Site-specific Phosphorylation Dynamics of the Nuclear Proteome during the DNA Damage Response*

Martin V. Bennetzen‡§, Dorthe Helena Larsen§¶‖, Jakob Bunkenborg‡, Jiri Bartek¶‖**, Jiri Lukas¶‖‡‡, and Jens S. Andersen‡§§

To investigate the temporal regulation of the DNA damage response, we applied quantitative mass spectrometry-based proteomics to measure site-specific phosphorylation changes of nuclear proteins after ionizing radiation. We profiled 5204 phosphorylation sites at five time points following DNA damage of which 594 sites on 209 proteins were observed to be regulated more than 2-fold. Of the 594 sites, 372 are novel phosphorylation sites primarily of nuclear origin. The 594 sites could be classified to distinct temporal profiles. Sites regulated shortly after radiation were enriched in the ataxia telangiectasia mutated (ATM) kinase SQ consensus sequence motif and a novel S*XX*Q motif. Importantly, in addition to induced phosphorylation, we identified a considerable group of sites that undergo DNA damage-induced dephosphorylation. Together, our data extend the number of known phosphorylation sites regulated by DNA damage, provides so far unprecedented temporal dissection of DNA damage-modified phosphorylation events, and elucidate the cross-talk between different types of post-translational modifications in the dynamic regulation of a multifaceted DNA damage response. ***Molecular & Cellular Proteomics 9:1314–1323, 2010.***

Ionizing radiation (IR)[1] covers a spectrum of energy-rich electromagnetic waves that are capable of ionizing molecules, thereby generating chemically reactive radicals. These radicals can in turn oxidize macromolecules and cause various types of DNA lesions. Among those, the DNA double strand breaks (DSBs) represent the most serious threat to genome integrity, which can be lethal or predispose to oncogenic transformation (1). To counteract the adverse effects of IR, cells are equipped with multifaceted genome surveillance pathways that are activated shortly after DNA damage, the DNA damage response (DDR) (2). The principal feature that fuels the genome surveillance pathways is a series of hierarchically organized post-translational modifications (PTMs) of proteins involved in DSB recognition, generation, and amplification of the primary signal and its propagation to various effectors that coordinate DNA repair with other key cellular events such as DNA replication, cell cycle progression, and gene expression (3).

The currently best studied DNA damage-induced PTM is protein phosphorylation, which in the case of the DSBs is triggered by the ATM kinase. Indeed, a large body of evidence documents that activation of ATM is the crucial upstream event required for the transmission and coordination of most, if not all, branches of the DSB-induced genome surveillance network (4, 5). The direct targets of ATM include proteins ranging from the initial sensing and enzymatic processing of DSB (such as the MRN complex), histones (H2AX), DDR signaling mediators (such as MDC1, 53BP1, RNF8, and RNF168), repair factors (*e.g.* BRCA1 and TopBP1), checkpoint transducers (*e.g.* CHK2), chromatin modifiers (*e.g.* KAP1), transcription factors and tumor suppressors (most notably p53), up to ATM itself. Indeed, autophosphorylation of ATM is crucial for dissociation of the inactive ATM dimers to active monomers and as such for the activity of the entire genome surveillance mechanism (6). Importantly, the recent systems biology approaches revealed that the current list of known DSB-induced phosphorylations is only a fragment of the immensely complex network of PTMs activated by geno-

From the ‡Center for Experimental BioInformatics, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, ¶Center for Genotoxic Stress Research, Danish Cancer Society, Strandboulevarden 49, DK-2100 Copenhagen, Denmark, and **Institute of Molecular and Translational Medicine, Palacky University, Olomouc, Czech Republic CZ-775 15

Received, December 18, 2009, and in revised form, February 12, 2010

Published, MCP Papers in Press, February 16, 2010, DOI 10.1074/mcp.M900616-MCP200

[1] The abbreviations used are: IR, ionizing radiation; ATM, ataxia telangiectasia mutated; DDR, DNA damage response; DSB, double strand break; ERLIC, electrostatic repulsion hydrophilic interaction chromatography; GO, gene ontology; H3K79me, methylation of histone H3 at lysine 79; MRN, Mre11-Rad50-Nbs1; PTM, post-translational modification; SILAC, stable isotope labeling by amino acids in cell culture; IPI, International Protein Index; ATR, Serine/threonine-protein kinase ATR ((Ataxia telangiectasia and Rad3-related protein)); PNKP, Polynucleotide kinase-3′-phosphatase; DNAPKcs, DNA-dependent protein kinase catalytic subunit; PPIG, Peptidyl-prolyl cis-trans isomerase G; RBP, RNA binding protein; RALY, RNA binding protein, autoantigenic; MGMT, 0–6 methylguanine-DNA methyltransferase; SET domain, (Su(var) 3–9, Enhancer-of-zeste, Trithorax) domain; MLL, Histone-lysine N-methyltransferase; CDYL, Chromodomain Y-like protein.

© 2010 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org

toxic stress. For instance, a large scale proteomics analysis identified several hundred proteins targeted by ATM or ATM-related kinases (7), and the list of the DSB-induced phosphorylations will inevitably expand after including kinases such as CHK1 and CHK2 operating downstream of ATM. Thus, the key challenge in the contemporary DNA damage field is to find efficient means to sort the expanding collection of the DNA damage-regulated phosphorylations into physiologically relevant modules and functionally characterize these modules to gain more holistic and integrating insight into the key mechanisms that protect the integrity of mammalian genomes.

To address this issue, we applied advanced mass spectrometry-based proteomics combined with affinity enrichment of phosphopeptides. Such an approach has emerged as currently the most powerful means to study protein phosphorylations on a truly proteome-wide scale. Importantly, it has been recently demonstrated that quantitative mass spectrometry can provide temporal resolution of site-specific PTMs (8). Stimulated by these advances and unprecedented analytical possibilities, we reasoned that an important step in sorting out the complex DSB-induced phosphorylations would be a detailed subclassification of these PTMs along a defined temporal scale after DNA damage. To this end, we isolated nuclei from SILAC-labeled (9) cells at multiple time points after exposure to IR and quantified site-specific changes in phosphorylation using high resolution mass spectrometry. Here we report on the temporal profiles of 594 regulated phosphorylation sites on 209 proteins. Together with the accompanying bioinformatics analysis, these data uncovered several hitherto unknown aspects of the dynamics of the DDR and revealed novel links between the DSB-induced phosphorylation and other functional modules involved in genome surveillance and maintenance.

EXPERIMENTAL PROCEDURES

Cell Culture, IR Exposure, Harvesting, and Purification of Nuclei—Epstein-Barr virus-transformed B-lymphocyte cells, GM00130 (Coriell Cell Repositories), are suspension cells with an intact DNA damage apparatus. GM00130 cells were cultured in RPMI 1640 medium + Glutamax (Invitrogen) containing 15% dialyzed fetal calf serum, 100 units of penicillin, and 100 μg/ml streptomycin. Two time course experiments were performed as outlined in Fig. 1*A*. For SILAC experiments, 2 $\times$ 3 populations of GM00130 cells were SILAC-encoded with stable isotope-encoded lysine and arginine. Cells were grown for 7 days in L-lysine- and L-arginine-deficient medium supplemented with normal isotopes, [$^2H_4$]lysine/[$^{13}C_6$]arginine, or [$^{13}C_6$,$^{15}N_2$]lysine/[$^{13}C_6$,$^{15}N_4$]arginine, also referred to as K0/R0, K4/R6, and K8/R10. Cells were irradiated using an x-ray generator (Pantak HF160; 150 kV; 15 mA; dose rate, 2.18 grays/min) and mixed after incubating for the appropriate time prior to harvest. Time points and IR dose were chosen on the basis of initial Western blot experiments (supplemental Fig. S1). For each sample, 2 $\times$ 10^8 cells were harvested by centrifugation and resuspended in Buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol + full inhibitor range (1 mM DTT, 5 μg/ml aprotinin, 5 μg/ml leupeptin, 0.1 mM PMSF, 0.1 mM sodium orthovanadate, 10 mM β-glycerophosphate, 1 μM okadaic acid, 1 mM sodium fluoride)). Cells were lysed in 0.02% Triton X-100 for 1 min, and nuclei were pelleted and washed twice in Buffer A (the last wash without protease inhibitors) and frozen in liquid nitrogen (supplemental Fig. S2).

Peptide Preparation—Starting material was approximately 630 million cell nuclei for experiments A and B. Phosphatase inhibitors (sodium fluoride, β-glycerophosphate, and sodium orthovanadate), the nuclease Benzonase (Merck), and denaturant 6 M urea, 2 M thiourea in 50 mM ammonium bicarbonate were added to the purified nuclei. All steps were performed at room temperature to avoid carbamoylation of amines. Reduction of cysteines was performed with DTT (5 mM) for 30 min followed by alkylation with iodoacetamide (11 mM) for 20 min in the dark. The proteins were first digested at 1:100 protease:protein with Lys-C (Wako) for 3.5 h, diluted 4 times with 50 mM ammonium bicarbonate with phosphatase inhibitors, and then digested with trypsin (Promega) overnight. Formic acid (Sigma) was added to stop the digestion. The mixture was spun at 10,000 rpm for 10 min, and the supernatant was filtered through a 0.45-μm filter (Millex-HV).

Phosphopeptide Enrichment Using ERLIC and TiO_2—The protocol for ERLIC was essentially as described (10) with some modifications.[2] For separation, the following solvents were used: Buffer A1, 20 mM NH_4 formate in 70% ACN; Buffer B1, 20 mM NH_4 formate in 10% ACN; Buffer A2, 1.0 M NH_4 formate in 10% ACN; and Buffer B2, 0.3 M phosphoric acid in 10% ACN. Peptides were reconstituted in ERLIC-Buffer A1 (20 mM NH_4 formate in 70% ACN), and the pH was adjusted to 3.8 with formic acid. The column was PolyWAX LP resin (PolyLC; 5 μm, 300 Å) packed into a 1.0-ml Tricorn 5/50 column (GE Healthcare). 16 fractions were collected during a 24-min isocratic part followed by collection of 35 fractions during the gradient part of the chromatographic run; all fractions were dried in a vacuum centrifuge. Phosphopeptide enrichment of each fraction was performed by reconstituting each sample in 150 μl of loading buffer (80% ACN, 5% TFA) and loading onto TiO_2 microcolumns (GL Sciences; 1 mg of TiO_2 beads/column) (11). Four washing steps were carried out to remove non-phosphorylated peptides using the solvents 5% TFA; 5% TFA, 80% ACN; 0.1% TFA; and 0.1% TFA, 80% ACN. The flow-through were collected followed by three elution steps using the solvents 0.5% NH_4OH, 20% ACN; 0.5% NH_4OH, 40% ACN; and 0.5% AcOH, 80% ACN into one vial. The flow-through was subjected to an additional phosphopeptide enrichment by TiO_2 chromatography for fractions where >40% of the identified peptides were phosphorylated.

Mass Spectrometry—The TiO_2 column eluates were dried using a SpeedVac and C_{18} Stage tip-purified (12). MS analysis was performed on an LTQ-Orbitrap (13) (Thermo Fisher Scientific) connected to an Agilent 1200 nanoflow HPLC system (Agilent) using a nanoelectrospray ion source (Proxeon Biosystems). The peptides were separated by reversed phase chromatography using an in-house made fused silica emitter (75-μm inner diameter) packed with Reprosil-Pur C_{18}-AQ 3-μm reversed phase material (Dr. Maisch GmbH). Peptides were loaded in 98% solvent A (0.5% acetic acid) followed by a 100-min linear gradient to 50% solvent B (80% acetonitrile, 0.5% acetic acid). Survey full-scan MS spectra (*m/z* range, 350–1800; resolution, 60,000 at *m/z* 400) were acquired on an LTQ-Orbitrap XL (Thermo Fisher Scientific) followed by fragmentation of the three most intense multiply charged ions using multistage activation with excitation at nominal mass losses of 97.97, 48.99, and 32.66 Da from the precursor *m/z* (14). Ions selected for MS/MS were placed on a dynamic exclusion list for 45 s. Real time internal lock mass recalibration was used during data acquisition (15).

MS Data Analysis—All raw files were processed with MaxQuant version 1.0.12.16 (16) into centroided data and submitted to database searching with Mascot version 2.2 (Matrix Science) using the follow-

[2] T. Y. Low, personal communication.

ing parameters: machine, Orbitrap/FT-ULTRA; SILAC, triplet; fixed modification, carbamidomethyl (Cys); variable modifications, acetyl (protein N terminus), oxidation (Met), deamidation (NQ), Gln → pyro-Glu (N-terminal Gln), and phospho-STY (STY); database (searched), IPIhumanDecoy (fasta file date, May 6, 2008; 152,616 sequences); enzyme, trypsin/p + DP; MS tolerance, 7 ppm; MS/MS tolerance, 0.6 Da; maximum missed cleavages, 3; top MS/MS peaks per 100 Da, 6. Protein identification after the database search utilized the Identify module of MaxQuant with the following settings: maximum peptide false discovery rate, 0.01; maximum peptide posterior probability, 1; base peptide posterior rate on Mascot score, 7; minimum peptides, 1; minimum peptide length; 6; minimum unique, 1; maximum protein false discovery rate, 0.01; requantify, on; keep low scoring versions of identified peptides, on; match between runs, 2 min.

For protein quantitation, only unmodified peptides and peptides modified by acetyl (protein N terminus), oxidation (Met), deamidation (NQ), and Gln → pyro-Glu (N-terminal Gln) were used. Thus, phosphorylated peptides are not used for quantitation at the protein level. If a counterpart to a phosphorylated peptide was identified, this counterpeptide was also not used for protein quantitation. According to the protein group assignment done by MaxQuant (16), both razor and unique peptides were used for protein quantitation. A minimum of two ratio counts were required to perform protein quantitation. For quantitation of phosphopeptides, the least modified peptides were used.

Temporal Profile Analysis—Changes in protein expression levels during the experiment were tested by measuring 380 protein expression profiles based on non-phosphorylated peptides found at all five time points (supplemental Fig. S3) and by Western blot analysis of H2AX and SMC1 (supplemental Fig. S1). In the case that more than one protein IPI identifier was associated with a given identified peptide, only the leading protein IPI identifier was included in the gene ontology (GO) term analysis. For the calculation of expression profiles, ratios were transformed to "-fold change" and standardized into Z-statistics where ± 1 were used to avoid mathematical discontinuity in the expression profile. Thus, the ratio at 1 h was fixed as 1. Ratios with $X > 1$ at a given time point were transformed to $Y = X - 1$, and ratios with $X < 1$ at a given time point were transformed to $Y = -1/X + 1$. The Y values (including $Y = 0$ for 1 h) for each profile were then standardized into Gaussian $N(0,1)$ distributed Z-statistics.

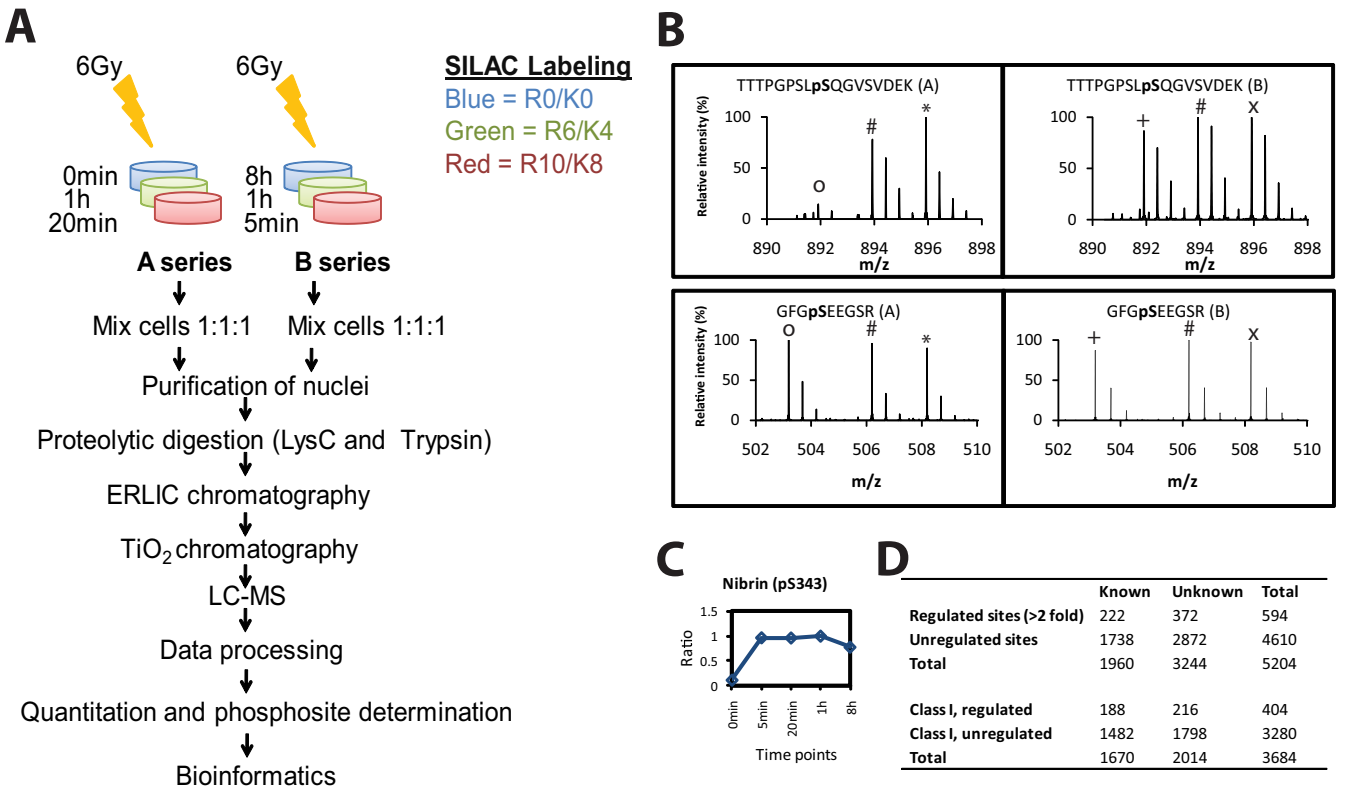
Validation of Phosphorylation Site Dynamics by Western Blotting—Cell lysates were prepared as above. Phosphorylation of indicated proteins was analyzed by Western blotting using phospho-Chk1 Ser^{317} (Cell Signaling Technology), phospho-SMC1 Ser^{966} (Abcam), SMC1 (Abcam), phospho-ATM Ser^{1981} (Rockland), γ-H2AX (Cell Signaling Technology), phospho-Nbs1 Ser^{343} (Cell Signaling Technology), and phospho-RSP6 Ser^{236} (Abcam).

RESULTS AND DISCUSSION

Identification and Classification of Phosphorylation Sites—To study the temporal dynamics of protein phosphorylation during the DDR, we applied high resolution mass spectrometry to identify and quantify phosphopeptides in proteins from unsynchronized GM00130 cells, an immortalized human B-cell line with an intact DDR. Time course experiments were performed using SILAC-encoded cell populations mixed in a ratio of 1:1:1 for each time series so proteins from different cell populations can be discriminated by mass (Fig. 1*A*). The radiation dose and time points were chosen on the basis of prior Western blot experiments, which defined conditions under which the dynamics of the known ATM-regulated phosphorylations were clearly discernible (supplemental Fig. S1). Proteins in nuclear lysates (supplemental Fig. S2) were digested with trypsin, and the resulting peptide mixtures were separated and enriched for phosphopeptides by ERLIC (10) and TiO_2 chromatography (11). The resulting peptide mixtures were subjected to nano-LC-MS analysis with multistage activated fragmentation (14) for phosphopeptide identification. Temporal profiles were determined from the intensity of the three isotope clusters, reflecting the relative amounts of phosphopeptide at each time point (Fig. 1*B*). The two experiments yielded a five-time point profile as illustrated for the Nbs1 phosphopeptide TTTPGPSLpS^{343}QGVSVDEK (where pS is phosphoserine) (Fig. 1*C*) and for additional known and novel phosphorylation sites associated with the DDR (supplemental Fig. S4).

A total of 7043 phosphorylation sites were identified at a false discovery rate of less than 1% using the MaxQuant software (16). For peptides containing more than a single Ser, Thr, or Tyr residue, the most likely site of phosphorylation was determined using a PTM scoring algorithm (17). Sites for further analysis were limited to those having a PTM localization probability greater than 0.75 (class I sites) and those quantified at all five time points (Fig. 1*D*). The majority of these sites (3244 of 5204; ~62%) were not annotated in the Swiss-Prot databases at the time of the MaxQuant data processing. Thus, this study contributes novel phosphorylation sites of primarily nuclear origin (84% of the regulated proteins have the Slim GO term "nucleus"). Phosphopeptides were considered regulated if their abundance for at least one time point was changed more than 2-fold compared with the 1-h reference time point. We categorized 4610 unregulated sites on 972 proteins and 594 regulated sites on 209 proteins (supplemental Table S1). It should be noted that 456 phosphorylation sites were regulated 1.5–2-fold, including sites associated with the DDR. Minimal changes in protein expression during the time course of the experiment were observed for most proteins as shown by the box plots in supplemental Fig. S3.

The mass spectrometry-derived time profiles were compared with Western blot analyses for selected regulated phosphorylation sites known to be involved in the DDR (Fig. 2, *A–C*). We observed consistency between the two methods for the phosphorylation sites Ser(P)^{343} (Nbs1), Ser(P)^{139} (H2AX), and Ser(P)^{236} (RSP6) representing clusters 1, 3, and 9, respectively (see below). By Western blot analysis, we also checked the profiles of Ser(P)^{1981} (ATM), Ser(P)^{317} (CHK1), and Ser(P)^{966} (SMC1) representing classical DNA damage-induced phosphorylation sites not covered by mass spectrometry possibly because they are found within peptides suboptimal for our analysis (Fig. 2*D*). These experiments validate that the profiles obtained by quantitative phosphoproteomics are representative for *in vivo* cellular signaling events. Moreover, testing regulated proteins against unregulated proteins, three Spike (18) annotated pathways were enriched: the ATM pathway ($p = 4.91e-11$), the DSB repair pathway ($p = 1.3e-10$), and the p53 pathway ($p = 0.00164$)



	Known	Unknown	Total
Regulated sites (>2 fold)	222	372	594
Unregulated sites	1738	2872	4610
Total	1960	3244	5204
Class I, regulated	188	216	404
Class I, unregulated	1482	1798	3280
Total	1670	2014	3684

FIG. 1. **Experimental strategy for characterization of protein phosphorylation dynamics.** *A*, three SILAC-encoded cell populations were irradiated with 6 grays (6 Gy) and maintained in culture for different lengths of time before the cells were harvested and mixed in a ratio of 1:1:1. Nuclei were purified from the combined cells followed by protein digestion and affinity enrichment of phosphopeptides by ERLIC and TiO_2 chromatography. The resulting phosphopeptides were analyzed by LC-MS/MS, and the data were processed by MaxQuant software and various bioinformatics methods. *B*, mass spectra of the regulated Nbs1 peptide TTTPGPSLpSQGVSVDEK and the non-regulated GFGpSEEGSR peptide from RNA-binding protein 8A from two separate experiments showing three isotope clusters representing separate time points (o, 0 min; #, 1 h; *, 20 min; +, 8 h; ×, 5 min). *C*, time profile for the Nbs1 phosphopeptide determined from isotope ratios. *D*, summary of the identified phosphopeptides.

(supplemental Fig. S5). These pathways constitute the core of the DDR and support the functional relevance of the identified phosphosites and their quantitative changes.

Phosphopeptide Dynamics during DNA Damage Response—To gain insight into the temporal regulation of the DDR, we clustered the time profiles using the MultiExperimentViewer software (19) using the hard clustering algorithm "*k*-means," which implies that a given profile can be associated with only one cluster. The transformed and standardized profiles for the 594 regulated phosphorylation sites were distributed to 10 clusters (Fig. 2*E*). We selected a group of 23 proteins that are linked to the DDR and examined the cluster distribution of their 105 regulated phosphorylation sites. The profiles for the majority of these sites were distributed to clusters 1 and 2 (Fig. 3), indicating that the majority of regulated sites are related to *early* responses to DNA damage as observed for *e.g.* Ser(P)^{343} on Nbs1 (Fig. 2*A*). The profiles for clusters 3–7 and 8–10 correspond to more transient and late responses, respectively. No obvious functional subgroup could be assigned to each of these clusters, but well characterized phosphorylation sites are in concordance with transient activation profiles. H2AX, for example, was rapidly phosphorylated on Ser(P)^{139} followed by a near complete reset to the initial phosphorylation level 8 h after radiation (Fig. 2*C*). One-third of the phosphosites were regulated by dephosphorylation, and proteins often referred to in the context of phosphorylation, *e.g.* 53BP1, SMC4, MDC1, and MCM4, are here associated with dephosphorylation events (Fig. 3 and supplemental Fig. S4). Thus, the DNA damage-induced dephosphorylation appears to have a more significant role than anticipated.

Multiple individual phosphorylation sites were typically identified for each protein. To examine whether these sites have distinct regulatory functions, we looked at their cluster distributions. Among the 209 proteins containing regulated sites, 14% were present in two or more clusters, including proteins found in both phosphorylation and dephosphorylation clusters (Fig. 3*B*). 53BP1, for example, contains 16 regulated sites (supplemental Fig. S4A) distributed mainly between clusters 1 and 2. The frequent observation of differential phosphorylation dynamics within the same protein indicates that the DDR involves complex molecular switching mechanisms.

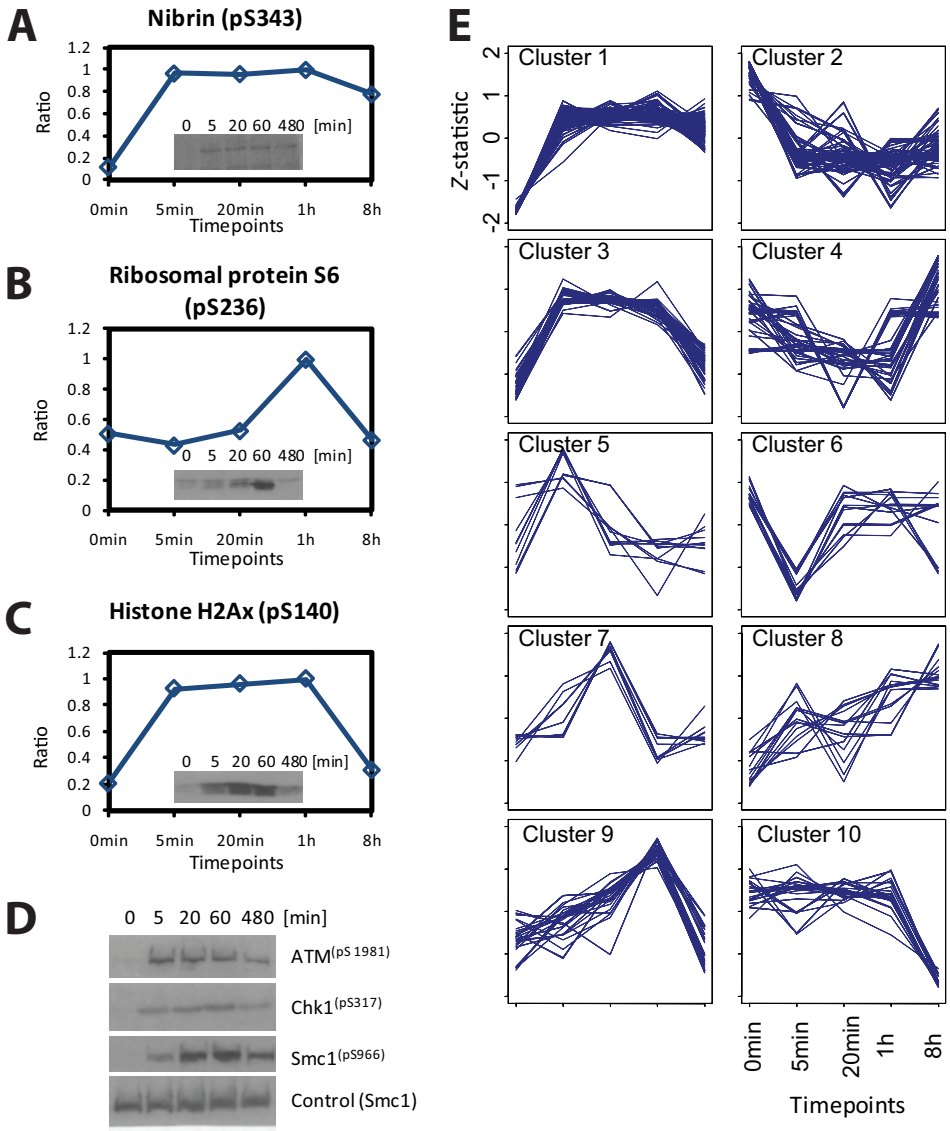
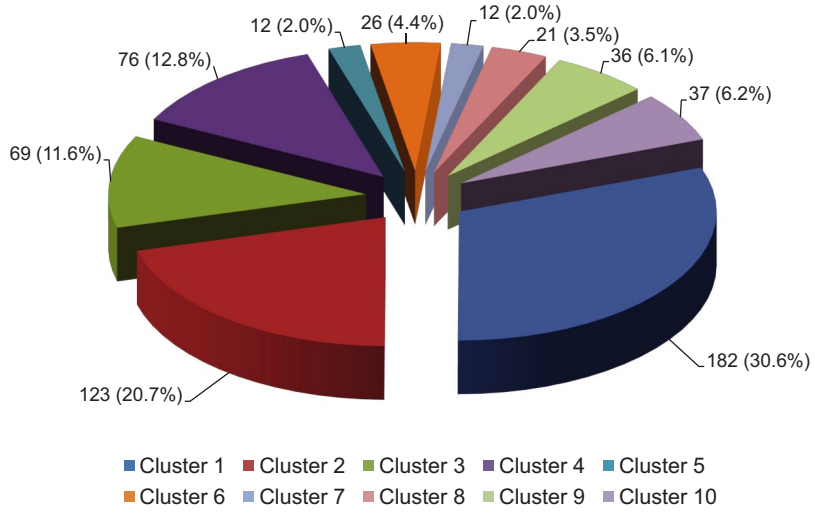


FIG. 2. **Validation and clustering of time profiles for phosphosites regulated during DNA damage response.** *A–C*, time profiles and corresponding Western blot analyses of selected phosphosites representing clusters 1, 9, and 3. *D*, Western blot analyses of phosphosites known to be induced by ionizing radiation. The SMC1 protein was used as loading control. *E*, clustering of time profiles for 594 regulated phosphosites.

Phosphorylation Consensus Sequences and Kinase Motifs—A recently reported large scale proteomics analysis of phosphosites modified in response to DNA damage has primarily taken advantage of targeted enrichment of ATM and ATR substrates using antibodies raised against the SQ and TQ consensus sequences (7). In contrast, here we applied a combination of ERLIC and TiO_2 affinity chromatography, a method expected to purify a more unbiased population of phosphopeptides (8). To assess the sequence motifs represented by these sites, we analyzed a 13-amino acid sequence window surrounding the central phosphorylated residue for regulated class I sites as compared with the unregulated class I sites using the Motif-X software (20). The consensus sequence for ATM kinase substrates was significantly overrepresented with 80 SQ motifs distributed on 67 proteins (Fig. 4, *A* and *C*). This shows consistency between our work and previous studies using similar approaches (7). The SQ motif was primarily found in clusters 1 and 3 (Fig. 4*D*), which represent sites phosphorylated immediately after the exposure of cells to ionizing radiation. This correlates with ATM being active in the early phases of the DDR. Most of the predicted ATM substrates represented in clusters 1 and 3 remained phosphorylated for the 8-h time window investigated here.

Interestingly, a new potential DNA damage-related phosphorylation motif, S*XX*Q, was overrepresented for 24 regulated class I sites (Fig. 4*B*), having a cluster distribution similar to that observed for the ATM motif (Fig. 4*D*). Using NetworKIN (21), no specific kinase family could be assigned to this motif

A



B

Protein / Cluster	1	2	3	4	5	6	7	8	9	10
ATM	4									
ATM-2	7									
BRCA1	3									
H2AFX			2							
MCM2	10	2								
MCM4		1								
MCM6	1									
MDC1	4	4								
MRE11A				3						
NBN	5							2		
NUMA1	2								2	2
PRKDC	1									
RAD23B		4								
RAD50								1		
RAD51AP1	1									
RAD9A				3						
RFC1	3		1							
SMC1A	3									
SMC3	5									
SMC4		2		1						
TOP2B	3			3						
TP53BP1	5	6			3			2		
XPC	4									

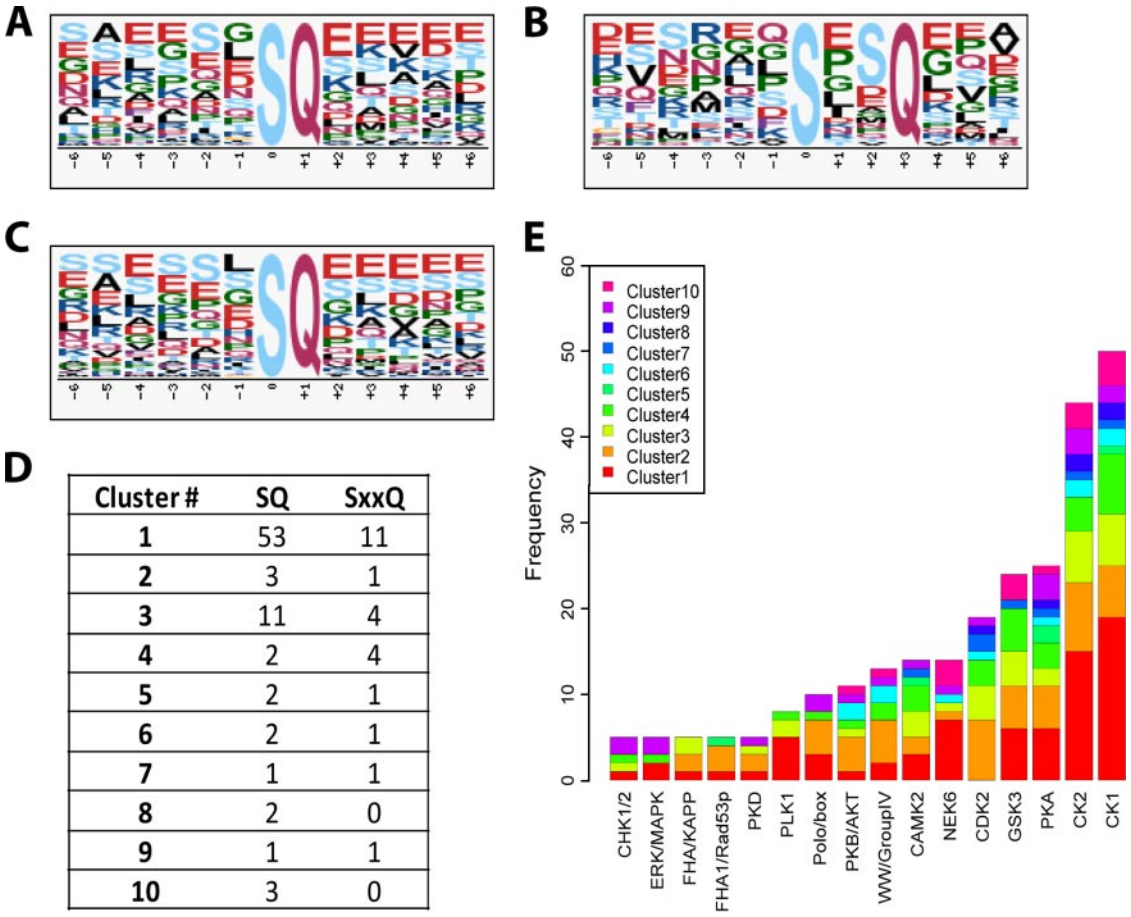
FIG. 3. **Cluster distribution of regulated phosphosites.** *A*, the number of phosphosites associated with each clusters. *B*, cluster distribution of known DNA damage response proteins and regulated sites associated with more than one cluster. The *numbers 1–10* denote the clusters, and the color indicates phosphorylation (*green*) and dephosphorylation (*red*). BRCA1 = Breast cancer type 1 susceptibility protein; NBN = Nibrin (NBS1); PRKDC = DNA-dependent protein kinase catalytic subunit ($DNAPK_{cs}$); XPC = DNA repair protein complementing XP-C cells.

by statistical means or by considering contextual features. It is thus possible that ATM itself participates in phosphorylating at least some of these residues.

Analysis of additional kinase motifs associated with all of the 404 regulated class I phosphorylation sites predicted that a broad spectrum of kinase families are responsive to DNA damage. More than half of the regulated sites (64%) can be assigned to one of the 24 kinase motifs contained in the Phosida database (22) of which 16 kinase motifs are represented by at least five phosphorylation sites among all regulated sites (Fig. 4*E*). The cluster distribution of the sequence motifs indicates that substrates of CK1, CK2, NEK6, and PLK1 are primarily phosphorylated following IR exposure, whereas substrates with PKB/AKT, WW Group IV domain binding and CDK2 motifs are mostly dephosphorylated. The inhibition of the latter kinase is in agreement with the reported silencing of cyclin-dependent kinases during the DNA damage-induced cell cycle checkpoints (2).

Functional Analysis and Clustering of Proteins Containing Regulated Phosphosites—To elucidate the global impact of the DDR on the nuclear proteome, we tested the 209 proteins containing regulated phosphorylation sites for GO terms (23) overrepresented relative to proteins containing unregulated phosphorylation sites. As expected, overrepresented GO terms were observed for proteins related to DNA damage and cell cycle checkpoints and to DNA repair processes (supplemental Fig. S6). Overrepresentation was also observed for the GO terms gene expression, transcription and RNA synthesis, nucleoside and DNA synthesis, DNA replication, and DNA metabolism as well as chromosome biogenesis. Regulation of these processes is consistent with repair of damaged DNA, transcriptional activation of genes encoding DNA repair and stress survival factors, and inhibition of genes encoding proteins not required during the DDR. Overrepresentation of the GO terms nucleic acid and DNA binding implies that many of the regulated proteins might directly bind DNA, probably at the site of damage or at regulatory transcription sites.

Using the Cytoscape plugin BINGO (24, 25), 118 GO terms were found to be significantly overrepresented when the regulated proteins in each cluster were tested relative to the human genome under the requirement that at least five proteins in at least one cluster share a given GO term (supplemental Fig. S7). The overrepresented proteins are involved in gene expression, DNA damage, cell cycle, and RNA processing, indicating that all of these processes are important for the DDR. Overrepresented GO terms were found mainly in clusters 1 and 2. A total of 81 significant GO terms were shared between the two clusters, indicating that these biological processes are regulated by both phosphorylation and dephosphorylation events.



Cluster #	SQ	SxxQ
1	53	11
2	3	1
3	11	4
4	2	4
5	2	1
6	2	1
7	1	1
8	2	0
9	1	1
10	3	0

FIG. 4. **Phosphorylation consensus sequences and kinase motifs.** *A–C*, consensus sequences overrepresented among sites regulated >2-fold (*A* and *B*) and 1.5–2-fold (*C*) as compared with unregulated sites during the DNA damage response. *D*, cluster distribution of the SQ and S*XX*Q phosphosite sequence motifs. *E*, distribution and frequency of best assigned kinase sequence motifs for class I sites derived from the Phosida database. Only motifs with five or more counts among all regulated sites are shown in *E*. CHK1/2 = Serine/threonine-protein kinase Chk1/Chk2; ERK/MAPK = Extracellular signal-regulated kinase/Mitogen-activated protein kinase; FHA/KAPP = Motif known to bind to Forkhead-associated domain found in kinase-associated protein phosphatase; FHA1/Rad53p = Motif known to bind to Forkhead-associated domain found in Rad53p; PKD = Protein kinase D; PLK1 = Polo-like kinase 1; PKB/AKT = Protein kinase B/alpha serine/threonine-protein kinase; WW/GroupIV = Motif known to bind to Goup4 of WW domains; CAMK2 = Calcium/calmodulin-dependent protein kinase type II; NEK6 = Never in mitosis A-related kinase 6; CDK2 = Cell division protein kinase 2; GSK3 = Glycogen synthase kinase-3; PKA = Protein kinase A; CK1/2 = Casein kinase 1/2.

Cross-talk between Protein Phosphorylation and Other Types of Modifications Controlling DNA Damage Response—To examine how phosphorylation regulates the biological processes discussed above, we interrogated each cluster for DNA, RNA, and protein modifiers using the DAVID software (26). Because the Pfam and InterPro domain features are included, a structural relationship with the functionality of the proteins in the functional clusters can be deduced (supplemental Table S2). The analysis highlighted five functional classes of proteins involved in post-transcriptional modifications; kinases, peptidases, and ubiquitin-modifiers; methyltransferases; acetyltransferases and chromatin modifiers; and RNA processing proteins (Fig. 5 and supplemental Table S2). Identification of these modifiers suggests that phosphorylation-dependent events are linked to other types of modifications to modulate proteins, DNA, and RNA in response to DNA damage.

Regulated phosphorylation sites were identified for the kinases OSR1, PNKP, $DNAPK_{cs}$, and ATM isoforms 1 and 2 as well as PRP4, TLK1, and CDC2L5, all reported to play a part in the DDR with the exception of OSR1 and PRP4 (7, 27–30). The first five mentioned are associated with cluster 1, indicating early phosphorylation of those proteins. Several novel phosphorylation sites were identified for these kinases (supplemental Table S1) that may modulate kinase activity. Kinases not previously linked to DDR include PRP4, which regulates mRNA processing and splicing, and OSR1, which is generally involved in environmental stress like osmotic stress responses and probably actin cytoskeleton regulation (31). PRP4 becomes dephosphorylated on 18 sites, highlighting the importance of controlling RNA processing at the onset of the DDR. This is further substantiated by the identification of regulated phosphosites for 20 RNA processing and splicing-

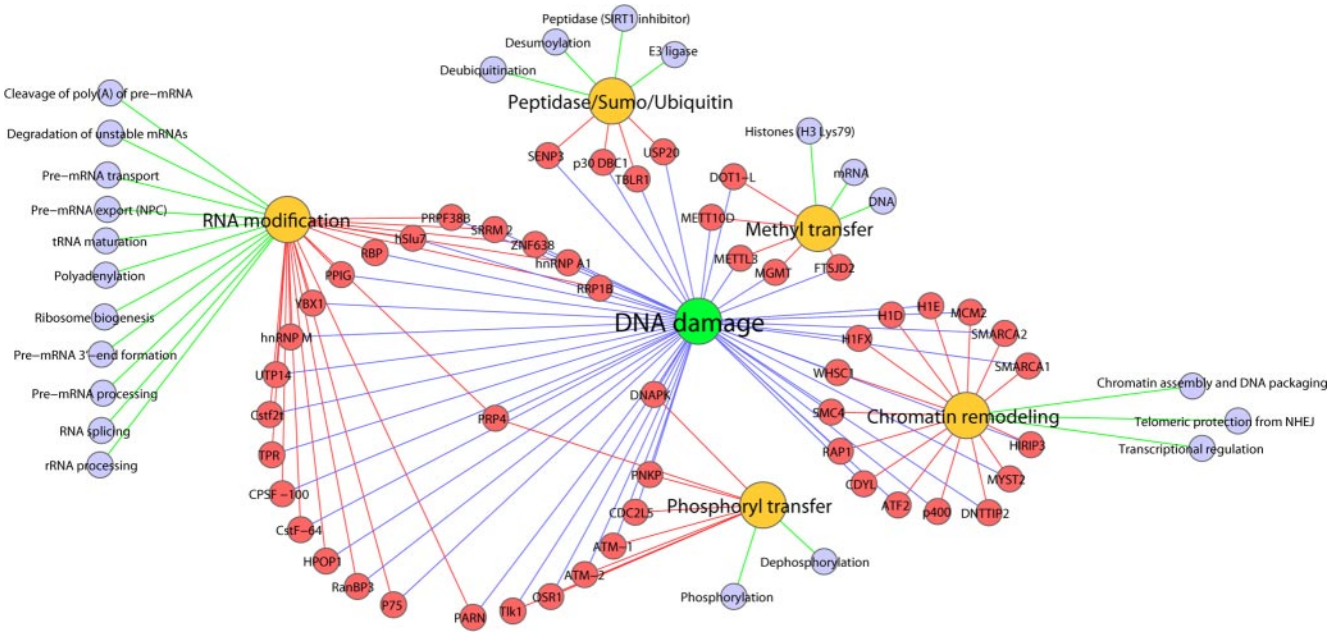


FIG. 5. **Diversifying the response: functional network of proteins with regulated phosphorylation sites and associated processes.** Color code for the *nodes* are as follows: proteins are *red*, modifications carried out by the proteins are *yellow*, and processes related to the proteins are *blue*. *Sumo*, small ubiquitin-like modifier. NHEJ = Non-homologous end joining; PNKP = Polynucleotide kinase-3′-phosphatase; DNAPK= DNA-dependent protein kinase d.

related proteins. Examples include the functional cluster of RNA splicing-associated proteins YBX1, PPIG, RBP (RALY), ZNF638, and SRRM2, all of which are associated with cluster 2, indicating that these splicing factors are dephosphorylated early during the DDR. These findings give novel insight into the role of RNA processing in the DDR, and we speculate that dephosphorylation of these factors may serve to regulate RNA processing for the period of DNA repair.

The functional class of methylation enzymes comprises the demethylase MGMT and the methyltransferases FTSJD2, METT10D, METTL3, and DOT1-L, all associated with cluster 1 (early phosphorylation) except for DOT1-L, which is associated with cluster 4 (early dephosphorylation). We find it intriguing that these enzymes are regulated by phosphorylation in response to DNA damage because methylation of histone tails is known to affect chromatin structure and accessibility. DOT-1L is a lysine histone methylase that in contrast to most other histone methylases does not contain a SET domain. It functions by methylating histone H3 (H3K79me), preferably when the histones are bound to DNA (32). Methylation of histone H3 activates transcription but has also been reported to modulate the recruitment of 53BP1 in response to DNA damage (33, 34). MGMT is a DNA demethylase that functions by demethylating DNA modified following genotoxic stress (35), and METTL3 has previously been shown to be phosphorylated upon DNA damage (7). METTL3 methylates internal adenosine residues in mRNAs, forming N^6-methyladenosine involved in mRNA splicing, transport, and translation (36). Moreover, two additional proteins, YY1 and MLL, might be linked to this functional class. YY1 is involved in methylation-based regulation of gene expression (37). YY1 is a transcriptional regulator that might activate or inhibit transcription depending on the cofactors it recruits, and YY1 is also known to be involved in the DDR by activation of DNA repair (38). The SET domain-containing MLL is by similarity a lysine histone methylase of the trithorax type causing transcriptional activation (39).

The functional class of acetyltransferases and chromatin modifiers has recently emerged as important for regulation of genotoxic stress. We found that these are associated with primarily *early* dephosphorylation and phosphorylation events (clusters 1–4) except for ATF2, which is phosphorylated late (cluster 7). Both histone acetylation and chromatin remodeling factors have important roles in the DDR by modulating chromatin surrounding DNA double strand breaks (40). We identified multiple regulated phosphorylation sites on proteins associated with acetylation of chromatin, including RAP1, SMARCC1, SMARCC2, SMARCA2, MYST2, CDYL, ATF2, and p400 (Fig. 5 and supplemental Table S2). These sites further support the existence of cross-talk between different post-translational modifications controlling the DDR. RAP1 (cluster 1) is known to bind telomere ends of non-histone-bound DNA in yeast followed by interaction with the sirtuins SIR3 and SIR4. The RAP1-SIR3-SIR4 complex facilitates recruitment of the deacetylase SIR2, which successively deacetylates histones, leading to chromatin condensation and thus transcriptional repression (41). In human cells, RAP1 has been shown to be phosphorylated during the DDR (7).

Human RAP1 is recruited to telomeres by interacting with TRF2, regulating telomere length. RAP1 also protects telomeric DNA from nonhomologous end joining at telomeric DNA ends (42), which may be the reason for its activation during the DDR. SMARCC1, SMARCC2, and SMARCA2 are components of the SWI/SNF complex involved in chromatin remodeling and transcriptional regulation (43). We found that SMARCC1 and SMARCC2 are dephosphorylated early (cluster 2) whereas SMARCA2 is phosphorylated early (clusters 1 and 3) during the DDR, indicating that the SWI/SNF complex is regulated by differential phosphorylation and dephosphorylation following DDR. p400 (early dephosphorylation; cluster 4) is a component of the NuA4 histone acetylase complex that comprises the catalytic subunit HTATIP/TIP60. The complex acetylates H2A and H4 causing transcriptional activation. Notably, it seems that the acetylase complex is essential for transcription of genes regulated by E2F1 and Myc involved in cell cycle progression and proliferation (44). Thus, dephosphorylation upon DNA damage might inhibit p400, thereby leading to transcriptional inhibition of Myc and E2F1 target genes and thus inhibiting cell cycle progression. The role and phosphorylation-dependent regulation of the modifying enzymes described above are not widely reported but are of potential importance in genotoxic stress responses. Taken together, the data show how early phosphorylation events may lead to signaling propagation to alter the structure of chromatin in response to DNA damage (supplemental Fig. S8).

In summary, our study of the temporal phosphorylation dynamics using quantitative mass spectrometry led to the identification of 372 novel phosphorylation sites regulated more than 2-fold during DDR. In addition and perhaps most importantly, we performed an unprecedented temporal analysis and subclassified regulated phosphorylation sites into 10 distinct temporal profiles, reflecting their association with different functional modules. Analysis of consensus sequence motifs indicated that several kinase families respond to DNA damage, and the frequent observation of dephosphorylation events suggests that phosphatases play a more important part early in the DDR than hitherto reported. GO term analysis and functional clustering suggested regulation of RNA biogenesis and other biological processes beyond those related directly to DNA repair. Interestingly, the identification of regulated phosphorylation sites on several methyl- and acetyltransferases and additional chromatin remodeling factors suggests elaborate cross-talk between these post-translational modifications in the spatiotemporal regulation of the DDR.

Acknowledgments—Lasse G. Falkenby and Kristoffer T. G. Rigbolt are acknowledged for computational assistance during the data and bioinformatics analysis, Teck Y. Low is acknowledged for providing an optimized protocol for ERLIC, and Matthias Mann and Jürgen Cox are acknowledged for providing an early version of MaxQuant.

* This work was supported in part by the European Community's Seventh Framework Programme (FP7/2007–2013) under Grants HEALTH-F4-2007-200767 for APO-SYS and HEALTH-F4-2008-201648 for Prospects.

S This article contains supplemental Figs. S1–S8 and Tables S1 and S2.

§ Both authors contributed equally to this work.

‖ Supported by the Danish Cancer Society, Danish National Research Foundation, Grant MSM6198959216, and Lundbeck Foundation Grant R13/A1287. Source of the MSM6198959216 grant: Czech Ministry of Education.

‡‡ To whom correspondence may be addressed. E-mail: jil@cancer.dk.

§§ To whom correspondence may be addressed. E-mail: jens.andersen@bmb.sdu.dk.

REFERENCES

1. Löbrich, M., and Jeggo, P. A. (2007) The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. *Nat. Rev. Cancer* **7,** 861–869
2. Bartek, J., and Lukas, J. (2007) DNA damage checkpoints: from initiation to recovery or adaptation. *Curr. Opin. Cell Biol.* **19,** 238–245
3. Zhou, B. B., and Elledge, S. J. (2000) The DNA damage response: putting checkpoints in perspective. *Nature* **408,** 433–439
4. Lavin, M. F. (2008) Ataxia-telangiectasia: from a rare disorder to a paradigm for cell signalling and cancer. *Nat. Rev. Mol. Cell Biol.* **9,** 759–769
5. Jackson, S. P., and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature* **461,** 1071–1078
6. Bakkenist, C. J., and Kastan, M. B. (2003) DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature* **421,** 499–506
7. Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., 3rd, Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S. P., and Elledge, S. J. (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316,** 1160–1166
8. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127,** 635–648
9. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1,** 376–386
10. Alpert, A. J. (2008) Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Anal. Chem.* **80,** 62–76
11. Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jørgensen, T. J. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* **4,** 873–886
12. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–670
13. Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M., and Graham Cooks, R. (2005) The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* **40,** 430–443
14. Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F., and Coon, J. J. (2004) A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal. Chem.* **76,** 3590–3598
15. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4,** 2010–2021
16. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26,** 1367–1372
17. Olsen, J. V., and Mann, M. (2004) Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 13417–13422
18. Elkon, R., Vesterman, R., Amit, N., Ulitsky, I., Zohar, I., Weisz, M., Mass, G.,

Orlev, N., Sternberg, G., Blekhman, R., Assa, J., Shiloh, Y., and Shamir, R. (2008) SPIKE—a database, visualization and analysis tool of cellular signaling pathways. *BMC Bioinformatics* **9,** 110
19. Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003) TM4: a free, open-source system for microarray data management and analysis. *BioTechniques* **34,** 374–378
20. Schwartz, D., and Gygi, S. P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat. Biotechnol.* **23,** 1391–1398
21. Linding, R., Jensen, L. J., Pasculescu, A., Olhovsky, M., Colwill, K., Bork, P., Yaffe, M. B., and Pawson, T. (2008) NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res.* **36,** D695–D699
22. Gnad, F., Ren, S., Cox, J., Olsen, J. V., Macek, B., Oroshi, M., and Mann, M. (2007) PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biol.* **8,** R250
23. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25,** 25–29
24. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13,** 2498–2504
25. Maere, S., Heymans, K., and Kuiper, M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21,** 3448–3449
26. Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4,** P3
27. Groth, A., Lukas, J., Nigg, E. A., Silljé, H. H., Wernstedt, C., Bartek, J., and Hansen, K. (2003) Human Tousled like kinases are targeted by an ATM- and Chk1-dependent DNA damage checkpoint. *EMBO J.* **22,** 1676–1687
28. Sunavala-Dossabhoy, G., Balakrishnan, S. K., Sen, S., Nuthalapaty, S., and De Benedetti, A. (2005) The radioresistance kinase TLK1B protects the cells by promoting repair of double strand breaks. *BMC Mol. Biol.* **6,** 19
29. Dephoure, N., Zhou, C., Villén, J., Beausoleil, S. A., Bakalarski, C. E., Elledge, S. J., and Gygi, S. P. (2008) A quantitative atlas of mitotic phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 10762–10767
30. Jilani, A., Ramotar, D., Slack, C., Ong, C., Yang, X. M., Scherer, S. W., and Lasko, D. D. (1999) Molecular cloning of the human gene, PNKP, encoding a polynucleotide kinase 3′-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. *J. Biol. Chem.* **274,** 24176–24186
31. Chen, W., Yazicioglu, M., and Cobb, M. H. (2004) Characterization of OSR1, a member of the mammalian Ste20p/germinal center kinase subfamily. *J. Biol. Chem.* **279,** 11129–11136
32. Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr. Biol.* **12,** 1052–1058
33. Huyen, Y., Zgheib, O., Ditullio, R. A., Jr., Gorgoulis, V. G., Zacharatos, P., Petty, T. J., Sheston, E. A., Mellert, H. S., Stavridi, E. S., and Halazonetis, T. D. (2004) Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* **432,** 406–411
34. Wysocki, R., Javaheri, A., Allard, S., Sha, F., Côté, J., and Kron, S. J. (2005) Role of Dot1-dependent histone H3 methylation in G1 and S phase DNA damage checkpoint functions of Rad9. *Mol. Cell. Biol.* **25,** 8430–8443
35. Dumenco, L. L., Allay, E., Norton, K., and Gerson, S. L. (1993) The prevention of thymic lymphomas in transgenic mice by human O6-alkylguanine-DNA alkyltransferase. *Science* **259,** 219–222
36. Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997) Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA* **3,** 1233–1247
37. Kim, J. D., Kang, K., and Kim, J. (2009) YY1's role in DNA methylation of Peg3 and Xist. *Nucleic Acids Res.* **37,** 5656–5664
38. Gordon, S., Akopyan, G., Garban, H., and Bonavida, B. (2006) Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* **25,** 1125–1142
39. Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D., and Hess, J. L. (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol. Cell* **10,** 1107–1117
40. van Attikum, H., and Gasser, S. M. (2009) Crosstalk between histone modifications during the DNA damage response. *Trends Cell Biol.* **19,** 207–217
41. Lustig, A. J. (1998) Mechanisms of silencing in Saccharomyces cerevisiae. *Curr. Opin. Genet. Dev.* **8,** 233–239
42. Bae, N. S., and Baumann, P. (2007) A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends. *Mol. Cell* **26,** 323–334
43. Yudkovsky, N., Logie, C., Hahn, S., and Peterson, C. L. (1999) Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators. *Genes Dev.* **13,** 2369–2374
44. Doyon, Y., and Côté, J. (2004) The highly conserved and multifunctional NuA4 HAT complex. *Curr. Opin. Genet. Dev.* **14,** 147–154