Modulation of the Chromatin Phosphoproteome by the Haspin Protein Kinase*

Alessio Maiolica†, Maria de Medina-Redondo§, Erwin M. Schoof¶, Apirat Chaikued¶¶, 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‡‡‡

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 (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 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 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 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 Ser137 of the histone variant macroH2A as a target of Haspin kinase activity. MacroH2A Ser137 resides in a basic stretch of about 40 amino acids that is required to stabilize extranucleosomal DNA, suggesting that phosphorylation of Ser137 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. *Molecular & Cellular Proteomics 13: 10.1074/mcp.M113.034819, 1724–1740, 2014.

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 electronnegative cleft for the histone H3 basic tails (5). Interestingly, the recognition of H3 has been shown to be modulated by methylation at H3 residue Lys^4, thus coupling Haspin activity with epigenetic mechanisms of chromatin regulation (5). Histone H3 that is phosphorylated at Thr^3 is so far the only well-characterized Haspin substrate (7). H3Thr^3 phosphorylation (H3Thr^{3ph}) 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 Ser^{137} (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—**

Haspin^652–798 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-2rbs, a dicistronic derivative of pGEX-6P vector generated in-house. The sequence 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 Ser^{137} (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—**

Haspin^652–798 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-2rbs, a dicistronic derivative of pGEX-6P vector generated in-house. The velocity of conformation was confirmed by sequencing. For the protein expression, BL21(DE3) Rosetta cells containing the pGEX CENP T^101 plasmid were grown in Terrific broth at 37 °C to an OD_{600} 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 mix-
Histone Acidic Extraction and Western Blot—293T cells used for acidic extraction of nucleosomal histones were transfected both with Haspin and empty plasmids using calcium phosphate method as reported before (16). Cellular extracts (80 to 150 μg) were separated on SDS-polyacrylamide gel electrophoresis and transferred overnight at 4 °C onto a polyvinylidene difluoride membrane (Amersham Biosciences). Membranes were blocked for 40 min in 5% dry milk in Tween-tris-buffered saline (TTBS) and incubated overnight at 4 °C with a primary antibody diluted in TTBS containing 1% bovine serum albumin.

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-DDB), which collects known phosphorylation sites from public resources such as PhosphoELM (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 assay were 250 μM ATP (5 μCi of γ-32P-ATP for SDS page only), 5 μM of histones H3 and H2A (Roche), and 5 μM 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—Recombinant 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 iotodotubercidin and 3 μM 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 mM 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 P22121 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 Mass Spectrometry Analysis—For each condition, ten 15-cm dishes of Hela cells were grown to ~80% confluence and treated with 1 μM 5-Itu. Chromatin proteins were prepared as described before.

Protein Digestion—Disulfide bonds were reduced with TCEP (Thermo) at a final concentration of 10 μM at room temperature for 1 h. Free thiols were alkylated with 10 mM iodoacetamide at room temperature for 30 min in the dark. The solution was subsequently diluted with 50 mM ammonium bicarbonate (AMIC) (pH 8.3) to a final concentration of 1.0 mM urea, 0.1% RapiGest (Waters) and digested overnight at 37 °C with sequencing-grade modified trypsin (Promega, Madison, WI) at a protein-to-enzyme ratio of 50:1. Peptides were desalted on a C18 Sep-Pak cartridge (Waters) and dried under vacuum. Phosphopeptides were isolated from 300–500 μg of total peptide mass with TiO2 as described previously (26). Briefly, the dried peptides were dissolved in an 80% acetonitrile 3.5% trifluoroacetic acid solution saturated with phthalic acid. The peptides were eluted twice with 150 μl 1% NH4OH.

Sample Preparation for Phosphoproteomics—Chromatographic separation of peptides was carried out with an Eksigent (Eksigent, Dublin, CA) and Proxenon (Thermo Scientific, San Jose) NanoLC system connected to a 15-cm fused-silica emitter with 75-μm inner diameter (BGB Analytik, Alexandria, VA) packed in-house with a Magic C18 AQ 3-μm resin (Michrom BioResources, Auburn, CA). The phosphopeptide samples were analyzed by LC-tandem MS (LC-MS/MS) run with a linear gradient ranging from 95% solvent A (98% H2O, 2% acetonitrile, 0.1% formic acid) to 35% solvent B (98% acetonitrile, 2% H2O, 0.1% formic acid) over 90 min at a flow rate of 300 nl/min. Shorter gradients of 60 min were instead used 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.

Chromatin Phosphoproteome Modulation by Haspin Kinase
Generation of Cell Lines Stably Expressing Haspin and Haspin Pull Down—Open reading frames of Haspin were retrieved from the Gateway (Invitrogen, Carlsbad, CA) adapted human open reading frames collection (horfome v5.1, Open Biosystems, www.openbiosystems.com). Open reading frames were introduced by LR recombination into a destination vector that was constructed by ligating the Gateway recombination cassette and an N-terminal His6-HA-StrepII-tag into the polylinker of the pcDNA3/FRT/TO vector (Invitrogen).

HEK293 cells were grown in DMEM supplemented with 10% fetal bovine serum (PAA Laboratories, Coelbe, Germany), 0.2 mM L-glutamine, 100 μg/ml hygromycin B and 15 μg/ml blasticidin S (all Invitrogen) and were plated on 15 mm cell culture dishes (Nunc). Protein expression was induced in medium lacking hygromycin and blasticidin by the addition of 1 μg/ml tetracycline 24 h prior to harvest. Cells were detached by pipetting, washed twice with ice-cold PBS and cell pellets were frozen in liquid nitrogen.

For each pull-down we used a cell pellet deriving from 10 × 15 cm 80% confluent both from cycling and prometaphase-arrested cells (0.33 μM nocodazole for 16 h). The cells were harvested and washed twice with PBS. The cell pellets obtained were lysed on ice in two volumes of lysis buffer (LB, 50 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.1% Nonidet P-40, phosphatase and protease inhibitor mixture, Roche). Insoluble material was removed by centrifugation at 13,000 rpm for 30 min at 4 °C. Strept-Tactin Sepharose beads (1200 μl, slurry) were transferred to a Bio-Spin chromatography column and washed twice with 1.5 ml of LB. The cell lysates were loaded two consecutive times on the Strept-Tactin columns. The beads were washed with 3 × 1 ml lysis buffer, and bound proteins were eluted with 3 × 300 μl freshly prepared 2 μM d-biotin in LB. Anti-HA agarose beads (80 μl) were washed with 2 × 1 ml LB buffer and centrifuged at 1000 rpm for 1 min at 4 °C. The supernatant was removed and the beads resuspended in 100 μl LB buffer. The anti-HA agarose beads were added to the biotin eluate and rotated for 1 h at 4 °C. The anti-HA agarose beads were washed with 3 × 1 ml LB buffer and then with 3 × 1 ml LB buffer without detergent and protease inhibitors.

Retained proteins were eluted twice with 200 μl of 100 mM glycine, pH 2.5. The pH in the sample was adjusted using AMBIC 1 N pH 8 and successively trypsin. The peptides originated were purified and analyzed by mass spectrometry as previously described. To identify specific Haspin binding partners we used a strategy based on spectral counts and control pull downs. The data set of control pull-downs consisted of 80 different protein purifications either performed cell expressing GFP or cells transfected with an empty vector. We retrieved spectral counts for every identified protein using the ABA-CUS software (36). For the set of proteins identified in the control pull-downs we selected the highest five spectral count values between all the 80 analyses performed. We estimated the probability for each identified protein of being Haspin specific interactors using SAINT (37). The SAINT algorithm creates separate distributions for true and false interactions and assigns the probability of a bona fide protein–protein interaction to each identified proteins in the data set. We considered Haspin specific interacting partners only proteins: (1) with spectral counts seven times higher than the average values of the negative controls; and (2) protein interactions with the highest possible SAINT probability (SAINT probability = 1).

RESULTS

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

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 besides 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 substrate 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 subdivided into three distinct structural layers: (1) the inner kinetochore 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 phosphopeptides 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 H3Thr10ph 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 Ser693 and Thr601 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 Ser69 and Ser62 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
5-ido tubercidin
Mitotic Chromatin purification
Phosphorylated peptides enrichment
LC-MS/MS
Identification and quantification

B

Novel Phosphorylation sites
Known Phosphorylation sites
158
1273
4375
5527
4 Phosphorylation sites
3 Phosphorylation sites
2 Phosphorylation sites
1 Phosphorylation sites

C

Kinetochore microtubules
Centromere

D

E

Log2(Height/Control) vs. Log2(Height/Haspin)

Log2 Fold Change

Log2 Fold Change

-10 0 10
-10 0 10

64
67
30
34
16
12
108
12
18
8

1.0E-30
1.0E-27
1.0E-24
1.0E-21
1.0E-18
1.0E-15
1.0E-12
1.0E-9
1.0E-6
1.0E-3
1.0E-0
1.0E+0

Molecular & Cellular Proteomics 13.7
The “RNA processing cluster” included subunits of the splicesome, 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, 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.

The “RNA processing cluster” included subunits of the splicesome, 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,
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 and RNA helicases and polymerases, as well as methyl and acetyl transferases (supplemental Tables S2, S3) and proteins that bind specific post translationally 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 Ser10 (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 substrate 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. Conversely, up-regulation of the phosphorylation sites would suggest that the activity of specific kinases is increased upon 5-ITu treatment.

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 × 10−4); ii) the CLK family consisting of the kinases Clk1, Clk2, and Clk3 (p value 2.86 × 10−9); and iii) the RSK family consisting of p90RSK, RSK2, and RSK3 kinases (p value 1.7 × 10−9) (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 down-stream 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 (54).

We purified a recombinantly expressed Haspin kinase domain and used it for the in vitro kinase reaction on degenerated 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-noD/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 Fig. S4) which likely creates ionic electrostatic repulsions with acidic substrates. In summary, the positional scanning oriented peptide library screening identified a preferred recognition sequence for Haspin that matches with the Haspin phosphorylation site on histone H3 tail.

**Structural Mechanisms of Substrate Recognition—**To gain structural insights on the basis of the Haspin substrate recognition motif identified, we cocrystallized the kinase catalytic domain with the H3 tail substrate sequence. The structure was refined to 1.9 Å resolution (Fig. 5B, S4 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 synthetized 101 peptide sequences selected within the predicted sites where Haspin was among the top
FIG. 5. Identification of Haspin consensus motif and Haspin substrates prediction.

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; 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 32P. 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)RSSRAGLPFVGR. 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 showsthat Haspin preferentially phosphorylates the protein on three and four residues respectively. The phosphorylated 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 ALLET ASPRKL SGQTR T(Pho) IAR.
10% most likely upstream kinases and, tested them as Haspin substrates in an in vitro phosphorylation experiment (supplemental 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 phosphorylate 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-Prophet 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 among the predicted Haspin phosphorylation sites between H2AThr16 and macroH2ASer137 (Fig. 5C) and, because macroH2ASer137-phosphate 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 found 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 digests of the respective phosphoproteins 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 mediated 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 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 S2). It has been described previously that centromeric enrichment of the CPC complex lead to an increased catalytic activity of Aurora B (8). It is thus not surprising that we measured a strong down-regulation of Aurora B dependent phosphosites on CPC members (Fig. 4B, 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 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 NetPhorest 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 (Pro¹³⁸) (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, 5D). 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, Thr⁵⁷, 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.

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, subunits 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 Wollscheid 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

11. De Antoni, A., Maffini, S., Knapp, S., Musacchio, A., and Santaguida, S.

10. Yamagishi, Y., Honda, T., Tanno, Y., and Watanabe, Y. (2010) Two histone

12. Wang, F., Ulyanova, N. P., Daum, J. R., Patnaik, D., Kateneva, A. V.,

9. Wang, F., Dai, J., Daum, J. R., Niedzialkowska, E., Banerjee, B., Stuken-

7. Dai, J. J., Sultan, S. S., Taylor, S. S. S., and Higgins, J. M. G. (2005) The

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

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

2. Ubersax, J. A., and Ferrell, J. E. (2007) Mechanisms of specificity in protein

1. Manning, G. (2002) The protein kinase complement of the human genome. Science 296, 1912–1934

6. Eswaran, J., Patnaik, D., Filipakopoulos, P., Wang, F., Stein, R. L., Murray, J. W., Higgins, J. M. G., and Knapp, S. (2009) Structure and functional characterization of the atypical human kinase haspin. Proc. Natl. Acad. Sci. 106, 20198–20203

7. Dai, J. J., Sultan, S. S., Taylor, S. S. S., and Higgins, J. M. G. J. (2005) The 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 Aurora B pathway are coupled to spatially regulate spindle assembly. Dev. Cell 12, 31–43

9. Wang, F., Dai, J., Daum, J. R., Niedzialkowska, E., Banerjee, B., Stukenberg, P. T., Gorsky, G. J., and Higgins, J. M. G. (2010) Histone H3 Thr-3 phosphorylation by haspin positions Aurora B at centromeres in mitosis. Science 330, 231–235

10. Yamagishi, Y., Honda, T., Tanno, Y., and Watanabe, Y. (2010) Two histone marks establish the inner centromere and chromosome bi-orientation. Science 330, 239–243

11. De Antoni, A., Mattini, S., Knapp, S., Musacchio, A., and Santaguida, S. (2012) A small-molecule inhibitor of Haspin alters the kinetochore functions of Aurora B. J. Cell Biol. 199, 269–284

12. Wang, F., Ulyanova, N. P., Daum, J. R., Patnaik, D., Kateneva, A. V., Gorsky, G. J., and Higgins, J. M. G. (2012) Haspin inhibitors reveal 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 Allis, C. D. (2008) A phosphorylated subpopulation of the histone variant macroH2A1 is excluded from the inactive X chromosome and enriched during mitosis. Proc. Natl. Acad. Sci. 105, 15303–15308

14. Adolph, K. W., Cheng, S. M., and Laemmli, U. K. (1977) Role of nonhistone proteins in metaphase chromosome structure. Cell 12, 805–816

15. The CENP-A NAC/CAD (2007) The CENP-A NAC/CAD kinetochore complex controls chromosome congression and spindle bipolarity. EMBO J. 26, 5033–5047

16. Gatti, M., Pinato, S., Maspero, E., Soffientini, P., Polo, S., and Penengo, L. (2012) A novel ubiquitin mark at the N-terminal tail of histone H2As promotes histone variant macroH2A1. J. Cell Biol. 200, 20204–20209

17. Mok, J., Kim, P. M., Lam, H. Y. K., Piccirillo, S., Zhou, X., Jeschke, G. R., Sheridan, D. L., Parker, S. A., Desai, V., Jwa, M., Cameroni, E., Niu, H., Goodman, M., Remenyi, A., Ma, J. L. N., Sheu, Y. J., Sassi, H. E., Sopko, R., Chan, C. S. M., De Virgilio, C., Hollingsworth, N. M., Lim, W. A., Stern, D. F., Stillman, B., Andrews, B. J. J., Gerstein, M. E., Snyder, M., and Turk, B. E. (2010) Deciphering protein kinase specificity through large-scale analysis of yeast phosphoprotein site motifs. Sci. Signal. 3, ra12–ra12

18. Dinkel, H., Chica, C., Via, A., Gould, C. M., Jensen, L. J., Gibson, T. J., and Diella, F. (2010) Phospho.ELM: a database of phosphorylation sites—update 2011. Nucleic Acids Res. 39, D261–D267

19. Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Res. 40, D261–D270

20. Wenschuh, H., Volkm-Engert, R., Schmidt, M., Schulz, M., Schneider-Mergener, J., and Reineke, U. (2000) Coherent membrane supports for parallel microsynthesis and screening of bioactive peptides. Biopolymers 55, 188–206

21. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. 50, 760–768

22. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2000) Likelihood-enhanced fast translation functions. Acta Crystallogr. D. Biol. Crystallogr. 61, 458–464

23. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. 60, 2126–2132

24. Mursudhov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D. Biol. Crystallogr. 53, 240–255

25. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D. Biol. Crystallogr. 66, 12–21

26. Bodenmiller, B., and Aebersold, R. (2010) Quantitative analysis of protein phosphorylation on a system-wide scale by mass spectrometry-based proteomics. Methods Enzymol. 470, 317–334

27. Keller, A., Eng, J., Zhang, N., Li, X.-J., and Aebersold, R. (2005) A uniform proteomics MS/MS analysis platform utilizing open XML file formats. Mol. Sys. Biol. 1, 2005.0017

28. Shteynberg, D., Deutsch, E. W., Lam, H., Eng, J. K., Sun, Z., Tasman, N., Mendoza, L., Moritz, R. L., Aebersold, R., and Nesvizhskii, A. I. (2011) iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. Mol. Cell. Proteomics 10, M111.007890

29. Lam, H., Deutsch, E. W., Eddes, J. S., Eng, J. K., King, N., Stein, S. E., and Aebersold, R. (2007) Development and validation of a spectral library searching method for peptide identification from MS/MS. Proteomics 7, 655–667

30. Sturm, M., Bertsch, A., Gripl, C., Hildebrandt, A., Hussong, R., Lange, E., Pfeifer, N., Schulz-Trieglaff, O., Zerck, A., Reinert, K., and Kohlbacher, O. (2008) OpenMS - an open-source software framework for mass spectrometry. BMC Bioinformatics 9, 163

31. Smyth, G. K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. 3, 1544–6115

32. Eden, E., Navon, R., Steinfeld, L., Lipson, D., and Yakhini, Z. (2009) GOrilla:
Molecular & Cellular Proteomics 13.7

Chromatin Phosphoproteome Modulation by Haspin Kinase

a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10, 48

33. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Mingeza, P., Bork, P., Mering, von, C., and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 41, D888–D891

34. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., and Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27, 431–432

35. Bensimon, A., Schmidt, A., Ziv, Y., Elkon, R., Wang, S. Y., Chen, D. J., Aebersold, R., and Shilo, Y. (2010) ATM-dependent and independent dynamics of the nuclear phosphoproteome after DNA damage. Sci. Signal. 3, r3–r5

36. Fermin, D., Basrur, V., Yocum, A. K., and Nesvizhskii, A. I. (2011) Abacus: a general framework for analyzing and visualizing sequencing data for label-free quantitative proteomic analysis. Proteomics 11, 1340–1345

37. Choi, H., Larsen, B., Lin, Z.-Y., Breitbart, A., Mellacheruvu, D., Fermin, D., Qin, Z. Y., Tsiers, M., Gingras, A.-C., and Nesvizhskii, A. I. (2011) SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat. Methods 8, 70–73

38. Banerjee, S., Sargsyan, S., Musacchio, A., and Villa, F. (2011) A general framework for inhibitor resistance in protein kinases. Chem. Biol. 18, 966–975

39. Kelly, A. E., Ghenoiu, C., Xue, J. Z., Zierhut, C., Kimura, H., and Funabiki, H. (2010) Survivin reads phosphorylated histone h3 threonine 3 to activate the mitotic kinase Aurora B. Science 330, 235–239

40. Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B., and Aebersold, R. (2007) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. Nat. Methods 4, 231–237

41. Beck, M., Schmidt, A., Malmstroem, J., Claassen, M., Ori, A., Szymborska, A., Herzog, F., Rinner, O., Ellenberg, J., and Aebersold, R. (2011) The quantitative proteome of a human cell line. Mol. Sys. Biol. 7, 1–8

42. Ohta, S., Bukowski-Williams, J.-C., Sanchez-Pulido, L., de Lima Alves, F., Wood, L., Chen, Z. A., Platani, M., Fischer, L., Hudson, D. F., Ponting, C. P., Fukagawa, T., Emrshaw, W. C., and Rappoliber, J. (2010) The protein composition of mitotic chromosomes determined using multi-classifier combinatorial proteomics. Cell 141, 810–821

43. Musacchio, A., and Salmon, E. D. (2007) The spindle-assembly checkpoint. Nature 449, 267–276

44. Mahadevan, L. C. (1999) The nucleosomal response associated with DNA replication. Cell 97, 321–330

45. Cesareni, G., Pawson, T., Turk, B. E., Yaffe, M. B., Brunak, S., and Mornon, J. P. (1999) The linear motif atlas for phosphorylation-dependent signaling. Sci. Signal. 1, ra2–ra2

46. Couture, J.-F., Collazo, E., and Trievel, R. C. (2006) Molecular recognition of histone H3 by the WD40 protein WDR5. Nature Struct. Mol. Biol. 13, 698–703

47. Chignola, F., Gaetani, M., Rebane, A., Org, T., Zucchelli, C., Spitaleri, A., Mannella, V., Peterson, P., and Musco, G. (2009) The solution structure of the first PHD finger of autoimmune regulator in complex with non-modified histone H3 tail reveals the antagonistic role of H3K4 trimethylation. Nucleic Acids Res. 37, 2951–2961

48. Southall, S. M., Wong, P.-S., Odo, Z., Roe, S. M., and Wilson, J. R. (2009) Structural basis for the requirement of additional factors for ML1 SET domain activity and recognition of epigenetic marks. Mol. Cell 33, 181–191

49. Varier, R. A., Ouchkhourov, N. S., de Graaf, P., van Schaijk, F. M. A., Eising, H. J. L., Wang, F., Higgins, J. M. G., Kops, G. J. P. L., and Timmers, H. M. (2010) A phospho/methyl switch at histone H3 regulates TRF2 association with mitotic chromosomes. The EMBO J. 29, 3967–3978

50. Markaki, Y., Christogianni, A., and Politou, A. S. (2009) Phosphorylation of histone H3 at Thr3 is a part of a combinatorial pattern that marks and configures mitotic chromatin. J. Cell Sci. 122, 2809–2819

51. Flanagan, J. F., Mi, L.-Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y., Minor, W., Rastinejad, F., and Khorsanizadeh, S. (2009) Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature 438, 1181–1185

52. Verschure, P. J., van der Kraan, I., de Leeuw, W., van der Vlag, J., Carpen, T. A., Belmont, A. S., and van Driel, R. (2005) In vivo HIP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation. Mol. Cell. Biol. 25, 4552–4564

53. Thomson, S., Clayton, A. L., Hazzalin, C. A., Rose, S., Barratt, M. J., and Mahadevan, L. C. (1999) The nucleosomal response associated with immediate-early gene induction is mediated via alternative MAP kinase cascades: MSK1 as a potential histone H3/HMG-14 kinase. The EMBO J. 18, 4779–4793

54. Lim, J.-H., Catez, F., Birger, Y., West, K. L., Prymakowska-Bosak, M., Postnikov, Y. V., and Bustin, M. (2004) Chromosomal protein hmg1 modulates histone h3 phosphorylation. Mol. Cell 15, 573–584
73. West, K. L., Ito, Y., Birger, Y., Postnikov, Y., Shirakawa, H., and Bustin, M. (2001) HMGN3a and HMGN3b, two protein isoforms with a tissue-specific expression pattern, expand the cellular repertoire of nucleosome-binding proteins. J. Biol. Chem. 276, 25959–25969.

74. Lee, J. W., Choi, H. S., Gyuris, J., and Brent, R. (1995) Two classes of proteins dependent on either the presence or absence of thyroid hormone for interaction with the thyroid hormone receptor. Mol Endocrinol. 9, 243–254.

75. Kettenbach, A. N., and Gerber, S. A. (2011) Rapid and Reproducible Single-Stage Phosphopeptide Enrichment of complex peptide mixtures: application to general and phosphotyrosine-specific phosphoproteomics experiments. Anal. Chem. 83, 7635–7644.

76. Costanzi, C., and Pehrson, J. R. (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature 393, 599–601.

77. Gamble, M., and Kraus, W. L. (2010) Multiple facets of the unique histone variant macroH2A: From genomics to cell biology. Cell Cycle 9, 2568–2574.

78. Chakravarthy, S., Patel, A., and Bowman, G. D. (2012) The basic linker of macroH2A stabilizes DNA at the entry/exit site of the nucleosome. Nucleic Acids Res. 40, 8285–8295.

79. Muthurajan, U. M., McBryant, S. J., Lu, X., Hansen, J. C., and Luger, K. (2011) The linker region of macroH2A promotes self-association of nucleosomal arrays. J. Biol. Chem. 286, 23852–23864.

80. Gascoigne, K. E., and Cheeseman, I. M. (2013) CDK-dependent phosphorylation and nuclear exclusion coordinately control kinetochore assembly state. J. Cell Biol. 201, 23–32.

81. Hofmann, J. C., Husedzinovic, A., and Gruss, O. J. (2010) The function of spliceosome components in open mitosis. Nucleus 1, 447–459.

82. Kadauke, S., and Blobel, G. A. (2013) Mitotic bookmarking by transcription factors. Epigenetics & Chromatin 6, 1–1.

83. Neumann, B., Walter, T., Hériché, J.-K., Bulkescher, J., Erle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U., Cetin, C., Sleckmann, F., Pau, G., Kabbe, R., Wünsche, A., Satagopam, V., Schmitz, M. H. A., Chapuis, C., Gerlich, D. W., Schneider, R., Els, R., Huber, W., Peters, J.-M., Hyman, A. A., Durbin, R., Pepperkok, R., and Ellenberg, J. (2010) Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature 464, 721–727.

84. Chan, F. L., Marshall, O. J., Saffery, R., Kim, B. W., Earle, E., Choo, K. H. A., and Wong, L. H. (2012) Active transcription and essential role of RNA polymerase II at the centromere during mitosis. Proc. Natl. Acad. Sci. 109, 1979–1984.

85. Ferri, F., Bouzina-Segard, H., Velasco, G., Hube, F., and Francastel, C. (2009) Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. Nucleic Acids Res. 37, 5071–5080.

86. MacLean, B., Tomazela, D. M., Shultman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26, 966–968.