Modulation of the chromatin phosphoproteome by the Haspin protein kinase

MAIOLICA, Alessio, et al.

Abstract

Recent discoveries have highlighted the importance of Haspin kinase activity for the correct positioning of the kinase Aurora B at the centromere. Haspin phosphorylates Thr3 of the Histone H3 which provides a signal for Aurora B to localize to the centromere of mitotic chromosomes. To date Histone H3 is the only confirmed Haspin substrate. We used a combination of biochemical, pharmacological and mass spectrometric approaches to study the consequences of Haspin inhibition in mitotic cells. We quantified 3964 phosphorylation sites on chromatin-associated proteins and identified a Haspin protein-protein interaction network. We determined the haspin consensus motif and the co-crystal structure of the kinase with the Histone H3 tail. The structure revealed a unique bent substrate binding mode positioning the Histone H3 residues Arg2 and Lys4 which are adjacent to the Haspin phosphorylated threonine into acidic binding pockets. This unique conformation of the kinase-substrate complex explains the reported modulation of haspin activity by methylation of Lys4 of the histone H3. In addition, the identification of the structural […]

Reference


DOI : 10.1074/mcp.M113.034819
PMID : 24732914

Available at:
http://archive-ouverte.unige.ch/unige:38573

Disclaimer: layout of this document may differ from the published version.
Modulation of the Chromatin Phosphoproteome by the Haspin Protein Kinase*

Alessio Maiolica‡, Maria de Medina-Redondo§, Erwin M. Schoof¶¶, Apirat Chaikuad∥∥, Fabrizio Villa**, Marco Gatti‡‡, Siva Jeganathan§§, Hua Jane Lou¶¶, Karel Novy‡, Simon Hauri‡, Umut H. Toprak§, Franz Herzog∥∥∥, Patrick Meraldi§, Lorenza Penengo‡‡, Benjamin E. Turk∥∥∥, Stefan Knapp∥∥∥, Rune Linding¶, and Ruedi Aebersold‡"ab

Recent discoveries have highlighted the importance of Haspin kinase activity for the correct positioning of the kinase Aurora B at the centromere. Haspin phosphorylates Thr\(^3\) of the histone H3 (H3), which provides a signal for Aurora B to localize to the centromere of mitotic chromosomes. To date, histone H3 is the only confirmed Haspin substrate. We used a combination of biochemical, pharmacological, and mass spectrometric approaches to study the consequences of Haspin inhibition in mitotic cells. We quantified 3964 phosphorylation sites on chromatin-associated proteins and identified a Haspin protein-protein interaction network. We determined the Haspin consensus motif and the co-crystal structure of the kinase with the histone H3 tail. The structure revealed a unique bent substrate binding mode positioning the histone H3 residues Arg\(^2\) and Lys\(^4\) adjacent to the Haspin phosphorylated threonine into acidic binding pockets. This unique conformation of the kinase-substrate complex explains the reported modulation of Haspin activity by methylation of Lys\(^4\) of the histone H3. In addition, the identification of the structural basis of substrate recognition and the amino acid sequence preferences of Haspin aided the identification of novel candidate Haspin substrates. In particular, we validated the phosphorylation of Ser\(^317\) of the histone variant macroH2A as a target of Haspin kinase activity. MacroH2A Ser\(^317\) resides in a basic stretch of about 40 amino acids that is required to stabilize extranucleosomal DNA, suggesting that phosphorylation of Ser\(^317\) might regulate the interactions of macroH2A and DNA. Overall, our data suggest that Haspin activity affects the phosphorylation state of proteins involved in gene expression regulation and splicing.


Eukaryotic protein kinases (ePK)\(^1\) constitute a large family of enzymes that coordinate virtually any cellular processes by the phosphorylation of their target proteins at specific sites (1, 2). Active kinases often modulate the activity of other enzymes, including other kinases, thus amplifying and extending an initial signal that affect sometimes thousands of proteins (3). This creates a highly complex network of feedback and forward loops where multiple kinases can mutually influence each other’s activity. Kinases adopt three molecular strategies to select and specifically phosphorylate their substrates in the crowded environment of a cell (2). First, tight control of cellular kinase localization assures that only proteins present in the close proximity of the kinase can be phosphorylated; second, the kinase specific activity can be regulated via post-translational modifications or the recruitment of cofactor molecules; and third, the recognition of specific consensus motifs on substrates ensures that phosphorylation only occurs at the intended site or sites (2).

The Haspin kinase is a member of the ePK family that structurally diverges from most ePKs (1, 4). The Haspin kinase

\(^{1}\) The abbreviations used are: ePK, eukaryotic protein kinase; IF, immunofluorescence; GO, gene ontology.
domain displays structural features that have never been observed in other ePK family members (5, 6). Specifically, the possibility of activation loop phosphorylation, a frequent regulatory mechanism to control kinase activity, is absent in Haspin (5). Haspin is characterized by an active conformation that is stabilized by a hydrophobic lock of the helix αC inducing a stable S conformation of the structurally unique activation segment. These specific structural features also create a structurally diverse substrate binding site comprising a highly electronegative cleft for the histone H3 basic tails (5). Interestingly, the recognition of H3 has been shown to be modulated by methylation at H3 residue Lys4, thus coupling Haspin activity with epigenetic mechanisms of chromatin regulation (5). Histone H3 that is phosphorylated at Thr3 is so far the only well-characterized Haspin substrate (7). H3Thr3 phosphorylation (H3Thr3ph) is required for the localization of Aurora B at the centromere (8–10). Inactivation of Haspin catalytic activity by ATP mimetic inhibitors induces Aurora B centromeric de-localization, leading to a loss of phosphorylation in chromatin associated Aurora B substrates (11, 12). To date, apart from this well-characterized centromeric function of Haspin activity, the broader cellular functions of the kinase and the phosphorylation events that control these remain essentially unknown.

In this study, we used an integrated biochemical, proteomic, pharmacologic, and structural biology approach to study the Haspin kinase, its substrates and the cellular consequences of its activity. Specifically, we determined a new mode of kinase substrate binding and identified a Haspin kinase substrate recognition motif. We identified 3964 phosphorylation sites in chromatin-associated proteins, quantified their response to Haspin inhibition, and verified the mitotic phosphorylation of MacroH2A Ser137 (13) as directly dependent by Haspin activity. Altogether, our data suggest that Haspin regulates the phosphorylation of proteins involved in mechanisms that control gene expression, including the modifications of histones, and provide evidence for novel molecular effects of Haspin activity on mitotic chromatin.

EXPERIMENTAL PROCEDURES

Reagents—Chemicals of the highest available purity were purchased from Sigma-Aldrich unless otherwise stated. The Haspin inhibitor 5-iodotubercidin (5-ITu) (5, 6) was obtained from Cayman Chemical.

Haspin and CENP-T Proteins Production and Purification—Haspin452–798 catalytic domain fragment was purified as previously reported (5).

Sequence encoding N-terminal CENP T fragment (2–101) was PCR amplified and cloned in the first cassette of pGEX-6P-2b vector generated in-house. The vector of construct was confirmed by sequencing. For the protein expression, BL21(DE3) Rosetta cells containing the pGEX CENP T101 plasmid were grown in Terrific broth at 37 °C to an OD600 of ~0.8. Cells were induced for expression with the addition of 0.25 mM IPTG at 20 °C, and were incubated overnight. Cell pellets were resuspended in buffer A (25 mM Tris/HCl, pH 8.0, 300 mM NaCl, 10% [v/v] glycerol, and 2 mM dithioerythritol) plus protease inhibitor mixture. Then, cells were lysed by sonication, and cleared by centrifugation at 85,000 × g for 60 min. The cleared lysate was incubated with GST beads (GE Healthcare) pre-equilibrated with buffer A and incubated for ~4 h. Beads were washed three times with buffer A and the bound protein was cleaved from GST fusion with overnight incubation of PreScission protease. The eluate was applied to a Heparin HP (GE Healthcare) column pre-equilibrated in the same buffer. Elution of bound protein was achieved by a linear gradient from 300 to 1000 mM NaCl in 20 bed column volumes. The protein fractions were pooled and concentrated in 3-kDa molecular mass cut-off Vivaspin concentrators (Sartorius). The concentrated protein was subjected to final purification step on a Superdex 200 10/300 column (GE Healthcare) equilibrated in size-exclusion chromatography buffer (25 mM Tris/HCl, pH 8.0, 150 mM NaCl, and 1 mM DTE). Eluted protein fractions were analyzed by SDS-gel and fractions containing CENP T101 were pooled, concentrated, frozen and stored at −80 °C.

Cell Culture—HeLa cells used for chromatin purification were grown in Dulbecco’s modified Eagle’s medium (DMEM, Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS, Hyclone, Logan, UT) and antibiotics. 293T cells were grown in DMEM supplemented with 10% fetal bovine serum ( Gibco) and 2 mM L-Glutamine. Nocodazole was used at concentrations of 3.3 μM and 0.33 μM, for 1-hour arrest or 16-hours arrest, respectively.

Mitotic Chromatin Preparation—Chromatin samples were prepared as previously reported (14). Briefly, HeLa cells were arrested in mitosis with 0.33 μM nocodazole for 16 h and successively treated with the proteasome inhibitor MG132 for 30 min. The Haspin inhibitor 5-ITu was added for 1.5 h and successively mitotic cells were harvested by mitotic shake-off in cell growing medium. Cells were centrifuged at 1000 rpm 3 min, resuspend in 10 ml of cold PBS, and placed on ice for 45 min. The cell suspension was then centrifuged at 3000 rpm for 10 min in a 4 °C cold centrifuge. The cell pellet was resuspended in a hypotonic buffer containing 50 mM NaCl, 5 mM HEPES pH 7.4, 5 mM MgCl2, 0.5 mM CaCl2, 0.1 mM PMSF, and phosphatase and protease inhibitors (Roche). After, the cell suspension was immediately centrifuged for 10 min and the cell pellet was lysed in 4 ml of the hypotonic buffer with the addition of 0.5% Nonidet P-40 and 1 mM PMSF. Cell lysis was achieved using homogenizer (15 ml Wheaton) and 12 gentle complete strokes. Sodium deoxycholate (to 0.1% final concentration) was successively added to the cell solution and other 24 additional strokes were performed. The cell extract was then transferred to a 15 ml tube (BD, Biosciences) and spun down at 400 rpm for 10 min at 4 °C. The supernatant was collected and centrifuged through 4 ml of 40% sucrose cushion at 2400 rpm for 30 min in a 4 °C cold centrifuge. The mitotic chromosomes present in the pellet were first washed three times with 500 μl of lysis buffer and, successively resuspended in a solution containing 8 M UREA and 0.1% RapiGest (Waters, Milford, MA) and sonicated at 4 °C for 3 min for 3 min pause.

Immunofluorescence and Antibodies—For immunofluorescence (IF) analysis HeLa cells were plated onto acid-treated coverslips. Chromatin sample preparations for IF analysis were centrifuged at 2500 rpm for 30 min over a 2.5 ml, 40% sucrose solution onto an acid-treated coverslip.

Samples were fixed at room temperature for 10 min in PTEM buffer (0.25% TritonX-100, 20 mM PIPES pH 6.8, 1 mM MgCl2, 10 mM EGTA, and 4% formaldehyde). The following antibodies were used: human CREST (1/400; Antibodies Incorporated). Mouse anti-phospho-H3 Ser10 (1/1000, Abcam, Cambridge, MA) rabbit monoclonal anti-phospho-CENP-A Ser7 (1/2000; Upstate Biotechnology, Charlot- tesville, VA), rabbit anti-AuroraB (1/2000(15)). Cross-adsorbed secondary antibodies (Invitrogen) were used. 3D image stacks of mitotic cells were acquired in 0.2 μm steps using a 60x oil-immersion NA 1.3 objective on an Olympus DeltaVision microscope (Applied Precision,
Identification of Haspin Consensus Motif—Phosphorylation motif determination by peptide library array screening was performed as described by Mok et al. (17). The array consisted of 200 peptides with the general sequence YAXXXXX-S/T-XXXXAGKK(biotin), where X is an equimolar mixture of the 17 amino acids excluding Cys, Ser, and Thr, arrayed in a microtiter plate. In each well, the peptide mixture had one of the X positions fixed as one of the 20 unmodified amino acids or phosphoThr or phosphoTyr. Peptides (50 μM) were incubated with 200 or 400 nM Haspin in 50 mM Tris (pH 7.5), 10 mM MgCl2, 150 mM NaCl, 1 mM EDTA, 1 DTT, 0.1% Tween 20, and 50 μM ATP with 0.03 μCi/μl [γ-32P]ATP for 2 h at 30 °C. Aliquots (200 nl) were then spotted onto a streptavidin membrane, which was washed, dried, and exposed to a phosphor screen as described. Radiolabel incorporation into peptides was quantified using QuantityOne software (BioRad). Data from two runs were normalized and averaged to generate the position-specific scoring matrix for use with NetPhorest.

Computational Prediction of Haspin Substrates with the NetPhorest Algorithm—For the computational predictions of possible Haspin targets, we deployed an updated version of the NetPhorest algorithm, which can predict substrates for 222 kinases in the human kinome (17). As input data for the predictions, we used an in-house curated database (KinomeXplorer-DB), which collects known phosphorylation sites from public resources such as Phospho.ELM (18), PhosphoSitePlus (19), and PhosphoGRID (http://www.phosphogrid.org/). In the current version, there are 64,232 phosphorylation sites available, for each of which we predicted the most likely phosphorylating kinases. From these predicted kinases, we filtered the phosphorylation sites for which Haspin was a predicted kinase, and set a probability cutoff of 0.1, which should equate to a false-positive rate of maximum 10%. These results were subsequently used as input for follow-up experiments. From the perturbation experiments, the observed and quantified phosphorylation sites were used as input for NetPhorest predictions, in order to determine which phosphorylation sites are potential Haspin substrate sites based on their sequence motif. Again, predictions were filtered for best scoring kinases and a cutoff of 0.1 was applied to maintain high accuracy.

In Vitro Kinase Assays—Kinase reactions were carried out in a solution containing 50 mM Tris (pH 7.6), 10 mM MgCl2, 150 mM NaCl, 1 mM EDTA, and 1 mM DTT. The peptides used as Haspin substrates were chemically synthesized by On-SPOT synthesis (JPT) (20), histones H3 and H2A were obtained from Roche and New England BioLabs, respectively.

Final substrate concentrations in the assays were 250 μM ATP (5 μCi of γ-32P ATP for SDS page only), 5 μM of histones H3 and H2A (Roche), and 5 nM Haspin kinase. Reaction were initiated by the addition of ATP and carried out at 30 °C for 15 min and finally terminated by adding SDS/PAGE loading buffer or UREA 8 M for the protein samples that were successively analyzed by mass spectrometry. Assay reactions were then separated on a 15% SDS/PAGE gel and the amount of transferred phosphate was visualized by autoradiography or processed for mass spectrometry analysis.

Crystallization, Data Collection, and Structure Determination—Re-combinant Haspin (aa 465–798) was purified as described (6). The protein was concentrated to 15 mg/ml for crystallization studies. The protein was pre-incubated with 1 mM iodoacetamide and 3 mM histone H3 peptide (aa 1–11; ARTKQTARKSTY), and the complex was crystallized by sitting drop vapor diffusion at 4 °C using the reservoir condition containing 20% PEG 3350, 0.2 M KSCN. Suitable crystals were cryo-protected with mother liquor supplemented with 20% ethylene glycol and 2.2 M peptide before flash-cooled in liquid nitrogen. Diffraction data were collected in-house on a Rigaku FRE SuperBright source, and processed and scaled with MOSFLM and Scala from CCP4 suite (21). The complex crystals belonged to a primitive orthorhombic P212121 spacegroup with a unit cell dimension of a = 50.5, b = 79.1, c = 100.8 Å, α = β = γ = 90°. Structure determination was achieved by molecular replacement using PHASER (22) and the previously published Haspin coordinates (6) as a search model. The structure was subjected to iterative cycles of manual model rebuilding in COOT (23) alternated with refinement using REFMAC (24). Geometric correctness of the final model was verified with MOLPROBITY (25). Data collection and refinement statistics are summarized in supplemental Table S8.

Sample Preparation for Phosphoproteomics—Chromatin proteins were used as input for identification of Haspin binding partners and in vitro kinase reactions on peptide and proteins. Mass spectrometric analysis was performed with a LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, San Jose) equipped with a nanoelectrospray ion source (Thermo Scientific, San Jose). Mass spectra were acquired in a data-dependent manner, with an automatic switch between MS and MS/MS scans. High-resolution MS scans were acquired in the Orbitrap (60,000 FWHM, target value 106) to monitor peptide ions in the mass range of 350–1650 m/z, followed by collision-induced dissociation MS/MS scans in the ion trap (minimum signal threshold 150, target value 104, isolation width 2 m/z) of the five most intense precursor ions. The precursor ion masses of scanned ions were dynamically excluded from MS/MS analysis for 10 s. Singly charged ions and ions with unassigned charge states were excluded from triggering MS2 events.
**Database Searching**—Raw data were converted to the open mzXML format with ReAdW (version 4.3.1). mzXML files were searched by the SEQUEST via Sorcerer Software 4.2.0 and Mascot (version: 2.4.1) against UniProtKB/Swiss-Prot protein databases (release 2012.11, containing 20,243 proteins) concatenated with reverse sequences. For in silico digestion, trypsin was used as the protease and was assumed to cleave after lysine (K) and arginine (R) unless followed by proline (P). Three missed cleavage sites and one nontryptic terminus were allowed per peptide. The precursor ion tolerance was set to 50 parts per million (ppm), and fragment ion tolerance was set to 0.5 Da. The data were searched allowing phosphorylation (+79.9663 Daltons) of serine, threonine, and tyrosine as a variable modification and carboxy-amidomethylation of cysteine (+57.0214 Daltons) residues as a fixed modification. Finally, The identification results were statistically analyzed with the PeptideProphet algorithm (v.4.6) (27) and the results combined using the iProphet algorithm (28). In all the data sets presented, the FDR was maintained below 1%, this was based on the number of the decoy hits at the PeptideProphet cut-off score used. Accuracy of phosphorylation site location was determined by PTMprophet algorithm, the same mass accuracy specified for the database search was also used for the calculation of phosphorylation site localization probabilities.

The annotated spectra provided in supplemental Fig. S7 were generated using an in house developed python script and SpectraST (29) to create consensus spectra for all the identified phosphorylated peptides. Only fragment assignments with a mass error below or equal to 0.5 Da were annotated in each spectrum.

**Label-free Quantification**—For label-free quantification of the identified peptides, mzXML files were processed with OpenMS (30) suite to detect and extract ion signals. Each signal feature was matched with the best peptide assignment obtained after database searching. The generated intensity maps contained all aligned features together with the corresponding peptide sequences. The intensities maps of three different replicate experiments were analyzed to generate final fold-change ratios in peptide amounts between treated and untreated samples. Each experiment consisted of two consecutive injections for each cellular condition (treated with 5-iTu and not treated cells). Feature intensity values that were missing in the data matrix were addressed as follow. All peptide identifications with three or one missing value for each condition were excluded for further analysis. When instead two missing values were absent in the same biological condition we set a nominal lower bound values consisting in the minimum measured intensity peak over the background ions; we used $5 \times 10^{-10}$. For the statistical analysis, the intensities for peptide features were log-transformed; the fold change and the confidence of the minimum measured intensity peak over the background ions; we used $5 \times 10^{-10}$.

**Functional Annotation**—For functional annotation and Gene Ontology (GO) enrichment analysis of protein sets, we used the annotation tools GOrilla (32). The p values were calculated after FDR correction for multiple testing (Benjamini-Hochberg). STRING database (version 9.05) was used to predicted protein–protein interaction within Haspin binding partners (33). The STRING predictions were based on: experiment and databases; the minimum STRING score was set to 0.4. The protein networks were represented using in Cytoscape (v. 2.6.2) (34).

**Kinase-substrate Relationship Prediction**—To predict protein kinases affected by 5-iTu treatment, we used IGPS (version 1.0) and NetworkKin algorithms (version 2.0). Enrichment tests used the hypergeometric distribution as previously shown (35). As a background in these tests, we used peptide identifications that were not affected by 5-iTu treatments ($-3<FC<-3$, $p$ value $<0.05$). The $p$ value cutoff for enrichment was set at $5 \times 10^{-5}$.

**Generation of a Data Set of Haspin Modulated, Chromatin Associated Phosphorylation Events**—A Robust Cellular Assay to Detect Chromosomal Haspin Activity—We employed the small-molecule inhibitor 5-iodotubercidin (5-iTu) (supplemental Fig. S1) to inhibit Haspin kinase activity and to study the effect of its inhibition in human
Chromatin Phosphoproteome Modulation by Haspin Kinase

Fig. 1. Immunofluorescence analysis of the mitotic cells and the purified mitotic chromatin. A, Schematic representation of AuroraB recruitment at centromeres. Haspin phosphorylates Thr3 of the histone H3 (H3Thr3) at the centromere. Survivin a component of the CPC recognizes H3Thr3ph and recruits other subunits of the complex including Aurora B (adapted from (11)). B, 5-ITu has been shown to inhibit phosphorylation of H3Thr3 at concentration 1 μM. HeLa cells were arrested in mitosis for 16 h with nocodazole (0.33 μM) and successively 10 μM MG132 was added for 15 min to prevent mitotic exit. The cells were treated with 1 μM 5-ITU to block Haspin activity. After 1.5 h the mitotic cells were isolated by shake-off and successively processed for immunofluorescence. The cells were stained with DAPI (DNA, blue), CREST sera (kinetochores, red), and H3Thr3ph (green). C, Hela cells were treated as in B and stained with DAPI (DNA, blue), CREST sera (kinetochores, red), and CENP-A Ser7 phosphorylation (green). D, Hela cells were prepared as in B and the mitotic chromatin purified. Mitotic chromatin samples were gently centrifuged and stained with DAPI (DNA, blue), CREST sera (kinetochores, red), and H3Thr3ph (green). E, Mitotic chromatin samples were prepared as in D and stained with DAPI (DNA, blue), CREST sera (kinetochores, red), and AuroraB (green).

mitotic HeLa cells. A specificity screen performed on a panel of 138 kinases in vitro revealed the selectivity of 5-ITU toward Haspin and the members of Cik kinase family (11). To confirm efficient Haspin inhibition in cells, we monitored the phosphorylation of Thr3 of histone H3 by IF (supplemental Fig. S2). Briefly, the cells were arrested in mitosis with nocodazole and after 16 h the proteasome inhibitor MG132 was added to the medium to prevent exit from mitosis. Finally, the cells were treated for 1.5 h with 5-ITu concentrations ranging from 0.1 to 10 μM (supplemental Fig. S2). In agreement with previously published studies (38), we observed a drastic reduction of H3Thr3ph in cells treated with as little as 1 μM 5-ITU (5) (supplemental Fig. S2, Fig. 1A, 1B). The loss of H3Thr3ph is thought to prevent the recruitment of Aurora B to the centromere (Fig. 1A). However, under the experimental conditions used, centromeric phosphorylation of the Aurora B substrate CENP-A at Ser7 (Fig. 1C) was not affected. This was probably because of incomplete Aurora B delocalization from the centromeres (9, 11, 12, 39).

It has been reported in the literature that Haspin also localizes to condensed chromatin during mitosis (7), implying that its direct substrates and the Haspin dependent signaling cascade might affect the phosphorylation state of chromatin-associated proteins beside H3Thr3ph. To investigate the effect of Haspin inhibition on the state of phosphorylation of chromatin-associated proteins, we developed a robust protocol for the mass spectrometric identification and quantification of phosphorylated peptides on a chromatin fraction purified from mitotic cells (Fig. 1D, 1E). First we established conditions that prevented detectable de-phosphorylation of phosphoproteins during sample workup. Phosphorylated proteins can be rapidly dephosphorylated by phosphatases after cell lysis, if phosphatase inhibitors or denaturing conditions are not employed. To determine whether phosphatases affected the phosphorylation state of the samples during the chromatin enrichment protocol we monitored the phosphorylation state of H3Thr3ph by IF. Chromatin samples purified from Haspin inhibited mitotic cells and from cells mock treated with DMSO, respectively, were gently centrifuged onto microcopy slides and the specific protein epitopes were detected with antibodies (Fig. 1D, 1E). Mock treated cells substantially preserved the H3Thr3ph signal, whereas phosphorylation of the Haspin histone H3 target in mitotic cells treated with 5-ITU was barely detectable, validating the applied protocol (Fig. 1D). Interestingly, while centromere staining of Aurora B was strongly reduced upon Haspin inhibition, its localization on chromosome arms remained unchanged (Fig. 1E). In conclusion, we established and validated a protocol for the purification of mitotic chromatin from cells subjected to different pharmacological regimens in a state that largely preserves the phosphorylation state of the proteins. This allowed us to
directly study the effect of Haspin inhibition by 5-ITu on sub-strate phosphorylation in human cells.

Identification of Sites of Phosphorylation of Chromatin Associated Proteins—We used quantitative mass spectrometry to identify Haspin modulated phosphorylation sites on chromatin-associated proteins. Purified mitotic chromatin samples from cells treated with the Haspin inhibitor 5-ITu and untreated cells, respectively, were prepared using the conditions optimized above, digested with trypsin, and the phosphorylated peptides were enriched using immobilized metal affinity chromatography prior to analysis by LC-MS/MS (Fig. 2A). We identified 5822 phosphorylated sites on 1347 proteins (Fig. 2B, supplemental Table S1) at 1% FDR. Phosphorylation site localization probability was controlled using the PTM-prophet algorithm (supplemental Table S1). The distributions of the individual phosphorylated residues across the different hydroxyl amino acids and the number of phosphorylation sites per peptide identified on chromatin proteins were similar to those observed in other phosphoproteomic studies focusing on different cellular subcompartments (Fig. 2B) (35, 40). Further, we identified 1217 previously unreported phosphorylation sites in the PhosphoSitePlus database (19), corresponding to ~10% of phosphosites identified in this study. To estimate the limit of detection of our phosphopeptide measurements we related the identified phosphoproteins to estimates of their cellular concentration (41). The results indicate that the identified phosphoproteins spanned a range of abundance from over a million to less than 500 copies per cell (supplemental Table S1). More than 40% of the phosphoproteins identified in this study are core components of mitotic chromatin (42). Detailed analysis of the kinetochore, a macromolecular structure larger than 1 MDa that assemble on the centromere of mitotic chromosomes, indicated that the gentle conditions used for the purification of mitotic chromatin (Experimental Procedures) preserved even labile protein-protein interactions (43, 44) and thus retained even loosely associated proteins in the chromatin fraction. The kinetochore consists of more than 80 proteins (44). It has traditionally been sub-divided into three distinct structural layers: (1) the inner kinetochores which hosts proteins in close contact with the centromere proteins and the DNA; (2) the outer kinetochore that functions as a bridging platform between the inner kinetochore and the corona; and (3) the corona where the proteins transiently localize (43). In this study, we identified phosphorylation sites on proteins localized in all three layers, including the kinase Aurora B and several CENP proteins that localize to the inner kinetochore, Ndc80 and the kinases Plk1 and Mps1 along with structural components of the outer kinetochore and Cdc20 and CENP-E, BUBR1, proteins of the corona (Fig. 2C; supplemental Table S1).

In conclusion, we established and applied a strategy for the mass spectrometric detection of proteins and their phosphorylation sites from purified mitotic chromatin. To the best of our knowledge, this is the largest existing dataset of phosphorylation sites on chromatin proteins in human cells.

Quantification of Phosphorylation Sites in Chromatin-associated Proteins—To identify phosphorylation sites modulated by 5-ITu treatment in the dataset acquired above we related the precursor ion intensities of phosphopetides in 5-ITu treated and non-treated mitotic chromatin samples. We used the OpenMS quantification tool (30) and statistically evaluated the data as indicated in (45). To achieve robust phosphorylation site quantification we eliminated from further analysis those phosphopeptides that were inconsistently detected in replicate analyses of the same sample (described in the Experimental Procedures section). We confidently quantified 3964 phosphorylation sites from 1125 proteins (supplemental Table S2). We considered phosphopeptides as regulated if they showed a larger than sixfold change in signal intensities between inhibitor treated and not treated cells and a p value lower or equal to 0.05. Using these criteria we identified 258 and 324 phosphosites belonging to 338 proteins that were down- or up-regulated, respectively, in response to 5-ITu treatment (Fig. 2D, supplemental Table S2). To gain insights in the cellular functions affected by the Haspin inhibitor treatment, we performed gene ontology (GO) analysis using the GOrilla software (32) which indicated 66 GO terms significantly (p value <1E-7) enriched in the 338 phosphoregulated proteins (supplemental Table S3, Fig. 3E). We manually inspected the function of the proteins grouped under each enriched GO term. In total we could identify three major functional clusters of Haspin modulated phosphoproteins: (1) mitotic proteins; (2) RNA processing proteins; and (3) histones and chromatin modifier proteins (Fig 2E, supplemental Table S3).

Among the 574 phosphorylation sites modulated by Haspin inhibition, 25 sites mapped to mitotic proteins localizing to the centromere. These included: MIS18BP1, NSL1, CENP-E, CENP-F, Hec1, Borealin, and INCENP (supplemental Table S2–S3) as well as BUB1 and Aurora B kinases and the BUBR1 pseudokinase. As expected, we observed down-regulation of phospho-sites on three of the four-chromosome passenger complex (CPC) protein subunits identified in this study: Aurora B, Borealin, and INCENP (9, 39). These data are consistent with the notion that the absence of H3Thr<sup>3ph</sup> caused by 5-ITu treatment reduces both Aurora B localization at the centromere and the localization of the other subunits of the CPC complex (Fig. 1A–1E). Interestingly, the phosphorylation sites Ser<sup>593</sup> and Thr<sup>601</sup> on the BUB1 kinase were also strongly down-regulated upon 5-ITu treatment. Furthermore, about half of the phosphorylation sites we quantified after Haspin inhibition were increased compared with non-treated cells, suggesting indirect or compensatory effects. Up-regulated sites include Ser<sup>69</sup> and Ser<sup>62</sup> on the Ndc80 protein as well as phosphorylation sites on three proteins required for cytokinesis, namely Septin 7 and Septin 9 and Anillin.
Chromatin Phosphoproteome Modulation by Haspin Kinase

A

- Control
- Haspin inhibitor

- Mitotic Chromatin purification
- Mitotic Chromatin purification

- Phosphorylated peptides enrichment

- LC-MS/MS
- Identification and quantification

B

- Novel Phosphorylation sites
- Known Phosphorylation sites

- Thr Phosphorylations
- Ser Phosphorylations
- Tyr Phosphorylations

C

- Kinetochore microtubules

D

- Log2(Interactant/Component)
- Log2(Field Change)

E

- LogP Footprint

Molecular & Cellular Proteomics 13.7
The "RNA processing cluster" included subunits of the spliceosome, ribonucleoproteins, RNA binding proteins and transcription factors. Interestingly, we identified induced or reduced phosphorylation in several components of the spliceosome including, SRSF1, SRSF2, SRSF3, SRSF6, SRSF7, and SRSF9. The phosphorylation sites Ser^{199} and Ser^{201} on SRSF1 were down-regulated, whereas the phosphorylations on Ser^{234} and Ser^{238} on the same protein were up-regulated. Further, the data indicated that all the quantified sites on SRSF6, SRSF7, and SRSF9, were up-regulated. In contrast

---

Fig. 2. Inhibition of Haspin activity affects the phosphorylation status of chromatin associate proteins. A, Workflow describing the mass spectrometry approach used for quantification of phosphorylation sites on chromatin proteins. Chromatin samples were prepared as in Figure 1D and the phosphorylated peptides were enriched by titanium dioxide chromatography (TiO2). B, Overview of the identified phosphorylated sites on chromatin proteins. C, The protocol used for the purification of chromatin preserves protein-protein interactions within the kinetochore. We identified several phosphorylated proteins that are stably associated to the centromere as well as proteins that just transiently localize to the corona of the kinetochore. Either individual proteins or subunits of protein complexes identified in this study are labeled in orange some of the other known constituents of the kinetochore are colored in gray. D, Volcano plot representing quantification of the phosphorylated peptides. Logarithmic ratio of phosphorylated peptide intensities is plotted against the negative logarithmic p value. Blue and red circles represent phosphorylation sites up- or down-regulated respectively (log_{2}FC ≤ -3 and ≥ 3 and a p value ≤ 0.05). E, Functional classifications of the regulated phosphorylated proteins. GOriilla algorithm (32) was used to retrieve statistically significant enriched Gene Ontology (GO) terms within the set of regulated phosphorylated proteins. Numbers on top of each bar indicate the number of proteins in each identified GO cluster. Full list of the enriched GO terms is available in table SS3.

Fig. 3. The phosphorylation of histone macroH2A Ser^{137} depends by the Haspin activity. A, Mass spectrometry quantification of the histone macroH2A doubly charged phosphorylated peptide 135-ALK(Phospho)PSQKKPVSK-146 in chromatin samples upon Haspin inhibitor treatment. Areas of the MS1 peaks were extracted using Skyline (86). B, HEK293 cells were either transfected with a cDNA encoding Haspin (lanes 2 and 3) or the empty vector (Lane 1). Samples in lanes 3 and 4 were treated with the Haspin inhibitor 5-ITu (10μM) for 1.5 h before cell lysis. Protein extract were subjected to acid extraction and analyzed by SDS-PAGE and immunodecorated with anti-phosphorylated-Ser^{137} of histone macroH2A, anti-macroH2A, anti-HA or anti-phosphorylated-Ser^{10} of the histone H3.
phosphorylation sites on SRSF2 and SRSF3 were down-regulated. The complex regulation of phosphorylation on splicing proteins is likely caused by a combination of direct and indirect effects of Haspin inactivation.

The “chromatin modifier” cluster contained proteins involved in enzymatic processing of DNA such as DNA helicases and polymerases, as well as methyl and acetyl transferases (supplemental Tables S2, S3) and proteins that bind specific posttranslationally modified histones. Interestingly, we observed phosphorylation changes on a number of chromobox proteins, such as CBX1, CBX3, and CBX8 (supplemental Tables S2, S3). These proteins control epigenetic repression of chromatin by affecting PTMs on histones (46). In addition, we observed significant down-regulation of the phosphorylation sites on several histone isoforms upon 5-ITu treatment (supplemental Tables S2, S3). In contrast, Ser122 of the H2AX and H4 Ser28 (47) were up-regulated upon inhibitor treatment (Fig. 2E, supplemental Table S2).

Phosphorylation at Ser137 of the histone macroH2A was strongly down-regulated upon 5-ITu treatment (Fig. 3A). Recently, it has been reported that the macro domain of histone macroH2A controls the phosphorylation levels of Ser10 and Thr3 of histone H3 in human cells (48). To test whether macroH2A Ser137 phosphorylation was directly dependent on Haspin activity, we over-expressed Haspin in the presence or absence of the inhibitor 5-ITu and quantified Ser137 phosphorylation using a phospho-specific antibody (Fig. 3B). The results showed a drastic reduction of macroH2A phosphorylation in HEK293 cells upon treatment with 5-ITu (Fig. 3B). In contrast, Haspin over-expression led to a significant increase in macroH2A Ser137 and histone H3 Ser10 phosphorylation that again diminished upon 5-ITu treatment (Fig. 3B). In combination these results strongly suggest a direct dependence of histone macroH2A Ser137 phosphorylation on Haspin catalytic activity (Fig. 3B). In summary, mass spectrometric quantification of phosphorylation sites on chromatin-associated proteins upon 5-ITu treatment confirmed the importance of Haspin kinase activity toward the phosphorylation of centromeric proteins. Furthermore, our data indicate potential novel connections between Haspin-dependent signaling and processes involved in gene transcription.

**Kinase-substrate Enrichment Analysis (KSEA) Reveals Inactivation of Aurora B, CLK, and RSK Kinases upon 5-ITu Treatment**

To gain functional insights into the phosphorylation sites regulated in the quantitative phosphoproteomic data set described above, we performed a kinase-substrate enrichment analysis (KSEA) (35, 49). Specifically, we first used the iGPS (50) and NetworKin (51) kinase-substrate prediction algorithms to identify those kinases that have predicted substrates phosphorylation sites among the 3964 quantified phosphorylation sites. We then further statistically filtered this initial kinase-substrate matrix to determine those kinases that were significantly associated with the 5-ITu modulated phosphorylation sites. Substrates down-regulated upon 5-ITu treatment implied that the activity of the respective kinase(s) were reduced, for example, by the inactivation of Haspin or the direct Haspin dependent inactivation of a downstream kinase.

Both the iGPS and Networkin algorithms did not detect a significant association between the up-regulated sites and specific kinases. In contrast, for the down-regulated sites the two algorithms identified three kinase families that were significantly associated ($p$ value lower than 0.0005) with the phosphopeptides in that group: i) the Aurora family comprising Aurora B, Aurora A, and Aurora C kinases ($p$ value $2.79 \times 10^{-6}$); ii) the CLK family consisting of the kinases Clk1, Clk2, and Clk3 ($p$ value $2.86 \times 10^{-4}$); and iii) the RSK family consisting of p90RSK, RSK2, and RSK3 kinases ($p$ value $1.7 \times 10^{-4}$) (Fig. 4A, supplemental Table S4). The down-regulation of Aurora B substrates was expected because 5-ITu treatment prevents Aurora B localization to the centromere (see above), thus preventing its activity at the centromere (Fig. 4B, supplemental Table S5). CLKs are dual specificity protein kinases involved in pre-mRNA processing (52). Inhibition of CLK family kinases can be interpreted as a direct effect of 5-ITu treatment as reported by a specificity screen showing that 5-ITu inhibits in vitro Clk2 (94% of the kinase activity) and less efficiently Clk3 (50% of the kinase activity) kinases (11).

Eighteen phosphorylation sites were predicted as CLK substrates; more than 50% of those were from proteins involved in gene expression and splicing mechanisms (Fig. 4B, supplemental Table S6). The RSK family consists of a group of highly conserved Ser/Thr kinases that regulate a range of cellular processes,
including cell growth, cell motility, cell survival, and cell proliferation (S3). Interestingly, substrate prediction suggested that RSK activity mostly targets the histones and histone associated (Fig. 4B, supplemental Table S7). The inhibition of RSK is unlikely to be caused by off-target 5-iTu inhibition, because the concentration of the inhibitor used in this study does not significantly affect RSK kinases activity in vitro (1 μM 5-iTu reduces only of the 3% RSKs kinase activities) (11). Therefore, we consider the down-regulation of predicted RSK substrates as an indirect effect of Haspin inhibition.

In summary, KSEA after Haspin inhibition correctly highlighted downstream effects of 5-iTu treatments such as the inhibition of both Aurora B and CLK family kinases (9, 11, 39). Furthermore our data also shows a significant down-regulation of the RSK predicted substrates.

Identification of a Haspin Consensus Motif by Positional Scanning Oriented Peptide Library Screening (PS-OPLS)—Substrate recognition by kinases depends on different factors, including spatial proximity, site accessibility and the amino acid sequence motifs surrounding the phosphorylation sites (2). We determined a preferred Haspin substrate motif by positional scanning oriented peptide library screening (Fig. 5A) as described in (S4).

We purified a recombinantly expressed Haspin kinase domain and used it for the in vitro kinase reaction on degenerate peptide libraries. The autoradiography pattern indicated a preference for poly-threonine compared with poly-serine peptides (bottom left Fig. 5A) suggesting a preference of Haspin for threonine residues. Also, the phosphorylation of peptides carrying threonine residues in the different positions tested could be explained by the preference of Haspin for threonine residues (squared in blue Fig. 5A). Haspin was most strongly selective for peptides having an Arg residue at position P−1, and displayed a substantial preference for Ala and Val at the P−2 position and for Lys at the P+1 position. We therefore defined the preferred recognition motif for Haspin as A/V-R-T/S-K-(X-n)oD/E (Fig. 5A). Strikingly, this motif is in complete agreement with the only presently known Haspin phosphorylation sequence centered around Thr3 of the histone H3 (supplemental Fig. S3). In addition, we also noted that acidic amino acids were strongly disfavored by Haspin at multiple positions near the phosphorylation site (Fig. 5A squared in red). This observation is consistent with the negative net charge within the Haspin active site (supplemental Table S8). The kinase domain adopted a similar conformation when compared with the apo-structure (5). The first seven residues were modeled into the experimental density revealing that three substrate residues Ala1 (position −2 in the PS-OPLS), Arg2 (−1) and phosphoacceptor site Thr3 (0) were anchored deep within the substrate pocket. Surprisingly, the bound peptide adopted an unusually sharp 180° turn at Lys4 (+1) projecting the C-terminal tail outward (Fig. 5B). This is in contrast to binding modes of substrates of other kinase-substrate complexes, which typically exhibit an elongated linear conformation. As a consequence of this unique substrate conformation, residues Arg2 (−1) and Lys4 (+1) are positioned into deep hydrophilic pockets explaining the strong selection for these two residues and position −1 and +1 in the degenerated library peptide array. The only other residue preferentially selected in position +1 was a tyrosine, which may functionally replace the lysine forming a hydrogen bond with Asp707. The structure explains also the strong selection for small hydrophobic residues at position −2 in the PS-OPLS experiment. In this position the amino acid side chain is oriented toward a small surface cavity excluding residues with bulkier groups than alanine or valine.

Prediction and In Vitro Phosphorylation of Potential Haspin Substrates—The strength of the interactions that stabilizes the binding of the substrates with the kinase reflect the signal intensities of the phosphorylated peptides measured in the PS-OPLS array and indicate the residues in the recognition sequences that are particularly important in Haspin substrates. This information, in turn is useful to predict substrates of a kinase (55, 56).

We used the NetPhorest algorithm to identify candidate Haspin direct substrates in the fraction of the down-regulated phosphorylation sites measured on the chromatin associated proteins. The results were filtered based on NetPhorest probability higher than 0.1, which led to a false positive rate (FPR) of the Haspin substrate predictions lower than 10%. Using this score cutoff we identified 11 candidate sites phosphorylated by Haspin (supplemental Table S9). These proteins include the splicing factors SRSF1, SRSF2, SRSF10, and the PRP4 kinase, which controls RNA splicing (supplemental Table S9). To extend Haspin substrate prediction to proteins that were not identified in the chromatin data set presented here, we used the newly established Haspin consensus recognition motif to computationally predict potential Haspin substrates from phosphorylation sites identified in the literature. We employed the NetPhorest algorithm to query a large collection of more than 64,000 identified phosphorylation sites (Experimental Procedures). From these the algorithm predicted 2926 sites with a NetPhorest probability higher than 0.1 and a FPR less 10% (supplemental Table S10). We tested specificity of Haspin activity on 101 predicted substrates sites (supplemental Fig. S5) using an in vitro phosphorylation experiment. We chemically synthesized 101 peptide sequences selected within the predicted sites where Haspin was among the top
A. Identification of Haspin motif by positional scanning oriented peptide library screening. Haspin reveals a strong selection for Ala in the Ser/Thr –2, Arg in –1, and Lys in +1 positions. Blue squared signals denote Haspin preferences toward threonine rather than serine residues. Negatively charged residues in the peptide substrate sequences prevent their phosphorylation (red squared region). Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

B. Stabilization of the histone H3 peptide within the Haspin active site of the Histone H3 is achieved through several direct or water-mediated interactions.

C. In vitro kinase reaction of histone H2A full-length protein. Sequence alignment of the Histone H2A and macroH2A sequence show good conservation around the phosphorylation sites H2AThr16 and macroH2A-Ser137 (bottom). Phosphorylation of Histone H3 and H2A was detected by autoradiography. Phosphorylated Histone H2A was subjected to intact molecular weight determination by mass spectrometry. The inset shows that Haspin preferentially phosphorylates the histone on one site. The phosphorylated histone H2A was digested with Arg-C protease and the peptides mixture analyzed by LC-MS/MS. Fragmentation spectrum of the phosphorylated peptide with sequence AKAKT(Pho)RSSRAGLQFPVGR. D, In vitro kinase reaction of CENP-T fragment encompassing residues 1–101. Phosphorylation of CENP-T and histone H3 was detected by autoradiography. Phosphorylated CENP-T 1–101 was subjected to intact molecular weight determination by mass spectrometry. The inset shows that Haspin preferentially phosphorylates the protein on three and four residues respectively. The phosphorilated CENP-T 1–101 protein fragment was digested with Arg-C protease and the peptide mixture analyzed by LC-MS/MS. We identified phosphorylation events on Thr14/27/57 and Ser72 residues. The spectrum shows the fragmentation of the peptide sequence ALLETASPRKLSGQTRT(Pho)IAR.
10% most likely upstream kinases and, tested them as Haspin substrates by in vitro phosphorylation experiment (supplementary Fig. S5, supplemental Table S11). We detected the newly formed phosphorylation sites on the peptide sequences by shotgun mass spectrometry. More than 90% of the peptides in the library contained at least two hydroxyl amino acids (supplemental Table S11). A kinase with poor substrate specificity would be expected to randomly phoshorylate in vitro every site available on the peptide substrates. In the experiment, Haspin showed high specificity for the selection of the substrate. The kinase phosphorylated 96 sites within the peptide library (supplemental Table S12). Thirty-five of these were confidently assigned to specific sites when multiple hydroxyl amino acids were present (PTM-Phorest probability equal to 1) and all of these confirmed sites matched the predictions. On average about 80% of all fragment ion spectra identified the predicted sites, confirming the specificity of the kinase reaction and the accuracy of the NetPhorest predictions (supplemental Fig. S5, supplemental Table S12). We further validated the phosphorylation sites associated with CENP-T (NetPhorest probability 0.23) and histone H2A (NetPhorest probability 0.11) by phosphorylating the respective proteins in vitro (Fig. 5C, 5D). We selected CENP-T because of its role in kinetochore assembly and likely colocalization with Haspin (57). We selected histone H2A because of the high degree of sequence conservation around the predicted Haspin phosphorylation sites between H2AThr16 and macroH2ASer137 (Fig. 5C) and, because macroH2ASer137ph is reduced after Haspin inhibition. Specifically, a fragment encompassing residues 1–101 of CENP-T, histone H2A and histone H3, respectively, were incubated with Haspin kinase domain in the presence of 32P ATP (Fig. 5C, 5D). We showed that Haspin very efficiently phosphorylated both CENP-T and histone H2A in vitro (Fig. 5C, 5D). To determine the number of phosphorylated residues on CENP-T and histone H2A, respectively, we determined the intact molecular weight of the proteins by mass spectrometry. The data showed that Haspin phosphorylated CENP-T preferentially on three or four sites, whereas histone H2A was predominantly phosphorylated at a single site (Fig. 5C, 5D). Mass spectrometric analysis of tryptic digestes of the respective phospho-proteins identified Thr14/27/57 and Ser72, for CENP-T 1–101 and Thr16 for H2A as the phosphorylated residues. (Fig. 5C, 5D). In summary, using computational predictions based on the newly identified Haspin consensus motif and mass spectrometric validation of the predictions we identified novel bona fide Haspin substrates. This study extends the number of putative Haspin substrates from 1 to 38.

Identification of Haspin Protein Interaction Network—We next performed affinity purification-mass spectrometry (AP-MS) analysis of the Haspin kinase to test whether some of the predicted substrates or other proteins physically associated with the Haspin kinase. Haspin was expressed as affinity tagged bait protein in FLP-in HEK293 cells using established protocols (58) and the purified complex was analyzed by standard LC-MS/MS (Fig. 6A). To distinguish true interactors from proteins nonspecifically associating with the isolated complexes we generated data from control samples using unrelated bait proteins (GFP) (36) and statistically filtered the data (37) (Experimental Procedures). The results identified 50 proteins that passed the filtering criteria and that were thus considered true interactors (Fig. 6A, supplemental Table S13). About 70% of these (33 of 50) were implicated with transcriptional regulation and 50% (25 of 50) were previously reported to copurify with components of the spliceosome (http://spliceosomedb.ucsc.edu/) (Fig. 6A).

Indeed, STRING analysis of the identified proteins reported a highly interconnected network of interactions between the spliceosome-associated subset of Haspin interactors (Fig. 6A, pink subgroup). Interestingly, in the chromatin data set 75 phosphorylation sites were observed on 18 of the 50 identified Haspin interactors. Ten phosphorylation sites mapping to five proteins were significantly regulated upon 5-ITu treatment (supplemental Table S14). However, none of sites identified on Haspin binding partners matched the Haspin consensus motif, suggesting that other kinases are responsible of these phosphorylation events. Although none of the interactors was scored as a possible Haspin substrate, we asked whether the protein complex could instead mediate the interaction between Haspin and its predicted high confidence substrates. Indeed, out of the 29 highly ranked predicted substrates 11 were previously shown to physically interact with one or more of the Haspin binding partners, (Fig. 6B) suggesting that Haspin binding partners could physically mediate the interaction between Haspin and its substrates. In conclusion, the identification of a Haspin interaction network, both reinforces and complements our substrate predictions; about 50% of the most confidently predicted Haspin substrates physically interact with Haspin binding partners, suggesting that Haspin might in vivo phosphorylate components of this complex network of interactors on the predicted sites.

DISCUSSION

In this study, we used the ATP analog 5-ITu to investigate the role of Haspin catalytic activity in mitotic cells. A specificity screen performed on a panel of 138 kinases in vitro revealed a tight selectivity of 5-ITu toward Haspin and the members of Clk kinase family (11). We developed an efficient and robust protocol to identify phosphorylation changes on chromatin associated proteins as a result of treatment with 5-ITu (Fig. 2A). We identified and quantified by mass spectrometry, 3964 phosphorylation sites mapping on 1125 proteins (Fig. 2B–2D, supplemental Table S2). As far as we are aware this represents the largest data set of phospho-sites available on chromatin proteins. The vast majority of changes in phosphorylation state that we observed in response to 5-ITu treatment affected mainly the three classes of proteins: (1) proteins involved in mitotic reg-
ulation; (2) proteins involved in RNA processing; and (3) histones and chromatin processing proteins (Fig. 2E). KSEA analysis of the phosphorylation sites down-regulated upon 5-ITu treatment of mitotic cells indicated that Haspin and the three kinase families Aurora, CLK, and RSK, were affected (Fig. 4). This could either be caused by a direct inhibition mediated by the 5-ITu, or caused by indirect effects secondary to Haspin or CLK inhibition. One known indirect effects of Haspin inactivation is the release of the CPC complex (AuroraB, Incenp, Borealin, and Survivin) from the centromere (Fig. 1A). In agreement with the literature we found that the phosphorylation sites identified on CPC complex components are down-regulated after 5-ITu treatment (Fig 2D, supplemental Table S5). Interestingly, we found that the Aurora B substrate Hec-1Ser69 is strongly up-regulated upon 5-ITu treatment (supplemental Table S2). This unexpected result suggests the existence of a delicate equilibrium at the centromere, between phosphorylation and dephosphorylation of proteins that likely involves, in addition kinases as well as protein phosphatases (11).

Although phosphorylation sites linked to mitosis could be confidently attributed to both Haspin and Aurora B activities, the phosphorylation of proteins implicated in the control of gene expression mechanisms cannot be unambiguously assigned to any specific kinase highlighted by the KSEA. Indeed, there are several lines of evidence that Aurora B, CLKs, RSKs, and Haspin kinase activities can affect, either directly or indirectly, the mechanisms that control gene expression (9,
To predict Haspin direct substrates within the down-regulated sites we determined the kinase’s consensus motif. The results of PS-OPLS revealed a strong preference for Ala in position P-2, Arg in P-1, and Lys in P+1 (Fig. 5A) as also recently reported by Kettenbach et al. (75). The Haspin preference for these residues finds a structural explanation in the co-crystal structure of Haspin with the histone H3 tail added as substrate (Fig. 5B, supplemental Fig. S4). This indicates a so far unique substrate-binding mode, where the substrate backbone formed a 180° turn in the Haspin active site placing residues Arg² (P-1) and Lys⁴ (P+1), which are important for consensus motif recognition, into deep hydrophilic pockets. Interestingly, these two residues flanking the substrate Thr³ in histone H3 are a hotspot of post-translational modifications such as methylation and acetylation, and thus have key regulatory function for chromatin structure and transcription. Previous studies showed that increased Lys⁴ methylation had an inhibitory effect on Haspin substrate recognition (6). Lys⁴ forms an ion-pair network with the two acidic residues Asp⁷⁰⁷ and Asp⁷⁰⁹. Lys⁴ methylation sterically hinders substrate binding. Based on the structure and the data measured on trimethylated Lys⁴ we predict that Lys⁴ hinders substrate binding. Based on the structure and the data measured on trimethylated Lys⁴ we predict that Lys⁴ acetylation will also weaken H3 substrate recognition by Haspin, and therefore phosphorylation of the Thr³. The identification of the Haspin consensus motif and the structural details of the kinase-substrate interaction guided the identification of potential novel substrates. We identified 11 bona fide novel Haspin substrates including several splicing proteins (supplemental Table S9). Ser¹₃⁷ of histone macroH2A was strongly down-regulated upon 5-ITu treatment. Even though the Net-Phorest probability was below the threshold set in this study we followed up on this phosphorylation site as a potential Haspin substrate because it was recently shown to be important for the phosphorylation of Ser¹⁰ and Thr³ of the histone H3 (48). Histone macroH2A is a histone variant originally found enriched in the inactive X chromosome of female mammals (76). Several studies demonstrate that histone macroH2A functions both as positive and as negative regulator of gene transcription (77). MacroH2A histone variant is about three time larger than other canonical histone proteins and it is composed of three segments: (1) a histone like domain, 64% protein sequence identical to histone H2A; (2) a highly positively charged flexible loop linking the histone fold to the macro domain; and (3) a macro domain that is exposed to the nucleoplasm (77). The DNA-linker interactions are stabilized by ionic bonds involving the DNA negative charges and the positive charges of the basic macroH2A stretch Thr¹²⁵, Pro¹⁶⁰ (78, 79), and these interactions are likely to be important for the control of chromatin condensation (78, 79). Interestingly, we noted that the consensus motif around Thr¹⁶ of histone H2 and Ser¹₃⁷ of histone macroH2A are conserved except for the residue in position +1. In that position histone H2A contains an arginine (Arg¹⁷) whereas histone macroH2A has a proline residue (Pro¹₁₉) (Fig. 5E). Our data clearly demonstrate that macroH2A Ser¹₃⁷ phosphorylation is dependent on the Haspin kinase (Fig. 3) and we speculate that this site is a Haspin direct substrate in vivo whose phosphorylation could modulate the degree of chromatin condensation and control DNA transcription. We validated CENP-T Thr⁵⁷ as bona fide Haspin substrate by in vitro kinase reaction of the CENP-T protein fragment 1–101 (Fig. 5C, S5). CENP-T is a component of the CCAN (constitutive centromere associated network) that plays a central role in the kinetochore assembly, mitotic progression and chromosome segregation (57). Haspin phosphorylates in vitro four CENP-T sites, Thr¹⁴/²⁷/⁵⁷ and Ser⁷² (Fig. 5C, supplemental Fig. S6). Recently, it has been demonstrated that the tight control of CENP-T phosphorylation, primarily by cyclin-dependent kinase (CDK), is important for the assembly and the disassembly of the kinetochores (80) during mitosis. In particular, CDK-dependent phosphorylation of CENP-T recruits Ndc80 complex to the kinetochore (80). It is therefore likely that Haspin activity could assist CDK kinase in the control of protein recruitment at the kinetochores during the different mitotic stages.

Altogether, the results presented here greatly expanded the number of Haspin bona fide direct substrate in cells and clarified the structural mechanisms of Haspin-substrate binding. Furthermore, a novel link between Haspin kinase activity and phosphorylation of proteins involved in the regulation of gene expression is presented.

For over half a century it has been assumed that transcription is globally silenced during mitosis. Recently, new unexpected links between mitosis and the general mechanisms that regulate gene transcriptions have been emerging (81–84). There is evidence of active transcription of centromere satellite regions during mitosis and suggestions that such enzymatic activity is essential to guarantee chromosome stability (84). Recently, it has been shown that specific RNA transcripts interact with CENPA nucleosomes, and that such proteins-RNA binding controls Aurora B localization and its activity at the kinetochore (85). Our data indicate that Haspin might regulate transcription by direct or indirect phosphorylation of splicing and transcription factors, sub-units of the transcription machineries, as well as histones. It is therefore likely that Haspin controls the recruitment of Aurora B to the kinetochore by the synergistic effect of its enzymatic activity, first phosphorylating the histone H3 Thr³ (9, 39) and second by maintaining an active RNA transcription during mitosis. This novel function should be considered in future studies and take into account for a comprehensive understanding of Haspin role in mitosis and more in general for a better understanding of the molecular mechanisms that link DNA transcription and cell cycle progression in human cells.

Acknowledgments—We thank members of the Aebersold, Gstaiger and Wollischl laboratories at IMSB for helpful discussions. Dr. Chris Barnes and Dr. Martin Junger for critical reading of the manuscript.
Chromatin Phosphoproteome Modulation by Haspin Kinase


Molecular & Cellular Proteomics 13.7

Independent Research (FSS) and the European Research Council

the European Union FP7 Grant No. 278568 "PRIMES".

work in the Linding Lab is supported by the Lundbeck Foundation,

the Human Frontier Science Program (HFSP), the Danish Council for

Independent Research (FSS) and the European Research Council

(ERC).

This article contains supplemental Figs. S1 to S7 and Tables S1
to S15.

To whom correspondence should be addressed: Department

of Biology, Institute of Molecular Systems Biology, HPT E 51, Auguste-

Piccard Court 1, CH-8093 Zurich, Switzerland. Fax: +41 44 633 10

51; E-Mail: aebersold@imsb.biol.ethz.ch.

Competing interests: The authors declare that they have no competing
interests.

REFERENCES


Science 298, 1912–1914


3. Bodenmiller, B., Wanika, S., Kraft, C., Urban, J., Campbell, D., Pedrioli,

P. G., Gerrits, B., Picotti, P., Lam, H., Vitek, O., Brusniak, M. Y., Roschi-
tzki, B., Zhang, C., Shokat, K. M., Schlipbach, R., Colman-Lerner, A.,

Nolan, G. P., Nussinovski, A. I., Peter, M., Loewith, R., Mering, von, C.,


system-wide responses to perturbations of kinases and phospho-

proteins in yeast. Sci. Signal. 3, rs4–rs4


conserved putative eukaryotic protein kinases. Protein Sci. 10, 1677–1684

5. Villa, F., Capasso, P., Tortorici, M., Forneris, F., de Marco, A., Mattevi, A.,

and Musacchio, A. (2009) Crystal structure of the catalytic domain of

Haspin, an atypical kinase implicated in chromatin organization. Proc.


Sci. 106, 20198–20203


kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation

and normal metaphase chromosome alignment. Gene Dev. 19, 472–488

8. Kelly, A. E., Sampath, S. C., Maniar, T. A., Woo, E. M., Chait, B. T., and

Funabiki, H. (2007) Chromosomal enrichment and activation of the Au-

rora B pathway are coupled to spatially regulate spindle assembly. Dev.

Cell 12, 31–43


phosphorylation by haspin positions Aurora B at centromeres in mitosis.

Science 330, 231–235


marks establish the inner centromere and chromosome bi-orientation.

Science 330, 239–243


(2012) A small-molecule inhibitor of Haspin alters the kinetochore func-

tions of Aurora B. J. Cell Biol. 199, 269–284


centromeric functions of Aurora B in chromosome segregation. J. Cell

Biol. 199, 251–268

13. Bernstein, E., Muratore-Schroeder, T. L., Diaz, R. L., Chow, J. C., Chang-

olikar, L. N., Shabanowitz, J., Heard, E., Pehrson, J. R., Hunt, D. F., and


macroph2A1 is excluded from the inactive X chromosome and enriched


proteins in metaphase chromosome structure. Cell 12, 805–816

15. The CENP-A NAC/CAD (2007) The CENP-A NAC/CAD kinetochore com-

plex controls chromosome congression and spindle bipolarity. EMBO J.

26, 5033–5047


(2012) A novel ubiquitin mark at the N-terminal tail of histone H2As

supported by an SNF project grant, the Swiss Cancer League and the

Novartis Foundation. SK is grateful for support from the SGC, a

grant#233266 and the Phos-

spectrometry. 

BMC Bioinformatics 9, 163


Genet. Mol. Biol. 3, 1544–6115


Molecular & Cellular Proteomics 13.7

19. Sturm, M., Bertsch, A., Gröpl, C., Hildebrandt, A., Hussong, R., Lange, E.,

Pfleger, N., Schulz-Trieglaff, O., Zerck, A., Reinert, K., and Kohlbacher, O.

(2008) OpenMS - an open-source software framework for mass spec-

trometry. BMC Bioinformatics 9, 163

20. SmithKline, Janssen, Lilly Canada, the Novartis Research Foundation,

Novartis Foundation. SK is grateful for support from the SGC, a

supported by an SNF project grant, the Swiss Cancer League and the
a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10, 48


44. Clough, T., Key, M., Ott, I., Schadow, G., and Vitkoe, O. (2009) Protein phosphorylation in label-free LC-MS experiments. *J. Proteome Res.* 8, 5275–5284


