020,** *11* (1), 1790.

20. Schauwecker, S. M.; Kim, J. J.; Licht, J. D.; Clevenger, C. V., Histone H1 and chromosomal protein hmgn2 regulate prolactin-induced stat5 transcription factor recruitment and function in breast cancer cells. *J. Biol. Chem.* **2017,** *292* (6), 2237-2254.

21. Alam, M. M.; Yang; Trivett, A.; Meyer, T. J.; Oppenheim, J. J., HMGN1 and r848 synergistically activate dendritic cells using multiple signaling pathways. *Front. Immunol.* **2018,** *9*, 2982.

22. Reeves, R., High mobility group (HMG) proteins: Modulators of chromatin structure and DNA repair in mammalian cells. *DNA repair* **2015,** *36*, 122-136.

23. Apelt, K.; Zoutendijk, I.; Gout, D. Y.; Wondergem, A. P.; van den Heuvel, D.; Luijsterburg, M. S., Human HMGN1 and hmgn2 are not required for transcription-coupled DNA repair. *Sci. Rep.* **2020,** *10* (1), 4332.

24. Mowery, C. T.; Reyes, J. M.; Cabal-Hierro, L.; Higby, K. J.; Karlin, K. L.; Wang, J. H.; Kimmerling, R. J.; Cejas, P.; Lim, K.; Li, H.; Furusawa, T.; Long, H. W.; Pellman, D.; Chapuy, B.; Bustin, M.; Manalis, S. R.; Westbrook, T. F.; Lin, C. Y.; Lane, A. A., Trisomy of a down syndrome critical region globally amplifies transcription via HMGN1 overexpression. *Cell Rep.* **2018,** *25* (7), 1898-1911.e1895.

25. Shimahara, H.; Hirano, T.; Ohya, K.; Matsuta, S.; Seeram, S. S.; Tate, S., Nucleosome structural changes induced by binding of non-histone chromosomal proteins HMGN1 and hmgn2. *FEBS Open Bio* **2013,** *3*, 184-191.

26. Kato, H.; van Ingen, H.; Zhou, B. R.; Feng, H.; Bustin, M.; Kay, L. E.; Bai, Y., Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc. Natl. Acad. Sci. USA* **2011,** *108* (30), 12283-12288.

27. Pogna, E. A.; Clayton, A. L.; Mahadevan, L. C., Signalling to chromatin through post-translational modifications of HMGN. *Biochim. Biophys. Acta.* **2010,** *1799* (1-2), 93-100.

28. Zhang, Q.; Wang, Y., HMG modifications and nuclear function. *Biochim. Biophys. Acta.* **2010,** *1799* (1-2), 28-36.

29. Zhang, Q.; Wang, Y., High mobility group proteins and their post-translational modifications. *Biochim. Biophys. Acta.* **2008,** *1784* (9), 1159-1166.

30. Louie, D. F.; Gloor, K. K.; Galasinski, S. C.; Resing, K. A.; Ahn, N. G., Phosphorylation and subcellular redistribution of high mobility group proteins 14 and 17, analyzed by mass spectrometry. *Protein Sci.* **2000,** *9* (1), 170-179.

31. Prymakowska-Bosak, M.; Misteli, T.; Herrera, J. E.; Shirakawa, H.; Birger, Y.; Garfield, S.; Bustin, M., Mitotic phosphorylation prevents the binding of HMGN proteins to chromatin. *Mol Cell. Biol.* **2001,** *21* (15), 5169-5178.

32. Lim, J. H.; West, K. L.; Rubinstein, Y.; Bergel, M.; Postnikov, Y. V.; Bustin, M., Chromosomal protein HMGN1 enhances the acetylation of lysine 14 in histone h3. *EMBO J.* **2005,** *24* (17), 3038-3048.

33. Postnikov, Y.; Bustin, M., Regulation of chromatin structure and function by HMGN proteins. *Biochim. Biophys. Acta.* **2010,** *1799* (1-2), 62-68.

34. Bonfiglio, J. J.; Fontana, P.; Zhang, Q.; Colby, T.; Gibbs-Seymour, I.; Atanassov, I.; Bartlett, E.; Zaja, R.; Ahel, I.; Matic, I., Serine ADP-ribosylation depends on hpf1. *Mol. Cell* **2017,** *65* (5), 932-940.e936.

35. Fierz, B.; Muir, T. W., Chromatin as an expansive canvas for chemical biology. *Nat. Chem. Biol.* **2012,** *8* (5), 417-427.

36. Muller, M. M.; Muir, T. W., Histones: At the crossroads of peptide and protein chemistry. *Chem. Rev.* **2015,** *115* (6), 2296-2349.

37. Boichenko, I.; Fierz, B., Chemical and biophysical methods to explore dynamic mechanisms of chromatin silencing. *Curr. Opin. Chem. Biol.* **2019,** *51*, 1-10.

38. Conibear, A. C.; Watson, E. E.; Payne, R. J.; Becker, C. F. W., Native chemical ligation in protein synthesis and semi-synthesis. *Chem. Soc. Rev.* **2018,** *47* (24), 9046-9068.

39. Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O., Native chemical ligation and extended methods: Mechanisms, catalysis, scope, and limitations. *Chem. Rev.* **2019,** *119* (12).

40. Thompson, R. E.; Muir, T. W., Chemoenzymatic semisynthesis of proteins. *Chem. Rev.* **2020,** *120* (6), 3051-3126.

41. Theillet, F. X.; Smet-Nocca, C.; Liokatis, S.; Thongwichian, R.; Kosten, J.; Yoon, M. K.; Kriwacki, R. W.; Landrieu, I.; Lippens, G.; Selenko, P., Cell signaling, post-translational protein modifications and NMR spectroscopy. *J. Biomol. NMR* **2012,** *54* (3), 217-236.

42. Bah, A.; Forman-Kay, J. D., Modulation of intrinsically disordered protein function by post-translational modifications. *J. Biol. Chem.* **2016,** *291* (13), 6696-6705.

43. Zheng, J. S.; Tang, S.; Qi, Y. K.; Wang, Z. P.; Liu, L., Chemical synthesis of proteins using peptide hydrazides as thioester surrogates. *Nat. Protoc.* **2013,** *8* (12), 2483-2495.

44. Marley, J.; Lu, M.; Bracken, C., A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **2001,** *20* (1), 71-75.

45. Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D., 1h, 13c and 15n chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **1995,** *6* (2), 135-140.

46. Harris, R. K.; Becker, E. D.; Cabral de Menezes, S. M.; Granger, P.; Hoffman, R. E.; Zilm, K. W., Further conventions for NMR shielding and chemical shifts (IUPAC recommendations 2008). *Pure Appl. Chem.* **2008,** *80* (1), 59-84.

47. Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D., The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **2005,** *59* (4), 687-696.

48. Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D., 1h, 13c and 15n random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* **1995,** *5* (1), 67-81.

49. Conibear, A. C.; Rosengren, K. J.; Becker, C. F. W.; Kaehlig, H., Random coil shifts of posttranslationally modified amino acids. *J. Biomol. NMR* **2019,** *73* (10-11), 587-599.

50. Williamson, M. P., Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc* **2013,** *73*, 1-16.

51. Stevens, A. J.; Sekar, G.; Shah, N. H.; Mostafavi, A. Z.; Cowburn, D.; Muir, T. W., A promiscuous split intein with expanded protein engineering applications. *Proc. Natl. Acad. Sci. USA* **2017,** *114* (32), 8538-8543.

52. Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Kent Wenger, R.; Yao, H.; Markley, J. L., Biomagresbank. *Nucl. Acids Res.* **2008,** *36* (Database issue), D402-D408.

53. Ferrage, F.; Piserchio, A.; Cowburn, D.; Ghose, R., On the measurement of 15n-{1h} nuclear overhauser effects. *J. Magn. Reson.* **2008,** *192* (2), 302-313.

54. Lowary, P. T.; Widom, J., New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **1998,** *276* (1), 19-42.

55. Ryan, D. P.; Tremethick, D. J., The interplay between H2A.Z and h3k9 methylation in regulating hp1alpha binding to linker histone-containing chromatin. *Nucleic Acids Res.* **2018,** *46* (18), 9353-9366.

56. Luger, K.; Rechsteiner, T. J.; Richmond, T. J., Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol. Biol.* **1999,** *119*, 1-16.

57. Luger, K.; Rechsteiner, T. J.; Flaus, A. J.; Waye, M. M.; Richmond, T. J., Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.* **1997,** *272* (3), 301-311.

58. Lim, J. H.; Catez, F.; Birger, Y.; Postnikov, Y. V.; Bustin, M., Preparation and functional analysis of HMGN proteins. *Methods Enzymol.* **2004,** *375*, 323-342.

59. Wan, Q.; Danishefsky, S. J., Free-radical-based, specific desulfurization of cysteine: A powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew.Chem Int.Ed Engl.* **2007,** *46* (48), 9248-9252.

60. Kulkarni, S. S.; Sayers, J.; Premdjee, B.; Payne, R. J., Rapid and efficient protein synthesis through expansion of the native chemical ligation concept. *Nat. Rev. Chem.* **2018,** *2* (4), 0122.

61. Lim, J. H.; Catez, F.; Birger, Y.; West, K. L.; Prymakowska-Bosak, M.; Postnikov, Y. V.; Bustin, M., Chromosomal protein HMGN1 modulates histone h3 phosphorylation. *Mol. Cell* **2004,** *15* (4), 573-584.

62. Zhang, X.; Perugini, M. A.; Yao, S.; Adda, C. G.; Murphy, V. J.; Low, A.; Anders, R. F.; Norton, R. S., Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein msp2. *J. Mol. Biol.* **2008,** *379* (1), 105-121.

63. Muir, T. W.; Sondhi, D.; Cole, P. A., Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci. USA* **1998,** *95* (12), 6705-6710.

64. Zou, Y.; Jiang, X.; Wang, Y., Identification of novel in vivo phosphorylation sites in high mobility group n1 protein from the mcf-7 human breast cancer cells. *Biochemistry* **2004,** *43* (20), 6322-6329.

65. Schutkowski, M.; Bernhardt, A.; Zhou, X. Z.; Shen, M.; Reimer, U.; Rahfeld, J. U.; Lu, K. P.; Fischer, G., Role of phosphorylation in determining the backbone dynamics of the serine/threonine-proline motif and pin1 substrate recognition. *Biochemistry* **1998,** *37* (16), 5566-5575.

66. Mardian, J. K.; Paton, A. E.; Bunick, G. J.; Olins, D. E., Nucleosome cores have two specific binding sites for nonhistone chromosomal proteins HMG 14 and HMG 17. *Science* **1980,** *209* (4464), 1534-1536.

Supplementary Data for

Site-specific modification and segmental isotope labelling of HMGN1 reveals long-range conformational perturbations caused by posttranslational modifications

Gerhard Niederacher[1], Debra Urwin[2], David J. Tremethick[2], K. Johan Rosengren[3], Christian F. W. Becker[1], and Anne C. Conibear[3*]

[1] University of Vienna, Faculty of Chemistry, Institute of Biological Chemistry, Währinger Straße 38, 1090 Vienna, Austria.

[2] The Australian National University, John Curtin School of Medical Research, Department of Genome Sciences, ACT 2601, Australia.

[3] The University of Queensland, School of Biomedical Sciences, Brisbane, QLD 4072, Australia.

[*]Corresponding author:

Anne C. Conibear

The University of Queensland, School of Biomedical Sciences, St Lucia 4072, Brisbane, Australia.

E-mail: a.conibear@uq.edu.au

Phone: + 61-7-3365-1738

Table of Contents

1. Gene constructs and sequences

1.1 Full length HMGN1 (His$_6$-TEV-HMGN1)

DNA sequence:

CAT ATG CAC CAC CAC CAC CAC CAC GAA AAC CTG TAT TTT CAG AGC CCG AAA CGC AAA GTG TCC TCT GCC GAA GGA GCA GCG AAA GAA GAA CCG AAA CGT CGC TCA GCT CGC TTA AGC GCG AAA CCA CCG GCG AAA GTT GAA GCGAAA CCG AAG AAA GCT GCA GCC AAA GAC AAG TCG AGC GAC AAG AAA GTC CAG ACC AAG GGC AAA CGTGGT GCC AAA GGC AAA CAA GCC GAA GTA GCG AAT CAG GAG ACT AAA GAG GAT CTG CCC GCA GAA AACGGG GAA ACG AAA ACC GAA GAG AGT CCT GCA TCG GAT GAA GCT GGT GAG AAG GAA GCG AAA AGC GAT TGA CTC GAG

Amino acid sequence:

MHHHHHHENLYFQSPKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDK SSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD

1.2 His$_6$-TEV-HMGN1_11-99

DNA sequence:

CAT ATG CAT CAC CAT CAC CAC CAT GAA AAT CTG TAT TTT CAG TGT GCT AAA GAG GAG CCT AAG CGT CGT TCG GCT CGT TTG AGT GCTAAG CCT CCG GCT AAG GTG GAA GCT AAG CCT AAA AAG GCT GCG GCC AAA GAT AAA TCG TCT GAC AAA AAA GTA CAA ACG AAG GGA AAA CGT GGG GCC AAA GGT AAA CAG GCG GAA GTG GCTAAC CAG GAG ACC AAG GAG GAC CTT CCT GCG GAA AAC GGT GAG ACA AAG ACA GAA GAA AGT CCT GCA TCT GAT GAG GCA GGA GAG AAA GAG GCG AAA TCG GAT TGA CTC GAG

Amino acid sequence:

MHHHHHHENLYFQCAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGK RGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD

1.3 HMGN1_1-65_Mxe_His$_7$_CBD

DNA sequence (Mxe-HIS$_7$-CBD not included):

CAT ATG CCG AAA CGC AAA GTC AGC AGT GCA GAA GGT GCA GCG AAA GAG GAA CCG AAA CGC CGT TCT GCT CGC CTG TCA GCC AAA CCT CCA GCG AAA GTT GAG GCG AAA CCG AAG AAA GCA GCC GCC AAG GAC AAA TCG TCC GAT AAG AAG GTG CAG ACC AAA GGGAAA CGT GGC GCT AAA GGC AAA CAA GCG GAA GTA TGC ATC ACG GGA GAT GCA CTA GT

Amino acid sequence:

PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGA
KGKQAEV(CITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLFH
SGEHPVYTVRTVEGLRVTGTANHPLLCLVDVAGVPTLLWKLIDEIKPGDYAVIQRSAFS
VDCAGFARGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELTDGRFYYAKVASV
TDAGVQPVYSLRVDTADHAFITNGFVSHATGLTGIHHHHHHHHSGLNSGLTTNPGVSAW
QVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ)

2. Table of HMGN1 variants generated in this study

Name	MW (Da)	Sequence
HMGN1_S0	10615.0	SPKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_S0_15N	10755.8	SPKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_S0_15N13C	11200.0	SPKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_unmodN	10527.8	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_unmodN_15N	10652.8	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_acK2	10569.8	P**K(ac)**RKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_acK2_15N	10694.9	P**K(ac)**RKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_pS6	10607.8	PKRKV**(pS)**SAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_pS6_15N	10732.8	PKRKV**(pS)**SAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_acK2_pS6	10649.8	P**K(ac)**RKV**(pS)**SAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_acK2_pS6_15N	10774.9	P**K(ac)**RKV**(pS)**SAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_unmod_C	10527.8	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_unmod_C_15N	10623.7	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD

HMGN1_pS85	10607.8	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGK RGAKGKQAEVANQETKEDLPAENGETKTEE**(pS)**PASDEAGEKEAKSD
HMGN1_pS88	10607.8	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGK RGAKGKQAEVANQETKEDLPAENGETKTEESPA**(pS)**DEAGEKEAKSD
HMGN1_pS88 _15N	10703.6	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGK RGAKGKQAEVANQETKEDLPAENGETKTEESPA**(pS)**DEAGEKEAKSD
HMGN1_pS85 ,88,98	10767.8	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGK RGAKGKQAEVANQETKEDLPAENGETKTEE**(pS)**PA**(pS)**DEAGEKEAK**(pS)**D
HMGN1_pS85 ,88,98_15N	10863.6	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGK RGAKGKQAEVANQETKEDLPAENGETKTEE**(pS)**PA**(pS)**DEAGEKEAK**(pS)**D
HMGN1_1-65	6810.0	PKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGK RGAKGKQAEV
HMGN1_11-99	9487.6	AAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAE VANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_65-99	3767.8	CNQETKEDLPAENGETKTEESPASDEAGEKEAKSD
HMGN1_65-99_pS88	3847.8	CNQETKEDLPAENGETKTEESPA**(pS)**DEAGEKEAKSD
HMGN1_65-99_pS85	3847.8	CNQETKEDLPAENGETKTEE**(pS)**PASDEAGEKEAKSD
HMGN1_65-99_pS85,88,98	4007.8	CNQETKEDLPAENGETKTEE**(pS)**PA**(pS)**DEAGEKEAK**(pS)**D
HMGN1_1-26	2838.3	PKRKVSSAEGAAKEEPKRRSARLSAK
HMGN1_1-26_acK2,pS6	2960.3	P**K(ac)**RKV**(pS)**SAEGAAKEEPKRRSARLSAK
Random coil_SP	472.5	Ac-GGSPGG-NH_2
Random coil_pSP	551.5	Ac-GG**(pS)**PGG-NH_2

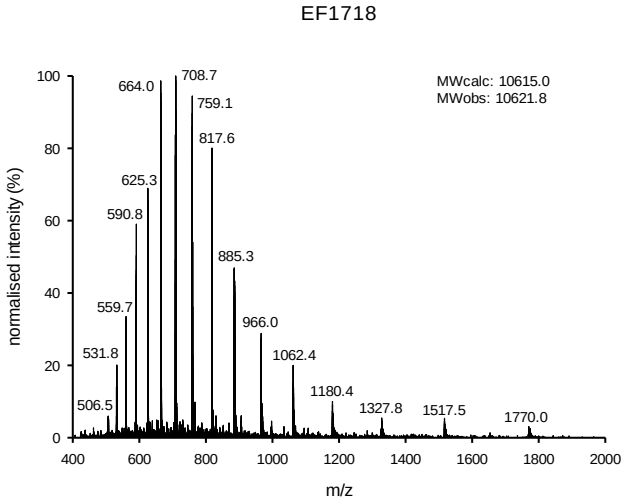
K(ac) = *N*-acetyllysine, **(pS)** = phosphoserine, Ac = N-terminal acetylation

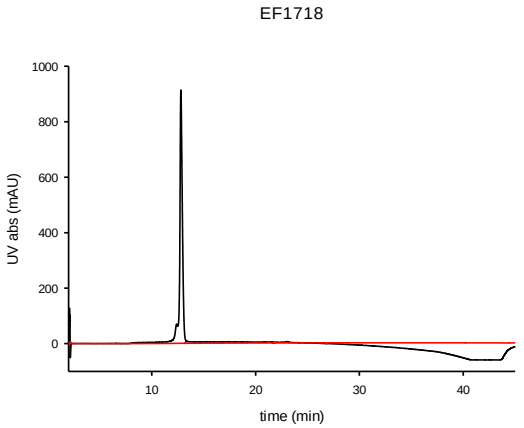
Residues in blue are ^{15}N-labelled, residues in green are ^{15}N/^{13}C-labelled.

3. Mass spectra and analytical HPLC analyses of HMGN1 variants

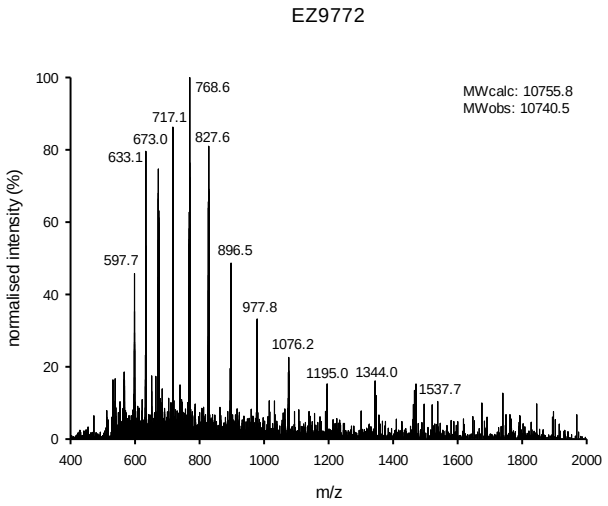
ESI mass spectra were obtained in positive ion mode and intensities are normalised to the highest intensity peak. Analytical RP-HPLC was carried out using a C4 analytical HPLC column and a 2%/min gradient of acetonitrile (0.045% TFA) in water (0.05% TFA). UV absorbance was detected at 214 nm (black traces) and 280 nm (red traces).

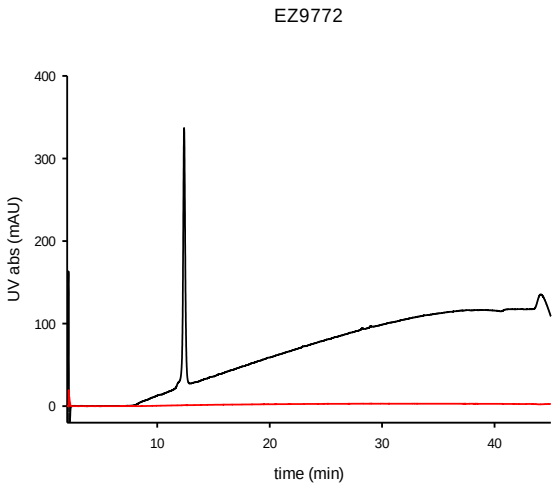
HMGN1_S0 (Yield after expression, purification, TEV cleavage and HPLC purification: ~ 1.1 mg/L of culture, 1.6 mg)



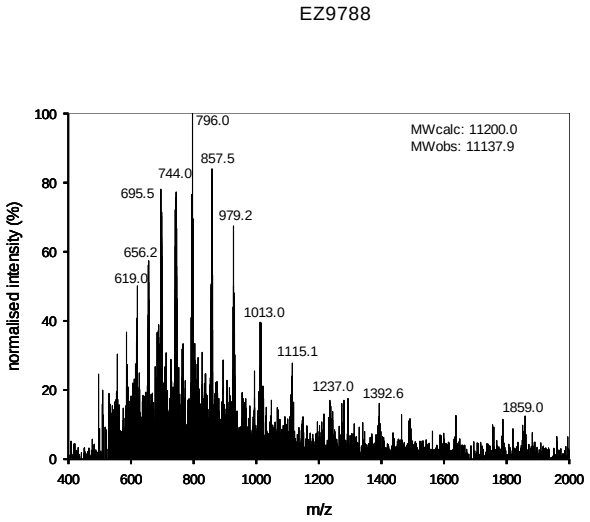


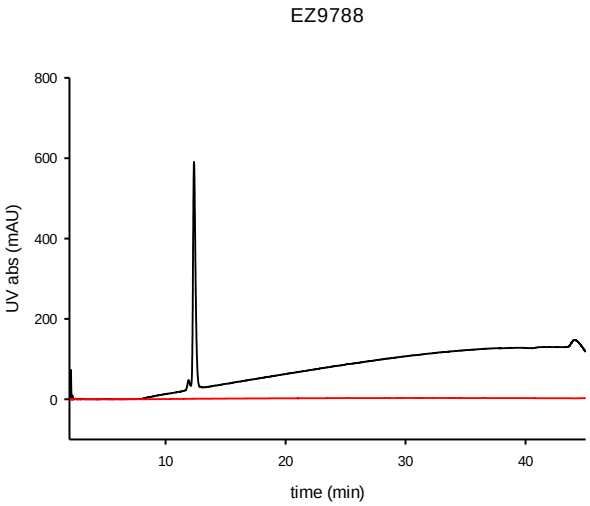
HMGN1_S0_15N (Yield after expression, purification, TEV cleavage and HPLC purification: ~ 1 mg/L of rich medium, 250 mL minimal medium, 1.2 mg)



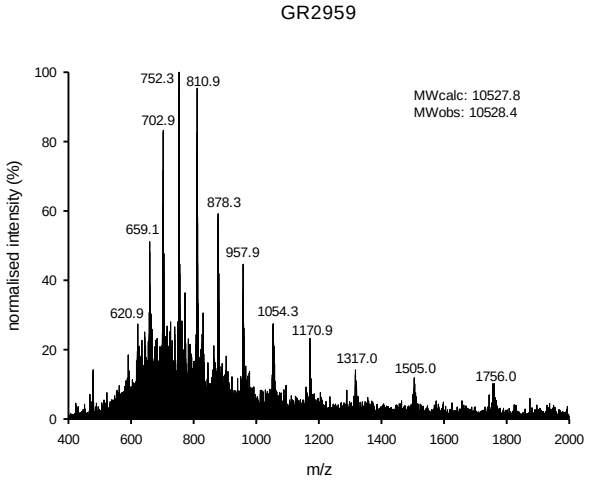


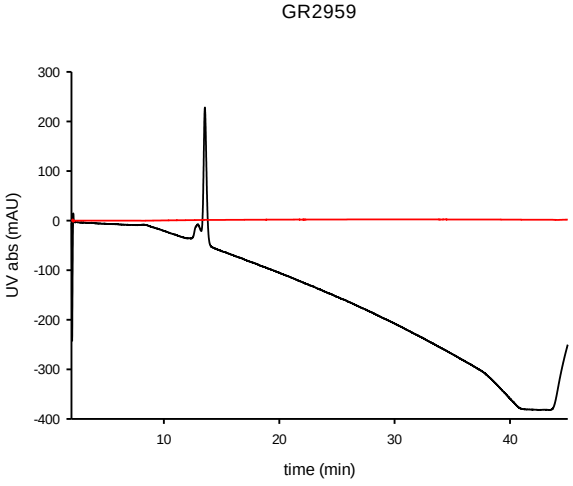
HMGN1_S0_15N13C (Yield after expression, purification, TEV cleavage and HPLC purification: ~ 0.8 mg/L of rich medium, 250 mL minimal medium, 1.7 mg)



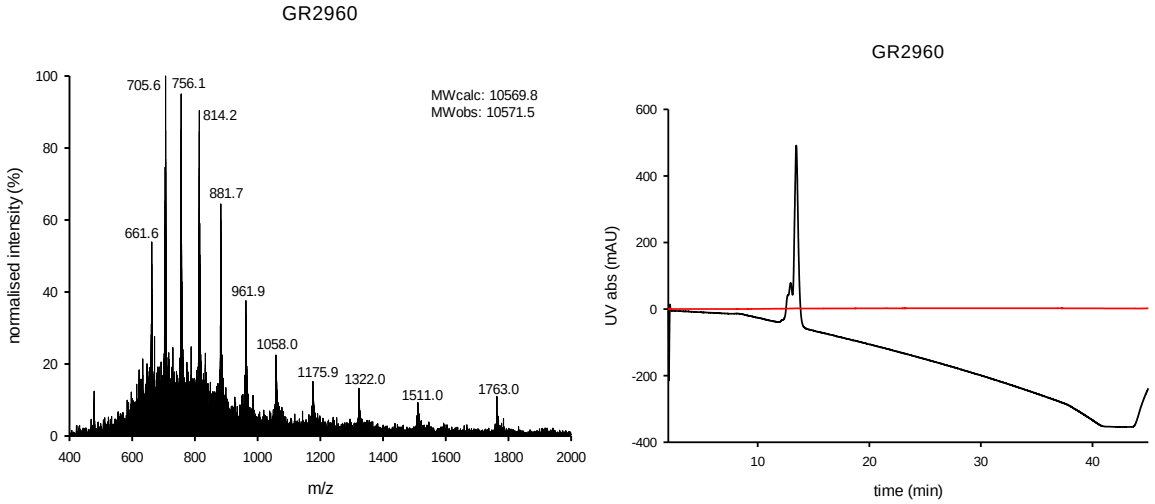


HMGN1_unmodN (Yield after ligation, desulfurisation and purification: ~ 59%, 3.2 mg for unlabelled; ~11%, 0.6 mg for ^{15}N-labelled)

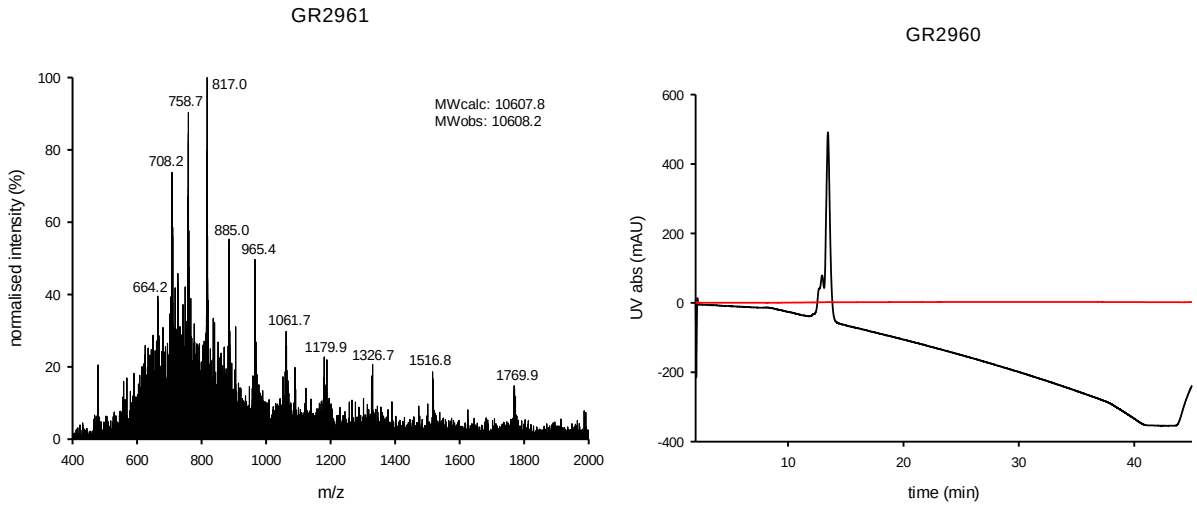




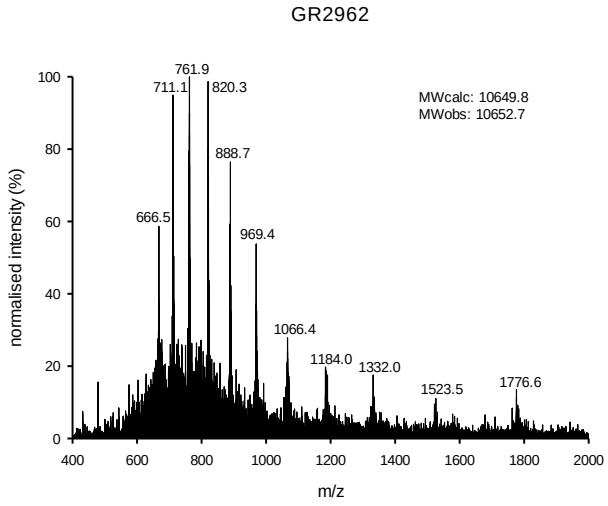
HMGN1_acK2 (Yield after ligation, desulfurisation and purification: ~ 86%, 4.7 mg for unlabelled; ~9%, 0.5 mg for ^{15}N-labelled)

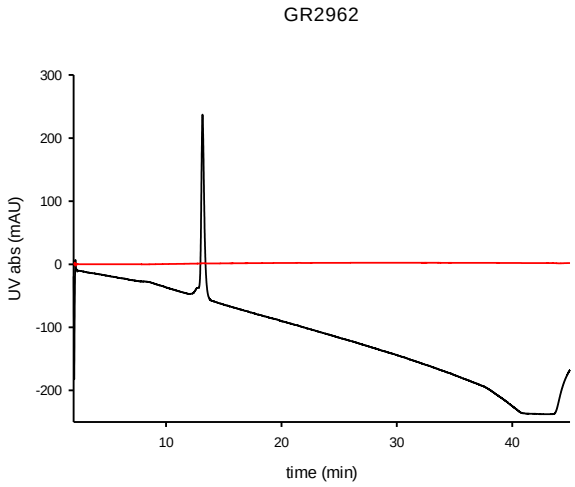


HMGN1_pS6 (Yield after ligation, desulfurisation and purification: ~ 60%, 3.3 mg for unlabelled; ~9%, 0.5 mg for ^{15}N-labelled)

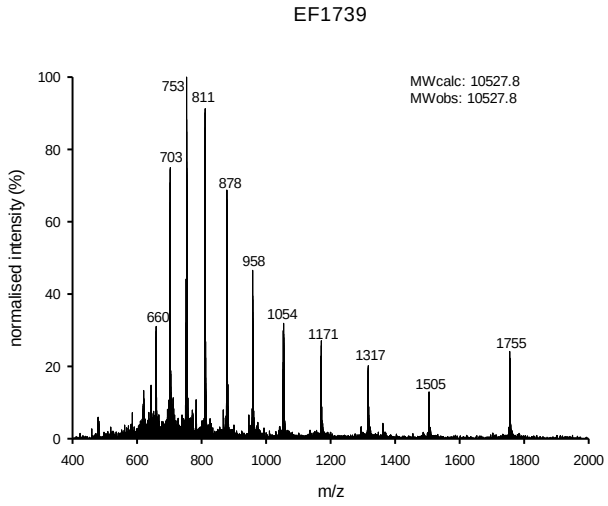


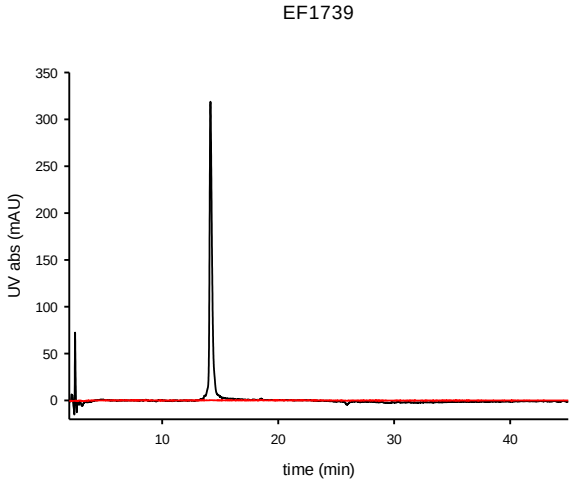
HMGN1_acK2_pS6 (Yield after ligation, desulfurisation and purification: ~ 33%, 1.8 mg for unlabelled; ~11%, 0.6 mg for ^{15}N-labelled)



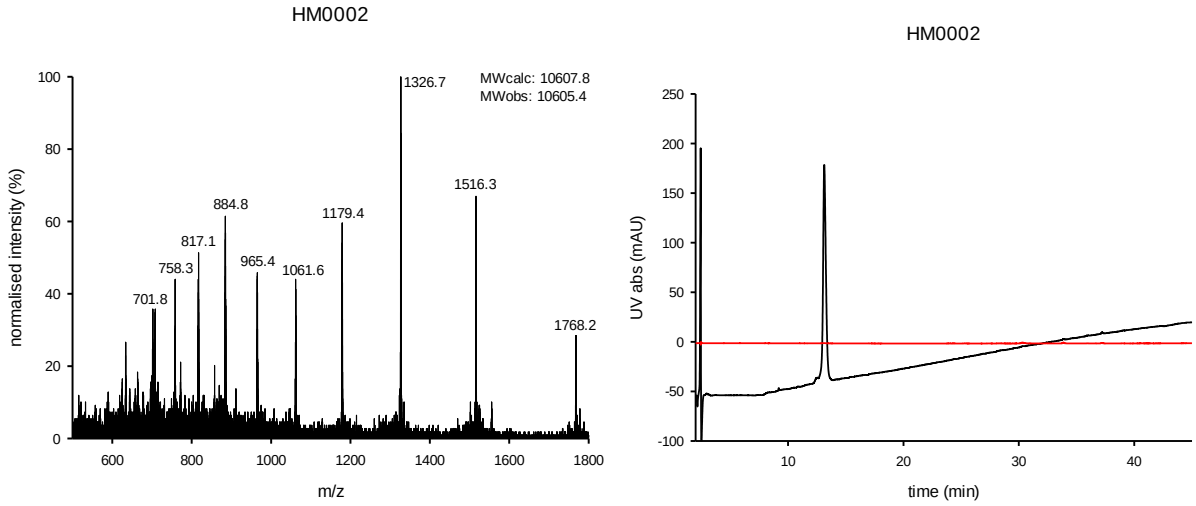


HMGN1_unmod_C (Yield after ligation, desulfurisation and purification: ~ 21%, 1.0 mg for unlabelled; ~42%, 0.3 mg for ^{15}N-labelled)

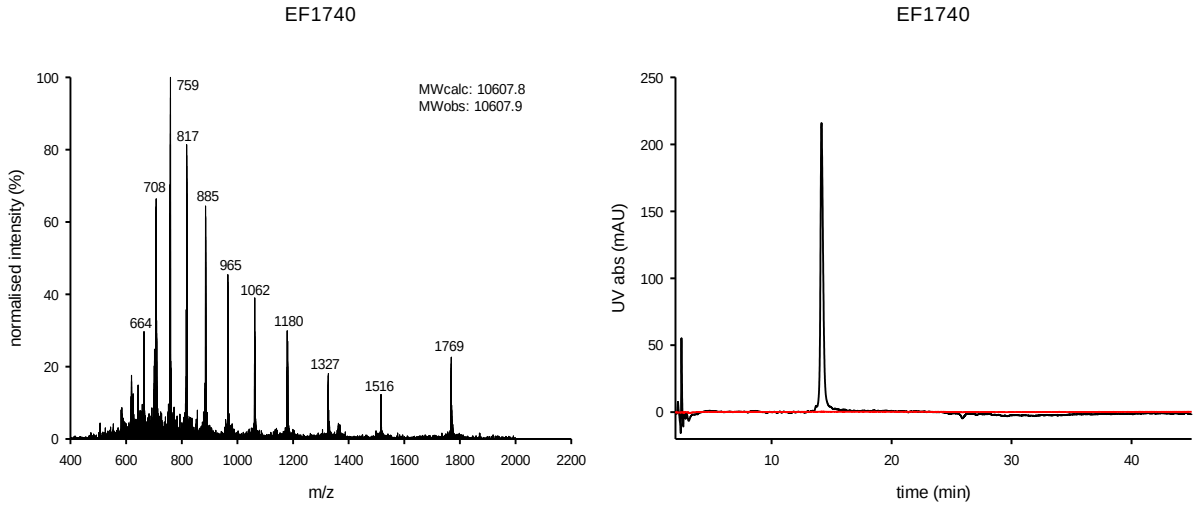




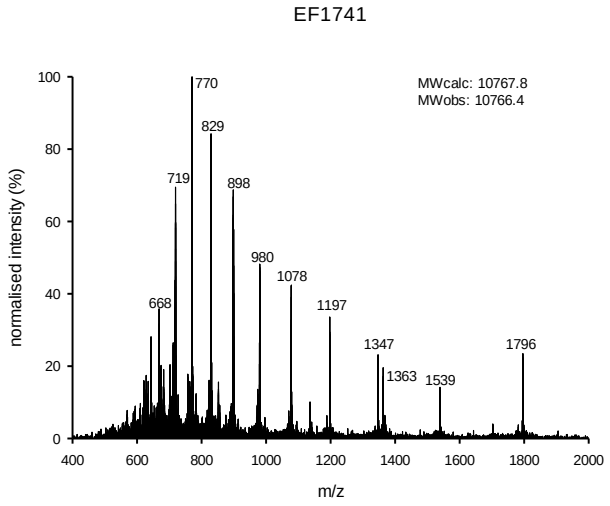
HMGN1_pS85 (Yield after ligation, desulfurisation and purification: ~ 12%, 0.6 mg for unlabelled)

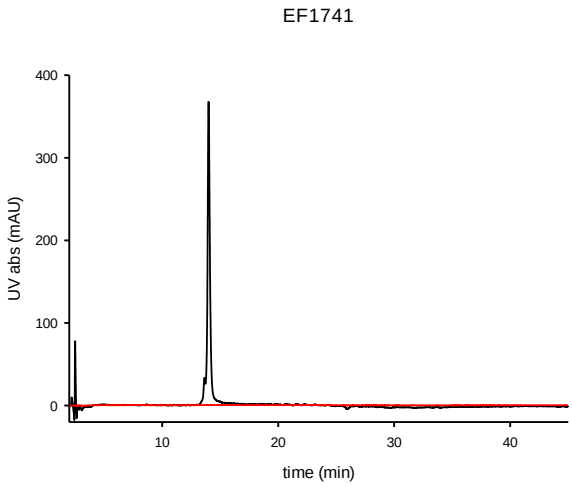


HMGN1_pS88 (Yield after ligation, desulfurisation and purification: ~ 21%, 1.0 mg for unlabelled; ~13%, 0.4 mg for ^{15}N-labelled)

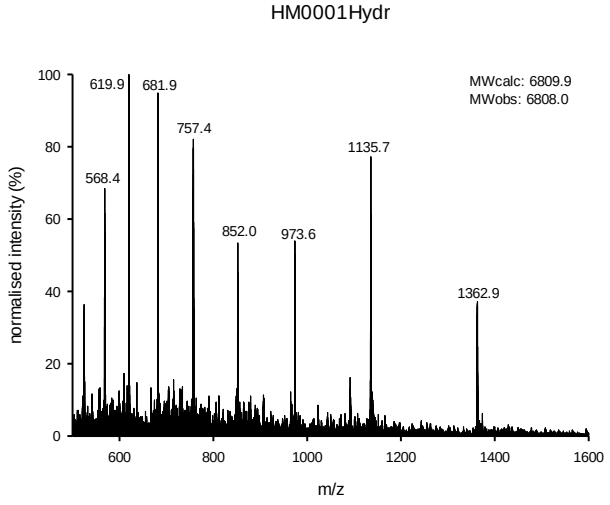


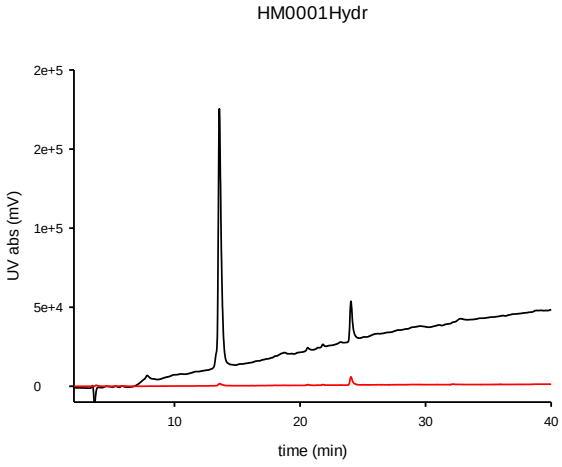
HMGN1_pS85,88,98 (Yield after ligation, desulfurisation and purification: ~ 29%, 1.4 mg for unlabelled; ~11%, 0.3 mg for ^{15}N-labelled)



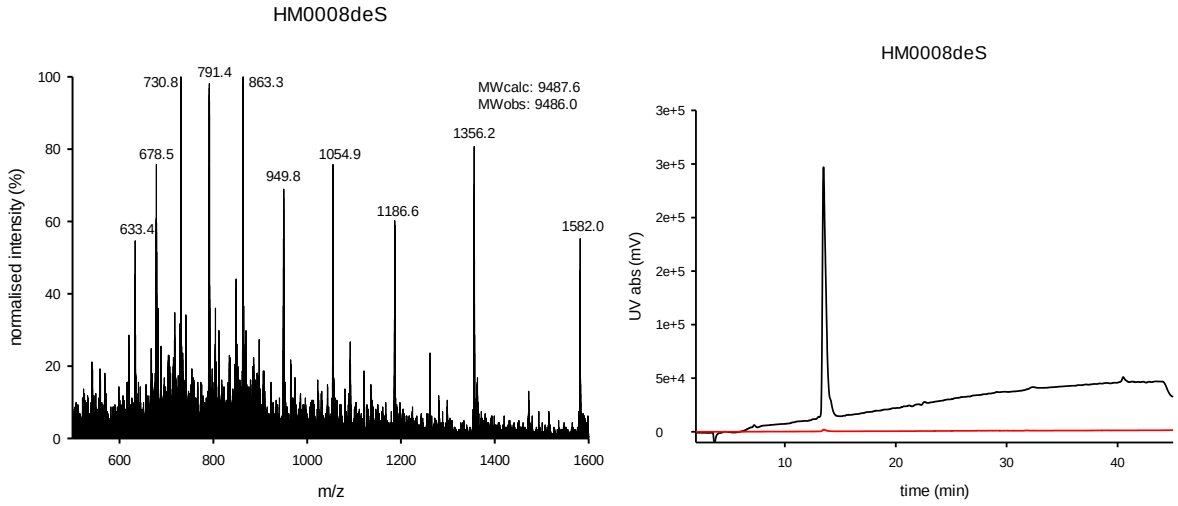


HMGN1_1-64 (Yield from expression, purification, intein cleavage and HPLC purification: 5.7 mg/L culture for unlabelled; 0.6 mg/L of rich medium, 250 mL minimal medium for ^{15}N-labelled ~ 16%, 0.5 mg after hydrolysis and repurification, unlabelled)

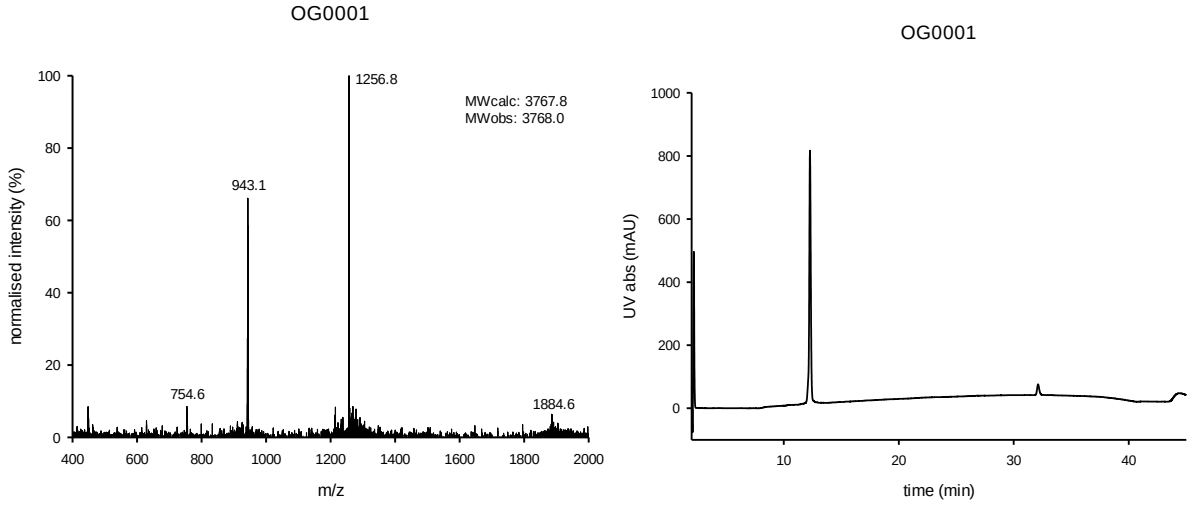




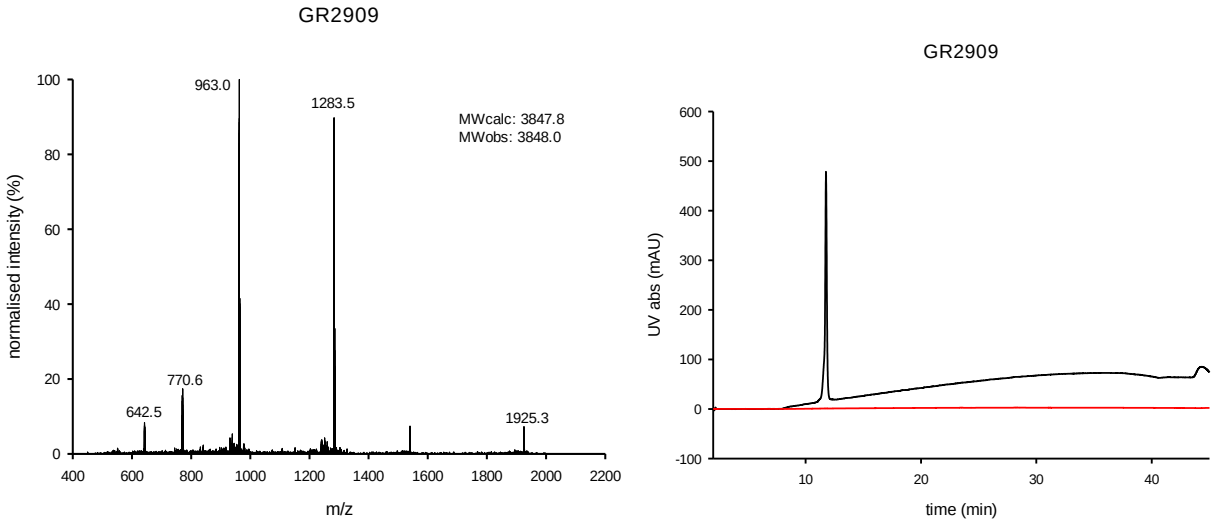
HMGN1_11-99 (Yield from expression, purification, TEV cleavage and HPLC purification: 6.7 mg/L culture for unlabelled; 3.8 mg/L of rich medium, 250 mL minimal medium for ^{15}N-labelled ~ 65%, 5.4 mg for desulfurisation and repurification, unlabelled)



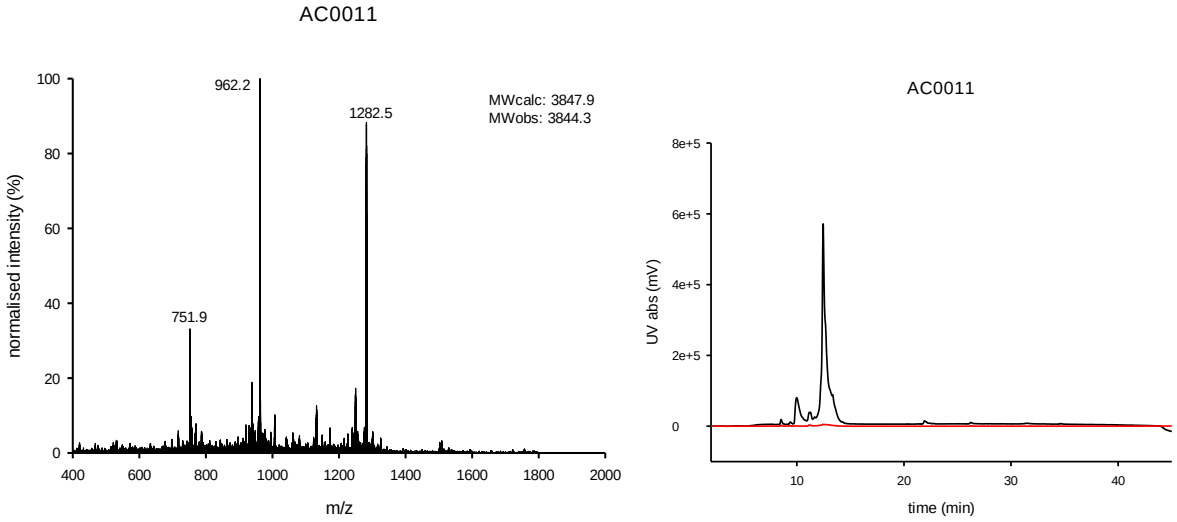
HMGN1_65-99 (Yield: 22% after HPLC purification, relative to calculated yield from synthesis scale of 0.05 mmol)



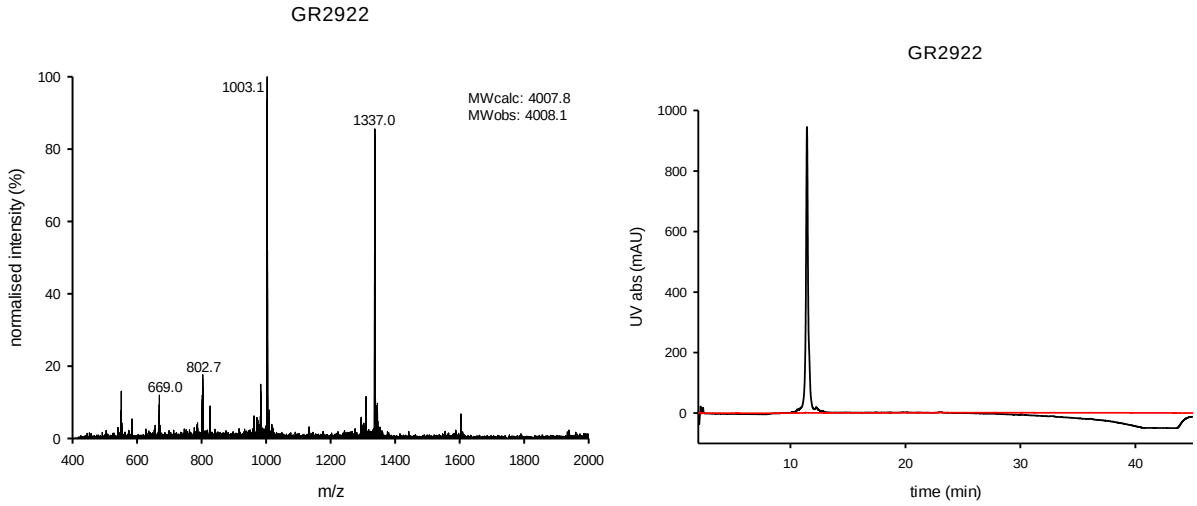
HMGN1_65-99_pS88 (Yield: 28% after HPLC purification, relative to calculated yield from synthesis scale of 0.1 mmol)



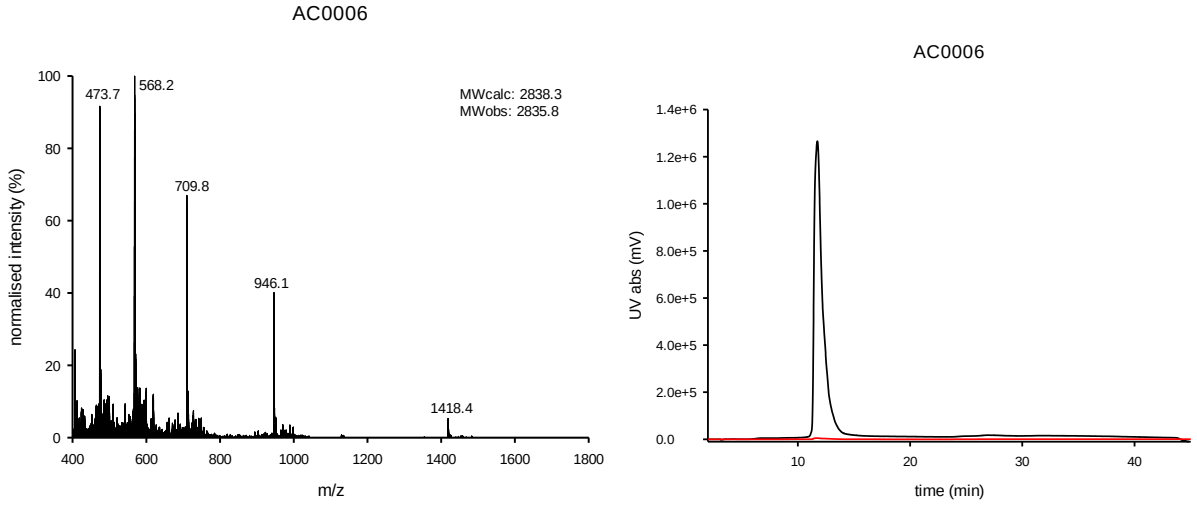
HMGN1_65-99_pS85 (Yield: 32% after HPLC purification, relative to calculated yield from synthesis scale of 0.1 mmol)



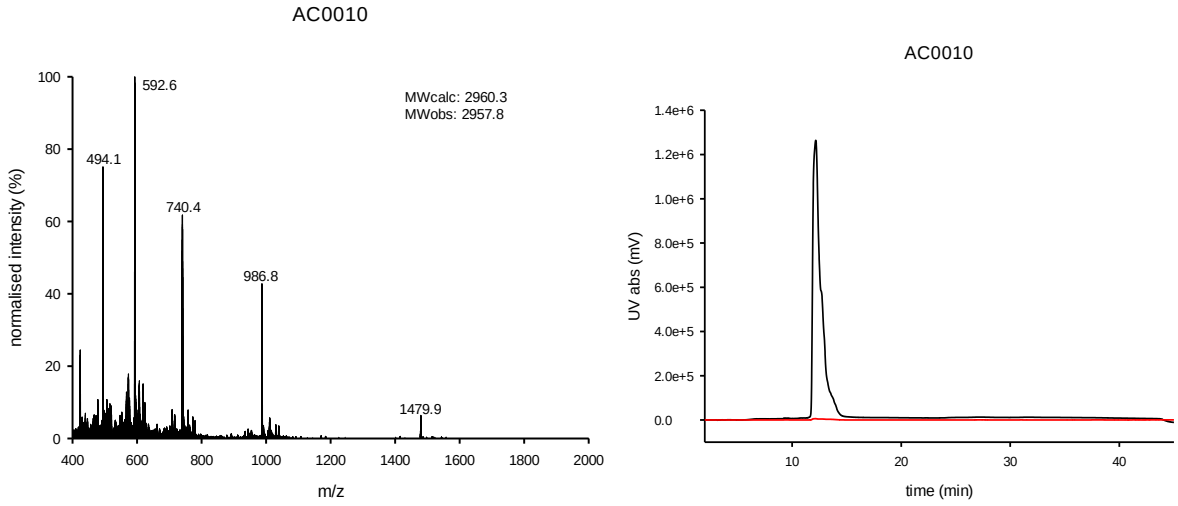
HMGN1_65-99_pS85,88,98 (Yield: 16% after HPLC purification, relative to calculated yield from synthesis scale of 0.1 mmol)



HMGN1_1-26 (Yield: 58% after HPLC purification, relative to calculated yield from synthesis scale of 0.05 mmol)

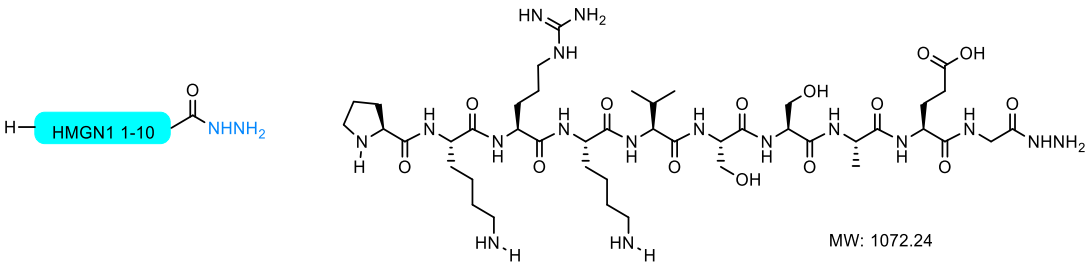


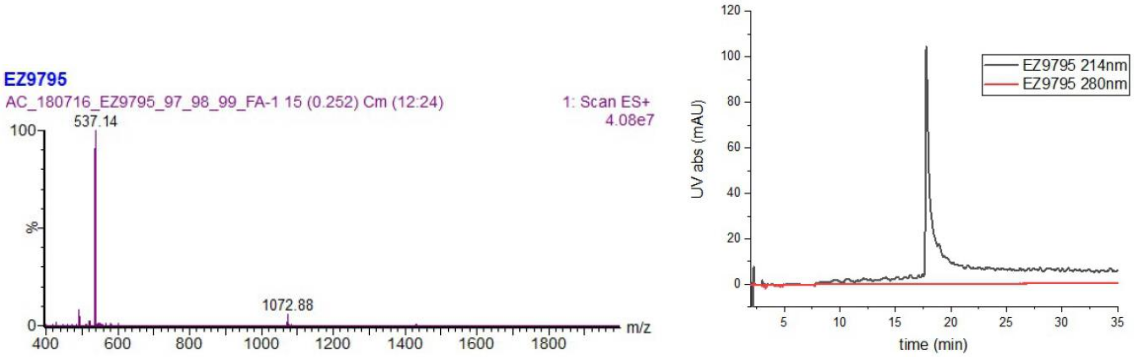
HMGN1_1-26_acK2,pS6 (Yield: 55% after HPLC purification, relative to calculated yield from synthesis scale of 0.05 mmol)



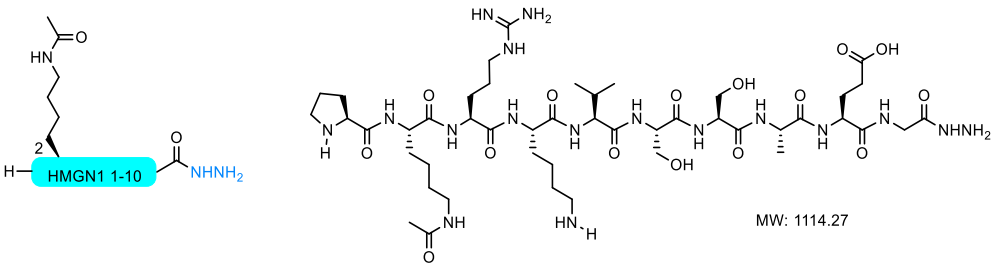
4. Chemical structures, mass spectra and analytical HPLC data for synthetic N-terminal HMGN1 hydrazide segments for ligation.

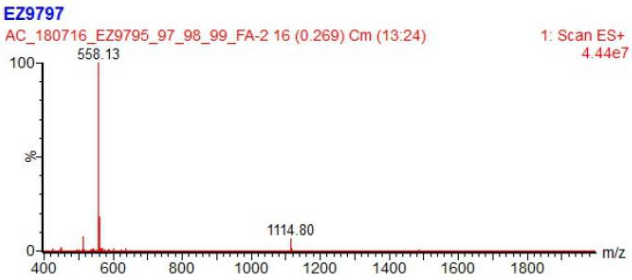
HMGN1_1-10-hydrazide (Yield: 54 % after HPLC purification, relative to calculated yield from synthesis scale of 0.05 mmol)

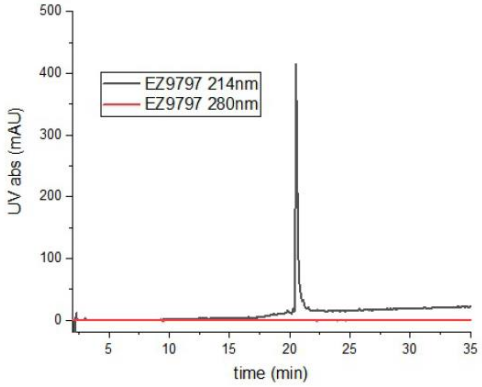




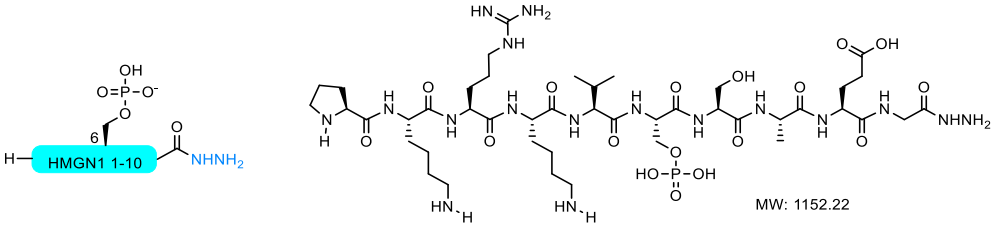
HMGN1_1-10_acK2-hydrazide (Yield: 37% after HPLC purification, relative to calculated yield from synthesis scale of 0.05 mmol)

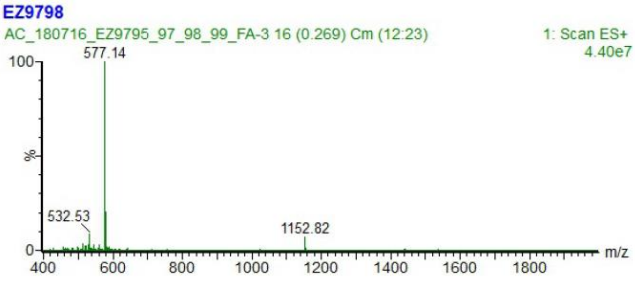


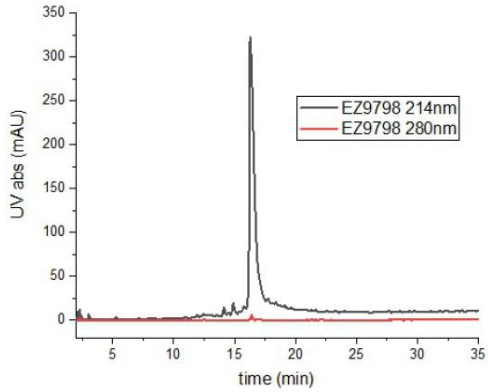




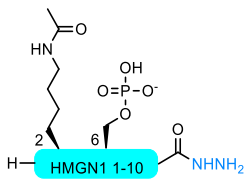
HMGN1_1-10_pS6-hydrazide (Yield: 30% after HPLC purification, relative to calculated yield from synthesis scale of 0.05 mmol)

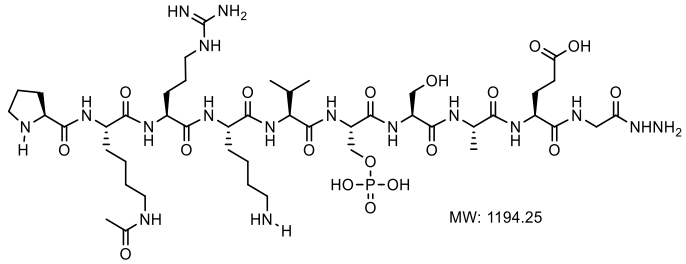


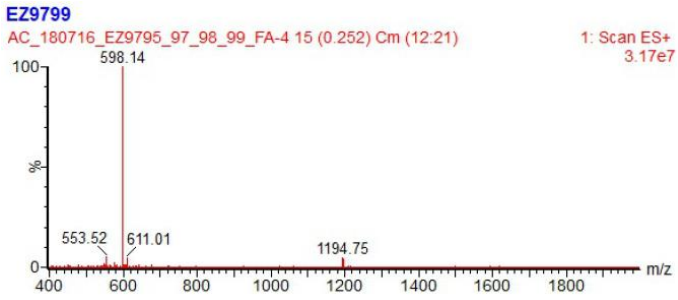


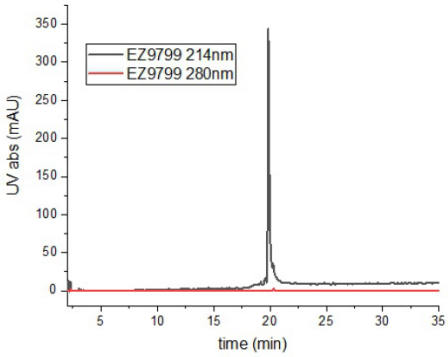


HMGN1_1-10_acK2,pS6-hydrazide (Yield: 17% after HPLC purification, relative to calculated yield from synthesis scale of 0.05 mmol)





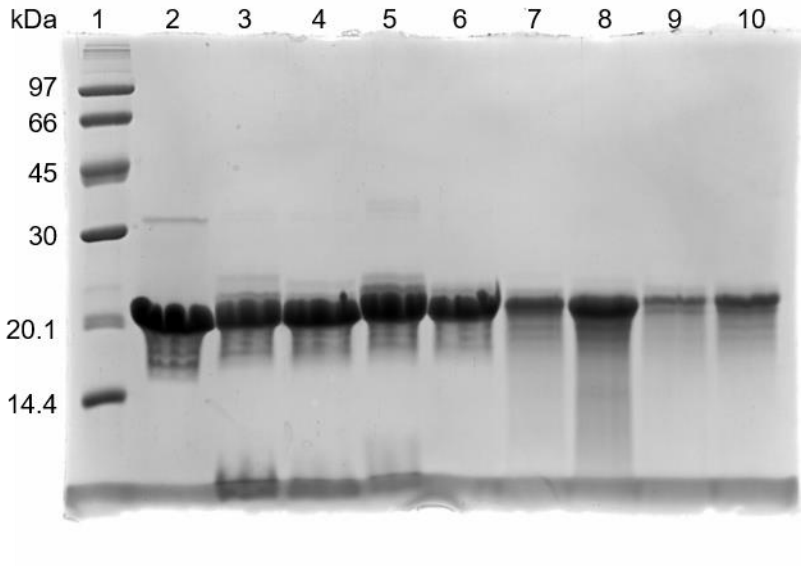




5. SDS PAGE gels

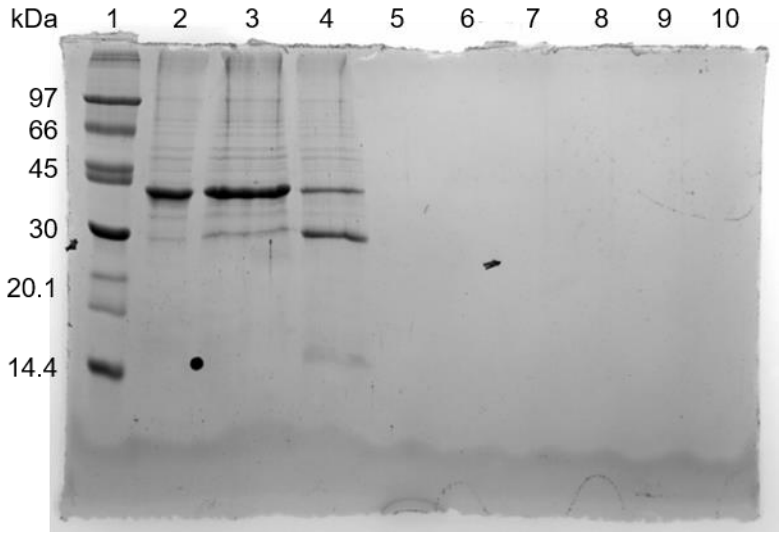
5.1 SDS PAGE gel (15%): Ligation and desulfurisation of HMGN1 bearing N-terminal PTMs

Lanes: 1 – marker; 2 – HMGN1_11-99_A11C before ligation; 3 – HMGN1_unmod_N ligation 4 h; 4 – HMGN1_acK2 ligation 4 h; 5 – HMGN1_pS6 ligation 4 h; 6 – HMGN1_acK2,pS6 ligation 4 h; 7 – HMGN1_unmod_N desulfurised; 8 – HMGN1_acK2 desulfurised; 9 – HMGN1_pS6 desulfurised; 10 - HMGN1_acK2,pS6 desulfurised.



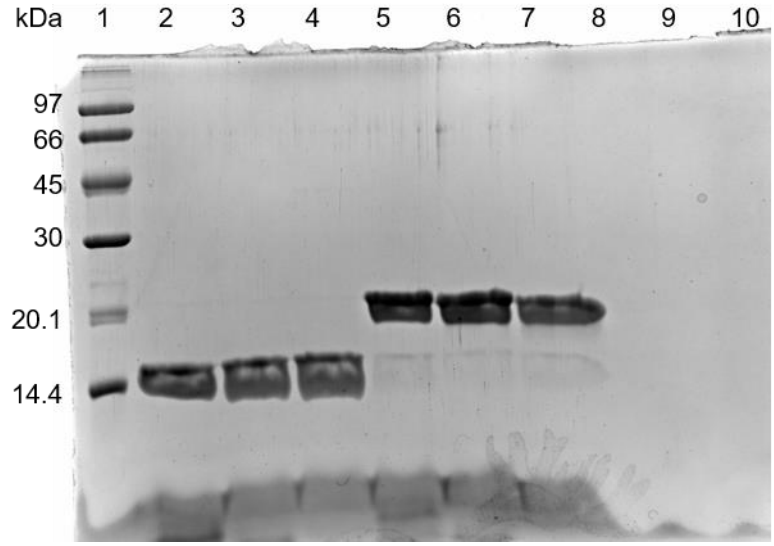
5.2 SDS PAGE gel (15%): Intein cleavage of HMGN1_1-65_MxeHis7CBD.

Lanes: 1 – marker; 2 – HMGN1_1-64_MxeHis7_CBD; 3 – intein cleavage with 250 mM MesNa 0 h; 4 – intein cleavage with 250 mM MesNa 16 h. HMGN1_1-64_MxeHis7_CBD = 35.9 kDa; Cleaved Mxe intein = 29.1 kDa; HMGN1_1-64 thioester = 6.9 kDa.



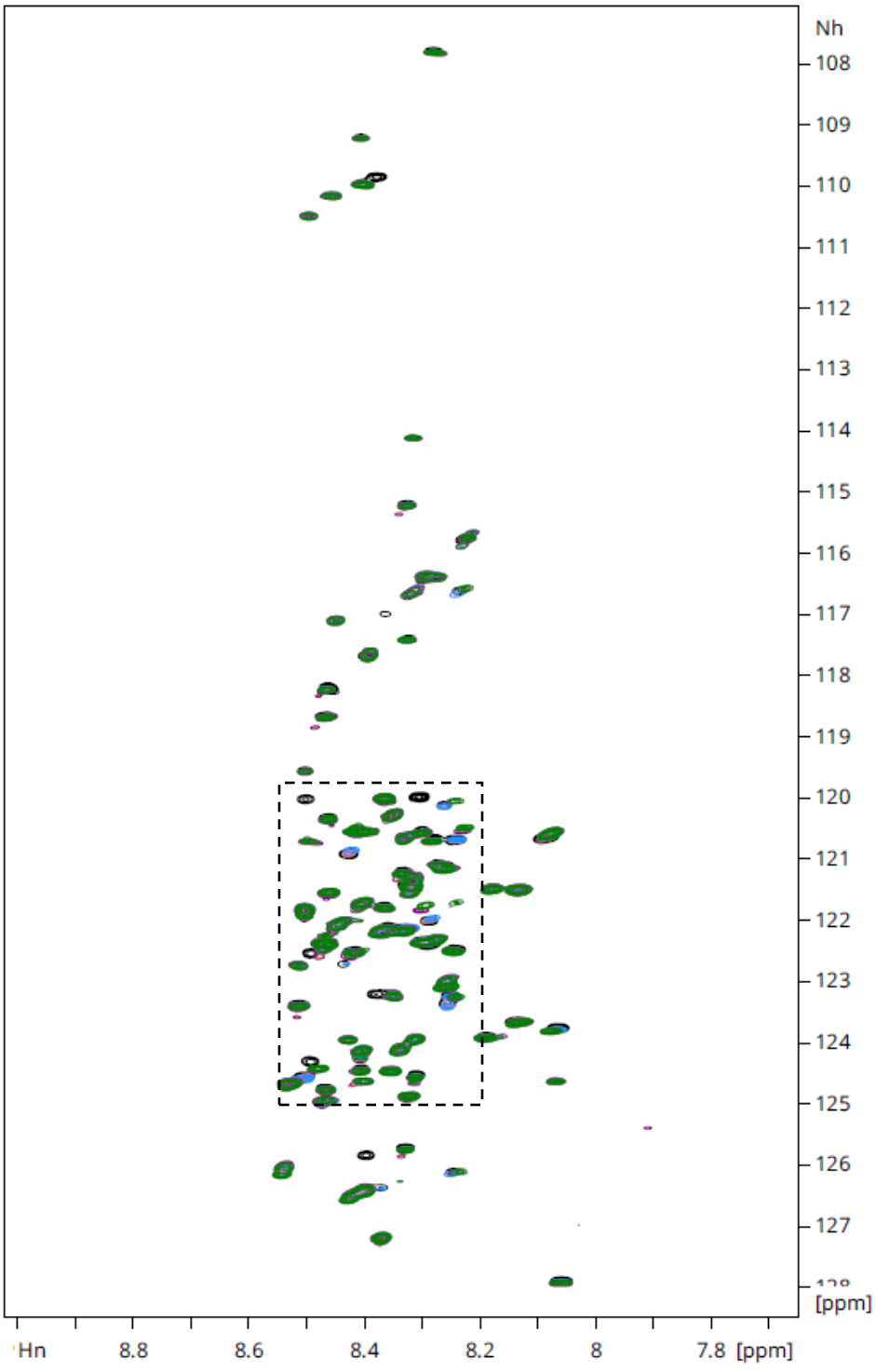
4.3 SDS PAGE gel (15%): Ligation of HMGN1 bearing C-terminal PTMs.

Lanes: 1 – marker; 2 – HMGN1_unmod_C ligation 0 h; 3 – HMGN1_pS88 ligation 0 h; 4 – HMGN1_pS85,88,98 ligation 0 h; 5 – HMGN1_unmod_C ligation 16 h; 6 – HMGN1_pS88 ligation 16 h; 7 – HMGN1_pS85,88,98 ligation 16 h.

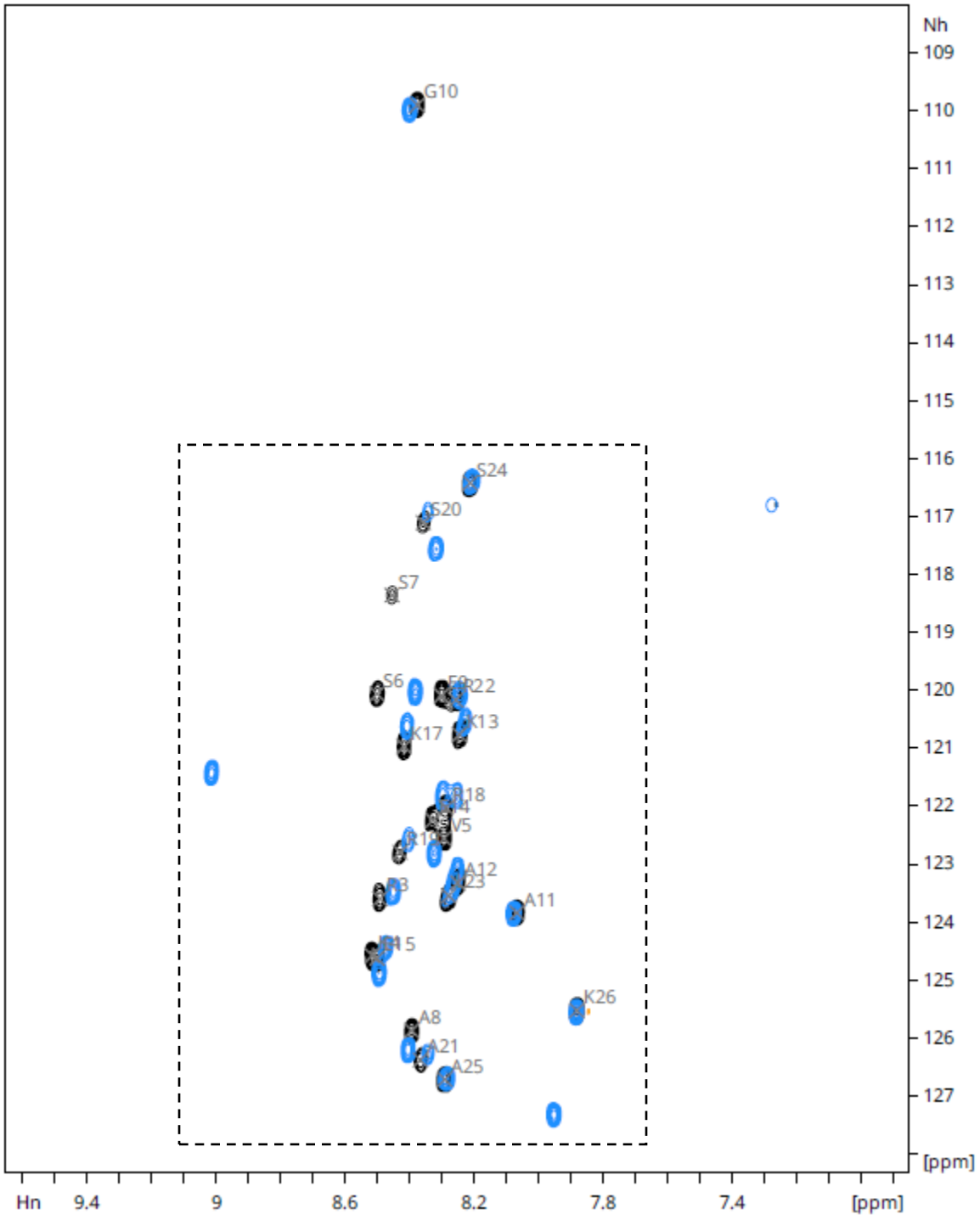


6. Full NMR spectra

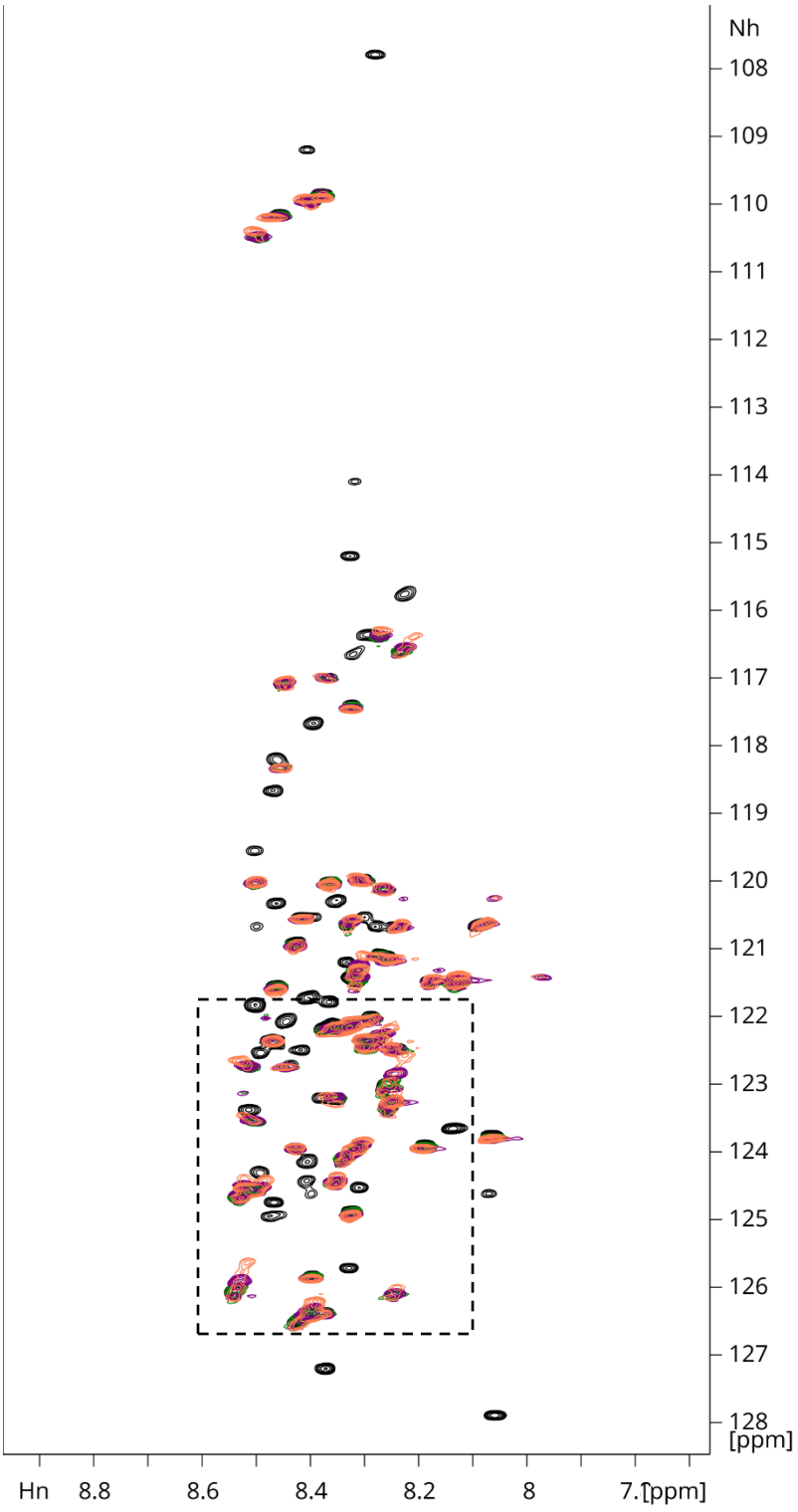
6.1 ^{15}N-HSQC spectra of HMGN1_1-99_15N13C (black), HMGN1_unmodN_15N (pink), HMGN1_acK2_15N (blue), HMGN1_pS6_15N (purple), and HMGN1_acK2,pS6_15N (green). Section shown in Figure 3b marked with a dashed box.



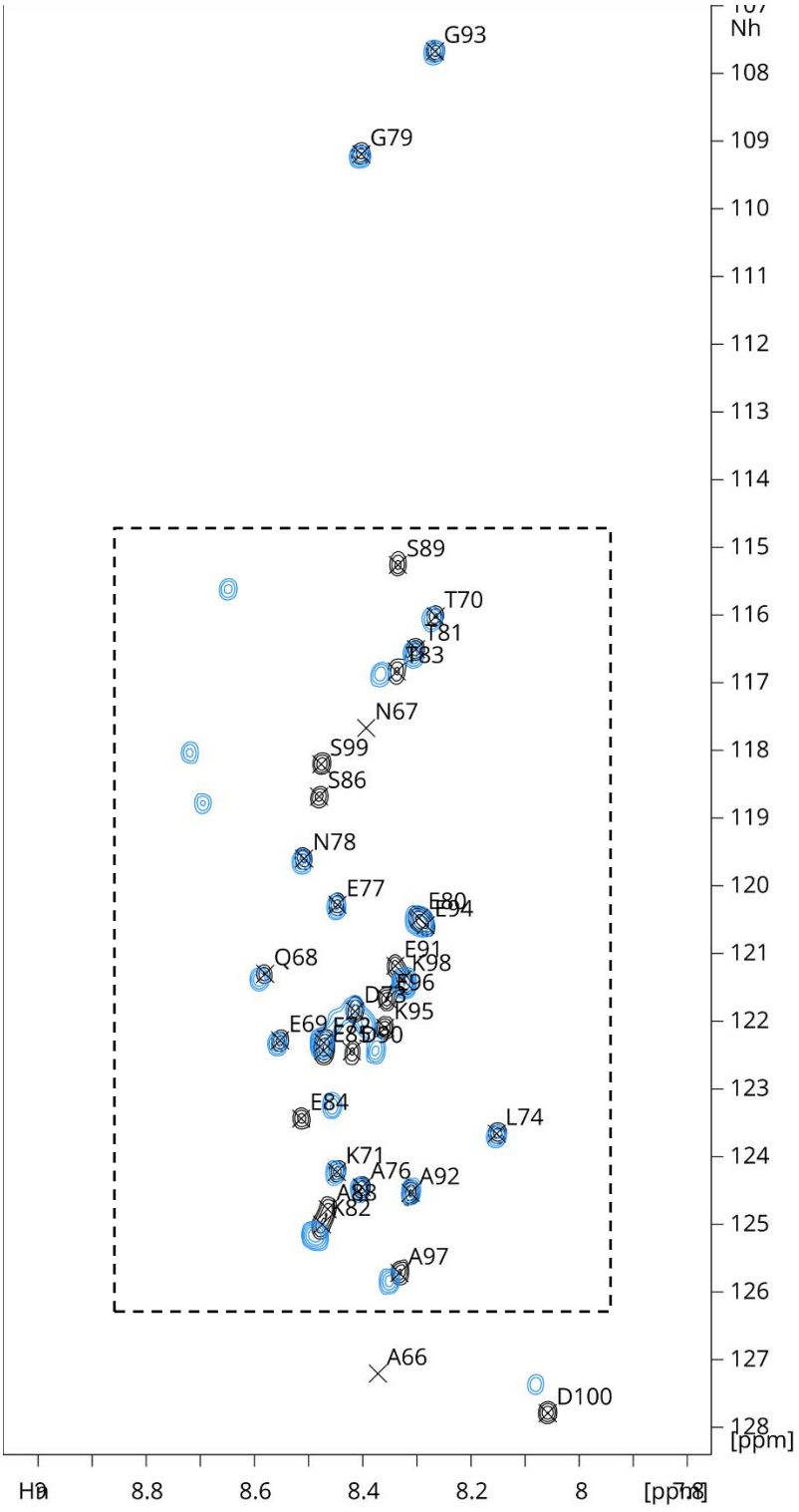
6.2 ^{15}N-HSQC spectra of HMGN1_1-26 (black) and HMGN1_1-26_acK2,pS6 (blue). Section shown in Figure 3c marked with a dashed box.



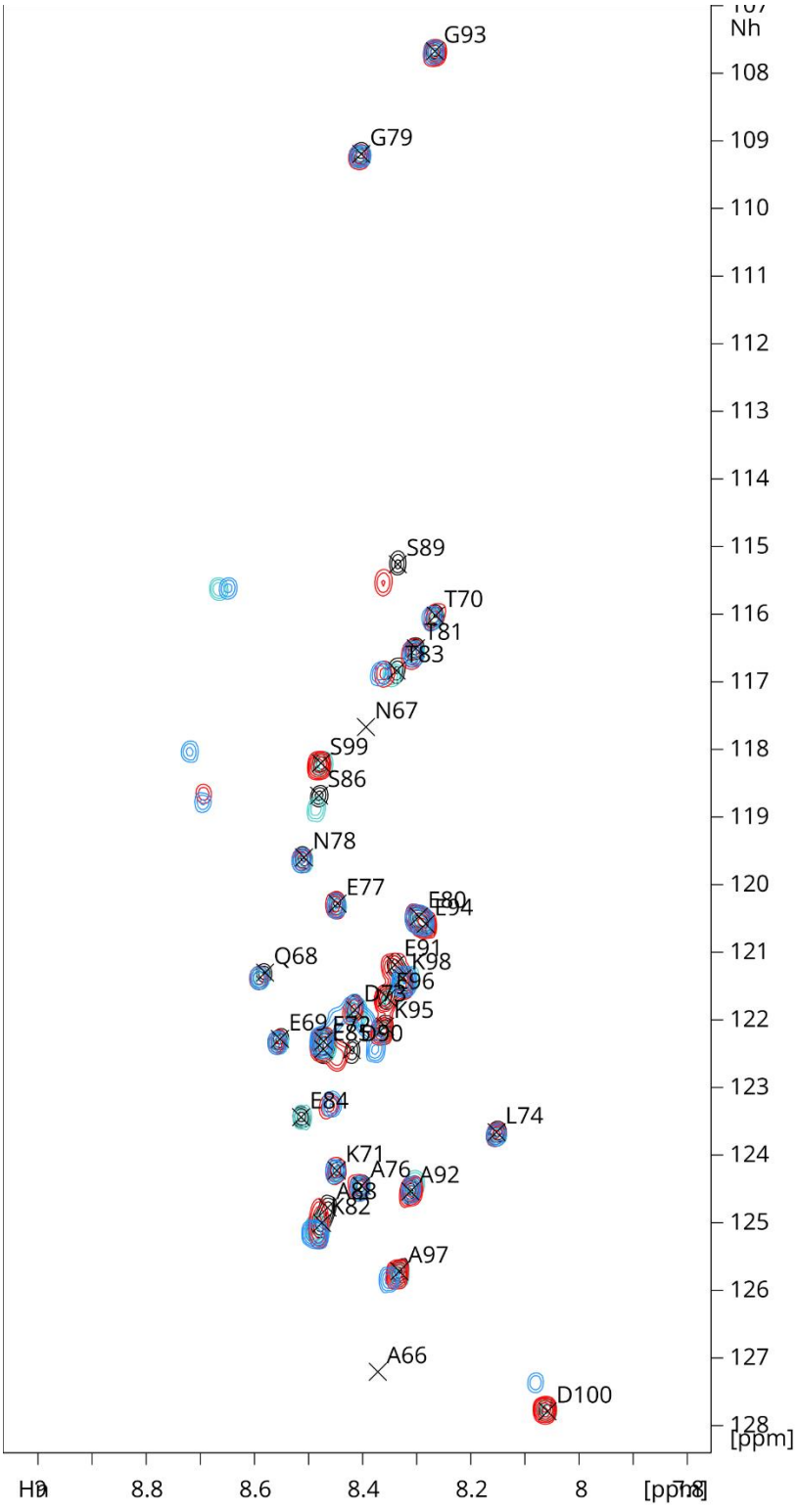
6.3 ^{15}N-HSQC spectra of HMGN1_1-99_15N13C (black), HMGN1_unmodC_15N (green), HMGN1_pS88_15N (purple), and HMGN1_pS85,88,98_15N (coral). Section shown in Figure 4b marked with a dashed box.



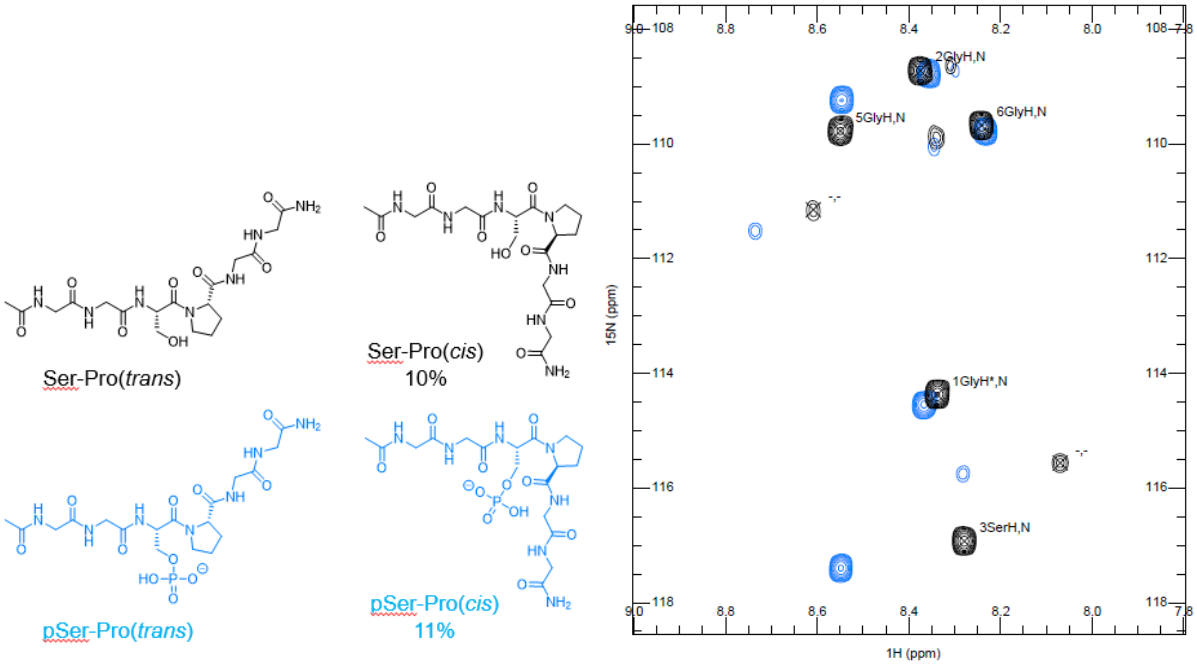
6.4 ^{15}N-HSQC spectra of HMGN1_65-99_A65C (black) and HMGN1_65-99_pS85,88,98_A65C (blue). Section shown in Figure 4c marked with a dashed box. Numbering is shown as (n+1) relative to the sequence numbering in Figure 4c.



6.5 ^{15}N-HSQC spectra of HMGN1_65-99_A65C (black), HMGN1_65-99_pS88_A65C (turquoise), HMGN1_65-99_pS85_A65C (red) and HMGN1_65-99_pS85,88,98_A65C (blue). Numbering is shown as (n+1) relative to the sequence numbering in Figure 4c.



6.6 Structures of short model peptides containing Ser-Pro and pSer-Pro residues in their *cis*- and *trans*-conformations. ^{15}N-HSQC spectra of Ser-Pro containing peptide (black) and pSer-Pro containing peptide (blue) showing shifts upon phosphorylation. Minor conformations corresponding to the *cis*-Pro population are visible as low intensity peaks.



7. Heteronuclear ^{1}H-^{15}N NOE ratios

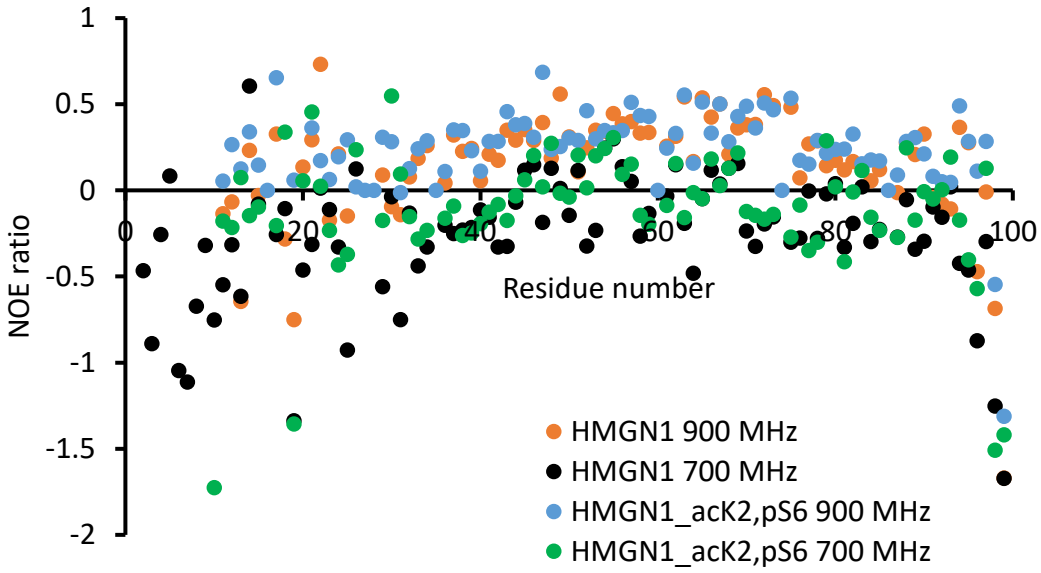


Figure S7a. ^{1}H-^{15}N heteronuclear NOE ratios for unmodified HMGN1 (HMGN1_S0_15N) and HMGN1_acK2,pS6_15N at 700 MHz and 900 MHz.

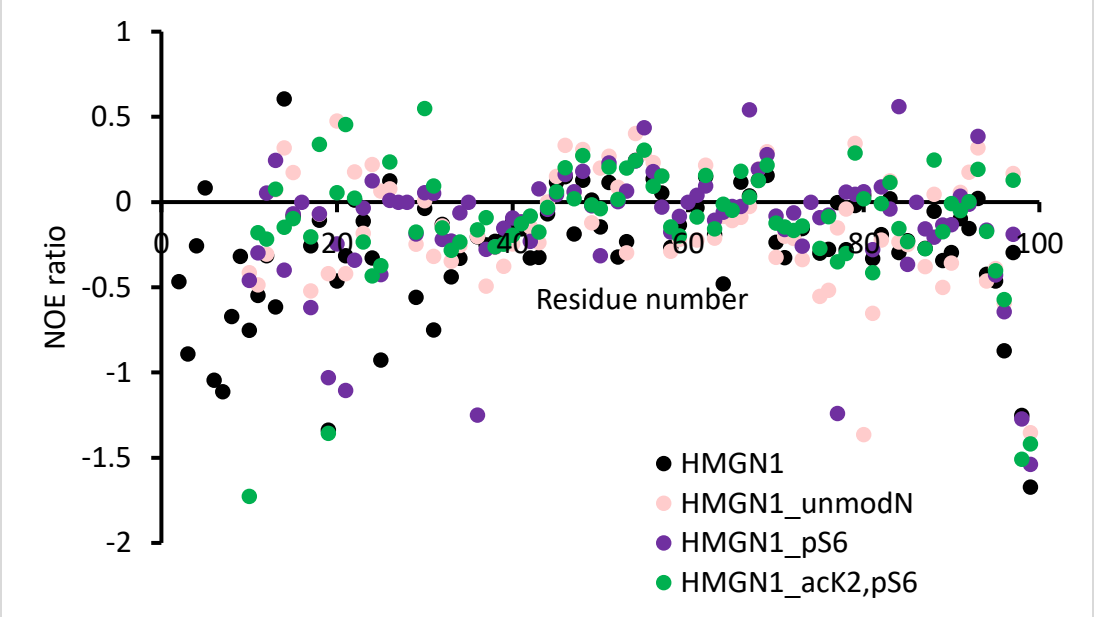


Figure S7b. ^{1}H-^{15}N heteronuclear NOE ratios for unmodified HMGN1 (HMGN1_S0_15N), HMGN1_unmodN_15N, HMGN1_pS6_15N and HMGN1_acK2,pS6_15N at 700 MHz.

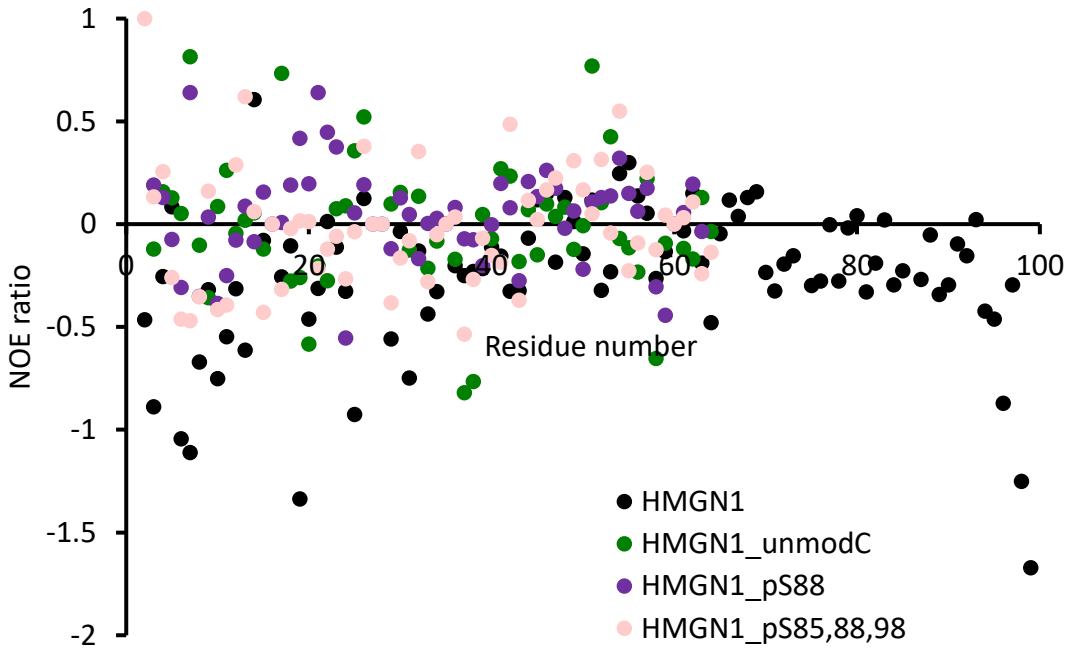


Figure S7c. ^{1}H-^{15}N heteronuclear NOE ratios for unmodified HMGN1 (HMGN1_S0_15N), HMGN1_unmodC_15N, HMGN1_pS88_15N and HMGN1_acK2,pS6_15N at 700 MHz.

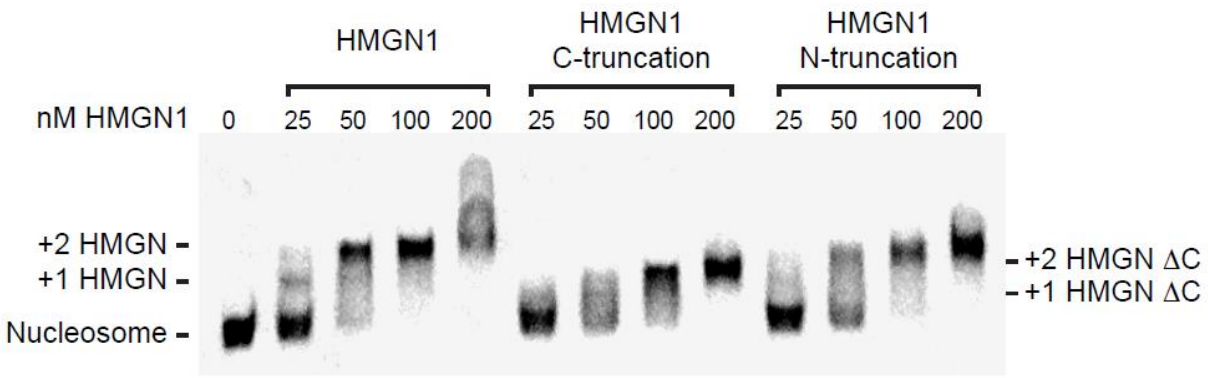
8. Backbone chemical shifts (ppm) of unmodified HMGN1

Residue	AA	NH	HN	CA	CB	CO
0	S					
1	P					
2	K	122.5	8.49	56.4	33.1	176.6
3	R	123.2	8.38	56.0	31.2	176.0
4	K	124.3	8.49	56.4	33.1	176.5
5	V	122.4	8.30	62.1	33.0	176.2
6	S	120.0	8.50	58.2	64.1	174.7
7	S	118.3	8.46	58.5	63.9	174.5
8	A	125.8	8.40	52.9	19.1	178.0
9	E	120.0	8.31	57.0	30.2	177.3
10	G	109.8	8.38	45.4		174.0
11	A	123.8	8.07	52.5	19.3	177.7
12	A	123.2	8.25	52.5	19.1	177.8
13	K	120.7	8.24	56.3	33.2	176.5
14	E	122.1	8.33	56.1	30.6	176.2
15	E	124.5	8.51	54.7	29.7	174.7
16	P					
17	K	120.9	8.43	56.6	32.8	177.0
18	R	122.0	8.29	56.3	30.8	176.5
19	R	122.7	8.44	56.5	30.8	176.5
20	S	117.0	8.37	58.5	63.8	174.5
21	A	126.4	8.37	52.8	19.3	177.8
22	R	120.1	8.26	56.4	30.7	176.5
23	L	123.4	8.26	55.2	42.3	177.4
24	S	116.6	8.24	58.2	63.9	174.0
25	A	126.1	8.25	52.3	19.4	177.3
26	K	122.3	8.28	54.1	32.5	174.1
27	P					
28	P					
29	A	124.4	8.36	52.3	19.3	177.7
30	K	121.4	8.32	56.2	33.2	176.4
31	V	123.0	8.26	62.2	33.0	176.0
32	E	126.1	8.54	56.1	30.6	175.8
33	A	126.5	8.42	52.3	19.3	177.4
34	K	122.1	8.34	54.1	32.5	174.5
35	P					
36	K	122.4	8.48	56.4	33.3	176.7
37	K	123.2	8.35	56.2	33.2	176.1
38	A	126.4	8.40	52.3	19.3	177.3

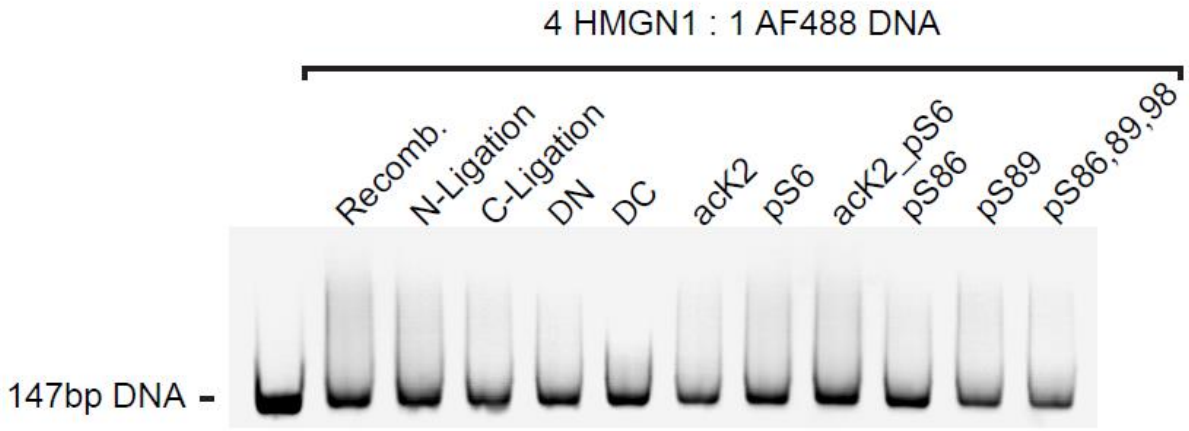
39	A	124.1	8.34	52.2	19.4	177.5
40	A	123.9	8.32	52.5	19.3	177.9
41	K	120.6	8.33	56.4	33.1	176.4
42	D	121.4	8.32	54.3	41.4	176.4
43	K	122.2	8.38	56.5	32.9	177.1
44	S	117.1	8.45	59.1	63.8	175.0
45	S	117.4	8.33	58.7	63.8	174.4
46	D	122.5	8.25	54.6	41.1	176.3
47	K	121.5	8.18	56.4	32.9	176.7
48	K	122.4	8.29	56.4	32.9	176.7
49	V	121.5	8.14	62.4	32.8	176.2
50	Q	124.7	8.53	55.8	29.7	176.1
51	T	116.4	8.28	62.0	69.9	174.5
52	K	124.0	8.43	56.6	33.1	177.0
53	G	110.2	8.46	45.2		174.0
54	K	122.5	8.27	56.3	33.1	176.9
55	R	122.7	8.51	56.4	30.8	176.9
56	G	110.5	8.50	45.2		173.9
57	A	123.9	8.19	52.5	19.5	177.9
58	K	120.5	8.41	56.5	33.1	177.2
59	G	109.9	8.40	45.3		174.2
60	K					
61	Q	121.5	8.46	56.1	29.4	175.8
62	A	124.9	8.33	52.7	19.3	177.7
63	E	120.0	8.36	56.7	30.3	176.6
64	V	120.6	8.09	62.2	32.9	175.9
65	A	127.2	8.37	52.7	19.3	177.5
66	N	117.7	8.40	53.5	38.8	175.2
67	Q	120.3	8.35	56.1	29.5	176.0
68	E	121.8	8.50	56.8	30.3	176.7
69	T	115.8	8.23	61.8	69.9	174.4
70	K	124.1	8.41	56.3	33.1	176.4
71	E	122.1	8.45	56.4	30.5	175.9
72	D	121.7	8.40	54.2	41.2	175.7
73	L	123.7	8.14	53.0	41.9	175.1
74	P					
75	A	124.4	8.41	52.5	19.3	178.1
76	E	120.3	8.47	56.6	30.2	176.4
77	N	119.5	8.51	53.4	39.1	175.8
78	G	109.2	8.41	45.5		174.2
79	E	120.5	8.30	56.6	30.5	176.8
80	T	116.4	8.29	62.1	69.8	174.4

81	K	124.9	8.47	56.1	33.2	176.6
82	T	116.6	8.32	61.9	69.9	174.5
83	E	123.4	8.52	56.4	30.5	176.2
84	E	122.4	8.47	56.4	30.5	176.3
85	S	118.7	8.47	56.5	63.3	179.5
86	P					
87	A	124.7	8.47	52.5	19.3	177.9
88	S	115.2	8.33	58.3	64.0	174.5
89	D	122.5	8.42	54.4	41.1	176.5
90	E	121.2	8.33	56.9	30.1	176.6
91	A	124.5	8.31	53.0	19.2	178.4
92	G	107.8	8.28	45.4		174.3
93	E	120.7	8.28	56.7	30.3	176.8
94	K	122.1	8.37	56.5	33.0	176.7
95	E	121.8	8.37	56.5	30.3	176.1
96	A	125.7	8.33	52.4	19.2	177.5
97	K	121.4	8.33	56.2	33.4	176.5
98	S	118.2	8.46	58.2	64.2	173.4
99	D	127.9	8.06	55.8	42.1	173.8

9. Nucleosome and DNA binding assay gels



S9a: Electrophoretic mobility shift assays showing mononucleosome binding of full length HMGN1 and truncated versions used for ligation reactions. Position of nucleosomes bound by one HMGN1 molecule (+1 HMGN), and two HMGN1 molecules (+2 HMGN) are shown in relation to the unbound nucleosome. Binding reactions were separated on 5% TBE-acrylamide gels and then scanned for fluorescence.



S9b: Electrophoretic mobility shift assays showing binding of full length, truncated and modified HMGN1 variants to 147 bp AlexaFluor488 labelled DNA containing the 601 nucleosome positioning sequence. Binding reactions were separated on 5% TBE-acrylamide gels and then scanned for fluorescence.