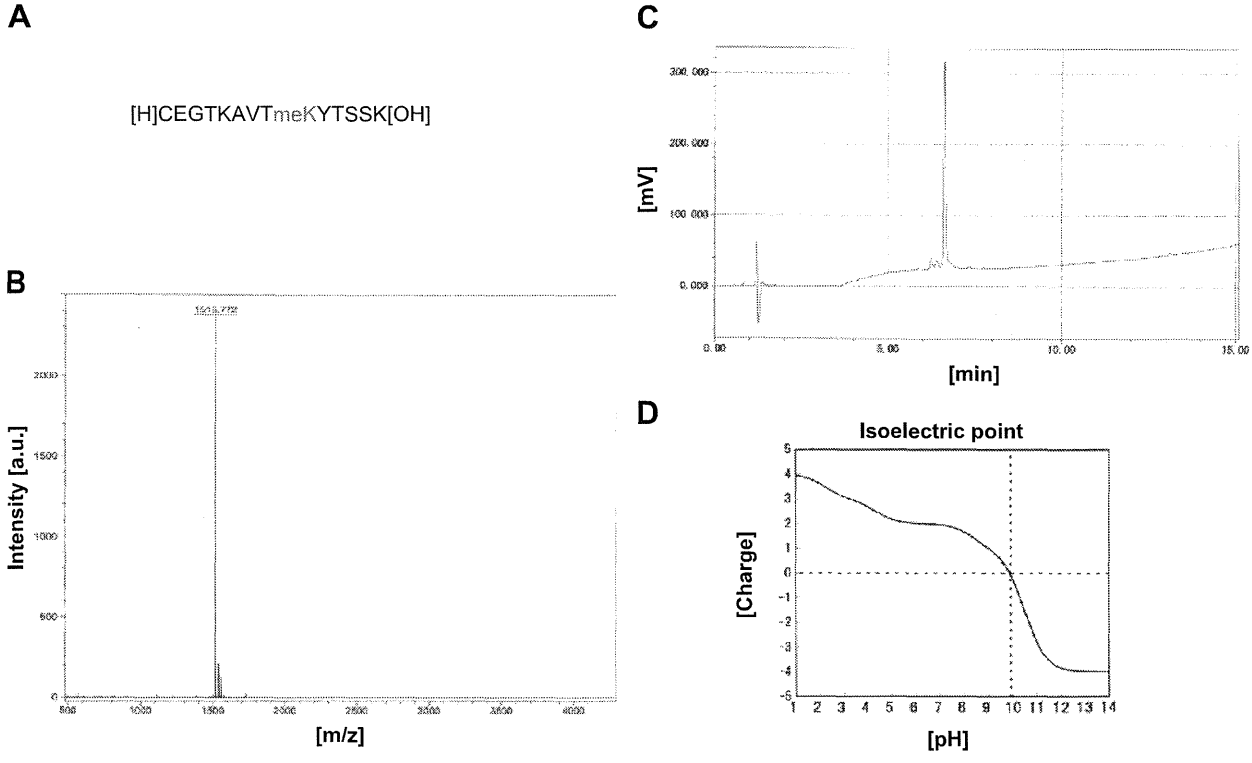


Figure W2. Quality check of peptides used as immunogens. (A) Amino acid sequence of the peptide. Ninth lysine residue is monomethylated. (B) Mass spectrum of the peptide. Bruker Daltonics flexAnalysis software (Billerica, MA) was used for the analysis. (C) Chromatogram of the peptide analyzed by the MultiStation LC-8020 system. (D) Isoelectric point figure of the peptide.

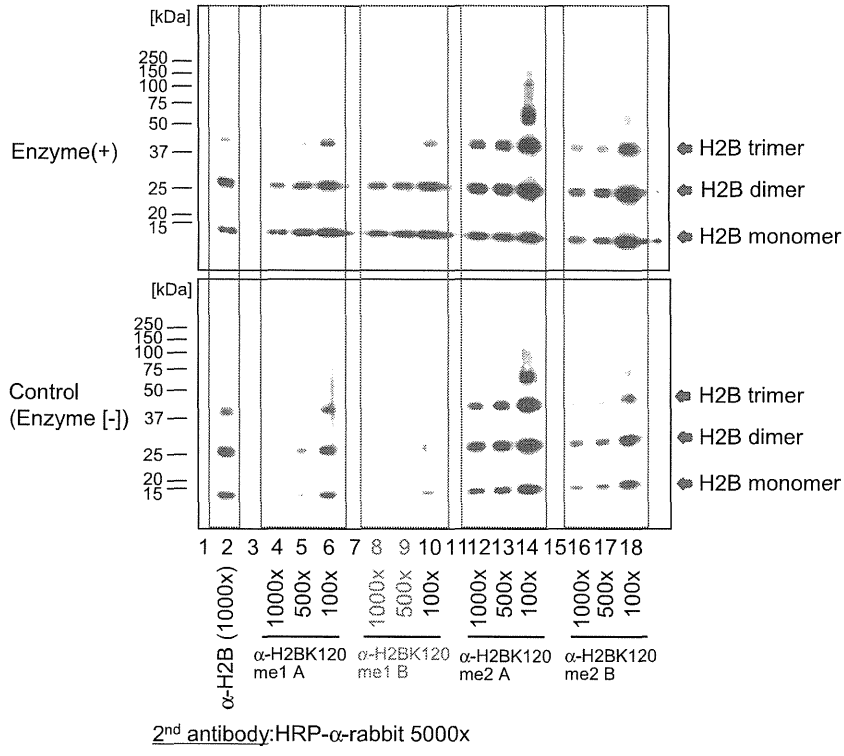


Figure W3. Slot blot analysis for the quality check of antibodies. *In vitro* methyltransferase assays were performed using recombinant EZH2 enzyme complex (No. 51004; BPS Bioscience) as an enzyme source. Samples were fractionated with SDS-PAGE and transferred to the nitrocellulose membranes. Slot blot analysis was conducted using anti-H2BK120 antibodies. An anti–histone H2B antibody was used as an internal control.

Table W5. Information of Certificated Cell Lines.

Name	Origin	Certification Institution	Tested Method	DNA Profile or Characteristics
CCD-18Co	human colonic fibroblast	ATCC	STR	Amelogenin: X CSF1PO: 8 D13S317: 12 D16S539: 12,13 D5S818: 12 D7S820: 8 THO1: 6,7 TPOX: 8,11 vWA: 15,17
HFL1	human fetal lung fibroblast	ATCC	STR	Amelogenin: X,Y CSF1PO: 10,12 D13S317: 11,12 D16S539: 9,11 D5S818: 12 D7S820: 9,10 THO1: 7,9 TPOX: 6,9 vWA: 17
SBC5	human small lung cancer	JCRB	STR	Amelogenin: X,Y TPOX: 9,12 CSF1PO: 10 D5S818: 10,11 D13S317: 8,10 D7S820: 8,11 D16S539: 12 vWA: 14,18 TH01: 6
RERF-LC-AI	human lung squamous cell carcinoma	JCRB	STR	Amelogenin: X,Y TPOX: 10,11 CSF1PO: 12 D5S818: 10.3 D13S317: 10 D7S820: 10,11 D16S539: 11 VWA: 17 TH01: 7,9
HCT116	human colorectal carcinoma	ATCC	STR	Amelogenin: X,Y CSF1PO: 7,10 D13S317: 10,12 D16S539: 11,13 D5S818: 10,11 D7S820: 11,12 THO1: 8,9 TPOX: 8,9 vWA: 17,22
SW480	human colorectal carcinoma	ATCC	STR	Amelogenin: X CSF1PO: 13,14 D13S317: 12 D16S539: 13 D5S818: 13 D7S820: 8 THO1: 8 TPOX: 11 vWA: 16
Alexander	human malignant liver cancer	JCRB	STR	D5S818: 12 D13S317: 11,12 D7S820: 9,11 D16S539: 13 vWA: 15,16 TH01: 7,8 Amelogenin: X TPOX: 8 CSF1PO: 10
RT4	human urinary bladder cancer	ATCC	STR	Amelogenin: X,Y CSF1PO: 10,12 D13S317: 8 D16S539: 9 D5S818: 11,12 D7S820: 9,12 THO1: 9,9.3 TPOX: 8,11 vWA: 14,17
MCF7	human breast adenocarcinoma	ATCC	STR	Amelogenin: X CSF1PO: 10 D13S317: 11 D16S539: 11,12 D5S818: 11,12 D7S820: 8,9 THO1: 6 TPOX: 9,12 vWA: 14,15
HeLa	human cervix carcinoma	ATCC	STR	Amelogenin: X,Y CSF1PO: 11,12 D13S317: 11,14 D16S539: 9,11 D5S818: 11,12 D7S820: 10,11 THO1: 8 TPOX: 8 vWA: 15

ATCC indicates American Type Culture Collection; JCRB, Japanese Collection of Research Bioresources.

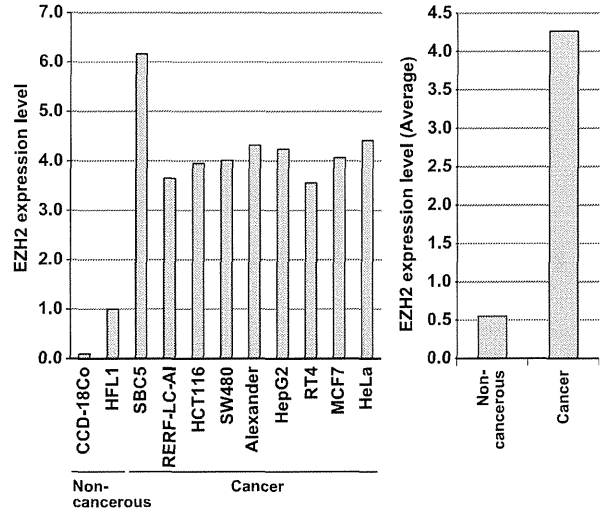


Figure W4. Quantitative analysis of EZH2 expression in various types of cell lines. X-ray films were scanned with GS-800 calibrated densitometer (Bio-Rad Laboratories). The intensity of each EZH2 signal was normalized by the corresponding ACTB signal and averaged.

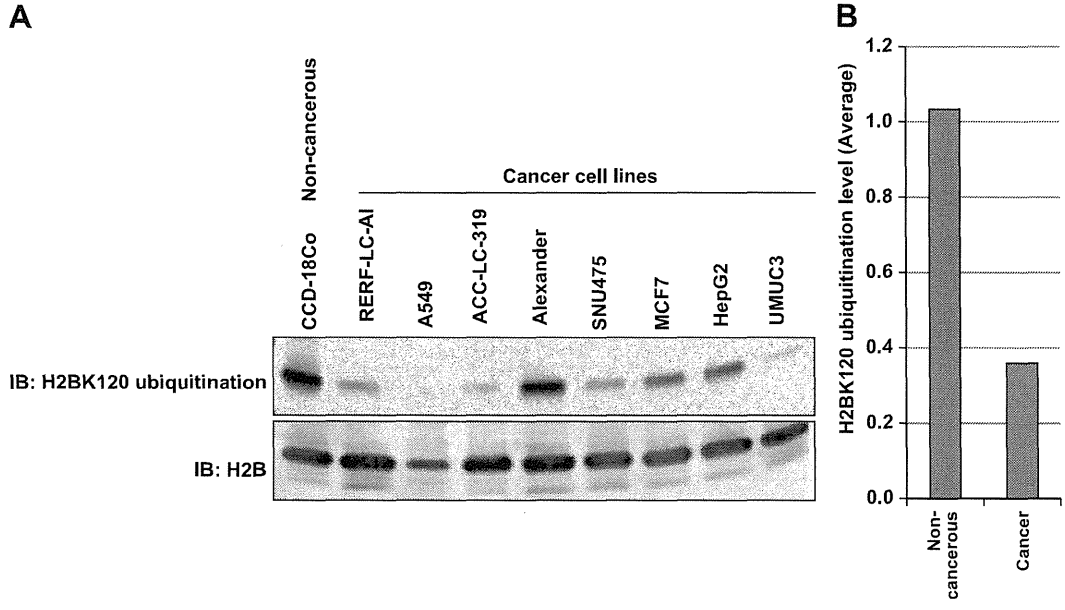


Figure W5. H2BK120 ubiquitination was decreased in cancer cells. (A) Validation of H2BK120 ubiquitination status in various cell lines. Lysates from a noncancerous cell line (CCD-18Co) and cancer cell lines (RERF-LC-AI, A549, ACC-LC-319, Alexander, SNU475, MCF7, HepG2, and UMUC3) were immunoblotted with anti–Ubiquityl-Histone H2BK120 antibody (No. 5546; Cell Signaling Technology) and anti–histone H2B (ab1790; Abcam) antibodies. (B) Quantitative analysis of H2BK120 ubiquitination levels. X-ray films were scanned with GS-800 calibrated densitometer (Bio-Rad Laboratories). The intensity of each H2BK120 ubiquitination signal was normalized by the corresponding H2B signal and averaged.

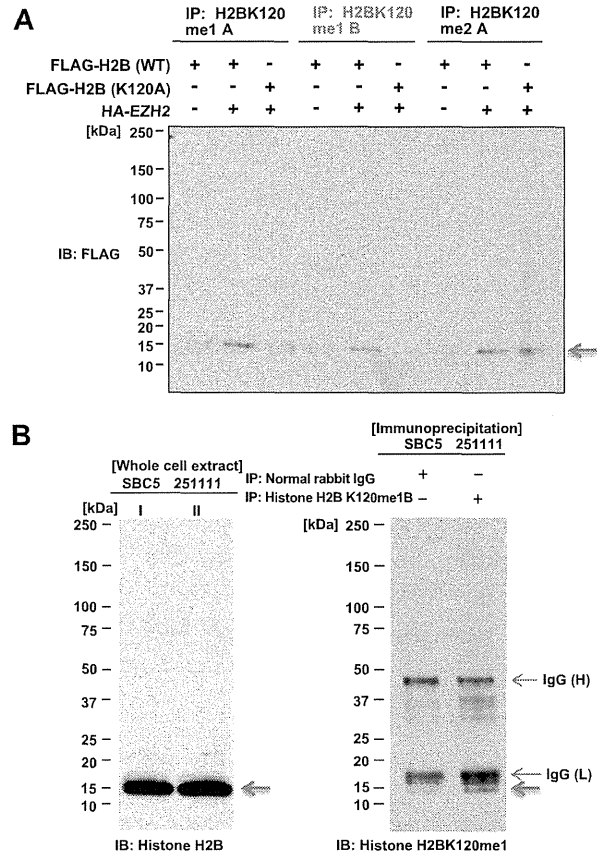


Figure W6. Validation of anti-H2BK120 methylation antibodies. (A) Wild-type FLAG-H2B (H2B-WT) and lysine 120–substituted FLAG-H2B expression vectors (K120A) were transfected into 293T cells. Cell lysates were immunoprecipitated with anti-H2BK120 methylation antibodies, and immunoblot analysis was performed using an anti-FLAG antibody (M2; Sigma-Aldrich). (B) SBC5 cells were lysed with CelLytic M, and cell lysates were immunoprecipitated with an anti-H2BK120 methylation antibody. Whole-cell extracts were blotted with an antihistone H2B antibody (Abcam; ab1790).

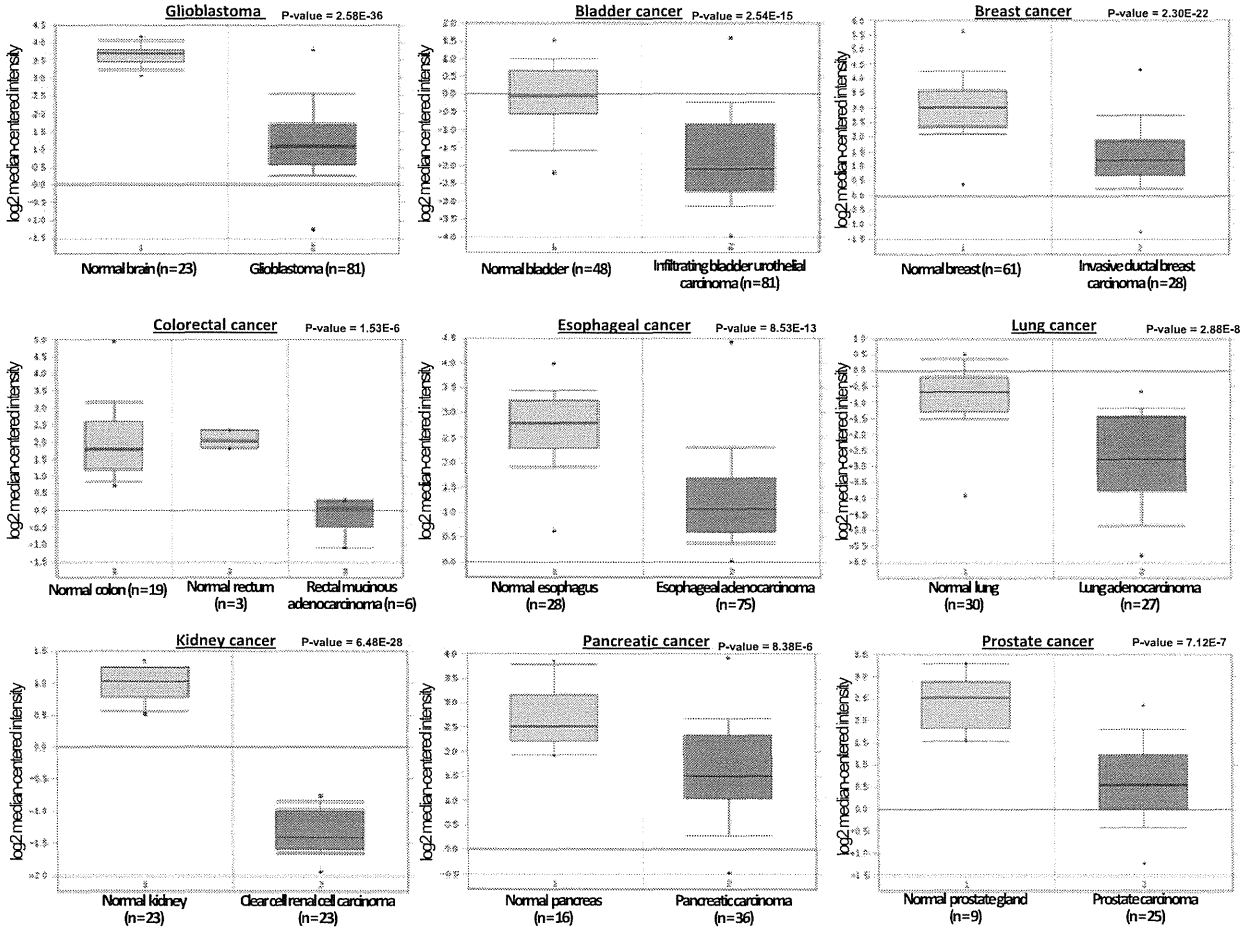


Figure W7. Expression of *GNAO1* in cancer tissues. Expression levels of *GNAO1* are downregulated in various types of cancer. Gene expression data in Oncomine were analyzed. The thick bars in the boxes are average expression levels, and the boxes represent 95% of the samples. The error bars are above or below the boxes, and the range of expression levels is enclosed by two dots.

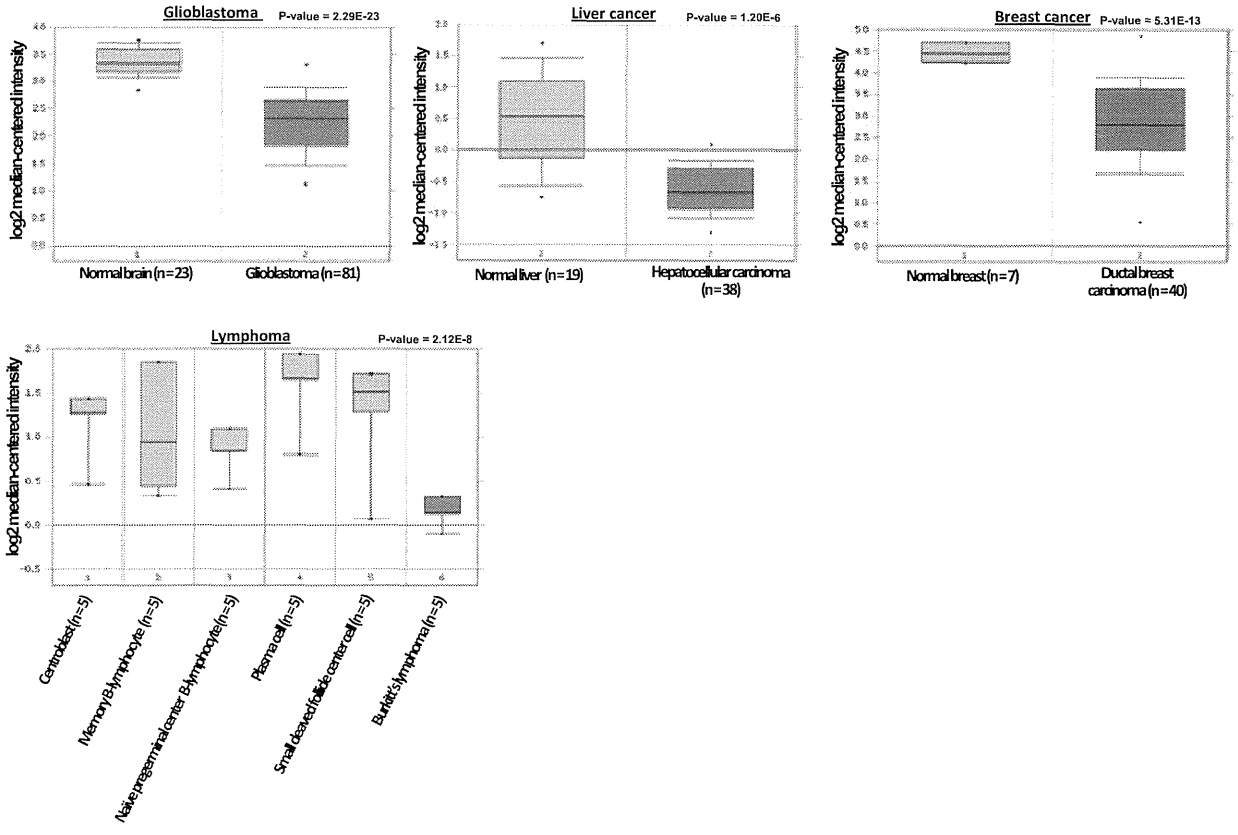


Figure W8. Expression of *ATG2B* in cancer tissues. Expression levels of *ATG2B* are downregulated in various types of cancer. Gene expression data in Oncomine were analyzed. The thick bars in the boxes are average expression levels, and the boxes represent 95% of the samples. The error bars are above or below the boxes, and the range of expression levels is enclosed by two dots.

Table W6. Comparison of H2BK120 Monomethylation and H3K27 Trimethylation Status in the Gene-Regulatory Regions of Downstream Candidates.

Gene Name	H2BK120me1*	H3K27me3*
KCNAB1	0.018	0.024
GNAO1	0.044	0.58435
ATG2B	0.04	NA†
MOV10	0.021	NA†

*Value is the ratio of tag number (siEZH2/siNC).
†NA; No significant methylation was detected in both siEZH2 and siNC samples.

analytical chemistry

Editors' Highlight

pubs.acs.org/ac

Quantitative Structural Characterization of Local N-Glycan Microheterogeneity in Therapeutic Antibodies by Energy-Resolved Oxonium Ion Monitoring

Atsuhiko Toyama,[†,‡] Hidewaki Nakagawa,[§] Koichi Matsuda,[‡] Taka-Aki Sato,[†] Yusuke Nakamura,[⊥] and Koji Ueda*[,§]

[†]Life Science Research Center, Shimadzu Corporation, Kanda-Nishikicho 1, Chiyoda-ku, Tokyo, Japan

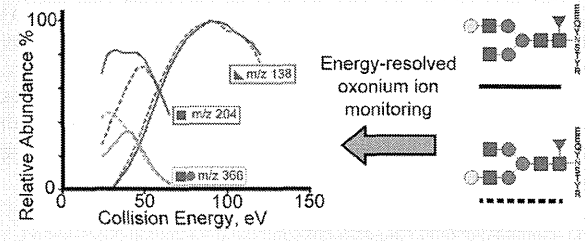
[‡]Institute of Medical Sciences, The University of Tokyo, Shirokanedai 4-6-1, Minato-ku, Tokyo, Japan

[§]Laboratory for Biomarker Development, Center for Genomic Medicine, RIKEN, Shirokanedai 4-6-1, Minato-ku, Tokyo, Japan

[⊥]Section of Hematology/Oncology, Department of Medicine Faculty, The University of Chicago, 5841 S. Maryland Ave, MC2115, Chicago, Illinois, United States

Supporting Information

ABSTRACT: Site-specific characterization of glycoform heterogeneity currently requires glycan structure assignment and glycopeptide quantification in two independent experiments. We present here a new method combining multiple reaction monitoring mass spectrometry with energy-resolved structural analysis, which we termed "energy-resolved oxonium ion monitoring". We demonstrated that monitoring the yields of oligosaccharide-derived fragment ions (oxonium ions) over a wide range of collision induced dissociation (CID) energy applied to a glycopeptide precursor exhibits a glycan structure-unique fragmentation pattern. In the analysis of purified immunoglobulin glycopeptides, the energy-resolved oxonium ion profile was shown to clearly distinguish between isomeric glycopeptides. Moreover, limit of detection (LOD) of glycopeptide detection was 30 attomole injection, and quantitative dynamic range spanned 4 orders magnitude. Therefore, both quantification of glycopeptides and assignment of their glycan structures were achieved by a simple analysis procedure. We assessed the utility of this method for characterizing site-specific N-glycan microheterogeneity on therapeutic antibodies, including validation of lot-to-lot glycoform variability. A significant change in the degree of terminal galactosylation was observed in different production lots of trastuzumab and bevacizumab. Cetuximab Fab glycosylation, previously known to cause anaphylaxis, was also analyzed, and several causative antigens including Lewis X motifs were quantitatively detected. The data suggests that energy-resolved oxonium ion monitoring could fulfill the regulatory requirement on the routine quality control analysis of forthcoming biosimilar therapeutics.



Foreseeing the patent expiry of several therapeutic antibodies that sold more than $30 billion worldwide in 2009,[1] there is growing interest over how biosimilar substitutes could win United States Food and Drug Administration (FDA) approval in the near future.[2] The first draft guidance on the evaluation of biosimilarity was published by the FDA in Feb 2012,[3] in which emphasis was placed on the importance of evaluating minor structural differences that can significantly affect the potency and safety of biopharmaceuticals, with specific reference to glycosylation patterns. This implicates the need for precise and routine structural characterization of biopharmaceuticals. With regard to therapeutic antibodies which have a common N-glycosylation consensus sequence at Asn_{297} in the conserved (Fc) region of heavy chain, some specific features of N-glycosylation have been characterized to affect their potency and safety. For example, the absence of core fucosylation showed enhanced antibody-dependent cellular cytotoxicity by 10-fold,[4] Moreover, recent studies revealed that nonhuman oligosaccharide motifs such as *N*-glycolylneuraminic acid (Neu5Gc)[5] and galactose-α1,3-galactose (α-Gal epitope)[6,7] are immunogenic and can cause anaphylaxis in patients expressing specific IgE.[8] Further characterization elucidated more specifically that immunogenicity of α-Gal epitope was primarily attributed to an extra N-glycan occurring within the antigen-binding (Fab) region.[9] These findings have raised the issue of antibody glycosylation to the level that a global picture of its heterogeneity and biological impact is urgently needed.

Therefore, we developed and describe herein a novel mass spectrometric (MS) method for analyzing the local microheterogeneity of antibody glycosylation. The method involves tryptic digestion of the antibody followed directly by MS analysis at the glycopeptide level, which enables simultaneous

Received: August 14, 2012
Accepted: September 24, 2012

dx.doi.org/10.1021/ac3023372 | Anal. Chem. XXXX, XXX, XXX−XXX

and independent characterization of multiple glycosylation sites. The common drawbacks experienced in MS analysis of glycopeptides are the lack of sensitivity of glycopeptides in a mixture with nonglycosylated peptides[10] and, importantly, the difficulty in verifying the glycan structures by MS/MS analysis of glycopeptides, which in most cases can only determine the oligosaccharide composition. To address these problems, we modified the methodology of multiple reaction monitoring (MRM) using a triple-quadrupole mass spectrometer into a novel concept which we termed "energy-resolved oxonium ion profiling". Oxonium ions are a series of oligosaccharide-derived fragment ions generated by collision induced dissociation (CID).[11,12] We iteratively scanned oxonium ions at various collision energy and plotted the ion yield against the collision energy. This "energy-resolved" approach is known to strengthen structural interpretation of analytes based on differential activation energy of fragmentation,[13−15] with potential application to characterization of oligosaccharide structures.[16,17] Hence, we examined whether the pattern of energy-resolved oxonium ion profile might be unique to the glycan structure of glycopeptides by acquiring the oxonium ion profiles of well-characterized standard glycopeptides. Next, we showed the lot-to-lot heterogeneity in trastuzumab (Herceptin) and bevacizumab (Avastin) as a worked example of routine analysis of therapeutic antibodies for quality control. Moreover, energy-resolved oxonium ion monitoring enabled simultaneous characterization of cetuximab glycoforms at both Fc and Fab glycosylation sites, demonstrating putative anaphylaxis-inducing glycan structures of nonhuman origin.

EXPERIMENTAL SECTION

Reagents. Trastuzumab (Herceptin) and bevacizumab (Avastin) were purchased from Chugai Seiyaku (Osaka, Japan). Cetuximab (Erbitux) was purchased from Merck Serono (Tokyo, Japan). Human immunoglobulin G (IgG), iodoacetamide, iodoacetic acid, and ammonium bicarbonate were purchased from Sigma-Aldrich (Saint Louis, MO). Dithiothreitol, urea, and PD-Miditrap desalting column was purchased from GE Healthcare (Buckinghamshire, UK). Trypsin Gold and ProteaseMAX was supplied by Promega (Madison, WI). All other unspecified reagents were products of Wako Pure Chemicals (Osaka, Japan).

Preparation of Therapeutic Immunoglobulin. Eighty micrograms of trastuzumab, bevacizumab, or cetuximab was dissolved in 33 μL of 8 M urea, and dithiothreitol and ammonium bicarbonate were added to a final concentration of 5 and 50 mM, respectively. After a 30 min incubation in 55 °C, reduced cysteine residues were alkylated in 50 mM iodoacetic acid at room temperature for 30 min. The product was diluted 10 times in 2.5 mM dithiothreitol/25 mM ammnonium bicarbonate before addition of 4 μg of mass spectrometric grade endoproteinase Lys-C (Wako Pure Chemicals) for digestion at 37 °C for 2 h. The Lys-C digest was further digested with addition of 4 μg of Trypsin Gold and allowed to react for an additional 4 h. The digest was desalted with OASIS HLB cartridge and eluted with 15% acetonitrile, and the eluate was diluted 5 times with 0.1% acetic acid for analysis by nanoHPLC-MRM.

Preparation of Glycopeptide Standards. For the preparation of glycopeptide standards, lyophilized powder of IgG was dissolved at 1.5 mg/mL in 50 mM Tris-HCl, pH 7.5, and 8 M urea. The solution was reduced and alkylated prior to trypsin digestion by 10 mM dithiothreitol followed by 50 mM iodoacetamide. The mixture was desalted by a PD-Miditrap desalting column to 100 mM ammonium bicarbonate. Twenty micrograms of Trypsin Gold was added to every 1 mg of IgG, and the mixture was incubated at 37 °C for 12 h. Tryptic digest was aliquoted after desalting by OASIS HLB cartridge (Waters Corporation, Milford, MA), and the aliquots were dried by SpeedVac for long-term storage. Homogeneous glycopeptides were prepared by two-dimensional HPLC fractionation using the Prominence HPLC system (Shimadzu Corporation, Kyoto, Japan). The first dimension was performed with 4.6 mm × 500 mm Cadenza CD-C18 column (Imtakt Corporation, Kyoto, Japan). Using 0.2% trifluoroacetic acid (TFA, Tokyo Chemical Industry, Tokyo, Japan) as solvent A and 75% acetonitrile/0.1% TFA as solvent B, IgG tryptic digest was separated by a 5−20% solvent B gradient in 90 min at a flow rate of 0.6 mL/min. Collected fractions were dried, reconstituted in 0.1% TFA, and further fractionated by a 150 mm × 4.6 mm SunShell C18 column (Chromanik Corporation, Osaka, Japan). Solvent A was 6% acetonitrile with 0.1% TFA, and Solvent B was 12% acetonitrile with 0.1% TFA; separation was performed over a 10−90% solution B gradient.

Fluorescent HPLC. A portion of each glycopeptide fraction was treated with 0.5 U of N-glycosidase F (Roche Diagnostics, Basel, Switzerland) dissolved in 0.01% ProteaseMAX solution and incubated at 37 °C for 8 h. Released glycans were dried in a SpeedVac and derivatized with 2-aminopyridine (2-AP) by reductive amination as described elsewhere. 2-AP labeled oligosaccharides were purified by a MassPREP HILIC μElution plate (Waters Corporation) according to the manufacturer's protocols but using 90% acetonitrile/100 mM triethylammonium bicarbonate for washing and 25% acetonitrile/100 mM triethylammonium bicarbonate for elution. The separation of 2-AP labeled oligosaccharides was carried out by RP-HPLC with a Shimpack CLC-ODS 0.6 × 150 mm column (Shimadzu Corporation). Solvent A was 10 mM sodium phosphate buffer pH 3.8, and Solvent B was the same buffer containing 0.5% 1-butanol. After equilibrating the column with 20% solvent B, the oligosaccharides were eluted by raising the concentration of solvent B to 50% in 75 min. 2-AP labeled oligosaccharides were detected by fluorescence using excitation and emission wavelengths of 320 and 400 nm, respectively. The retention times were normalized and expressed as glucose units using PA-glucose oligomer (Takara Bio Inc., Shiga, Japan), and glycan structures were verified according to the HPLC mapping data as described elsewhere.[18]

Quantification of Glycopeptide Standards. Glycopeptide fractions, which were successfully purified at high yield, and structural homogeneity were next subjected to quantification by UV absorbance following HPLC separation. The accuracy of quantification was guaranteed using lyophilized synthetic peptide EEQYNSTYR as the standard for calibration curve. Two species of glycopeptides were successfully quantified and were used as calibration standards for assessing the quantification by oxonium ion monitoring (Figure 2b). Standard concentration was adjusted to 100, 33.3, 11.1, 3.70, 1.23, 0.41, 0.14, 0.045, 0.015, and 0.005 fmol/μL with 0.1% TFA containing 50 fmol/μL (final concentration) of EEQYNSTYR as the internal standard and analyzed by nanoHPLC-MRM as described in the next subsection.

Multiple Reaction Monitoring of Oxonium Ions. The energy-resolved oxonium ion profiling and glycopeptide quantification by m/z 138 oxonium ion monitoring was performed using a 4000QTRAP triple quadrupole mass

dx.doi.org/10.1021/ac3023372 | Anal. Chem. XXXX, XXX, XXX−XXX

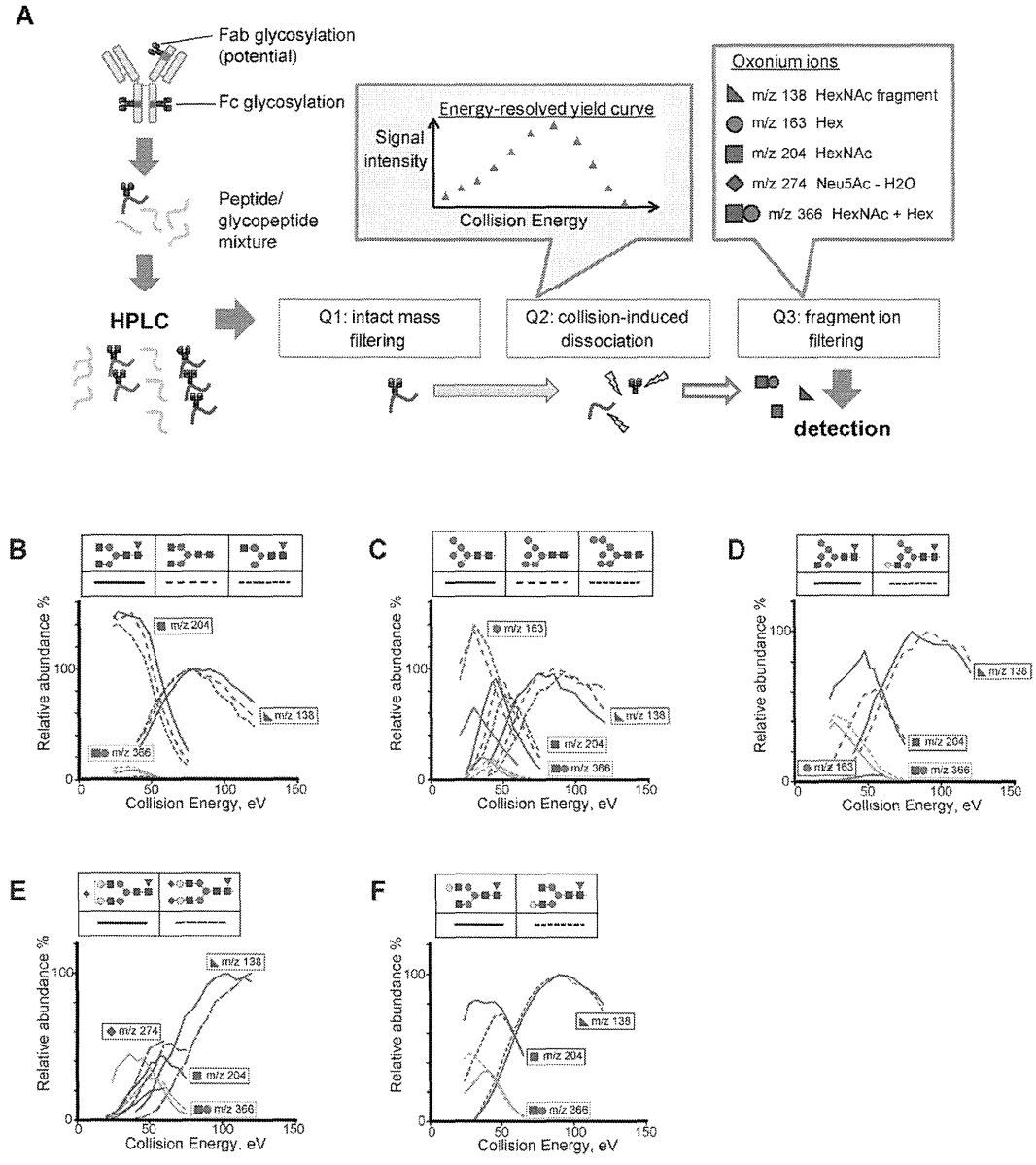


Figure 1. The energy-resolved oxonium ion monitoring and profiling. (a) An overview of glycoform analysis by multiple reaction monitoring (MRM) mass spectrometry. Glycoprotein sample (immunoglobulin) is digested by trypsin into peptide/glycopeptide mixture and is directly subjected to MS analysis. In the mass spectrometer, the first quadrupole (Q1) isolates the glycopeptide ions with unique intact mass, which are subsequently guided into the second quadrupole (Q2) where they undergo collision-induced dissociation. The kinetic energy of the precursor ions, which governs the energy and rate of dissociation, can be controlled by changing the collision energy applied in the Q2 cell. The third quadrupole (Q3) filters the product ions to selectively detect oligosaccharide-derived oxonium ions (inset) at high-sensitivity. (b–f) The energy-resolved oxonium ion profiles of representative glycopeptides displayed as yield curves. Similar profiles are observed for glycoforms sharing the same motifs, such as nongalactosylated type (b), high-mannose type (c), hybrid type (d), and sialylated type (e). The monogalactosylated glycopeptide isomers (f) have been purified to >90% by HPLC. Blue square, GlcNAc; green circle, mannose (Man); yellow circle, Gal; red triangle, Fuc; purple diamond, Neu5Ac. The peptide portion EEQYNSTYR has been omitted from structure display for simplicity.

spectrometer (AB Sciex, Foster City, CA) with Agilent 1200 nano-HPLC system (Agilent Technologies, Palo Alto, CA). Glycopeptide samples were separated by a 75 μm × 200 mm electrospray ionization (ESI) sprayer tip packed with 3 μm C18 resin (Nikkyo Technos, Tokyo, Japan). The compositions of solvents A and B were 0.1% formic acid and 70% acetonitrile/ 0.1% formic acid, respectively, and the flow rate was 250 nL/ min. The column was equilibrated with 20% solvent B, and glycopeptides were separated isocratically for 3 min, after which the remainder was eluted by a 5 min gradient from 20% to 95% solvent B. The fixed instrumental settings were as follows: 2200 V ionization spray voltage; 12 psi curtain gas (N_2); CAD = 4; 70 V declustering potential; 10 V entrance potential; Q1 resolution, HIGH; Q3 resolution, LOW; and 2 ms pause in between. For energy-resolved data acquisition, analysis transitions were designed such that there were several transitions for the same Q1 and Q3 mass combination, differing only by collision energy settings (in steps of 3 or 5

dx.doi.org/10.1021/ac3023372 | Anal. Chem. XXXX, XXX, XXX–XXX

eV). Dwell times were adjusted to the total cycle time of 1−1.5 s. For relative quantification of glycopeptide microheterogeneity, precursor ion m/z was derived in-silico, covering all possible glycan compositions, and a single quantitative transition was monitored for each precursor ion (Q3 m/z = 138, collision energy (CE) = predicted optimum collision energy, as described in text). Only for generating the calibration curves for determining the sensitivity and dynamic range of detection (Figure 2b), synthetic peptide EEQYNSTYR was coinjected, and the following internal standard transitions were monitored for normalization: Q1 m/z = 595.4, Q3 m/z = 640.4, CE = 38; Q1 m/z = 595.4, Q3 m/z = 526.4, CE = 30.

Data Analysis. The acquired raw data was processed with MultiQuant version 2.02 (AB Sciex, Foster City, CA). The mass chromatogram of each transition was integrated after Gaussian Smoothing (σ = 1 point) and quantified by peak area. For energy-resolved oxonium ion profiling, the data set for each glycopeptide was normalized with respect to the highest intensity of Q3 m/z = 138 and expressed as percentage. Analyses were performed in triplicates, and average normalized percentage was computed. Optimum collision energy was defined as the voltage setting to achieve the maximum percentage, unless a series of 98−100% was observed as a plateau, in which case the average collision energy was regarded as optimum. For relative quantification of glycopeptide microheterogeneity, the percentage composition of each component was calculated with respect to the sum of all components combined. The data for plotting the calibration curves (Figure 2b) was generated by normalizing the glycopeptide peak area by the average of two internal standard transitions.

RESULTS AND DISCUSSION

Energy-Resolved Oxonium Ion Profile. Figure 1a shows the scheme to acquire an energy-resolved yield curve of oxonium ions using a triple quadrupole mass spectrometer. Briefly, the collision energy applied in the Q2 cell was changed in stepwise fashion in repetitive measurements of short time intervals for each mass filter settings. We termed simultaneous acquisition of oxonium ion yield curves as energy-resolved oxonium ion monitoring technology and examined whether this profile could be utilized for determination of glycan structures on a particular glycopeptide. We first prepared a series of glycopeptides from human IgG by chromatographic separation. The isolated glycopeptides shared the same amino acid sequence EEQYNSTYR but bore different glycan structures. For each glycopeptide, energy-resolved yield curves of 5 selected oxonium ions were simultaneously acquired and displayed on the same plot, namely, m/z 138 (internal fragment of *N*-acetylglucosamine, GlcNAc), m/z 204 (GlcNAc), m/z 366 [Hexose (Hex) + GlcNAc], m/z 163 (Hexose), and m/z 274 (*N*-acetylneuraminic acid, Neu5Ac − H_2O). The results as displayed in Figure 1b−f convey an interesting feature of the oxonium ion profiles, that they represent the glycan structure as a whole rather than simple summation of possible bond breakage events. The non-galactosylated glycoforms (Figure 1b) are characterized by an explicitly higher yield of m/z 204 (GlcNAc) than other glycoforms bearing the same number of GlcNAc. Moreover, the high yield of oxonium ion m/z 163, which is unique to the high-mannose type (Figure 1c), cannot be explained by the number of mannosyl linkages alone because hybrid type (Figure 1d) structures hardly produce m/z 163. A possible explanation of this is that mannose or galactose, which is in proximity to GlcNAc, lacks the labile proton needed for oxonium ion formation.

Most striking was the different profile acquired from independent analysis of isomeric glycopeptides with glycan composition 4[Hex]4[GlcNAc]1[Fuc] (where Fuc is fucose; structure validation by fluorescent HPLC supplied in Figure S1, Supporting Information), differing by terminal galactose linked either to α1-3 branch or α1-6 branch (Figure 1f). These glycans share the same types of glycosidic linkages but exhibited distinct energy-resolved patterns, suggesting that galactose present on α1-6 branch contributed to higher yield of m/z 366 oxonium ion at low collision energy than when present on α1-3 branch. Previous examples of energy-resolved mass spectrometry suggested that the maximum of yield curve indicates the ease of CID event to generate the product ion.[15] Extending this view, the different maxima observed in m/z 366 curves (Figure 1f) are indicative of one bond being more stable than the other, despite that the bonds responsible for m/z 366 production, GlcNAc β(1-2)Man, are locally identical in both structures.

From these results, we conclude that bond stability in glycans is governed by the overall spatial arrangement, and the energy-resolved oxonium ion profile reflects the shape of the glycan rather than its composition. The notion of precise glycoform identification by relative quantification of fragment ions is not a novel one,[19] but its application has been limited to detached N-glycans. Here, we showed that energy-resolved oxonium ion monitoring can identify glycoforms attached to peptides by focusing on low molecular weight oxonium ions, which occur at the same m/z for any glycopeptides, and differentiating between structures by collision energy-resolved profiles. Multiple reaction monitoring by triple-quadrupole mass spectrometer complies with this purpose because it does not exert a low-mass cutoff effect which otherwise alters the signal yield, depending on the size of intact glycopeptides. Therefore, we expect to observe the same energy-resolved oxonium ion profiles for the same glycoform regardless of the peptide amino acid sequence. Furthermore, this energy-resolved approach can potentially be used to estimate the percentage composition of isomeric glycopeptides without requiring chromatographic separation. The energy-resolved yield curves of 1:1 and 1:2 mixtures of the two isomers showed curve shapes that fell between the two reference curves (Figure S2, Supporting Information).

High-Sensitivity Quantification of Glycopeptides. A notable observation from various energy-resolved profiles was that the yield curve shapes of oxonium ion m/z 138 were similar among the tested glycan structures. We speculated that m/z 138 originates from the GlcNAc of the core structure which was conserved among all N-glycans and, therefore, oxonium ion m/z 138 under optimum collision energy would serve as an ideal reporter ion for quantifying glycopeptides of various glycan structures. To clarify this, we plotted the collision energy at which m/z 138 was maximal against the m/z of various glycopeptides (Figure 2a). The result suggested a strong linear correlation (R^2 = 0.842), which was independent of the number of GlcNAc in the glycan composition or the peptide amino acid sequence. From these results, it seemed adequate to assume that the collision energy to yield maximum m/z 138 signal can be predicted from the m/z of glycopeptide and that m/z 138 signal acquired in this way would best represent the quantity of the glycopeptide. Furthermore, since the relative intensity of m/z 138 at its maximum was generally

dx.doi.org/10.1021/ac3023372 | Anal. Chem. XXXX, XXX, XXX−XXX

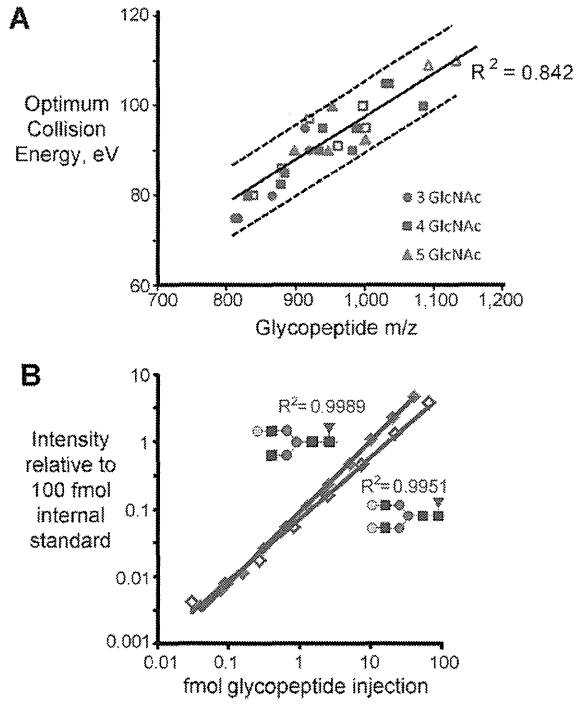


Figure 2. Monitoring oxonium ion m/z 138 for high-sensitivity quantification of glycopeptides. (a) The energy-resolved oxonium ion profiles were acquired for 28 glycopeptides, and the collision energy that yielded highest signal intensity of m/z 138 was plotted against the m/z of precursor glycopeptide. Blue circle, glycopeptide containing 3 GlcNAc; red square, containing 4 GlcNAc; green triangle, containing 5 GlcNAc; filled points, peptide sequence EEQYNSTYR; open points, peptide sequence MNSLQSNDTAIYYCAR. Correlation coefficient R^2 is 0.842, and all the points fall within 95% confidence interval (broken lines). (b) Two purified homogeneous glycopeptides, 4[Hex]4-[HexNAc]1[Fuc] and 5[Hex]4[HexNAc]1[Fuc] (where HexNAc is *N*-acetylhexosamine), were serially diluted and subjected to sensitivity determination. m/z 138 was monitored at the optimum collision energy. Signal was normalized to 100 fmol of internal standard EEQYNSTYR (synthetic peptide with no glycosylation). The quantitative range of detection was determined to be 30 amol to 30 fmol injection with R^2 = 0.9989 and 0.9951.

higher than other oxonium or peptide-derived fragment ions, it was expected that monitoring m/z 138 would enhance the sensitivity of glycopeptide detection. This was demonstrated by the standard curve of two glycopeptide standards, each composed of different and homogeneous glycan structure and concentration determined by UV absorbance (Figure 2b). The linear quantitative range of the glycopeptides were 30 amol to 30 fmol injection, spanning 4 orders of magnitude. Surprisingly, the sensitivity of glycopeptide detection was about 10 times more than the synthetic nonglycopeptide EEQYNSTYR which was coinjected as an internal standard. This result is contrary to the common notion that glycopeptides have relatively low ionization efficiency[20] and implicates the overwhelming capability of our technology in detecting glycopeptides at high sensitivity and quantitativeness in mixture with non-glycopeptides. Moreover, the yield of m/z 138 is maximized at much higher collision energy than conventionally used for MRM.[21] Under this optimum condition, production of other oxonium ions or peptide backbone fragmentation is suppressed, which effectively reduces the background signal and further contributes to selective detection of glycopeptides.

Lot-to-Lot Heterogeneity of Trastuzumab and Bevacizumab. We next evaluated the applicability of the oxonium ion monitoring technology, taking advantage of the highly quantitative detection of glycopeptides, to routine lot-to-lot quality control of therapeutic antibodies by analyzing four lots of trastuzumab (Herceptin) and bevacizumab (Avastin). The glycoform heterogeneity was quantified, and the relative abundances of three most abundant glycoforms were presented in Figure 3 to illustrate the difference among the production

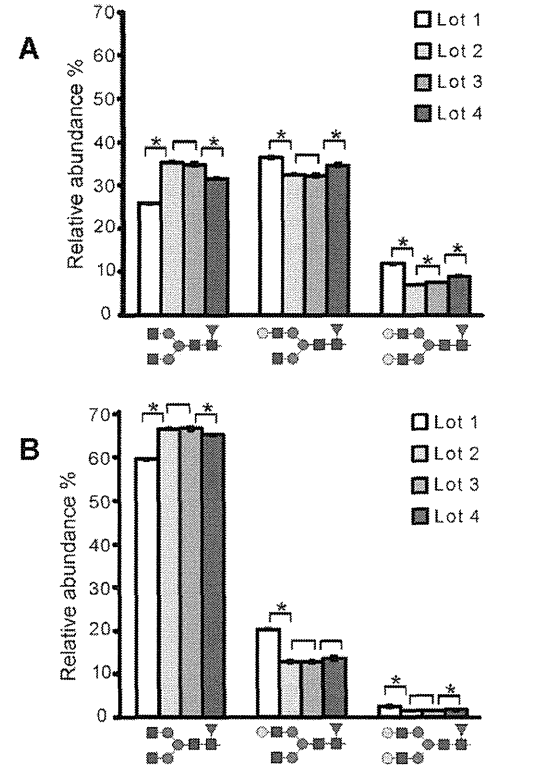


Figure 3. Evaluation of lot-to-lot heterogeneity of trastuzumab and bevacizumab. The glycoform heterogeneity of four lots of each product was quantified by m/z 138 oxonium ion monitoring, and the relative abundances of agalactosylated, monogalactosylated, and digalactosylated structures are displayed for (a) trastuzumab and (b) bevacizumab. The lots are numbered in chronological order of production, and the time-course variation was evaluated by comparing two consecutive lots. Error bars are standard deviation of triplicate runs. Asterisks represent where the difference between consecutive lots was statistically significant ($p < 0.05$ by Student's t test, $n = 3$).

lots. The quantification results of other glycoforms were given in Tables S1 and S2, Supporting Information. Analysis was performed in three independent preparations for each lot, and statistical significance of the difference between two consecutive lots was evaluated by Student's t test. The mean coefficient of variation of triplicate analysis was 4.7% for all measurements with more than 0.1% relative abundance, and the quantitative linearity over 4 orders of magnitude enabled the detection of minor glycoforms of as low as 0.05% relative abundance. As a result, the frequency of galactosylation onto the biantennary GlcNAc was shown to be prone to significant lot-to-lot variation in both trastuzumab and bevacizumab (Figure 3a,b,

dx.doi.org/10.1021/ac3023372 | *Anal. Chem.* XXXX, XXX, XXX–XXX

respectively, $P < 0.05$ by Student's t test). Convincingly, the observed variations in glycoform composition were interrelated in the way that the decrease in galactosylated form was compensated by the increase in agalactosylated form. The smallest statistically significant variation observed here was 1.2-fold between Lot 3 and Lot 4 of bevacizumab in the abundance of digalactosylated form (1.29% and 1.51%, respectively, $P < 0.05$ by Student's t test). Surprisingly, we observed that bevacizumab contained a fraction of nonglycosylated form (EEQYNSTYR detected without glycosylation) and that this fraction also varied among lots. These results outline the continual lot-to-lot variation of glycoform heterogeneity in both trastuzumab and bevacizumab and demonstrate that m/z 138 oxonium ion monitoring can accurately evaluate whether the glycoforms of therapeutic antibodies between lots are identical.

Analysis of Immunogenic Glycoforms on Fc and Fab of Cetuximab. We next applied energy-resolved oxonium ion monitoring for site-specific glycoform characterization of therapeutic antibody cetuximab for identification of potentially immunogenic glycoforms. Figure 4a shows the MRM chromatogram of cetuximab digest, monitoring m/z 138 for the quantification of all possible glycoforms. A series of glycopeptides derived from Fc region (EEQYN$_{297}$STYR) or Fab region (MNSLQSN$_{88}$DTAIYYCAR) were eluted at 13–14 min or 17–19 min, respectively. Glycoforms on both sites were thus characterized simultaneously in a single analysis. The most major glycoform in Fc glycan was agalactosylated biantennary structure, and relatively high level of hybrid and high-mannose structures were also detected (results summarized in Figure S3, Supporting Information). Focusing on the characterization of its Fab glycan, which has been reported as the primary cause of anaphylaxis by cetuximab,[9] we quantified all theoretically available combinations of glycoforms with potentially immunogenic glycan motifs and verified their structures by energy-resolved oxonium ion profiling. As a result, we discovered that a Lewis X antigen was also present on cetuximab along with previously reported α-Gal and Neu5Gc antigens (Figure 4b). The Lewis X motif was verified by the energy-resolved yield curves including the structure-specific oxonium ion of m/z 512 (Gal-GlcNAc-Fuc), which clearly discriminated the two almost isomeric glycopeptides with Lewis X or Neu5Gc motifs. The result of quantification was summarized in Table 1, showing the panel of biantennary and triantennary glycan structures with Lewis X, α-Gal, and/or Neu5Gc modification. Since the total number of modification never exceeded the number of antennae, it is most likely that addition of fucose to antennal GlcNAc (Lewis X), addition of galactose to terminal galactose (α-Gal), and addition of Neu5Gc to terminal galactose (Neu5Gc) occur in a competitive fashion, resulting in the occurrence of all possible combinations of the antigenic motifs. Considering the fact that 1–2% of patients administered with cetuximab experience severe hypersensitivity reactions[22] and potentially at much higher rate,[23] it is important to fully elucidate the cause of this hypersensitivity and strictly regulate it to minimize the risk. The implications of the variety of immunogenic glycoforms which has previously been overlooked would require revisiting.

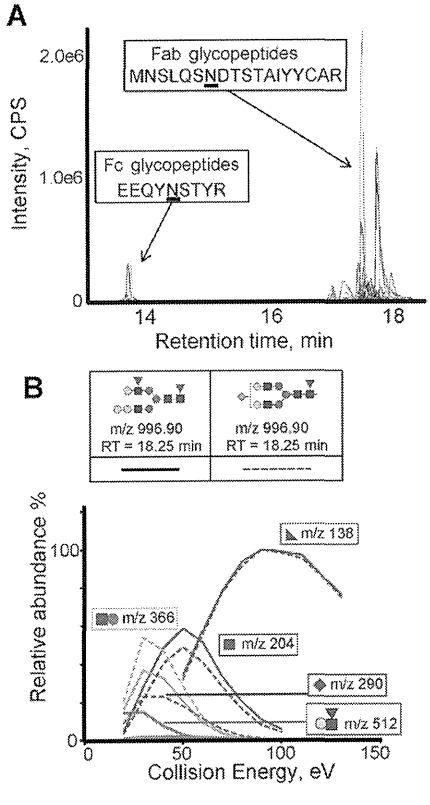


Figure 4. Characterization of glycoform heterogeneity of cetuximab. (a) Overlaid MRM chromatogram showing the quantification of glycoform heterogeneity in cetuximab digest by monitoring product ion m/z 138 for 40 different glycan structures. Fc glycopeptides elute at 14 min and Fab glycopeptides elute at 18 min, with minor variation in retention time by glycoforms. (b) The energy-resolved oxonium ion profiles of glycopeptides with m/z 996.90 and m/z 996.65 were acquired for verification of their structures, represented by solid lines and broken lines, respectively. The deduced structures are shown in the inset. Blue square, GlcNAc; green circle, Man; yellow circle, Gal; red triangle, Fuc; blue diamond, Neu5Gc. The peptide portion MNSLQSNDTAIYYCAR is omitted for simplicity.

CONCLUSIONS

We developed the new technology that combined multiple reaction monitoring mass spectrometry and energy-resolved structural analysis. It focuses on how the yields of commonly observed fragment ion signals change under various fragmentation conditions and uses this information as a new cue to replace the conventional notion of finding structure-specific signature ions for glycoform identification.[24,25] The advantage of this technology over other methods is 2-fold. First, the collision energy-resolved oxonium ion profiles uniquely represent glycan structures on tryptic glycopeptides. It is therefore suited for site-specific analyses for characterizing glycoproteins with multiple glycosylation sites. Moreover, the depth of structural information covers differentiation of isomeric structures, which currently requires glycan detachment. By accumulating more reference data of glycopeptides with validated glycan structures, it should be possible in the near future to program a matching system to assign unknown data with the glycan structure. Second, detection of glycopeptides by monitoring oxonium ion m/z 138 at optimum collision energy showed superior sensitivity and quantitative range to all currently available methods. This feature enables analysis using minute sample quantity or detection of very minor glycoforms to unravel glycoform microheterogeneity, as demonstrated herein by the analysis of therapeutic antibodies. The energy-resolved oxonium ion monitoring technology thus

dx.doi.org/10.1021/ac3023372 | Anal. Chem. XXXX, XXX, XXX–XXX

Table 1. Combinations of Potentially Immunogenic Glycan Motifs Found on Cetuximab[a]

	composition			% abundance	
	α-Gal	Lewis X	Neu5Gc	biantennary	triantennary
antenna occupancy = 1	1			6.4	1.1
		1		0.7	0.2
			1	7.1	0.9
antenna occupancy = 2	2			25.9	1.7
	1	1		2.5	1.0
	1		1	18.5	1.3
		2		0.2	0.1
		1	1	1.0	0.2
			2	3.3	0.3
antenna occupancy = 3	3				2.5
	2	1			1.0
	2		1		1.8
	1	2			0.2
	1	1	1		0.5
	1		2		0.6
		3			0.0
		2	1		0.1
		1	2		0.2
			3		N.D.

[a]Composition indicates the number of α-Gal, Lewis X, and Neu5Gc motifs in the glycan structure. N.D.: no signal detected.

shows novelty in both qualitative and quantitative aspects of glycoprotein analysis. It can potentially alleviate a burden on multiple, orthogonal analyses for site-specific characterization of glycan microheterogeneity. Primarily, it is applicable to evaluation of biosimilar pharmaceuticals at a routine basis; however, it can potentially be applied to any glycoproteins, with or without prior enrichment of the target.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary Figures S1–S3 and Supplementary Tables S1 and S2. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: k-ueda@riken.jp.

Notes

The authors declare the following competing financial interest(s): A.T. and T.A.S. are employees of Shimadzu Corporation, see grant information below.

■ ACKNOWLEDGMENTS

This study was supported in part by the grants from Toppan Printing Co., Ltd., Tokyo, Japan, and Shimadzu Corporation, Kyoto, Japan. This work was supported by funding to K.U. from The 3rd Term Comprehensive 10-year Strategy for Cancer Control, The Ministry of Health Labour and Welfare Japan.

■ REFERENCES

(1) Ledford, H. *Nature* **2010**, *468* (7320), 18–19.
(2) Weise, M.; Bielsky, M. C.; De Smet, K.; Ehmann, F.; Ekman, N.; Narayanan, G.; Heim, H. K.; Heinonen, E.; Ho, K.; Thorpe, R.; Vleminckx, C.; Wadhwa, M.; Schneider, C. K. *Nat. Biotechnol.* **2011**, *29* (8), 690–693.
(3) FDA issues draft guidance on biosimilar product development; http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm291232.htm.
(4) Shinkawa, T.; Nakamura, K.; Yamane, N.; Shoji-Hosaka, E.; Kanda, Y.; Sakurada, M.; Uchida, K.; Anazawa, H.; Satoh, M.; Yamasaki, M.; Hanai, N.; Shitara, K. *J. Biol. Chem.* **2003**, *278* (5), 3466–3473.
(5) Ghaderi, D.; Taylor, R. E.; Padler-Karavani, V.; Diaz, S.; Varki, A. *Nat. Biotechnol.* **2010**, *28* (8), 863–867.
(6) Qian, J.; Liu, T.; Yang, L.; Daus, A.; Crowley, R.; Zhou, Q. *Anal. Biochem.* **2007**, *364* (1), 8–18.
(7) Bosques, C. J.; Collins, B. E.; Meador, J. W., 3rd; Sarvaiya, H.; Murphy, J. L.; Dellorusso, G.; Bulik, D. A.; Hsu, I. H.; Washburn, N.; Sipsey, S. F.; Myette, J. R.; Raman, R.; Shriver, Z.; Sasisekharan, R.; Venkataraman, G. *Nat. Biotechnol.* **2010**, *28* (11), 1153–1156.
(8) Chung, C. H.; Mirakhur, B.; Chan, E.; Le, Q. T.; Berlin, J.; Morse, M.; Murphy, B. A.; Satinover, S. M.; Hosen, J.; Mauro, D.; Slebos, R. J.; Zhou, Q.; Gold, D.; Hatley, T.; Hicklin, D. J.; Platts-Mills, T. A. *N. Engl. J. Med.* **2008**, *358* (11), 1109–1117.
(9) Lammerts van Bueren, J. J.; Rispens, T.; Verploegen, S.; van der Palen-Merkus, T.; Stapel, S.; Workman, L. J.; James, H.; van Berkel, P. H.; van de Winkel, J. G.; Platts-Mills, T. A.; Parren, P. W. *Nat. Biotechnol.* **2011**, *29* (7), 574–576.
(10) Wada, Y.; Azadi, P.; Costello, C. E.; Dell, A.; Dwek, R. A.; Geyer, H.; Geyer, R.; Kakehi, K.; Karlsson, N. G.; Kato, K.; Kawasaki, N.; Khoo, K. H.; Kim, S.; Kondo, A.; Lattova, E.; Mechref, Y.; Miyoshi, E.; Nakamura, K.; Narimatsu, H.; Novotny, M. V.; Packer, N. H.; Perreault, H.; Peter-Katalinic, J.; Pohlentz, G.; Reinhold, V. N.; Rudd, P. M.; Suzuki, A.; Taniguchi, N. *Glycobiology* **2007**, *17* (4), 411–422.
(11) Huddleston, M. J.; Bean, M. F.; Carr, S. A. *Anal. Chem.* **1993**, *65* (7), 877–884.
(12) Kováčik, V.; Hirsch, J.; Kováč, P.; Heerma, W.; Thomas-Oates, J.; Haverkamp, J. *J. Mass Spectrom.* **1995**, *30* (7), 949–958.
(13) McLuckey, S. A.; Glish, G. L.; Cooks, R. G. *Int. J. Mass Spectrom. Ion Phys.* **1981**, *39* (2), 219–230.
(14) Fetterolf, D. D.; Yost, R. A. *Int. J. Mass Spectrom. Ion Process* **1982**, *44* (1–2), 37–50.
(15) Bursey, M. M.; Nystrom, J. A.; Hass, J. R. *Anal. Chim. Acta* **1984**, *159* (0), 265–274.
(16) Laine, R. A.; Pamidimukkala, K. M.; French, A. D.; Hall, R. W.; Abbas, S. A.; Jain, R. K.; Matta, K. L. *J. Am. Chem. Soc.* **1988**, *110* (21), 6931–6939.

dx.doi.org/10.1021/ac3023372 | *Anal. Chem.* XXXX, XXX, XXX–XXX

(17) Daikoku, S.; Ako, T.; Kato, R.; Ohtsuka, I.; Kanie, O. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (10), 1873–1879.
(18) Tomiya, N.; Awaya, J.; Kurono, M.; Endo, S.; Arata, Y.; Takahashi, N. *Anal. Biochem.* **1988**, *171* (1), 73–90.
(19) Kameyama, A.; Nakaya, S.; Ito, H.; Kikuchi, N.; Angata, T.; Nakamura, M.; Ishida, H. K.; Narimatsu, H. *J. Proteome Res.* **2006**, *5* (4), 808–814.
(20) Jebanathirajah, J.; Steen, H.; Roepstorff, P. *J. Am. Soc. Mass Spectrom.* **2003**, *14* (7), 777–784.
(21) MacLean, B.; Tomazela, D. M.; Abbatiello, S. E.; Zhang, S.; Whiteaker, J. R.; Paulovich, A. G.; Carr, S. A.; MacCoss, M. J. *Anal. Chem.* **2010**, *82* (24), 10116–10124.
(22) Cunningham, D.; Humblet, Y.; Siena, S.; Khayat, D.; Bleiberg, H.; Santoro, A.; Bets, D.; Mueser, M.; Harstrick, A.; Verslype, C.; Chau, I.; Van Cutsem, E. *N. Engl. J. Med.* **2004**, *351* (4), 337–345.
(23) O'Neil, B. H.; Allen, R.; Spigel, D. R.; Stinchcombe, T. E.; Moore, D. T.; Berlin, J. D.; Goldberg, R. M. *J. Clin. Oncol.* **2007**, *25* (24), 3644–3648.
(24) Robbe, C.; Capon, C.; Coddeville, B.; Michalski, J.-C. *Rapid Commun. Mass Spectrom.* **2004**, *18* (4), 412–420.
(25) Hashii, N.; Kawasaki, N.; Itoh, S.; Harazono, A.; Matsuishi, Y.; Hayakawa, T.; Kawanishi, T. *Rapid Commun. Mass Spectrom.* **2005**, *19* (22), 3315–3321.

dx.doi.org/10.1021/ac3023372 | *Anal. Chem.* XXXX, XXX, XXX–XXX

nature COMMUNICATIONS

ARTICLE

Received 5 Dec 2011 | Accepted 11 Jan 2012 | Published 14 Feb 2012

DOI: 10.1038/ncomms1676

Regulation of histone modification and chromatin structure by the p53–PADI4 pathway

Chizu Tanikawa[1], Martha Espinosa[1,2], Akari Suzuki[3], Ken Masuda[1], Kazuhiko Yamamoto[3,4], Eiju Tsuchiya[5,6], Koji Ueda[7], Yataro Daigo[1,8], Yusuke Nakamura[1] & Koichi Matsuda[1]

Histone proteins are modified in response to various external signals; however, their mechanisms are still not fully understood. Citrullination is a post-transcriptional modification that converts arginine in proteins into citrulline. Here we show *in vivo* and *in vitro* citrullination of the arginine 3 residue of histone H4 (cit-H4R3) in response to DNA damage through the p53–PADI4 pathway. We also show DNA damage-induced citrullination of Lamin C. Cit-H4R3 and citrullinated Lamin C localize around fragmented nuclei in apoptotic cells. Ectopic expression of PADI4 leads to chromatin decondensation and promotes DNA cleavage, whereas *Padi4*$^{-/-}$ mice exhibit resistance to radiation-induced apoptosis in the thymus. Furthermore, the level of cit-H4R3 is negatively correlated with p53 protein expression and with tumour size in non-small cell lung cancer tissues. Our findings reveal that cit-H4R3 may be an 'apoptotic histone code' to detect damaged cells and induce nuclear fragmentation, which has a crucial role in carcinogenesis.

[1] Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo 1088639, Japan. [2] Department of Infectomics and Molecular Pathogenesis, Center for Research and Advanced Studies, Mexico City 07360, Mexico. [3] Laboratory for Rheumatic Diseases, Center for Genomic Medicine, RIKEN, Yokohama 2300045, Japan. [4] Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo, Tokyo 1138655 Japan. [5] Department of Pathology, Saitama Cancer Center, Saitama 3620806, Japan. [6] Molecular Pathology and Genetics Division, Kanagawa Cancer Center Research Institute, Kanagawa 2418515, Japan. [7] Laboratory for Biomarker Development, Center for Genomic Medicine, RIKEN, Yokohama 2300045, Japan. [8] Department of Medical Oncology, Shiga University of Medical Science, Otsu 5202192, Japan. Correspondence and requests for materials should be addressed to K.M. (email: koichima@ims.u-tokyo.ac.jp).

In response to DNA damage, various protein kinases phosphorylate p53 at its amino terminus leading to the stabilization and accumulation of p53 protein[1,2]. As a guardian of the genome, p53 transactivates a number of its target genes to induce cell-cycle arrest or apoptosis and subsequently eliminates cells with risk of malignant transformation[3]. Morphological changes such as chromatin condensation, DNA laddering, nuclear lamina breakdown and nuclear fragmentation are characteristic of apoptotic cell death, although the role of p53 during these dynamic nuclear events has not been elucidated.

We have isolated a number of p53-target genes including p53AIP1, p53R2, p53RDL1 and XEDAR[4–9], and implicated the molecular mechanisms by which p53 regulated the cell fate, death or survival, by regulating the expression levels of these genes. We recently reported a p53-regulated posttranscriptional modification of many proteins through the transcriptional regulation of PADI4 (ref. 10) that converts an arginine residue in proteins to a citrulline residue, known as citrullination or deimination[11]. This reaction is mediated by the Ca^{2+}-dependent enzyme peptidylarginine deiminase. Citrullination of proteins causes a loss of positive charge and results in a significant biochemical effect. Bacterial infection has been indicated to stimulate PADI4-mediated hypercitrullination of histones in neutrophils and subsequently promoted highly decondensed extracellular chromatin structure called NET (neutrophil extracellular trap) that could capture and kill microorganisms[12,13]. On the other hand, citrullinated peptides/proteins are recognized as non-self proteins and subsequently activate immune systems. Citrullinated myelin basic protein in the brain is considered as a putative autoantigen associated with multiple sclerosis[14]. In addition, autoantibodies against citrullinated peptides derived

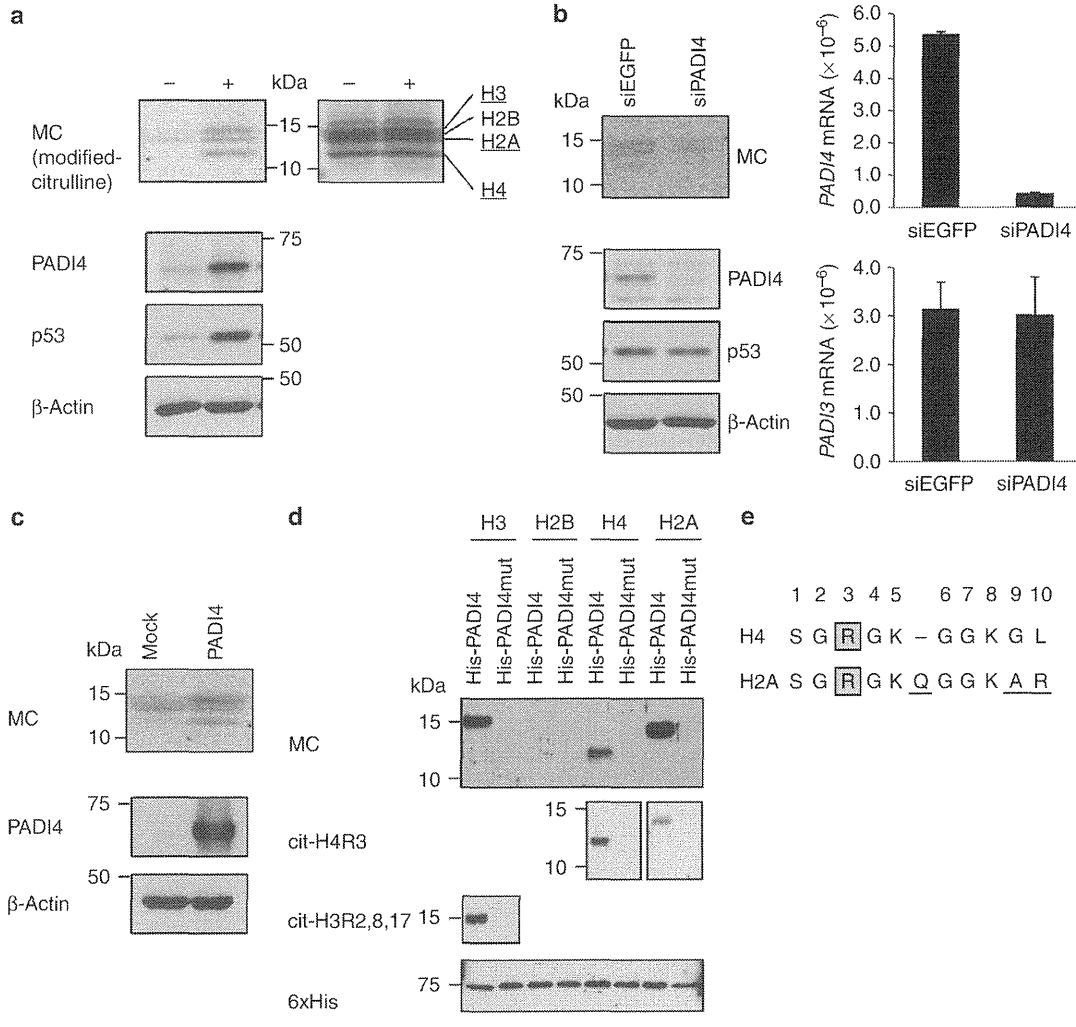


Figure 1 | Citrullination of core histones in the damaged cells. (**a**) Histones extracted from U-2 OS cells 48 h after ADR treatment (lane 2) were immunoblotted with anti-modified citrulline (anti-MC) antibody (left panel). U-2 OS cells without ADR treatment were used as a negative control (lane 1). CBB stain is shown in the right panel to ensure equal loading. Whole-cell extracts were subjected to immunoblotting with anti-PADI4, anti-p53 or anti-β-actin antibody. (**b**) siRNA against PADI4 was trasfected 10 h before DNA damage treatment. siEGFP was used as a control. Purified histones were immunoblotted with anti-MC antibody (upper left panel). Whole-cell extracts were subjected to immunoblotting with anti-PADI4, anti-p53 or anti-β-actin antibody (lower left panel). Quantitative PCR analysis of *PADI4* or *PADI3* expression (right panel). Error bars represent range ($n = 2$). *GAPDH* was used for normalization of expression levels. (**c**) Histones were extracted from PADI4-introduced or mock-transfected HEK293T cells, and immunoblotted with anti-MC antibody. Whole-cell extracts were subjected to immunoblotting with anti-PADI4 or anti-β-actin antibody. (**d**) *In vitro* citrullination followed by western blotting with anti-MC, anti-citrullinated histone H4R3 (cit-H4R3), anti-citrullinated histone H3R2,8,17 (cit-H3R2,8,17), or anti-His-tag (6×His) antibody. Each recombinant histone was incubated with wild-type or mutant PADI4 protein in the presence of 10^{-3} M of calcium at 37 °C for 1 h. (**e**) Alignment of amino-acid sequences of the N-terminal regions of histone H2A and H4.

from filaggrin or fibrin have been established as specific markers for rheumatoid arthritis[15–17]. Furthermore, we have revealed that a functional haplotype of the *PADI4* gene is associated with rheumatoid arthritis[18]. Thus, activated PADI4 function has been shown to be associated with the pathogenesis of rheumatoid arthritis through induction of the antigenicity of proteins. Here we investigate the physiological role of p53–PADI4 signalling in the DNA damage response by using PADI4 deficient mice as well as clinical cancer tissues. In this study, we revealed the crucial role of PADI4 in the apoptotic pathway and human carcinogenesis.

Results

Citrullination of core histones in response to DNA damage. As histones H3 and H4 were shown to be citrullinated by PADI4 (refs 19,20), we assessed whether core histones are citrullinated in response to DNA damage. Acid-extracted histones from adriamycin (ADR)-treated U-2 OS cells were immunoblotted with an antibody against a chemically modified form of citrulline (anti-MC antibody). Western blot analysis revealed the citrullination of three of four core histones in a PADI4-dependent manner (Fig. 1a,b). Ectopic expression of PADI4 also induced the citrullination of at least two core histones (Fig. 1c). To further characterize this modification, purified histones were treated with recombinant PADI4. Antibodies against specific citrulline residues (cit-H3R2,8,17 or cit-H4R3) as well as anti-MC antibody could detect PADI4-treated histones H3 and H4, respectively (Fig. 1d). Interestingly, anti-cit-H4R3 antibody cross-reacted with citrullinated H2A, probably owing to the high similarity between histones H4 and H2A (Fig. 1d,e).

p53–PADI4 citrullinates histone H4R3. Western blotting of core histone proteins purified from ADR-treated U-2 OS cells indicated citrullination of H4R3 by DNA damage (Fig. 2a), but the citrullination was blocked by small interfering RNA (siRNA) against p53 or PADI4 (Fig. 2b). Concordantly, ectopic introduction of PADI4 into cells induced the citrullination of H4R3 (Fig. 2c). Although PADI4 was reported to function as demethylase of mono-methyl-H4R3 *in vitro*[19,20], PADI4 activation did not affect H4R3 methylation status in our experimental system (Supplementary Fig. S1).

We also found the citrullination of histone H3 at its N terminus in non-treated U-2 OS cells and HEK293T cells (Supplementary Fig. S1), but these modifications were hardly detectable with anti-MC antibody, probably because of the low sensitivity of this antibody to the citrullination at its N terminus of histone H3 (Fig. 1a,c). Although citrullination of H3 was slightly increased by ectopic expression of PADI4 or by ADR treatment, siPADI4 did not affect H3-citrullination level (Supplementary Fig. S1). Because citrullination of H3 was observed in HEK293T and U-2 OS cells without any stresses, we suspect that other PADI family members might also be involved in H3 citrullination in these cells.

Localization of citrullinated histone H4 at fragmented nuclei. To investigate the physiological role of PADI4-mediated histone citrullination, we examined chromatin structure and subcellular localization of cit-H4R3 in PADI4-introduced HEK293T cells. Immunocytochemical analysis revealed the weak and irregular staining of nuclei with 4,6-diamidino-2-phenylindole (DAPI), a characteristic feature of chromatin decondensation (Fig. 3a). In contrast, cit-H4R3 displayed a dot-like staining pattern around the decondensed nuclei in cells expressing HA-PADI4 (Fig. 3a). Fractionation of chromatin obtained from PADI4-introduced HEK293T cells[21] suggested that cit-H4R3 was enriched in S1 (soluble in divalent metal; transcriptional active) and S2 (soluble in EDTA; transcriptional inactive) fractions, compared with P fraction (insoluble compact chromatin) (Fig. 3b). As apoptosis is associated with dynamic changes of nuclear structure, we examined the localization cit-H4R3 in apoptotic cells. At 48 h after treatment with ADR, cit-H4R3 was localized around the fragmented nuclei in U-2 OS cells, but it was not detected in the cells, additionally treated with siRNA against PADI4 (Fig. 3c,d). Similar results were observed in ADR-treated MCF7 cells (Supplementary Fig. S2a,b). In addition, cit-H4R3 expression was significantly correlated with the level of apoptosis induction demonstrated by the TdT-mediated dUTP nick

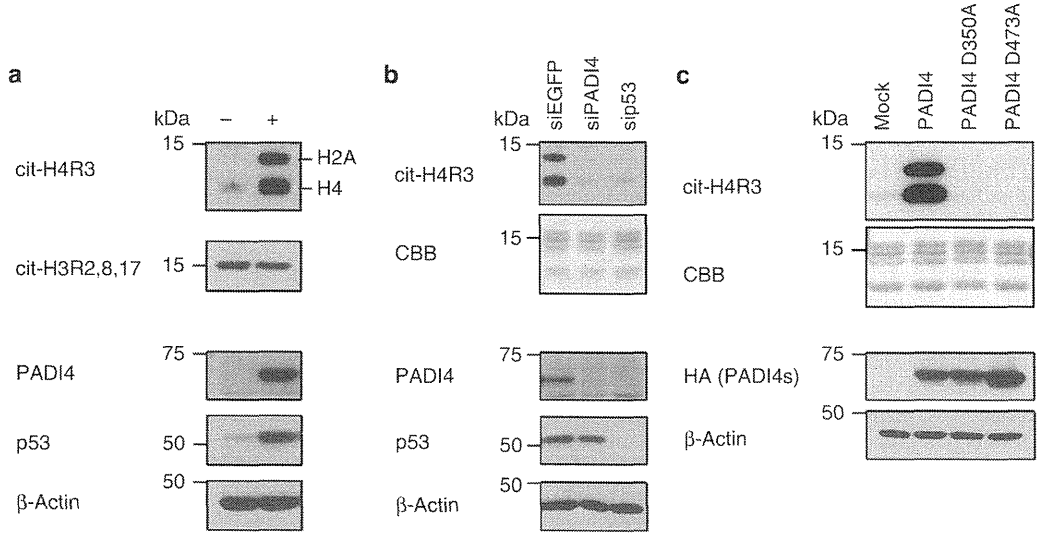


Figure 2 | Citrullination of histone H4R3 in the damaged cells. (**a**) Western blotting for histones extracted from ADR-treated U-2 OS cells (lane 2) using antibodies against citrullinated histone H4R3 (cit-H4R3) or citrullinated histone H3R2,8,17 (cit-H3R2,8,17). Whole-cell extracts were subjected to immunoblotting with anti-PADI4, anti-p53 or anti-β-actin antibody. U-2 OS cells without ADR treatment were used as a control (lane 1). (**b**) siRNA against PADI4 or p53 was transfected 10 h before ADR treatment. siEGFP was used as a control. Extracted histones were immunoblotted with anti-cit-H4R3 antibody. Whole-cell extracts were subjected to immunoblotting with anti-PADI4, anti-p53 or anti-β-actin antibody. CBB staining was shown to ensure equal loading. (**c**) Histones were extracted from HEK293T cells transfected with plasmid-expressing HA tagged-wild-type PADI4, mutant PADI4 (D350A or D473A), or mock and immunoblotted with anti-cit-H4R3 antibody. Whole-cell extracts were subjected to immunoblotting with anti-HA or anti-β-actin antibody.

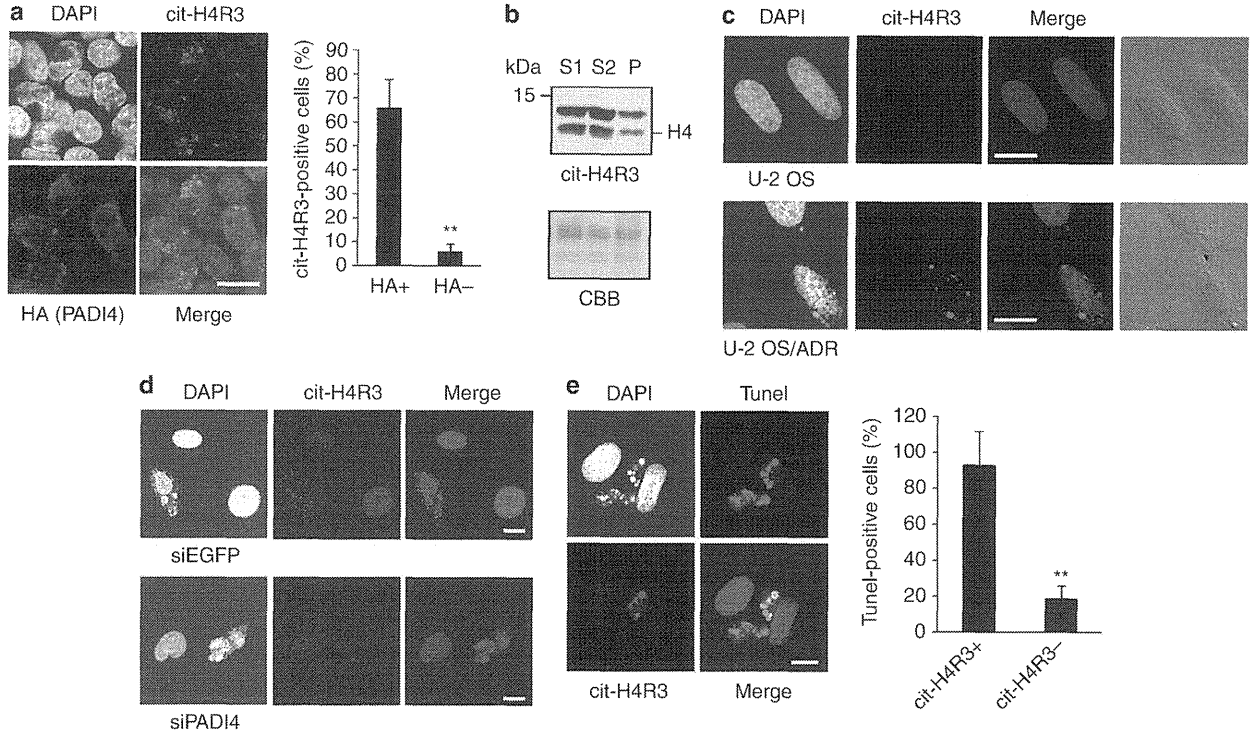


Figure 3 | Localization of cit-H4R3 around the fragmented nuclei during apoptosis. (**a**) Representative image of HA-PADI4-transfected HEK293T cells stained with anti-citrullinated histone H4R3 (cit-H4R3) antibody (Alexa Fluor 488) and anti-HA antibody (Aleza Fluor 594) (left panel).The proportion of cit-H4R3-positive cells stratified by HA-PADI4 expression was indicated (right panel). Error bars represent s.d. ($n=4$). $^{**}P<0.01$ by Student's t-test. Scale bars, 20 μm. (**b**) Chromatin fractionation followed by western blotting. Isolated nuclei from HEK293T cells transfected with PADI4-expressing plasmids were digested with MNase. Digested nuclei were fractionated into three fractions: transcriptionally active (S1), transcriptionally inactive (S2) and insoluble compact chromatin (P). Aliquots from S1, S2 and P fractions were separated by SDS-PAGE and then analysed by western blotting using anti-cit-H4R3 antibody. (**c**) Representative image of ADR-treated U-2 OS cells stained with anti-cit-H4R3 antibody (Alexa Fluor 488). Cells were treated with 2 μg ml^{-1} of ADR for 2 h and fixed 48 h later. Scale bars, 20 μm. (**d**) Representative image of ADR-treated U-2 OS cells stained with anti-citrullinated histone H4R3 antibody (Alexa Fluor 488). Each cell line was treated with 2 μg ml^{-1} of ADR for 2 h and fixed 48 h later. Each siRNA was transfected 10 h before the ADR treatment. siEGFP was used as a control. Scale bars, 20 μm. (**e**) Co-staining of U-2 OS cells with TUNEL (FITC) and anti-cit-H4R3 antibody (Alexa Fluor 594) at 48 h after treatment with 2 μg ml^{-1} of ADR for 2 h (left panel). The proportion of TUNEL positive cells stratified by cit-H4R3 expression was indicated (right panel). Error bars represent s.d. ($n=7$). $^{**}P<0.01$ by Student's t-test. Scale bars, 20 μm.

end labeling (TUNEL) assay in ADR-treated U-2 OS cells (Fig. 3e). These findings indicated that cit-H4R3 was associated with the decondensed soluble chromatin in apoptotic cells.

PADI4 promotes DNA fragmentation. In neutrophils, histone hypercitrullination is known to mediate NET formation that is associated with highly decondensed chromatin structures[22]. As chromatin decondensation was shown to increase the accessibility of DNase, we hypothesized that PADI4 promotes nuclear fragmentation. Micrococcal nuclease (MNase), which can digest chromatin at the linker DNA region into solubilized poly- and mono-nucleosome-size fragments, was used to evaluate the accessibility of DNA as well as chromatin compaction. MNase treatment of purified chromatin produced fragmented-nucleosomes more effectively in cells transfected with wild-type PADI4 than those with mutant PADI4 or mock (Fig. 4a). This result was confirmed by SDS–PAGE in which core histones from soluble nucleosome were stained with Coomassie brilliant blue (CBB) (Fig. 4b). ADR treatment also enhanced DNA cleavage (evaluated by histones in soluble nucleosome), but pretreatment with the siPADI4 remarkably suppressed MNase-mediated DNA digestion (Fig. 4c). Taking these observations into account, our findings revealed that PADI4 increased the accessibility of DNA and subsequently promoted the DNA fragmentation.

p53–PADI4 citrullinates Lamin C at its carboxy terminus. Next, we introduced PADI4 into HEK293T cells and performed immnunocytochemical analysis using anti-MC antibody. The results clearly indicated the citrullination of nuclear membrane proteins in cells expressing HA-PADI4 (Fig. 5a). Lamin family proteins are the main components of the nuclear lamina and are essential for various nuclear processes including apoptosis. Western blotting using nuclear extracts from PADI4-introduced HEK293T cells revealed the band shift of Lamin C protein, but not those of other Lamin proteins (Fig. 5b). Furthermore, we detected the citrullination of immunopurified Lamin C protein from PADI4-introduced HEK293T cells using anti-MC antibody (Fig. 5c). Because Lamins A and C share 566 amino acids of their N-terminal portions, six carboxy-terminal amino acids (VSGSRR) specific for Lamin C is likely to contain a target(s) of citrullination (Supplementary Fig. S3a). Hence, we constructed plasmids expressing various types of mutant Lamin C and examined the citrullination of each mutant protein. Western blotting with anti-MC antibody revealed that substitution of both 571st and 572nd arginine residues caused a significant reduction of protein citrullination (Fig. 5d). Accordingly, we raised an antibody against citrullinated Lamin C at these two arginine residues (Supplementary Fig. S3b). This antibody clearly recognized the citrullinated Lamin C in PADI4-introduced HEK293T cells as

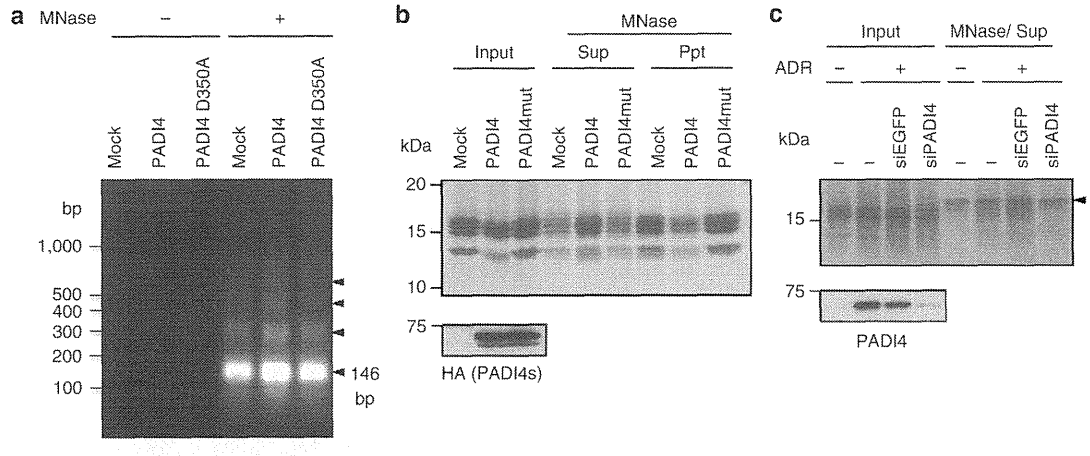


Figure 4 | Regulation of DNA cleavage by PADI4. (**a**) Chromatin was isolated from HEK293T cells transfected with PADI4 expressing plasmid. Chromatin extracted from equal numbers of cells was digested with MNase at 37 °C for 2 min. Soluble DNA released from MNase-treated chromatin were subjected to agarose gel electrophoresis. Arrow heads denote mono- and poly-nucleosomal DNAs. (**b**) Chromatin was isolated from HEK293T cells transfected with PADI4-expressing plasmid. Chromatin extracted from equal numbers of cells was digested with MNase at 37 °C for 2 min. Soluble core histones released from MNase-treated chromatin were subjected to SDS-PAGE. CBB staining for input chromatin is shown as controls. Input chromatin was subjected to immunoblotting with anti-HA antibody. (**c**) Chromatin was isolated from U-2 OS cells treated with ADR. Each siRNA was transfected 10 h before DNA damage treatment. siEGFP was used as a control. Chromatin extracted from equal numbers of cells was digested with MNase at 37 °C for 2 min. Soluble core histones released from MNase-treated chromatin were subjected to SDS-PAGE. CBB staining for input chromatin is shown as controls. Input chromatin was subjected to immunoblotting with anti-PADI4 antibody. Arrow head indicates MNase.

well as ADR-treated U-2 OS cells (Fig. 5e). In addition, immunocytochemistry revealed that the citrullinated Lamin C was localized predominantly around the fragmented nuclei in the damaged U-2 OS (Fig. 5f) and MCF7 cells (Supplementary Fig. S4a). Furthermore, cit-Lamin C expression was significantly correlated with the apoptotic cells in ADR-treated U-2 OS cells (Supplementary Fig. S4b), suggesting the involvement of citrullination of Lamin C in the nuclear fragmentation during apoptosis. Taken together, PADI4 could mediate the critical steps of apoptotic cell death through the citrullination of the core histones and Lamin C.

***Padi4*$^{-/-}$ mice exhibits apoptosis resistance.** To further investigate the physiological function of Padi4, we generated *Padi4*$^{-/-}$ mice (Supplementary Fig. S5). *Padi4*$^{-/-}$ mice were grossly normal in terms of survival, physical appearance and organ morphology. Then, we prepared messenger RNAs of thymuses from γ-ray-irradiated mice. The result of quantitative real-time PCR demonstrated that *Padi4* mRNA was significantly increased after DNA damage in wild-type mice but not in *Padi4*-deficient mice, compared with *p21*^{Waf1} as a positive control (Fig. 6a). We also detected the γ-ray-induced PADI4 expression in U-2 OS cells in a dose-dependent manner (Supplementary Fig. S6a). Immunohistochemical analysis of *in vivo* citrullination of H4R3 in *Padi4*$^{+/+}$ mice exhibited the significant signal of cit-H4R3 in thymus at 24 h after γ-ray irradiation, although cit-H4R3 was not detected in *Padi4*$^{-/-}$ mice (Fig. 6b). Concordantly, the numbers of TUNEL-positive and cleaved caspase-3-positive-cells were decreased in *Padi4*$^{-/-}$ mice (Fig. 6c,d). We also found that X-ray irradiation induced expressions of mouse *Padi4* and human *PADI4* in a p53-dependent manner (Supplementary Fig. S6b–d). Subsequently, we examined by western blotting, the level of cleaved caspase-3 in X-ray-irradiated thymus and found the significant reduction of cleaved caspase-3 in *Padi4*$^{-/-}$ mice (Fig. 6e), further supporting a critical role of Padi4 in the apoptotic process.

Frequent missense mutations of *PADI4* in cancer cell lines. To verify the biological and clinical significance of PADI4 in human carcinogenesis, we performed DNA sequencing of the *PADI4* gene in 22 breast, 21 colorectal and 37 lung cancer cell lines (Supplementary Table S1). We identified six missense mutations in six cell lines, among which the wild-type *PADI4* allele was lost in three cell lines (Supplementary Fig. S7a). As these mutations were not likely to affect the mRNA expression level (Supplementary Fig. S7b), we constructed plasmid vectors expressing either of six mutant PADI4 proteins and transfected either of these plasmids into HEK293T cells. Although T79 mutation did not affect enzymatic activity, five of six mutations lead to the remarkable decrease in enzymatic activity (Fig. 7a). Interestingly, T79 was not strictly conserved, while other five mutations were highly conserved among 14 mammalian species (Supplementary Table S2). These findings suggested the complete loss of PADI4 activity in MDA-MB-435S and SK-BR-3 cells. In concordance with this result, ectopic expression of p53 using adenovirus vector expressing p53 (Ad-p53) in MDA-MB-435S cells did not induce H4R3 citrullination, while H4R3 citrullination was observed in MDA-MB-231 cells in which wild-type PADI4 was retained (Fig. 7b). Moreover, citrullination of H4R3 was significantly associated with apoptosis induction of Ad-p53-infected MDA-MB-231 cells (Supplementary Fig. S7c), indicating a critical role of PADI4-mediated histone H4 citrullination in a p53-related pathway. Similarly, Ad-p53 infection enhanced citrullination of Lamin C only in cells in which both *PADI4* alleles are intact (Supplementary Fig. S7d). Because PADI4 was expected to function as a homodimer[23], our findings suggested that mutant PADI4 would exhibit dominant-negative activity. Although *PADI4* was mutated in LoVo cells, citrullination of Lamin C was remarkably increased in these cells. We assume that *PADI1*, *PADI2* and *PADI3* belonging to the PADI family might mediate the citrullination of Lamin C in LoVo cells because of relatively high levels of expression of these three members (Supplementary Fig. S7e).

PADI4 is a potential tumour suppressor. We examined the citrullination of H4R3 by tissue microarrays consisting of 309 NSCLC tissues (Fig. 7c). Clinicopathological analysis revealed that cit-H4R3 was associated with smaller tumour size ($P = 0.0136$ by Student's *t*-test, Table 1). We also found that 5-year survival rate in patients with cit-H4R3 expression (58.1%) was slightly longer than that in