KiPIK Screening: A Novel Method to Identify Kinases Responsible for Phosphorylation events of Interest

NIKOLAUS ALLAN WATSON

Thesis submitted for the degree of Doctor of Philosophy

Institute of Cell and Molecular Biosciences

Newcastle University

Submitted November 2018

Protein phosphorylation is one of the most abundant forms of post-translational modification and regulates nearly every aspect of cell biology. Understanding the function of a particular phosphorylation event depends to a great extent on identifying the kinase responsible for catalysing it. However, while advances in mass spectrometry based phosphoproteomics have seen an explosion in the ability to detect phosphorylation events occurring in cells, methodological limitations make identifying the kinase responsible for specific phosphorylation events challenging.

This thesis explores this problem, beginning with a discussion of the determinants of kinase-substrate specificity in a cell. This is followed by a review of methodologies currently available for identifying kinases responsible for specific phosphorylation events, and a chapter exploring the utility of one of these techniques (siRNA kinome screening) for identifying kinases required for specific histone phosphorylation events in mitosis.

We then report the development of KiPIK screening (Kinase Inhibitor Profiling to Identify Kinases), a novel general method for identifying the kinase responsible for a phosphorylation event of interest. The method exploits the fact that in recent years large numbers of kinase inhibitors have been profiled for inhibitory activity on near-kinome-wide panels of recombinant kinases. The method treats the inhibitory information for each kinase as a 'fingerprint' for the identification of kinases acting on target phosphorylation sites in cell extracts. In this thesis we detail the development of the technique and validate it on diverse known kinase-phosphosite pairs, including two mitotic histone phosphorylations carried out by Haspin and Aurora B, EGFR autophosphorylation, and the phosphorylation of integrin β1 by Src-family kinases. Finally, we use it to identify the kinase responsible for an as yet unassigned mitotic phosphosite on the Chromosomal Passenger Complex component INCENP.

KiPIK screening is broadly applicable and technically straightforward. In addressing the methodological insufficiency in this fundamental area, it has the potential to benefit research widely.

Acknowledgements

I would like to thank my supervisor Jon for his excellent mentorship. I feel very privileged to have had his input and guidance throughout my PhD. I would also like to thank Chris, Becky, Beth, Diana, Onur, Mark, Tyrell for being great people to work and socialise with.

Thanks also to my parents Erika and Aedan, sisters, Flora and Roisin, for their interest and support. Also huge thanks to Julia for her love and support throughout.

Table of Contents

List of figures

Chapter 1. Introduction

1.1 Overview

In recent years advances in mass spectrometry-based phosphoproteomics have led to an explosion in the detection of phosphorylation events occurring in cells (von Stechow et al., 2015). However, understanding the biology controlled by a particular phosphorylation event depends to a great extent on being able to identify the kinase responsible for catalyzing it. Unfortunately, the methods available for assigning kinase-phosphosite dependencies are limited. Consequently, our ability to benefit from our expanding knowledge of cellular phosphorylation events is severely hampered.

The problem of assigning kinases to identified phosphorylation sites is the focus of this thesis.

The sections of this Introduction include: a general introduction to protein phosphorylation in the cell; an overview of the mechanisms by which kinase-substrate specificity is achieved in the cell; and a review of available methodologies for determining physiological kinases for specific substrates.

1.2 Protein phosphorylation in the cell

1.2.1 *Overview*

Protein phosphorylation is one of the most abundant forms of post-translational modification and regulates nearly every aspect of cell biology. By some estimates, more than 75% of the human proteome undergoes phosphorylation (Sharma et al., 2014). Acting as a reversible molecular switch, it serves to modulate the biology of individual proteins and the cellular processes in which they are involved. It regulates protein-protein interactions, activation states, serves as the main conduit in many of

a cell's signalling processes, and facilitates the transformations in cellular biology required as a replicating cell moves through the cell cycle.

1.2.2 *Mechanism*

Protein phosphorylation is the post-translational modification of a protein by covalent attachment of a phosphoryl group to one (or more) of its amino acid residues. In a cell, phosphorylation comes about as the result of an enzymatic reaction catalysed by a specialised family of proteins known as kinases.

Kinases are a diverse family of proteins (see 1.2.4) but share a conserved catalytic core which contains an ATP binding cleft and substrate binding site. Kinases catalyse the transfer of γ-phosphate from a donor ATP molecule onto their substrate; resulting in the release of ADP and a phosphorylated substrate protein (de Oliveira et al., 2016). In eukaryotes, protein phosphorylation is typically restricted to serine, threonine and tyrosine residues (Ubersax and Ferrell, 2007).

1.2.3 *Functional consequences*

Phosphoryl groups are dianionic at physiological pH and capable of forming extensive hydrogen bond networks. Consequently phosphorylation can have profound effects on the structure and function of proteins (Johnson and Lewis, 2001).

Conformational changes

One common result of phosphorylation is the alteration of molecular interactions within a protein, which can cause extensive conformational change in the phosphorylated protein. Kinases themselves are often regulated in this manner. A highly conserved feature of protein kinases is an 'activation loop' whose phosphorylation is usually critical in inducing a conformational change around the active site of the kinase. This conformational change repositions catalytically

important residues in the kinase active site crucial for its enzymatic activity (Kornev et al., 2006).

Protein-protein interactions

Phosphorylation induced changes in protein conformation and charge can also extensively modulate protein-protein interactions (Child and Mann, 2006). Moreover, there are conserved phosphorylation binding domains (PBDs) found in hundreds of proteins in the cell. PBDs typically recognise and bind specific short linear amino acid motifs in a phosphorylation dependent manner, thereby functioning as phosphorylation recognition modules for the proteins in which they are found (Yaffe and Elia, 2001) (Schlessinger and Lemmon, 2003).

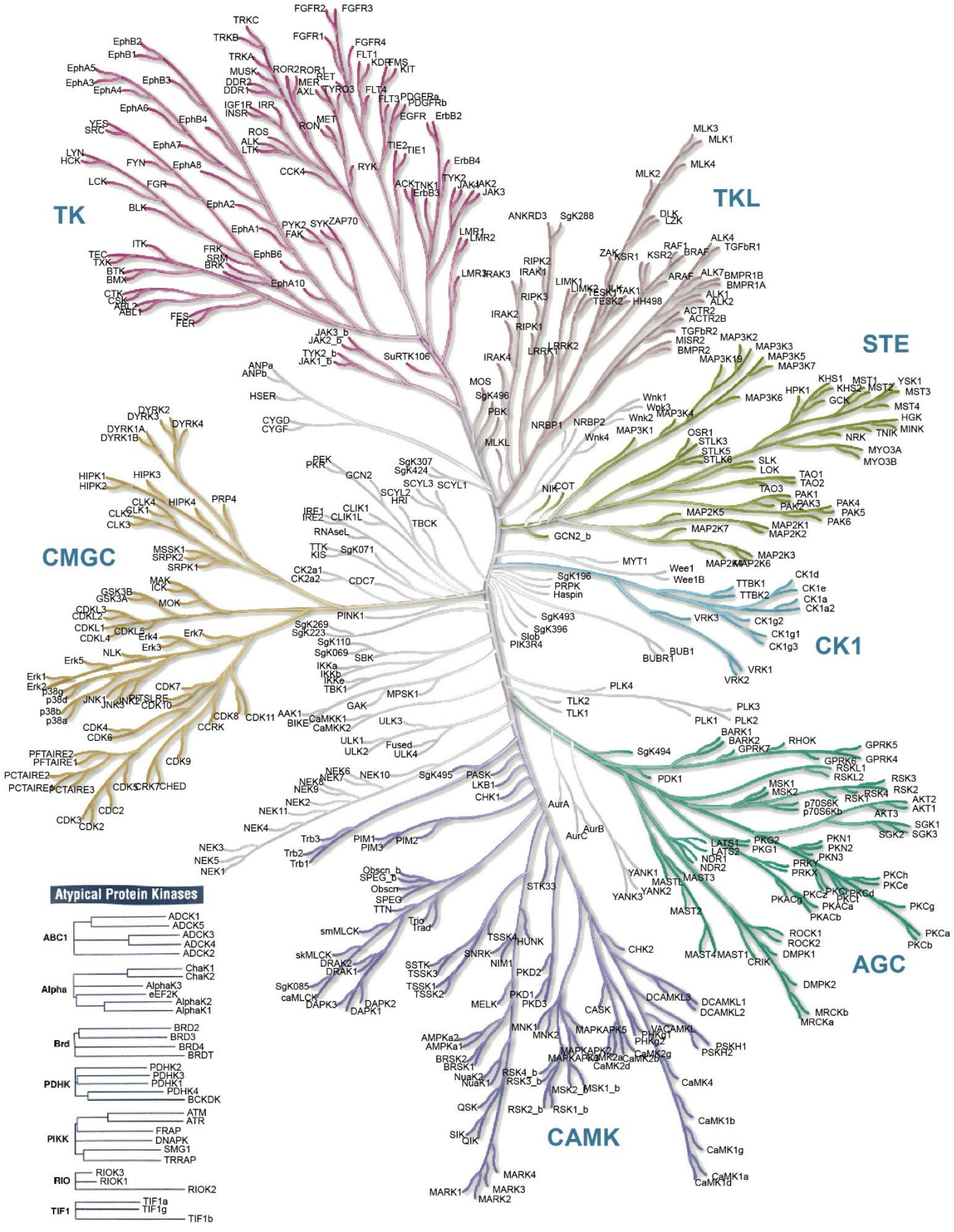
Systemic effects

By regulating the conformation and interactions a protein makes, phosphorylation facilitates dynamic changes in the function of proteins and the larger cellular processes in which they are involved.

1.2.4 *The human kinome*

The human kinome comprises more than 500 genes. This number represents almost 2% of the genome and underlines the importance of kinases in regulating cellular processes (Manning et al., 2002).

The kinome has been divided into 7 major subfamilies through phylogenetic analysis: AGC, CAMK, CK1, CMGC, STE, TK, TKL. These groupings were arrived at primarily by sequence comparison of their catalytic domains. Consequently, there are often similarities in substrate binding preferences between kinases with similar subfamily groupings (see 1.3.2). In addition there are a number of atypical protein kinases which lack sequence similarity in their catalytic domain but have demonstrated protein kinase activity (Taylor and Kornev, 2011).



"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"

Figure 1.1: Phylogenetic tree of the human kinome. Reproduced courtesy of Cell Signalling Technology Inc.

1.2.5 *Phosphatases and dephosphorylation*

Protein phosphorylation events can be reversed through the dephosphorylating activity of phosphatases. The dynamic nature of protein phosphorylation is essential to its utility in the cell; consequently the state of most phosphorylation events depends on a balance between competing kinase and phosphatase activities (Hornberg et al., 2005, Shi, 2009).

1.3 Factors determining kinase-substrate specificity

1.3.1 *Overview*

When approaching the problem of which kinase is responsible for phosphorylating a particular phosphorylation site it is important to first understand the factors that determine the substrate specificity of kinases in a cell.

The factors that determine kinase-substrate specificity can broadly be grouped into two major categories: structural differences in the catalytic domain between kinases, and regulation of local substrate concentrations (Johnson, 2011, de Oliveira et al., 2016, Ubersax and Ferrell, 2007, Miller and Turk, 2018).

1.3.2 *Structural differences around kinase catalytic sites*

Structural differences around the catalytic site between kinases is an important determinant of substrate specificity. These differences typically include variations in the shape, charge and hydrophobicity of the substrate binding site; and have the effect of greatly increasing the binding potential of substrates with complementary characteristics while excluding others (see Figure 1.2).

Most phosphorylation sites occur in relatively unstructured regions of proteins (Zanzoni et al., 2011). Consequently, it is these unstructured regions which have to accommodate themselves to a kinase's catalytic site for phosphorylation to take place. As such, the characteristics determined by the linear amino acid sequence directly flanking the phosphorylated residue are of great importance in determining substrate specificity. For each kinase particular peptide sequences are more or less optimal for binding to the active site and undergoing phosphorylation.

Efforts to characterise which linear sequences are more of less suited to phosphorylation by particular kinases have been underway for some time with the hope that rules for substrate binding could be determined for each kinase. Experiments with libraries of peptides have aided in determining optimum peptide

sequences for a number of kinases (Songyang et al., 1994, Hutti et al., 2004). However, the cellular substrates of particular kinases can deviate extensively from these optimal motifs and from each other (Miller and Turk, 2018).

Work from Mok et al. (2010) used a ‘positional scanning peptide library’ to identify individual substrate residues that were particularly important or required for phosphorylation by 61 recombinant kinases. The library consisted of 200 distinct short peptide mixtures spotted onto a membrane. The peptides consisted of a fixed phospho acceptor residue flanked by 4-5 residues either side. Each of the 200 separate mixtures had in addition to the phosphoacceptor one fixed residue in a flanking position, whilst all the other flanking residues were degenerate mixtures. The results of in vitro kinase reactions with the recombinant kinases on the library peptides indicated that only a few kinases had stringent requirements for particular amino acids at multiple residue positions. Most kinases had few or even no absolutely required residues in fixed positions but displayed subtle variations in phosphorylation efficiency between peptide mixtures containing particular fixed residues (Mok et al., 2010).

A key problem in our understanding of how substrate linear sequences contribute to specificity is the difficulty in measuring and defining how combinations of different residues contribute to catalytic site binding. In a study by (Joughin et al., 2012) the authors sought to address this issue by examining interpositional dependence between flanking residues in several sets of experimentally verified substrates (of specific kinases). Interpositional dependence was examined by measuring the frequency of co-occurrence of pairs of amino acids at particular flanking positions (within substrates of the same kinase) and comparing this to what would be expected by chance. Strikingly only a handful of statistically significant co-occurrences were detected. The authors propose that either kinases largely recognise each amino acid separately (which they deem biophysically implausible), or that individual co-occurrences contribute an energetic effect size to substrate binding that is too small for detection on the limited set of substrates the authors had available (Joughin et al., 2012).

Currently the proportion of kinase-substrate specificity which is conferred by substrate linear sequence around the binding site is a matter of debate (Miller and Turk, 2018, de Oliveira et al., 2016). Whether a sufficiently detailed understanding of the contribution of specific substrate residues, or residue combinations to catalytic domain binding for each kinase would allow accurate predictions of kinase-phosphosite dependencies is unclear.

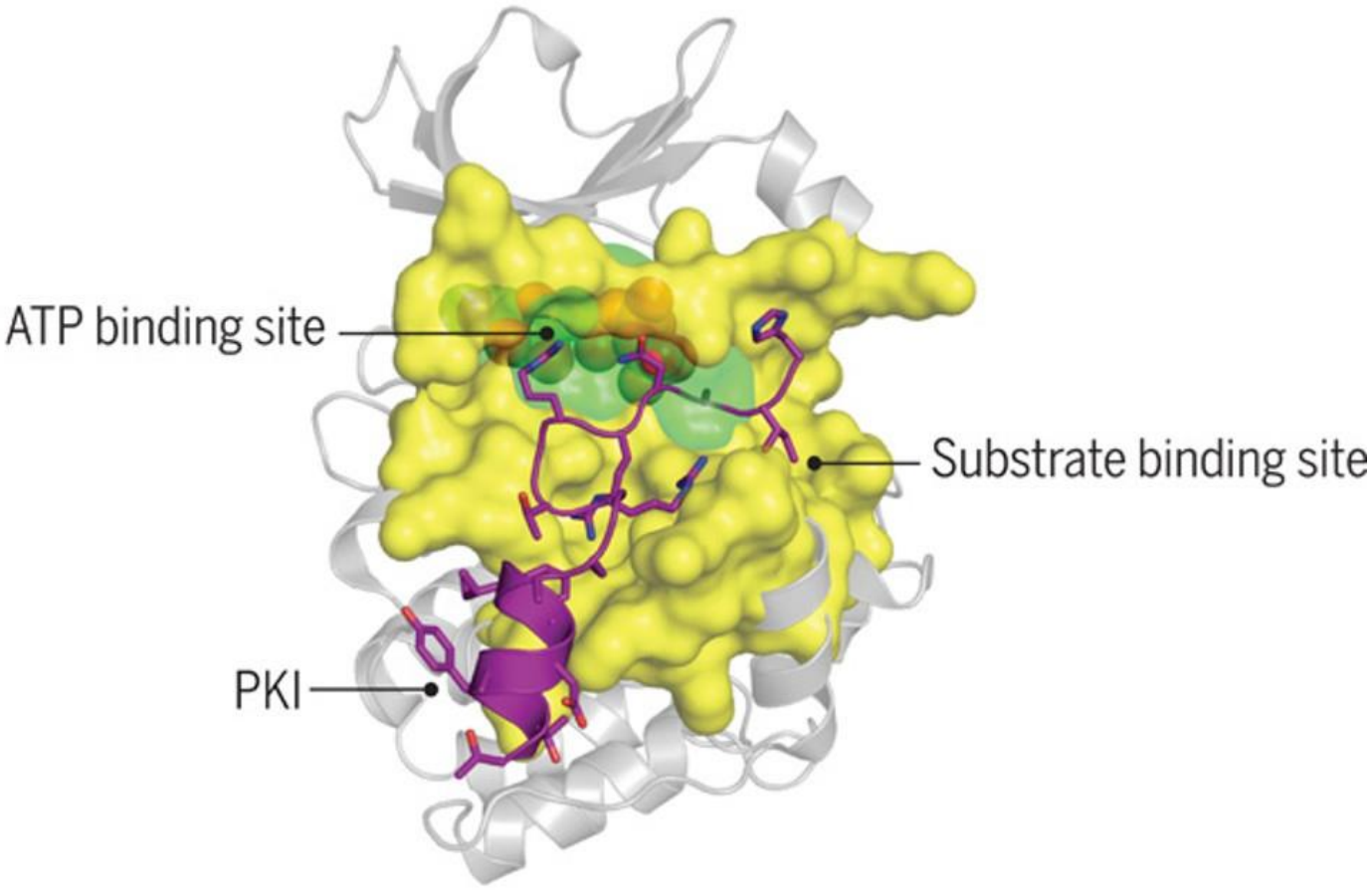


Figure 1.2: Local interactions of substrate residues around the active site of a kinase are important determinants of specificity. Structure of the catalytic domain pf PKA bound to synthetic peptide inhibitor PKI (purple). The synthetic peptide PKI (purple) binds the substrate binding site of PKA (yellow) with high affinity due to a complimentary linear amino acid sequence. Adapted from Oliveira et al. (2016)

1.3.3 *Regulation of local concentrations of kinases and substrates*

Another important determinant of kinase-substrate specificity in cells is the local concentration of substrate and kinase. This is affected at the whole cell level by overall protein abundance of both kinase and substrates - broadly determined by rates of synthesis and degradation of each. Differences in sub-cellular localisation can also have a decisive effect on whether particular kinases have access to particular substrates. Membrane-bound organelles for example can sequester or exclude particular proteins thereby facilitating or limiting the contact between particular proteins. The exclusion of mitotic kinases from the nucleus prior to nuclear envelope breakdown is a prominent example (Fields and Thompson, 1995).

The regulation of protein-protein interactions hardwired into a cell via protein-protein interaction domains also has profound effects on the subcellular localisation of kinases and their substrates.

Docking domains

Besides the substrate binding mediated by interactions at the catalytic site kinases frequently have additional binding sites that further regulate their protein-protein interactions (Ubersax and Ferrell, 2007, de Oliveira et al., 2016, Miller and Turk, 2018). In some instances, these can increase kinase-substrate binding directly by providing additional interaction sites between them. Mitogen activated protein kinases (MAPKs), for example, share a common docking (CD) domain which recognise specific motifs found in their substrates (Peti and Page, 2013, Tanoue et al., 2000). Another prominent example is the polo-box domain (PBD) shared by polo-like kinases. The PBD of polo kinases functions as a phosphoserine/threonine binding domain and plays an important role in substrate recognition and subcellular localisation of the polo kinases (van de Weerdt et al., 2008).

Adaptor proteins

The binding domains of kinases also frequently promote kinase-substrate interactions by means of intermediary binding partners. For example, a conserved mechanism of substrate recruitment by cyclin dependent kinases (CDKs) is through

remote interactions with their cyclin partners. A hydrophobic patch found on cyclins mediates CDK/cyclin interactions with substrates containing an RXL motif sequence (Brown et al., 1999) .

Multi-protein complexes

Kinase binding domains are also important in the formation of multi-protein complexes. These complexes often consist of multiple adaptor and scaffolding proteins which coordinate the enzymatic activity of the kinase by the integration of their own protein binding properties into the multi-protein complex.

The effect that multi-protein complex inclusion can have on substrate selection by kinases can be profound. A dramatic example of this is found with the kinase mTOR. mTOR is found in two distinct complexes in cells, mTORC1 and mTORC2. Remarkably, despite sharing the same catalytic subunit unique adaptor proteins in the 2 complexes result in phosphorylation of entirely distinct substrates (Tatebe et al., 2017, Wullschleger et al., 2006).

Similarly, despite sharing a highly conserved catalytic domain the Aurora kinases (Aurora A and Aurora B) have very distinct localisations and functions in mitosis. Their divergent N termini mediate inclusion into the chromosomal passenger complex (CPC) in the case of Aurora B (Carmena et al., 2012), while association with the microtubule binding protein TPX2 results in centrosomal localisation of Aurora A throughout mitosis (Kufer et al., 2002). The distinct localisation of the two kinases results in distinct substrate selection despite the similarity of their catalytic domains. Indeed Li et al. (2015) demonstrated that the catalytic domain of each could substitute for the mitotic functions of the other by generating chimeric proteins (Carmena et al., 2009, Li et al., 2015).

In mitosis, precise and dynamic control of the localisation of multiple kinases is essential (Alexander et al., 2011). The CPC, for example, positions Aurora B at multiple locations as cells progress through mitosis via association with several different scaffolding/adaptor proteins; in each of these locations Aurora B preferentially phosphorylates distinct substrates. Early in mitosis chromatin arms

serve as a scaffold for the CPC via interactions of CPC subunits with HP1α (Liu et al., 2014, Ruppert et al., 2018, Wang et al., 2010); here extensive phosphorylation of H3S10 and H3S28 by Aurora B occurs (Crosio et al., 2002, Fischle et al., 2005). As mitosis progresses, the CPC becomes concentrated at centromeres due to the H3S10ph induced displacement of HP1 from chromosome arms (Nozawa et al., 2010) as well as the concentration of CPC binding sites H3T3ph and SGO1 at the centromere (Kawashima et al., 2010, Wang et al., 2010) (Yamagishi 2010, Kelly 2010). In this location, Aurora B phosphorylates numerous targets associated with its role in regulating kinetochore-microtubule attachments (DeLuca et al., 2011, Carmena et al., 2012). At the onset of anaphase, the CPC is dramatically relocalised to the spindle midzone via association with Mklp2 and microtubules (van der Horst et al., 2015, Gruneberg et al., 2004) (See Chapter 5 for more detail). This relocalisation again favours preferential phosphorylation of a different subset of Aurora B substrates.

The function and delineation of classical intracellular signalling pathways also depend on the formation of multi-protein complexes for the necessary coordination and association of their components. Typically, multi-domain scaffolding proteins facilitate co-localisation of pathway components concentrating kinases and their substrates and allowing efficient signal transduction (Scott and Pawson, 2009, Good et al., 2011).

Substrate availability as determined by abundance and localisation clearly has a substantial effect on kinase-substrate selection and phosphorylation efficiency. Some have even suggested that regulated protein-protein interactions could be a primary determinant in a cell for which proteins are phosphorylated by a particular kinase – with the specificity conferred by kinase catalytic site characteristics serving the subsidiary role of defining which residues within interacting proteins are phosphorylated (Miller and Turk, 2018).

1.4 Identifying the kinase responsible for a specific phosphorylation event

1.4.1 *Introduction*

Most of the experimental techniques that have been developed to assign kinase-phosphorylation site dependencies focus on identifying candidate substrates for a particular kinase.

Methods that allow an experimenter to identify the kinase responsible for a particular phosphorylation site are much more limited. Indeed, we are not aware of a method that has found widespread use that can selectively identify direct upstream kinases for any substrate. Broadly, existing approaches can be divided into three categories: (i) *in silico* predictions; (ii) screens in intact cells, and (iii) biochemical methods, often using cell extracts or recombinant kinases.

1.4.2 *In silico prediction*

In silico methods attempt to make kinase-phosphorylation site dependency predictions by integrating information from datasets of known kinase-phosphorylation site relationships. Most of these methods base their predictions upon the short linear amino acid sequences surrounding phosphosites and therefore rest on the assumption that kinase catalytic site recognition of these sequences is the primary determinant substrate specificity (as discussed in 1.3.2).

Early approaches for kinase-phosphorylation site prediction took the form of simple consensus motifs compiled from the literature. A consensus motif is a simple description of amino acids frequently present at particular positions around the phosphorylation sites of a particular kinase. ‘S/T-P-X-R/K’ for example is a frequently described consensus motif for CDK1 (Amanchy et al., 2007). These motifs are appealing in their simplicity but, because of their low information content,

have limited use as predictive tools. Frequently a phosphorylation site motif could potentially fit the consensus motif of several kinases (many are not mutually exclusive), and genuine substrates often only partially fulfil the motif of the phosphorylating kinase.

More recently a number of groups have developed machine-learning based tools to predict kinase-phosphosite dependencies. Broadly these machine learning approaches work by training computer systems on curated datasets of substrates classified by phosphorylating kinase. The system generates a multi-dimensional predictive model which best-fits the training data into the classifications it has been given. Once trained, query data (phosphorylated sites) can be fed into the model and assigned a classification (kinase).

The largest repository of curated phosphorylation sites with experimentally verified kinases is Phospho.ELM (Puntervoll et al., 2003). With very few exceptions, the linear sequence data of substrates with assigned kinases (found on PhosphoELM) forms the basis of the training datasets used by machine learning based kinase-substrate prediction tools. However, a roadblock in the development of powerful machine-learning algorithm trained predictive models is the low number of experimentally verified phosphorylation sites assigned to many kinases; because of this many of the current models are limited in their kinome coverage.

PPSP is a prominent example of a machine-learning based tool for predicting kinase specific phosphorylation sites. Using kinase-phosphosite assigned data from PhosphoELM PPSP bases its predictions on the 8 amino acids surrounding a phosphorylated residue. Substrate sequences (for training or query) are first converted by a substitution matrix known as BLOSUM62 which scores amino acids based on biophysical similarity (consequently substrate sequences that are more biophysically similar are recognised as such by their machine learning algorithm). After BLOSUM62 substitution the authors implemented a Bayesian decision theory (BDT) algorithm for training predictive models on the PhosphoELM data and were able to train models covering 70 protein kinase groups (Xue et al., 2006).

In another study Miller et al. (2008) describe the development of NetPhorest, a predictive tool based primarily on machine-learning models developed using an artificial neural network algorithm (ANN). The authors used a kinase phylogenetic tree as a basis for selecting and organising the training data they extracted from PhosphoELM. Curating the limited assigned kinase-phosphosite data (from PhosphoELM) in this way allowed them to scale back the resolution of their predictions in some instances to kinase families rather than individual kinases – thereby increasing kinome coverage (but decreasing resolution). It also allowed them in other cases (where more assigned kinase-phosphosite data was available) to select negative control examples from kinases within the same family; increasing the resolution of their predictive models at these parts of the kinome. For example, when training a model to discriminate PKC family substrates, all PKC family assigned phosphosites would be used as positive control examples and all non-PKC assigned phosphosites as negative examples. While, when training models to specifically discriminate PKCα substrates all non-PKCα yet PKC family substrates were used as negative controls (if sufficient interfamily assigned sites were available). The authors also integrated motif predictions from in vitro experiments using positional scanning peptide libraries (PSPL) (as described in 1.3.2) to increase the kinome coverage of their tool (NetPhorest), selecting the best performing classifier (ANN trained or PSPL) for a kinase when there was overlap. Overall they were able to generate predictors for 60 subfamilies of kinases (covering 179 kinases in total) (Miller et al., 2008)

In a subsequent study from the same group the authors sought to improve on NetPhorest by integrating information from STRING (a database of known and predicted protein-protein interactions). Most kinase-substrate interactions are very transient and therefore few have been demonstrated experimentally (Xue et al., 2013). Nevertheless by combining likelihood ratios derived from NetPhorest and those from STRING the authors report a statistically significant improvement in prediction accuracy over NetPhorest (Horn et al., 2014)

A novel tool developed by Brinkworth et al. (2003) takes a different approach to kinase-phosphorylaton site specificity prediciton. The basis of this tool (known as Predikin) is the conserved nature of the catalytic domains of Ser/Thr kinases. By

analysis of crystal structures the authors determined that clusters of residues in conserved positions in the catalytic domains of Ser/Thr kinases make interactions with the 3 residues upstream and downstream of a substrate phosphosite (Brinkworth et al., 2003). Predikin first identifies these specificity determining residues (SDRs) by sequence alignment (for any Ser/Thr kinase). It then calculates for each SDR cluster the probability of each amino acid occurring in the substrate (at the position that SDR cluster interacts with). To do this, the program identifies kinases with similar SDR clusters and their substrates from experimentally determined database entries. The substrates are used to construct a position weighted matrix describing the predicted amino acid frequencies occurring in substrates at each SDR interacting position (-3 to +3 positions relative to the phosphosite) (Brinkworth et al., 2003, Ellis and Kobe, 2011). This technique has the advantage that large numbers of verified substrates are not required for each kinase in order to infer specificities of further substrates. The known substrates of divergent kinases that share similar SDRs in particular positions can be used to infer on a position by position basis the likely substrate amino acids for a query kinase. However, the fact that it calculates substrates on a position by position basis is also a weakness as information on interpositional dependence is lost (as discussed in 1.3.2).

Computational approaches present great advantages in terms of cost and scalability for kinase-phosphosite prediction. However, as they are based on inference rather than experimentation their accuracy and utility is highly dependent on the quality and quantity of training data available. The current lack of experimentally verified substrates for many kinases leads to major limitations in terms of kinome coverage for computational approaches. It is also clear that at their current state of development false positives and negatives are frequent, even when combined with contextual information such as protein-protein interaction data. Clearly these approaches can only become more powerful as more experimentally verified high-confidence kinase-phosphosite dependencies are defined and available as model training data.

1.4.3 *Screens in intact cells*

This category includes techniques such as kinome-wide RNAi, CRISPR/Cas9 or overexpression screens. Such screens are tremendously useful to biologists, but they often identify pathways or networks of kinases that are indirectly required for phosphorylation of a particular substrate rather than (or in addition to) the direct kinase(Friedman and Perrimon, 2006, Ramakrishnan and Rice, 2012, Papageorgiou et al., 2015, Azorsa et al., 2010) Indeed, these indirect effects can make it difficult to identify the direct kinase for a particular phosphorylation site. For example, if RNAi for a kinase impedes cell cycle progression, it is difficult to determine whether its effect on a substrate is direct (Moffat et al., 2006)

1.4.4 *Biochemical methods*

In the third category are the handful of biochemical approaches that have been developed to identify the kinases of specific substrates.

One approach is to screen recombinant kinases *in vitro* for their ability to phosphorylate a purified substrate of interest. Although kinases often display substrate promiscuity in a purified *in vitro* reaction (Peck, 2006, Cheng et al., 1993), by screening several kinases on the same substrate, differences in ability or efficiency of phosphorylation can be directly compared. Jansson et al. (2008) purified 180 GST-tagged human kinases and performed head-to-head kinase reactions on a peptide substrate containing a residue of interest (which undergoes phosphorylation *in vivo)*. They were able to successfully identify 2 kinases from the same family which phosphorylated the peptide with good efficiency and subsequently demonstrated that this role was conserved *in vivo* (Jansson et al., 2008). However, the broad utility of this approach is limited by concerns of specificity, expense and kinome coverage.

Another approach to kinase identification is tracking the kinase activity of interest through chromatographic separation of cell lysate. Ji et al. (2010) sought to identify the kinase responsible for an important activating phosphorylation on PLK1 (T210ph) by this means. Beginning with 7-8 $x10^{10}$ cells they performed multiple

rounds of chromatography on fractions of cell lysate able to phosphorylate a kinase-dead mutant Plk1 substrate. Subsequently they sent proteins in their final fraction for mass spectrometric analysis and identified several candidate kinases (Ji et al., 2010).

A number of strategies have also been developed that are designed to allow upstream kinases to be co-purified with a substrate of interest. Kinase-substrate co-purification is usually very difficult because of the transient nature of the kinase-substrate interaction. Following phosphorylation the negative charge of the added phosphate group often repels the phosphorylating kinase, quickly disrupting the interaction with the substrate (Xue et al., 2013).

Dedigama-Arachchige and Pflum (2016) present a relatively simple strategy to cross-link substrates to nearby proteins following phosphorylation. The technique known as K-CLASP (Kinase Catalyzed CrossLinking And Streptavidin Purification) utilises an ATP analogue ATP-ArN_3 (ATP-arylazide) which is accepted as a co-substrate by diverse kinases. The γ-phosphate of ATP-ArN_3 functions as a cross-linker when stimulated with UV. To co-purify proteins interacting with a substrate of interest, a biotinylated substrate of interest is subject to a kinase reaction in cell lysates in the presence of ATP-ArN_3 and UV. Phosphorylation of the substrate adds the cross-linker to the phosphorylated site which causes UV induced cross-linking to proteins nearby (including the phosphorylating kinase). This substrate can then be streptavidin purified and interacting proteins analysed by MS. Although the authors were able to co-purify the phosphorylating kinase of a peptide they tested the specificity was low; 324 proteins were enriched in total (Dedigama-Arachchige and Pflum, 2016).

Another approach developed by Maly et al. (2004) and later by Statsuk et al. (2008) utilises two modified components to chemo-selectively link a substrate to its phosphorylating kinase. The serine/threonine of a phosphorylated site of interest is replaced by a cysteine residue, which is capable of crosslinking with proximal lysines. Selectivity of the crosslinking with the phosphorylating kinase is ensured by the addition of a kinase ATP site binding chemical crosslinker. The crosslinker

composes a general protein kinase inhibitor moiety and a cross-linking heterocyclic dialdehyde moiety. When a cysteine substituted substrate binds to the active site of its phosphorylating kinase the *in situ* crosslinker catalyses the covalent crosslinking of the substrate with a lysine in the kinase active site. Substrate bound proteins can then be detected following co-purification. However, while the authors were able to demonstrate the effectiveness of this approach *in vitro* with purified components, detection limits of their cross-linking reaction prevented their demonstrating its utility in cell lysates. Therefore, significant further development is required before this approach could be utilised for non-biased kinase identification.

Zeng et al. (2017) describe a newly developed approach utilising bimolecular fluorescence complementation (BiFC) to facilitate co-purification of substrates with interacting kinases. In BiFC, 2 fragments of a fluorescent GFP molecule are fused to candidate interacting proteins. When expressed together in the same cell, any physiological interactions that occur between the fusion proteins will bring together the 2 GFP fragments. The GFP fragments readily bind when in proximity forming a complete GFP molecule; resulting in a stable interaction between the candidate proteins and emitting fluorescence. To adapt this technology for kinase identification, the authors co-expressed one fragment of GFP fused to their substrate of interest and the other fragment to each kinase in an overexpression vector library of 559 kinases. The authors pooled the library of overexpression vectors and purified fluorescence complemented complexes using a GFP antibody which only detected reformed GFP molecules. Subsequent MS analysis allowed them to identify kinases whose interaction was enriched in substrate co-expressing cells relative to control co-expressing (background determined by SILAC). Via this methodology, the authors identified 23 interacting kinases for their substrate of interest (Zeng et al., 2017). A major limitation of this technique is that it is not phosphosite specific. Moreover, kinases can be detected that interact with proteins yet have no role in phosphorylating them (for example downstream kinases in a signalling cascade). Another concern is that overexpression of proteins can lead to false positives.

1.4.5 *Chemical genetic approaches for Identifying substrates of specific kinases*

A notable chemical genetic strategy developed in the Shokat lab has proven successful in assigning kinase-substrate specificities from the other direction – that is by starting with a query kinase and identifying its direct substrates. This approach relies on 2 innovations to allow the differential tagging and isolation of phosphorylated substrates of a known kinase. The first of these is the mutation of sterically important residues within the ATP binding pocket of the query kinase that allow it to accommodate bulky ATP analogs which cannot be used by the rest of the kinome (Shah et al., 1997). In order to allow efficient isolation of the direct substrates of this so-called AS kinase (analog sensitive) the ATP analog used is further modified by substituting the terminal oxygen of the γ-phosphate with sulphur. In this way, when bulky, sulphur modified ATP (A*TPγS) is provided, direct substrates of the AS kinase are thiophosphorylated, whilst substrates of non-AS kinases are not. The reactive sulphur (in place of oxygen) on direct AS substrates serves as a differentiating starting point for the isolation of these substrates from other cellular proteins. In one approach, P-nitrobenzylemesylate (PNBM) is used to alkylate the thiophosphate (and other cellular nucleophiles), creating a unique cellular epitope (thiophosphate ester) on direct AS kinase substrates. An antibody that specifically recognises thiophosphate esters is then used to isolate the AS kinase substrate proteins (Allen et al., 2007). An alternative enrichment strategy developed by the same group allows the identification of specific phosphopeptides phosphorylated by query AS kinases. In this approach, following A*TPγS mediated thiophosphorylation, trypsin is used to digest all proteins, and peptide products allowed to react with iodacetyl-agarose beads. Thiol-containing groups form covalent bonds with these beads and unbound peptides are then washed away. Thiophosphopeptides are then specifically liberated by oxidation using the peroxide agent Oxone (whereas other thiol containing peptides such as thioether remain stable). Phosphopeptide substrates of the AS kinase can then be detected by mass spectrometry (Blethrow et al., 2008). AS mutants have been successfully developed for numerous kinases and this approach represents a powerful method to confidently identify their direct substrates (Hertz et al., 2010). A major limitation of this approach is that the toxicity of thiophosphates (they cannot be removed by phosphatases) limits its use in intact cells or organisms. As with all

phosphoproteomic approaches detection of low abundance phosphorylation events is also a challenge (Blethrow et al., 2008).

In another chemical genetic approach developed by the same group mutations are introduced in highly conserved 'gatekeeper' residues to confer sensitivity to specific kinase inhibitor analogues (Bishop et al., 2000). The mutations introduced allow inhibition of query kinases with bulky inhibitor analogues at low nanomolar concentrations (85-fold to 400-fold greater sensitivity than the most sensitive wild type kinases). The rapid, monoselective inhibition of specific kinases that this method facilitates can be used to explore of confirm the dependency of particular phosphorylation events on query kinases. A strength of this approach is that it can be used in cells without toxicity; however a cell line in which the endogenous gene has been replaced is required for complete inhibition of the target kinase with this method.

1.5 Mitotic phosphorylation events as test cases in the analysis of methodologies for assigning kinase-phosphorylation site relationships

1.5.1 *Overview*

Throughout this thesis mitotic phosphorylation events are used as test cases in the analysis of methodologies for identifying kinases-phosphosite dependencies. Mitotic phosphorylation is dynamically regulated by the interdependent actions of multiple mitotic kinases. This interdependence can complicate the identification of direct kinase-phosphosite dependencies and highlights some of the strengths and weaknesses of the methodologies assessed herein.

It is becoming increasingly clear that phosphorylation of the chromatin itself is a key facilitator of cell division and, in many instances, phosphorylated histones form key components of the regulatory signalling networks governing mitotic progression. Below we review some of the roles and regulation of important mitotic histone phosphorylation events, several of which are explored in the results sections.

1.5.2 *Histone Phosphorylation in Mitosis*

Several histone phosphorylation events occur in mitosis, most of which are first detected in late G2 or early prophase and decline rapidly in anaphase, but which have distinctive localizations on chromosomes. These include H3T3ph catalyzed by Haspin, which originates on chromosome arms but becomes focused at inner centromeres as mitosis progresses; H3S10ph and H3S28ph catalyzed by Aurora B, which first appear at centromeres but spread over the chromosome arms; H2AT120ph produced by BUB1, which is found in chromatin beneath the kinetochores; and CENP-AS7ph generated by Aurora kinases at centromeres (Hsu et al., 2000, Wang and Higgins, 2013, Crosio et al., 2002, Goto et al., 2002, Dai et al., 2005, Kawashima et al., 2010, Zeitlin et al., 2001, Kunitoku et al., 2003). Aurora B, interestingly, as well as being a direct kinase for H3S10, H3S28 and CENP-AS7, appears to act as a "master regulator" of histone kinases in mitosis, as it has roles in coordinating the activity of other mitotic histone kinases including Haspin and perhaps BUB1 (Wang et al., 2011, Brittle et al., 2007). Below we have described known histone phosphorylation modifications in more detail and provided functional context.

Displacement

A notable event in the early stages of mitosis is the displacement from chromosomes of many of the proteins found on chromatin during interphase. Recent studies have implicated several of the histone phosphorylation marks that occur during mitosis as causal agents contributing to this displacement.
For example, early mitosis is notable for widespread repression of transcription, and Varier et al. present evidence that a phospho-methyl switch contributes to this repression by preventing transcription initiation around phosphorylated H3T3 during mitosis (Varier et al., 2010). They show that, *in vitro*, H3T3ph causes inhibition of the interaction between the reader protein TAF3 and its histone target, H3K4me3. TAF3 is a subunit of the transcription factor complex TFIID (itself a component of the preinititation complex of RNA pol II) and so inhibition of its binding to chromatin would be predicted to inhibit transcription. In line with this, ectopic overexpression of the H3T3 kinase Haspin, which increases global H3T3ph in interphase (Dai et al., 2005), causes inhibition of TAF3 mediated transcription activation, and depletion of Haspin causes retention of TFIID on chromosomes during mitosis (Varier et al.,

2010). The purpose of transcriptional repression during cell division is currently unresolved but it may be required to allow proper chromosome condensation, or to provide a window for resetting of gene expression programs (Egli et al., 2008, Wang and Higgins, 2013). Interestingly, *in vitro* studies indicate a potent displacing effect of H3T3ph on other proteins that bind the N-terminal tail of H3 surrounding H3K4 (when either methylated or unmethylated, suggesting that a similar mechanism may remove multiple proteins from chromatin during cell division (Wang and Higgins, 2013).

There is evidence that the archetypal mitotic histone modification, H3S10 phosphorylation, also operates as part of a phospho-methyl switch. H3S10ph, catalyzed by Aurora B, appears in late G2 or early mitosis and is strong throughout the chromatin by late prophase. The neighboring residue H3K9 when di- or trimethylated, can recruit the Heterochromatin Protein HP1 via its chromodomain (Bannister et al., 2001, Nakayama et al., 2001). HP1 recruitment is crucial for heterochromatin formation but the bulk of HP1 is displaced in mitosis, despite no detectible loss of H3K9 methylation. H3S10ph can weaken HP1 binding to methylated H3K9 in vitro, and mitotic displacement of HP1 from chromosomes in mammalian cells is prevented by Aurora B inhibition. This suggests that a phospho-methyl switch operates in mitosis to cause the displacement of HP1 from most of the chromatin via H3S10ph (Figure 1.3) (Fischle et al., 2005, Hirota et al., 2005).

Less well established is the function of H3S28ph in mitosis, another modification generated by Aurora B that, notably, is found within the same ARKS motif as H3S10. Trimethylation of the neighbouring residue H3K27 recruits the Polycomb Repressive Complexes PRC1 and PRC2, and a similar phospho-methyl switch mechanism might promote the dissociation of these proteins during mitosis. However, the evidence for this is more indirect, relying on studies performed during interphase and in vitro (Fonseca et al., 2012, Gehani et al., 2010, Lau and Cheung, 2011). Nevertheless, there are interesting functional implications of such mechanisms. For example, the transcriptional status of repressed Polycomb target genes may need to be maintained through mitosis from one cell generation to the next to maintain cell lineage identity. H3S28ph, and indeed other mitotic marks (*e.g.* H3T3ph), may serve as temporary countermarks through mitosis; allowing the

displacement of transcription factors and other proteins, while retaining underlying epigenetic signals (*e.g.* H3K27me3) required to re-establish transcriptional programs following mitotic exit (Wang and Higgins, 2013).

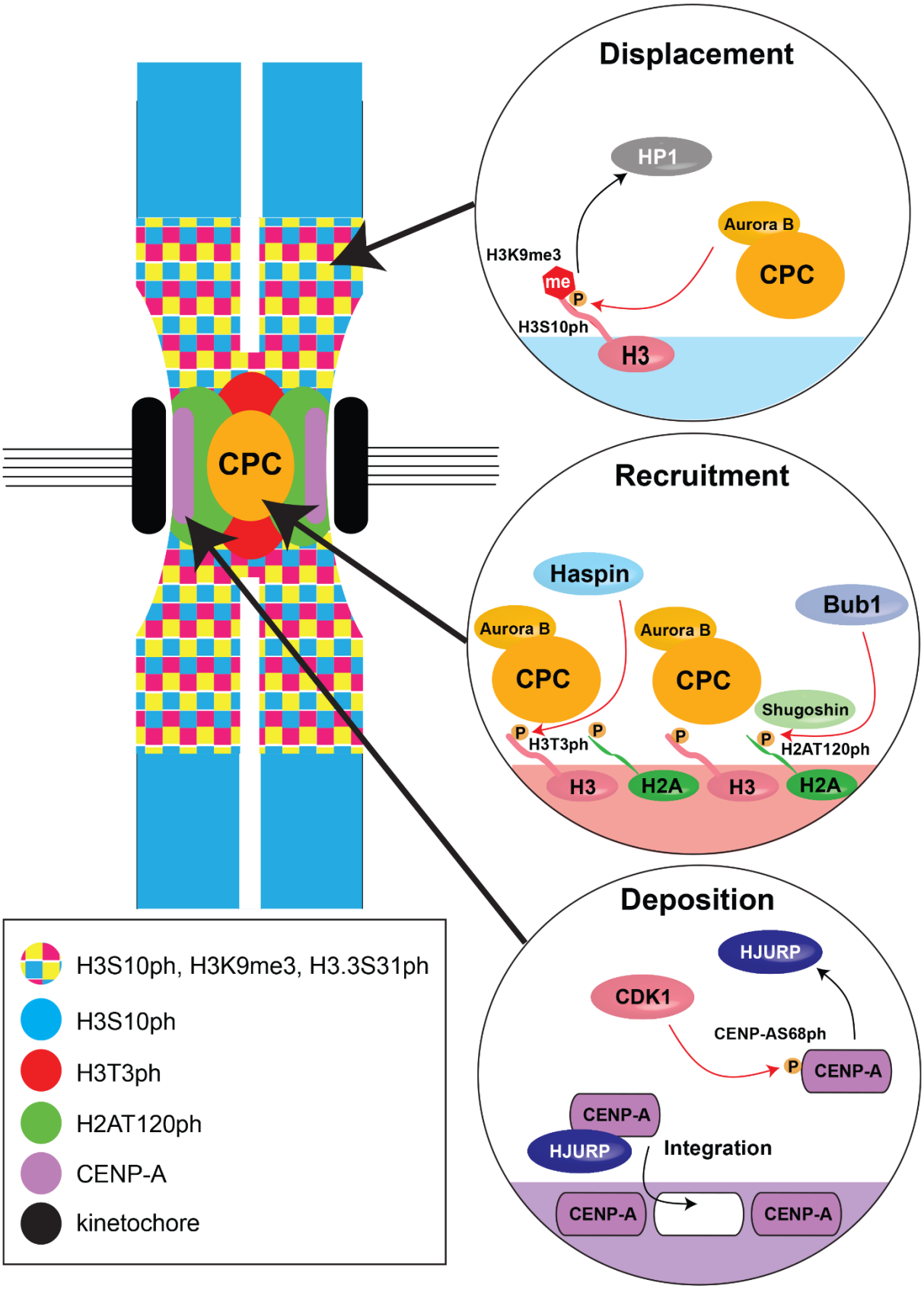


Figure 1.3: Schematic illustrating the location of histone marks on mitotic chromosomes. Also depicting the phosphorylating kinases and functional consequences for a selection of these modifications.

Although appealing, phospho-methyl switching might be more complicated than first envisioned. For example, phosphorylation may be required, but insufficient, to displace HP1 and PRC1: in some experiments H3S10ph or H3S28ph do not decrease HP1 or PRC1 binding to H3K9me2/3 or H3K27me3 *in vitro* (Mateescu et al., 2004, Vermeulen et al., 2010). It is possible that the binding of these proteins to chromatin involves domains other than their methylation-specific chromodomains (Hale et al., 2006, Terada, 2006), and H3K14ac may also contribute to the release of HP1 from H3K9me3 (Mateescu et al., 2004). Histone phosphorylation may also influence the extent of adjacent methylation in cells, for example by altering the recognition of target sites by histone methyltransferases. Further work is needed to fully understand the workings of these proposed molecular switches.

Landmarks

Histone phosphorylation marks have roles in defining landmarks on the chromatin. During mitosis this is particularly important at the centromere where differential phosphorylation helps to define different regions in and around the centromere and kinetochore. The establishment of this phosphorylation pattern enables key regulatory proteins to be recruited to the right places and at the right times.

The main function of the centromere during cell division is as the site of attachment for microtubules of the mitotic spindle. The Aurora B kinase, which functions as part of a multi-protein complex known as the Chromosomal Passenger Complex (CPC), is fundamental to this process, with key roles in establishing kinetochore structure, ensuring the fidelity of microtubule-kinetochore attachments and in mitotic checkpoint signalling. Recent studies have demonstrated that H3T3ph functions as a direct binding site for the CPC via the BIR domain of its subunit Survivin (Figure 1.3) (Wang et al., 2010, Kelly et al., 2010, Yamagishi et al., 2010). H3T3ph becomes enriched at inner centromeres during mitosis and it is believed that the H3T3ph-recruited population of Aurora B contributes to erroneous microtubule attachment correction and mitotic checkpoint signalling (Wang et al., 2012, De Antoni et al., 2012).

The mechanisms that focus the H3T3ph signal at inner centromeres during cell division are complicated and not yet fully understood. H3T3ph is restricted to mitosis as Haspin activation depends on a priming phosphorylation on residue T128 by the master mitotic regulator Cyclin B-CDK1. This phosphorylation provides a docking site for the Polo-like-kinase PLK1 which performs the multiple further phosphorylations of Haspin apparently necessary to fully activate the kinase on chromosomes (Zhou et al., 2014, Ghenoiu et al., 2013). Another important component in generating the H3T3ph signal is a positive feedback loop established by Aurora B recruitment to H3T3ph. Wang et al.(2011) demonstrated that, while Haspin recruits and helps activate Aurora B on chromatin, Aurora B functions to activate Haspin through phosphorylation (Wang et al., 2011, Wang et al., 2012). Moreover, Aurora B also can function in a complementary manner to prevent dephosphorylation of H3T3 by antagonizing the chromosomal binding of the phosphatase Repo-Man-PP1γ (Qian et al., 2011). These amplification loops drive robust H3T3ph generation on chromatin in mitosis, though they do not in themselves explain how H3T3ph (or the CPC) accumulates at centromeres.

H2AT120ph is thought to act together with H3T3ph to guide the CPC to centromeres (Figure 1.3). The kinetochore kinase BUB1 generates H2AT120ph in distinctive patches underlying each kinetochore (Kawashima et al., 2010). The H2AT120ph signal draws the Shugoshin protein to the centromere (possibly through indirect means), and Shugoshin directly binds the CPC (Kawashima et al., 2010, Tsukahara et al., 2010). The H2AT120ph region underlying kinetochores partially overlaps with inner centromeric H3T3ph and one proposal is that the area of intersection between the two marks specifically localizes the CPC (Yamagishi et al., 2010). Alternatively the BUB1-H2AT120ph-Shugoshin pathway may trigger the Haspin-H3T3ph-CPC feedback loop more strongly at centromeres to provide an increased CPC localization signal at the centromere (Wang et al., 2011). In either case, this provides an interesting example of how crosstalk between modifications on two different histones defines a specific chromosomal domain.

Building the kinetochore

The location of centromeres in many organisms is not determined by DNA sequence, but rather is determined epigenetically by the presence of nucleosomes containing the H3 histone variant CENP-A. CENP-A therefore defines the location at which the microtubule-binding outer kinetochore is built. Interestingly the deposition of CENP-A into chromosomes happens, unlike canonical histones, independently of replication. In human cells, mRNA and protein levels of CENP-A peak during late G2 but CENP-A is not integrated into chromatin until mitotic exit. Deposition relies on the activity of the assembly factor HJURP, which acts as a molecular chaperone for pre-nucleosomal CENP-A (Stellfox et al., 2012). Yu et al. demonstrated recently that phosphorylation of CENP-A at S68 by CDK1 inhibits the interaction of CENP-A with HJURP (Figure 1. 3). This mechanism ensures that integration of CENP-A is delayed until mitotic exit when CDK1 levels fall (Yu et al., 2015).

Other studies indicate roles for CENP-A phosphorylation in centromere structure and function. Bailey et al. report phosphorylation of S17 and S19 in human CENP-A and present evidence that these modifications promote the formation of intramolecular bridges between CENP-A tails that prevent hypercondensation of CENP-A nucleosomes. Overexpression of CENP-A S17A/S19A mutants in cells resulted in mitotic defects, and it was proposed that these phosphorylations are important for kinetochore integrity (Bailey et al., 2013). The kinase responsible is currently unknown.

“Orphan” histone phosphorylation marks in mitosis

A number of other “orphan” histone phosphorylations occur during cell division, for which the kinases required and functions are unknown. For example, H3.3S31ph shows a strong increase in signal at pericentromeric regions during mitosis (Hake et al., 2005). Chang *et al*. recently showed a role for CHK1 in H3.3S31ph in cancer cells displaying alternative lengthening of telomeres (ALT). However, H3.3S31ph has an unusual distribution in these cells, covering the whole of the chromosome arms in mitosis rather than being restricted to pericentromeric regions. Interestingly,

while CHK1 depletion resulted in significant loss of this erroneous H3.3S31ph distribution, the typical pericentromeric population remained intact, indicating CHK1 is unlikely to be the usual mitotic kinase (Chang et al., 2015).

H3T11ph is another modification with a strong mitotic centromeric signal but, although a candidate kinase has been proposed, this has never been confirmed in cells (Preuss et al., 2003). A yeast study by Govin *et al*. reveals a role for H3T11ph in meiosis: H3T11A mutation compromises sporulation. The kinase responsible appears to be the CAMK-family kinase Mek1 (Govin et al., 2010), but this is meiosis-specific and does not have a known homologue in metazoans, leaving open the identity of cell division H3T11 kinases in higher organisms (Discussed further in 3.1.5). Nevertheless, it seems certain that these understudied histone phosphorylation events will have important functions during cell division.

1.6 Aims of this Project

The aims of this project are as follows:

1. Explore the utility of traditional genetic screens for identifying the kinase responsible for specific phosphorylation events
2. Develop an alternative methodology for identifying the kinase responsible for specific phosphorylation events that overcomes some limitations of current techniques.

Chapter 2. Materials and Methods

2.1 siRNA screens

Hela cells were grown in 384 well clear-bottomed plates and treated with an siRNA kinome library (Dharmacon). Interferin-HTS was used for transfection, according to the manufacturer's instructions. 42 hours after siRNA transfection, cells were accumulated in mitosis for 6 hours by addition of 200 nM Nocodazole or 300 nM Taxol, and immunofluorescently stained. Average fluorescence intensity (integrated) for mitotic cells (defined as cells positive for either stained mitotic histone mark) was measured using a widefield Nikon microscope (Eclipse Ti-E) equipped with High-Content Analysis software. Experiments were carried out in quadruplicate and Excel was used to calculate the standard score for each repeat (n=4); p-values were calculated using one-sample t-test for each siRNA treatment. Appoximately 300 cells were measured for each siRNA treatment.

2.2 Other high-content imaging-based quantitation of cell populations

HeLa cells were grown in 96 well clear-bottomed plates. Following treatment as described in the relevant section, they were imaged using a widefield Nikon microscope (Eclipse Ti-E) equipped with High-content analysis software. Average fluorescence intensity (integrated) for mitotic cells (defined as cells positive in either stained mitotic histone mark).

2.3 Immunoblotting

HeLa cells were lysed, or *in vitro* kinase reactions stopped, in NuPAGE LDS sample buffer and boiled at 95°C for 5 minutes. Samples were then loaded onto 4-12% NuPAGE Bis-Tris gels and run for 1 hour at 180V. Proteins were then transferred onto a PVDF membrane in transfer buffer (1 hour, 70V, wet transfer). The membrane was then blocked with 2.5% milk in PBS-Tween for 1 hour, and then incubated with the appropriate primary antibodies overnight at 4°C (diluted in 2.5% milk in PBS-Tween). Next, the membrane was washed 3 times for 5 minutes with PBS-Tween, and then incubated for 1 hour with appropriate secondary antibodies (diluted in 2.5% milk in PBS-Tween). This was followed by 3 more washes with

PBS-Tween and 1 wash with PBS (all 5 minutes). ECL western blotting substrate was then applied to the membrane and signal detected with X-ray film.

2.4 Peptide ELISA

Pierce High Capacity Streptavidin coated 384-well plates were used. Prior to biotinylated peptide addition, plates were washed 3x with 60 µl/well TBS-Tween. Biotinylated peptides were then added and incubated at room temperature for 2 hours, then washed 3x with 60 µl/well TBS-Tween. Then primary antibodies added, 40 µl/well, diluted in TBS-Tween with 0.1% BSA, then incubated at room temperature for 2 hours. Then plates were washed 3x with 60 µl/well TBS-Tween. 40 µl/well of secondary antibodies in TBS-Tween with 0.1% BSA was then added, and incubated for 1 hour at room temperature. Then plates were washed 3x with 60 µl/well TBS-Tween, and 40 µl/well TMB substrate was added and blue colour change allowed to develop (for at least 10 minutes). Colour development was stopped by addition of 20 µl/well of 2N H_2SO_4.

2.5 Cell extract preparation

Mitotic extract

Cells were grown to confluence in T300 tissue culture flasks and treated with 300 nmol Nocodazole for 12 hours. Mitotic cells were collected by shake off, washed 1x with PBS and lysed in chilled P buffer on ice (1 ml of P buffer/ T300 flask of mitotic cells). Lysate was immediately flash frozen in liquid nitrogen.

A431 extract (EGF stimulated)

A431 cells were grown to confluence in T300 tissue culture flasks. Cells were trypsinised and then collected in prewarmed DMEM media supplemented with 50 ng/ml of human EGF (Cell signalling). Cells were then returned to a 37ºC incubator for a 5 minute incubation with EGF. Cells were then spun down at 1200 rpm, washed with 1x prewarmed PBS, and lysed in chilled P Buffer on ice (1 ml of P buffer/ T300 flask). Lysate was immediately flash frozen in liquid nitrogen.

2.6 KiPIK extract calibrations

All pipetting was performed with a Biomek FX liquid handling robot (Beckman Coulter). Reactions were performed in 384 well microplates.

384 well reaction plates are prepared as in 2.7, but no inhibitors were added to wells. Instead, wells were prepared either with or without peptides, after which 5, 2, 1 or 0.5% cell extract (final concentration) was added synchronously from the same solution of freshly thawed cell extract in KiPIK buffer. Different % extract was applied to different wells by the addition of different volumes, reciprocal amounts of KiPIK buffer were preloaded into well before extract addition, bringing the final volume to 35 µl per well (final concentration: +/- 0.1 µM Peptide; 0.2 mM ATP; 5, 2, 1 or 0.5% cell extract; diluted in KiPIK buffer).

After extract addition, reaction plates were immediately transferred to a 30°C incubator, and incubated for 30 minutes.

During the incubation Pierce High Capacity Streptavidin coated 384-well plates were washed 3x with 60 µl TBS-Tween per well. After this, 10 µl of 500 mM EDTA diluted in dH_2O was added to each well of the Streptavidin plate.

After the 30 minute incubation, the reaction plate solution was transferred into the Streptavidin coated plate and an ELISA performed (described in 2.4) to measure phosphorylation levels.

2.7 KiPIK screen

All pipetting was performed with a Biomek FX liquid handling robot (Beckman Coulter). Reactions were performed in 384 well microplates.

First 10 µl of kinase inhibitors were plated into the 384-well microplates in duplicate, diluted to 35 µM in Kinase reaction buffer. Next, 10 µl of KiPIK buffer with 0.35 µM peptide and 0.7 µM ATP were added to each well. Plates were then placed on ice. Next, cell extract was thawed and diluted into chilled kinase reaction buffer (at 2.3x the appropriate final % extract concentration, determined by calibration), vortexed briefly, and 15 µl added synchronously to each well of the peptide, inhibitor, and

ATP-containing reaction mix, bringing the total volume to 35 µl (final concentration: 10 µM inhibitor, 0.1 µM peptide, 0.2 mM ATP, % cell extract determined by calibration, diluted in KiPIK buffer). Each plate also included DMSO (60 wells) (DMSO instead of inhibitor added) and EDTA control wells (10 µl of 500 mM EDTA in place of 10 µl inhibitors). After extract addition, reaction plates were immediately transferred to a 30°C incubator, and incubated for 30 minutes.

During the incubation, Pierce High Capacity Streptavidin coated 384-well plates were washed x3 with 60 µl TBS-Tween per well. After this, 10 µl of 500 mM EDTA diluted in dH_2O was added to each well of the Streptavidin plate.

After the 30 minute incubation, the reaction plate solution was transferred into the Streptavidin coated plate and an ELISA performed (described in 2.4) to measure phosphorylation levels.

Following ELISA, a standard score was calculated for each inhibitor treatment in Excel (using the mean of duplicates). Standard score = (Inhibitor absorbance - mean absorbance of DMSO controls)/ standard deviation of DMSO controls. Next % change (% inhibition) of each inhibitor was calculated with the lowest standard score on the plate defined as 100% inhibition (either EDTA control wells or the lowest scoring inhibitor) (calculation: %inhibition = (inhibitor standard score / lowest standard score)*100). % inhibition scores for all inhibitors were then compiled as the inhibition 'fingerprint' of the phosphorylation event probed.

Using Prism (GraphPad), Pearson's correlation was then calculated for this 'fingerprint' against the inhibition profiles of each of the kinases which were profiled *in vitro* against that inhibitor library.

2.8 Cell culture

HeLa Kyoto or A431 cells were grown in DMEM medium supplemented with 5% (v/v) FBS and 100 U/ml Penicillin-Streptomycin, at 37^0C and 5% CO_2 in a humid incubator. Cells were passaged as required every few days with trypsin.

2.9 Indirect immunofluorescence

2.9.1 Sample preparation

HeLa cells were grown on polylysine-coated glass coverslips (Poly-L-Lysine, Sigma, or on Greiner uClear 384 well microplates for Kinome siRNA screens (2.1), or uClear 96 well microplates for other high-content assays (2.2).

Cells were fixed for 10 minutes with 2% paraformaldehyde in PBS. They were washed twice in PBS, and then permeabilized for 2 minutes with 0.5% Triton (in PBS). After washing twice in PBS, the cells were incubated for 1 hour in blocking buffer (5% milk in PBS-Tween) at room temperature with shaking. They were then incubated for 1 hour with primary antibodies at 37^0C (as indicated) diluted in blocking buffer. After washing twice with PBS-Tween, the plates were washed twice with PBS, then incubated for 45 minutes with secondary antibodies at 37^0C (as indicated) diluted in blocking buffer. After washing twice with PBS tween, twice with PBS, and once with milliQ H_2O, coverslips were mounted on microscope slides (or in 384 well microplates, where appropriate) using Prolong gold (for cover slips) (Invitrogen) or Fluoromount-G with DAPI (for microplates) (Invitrogen).

2.9.2 Imaging

Cells grown on coverslips were imaged with a Zeiss Axio Imager microscope using a Plan-Apochromat 100x/1.40 Oil objective. Optical sections were acquired every 0.1 µM using an AxioCam MR R3 camera. Image stacks are displayed as maximum intensity projections.

2.9.3 High-content imaging

High content imaging was carried out using a Nikon Eclipse Ti-E inverted microscope utilising the Nikon High Content Analysis (HCA) software package.

2.10 RNA interference

siRNA kinome screens utilised Interferin-HTS transfection reagent (PolyPlus) according to the manufacturer's instructions.

All other siRNA transfections were carried out using Lipofectamine RNAiMAX (Invitrogen), according to manufacturer's instructions, and at a concentration of 50 nM. Cells were harvested for immunoblotting or fixed for immunofluorescence 48 hours post siRNA treatment.

2.11 *In vitro* kinase reactions

Reactions were carried out in 50 µl Kinase reaction buffer with 0.25 µg of each substrate, and 0.05 µg of recombinant kinase

2.12 Buffers

"P" buffer

50 mM Tris, 0.25 M NaCl, 0.1% Triton X100, 10 mM $MgCl_2$, 2 mM EDTA, 1 mM DTT, pH7.5 (+ 1/1000 protease inhibitor cocktail [Sigma P8340], 1 mM PMSF, 0.1 µM okadaic acid, 10 mM NaF, 20 mM β-glycerolphosphate), phosphostop tablets (Roche)(1 tablet / 10 ml). PMSF added just before use. Phosphostop was added just before use.

KiPIK buffer

50 mM Tris, 10 mM $MgCl_2$, 1 mM E**G**TA, 10 mM NaF, 20 mM β-glycerolphosphate, pH to 7.5 with HCL. Just before use (with cell extract), 1mM PMSF, 1 tablet of Phosphostop / 10ml, pH 7.5

Kinase reaction buffer

20mM HEPES, 0.14 M NaCl, 3 mM KCl, 1 mM ATP, 5 mM $MgCl_2$ pH 7.4

Transfer buffer

PBS-Tween

TBS-Tween

2.13 List of peptides

INCENP peptide. Biotin-GPREPPQSARRKRSY

EGFR peptide. Biotin-ADEYLIPQQ

H3 1-21 peptide C-terminal biotin. Epicypher

H3 21-44 peptide C-terminal biotin. Eurogentec

Integrin β1 tail peptide. Biotin-GGKSAVTTVVNPKYEGK. Eurogentec

2.14 List of antibodies

Primary Antibodies

Target (name)	Species	Source	Dilution
H3T11ph (Fw_B)	Rabbit		1:2000(IF/WB) 1:10K (KiPIK ELISA)
H3S10ph (3H10)	Mouse	Millipore (05-806)	1:2000
H2BS6ph	Rabbit		1:1000
H3T3ph (16B2)	Mouse (monoclonal)	H. Kimura. Tokyo Inst. Technology	1:2500 (IF)
H3T3ph (8634)	Rabbit	Higgins Lab	1:2000 (ELISA)
DDR2	Goat	R&D systems (AF2538)	1:1000
β-Actin	Mouse (monoclonal)	Sigma-Aldrich (A2228)	1:2000
ACA	Human		1:2000
H3S28	Mouse (monoclonal)	H. Kimura. Tokyo Inst. Technology	1:2000
phospho-Tyrosine	Mouse (monoclonal)	Cell Signalling (9411S)	1:10,000
INCENP S446ph	Rabbit	JM Peters. IMP Vienna	1:2000 (IF)1:1000(WB)
INCENP TSSph	Rabbit	M. Lampson. Univ. Pennsylvania	1:1000
Aurora B (SAB.1)	Sheep	S. Taylor. Univ. Manchester	1:2000

Secondary Antibodies

Target species	Source	Dilution
Rabbit (Alexa Fluor 488)	Invitrogen	1:2000
Mouse (Alexa Fluor 594)	Invitrogen	1:2000
Mouse (HRP-linked IgG)	Cell Signalling	1:1000
Rabbit (HRP-linked IgG)	Cell Signalling	1:1000
Goat (HRP-linked IgG)	R&D systems	1:1000
Sheep (Alexa Fluor 647)	Invitrogen	1:1000
Human (Alexa Fluor 647)	Invitrogen	1:1000

2.15 Recombinant proteins

CDK1/CycB Kinase. purified full length GST fusion protein. #7518 Cell Signalling.
Aurora B active kinase. 14-835 Millipore
Purified recombinant fragment INCENP (369-583 aa) NBP2-37471. Novus Bio.
INCENP-GST (826-919 aa). 12-534. Upstate
INCENP-HIS (1-405 aa). Produced in Higgins lab by Debasis Patnaik.

2.16 siRNA

DDR2 mission siRNA #1. SIHK0565 Sigma-Aldrich
DDR2 mission siRNA #2. SIHK0566 Sigma-Aldrich
NT5M mission siRNA #1. SIHP0373 Sigma-Aldrich
NT5M mission siRNA #2. SIHP0374 Sigma-Aldrich
MAPKAPK3 mission siRNA #2. SIHK1232 Sigma-Aldrich
BUB1B mission siRNA #1. SIHK0210 Sigma-Aldrich
TTK siRNA. J-004105-12-0002 GE healthcare

human GSG2 siRNA duplex. IDT

Negative control no. 2 siRNA. AM4637. Life Technologies

2.17 Inhibitor profiling datasets

PKIS1 (as described in (Elkins et al., 2016))

PKIS2 (as described in (Drewry et al., 2017))

Chapter 3. Kinome-wide siRNA screening as a method to identify mitotic histone kinases

3.1 Introduction

3.1.1 Overview

We were interested in identifying the kinases responsible for unassigned mitotic histone phosphorylation events and decided to test the utility of siRNA screening in intact cells as a means to this end.

3.1.2 High-content imaging based siRNA library screening to identify genes involved in the regulation of specific phosphorylation events

Microscopy based high-content imaging siRNA screens have been used by several research groups to investigate genes involved in phosphorylation events of interest (Azorsa et al., 2010, Papageorgiou et al., 2015, Boutros et al., 2015).
As a genetic screen performed in intact cells, this type of approach has the advantage that it deals with phosphorylation events occurring in their physiological context. We decided to explore the utility of this approach for identifying kinases involved in specific mitotic histone phosphorylation events.

3.1.3 H3T3ph and H3S10ph

As covered earlier (see 1.5.2) the mitotic kinase of H3T3ph has been identified as Haspin (Dai et al., 2005), while Aurora B is responsible for mitotic H3S10ph (Hirota et al., 2005). We used these 2 phosphorylation sites as positive control cases for our siRNA screens below.

3.1.4 H2BS6ph

H2BS6ph is a novel histone phosphorylation site recently identified by our collaborators (M. Seibert, L. Schmitz, Justus-Liebig-University, Giessen, Germany). They report a strong H2BS6ph signal between prophase and anaphase with enrichment at inner centromeres. An siRNA kinome screen was performed (below) in hopes of gaining further insight into the regulation of H2BS6ph by mitotic kinases.

3.1.5 H3T11ph

H3T11ph in mitosis

Phosphorylation of H3T11 was first reported by Preuss *et al.* (2003). Using an *in vitro* kinase reaction, they observed the serine/threonine kinase Dlk (death-associated protein (DAP)-like kinase) was capable of phosphorylating core histones H3, H4 and H2A, and via phosphoamino acid analysis determined that in all 3 cases this occurred on threonine rather than serine residues. They determined by sequence analysis that on Histone 3 only Thr11 fulfilled the requirement for a Dlk phosphorylation site and subsequently raised antibodies against this mark using a synthetic phosphopeptide (Preuss et al., 2003).

They then used this antibody to probe the *in vivo* relevance of H3T11ph and determined by western blotting that it did occur in cells but seemed restricted to mitosis, evidenced by massive enrichment in nocodazole arrested cells (Preuss et al., 2003). They followed this up with immunofluorescence analysis and observed H3T11ph particularly between prophase and early anaphase. Furthermore, they report enrichment of the signal around the centromere. Finally, they investigated the localisation of Dlk and observed a similar pattern of chromatin localisation during mitosis as with H3T11ph. They argue that this, combined with the *in vitro* kinase data, strongly suggest a role for Dlk as a centromere-specific histone H3 kinase (Preuss et al., 2003). However, they present no experiments demonstrating a dependence of H3T11ph on Dlk activity in cells. Additionally, there have not been any follow up papers further exploring or confirming a role for Dlk in mitotic H3T11ph. Consequently, the identity of the H3T11ph mitotic kinase is in doubt.

H3T11ph in interphase

Subsequent reports on H3T11ph have confirmed its biological relevance but focused instead on interphase functions and its role in disease states.
Prominent among these is a paper by Shimada *et al*. (2008) which, contrary to initial reports (Preuss et al., 2003), identified phosphorylation of H3T11 throughout the cell cycle, not just restricted to mitosis. They determined this by immunoblotting synchronized cells over a period of 24 hours and found significant H3T11ph

throughout the cell cycle. They went on to identify Chk1 as a kinase capable of phosphorylating H3T11 *in vitro* and found a close correlation between H3T11 phosphorylation in cells and Chk1 chromatin association. Depleting Chk1 by cre/flox recombination resulted in loss of H3T11ph in both untreated and serum starved quiescent cells but, interestingly, not in nocodazole treated cells. Moreover, they demonstrate rapid H3T11ph loss alongside release of Chk1 from the chromatin following DNA damage by UV irradiation. They go on to show increased affinity of the acetyltransferase GCN5 for a phosphorylated H3T11 peptide *in vitro* and correlate decreased H3T11ph upon UV treatment with decreased H3K9 acetylation in cells. Finally they demonstrate, by ChIP analysis, decreased GCN5 residency and H3K9ac at CDK1 and Cyclin B promoters following Chk1 depletion. They propose a model whereby DNA damage causes reduced H3T11ph due to release of Chk1 from chromatin, resulting in decreased GCN5 recruitment and acetylation of H3K9 at the promotor regions of CDK1 and Cyclin B (Shimada et al., 2008).

Another report identifies a role for H3T11ph in androgen receptor (AR) signalling (Metzger et al., 2008). The authors demonstrate that PRK1 is able to phosphorylate H3T11 on bacterially expressed H3 fragments and isolated nucleosomes. PRK1 associates with androgen receptor target promotors upon ligand activation. Using ChIP and a PRK1 kinase inhibitor, the authors demonstrate that PRK1 specifically phosphorylates H3T11 at these sites *in vivo*. Similar to the paper by Shimada et al. (2008), they propose that this modification impacts the modifications of neighbouring H3K9 and H3K14 residues resulting in changes in AR target transcription.

More recently, a report from Yang et al. (2012) indicates the tumour cell specific kinase PKM2 phosphorylates H3T11 in response to epidermal growth factor (EGF) signalling. The authors describe *in vitro* phosphorylation of H3T11 by PKM2 and demonstrate EGF induced H3T11 phosphorylation dependent on catalytically active PKM2 *in vivo* (Yang et al., 2012). They follow this up with experiments supporting a model similar to the previous papers in which H3K9ac is enhanced by the displacement of histone deacetylase driven by H3T11ph.

With the mitotic kinase for H3T11ph in doubt, we decided to perform an siRNA screen to identify kinases involved specifically in the mitotic phosphorylation of H3T11ph (described below).

3.1.6 Aims

1. To investigate the effectiveness of kinome-wide siRNA library screening as a method for identifying kinases involved in specific histone phosphorylation events.
2. Identify kinases involved in the phosphorylation of H3T11 in mitosis
3. Identify kinases involved in the phosphorylation of H2BS6 in mitosis

3.2 Results

3.2.1 siRNA screen design

We designed our siRNA screening approach such that each screen included 2 distinct mitotic histone phosphorylation marks; one for which the mitotic kinase was well established (either H3T3ph or H3S10ph), and the other an 'orphan' modification whose mitotic kinase was unknown (H3T11ph and H2BS6ph). Consequently, for each screen presented, the 2 modifications serve as an internal control for each other.

Figure 3.1 outlines the methodology. Briefly, cells were plated in 384 well plates and transfected with an siRNA kinome library the following day (1 kinase per well). 42 hours after siRNA treatment, Nocodazole or Taxol (see figure legends) was added and the cells were incubated for a further 6 hours, during which time a large mitotic population accumulated. Cells were then fixed in 2% PFA, DAPI stained and immunofluorescently stained for the 2 histone phosphorylation marks. Widefield images were then collected at 3 wavelengths corresponding to DAPI and each of the histone phosphorylation signals.

These images were then processed by image analysis software to measure the average integrated intensity of mitotic cells in each siRNA knockdown treated well, for each of the histone phosphorylation fluorescence signals. For the purposes of

automated image analysis, a binary mask over each image was first created to define ‘cells’ by thresholding objects by DAPI intensity and size. We found that the most robust way to automatically define ‘mitotic cells’ was as ‘cells’ with above background levels of either histone phosphorylation mark fluorescence signal. Finally, an average integrated intensity of ‘mitotic cells’ was calculated for each siRNA treatment and p-values determined by one sample t-test for 4 repeats.

The screens are displayed below as volcano plots indicating, for each kinase, the effect its knockdown had on average mitotic cell intensity (x-axis: displayed as standard deviations from the screen mean) and the p-value results of one sample t-tests (y-axis: p values displayed as –log).

Figure 3.1: Schematic illustrating siRNA kinome screen workflow

3.2.2 H3T3ph and H2BS6ph siRNA screen

Kinome screen

First, we performed an siRNA kinome screen to probe for kinases involved in the phosphorylation of H3T3 or H2BS6 in mitosis. For staining H3T3ph, we used a mouse monoclonal antibody that has been verified for specificity in our lab and others (16B2, a gift from Hiroshi Kimura, Tokyo Institute of Technology, Japan). Our collaborators (M. Seibert, L. Schmitz, Justus-Liebig-University, Giessen, Germany) raised a rabbit polyclonal antibody against H2BS6ph peptides and verified its specificity on peptides and immunoprecipitated H2B (data not shown). An siRNA screen was performed and costained with both antibodies.

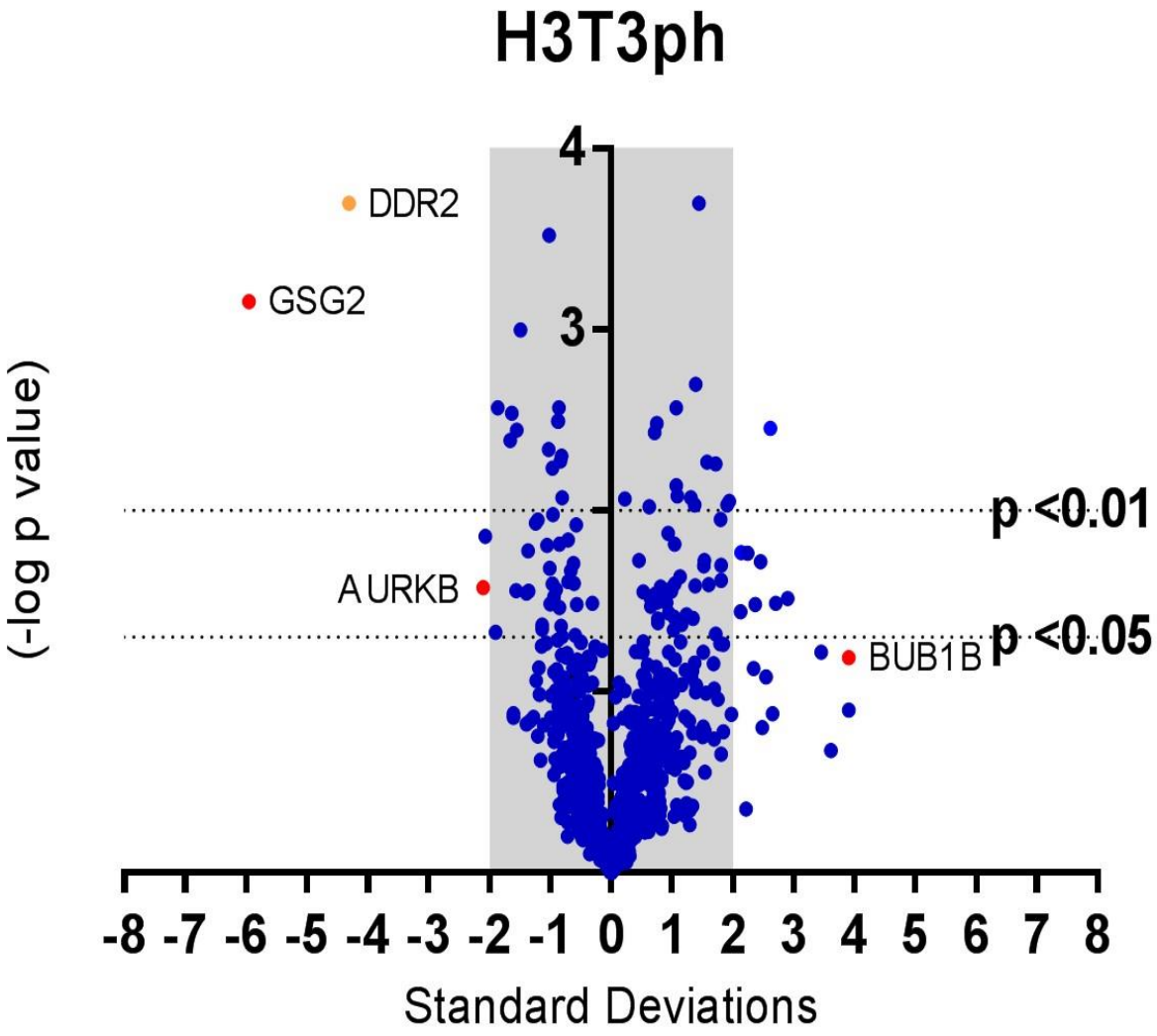
H3T3ph siRNA screen results

The majority of kinases had a relatively small effect on average H3T3ph intensity in arrested mitotic cells following siRNA knockdown; with the majority falling within 2 standard deviations of the mean (as expected for 95% of samples if assuming a normal distribution) (Fig 3.2).

The screen indicated GSG2 (gene name for Haspin kinase) knockdown caused the greatest reduction in H3T3ph. While knockdown of DDR2 and Aurora B were the next most effective in reducing H3T3ph. Haspin is well established as the H3T3 mitotic kinase (Dai et al., 2005), providing confidence in the utility of the screening approach. Furthermore, Aurora B has been shown both to have an essential role in the activation of Haspin (Wang et al., 2011) and a role in preserving H3T3ph from dephosphorylation by phosphatases (Qian et al., 2013). The effect of DDR2 knockdown on H3T3ph was unexpected as there are no reports of a role for DDR2 in H3T3ph in the literature.

siRNA knockdown of BUBR1 caused a large increase in the intensity of H3T3ph, although this was just below the 0.05 level of significance,

Kinase	Change (SD)	p (-log)
GSG2	-5.95	3.15
DDR2	-4.30	3.70
AURKB	-2.11	1.58
GALK1	-2.07	1.86
EDG2	-1.94	
NRBP2	-1.90	1.33
TJP2	-1.86	2.57
C9ORF12	-1.65	2.39
TACR2	-1.63	2.54
ADRA2B	-1.61	0.86
PLK1	-1.60	0.88
HIPK4	-1.56	1.56
PI4K2B	-1.55	2.44
HRI	-1.49	3.00
NEK7	-1.39	0.82
WEE1	-1.39	1.54
ICK	-1.36	1.78
NTRK1	-1.36	1.56
BAI1	-1.33	0.84
DTYMK	-1.28	0.86
GSK3B	-1.24	1.93
PANK4	-1.23	1.06
GPR75	-1.20	1.95
MAPK11	-1.20	0.75
JIK	-1.19	1.13



Kinase	Change (SD)	p (-log)
CMKLR1	1.81	0.65
RPS6KA2	1.81	1.70
PPP1R3A	1.81	1.62
RIOK1	1.84	0.78
PDXK	1.85	1.26
PIM1	1.91	2.03
PPP1R1B	1.94	2.05
MGC8407	1.98	0.88
CHKA	2.13	1.44
SRC	2.14	1.77
MINK	2.22	0.35
CTDSP1	2.25	1.76
PRKCL2	2.35	1.13
TNK1	2.37	1.48
UHMK1	2.46	1.72
ACVR1C	2.49	0.80
CAMK1G	2.55	1.08
EPHA6	2.62	2.46
TEX14	2.65	0.88
EPHA3	2.71	1.49
PYCS	2.90	1.51
CSF1R	3.46	1.22
ITK	3.62	0.68
DGKA	3.91	0.90
BUB1B	3.91	1.19

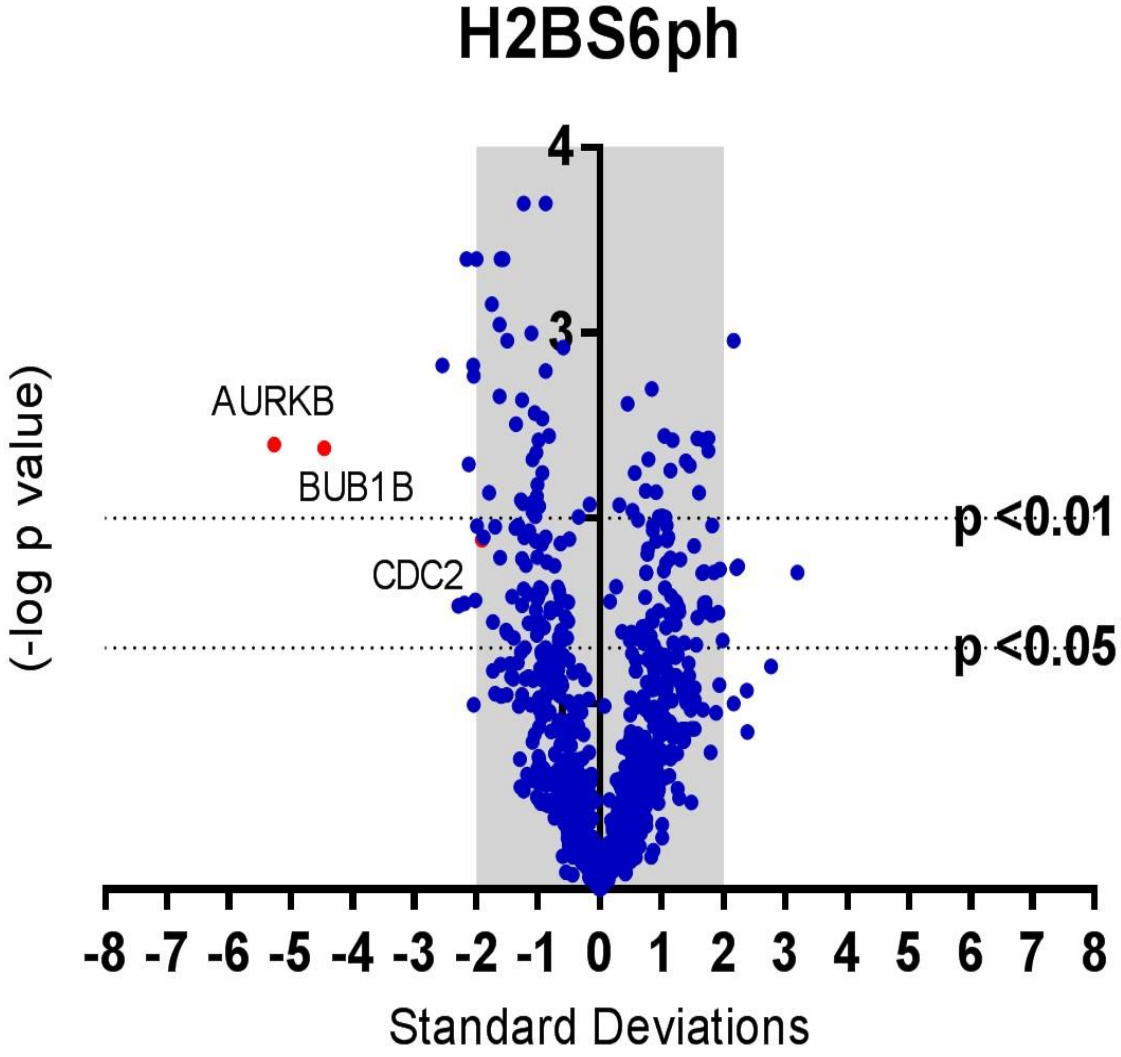
Figure 3.2: : Volcano plot of siRNA kinome screen results (as described in 3.3.1), depicting mean standard deviations from screen average of H3T3ph intensity (integrated) of siRNA treated cells from n=4 repeat wells (x-axis), and the probability of a difference from screen average (-log p values , y-axis; calculated by one-sample t-test). Red (mitotic kinases highlighted by screen), Orange (unexpected high confidence hit)

H2BS6ph siRNA screen results

The mitotic kinases Aurora B and BUBR1 were unambiguously highlighted by the siRNA screen (Fig 3.3). Their knockdown caused a clear reduction in H2BS6ph intensity, resulting in mean mitotic cell intensities more than 4 standard deviations lower than the average for the screen. Knowing that both Aurora B and BUBR1 have highly significant roles in mitosis and were unbiasedly highlighted by the screen increased our confidence in these hits.

Follow up on the results of the screen by our collaborators also led to interest in the role of CDK1 (gene name CDC2) in H2BS6ph (see discussion 3.3.2).

Kinase	Change	p (-log)
AURKB	-5.27	2.40
BUB1B	-4.46	2.38
IGF1R	-2.55	2.82
GPR123	-2.29	1.53
CDKL4	-2.19	1.54
MATK	-2.15	3.40
ADRA2B	-2.12	2.29
PCK2	-2.05	2.82
TACR2	-2.04	2.77
GPR142	-2.04	0.99
EMR3	-2.01	1.56
CDK11	-2.00	3.40
DTYMK	-1.99	1.96
CDC2	-1.91	1.89
PI4K2B	-1.88	1.90
WEE1	-1.79	2.14
DLG1	-1.75	3.15
TJP2	-1.73	1.44
RXFP3	-1.73	1.18
CDKN2C	-1.69	1.95
TRIB3	-1.69	1.06
EDG2	-1.65	
GPR24	-1.62	
STK32A	-1.62	2.66
CDK10	-1.62	3.05



Kinase	Change	p (-log)
FGFR2	1.69	1.71
EIF2AK4	1.69	1.54
CSNK2B	1.70	1.53
GRK7	1.71	1.54
RIPK1	1.74	1.49
UMPK	1.75	2.43
PPP1R1B	1.75	2.37
HTR1A	1.79	0.74
LMTK3	1.82	1.48
PFKL	1.82	1.96
SSTR5	1.84	1.71
PRKCL2	1.88	0.95
SRC	1.91	1.49
STK39	1.94	1.10
PPP1R3A	1.94	1.72
HUNK	1.98	1.34
CDK8	1.99	
STK11	2.16	2.96
G6PC3	2.16	1.00
RELA	2.21	1.73
CSNK1G1	2.24	1.74
ACVR1C	2.38	1.07
DGKA	2.39	0.85
MINK	2.77	1.20
CSF1R	3.20	1.71

Figure 3.3: : Volcano plot of siRNA kinome screen results (as described in 3.3.1), depicting mean standard deviations from screen average of H2BS6ph intensity (integrated) of siRNA treated cells from n=4 repeat wells (x-axis), and the probability of a difference from screen average (-log p values , y-axis; calculated by one-sample t-test). Red (mitotic kinases highlighted by screen)

3.2.3 Further investigation of DDR2 knockdown effects on H3T3ph

The prominence of DDR2 as a hit in the H3T3ph screen led us to investigate the possible role of DDR2 in mitotic phosphorylation of H3T3. To confirm the results of the screen, we repeated the siRNA knockdown of DDR2 (with a new different siRNA: siDDR2 #1) and measured the integrated intensity for a population of mitotic cells (defined by detectable levels of H3S10ph or H3T3ph) (Fig 3.4 A). The effect size we observed was small in comparison to GSG2 knockdown, although it was statistically significant (unpaired t-test). Conversely DDR2 knockdown had no significant effect on H3S10ph intensity in the same cell population.

To ensure we had knocked DDR2 down effectively, we performed a western blot of cells treated in the same way (Fig 3.4 B) and probed for DDR2. Although it was clear we had dramatically reduced the amount of DDR2, the effect size on H3T3ph was again very small. As DDR2 is a receptor tyrosine kinase whose activating ligand is collagen (Leitinger, 2011) we also investigated whether growing cells on collagen might amplify a DDR2 knockdown effect (Fig 3.4 C). Curiously we observed not only a larger effect on H3T3ph but also a clear reduction in H3S10ph. The lack of effect of Dasatanib (an inhibitor of DDR2 kinase activity) was curious, although possibly a result of the very slow and sustained kinetics of DDR2 kinase activity following collagen binding (Vogel et al., 1997).

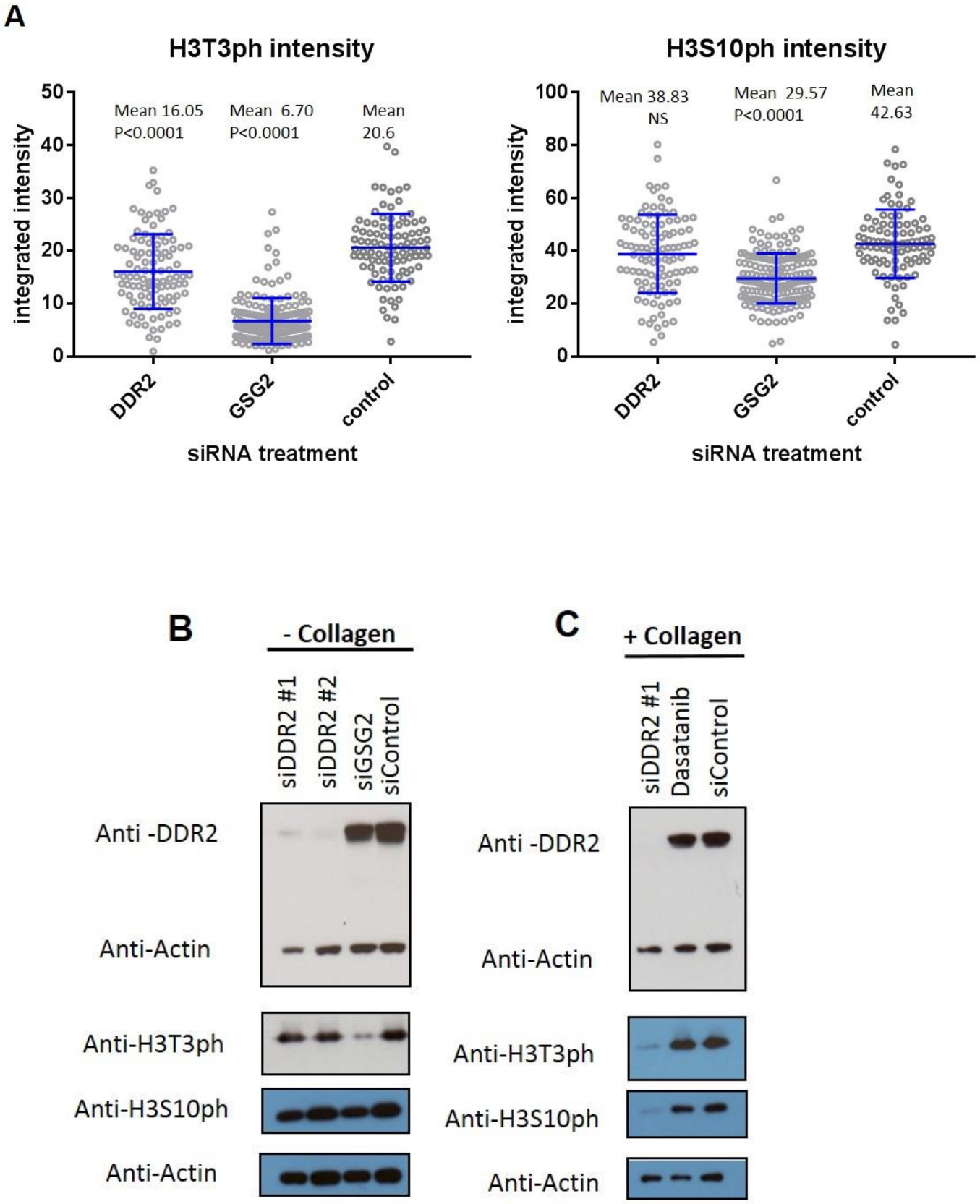

A
H3T3ph intensity
integrated intensity
50
40
30
20
10
0
Mean 16.05
P<0.0001
Mean 6.70
P<0.0001
Mean
20.6
DDR2
GSG2
control
siRNA treatment
H3S10ph intensity
integrated intensity
100
80
60
40
20
0
Mean 38.83
NS
Mean 29.57
P<0.0001
Mean
42.63
DDR2
GSG2
control
siRNA treatment
B
- Collagen
siDDR2 #1
siDDR2 #2
siGSG2
siControl
Anti -DDR2
Anti-Actin
Anti-H3T3ph
Anti-H3S10ph
Anti-Actin
C
+ Collagen
siDDR2 #1
Dasatanib
siControl
Anti -DDR2
Anti-Actin
Anti-H3T3ph
Anti-H3S10ph
Anti-Actin

Figure 3.4: **A** H3T3ph and H3S10ph levels (by integrated intensity) in mitotic Hela cells after 48 hour treatment of siRNA against DDR2, GSG2, or control (cells arrested in 200 nM Nocodazole for 6 hours). **B** Western blots of Hela cell lysates stained for DDR2, H3T3ph, H3S10ph and actin following 48 hour siRNA treatment and 6 hours Nocodazole treatment (200 nmol) **C** As (B) but cells grown on collagen I coated dishes (Dasatanib at 500 nmol for 48 hours)

3.2.4 H3S10ph and H3T11ph siRNA screen

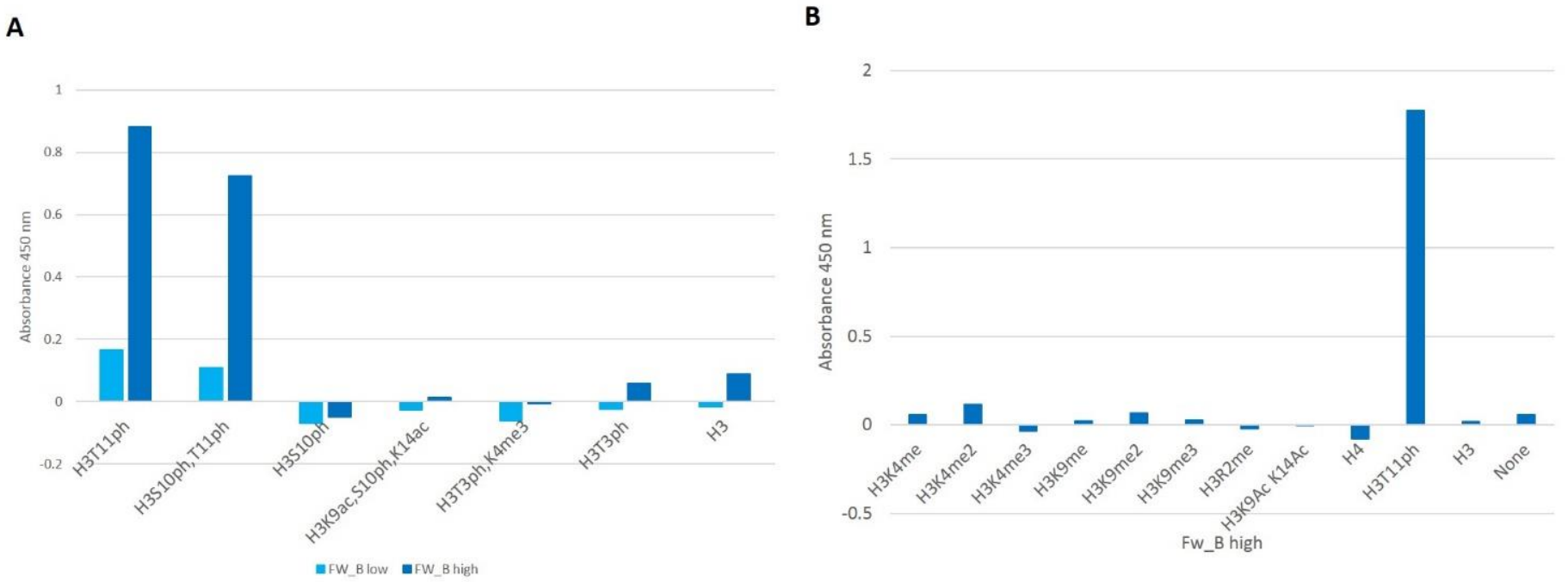
Analysis of H3T11ph antibodies

To aid our investigations into H3T11ph we had a polyclonal antibody raised in rabbits against a H3T11ph modified peptide (antibody Fw_B). Prior to beginning our investigations, it was important to verify the specificity of the new antibody. The need for this was particularly underlined by a report demonstrating broad cross-reactivity across a range of H3 tail modifications for another commercially available H3T11ph antibody (Natalya Nady, 2008). Our approach was to set up a synthetic peptide array of immobilized H3 N-terminal tail peptides corresponding to a range of H3 modifications and combinations of modifications. We then probed our untested H3T11ph primary antibody against the array to assess the titre and any cross-reactivity with other H3 tail modifications by ELISA.

The results indicated a high titre for the antibody. Additionally, they indicated no non-specific cross-reactions with any of the peptide modifications tested. Importantly, the antibody appears uninhibited in H3T11ph recognition by concurrent phosphorylation of the neighbouring H3S10 residue.

Western blotting with the antibody also resulted in a strong band at the expected size of Histone 3 when cells were arrested in mitosis, and no signal was detectable in asynchronous cells (data not shown). Immunofluorescence in fixed cells revealed an unusual pattern of mitotic staining for H3T11ph. While H3T3ph and H3S10ph is observable in all mitotic cells from late G2 up until anaphase (at which point they both decline), H3T11ph staining was very inconsistent in unperturbed mitotic cells.

Small numbers of intensely stained mitotic cells were observable but in most mitotic cells H3T11ph was undetectable by immunofluorescence. In contrast if cells were treated with inhibitors which arrest cells in mitosis (nocodazole or taxol) H3T11ph became observable in almost all mitotic cells. H3T11ph staining was significantly enriched around the centromere, but varied substantially in intensity between mitotically arrested cells (Figure 3.5 C)


A
Absorbance 450 nm
1
0.8
0.6
0.4
0.2
0
-0.2
H3T11ph
H3S10ph,T11ph
H3S10ph
H3K9ac,S10ph,K14ac
H3T3ph,K4me3
H3T3ph
H3
FW_B low
FW_B high
B
Absorbance 450 nm
2
1.5
1
0.5
0
-0.5
H3K4me
H3K4me2
H3K4me3
H3K9me
H3K9me2
H3K9me3
H3R2me
H3K9Ac K14Ac
H4
H3T11ph
H3
None
Fw_B high

C

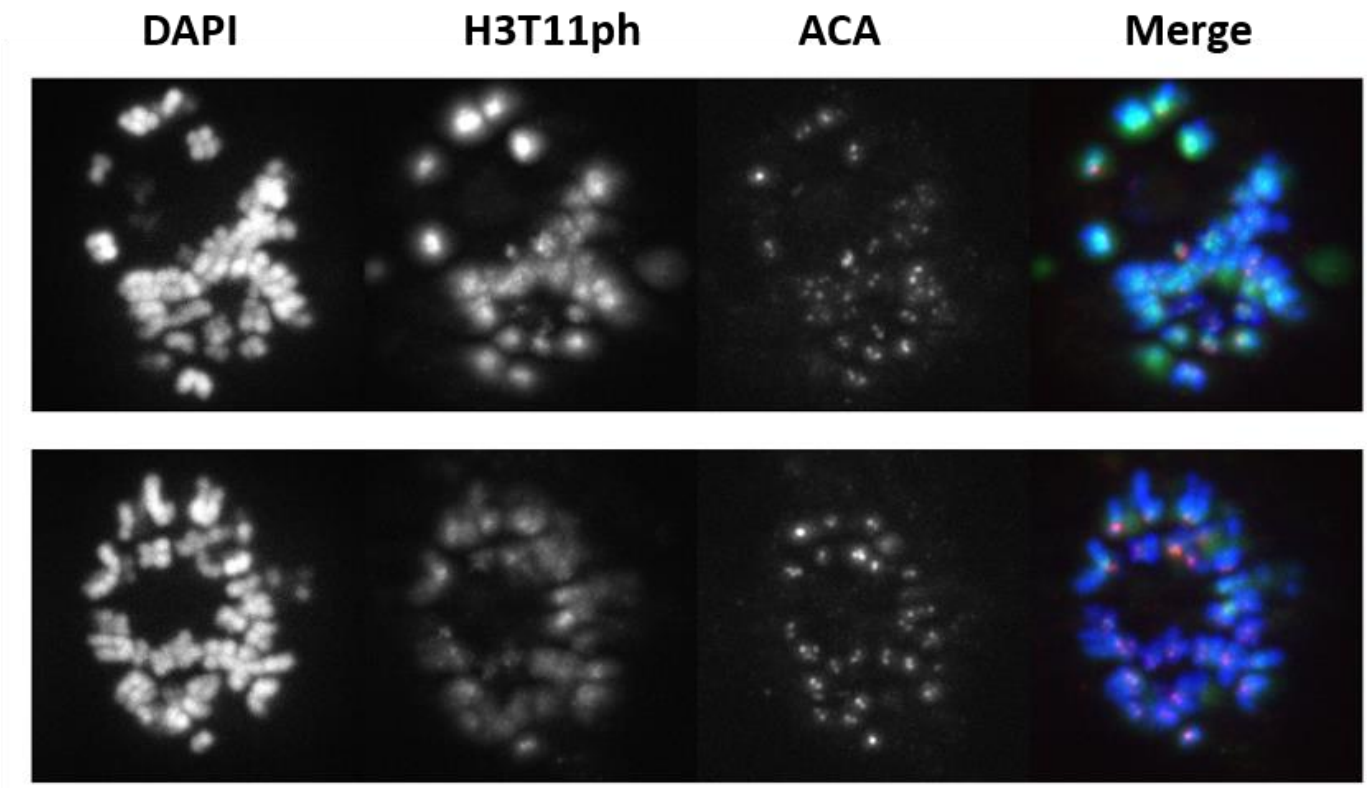


Figure 3.5: **A** H3T11ph antibody (Fw_B) was probed for biding to synthetic peptide chains corresponding to the first 22 amino acids of the N-terminal tail of histone H3 by ELISA. Each peptide carried a single or multiple modifications as labelled and was immobilized on a streptavidin coated plate via c-terminal conjugated biotin. Fw_B low at 1:1250000 dilution, FW_B high at 1:250000 dilution **B** as in (A). **C** Mitotic chromosome spreads were stained with DAPI and immunofluorescently probed with Fw_B and ACA antibodies

Kinome screen

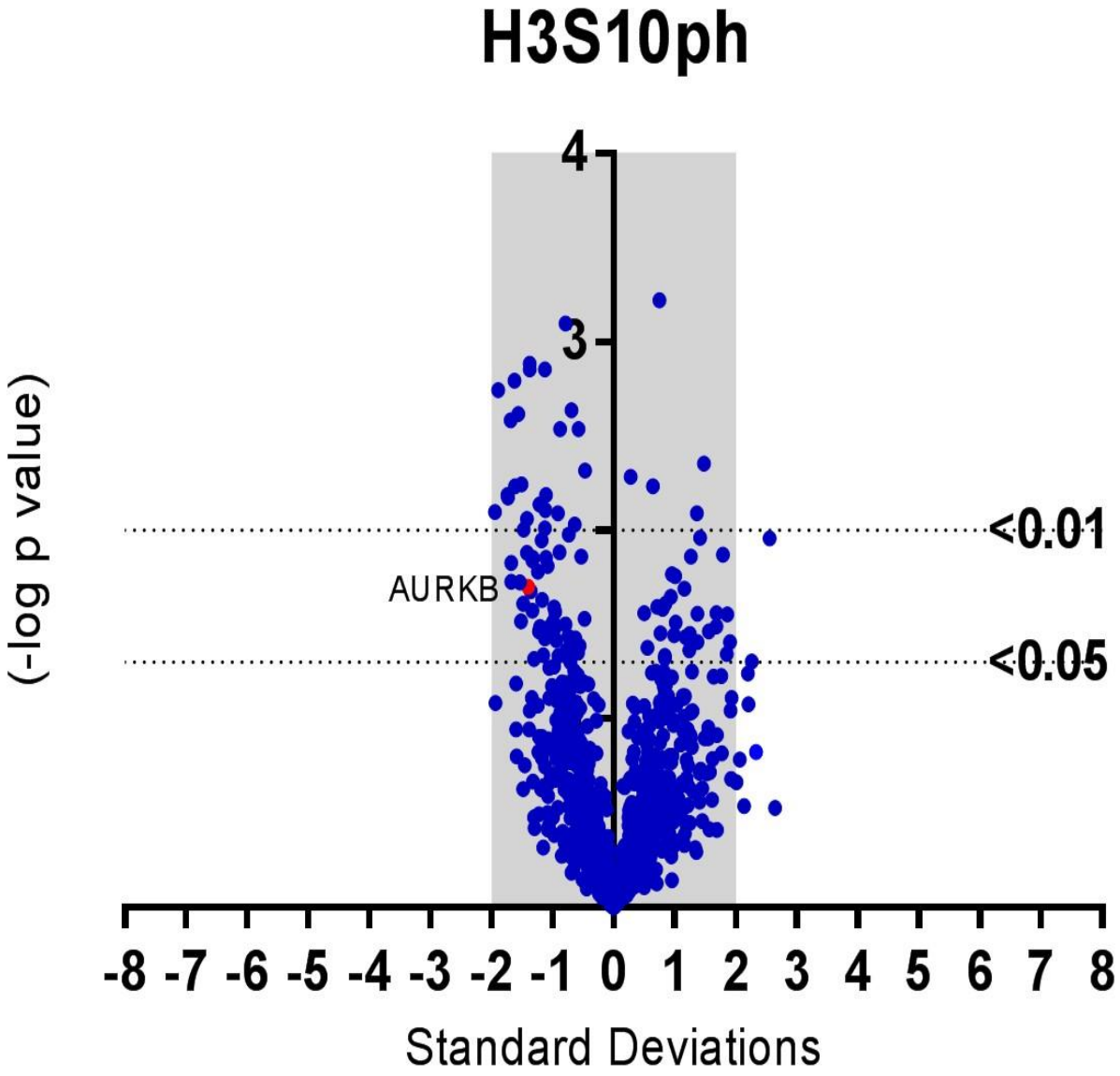
An siRNA kinome screen was then carried out to assess the impact of knockdown of each kinase in the human kinome on H3T11ph (Figure 3.7) and H3S10ph (Figure 3.6) in Taxol arrested mitotic cells. Cells were costained with FW_B and a well validated commercial monoclonal antibody against H3S10ph (05-806, Millipore).

H3S10ph siRNA screen results

The changes we measured in H3S10ph were within a narrower range than for the other histone phosphorylation marks screened. Only PI4KII caused a greater than 2 standard deviations change in H3S10ph and scored a P value <0.05 by one sample t-test (Figure 3.6)

Surprisingly, siRNA knockdown of Aurora B, which has been well characterised as a direct kinase of H3S10ph in mitosis, caused a detectable but only modest decline in H3S10ph intensity measurements (Hsu et al., 2000, Crosio et al., 2002).

Kinase	Change (SD)	p (-log)
ADCK2	-1.944	2.09691
ROCK2	-1.932	1.084073
TJP2	-1.888	2.744727
RAGE	-1.736	2.187087
NME3	-1.725	2.173925
GABBR1	-1.681	2.585027
AGTR2	-1.674	1.728158
ZAP70	-1.673	1.826814
CKMT2	-1.623	2.79588
GPR	-1.613	2.236572
RPS6KC1	-1.599	1.189096
GPR116	-1.592	0.946537
ADRA1B	-1.587	0.803271
PIP5K3	-1.562	2.619789
HRI	-1.535	1.725842
OSR1	-1.514	1.517126
CDK5R2	-1.51	2.244125
MAP2K1	-1.482	1.610834
LMTK2	-1.481	0.628009
LRRK1	-1.467	2.004365
ADRB3	-1.45	0.75597
F2RL3	-1.419	2.060481
NRBP2	-1.419	1.882729
AURKB	-1.401	1.701147
EDG1	-1.385	0.94424



Kinase	Change (SD)	p (-log)
CSNK1G1	1.617	0.572352
EMPTY	1.624	0.78728
BRAF	1.641	1.226214
PIM1	1.68	1.562249
ADK	1.688	1.490797
PRPS1L1	1.691	0.915781
PRKAG3	1.693	0.411728
EPHA4	1.762	1.228413
BDKRB2	1.777	0.818156
BCR	1.794	1.872895
CAMK4	1.855	1.344862
AIP1	1.859	1.554396
CSF1R	1.905	1.406714
PRKG2	1.912	1.04624
PPP1R3C	1.928	0.682146
ADCK1	1.932	1.110698
RXFP3	2.015	0.664943
PFTK1	2.065	0.785951
PRKAR2B	2.135	0.536704
MAP3K8	2.204	1.242604
DAPK1	2.213	1.078314
KSR2	2.261	1.306273
BUB1B	2.333	0.823619
PI4KII	2.556	1.958607
TNK1	2.648	0.529296

Figure 3.6: Volcano plot of siRNA kinome screen results (as described in 3.3.1), depicting mean standard deviations from screen average of H3S10ph intensity (integrated) of siRNA treated cells from n=4 repeat wells (x-axis), and the probability of a difference from screen average (-log p values , y-axis; calculated by one-sample t-test)

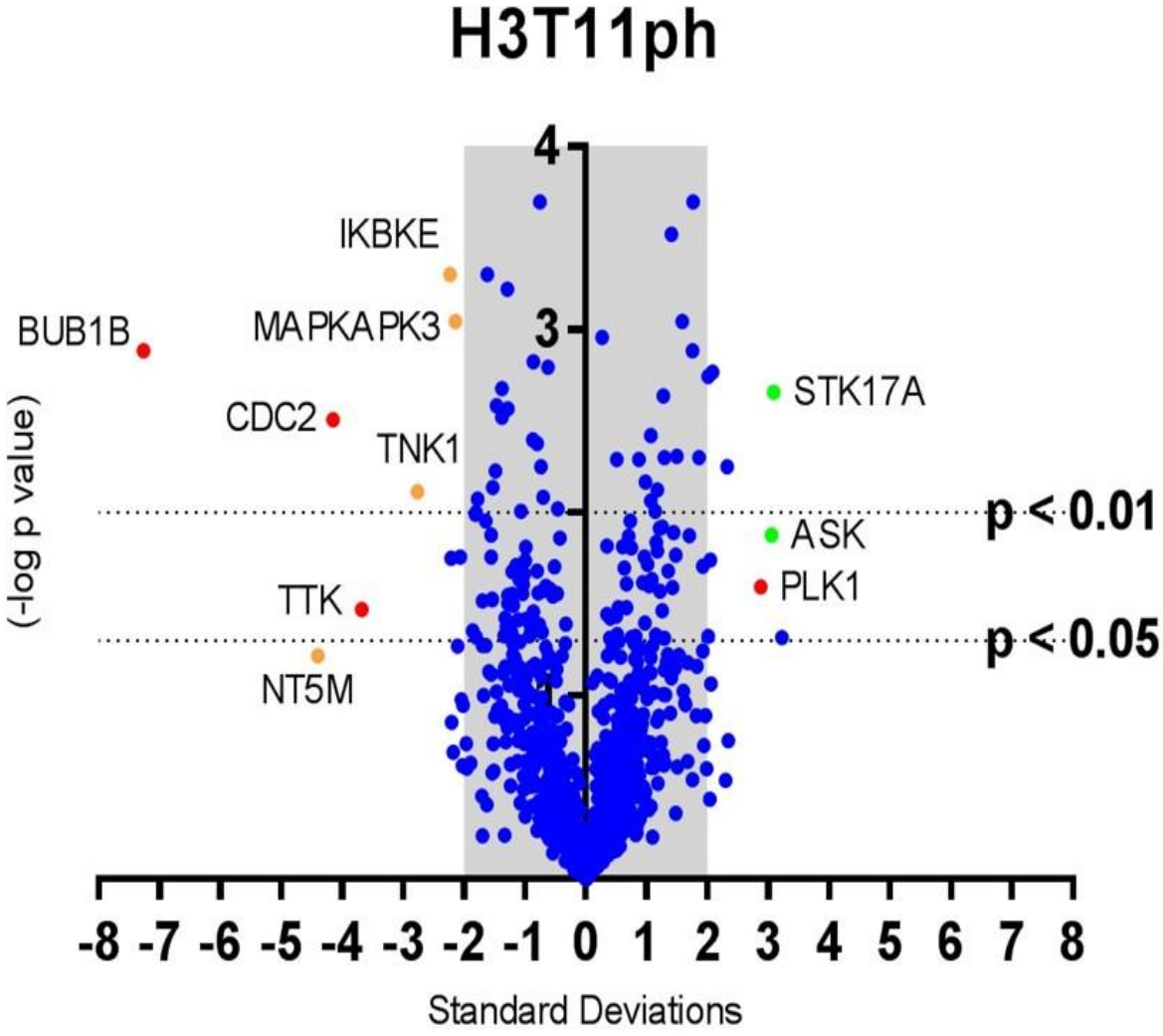
H3T11ph siRNA screen results

The siRNA screen revealed several genes which had a clear impact on H3T11ph intensity relative to the rest of the kinome. Of those with a statistical significance of p=<0.05, knockdown of BUB1B, CDC2 and TTK (genes encoding for BubR1, CDK1, MPS1 respectively) caused the greatest reduction in H3T11ph. Interestingly, these 3 genes are well known mitotic regulators. These were therefore considered our primary candidates as potential H3T11ph regulators.

Also if interest, based on statistical significance and deviation from the standard range, were NT5M, TNK1, MAPKAPK3 and IKBKE.

Genes whose knockdown caused an upregulation in H3T11ph included another prominent mitotic regulator, PLK1. Additionally, it was interesting to note that ASK and STK17A, whose knockdown also resulted in high confidence increases in H3T11ph, both have proposed pro-apoptotic functions (discussed in 3.3.4).

Kinase	Change (SD)	p (-log)
BUB1B	-7.268	2.886057
NT5M	-4.404	1.217527
CDC2	-4.151	2.508638
TTK	-3.685	1.47237
TNK1	-2.766	2.113509
IKBKE	-2.234	3.30103
ITK	-2.213	1.752027
CDK6	-2.209	0.854804
CTDSP1	-2.176	0.690157
MAPKAPK3	-2.141	3.045757
ITPKA	-2.109	1.271646
RXFP3	-2.067	1.759451
AK5	-2.045	0.978811
MAPK9	-2.027	0.617623
MTMR4	-2.02	0.948462
PKMYT1	-1.966	0.737312
MYLK	-1.956	0.605198
NME1	-1.894	0.633204
CDC2L5	-1.856	1.353596
PANK2	-1.814	1.9914
PTPDC1	-1.808	1.320572
CDKN2C	-1.806	2
CHEK1	-1.777	2.075721
EDG1	-1.709	0.453211
GSG2	-1.699	1.519993



Kinase	Change (SD)	p (-log)
VRK2	1.699	1.876148
DLG4	1.75	0.542421
STK36	1.756	2.886057
HSPB8	1.764	3.69897
MARK3	1.811	0.891773
MARK1	1.827	1.161151
CAMKIINALPHA	1.859	2.30103
KHK	1.922	1.244125
PRKCH	1.923	1.707744
PRKAG2	1.943	0.72723
DOT1L	1.967	0.891435
FGFR3	1.987	0.603104
DAPK2	2.01	1.326058
STK24	2.012	2.744727
FLJ32332	2.037	0.435926
PTK2	2.05	1.742321
MAP3K3	2.058	1.064493
PRKWNK1	2.077	2.769551
PRKAB2	2.295	0.539854
OPN1SW	2.317	2.251812
LGR5	2.337	0.754241
PLK1	2.876	1.595166
ASK	3.051	1.879426
STK17A	3.087	2.657577
HUNK	3.219	1.316953

Figure 3.7: Volcano plot of siRNA kinome screen results (as described in 3.3.1), depicting mean standard deviations from screen average of H3T11ph intensity (integrated) of siRNA treated cells from n=4 repeat wells (x-axis), and the probability of a difference from screen average (-log p values , y-axis; calculated by one-sample t-test). Red (mitotic kinases highlighted by screen), Orange (secondary candidates), Green (kinases with proposed apoptotic functions)

3.2.5 Follow up on H3T11ph hits

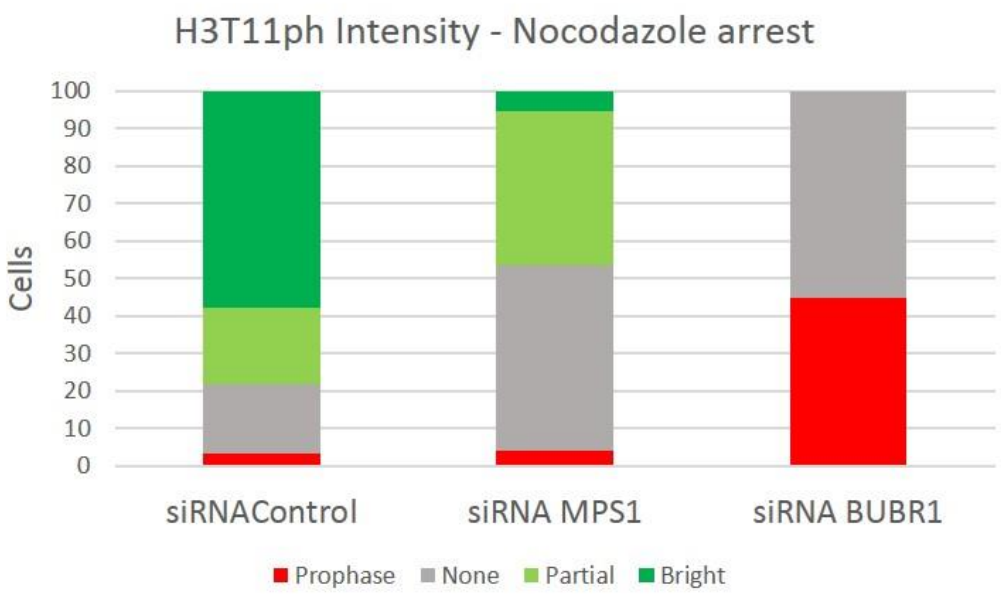
Primary candidates (BUBR1, CDK1, MPS1)

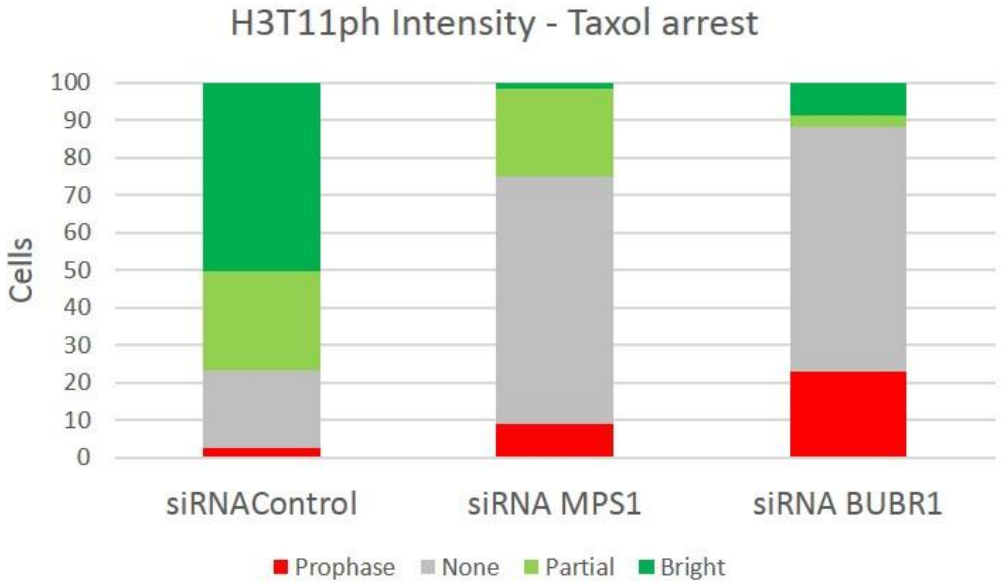
The siRNA screen highlighted BUBR1 as our top candidate for a H3T11ph regulating kinase. However, while BUBR1 is well established as having important functions in mitosis, Suijkerbuijk et al. (2012) present strong evidence that in humans it functions as a pseudokinase with no catalytic activity. It therefore seemed likely that BUBR1 was indirectly regulating H3T11ph. Both BUBR1 and MPS1 (gene name TTK) have central roles in the formation of the mitotic checkpoint complex. We therefore suspected that the highlighting of both BUBR1 and MPS1 in H3T11ph regulation could be due to their role in the mitotic checkpoint.

Taxol induces a mitotic arrest by stabilising microtubules and preventing the spindle assembly checkpoint (SAC) being satisfied; it therefore relies on the mitotic checkpoint complex to mediate this arrest. We therefore suspected that both BUBR1 and MPS1 knockdown were compromising our Taxol induced mitotic arrest. Confirming this, it was clear that far fewer mitotic cells were being detected in BUBR1 or MPS1 siRNA treated wells (approx. 25% and 50% of screen mean respectively). However our image analysis parameters ensured that although fewer cells were being measured these cells were in mitosis (H3S10ph or H3T11ph positive).

To see if the effect of BUBR1 or MPS1 knockdown on H3T11ph could be separated from their roles in the SAC, we repeated the siRNA knockdown of MPS1 and BUBR1 but arrested cells in either nocodazole, taxol or MG132. In contrast to

nocodazole and taxol, MG132 causes cells to arrest in mitosis in a SAC independent manner. The strong loss of H3T11ph observed in our screen was repeated when cells were arrested in either nocodazole or taxol but, in contrast, a dramatic rescue of H3T11ph intensity was observed if MG132 was used (Figure 3.8). We therefore concluded that the reduction in H3T11ph caused by knockdown of these genes was indirect and likely a result of their central roles in the mitotic checkpoint (discussed further in 3.3.4).


H3T11ph Intensity - Nocodazole arrest
Cells
100
90
80
70
60
50
40
30
20
10
0
siRNAControl
siRNA MPS1
siRNA BUBR1
Prophase
None
Partial
Bright


H3T11ph Intensity - Taxol arrest
Cells
100
90
80
70
60
50
40
30
20
10
0
siRNAControl
siRNA MPS1
siRNA BUBR1
Prophase
None
Partial
Bright

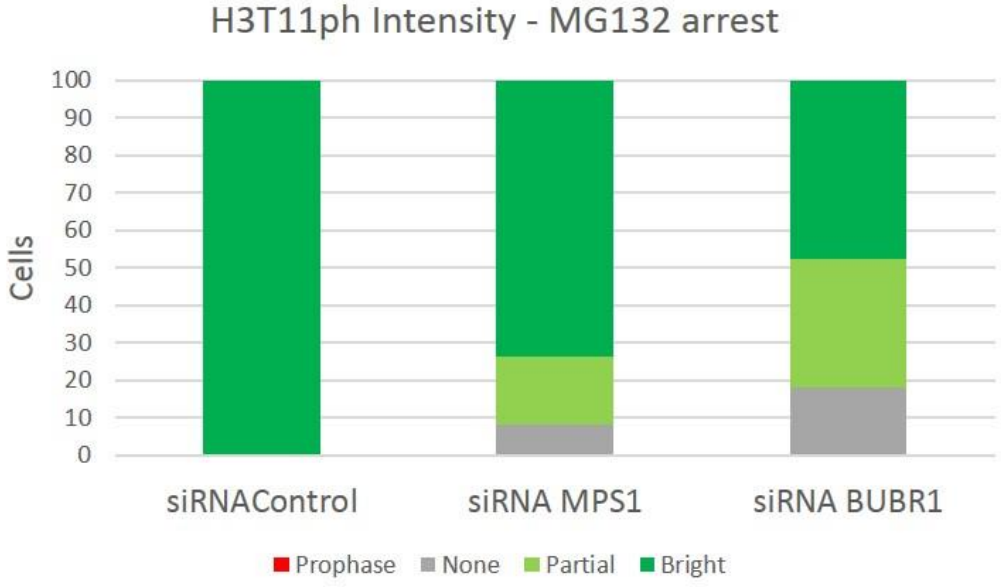

H3T11ph Intensity - MG132 arrest
Cells
100
90
80
70
60
50
40
30
20
10
0
siRNAControl
siRNA MPS1
siRNA BUBR1
Prophase
None
Partial
Bright

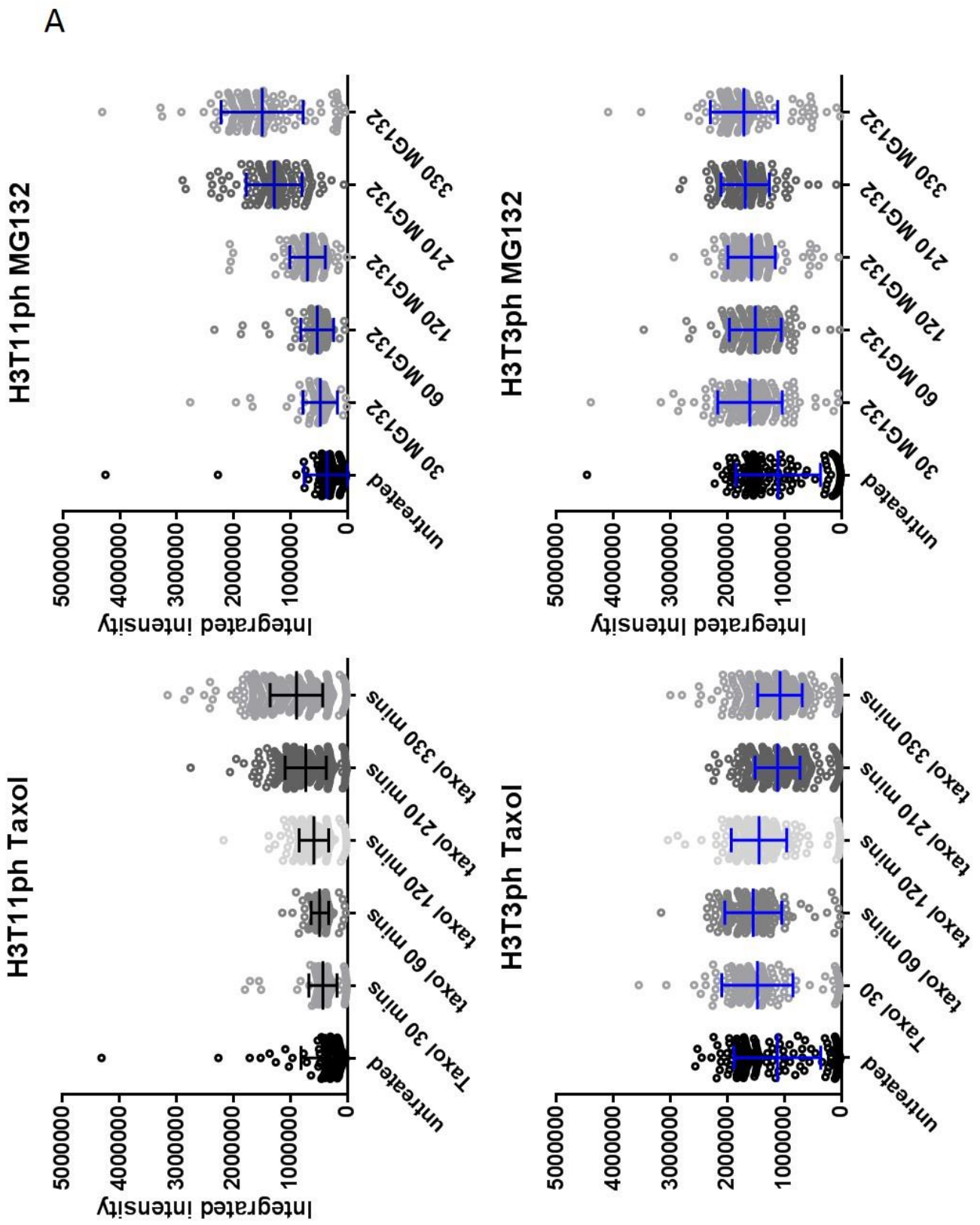
Figure 3.8: Hela cells were treated for 48 hours with siRNA against control, MPS1, or BUBR1. 6 hours before fixation Nocodazole (200 nM), Taxol (300 nM), or MG132 (10 μM) was added to arrest cells in mitosis. Cells were fixed and stained for H3T11ph. Slides were visually scanned for mitotic cells and the first 100 observed were scored for H3T11ph intensity (either: none, partial, bright, or prophase (prophase had no H3T11ph signal))

We were curious that such a dramatic difference in H3T11ph intensity was observed following BUBR1 or MPS1 knockdown despite the fact that the cells we were measuring were still in mitosis (had not yet exited due to SAC compromise). It occurred to us that in addition to reduced mitotic cell numbers rapid mitotic exit due to BUBR1 or MPS1 loss would also result in a population of mitotic cells which had spent less time in mitosis.

To explore if this effect on time spent in mitosis could account for the reduced H3T11ph intensity, we arrested cells in taxol for increasing lengths of time and analysed the mitotic population by high-content imaging. There was a clear increase in H3T11ph intensity across the population of mitotic cells as time spent in taxol arrest increased (Figure 3.9). Moreover, when performing the same experiment using MG132 to arrest cells (which prevents the entry into mitosis of new mitotic cells) it was clear that the entire population was increasing in H3T11ph intensity as mitotic arrest time increased. Interestingly, it appears evident when plotting individual cell intensities in rank order (from the taxol arrest experiment) that this increase in H3T11ph is linear and in direct proportion to time spent in mitotic arrest (Figure 3.9 B). Costaining the same cells for H3T3ph revealed that this intensity increase over time was unique to H3T11ph; in fact, H3T3ph appeared to decrease in cells at 210 minutes or 330 minutes compared to initial mitotic arrest levels (30 mins in Taxol). Interestingly conversely to H3T11ph the time induced change in intensity of H3T3ph were rescued by MG132 (Figure 3.9 A) (see discussion).

Based on this finding we suspected that CDK1 (gene name CDC2) was also being highlighted in our screen due to its knockdown shortening the length of mitotic

arrest (see discussion). To rule out a direct role we tested in mitotic cell extracts if CDK1 inhibition would prevent H3T11ph phosphorylation (see Chapter 4: Figure 4.2), and determined that it did not.


A
H3T11ph MG132
H3T3ph MG132
H3T11ph Taxol
H3T3ph Taxol
Integrated Intensity
5000000
4000000
3000000
2000000
1000000
0
untreated
30 MG132
60 MG132
120 MG132
210 MG132
330 MG132
Taxol 30 mins
Taxol 60 mins
Taxol 120 mins
Taxol 210 mins
Taxol 330 mins

B

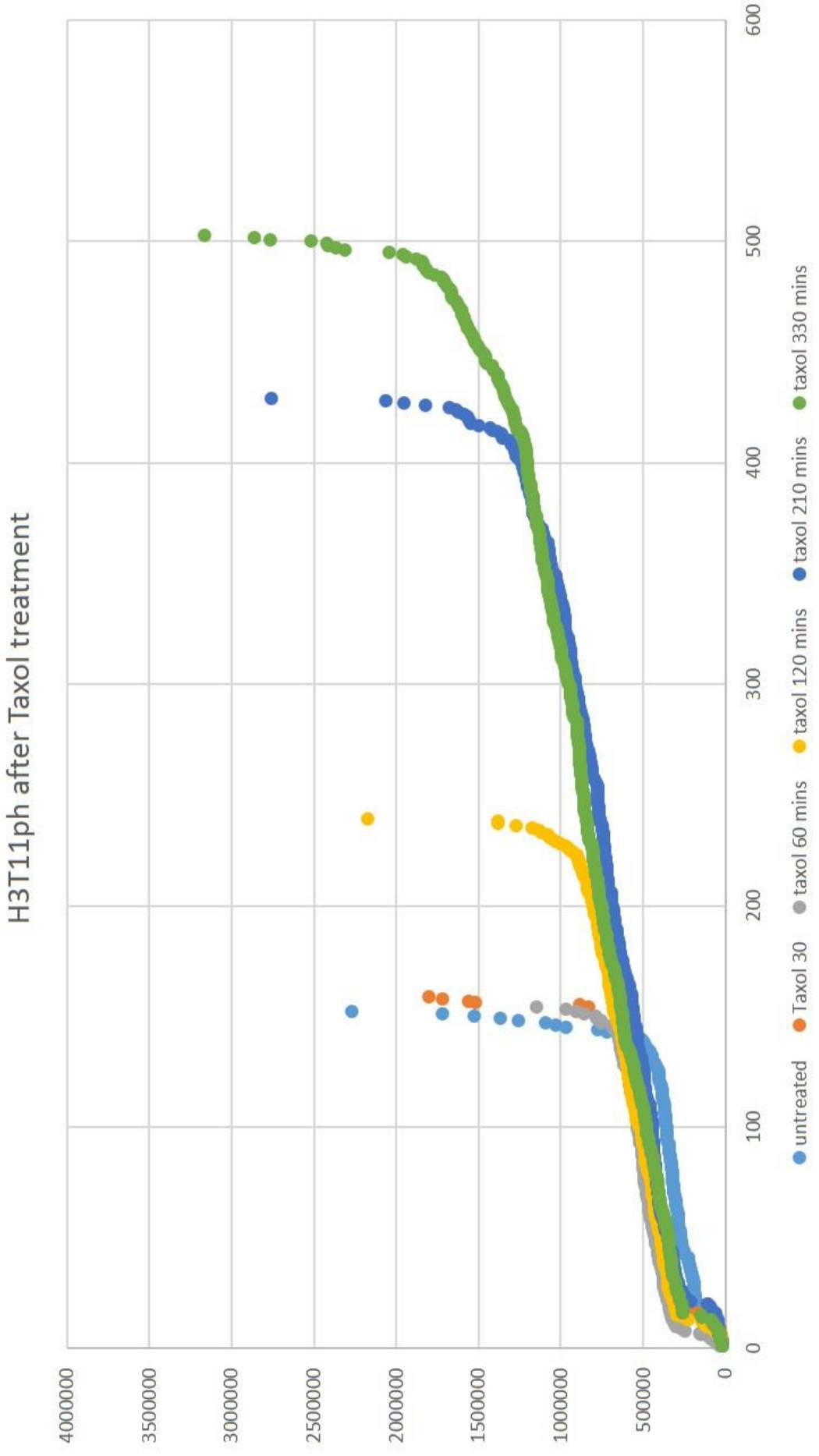

H3T11ph after Taxol treatment
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
0
100
200
300
400
500
600
untreated
Taxol 30
taxol 60 mins
taxol 120 mins
taxol 210 mins
taxol 330 mins

Figure 3.9: **A** Hela cells seeded in parallel were grown in 96 well plates and arrested for 30, 60, 120, 210 or 330 minutes in Taxol or MG132 (or untreated). Cells were fixed and stained for H3T11ph and H3T3ph. Mitotic cell H3T11ph and H3T3ph intensities (integrated) were collected using a microscope equipped with high-content image analysis software (mitotic cells were defined by positive H3T11ph or H3T3ph signal). **B** as in (A), individual cell intensities (H3T11ph) for each arrest time point from Taxol treatment were plotted in rank order (smallest first).

Secondary candidates (NT5M, TNK1, MAPKAPK3 and IKBKE)

As we were unable to identify a direct kinase responsible for H3T11ph from our primary candidates we subsequently investigated possible roles for other kinases less prominently highlighted by the screen.

We repeated the screen parameters with different siRNA against NT5M and MAPKAPK3 (Figure 3.10). The effects on H3T11ph appeared relatively modest compared to GSG2 on H3T3ph - we therefore determined that they were unlikely to be direct kinases for H3T11ph. Follow up experiments on TNK1 and IKBKE have yet to be performed but as TNK1 is a tyrosine kinase and the effect of IKBKE in the screen was modest they are unlikely to have a direct role in H3T11ph.

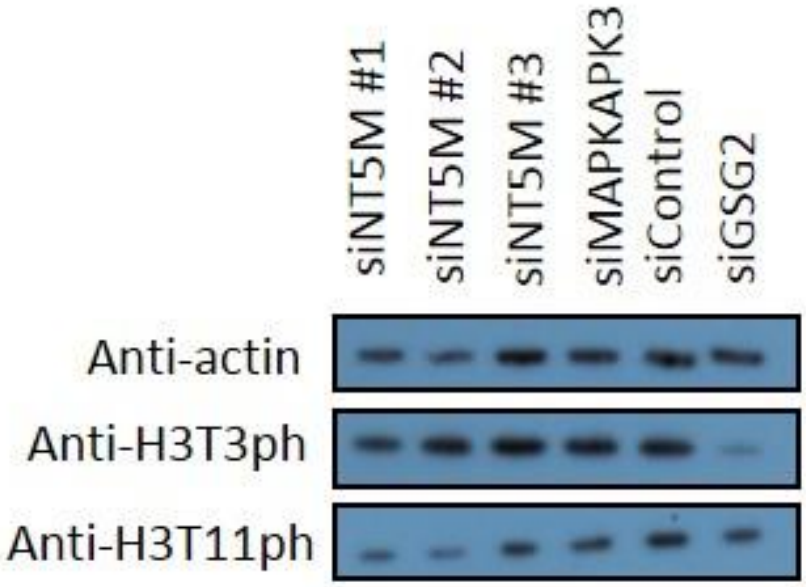


Figure 3.10: Western blots of Hela cells treated for 48 hours with indicated siRNA. Arrested in 300 nmol taxol prior to cell lysis

3.3 Discussion

3.3.1 siRNA screen identification of kinases regulating H3T3ph in mitosis

Haspin (gene name GSG2) is well established as the kinase directly responsible for mitotic phosphorylation of H3T3ph (Dai et al., 2005). The results of our unbiased siRNA screen of the kinome were in line with this, clearly identifying Haspin knockdown as causing greatest reduction in H3T3ph levels. The unambiguous identification of the direct kinase for H3T3ph indicates that the screen was performing well.

Aurora B and BUBR1, kinases with prominent roles in mitosis, were also highlighted by the screen (although with converse effects on H3T3ph). Aurora B has been shown to regulate H3T3ph indirectly through an essential role in the activation of Haspin (Wang et al., 2011) and a role in preserving H3T3ph from dephosphorylation by phosphatases (Qian et al., 2013). The reduction in H3T3ph upon Aurora B knockdown fits well with these reports.

The reasons for increased H3T3ph signal upon BUBR1 knockdown are less clear as an inhibitory role for BUBR1 on H3T3ph has not been reported. The results of (Figure 3.9) indicate that H3T3ph is susceptible to dephosphorylation in an extended mitosis. It was clear that BUBR1 knockdown was compromising our mitotic arrest (only 25% of mitotic cell numbers counted compared to screen mean); therefore the cells measured in this siRNA treatment will have spent less time in mitosis and not incurred H3T3ph dephosphorylation as a result of extended mitosis. The reasons for an extended mitosis causing decreased H3T3ph (and increased H3T11ph) are unknown but in the case of H3T3ph is presumably reflective of increased phosphatase activity on H3T3 relative to kinase activity. In line with reports that slow cyclin B destruction occurs during an active checkpoint (in a proteasome dependent manner) (Brito and Rieder, 2006), the decrease in H3T3ph was rescued if mitotic arrest was induced by the proteasome inhibitor MG132 rather than by spindle poisons (Figure 3.9A). CDK1/Cyclin B has an important role in the activation of Haspin (Ghenoiu et al., 2013) and also inhibits the H3T3ph phosphatase PP1-RepoMan (Qian et al., 2015), therefore its slow degradation over an extended mitosis might be expected to cause a reduction in H3T3ph.

Another possible contribution to increased H3T3ph levels (following BUBR1 siRNA) could be due to BUBR1's role as a binding site for the phosphatase PP2A-B56 (Suijkerbuijk et al., 2012). Although PP1γ/Repo-Man is thought to be the major H3T3ph phosphatase in mitosis (rather than PP2A)(Qian et al., 2011), PP2A has an essential role in targeting Repo-Man to chromatin through dephosphorylation of an Aurora B phosphosite on Repo-Man (S893) which when phosphorylated prevents it binding chromatin (Qian et al., 2013). Possibly, if this PP2A binding site was lost due to BUBR1 knockdown, less dephosphorylated PP1γ/Repo-Man would be available for H3T3ph dephosphorylation and H3T3ph would therefore be increased. Follow up experiments using MG132 or degradation resistant Cyclin B could be performed to assess if BUBR1 siRNA contributed to increased H3T3ph by a mechanism other than shortening mitosis through compromising the SAC.

The effect of DDR2 on H3T3ph was unexpected and has not been reported previously. The t-test performed on the siRNA screen results indicated a very high level of confidence in a difference in DDR2 siRNA treated H3T3ph levels (compared to screen average). Moreover, a DDR2 siRNA induced reduction in H3T3ph intensity was reproducible using a different siRNA, although the effect size was small in comparison to Haspin knockdown. Interestingly, if cells were grown on collagen (the ligand for DDR2) the effect of DDR2 siRNA on H3T3ph levels appeared much more pronounced. However, there was also a clear effect on H3S10ph in this case. The fact H3S10ph was also reduced could indicate that proliferation rate and therefore % of cells in mitosis was reduced by DDR2 knockdown. In line with this there are several studies in the literature reporting a role for DDR2 in proliferation (Labrador et al., 2001, Olaso et al., 2002, Marquez and Olaso, 2014).

3.3.2 siRNA screen identification of kinases regulating H2BS6ph in mitosis

Our investigations into H2BS6ph were part of a collaborative study led by Markus Seibert and Lienhard Schmitz at Justus-Liebig-University, Giessen, Germany. Their further work has confirmed the importance of the top hit in our screen, Aurora B, for regulating H2BS6ph in cells. They determined that Aurora B has a crucial but

indirect role in facilitating H2BS6ph through the inhibition of PP1 – in fact Aurora B was dispensable for H2BS6ph if PP1 phosphatases were depleted by RNAi (Seibert et al. 2019). This is in line with several reports in the literature describing a role for Aurora B phosphorylation in preventing PP1 chromatin binding (Qian et al., 2013, Nasa et al., 2018).

Interestingly, Seibert and Schmitz identified CDK1 (gene name CDC2) as the likely direct kinase responsible for H2BS6 phosphorylation in mitosis. In our screen CDC2 was #14 in kinases whose knockdown caused reduction in H2BS6ph. Because CDK1/Cyclin B is necessary for cells to enter and maintain a mitotic state (Vassilev et al., 2006) the cells we measured had necessarily an incomplete knockdown of CDK1, possibly accounting for its relative lack of prominence in H2BS6ph reduction. Cells in which CDK1 was knocked down very effectively would have been unable to enter or maintain mitosis, and therefore were not measured.

BUBR1 siRNA also significantly reduced H2BS6ph intensity. We hypothesize that this could also be due to BUBR1 siRNA's effect of shifting our mitotic population to cells that more recently entered mitosis (via compromising the SAC). Early in mitosis Aurora B is observable along the length of the chromosome arms, and becomes concentrated at centromeres as mitosis progresses (Hindriksen et al., 2017). BUBR1 RNAi, in shifting our mitotic population to cells that more recently entered mitosis (via compromising the SAC), might be skewing the cells we measure to those where a PP1 inhibiting critical mass of Aurora B has not yet accumulated at the centromere (where H2BS6ph is observed). In line with this immunofluorescence staining of a population of mitotic cells reveals a strikingly close relationship between Aurora B accumulation at the centromere and H2BS6 phosphorylation (data not shown).

3.3.3 siRNA screen identification of kinases regulating H3S10ph in mitosis

The results of our H3S10ph screen were within a much narrower range than for the other phosphorylation sites screened. Knockdown of only one kinase caused a change in H3S10ph intensity of more than 2 SD from the screen mean (PI4KII). While knockdown of the well-established direct kinase for H3S10ph in mitosis,

Aurora B, resulted in a reduction in H3S10 intensity by 1.4 SD (Crosio et al., 2002, Fischle et al., 2005). Among known Aurora B substrates, H3S10ph, in our hands and others, is difficult to knockdown by Aurora B siRNA (Girdler et al., 2006), although it can be depleted with Aurora B targeting small molecule inhibitors (data not shown). The apparent favourability of H3S10 as a substrate of Aurora B likely accounts for our seeing an effect of Aurora B depletion on H3T3ph and H2BS6ph, yet not for H3S10ph. In the absence of prominent effects, it was difficult to conclude anything about the regulation of H3S10ph by kinases in mitosis based on our screen.

3.3.4 siRNA screen identification of kinases regulating H3T11ph in mitosis

The most prominent hits from the siRNA screen (BUBR1, TTK, and CDC2) appeared to be regulating H3T11ph indirectly by reducing the length of mitotic arrest. BUBR1 and MPS1 (gene name TTK) are both required for proper functioning of the spindle assembly checkpoint (Stucke et al., 2002) (Sudakin et al., 2001). Their knockdown would therefore compromise a taxol mediated mitotic arrest (which depends on SAC signalling). In confirmation of their indirect role in regulating H3T11ph we observed a rescue of H3T11ph signal in BUBR1 or TTK siRNA treated cells when MG132 was used to arrest cells in mitosis (it causes a mitotic arrest independent of the SAC). Similarly, CDK1 activity (gene name CDC2) is also necessary for the SAC (Rattani et al., 2014).

Having established that H3T11ph has the interesting characteristic of slowly accumulating in intensity during an extended mitosis, several other kinases highlighted by the screen may also have arisen due to this feature. Several groups have demonstrated that extending mitosis for several hours leads to apoptosis in increasing numbers of cells (apoptotic death begins after about 5 hours), as a result of the gradual degradation of anti-apoptotic protein MCL-1 (Haschka et al., 2015, Topham and Taylor, 2013). Consequently, as H3T11ph appears to increase in a linear manner in an extended mitosis, and our siRNA screen entailed a 6 hour taxol arrest, one might expect knockdown of kinases which contribute to this apoptotic response to cause a net increase in H3T11ph in our experimental setup (as cells would survive for longer). In line with this supposition, both ASK and STK17A have

known apoptotic functions. ASK induces apoptosis via JNK and p38 signalling pathways both of which are known to lead to phosphorylation-induced ubiquitin mediated degradation of MCL-1 (Topham and Taylor, 2013). Furthermore, Taxol arrest has specifically been shown to induce ASK dependent apoptosis (Wang et al., 1999). STK17A has been much less studied but also induces apoptosis by a mechanism that is not yet clearly defined (Inbal et al., 2000, Oue et al., 2018). The highlighting of STK17A in our screen suggests that it could also play a role a role in the pathways that lead to cell death in an extended mitosis (Topham and Taylor, 2013).

The increase in H3T11ph after PLK1 siRNA might also be due to it causing an alteration in the mitotic arrest time of cells recorded in our screen. While a complete absence of PLK1 activity prevents mitotic entry (Gheghiani et al., 2017), incomplete inhibition (as might be expected by siRNA knockdown) leads to a prometaphase arrest of mitotic cells via the SAC (Petronczki et al., 2008). Consequently, it is likely that a great proportion of PLK1 siRNA treated cells in our screen became arrested in mitosis prior to taxol addition and therefore had been in mitosis longer. As a H3T11ph upregulating hit, PLK1 is by definition an indirect regulator of H3T11ph. The veracity of this explanation could be explored by repeating PLK1 inhibition in a SAC-depleted background (such as Mad2 deficient cells) and using MG132 for mitotic arrest. This would prevent PLK1 siRNA inducing a mitotic arrest and thereby ensure that any effects of PLK1 on H3T11ph were not the result of differences in time spent in mitosis.

The other kinases that we identified as secondary candidates (NT5M, TNK1, MAPKAPK3, IKBKE) would require further work to understand their possible roles in H3T11ph regulation.

In conclusion, although we gained insight into the regulation of H3T11ph we were unable to identify a clear candidate for the mitotic H3T11ph kinase by siRNA kinome screening.

Chapter 4. KiPIK screening: a novel method to identify the kinase responsible for a phosphorylation event of interest

4.1 Introduction

4.1.1 Overview

We were disappointed by the limitations in available techniques for identifying the direct kinase of a phosphorylation site of interest. We devised a novel methodology to overcome some of these limitations and assessed its efficacy.

4.1.2 Small molecule kinase inhibitors

Small molecules that inhibit the enzymatic activity of kinases have been the focus of extensive research and development over the past 30 years. Most commonly these inhibitors function by occupying the adenosine binding pocket (required for ATP binding) of the kinases they target and are therefore termed ATP competitive inhibitors. Non-ATP-competitive inhibitors which exert inhibitory effect by binding outside the ATP binding site also exist but account for less than 1% of reported kinase inhibitors (Dar and Shokat, 2011, Breen and Soellner, 2015).

4.1.3 Specificity and profiling of kinase inhibitors

Because small molecule kinase inhibitors have almost all been developed to target the ATP binding site, and this is highly structurally conserved between kinases, polypharmacology is common in kinase inhibitors. In recent years improvements in the availability and kinome coverage of panels of recombinant kinases have allowed researchers to profile the specificity of large numbers of kinase inhibitors against hundreds of kinases *in vitro* (Bain et al., 2007, Davis et al., 2011, Elkins et al., 2016).

4.1.4 Potential of kinase Inhibitor specificity information for identifying kinase involvement in specific phosphorylation events

We wondered whether the expanding profiling information on kinase inhibitor specificities might allow kinase inhibitors to be used as exploratory tools for identifying kinases responsible for specific phosphorylation events. In an ideal situation, monospecific kinase inhibitors would be available for every kinase in the kinome and one could identify the kinase acting on a phosphosite of interest by seeing which of these inhibitors depleted the phosphorylation. Although monospecific inhibitors are not available, quantitative inhibitory profiling information for hundreds of inhibitors (across hundreds of *in vitro* kinases) is available (as described in 4.1.3). If one were able to test large numbers of these profiled inhibitors for inhibitory action against a specific phosphorylation event their detailed specificity information might allow identification of the phosphorylating kinase by triangulation or correlation.

However, if kinase inhibitors were applied *in vivo* as a screening tool to identify kinases one would encounter the problems associated with indirect effects inherent in all screening approaches in intact cells (see 1.4.3 and chapter 3). Moreover, the non-specificity of each inhibitor would likely undergo convolution as a result of interaction with the indirect effects inherent in screens in intact cells; greatly complicating subsequent attempts to identify the direct kinase with the inhibitor profiling data.

To mitigate these concerns, we decided to focus on developing a technique using cell extracts rather than intact cells. There is good evidence that kinase reactions using cell extracts preserve physiological kinase-phosphosite dependencies (Yu et al., 2009, Deibler and Kirschner, 2010, Wang et al., 2011). Therefore, if we could identify the kinase directly responsible for a phosphorylation event in extracts it seemed likely these would reflect *in vivo* dependencies; a notion we intended to test on several positive controls.

4.1.5 *The KiPIK method*

The method we have developed, and termed KiPIK screening (Kinase inhibitor Profiling to Identify Kinases), proceeds in brief as follows (see 2.7 for detailed methodology). Cell conditions in which phosphorylation of the protein residue of interest is robust are identified. For example, cells at a particular cell cycle stage can be enriched (e.g in mitosis, see 4.2.1 below), or signalling pathways leading to phosphorylation can be triggered (e.g. EGF stimulation, see 4.2.9). Whole cell extracts of these cells are prepared in the presence of phosphatase inhibitors to "freeze" kinases in their active state. These cell extracts are then used as a source of all potentially relevant kinases for *in vitro* kinase reactions in the presence of a substrate (containing the phosphorylation target residue of interest). Multiple of these *in vitro* reactions are performed in parallel using the same preparation of cell lysate, each in the presence of one member of a profiled panel of kinase inhibitors. This yields a unique pattern of inhibition (of the phosphorylation event of interest) that can be compared to the known inhibition patterns of all kinases tested in the profiling panel.

We have used biotinylated peptides encompassing phosphosites of interest as substrates and found these sufficient to recapitulate the *in vivo* phosphosite dependencies tested.

The development and validation of KiPIK screening is detailed below.

4.1.6 *Aims*

1. Develop an assay that utilises kinase inhibitor profiling data to provide information on kinases acting on specific phosphorylation sites.
2. Assess the effectiveness of this assay by testing it on characterised kinase-phosphosite pairs

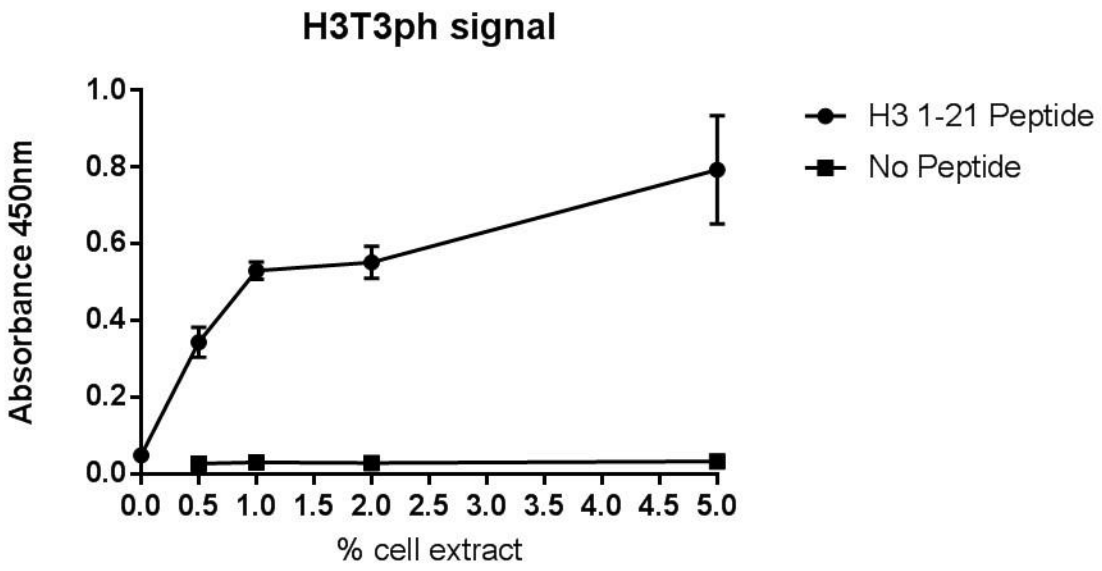
4.2 Results

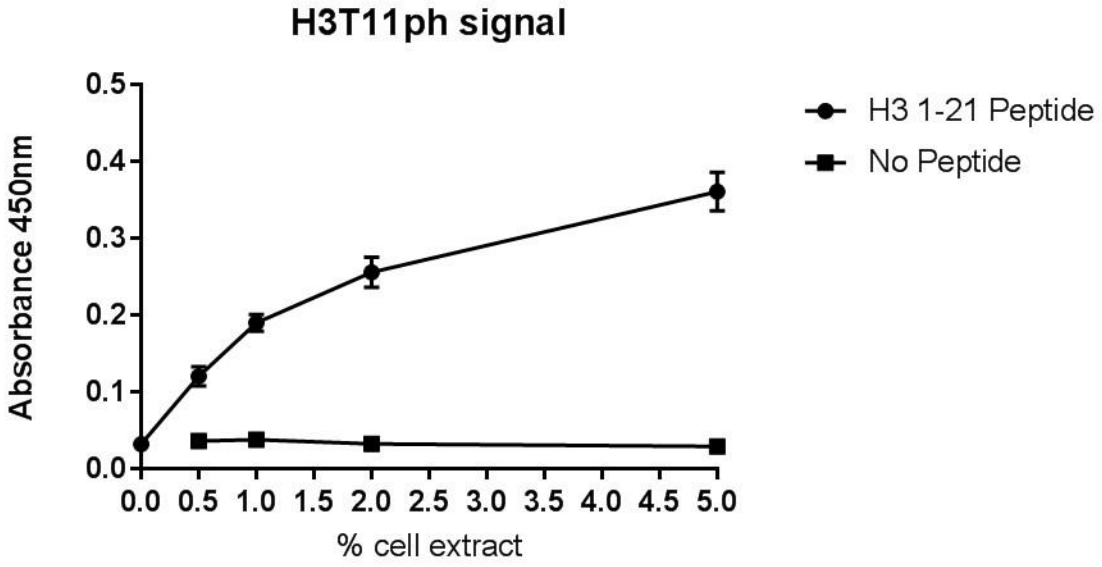
4.2.1 Cell extract can be used to phosphorylate peptides on residues corresponding to known in vivo phosphorylation sites

We began by assessing the feasibility of using cell extracts to generate phosphorylation on peptides which corresponded to *in vivo* phosphorylated regions of Histone H3. Briefly, mitotic cell extracts were prepared and kinase reactions carried out by incubating the extract in kinase reaction buffer with unphosphorylated peptides. After the kinase reaction, peptides were immobilised on streptavidin coated plates and phosphorylation levels determined by ELISA.

Our ELISA results (Figure 4.1) indicated clear extract dependent phosphorylation of H3T3, H3T11 and H3S28. For both H3T3 and H3T11, the phosphorylation signal was also wholly dependent on the presence of the H3 1-21 peptide substrate. For H3S28, the background signal was higher – with an increase in detected phosphorylation with increasing extract independent of the peptide substrate. This was likely due to an imperfect wash procedure for removing extract proteins from the immobilised biotinylated peptide substrate. Nevertheless, the H3S28ph signal measured with the addition of peptide substrate was higher, indicating that the peptide was being phosphorylated by the kinases in the cell extract. In general, the highest signal to noise ratio was found when 5% of the kinase reaction was extract.

We had therefore established that we could use cell extract as a source of kinases for generating *de novo* phosphorylation on peptide residues analogous to multiple known *in vivo* phosphorylation sites on Histone H3.


H3T3ph signal
Absorbance 450nm
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
% cell extract
H3 1-21 Peptide
No Peptide


H3T11ph signal
Absorbance 450nm
0.5
0.4
0.3
0.2
0.1
0.0
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
% cell extract
H3 1-21 Peptide
No Peptide

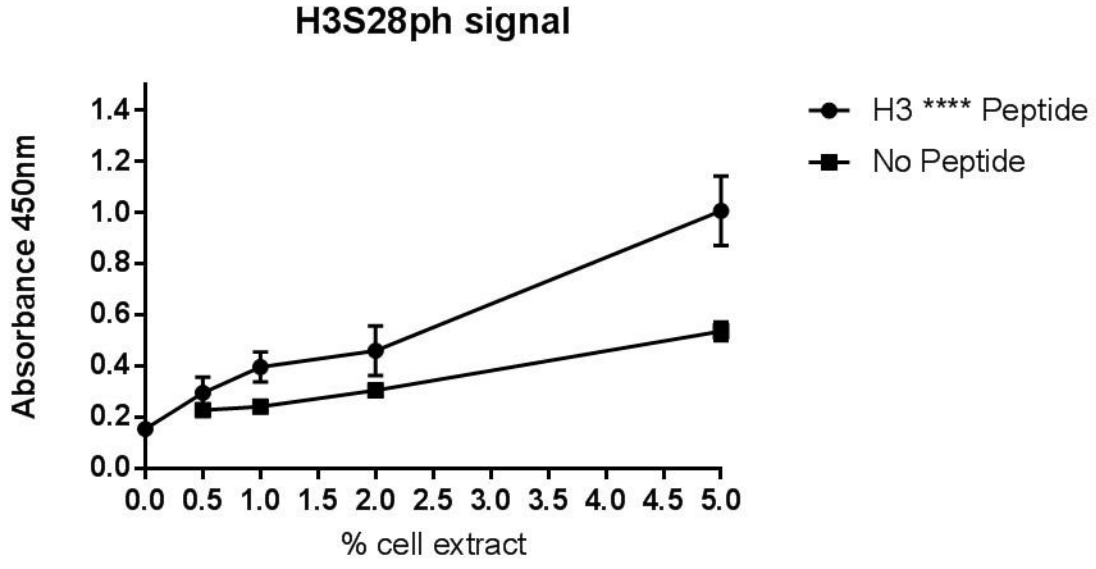

H3S28ph signal
Absorbance 450nm
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
% cell extract
H3 **** Peptide
No Peptide

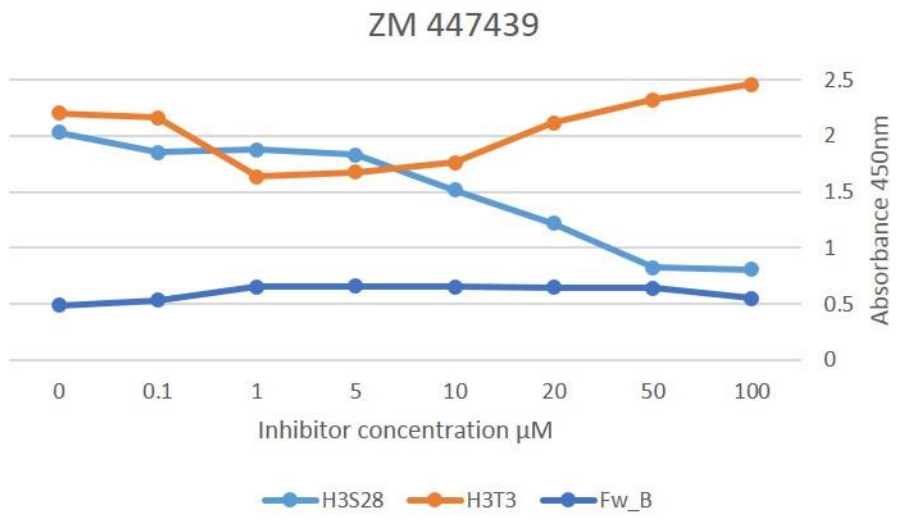
Figure 4.1: **Phosphorylation of H3 peptides with mitotic cell extract** : Peptides corresponding to residues 1-21 or 22-42 of histone H3 were incubated in kinase buffer (containing phosphatase and protease inhibitors) for 30 mins at 30°C with 0%, 0.5%, 1%, 2%, or 5% mitotic Hela cell extract. After 30 minutes peptides were immobilized on streptavidin plates and probed with antibodies for H3T3ph, H3T11ph, or H3S28ph. Signal was quantified by ELISA.

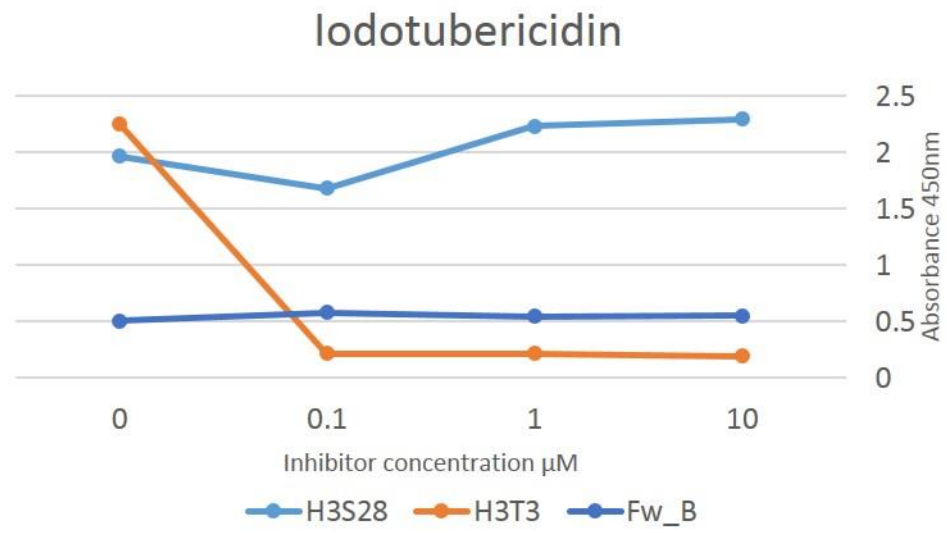
4.2.2 *Kinase Inhibitors can indicate target kinase involvement (or not) in a cell extract based kinase reaction*

Next we sought to determine whether the phosphorylation generated on peptides could be inhibited with small molecule kinase inhibitors. To assess this, we challenged each of our kinase reactions (as in Figure 4.1, 5% extract for all) with increasing concentrations of 3 small molecule kinase inhibitors (Figure 4.2). ZM447439 is a potent Aurora B kinase inhibitor (Ditchfield et al., 2003), and can be used in cells to inhibit phosphorylation of known Aurora B regulated phosphosites such as H3S28ph. 5-Iodotubericidin is a potent inhibitor of Haspin (De Antoni et al., 2012, Wang et al., 2012). RO-3306 is a potent CDK1 inhibitor but also has strong inhibitory activity against Haspin (Patnaik and Higgins, unpublished; LINCS 2018).

The H3T3ph signal was highly sensitive to the addition of 5-Iodotubericidin or RO-3306, in line with their known inhibitory activity against Haspin. H3S28ph had little response to these inhibitors (except perhaps RO-3306 at very high concentrations), yet was clearly potently inhibited by ZM447439. H3T11ph appeared uninhibited by any of the 3 compounds tested (as discussed earlier (3.1.5), the mitotic kinase for H3T11ph is unknown).

The sensitivity of H3T3ph and H3S28ph specifically to inhibitors of their known *in vivo* kinases indicated that these specificities were being retained in the extract assay. Therefore, if we could determine the kinases acting on other phosphorylation sites in this type of *ex vivo* assay, then we would have a method to identify the in vivo kinases for “orphan” phosphorylation sites.


ZM 447439
2.5
2
1.5
1
0.5
0
Absorbance 450nm
0
0.1
1
5
10
20
50
100
Inhibitor concentration µM
H3S28
H3T3
Fw_B


Iodotubericidin
2.5
2
1.5
1
0.5
0
Absorbance 450nm
0
0.1
1
10
Inhibitor concentration µM
H3S28
H3T3
Fw_B

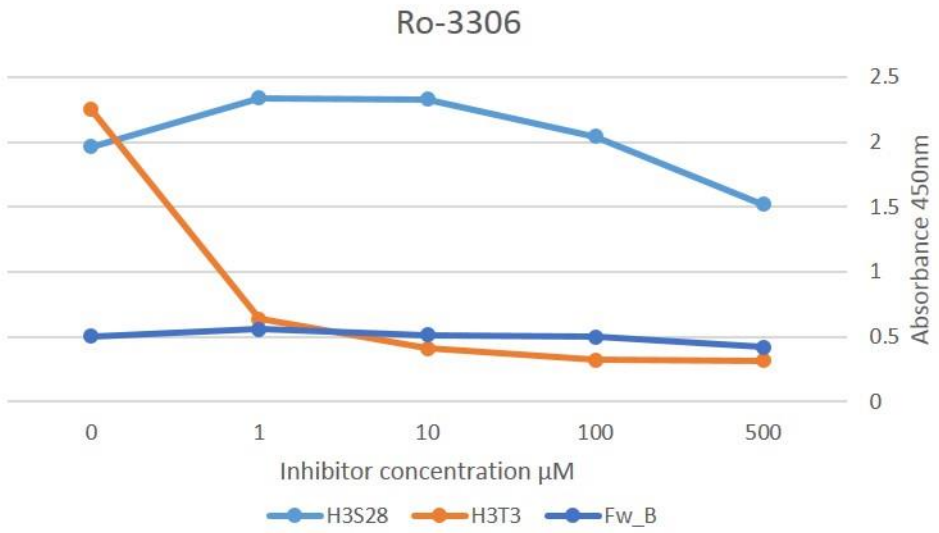

Ro-3306
2.5
2
1.5
1
0.5
0
Absorbance 450nm
0
1
10
100
500
Inhibitor concentration µM
H3S28
H3T3
Fw_B

Figure 4.2: **Inhibition of H3 mitotic extract phosphorylation with small molecule inhibitors** : Peptides corresponding to residues 1-21 or 22-42 of histone H3 were incubated in kinase buffer (containing phosphatase and protease inhibitors) for 30 mins at 30°C with 5% mitotic Hela cell extract and indicated concentrations of ZM-447438, Iodotubericidin or RO-3306. After 30 minutes peptides were immobilized on streptavidin plates and probed with antibodies for H3T3ph, H3T11ph, or H3S28ph. Signal was quantified by ELISA.

4.2.3 Screening panels of kinase inhibitors in parallel produces distinct and reproducible patterns of inhibition for specific phosphorylation events

A screening approach based on inhibitor profiling data would be far more powerful if multiple kinase inhibitors could be compared. To obtain comparable inhibition data for multiple kinase inhibitors, it is important that *ex vivo* assays are performed in parallel using the same cell extract. For this purpose, a Biomek FX liquid handling robot was used to enable large numbers of *ex vivo* assays to be performed in parallel using the same freshly thawed batch of cell extract.

For a pilot study, a small array of 46 kinase inhibitors was assembled (with multiple DMSO controls) and parallel *ex vivo* kinase reactions were performed on the H3 1-21 peptide in the presence of each inhibitor. Phosphorylation levels of H3T3ph and H3T11ph were determined by ELISA as described earlier. Notably as H3T3 and H3T11 are both within a short region of H3 the same peptide substrate was used in each case. 10 µM was chosen as the inhibitor concentration as this appeared the minimum required to achieve significant inhibition with all 3 inhibitors tested thus far.

A unique and reproducible pattern of inhibition was produced for each of these phosphorylation events (Figure 4.3B). For H3T11ph there were 4 compounds which were clearly able to inhibit the phosphorylation: Staurosporine, SU1438, CHR6494 and NSC-95397. H3T3 phosphorylation was inhibited by a much broader range of the compounds tested. Interestingly the majority of the compounds which inhibited H3T3ph are known Haspin inhibitors. Harmine, harmol and relative LDN211898 are known Haspin inhibitors, as is LDN192960 (Cuny et al. 2012) (Cuny et al. 2010).

RO-3306 is a potent CDK1 inhibitor but also has strong inhibitory activity against Haspin (Patnaik and Higgins, unpublished; LINCS 2018). H89 has IC50 of 0.47 μM for Haspin, and Cdk1 inhibitor III has 0.59 μM IC50 (Patnaik and Higgins, unpublished).

Thus the results of our test array were consistent with the notion that the same kinase acting on H3T3 *in vivo* (Haspin) was also responsible for the phosphorylation occurring in our *ex vivo* assay.

A

DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO
DMSO	ATMi Tocris	IKK (BMS-345541)	DMSO	MK2 (PF-3644022)	CDC7 (XL413)	CHK1 (CCT 244747)	DNA-PK (NU 7441)	p38 (Merck S13203580)	ATR (VE821)	MAPK (V0126)	DMSO
DMSO	Wee1 (MK1775)	IKK-Z inhib IV (TPCA)	DMSO	Roscovotine	SU 1438	MPS1	NSC-95397	SB 202474	GSK 461364	MLN 8237	DMSO
DMSO	MK 5108	H89	DMSO	CDK1 inhib III	Staurosporine	CHR 6494	Alsterpaullone	Indurubin 3'-Monoxin	DMAP	TBB	DMSO
DMSO	VX 680	Harmol	DMSO	Harmine	LDN 192960	LDN 2118298	HA 1077	ZM 447439	LDN 192960	BI-2536	DMSO
DMSO	Hesperadin	CHK2 (cct241533)	DMSO	RO-3306	iodotubercidin	SU 1411	SU 1433	SU 1266	SU 1261	DMSO	DMSO
DMSO	TLIO-163	ALW-II-49-7	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO
DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO	DMSO

B

0.471	0.391	0.41	0.424	0.401	0.383	0.432	0.382	0.371	0.424	0.426	0.507
0.392	0.397	0.428	0.43	0.388	0.377	0.481	0.444	0.439	0.412	0.396	0.422
0.443	0.451	0.42	0.417	0.422	0.199	0.455	0.163	0.427	0.61	0.425	0.443
0.413	0.341	0.491	0.416	0.434	0.141	0.129	0.396	0.398	0.437	0.407	0.414
0.412	0.481	0.388	0.398	0.439	0.342	0.406	0.384	0.525	0.39	0.613	0.48
0.431	0.587	0.447	0.374	0.338	0.336	0.4	0.454	0.397	0.442	0.468	0.416
0.377	0.383	0.394	0.407	0.391	0.354	0.394	0.404	0.386	0.385	0.41	0.413
0.385	0.327	0.407	0.359	0.327	0.326	0.337	0.305	0.333	0.301	0.324	0.378

0.357	0.415	0.358	0.375	0.394	0.35	0.361	0.457	0.454	0.472	0.486	0.535
0.361	0.452	0.521	0.427	0.481	0.423	0.511	0.499	0.547	0.539	0.598	0.561
0.357	0.583	0.465	1.197	0.72	0.12	0.546	0.059	0.533	0.651	0.553	0.578
0.393	0.452	0.666	0.529	0.549	0.251	0.09	0.572	0.529	0.583	0.519	0.65
0.439	0.474	0.519	0.51	0.624	0.565	0.531	0.561	0.638	0.543	0.767	0.603
0.343	0.555	0.611	0.684	0.493	0.568	0.442	0.507	0.371	0.637	0.566	0.532
0.323	0.529	0.355	0.521	0.541	0.501	0.456	0.5	0.578	0.645	0.612	0.497
0.436	0.382	0.418	0.46	0.44	0.38	0.383	0.408	0.461	0.45	0.495	0.515

0.275	0.293	0.312	0.251	0.295	0.306	0.367	0.366	0.348	0.386	0.434	0.395
0.256	0.324	0.285	0.325	0.389	0.276	0.335	0.401	0.335	0.285	0.383	0.438
0.252	0.293	0.259	0.361	0.321	0.078	0.399	0.333	0.279	0.435	0.455	0.509
0.279	0.103	0.106	0.288	0.099	0.089	0.083	0.363	0.302	0.4	0.199	0.476
0.323	0.294	0.125	0.306	0.109	0.07	0.076	0.264	0.421	0.09	0.592	0.511
0.274	0.161	0.282	0.327	0.092	0.085	0.092	0.132	0.111	0.112	0.492	0.457
0.357	0.474	0.32	0.455	0.424	0.401	0.336	0.453	0.435	0.483	0.462	0.44
0.383	0.365	0.389	0.358	0.352	0.328	0.326	0.36	0.394	0.412	0.428	0.442

0.302	0.318	0.292	0.292	0.283	0.306	0.291	0.347	0.324	0.363	0.358	0.382
0.29	0.333	0.328	0.3	0.359	0.299	0.284	0.395	0.334	0.314	0.363	0.365
0.435	0.34	0.468	0.519	0.385	0.098	0.412	0.291	0.25	0.358	0.356	0.389
0.275	0.131	0.191	0.531	0.147	0.125	0.075	0.333	0.243	0.327	0.181	0.428
0.333	0.373	0.21	0.354	0.124	0.169	0.101	0.219	0.354	0.097	0.426	0.409
0.319	0.152	0.333	0.411	0.112	0.11	0.117	0.117	0.139	0.102	0.383	0.386
0.326	0.325	0.267	0.342	0.341	0.306	0.347	0.341	0.344	0.345	0.358	0.343
0.354	0.291	0.306	0.303	0.32	0.294	0.321	0.306	0.324	0.324	0.339	0.373

C

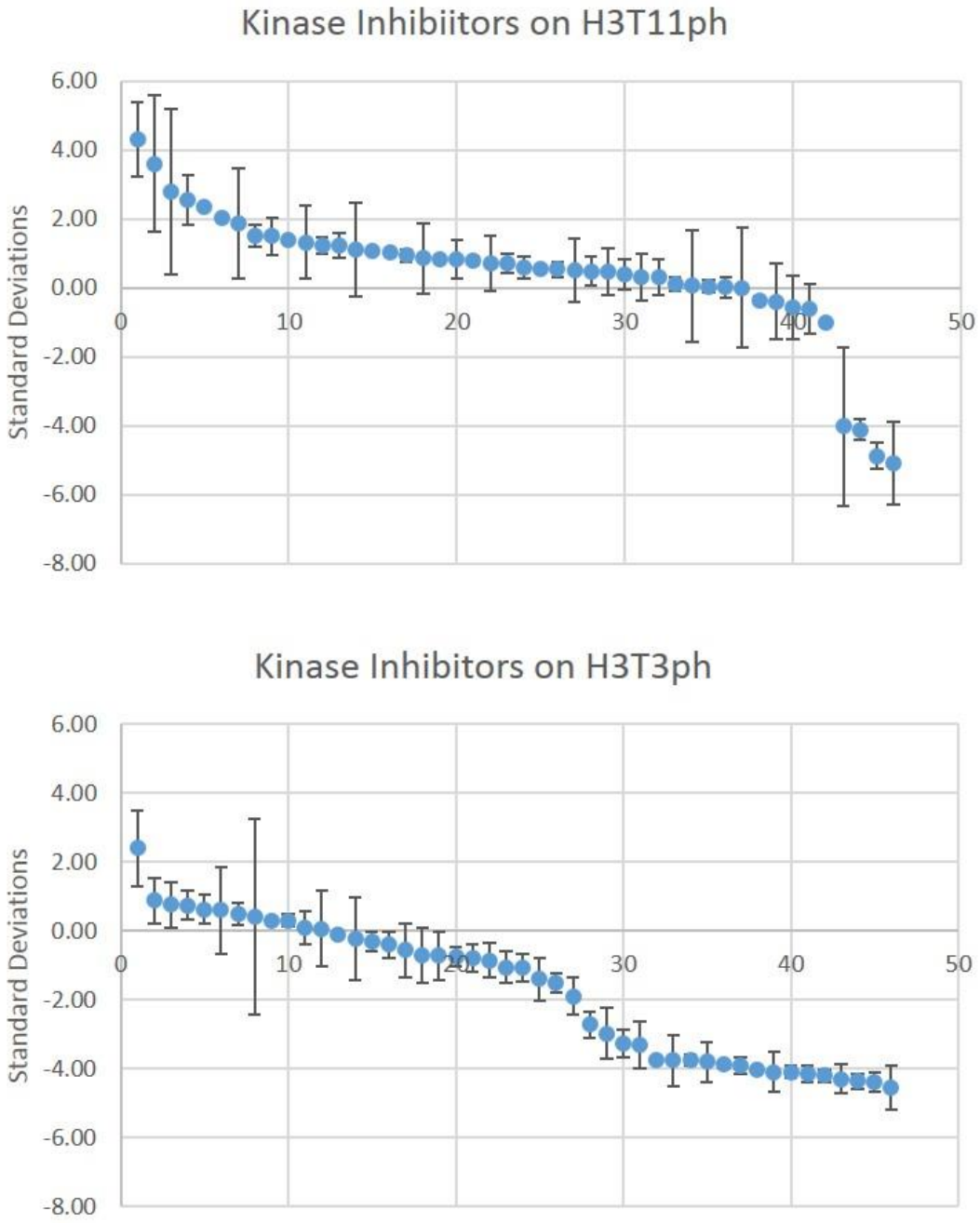


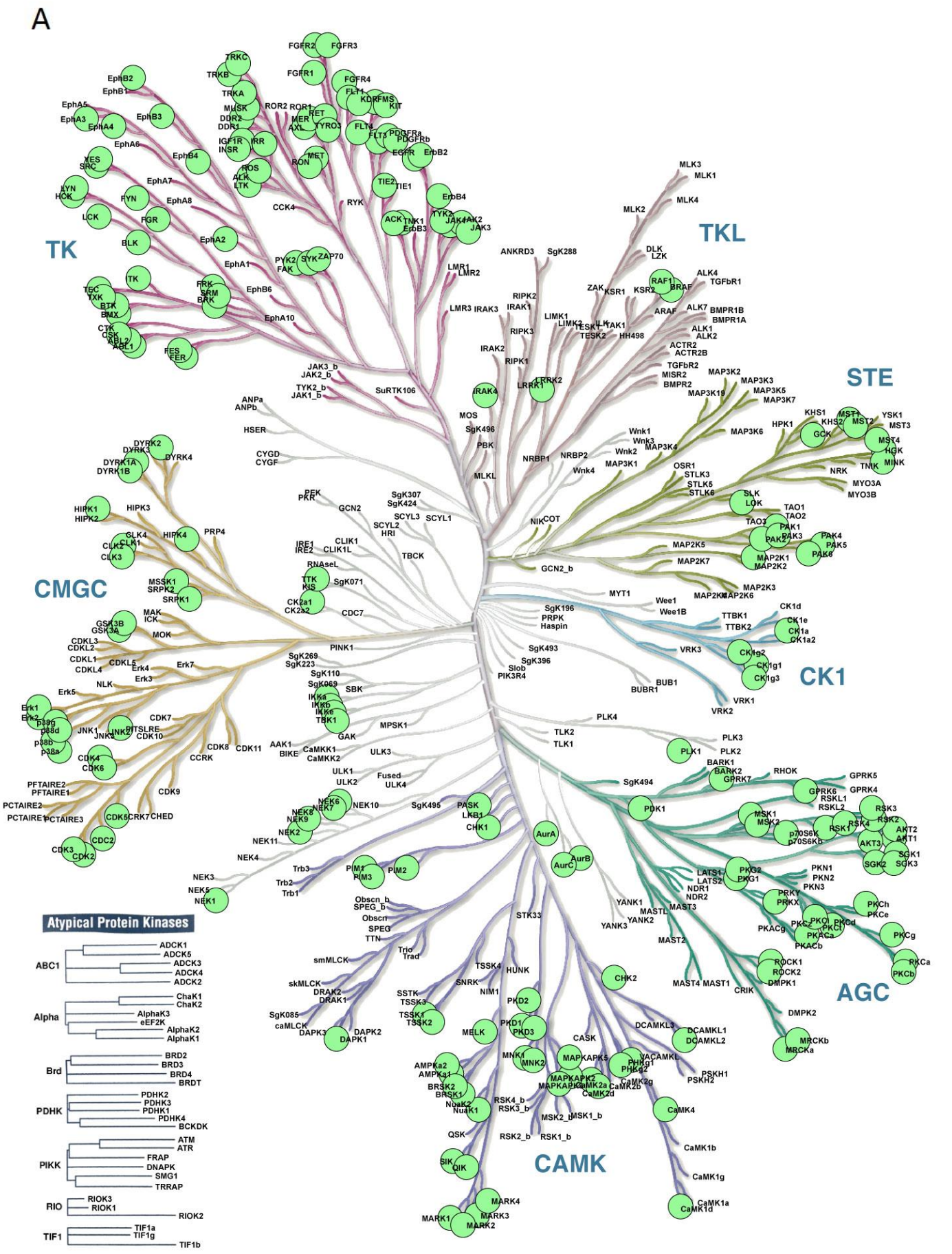
Figure 4.3: **Inhibitor test array of H3T3ph and H3T11ph inhibition**: **A** Layout of test array **B** Inhibitors (final conc. 10 µM , arrayed as in **A**) were added to H3 1-21 peptides in kinase buffer (containing phosphatase and protease inhibitors). 5% mitotic extract was added and plates incubated for 30 mins at 30°C. H3 1-21 peptides were then immobilised on streptavidin plates (via C-terminal biotin tag) and probed with antibodies for H3T3ph or H3T11ph. Signal was quantified by ELISA. Signal for H3T11ph probed plates (upper 2), H3T3ph probed plates (lower 2). **C** standard deviations of inhibitors from mean DMSO signal (data from B)

4.2.4 Correlation analysis between ex-vivo kinase reactions and large kinase inhibitor profiling datasets can identify known in vivo kinases

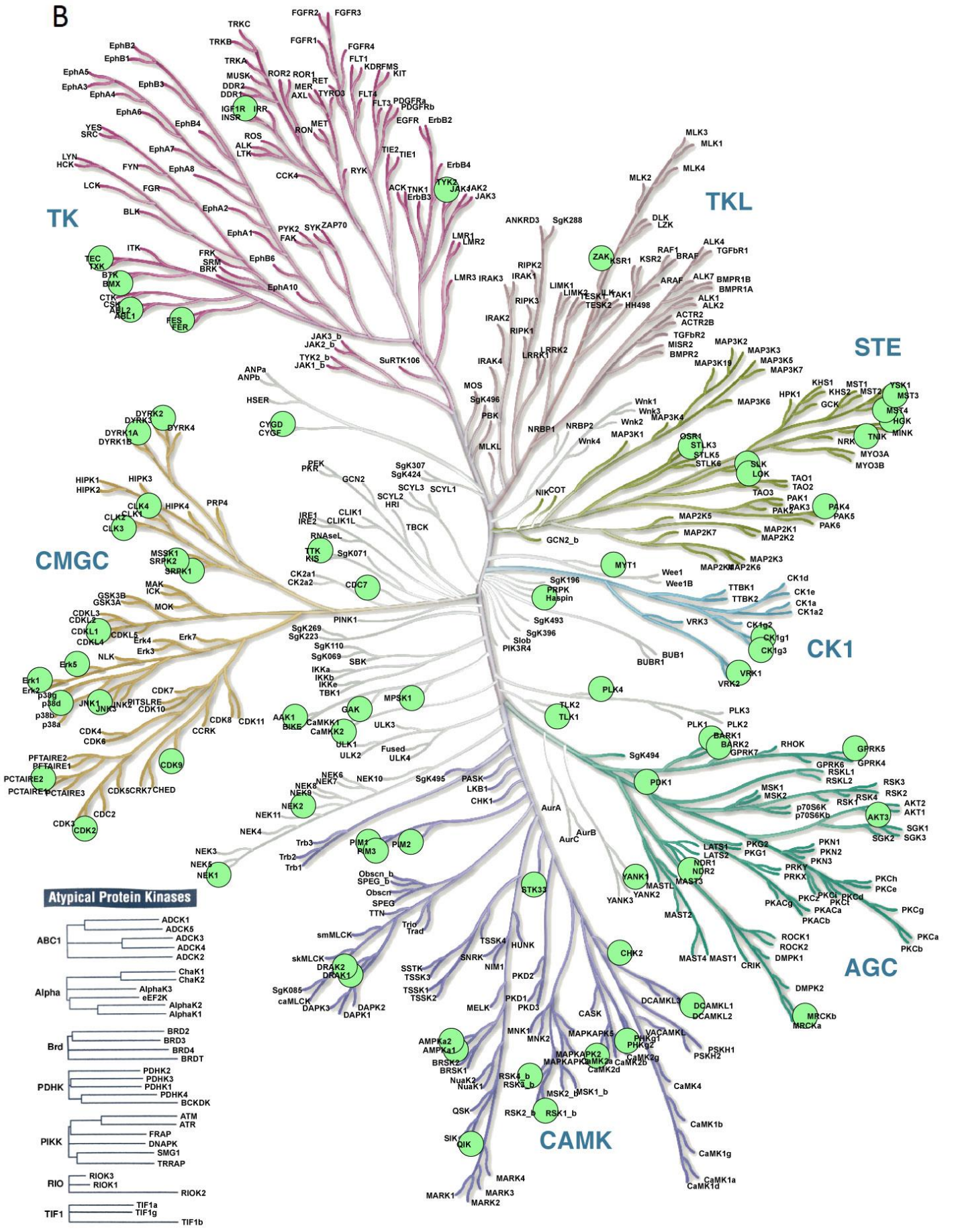
4.2.5 Published Kinase inhibitor set (PKIS 1)

The results of our test array indicated that our assay was successfully highlighting compounds that inhibit Haspin, and that Haspin was the kinase phosphorylating H3T3 on our H3T3 1-21 peptide in our ex-vivo assay. We therefore reasoned that, using a large library of inhibitors which had been profiled on multiple kinases (including Haspin), it should be possible to pick Haspin out 'blind' as the kinase phosphorylating H3T3.

To test this possibility, we acquired PKIS 1, a set of small molecule kinase inhibitors assembled by the Structural Genomics Consortium from published GlaxoSmithKline compounds (Elkins et al., 2016). We received 317 compounds which had been profiled *in vitro* on 2 panels of recombinant kinases, the Nanosyn enzyme assay panel and a 'DSF' (differential scanning fluorimetry) panel (the Nanosyn assay assesses inhibition of enzymatic activity by in vitro kinase assays; whereas the DSF assay is a binding assay based on thermal denaturation). The Nanosyn enzyme panel contains 196 unique members of the human protein kinome; these were each assayed with all 317 inhibitors at 2 concentrations of 0.1 µM and 1 µM. The DSF panel (assayed at 1 µM) contains 68 kinases (32 of which are in common with the Nanosyn panel), giving a total coverage of 232 kinases (Fig 4.4).



"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"



"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"

Figure 4.4: **Kinome coverage of PKIS1 inhibition datasets : A** Kinases profiled on the Nanosyn assay with PKIS1 compounds **B** Kinases profiled on the DSF assay with PKIS1 compounds.

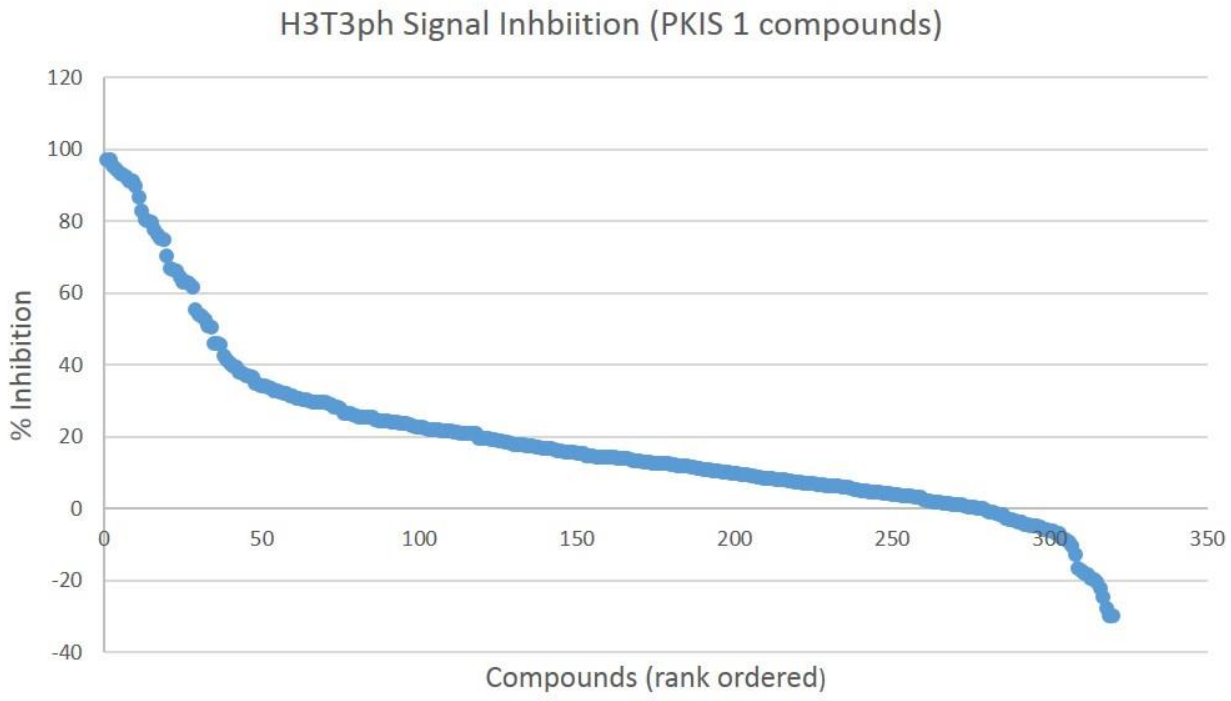
4.2.6 KiPIK screening identifies Haspin as H3T3ph kinase

We performed parallel *ex vivo* kinase reactions on our H3 1-21 peptide in the presence of the 317 small molecule kinase inhibitors of PKIS 1. For each compound we calculated a % inhibition where 100% inhibition was defined by the signal measured in the wells of the most inhibitory compound (or EDTA), and 0% determined by the signal measured in control wells where DMSO was added rather than a PKIS 1 compound. Calculated this way, 34 compounds inhibited our H3T3ph signal >50% while the majority had little effect. Based on the results of our test array we reasoned that the PKIS 1 compounds giving high inhibition scores were likely to be Haspin inhibitors; whilst those causing no reduction in H3T3ph signal in our assay were likely to be compounds that did not inhibit Haspin. If this were true we reasoned that the pattern of inhibition produced by our *ex vivo* screening of the PKIS 1 compounds ought to be more similar to their *in vitro* inhibition profile for Haspin than for the other kinases they were profiled on.

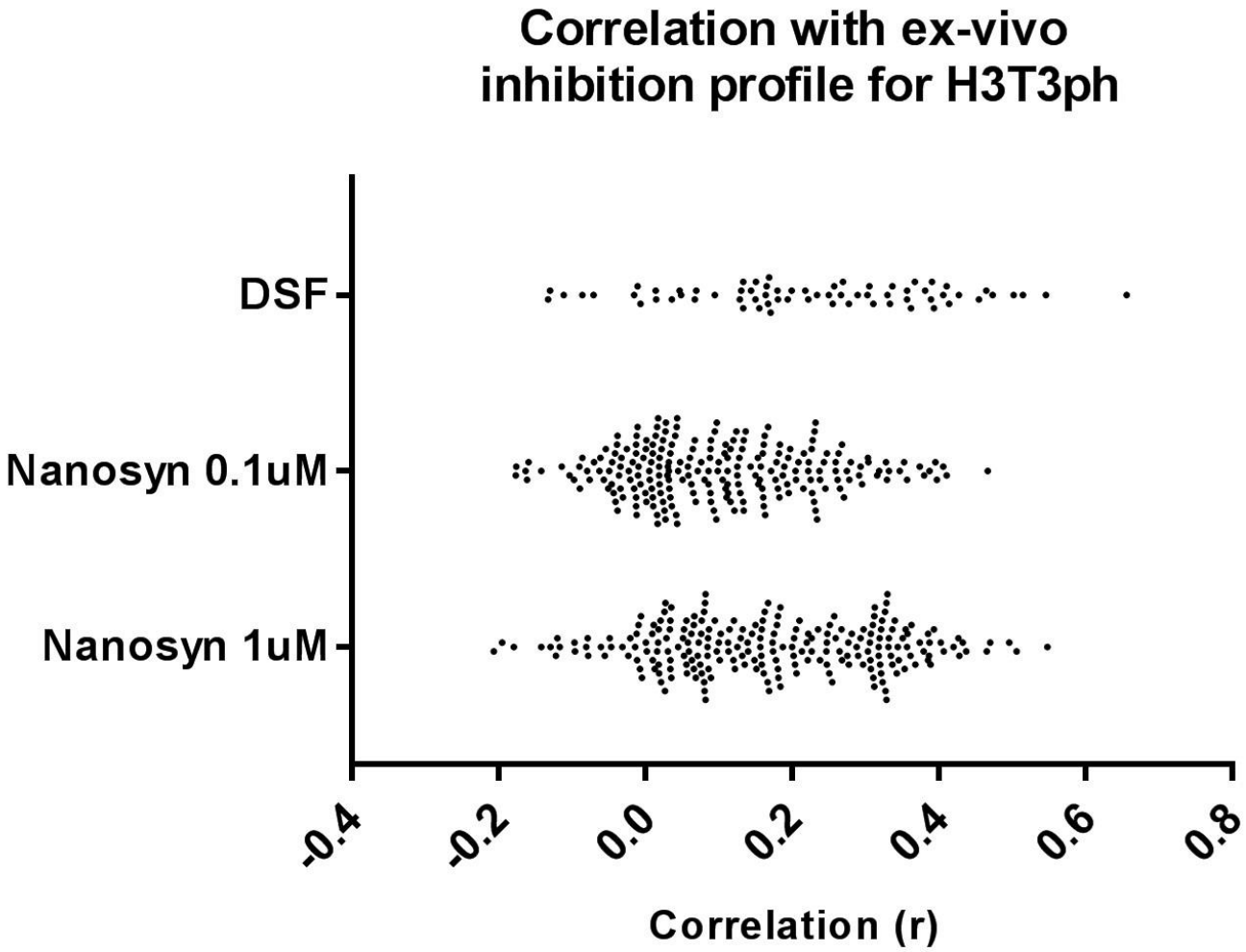
To test this, we performed a Pearson's test of correlation between our *ex vivo* generated inhibition scores and the inhibition profiles the same compounds produced for each kinase they were tested on *in vitro*.

The highest correlation calculated for our experimentally generated inhibition dataset across the kinase inhibition profiles for both recombinant kinase panels was for Haspin (Figure 4.5B). Of note, the Nanosyn panel did not contain Haspin. Haspin is well established as the H3T3ph kinase *in vivo* (Dai et al., 2005, Eswaran et al., 2009)

A


H3T3ph Signal Inhbiition (PKIS 1 compounds)
% Inhibition
120
100
80
60
40
20
0
-20
-40
0
50
100
150
200
250
300
350
Compounds (rank ordered)

B



Nanonsyn 1uM		Nanosyn 0.1uM		DSF	
DYRK1A	0.5475	DYRK1A	0.4664	HASPIN	0.6557
DYRK1B	0.5055	CDK5/p35	0.41	PIM3	0.5455
PIM3	0.4954	DYRK1B	0.4057	CLK4	0.5144
PIM1	0.4696	CDK2/cyclinE	0.3968	PIM1	0.5014
MELK	0.465	CDK1/cyclinB	0.3947	PIM2	0.4731
PIM2	0.4369	CDK2/cyclinA	0.3871	STK17B	0.464
CLK2	0.4322	CDK3/cyclinE	0.3796	STK38L	0.4541
DYRK2	0.4279	CDK4/cyclin D	0.3736	AAK1	0.4272
LRRK2-G2019S	0.4266	ARK5	0.3563	CLK3	0.4135
IKKA	0.4196	CLK2	0.3521	CAMK2A	0.4106
PI4-K-b	0.4062	LRRK2-G2019S	0.3464	PRKAA2	0.4056
CDK2/cyclinE	0.4024	HIPK1	0.3318	TYK2	0.393
CDK1/cyclinB	0.3928	MAP4K4	0.3288	DYRK1A	0.39
CDK5/p35	0.3885	MSK2	0.3278	CDK2	0.3887
PRKG2	0.3868	MELK	0.3183	CDK9	0.3809
CDK4/cyclinD	0.3861	PIM1	0.3155	STK17A	0.3668
CLK3	0.385	CDK6/cyclin D3	0.3037	PHKG2	0.3614
CDK3/cyclinE	0.3847	MSK1	0.3031	DYRK2	0.3566
MSK2	0.3791	PRKG1	0.2949	CDC7	0.3564
HIPK1	0.3734	RSK1	0.2929	PRKAA1	0.3365

C

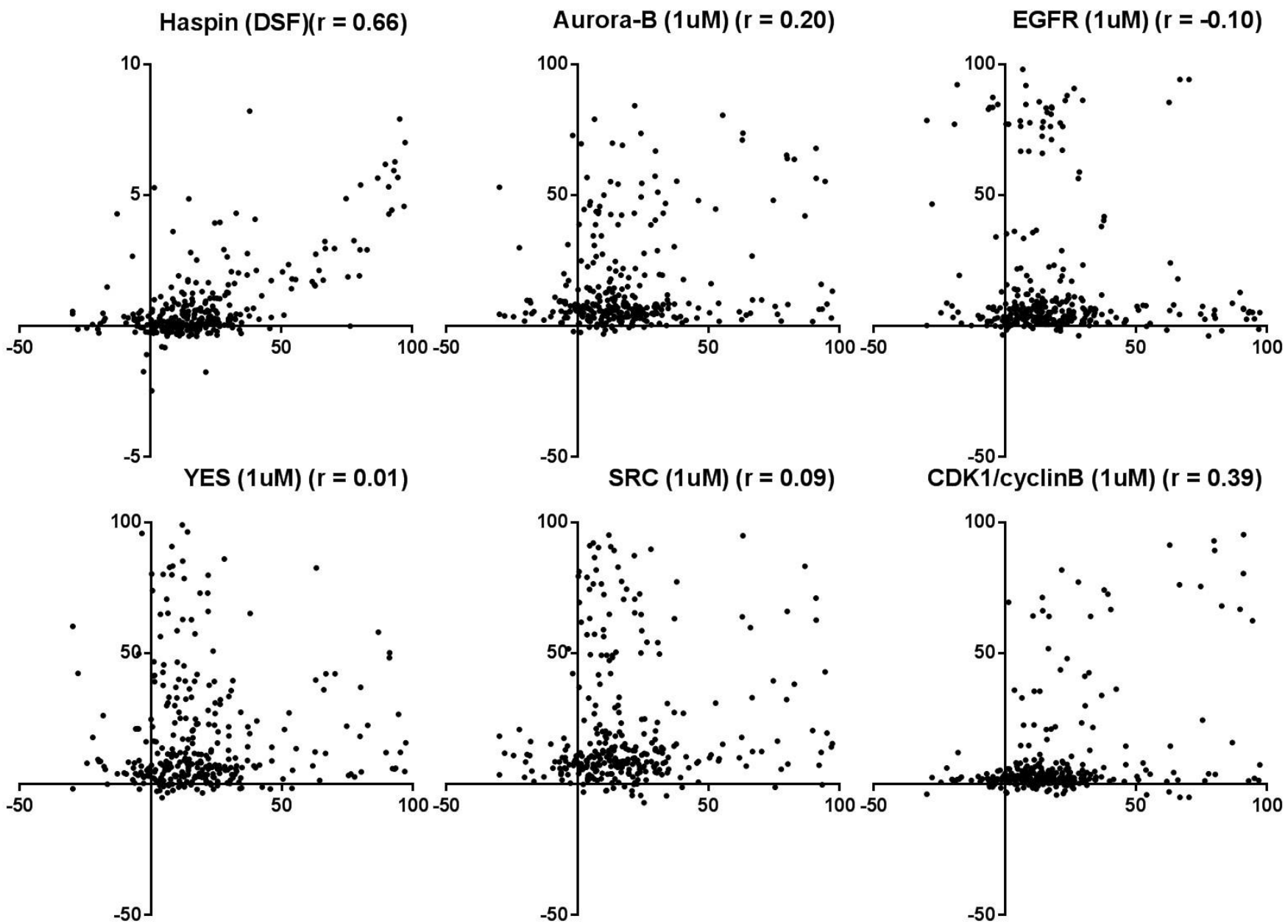

Haspin (DSF)(r = 0.66)
Aurora-B (1uM) (r = 0.20)
EGFR (1uM) (r = -0.10)
YES (1uM) (r = 0.01)
SRC (1uM) (r = 0.09)
CDK1/cyclinB (1uM) (r = 0.39)
10
5
-5
100
50
-50
-50
50
100

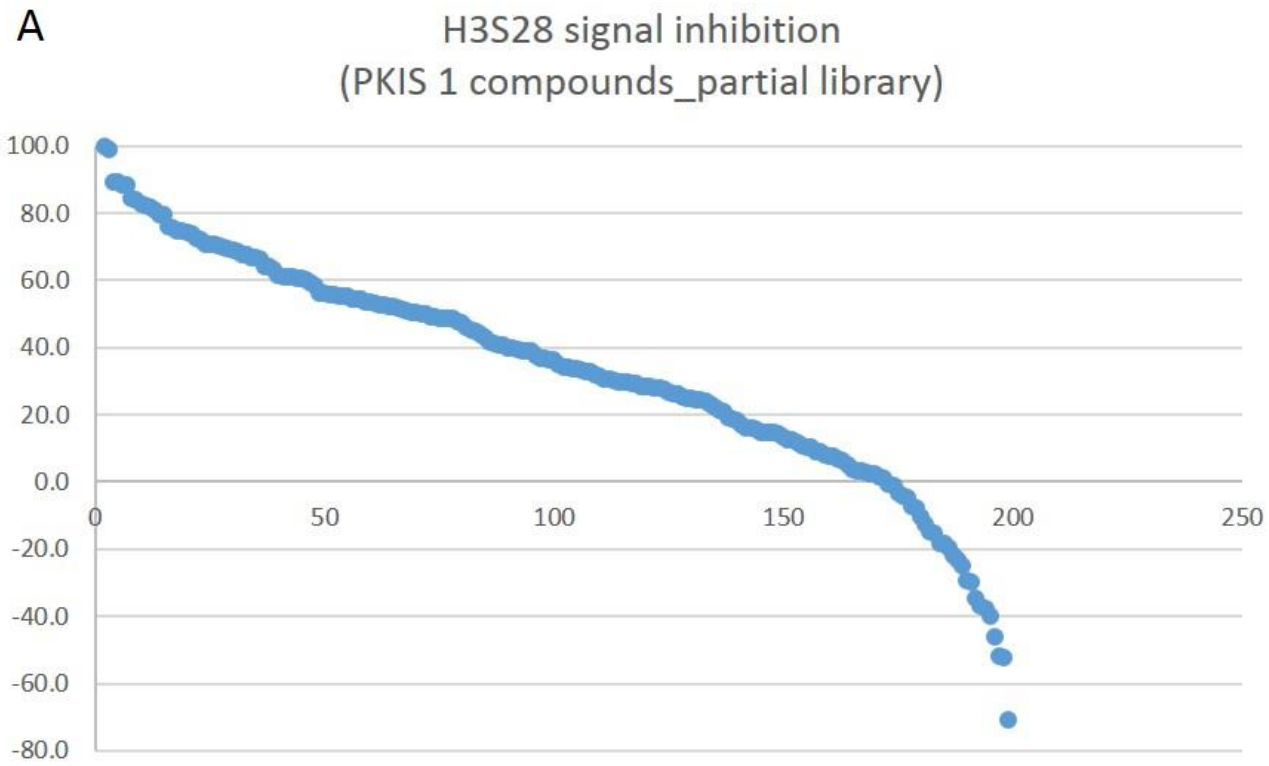
Figure 4.5: **H3T3ph KiPIK screen with PKIS1 compounds** : **A** % inhibition of H3T3ph signal by PKIS1 compounds **B** Pearson correlation of recombinant kinase *in vitro* inhibition profiles with H3T3ph lysate signal inhibition profile (PKIS1 compounds), (right) top 20 correlating kinases from 3 datasets of inhibition profiles **C** Scatter plots of PKIS1 compounds. (x-axis) H3T3ph signal inhibition in lysate reactions, against (y-axis) inhibition of recombinant kinase *in vitro* (from inhibitor profiling dataset: Nanosyn 1 µM, except Haspin from DSF 1 µM dataset)

4.2.7 KiPIK screening identifies Aurora (B) as H3S28ph kinase

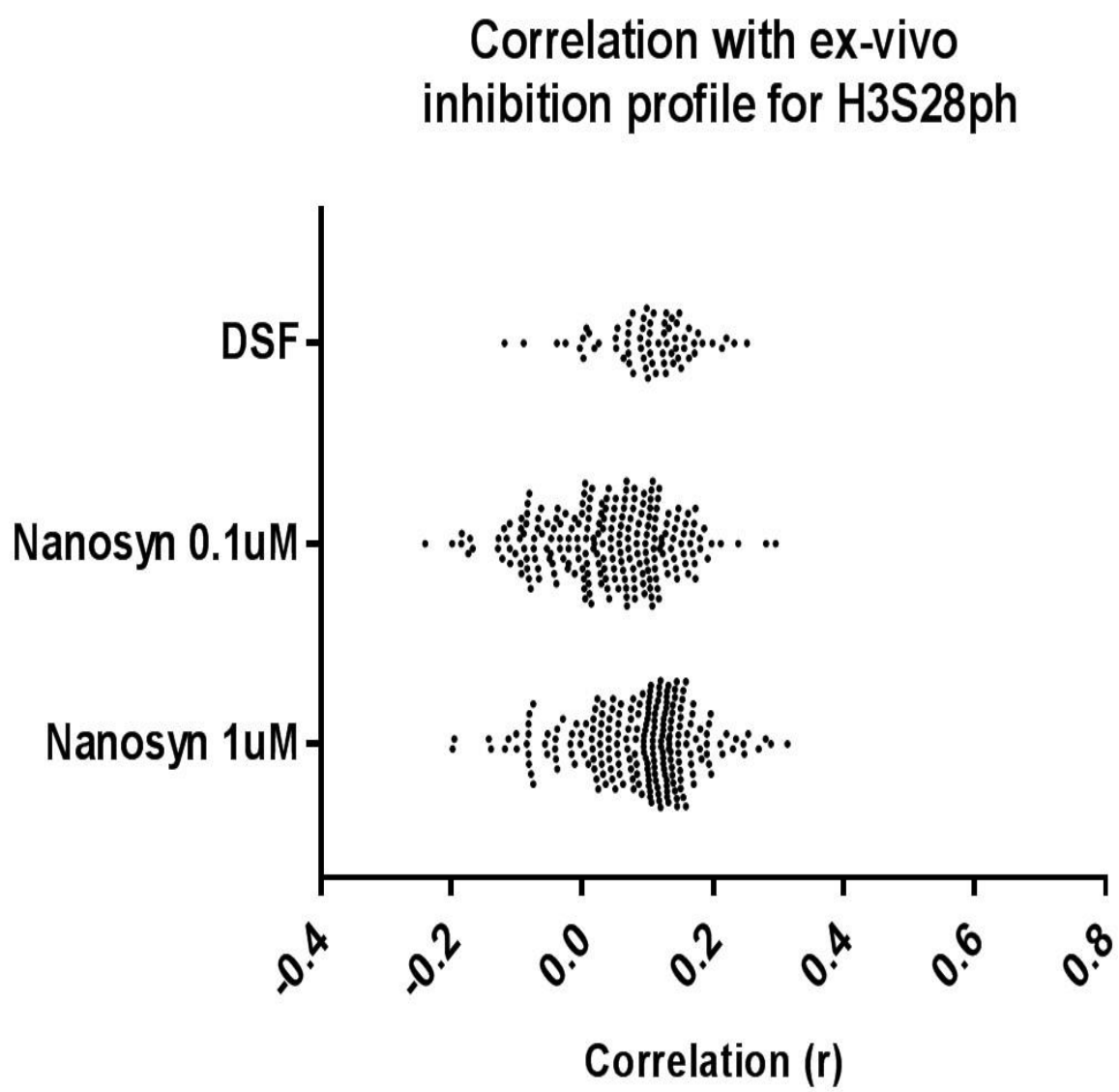
We performed the same screening technique for H3S28ph. Unfortunately, a systematic pipetting error occurred during the ELISA stage of this assay, the compromised wells had to be discarded from the results (118 inhibitors) leaving us with data for 199 of the 317 inhibitors in our PKIS1 panel. Despite this reduced dataset, Pearson's correlation clearly identified the Aurora kinases (from the Nanosyn panels) as most likely responsible for the phosphorylation signal (Figure 4.6). Note that the DSF panel does not contain Aurora kinases. Aurora B is well established as the H3S28 kinase in mitosis (Crosio et al., 2002, Goto et al., 2002).

A repeat screen was performed (this time with all 317 inhibitors). As before the Aurora kinases were clearly identified as most likely responsible for the H3S28 phosphorylation signal (Figure 4.6).

A

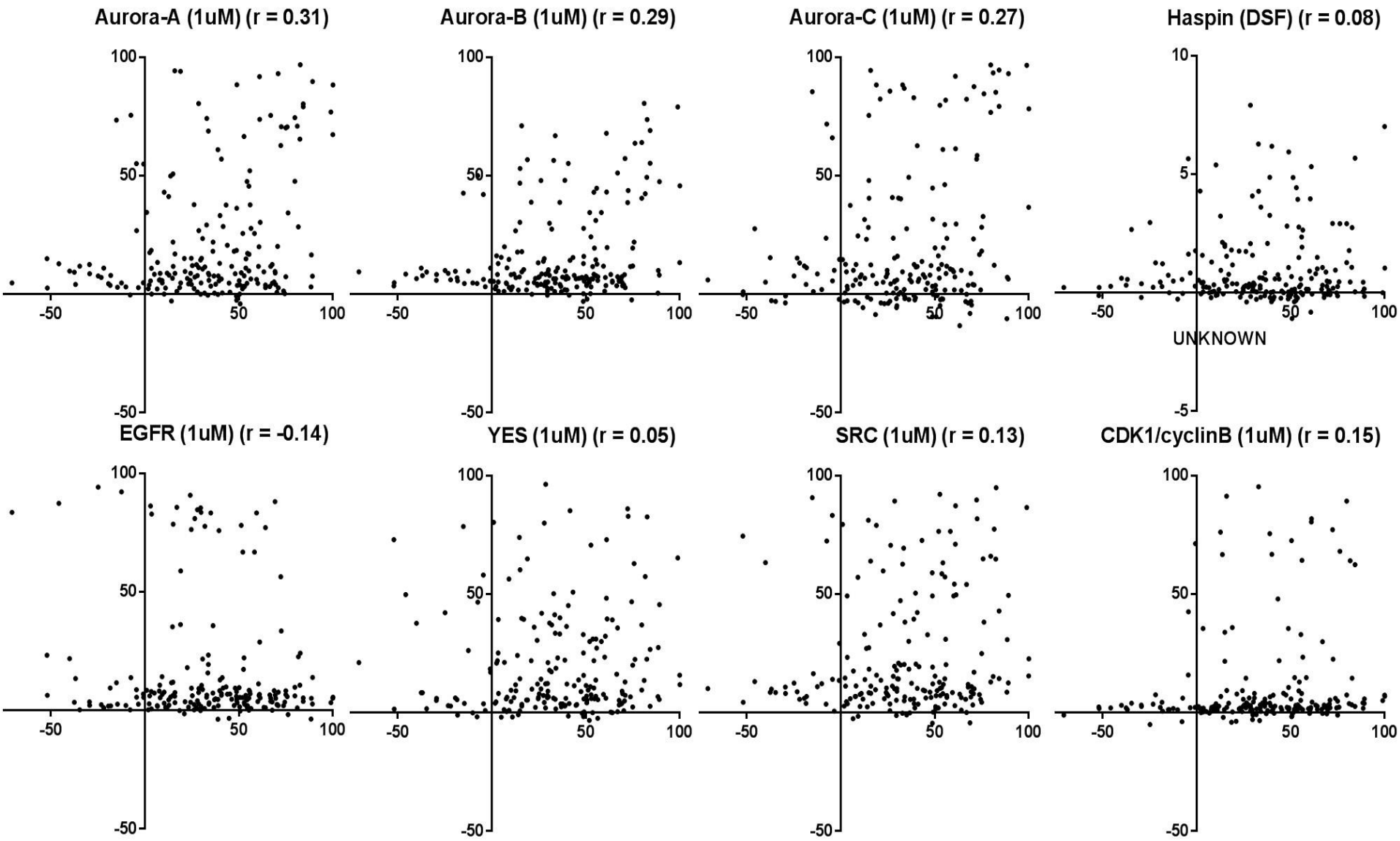

H3S28 signal inhibition
(PKIS 1 compounds_partial library)
100.0
80.0
60.0
40.0
20.0
0.0
-20.0
-40.0
-60.0
-80.0
0
50
100
150
200
250

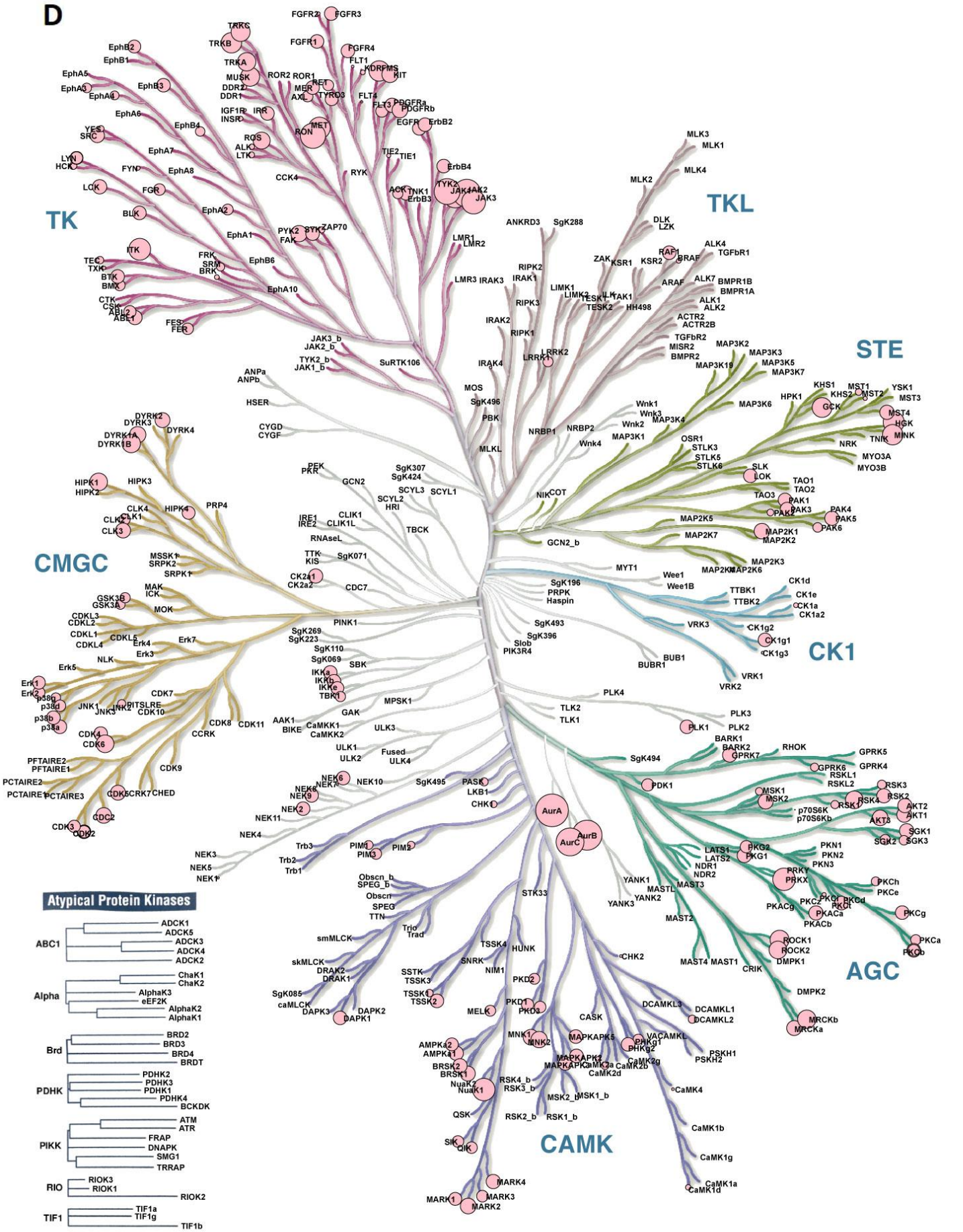
B

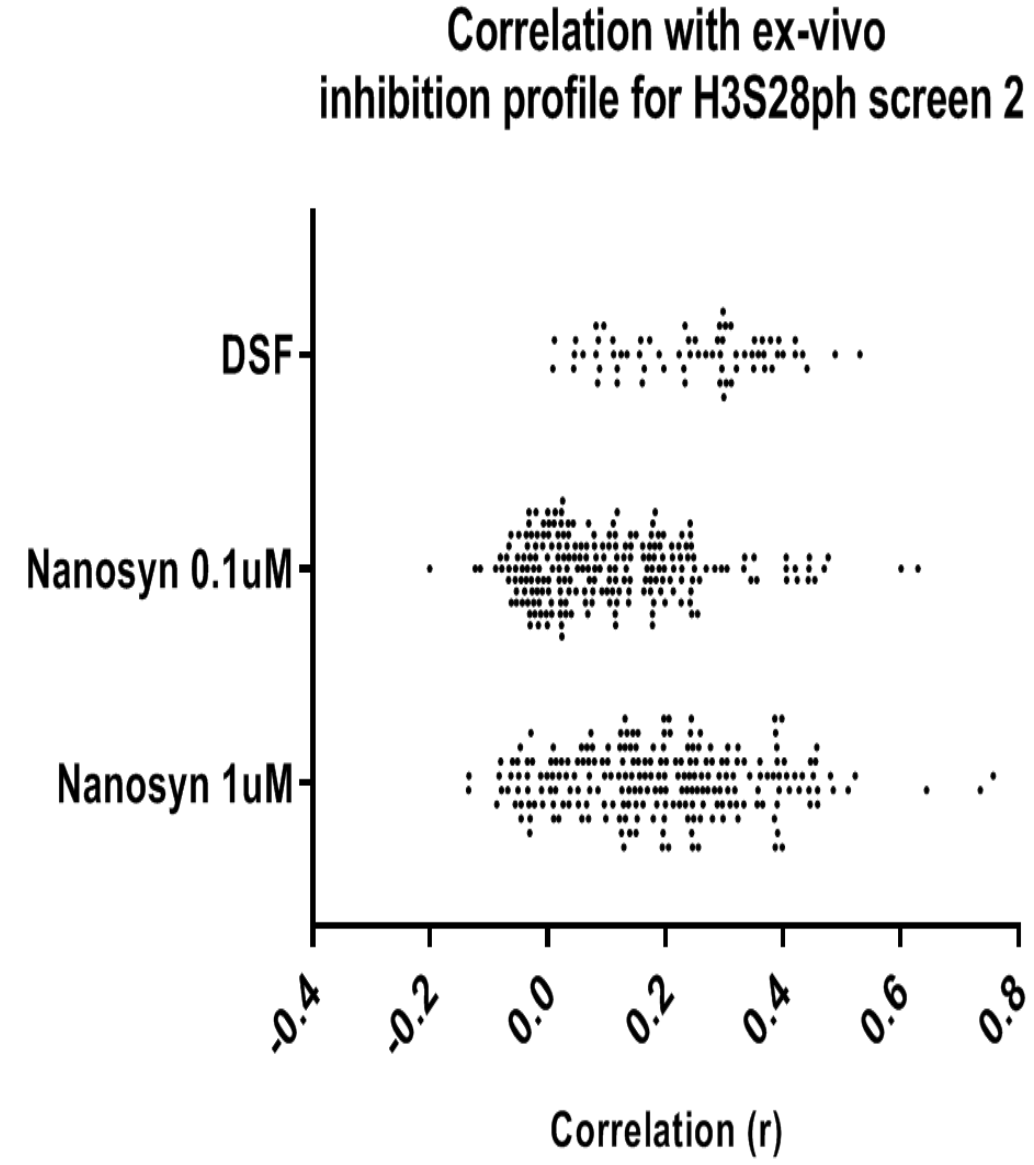


Nanosyn 1uM		Nanosyn 0.1uM		DSF	
AURORA-A	0.314	AURORA-C	0.296	STK38L	0.252
AURORA-B	0.288	AURORA-B	0.281	AAK1	0.233
JAK2	0.280	AURORA-A	0.238	PRKAA2	0.221
AURORA-C	0.270	JAK3	0.212	TYK2	0.213
TYK2	0.254	PRKX	0.198	ADRBK1	0.199
RON	0.248	RON	0.192	CDK2	0.184
JAK1	0.247	JAK2	0.186	STK33	0.177
JAK3	0.235	CDK6/cyclinD3	0.180	MAPK3	0.172
PRKX	0.230	FMS	0.180	GAK	0.171
ITK	0.220	MNK2	0.173	MAPK8	0.163
ARK5	0.214	TRKC	0.173	TNIK	0.163
MET	0.211	ARK5	0.172	ABL2	0.157
FMS	0.197	TRKB	0.171	PHKG2	0.156
MAP4K4	0.196	MUSK	0.169	MST4	0.151
MAP4K2	0.194	MKNK1	0.168	STK17B	0.148
ROCK1	0.194	MRCK-b	0.165	TTK	0.146
TRKB	0.191	ROCK1	0.162	PAK4	0.144
MRCK-b	0.191	ROCK2	0.162	CLK3	0.143
MINK	0.190	ABL-T315I	0.158	PCTK2	0.142
ROCK2	0.182	TRKA	0.157	MAPK7	0.139

C


Aurora-A (1uM) (r = 0.31)
Aurora-B (1uM) (r = 0.29)
Aurora-C (1uM) (r = 0.27)
Haspin (DSF) (r = 0.08)
UNKNOWN
EGFR (1uM) (r = -0.14)
YES (1uM) (r = 0.05)
SRC (1uM) (r = 0.13)
CDK1/cyclinB (1uM) (r = 0.15)


D
TK
TKL
STE
CMGC
CK1
AGC
CAMK
Atypical Protein Kinases
ABC1
Alpha
Brd
PDHK
PIKK
RIO
TIF1
"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"



Nanosyn 1uM		Nanosyn 0.1 uM		DSF	
AURORA-B	0.757	AURORA-C	0.629	PRKAA2	0.530
AURORA-C	0.735	AURORA-B	0.601	PLK4	0.488
AURORA-A	0.644	AURORA-A	0.476	STK16	0.440
MAP4K2	0.522	FMS	0.469	BMX	0.433
FMS	0.511	KIT	0.454	STK38L	0.420
KIT-T6701	0.485	PDGFRα-V561D	0.445	AAK1	0.420
FLT3	0.480	MAP4K2	0.444	PRKAA1	0.400
JAK2	0.458	PDGFRβ	0.442	FES	0.392
KIT-V560G	0.458	FLT3	0.424	TNIK	0.390
PDGFRα-V561D	0.457	KIT-T6701	0.420	CAMKK2	0.381
PDGFRα-T674I	0.455	PDGFRα	0.406	STK33	0.379
PDGFRβ	0.451	KIT-V560G	0.405	ABL2	0.368
KIT	0.447	PDGFRα-T674I	0.405	PDPK1	0.367
ARK5	0.446	TRKC	0.355	STK17B	0.364
PDGFRα	0.443	TRKB	0.350	MST4	0.358
TRKB	0.432	ARK5	0.345	CDK2	0.351
TRKC	0.427	RET-V804L	0.335	NEK2	0.349
RET-V804L	0.419	RET-Y791F	0.333	STK10	0.347
ROS	0.411	TNK1	0.304	STK24	0.334
FLT-3-D835Y	0.406	RET	0.295	CDK9	0.321

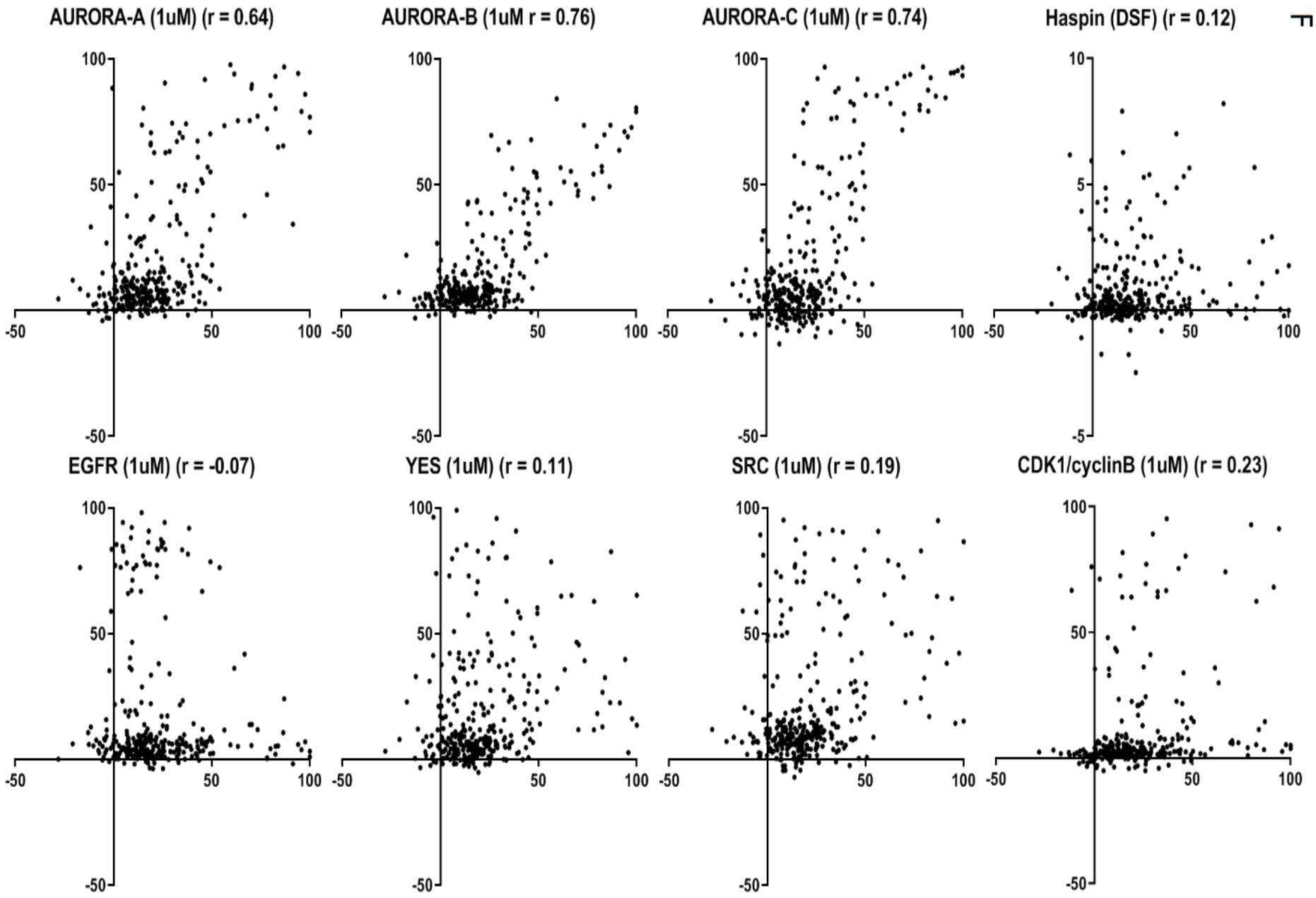

AURORA-A (1uM) (r = 0.64)
AURORA-B (1uM r = 0.76)
AURORA-C (1uM) (r = 0.74)
Haspin (DSF) (r = 0.12)
EGFR (1uM) (r = -0.07)
YES (1uM) (r = 0.11)
SRC (1uM) (r = 0.19)
CDK1/cyclinB (1uM) (r = 0.23)

Figure 4.6: H3S28ph KiPIK screen with PKIS1 compounds : **A** % inhibition of H3S28ph signal by PKIS1 compounds **B** Pearson correlation of recombinant kinase *in vitro* inhibition profiles with H3S28ph lysate signal inhibition profile (PKIS1 compounds), (right) top 20 correlating kinases from 3 datasets of inhibition profiles **C** Scatter plots of PKIS1 compounds. (x-axis) H3S28ph signal inhibition in lysate reactions, against (y-axis) inhibition of recombinant kinase *in vitro* (from inhibitor profiling dataset: Nanosyn 1 µM, except Haspin from DSF 1 µM dataset) **D** Pearson correlation coefficients (as in **B**), for kinases from Nanosyn 1 µM dataset, mapped onto phylogenetic tree of the human kinome. Circle size corresponds to correlation coefficient score. **E** as (B), repeat screen. **F** as (C) repeat screen

4.2.8 KiPIK screening works on phosphorylation sites regulated by diverse kinases and site-specific antibodies are not a requirement

It was important to test if KiPIK screening would work on phosphorylation sites regulated by different types of kinases. As both test cases thus far were histone sites phosphorylated by mitotic serine/threonine kinases, we decided to test KiPIK screening on phosphorylation sites regulated by tyrosine kinases. The next test cases we screened are reportedly regulated by a membrane bound receptor tyrosine kinase (autophosphorylation of EGFR at Y1016ph) and a cytosolic non-receptor tyrosine kinase (Src family kinase phosphorylation of Integrin β1A Y795).

Additionally, a limitation to our method was the need for a phospho-specific antibody against the phosphorylation site being screened. Producing such an antibody can be costly and time consuming so we wanted to assess if a generic phospho antibody could be used. A generic phospho-Tyrosine antibody (P-Tyr-100, Cell signalling) was used for both the EGFR Y1016 and Integrin β1A Y795 screens below.

4.2.9 KiPIK screening identifies EGFR as EGFR Y1016ph kinase

We acquired a biotinylated peptide corresponding to residues 1013-1021 of EGFR. This peptide contained a single Tyrosine corresponding to Y1016, which is reportedly an EGFR autophosphorylation site (Walton et al., 1990, Rotin et al., 1992). To generate a robust phosphorylation signal, we decided to use A431 cells

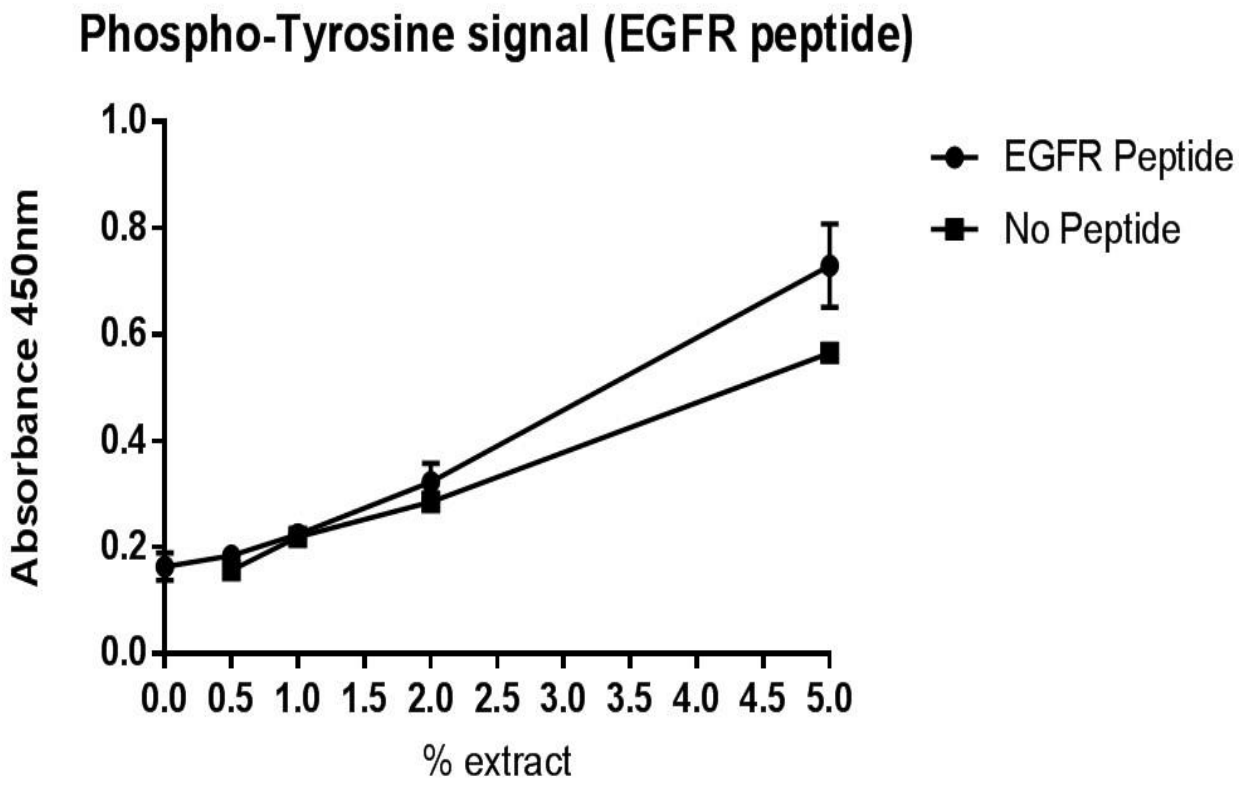
and to stimulate them with EGF for 5 minutes prior to harvesting the cells for cell extract. A431 cells express high levels of EGFR.

We performed a calibration experiment to determine the optimum concentration of A431 extract to use for our screening, and to ensure a clear phosphorylation signal could be detected. The results of this calibration indicated a robust tyrosine phosphorylation signal which increased with extract concentration. However, it was also clear that a large fraction of this signal was EGFR peptide substrate independent (Figure 4.7A). We were concerned that this background signal might distort the results but reasoned that if it composed a mixture of residues bound to the plate (due to imperfect extract wash-off), then our relatively dominant peptide phosphorylation signal may be resolvable.

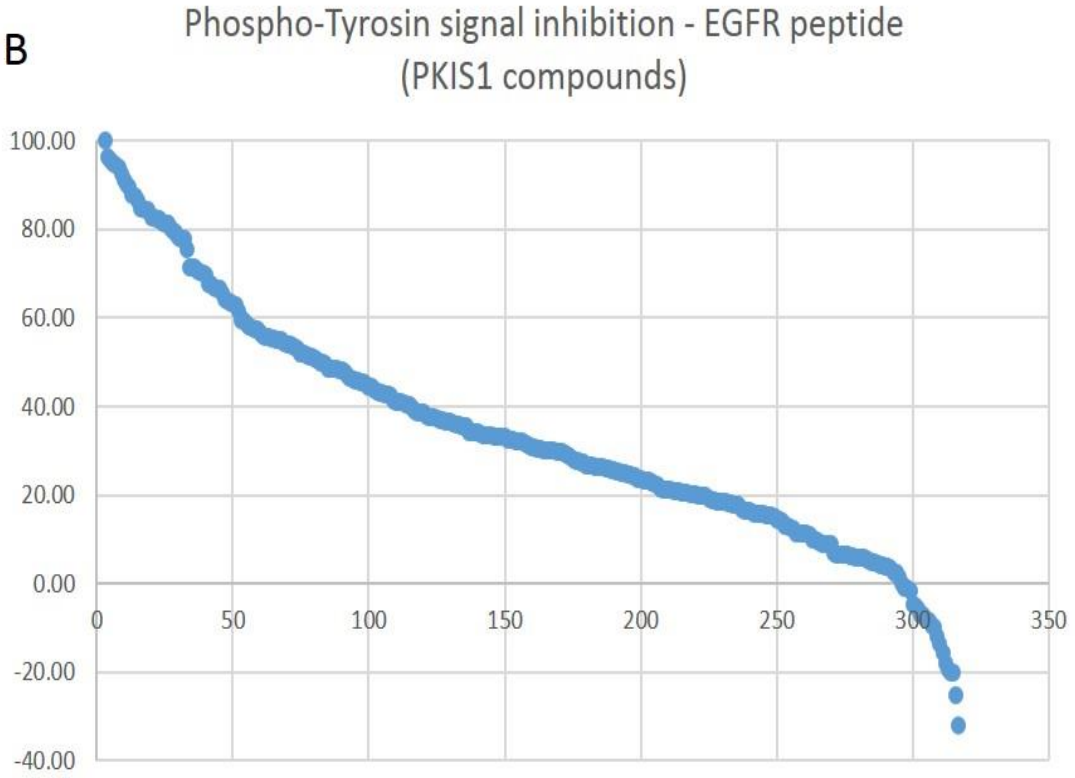
The results of the screen clearly identified EGFR as the most likely phosphorylating kinase (Figure 4.7C). ERBB2 and ERBB4 were the next most correlated hits. As these 2 kinases are the phylogenetically nearest of all kinases to EGFR, and their small molecule inhibition profiles are similar, it is probable that their high correlation scores are a reflection of this rather than their direct involvement in phosphorylating this EGFR autophosphorylation site.

The resolution from the rest or the kinome in terms of correlation scores was very good, adding to our confidence of this result. Additionally, strong correlation is clear from a visual inspection of the scatter plot of *ex vivo* inhibition scores vs *in vitro* inhibition scores (Figure 4.7D).

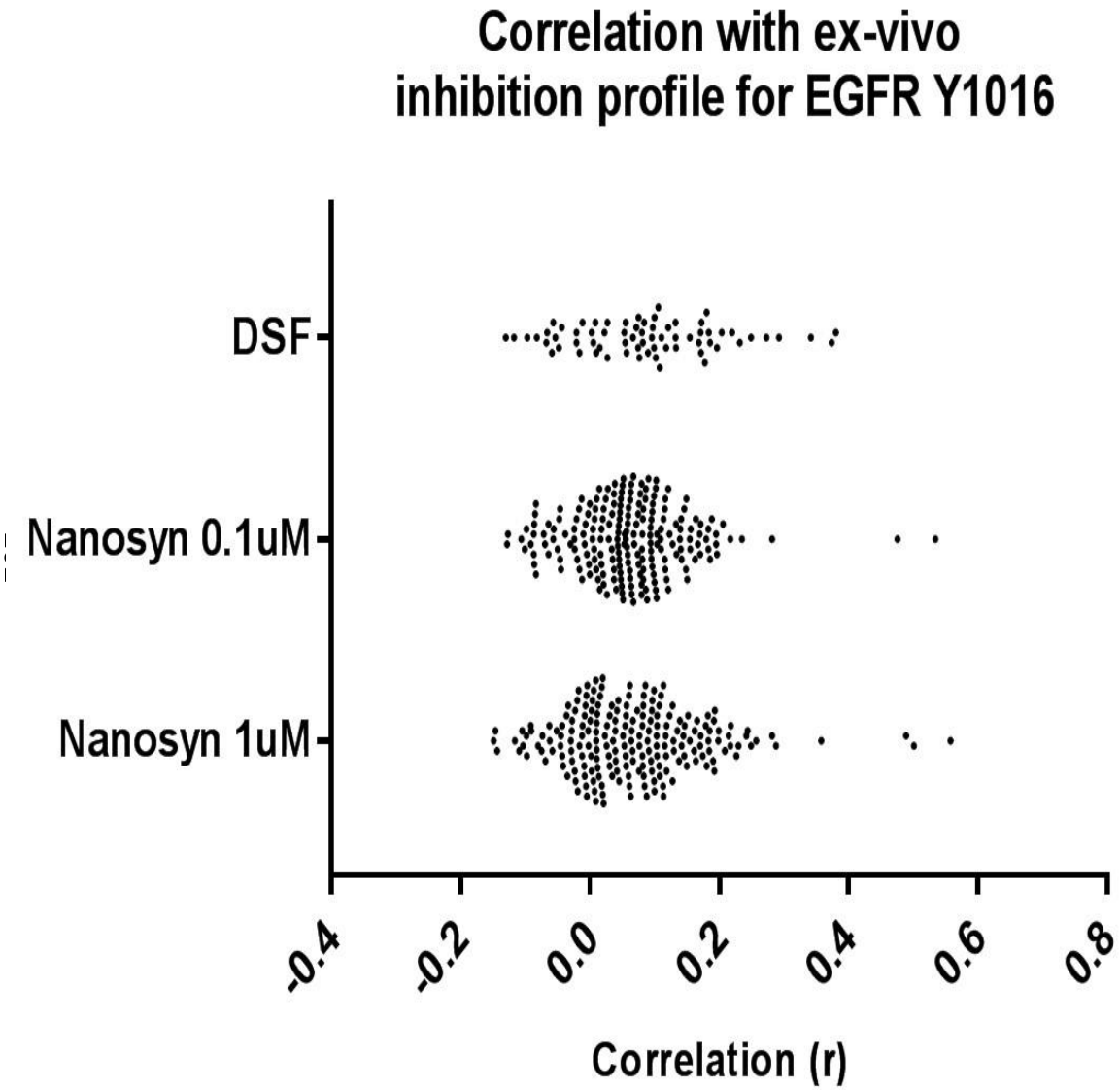
A


Phospho-Tyrosine signal (EGFR peptide)
Absorbance 450nm
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
% extract
EGFR Peptide
No Peptide

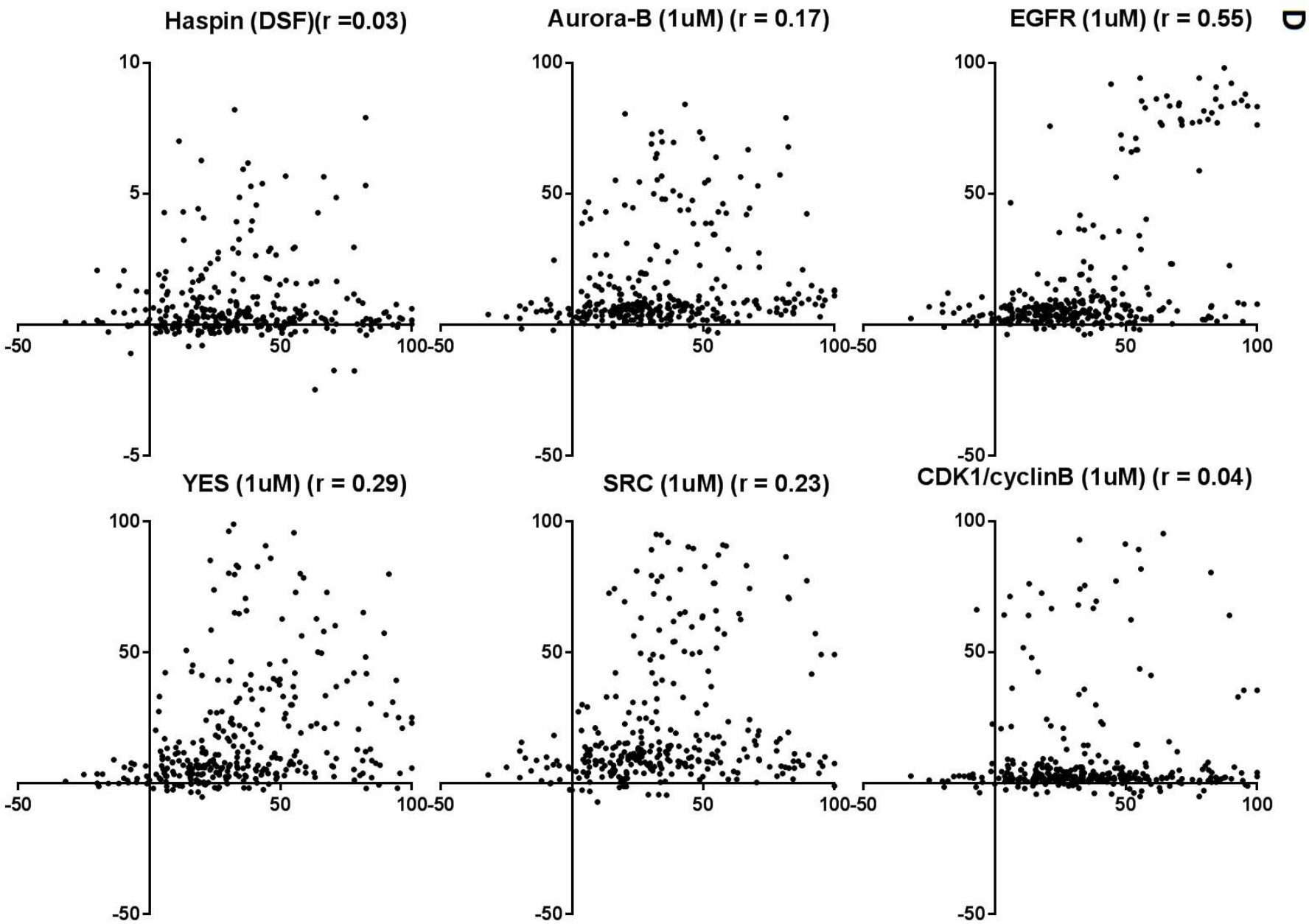
B


Phospho-Tyrosin signal inhibition - EGFR peptide
(PKIS1 compounds)
100.00
80.00
60.00
40.00
20.00
0.00
-20.00
-40.00
0
50
100
150
200
250
300
350

C



Nanosyn 1uM		Nanosyn 0.1uM		DSF	
EGFR	0.5578	EGFR	0.5346	BMX	0.3801
ERBB4	0.5012	ERBB4	0.4758	TEC	0.3737
ERBB2	0.4888	ERBB2	0.2814	NEK2	0.342
MER	0.3579	MER	0.2348	MST4	0.2921
YES	0.2876	PDGFRa	0.2167	STK10	0.273
LCK	0.2817	LCK	0.2055	CHEK2	0.2493
LYNA	0.2571	TYRO3	0.1961	FES	0.2317
LYNB	0.2489	RET	0.1953	STK24	0.2199
FGR	0.2439	LYNA	0.195	TLK1	0.2038
CSK	0.2419	EPHB2	0.1883	SRPK2	0.1961
SRC	0.2297	LYNB	0.1857	STK39	0.1856
HCK	0.2264	FLT1	0.1829	PIM2	0.1836
ARG	0.2175	KDR	0.1827	GUCY2D	0.1806
FYN	0.2167	FGR	0.1795	CDKL1	0.1775
EPHA4	0.2142	SRC	0.171	GRK5	0.1726
BLK	0.2086	DDR2	0.1701	PIM1	0.1718
EPHB2	0.1988	HIPK4	0.1699	ABL2	0.1712
MET	0.1986	ABL1	0.1667	STK17B	0.1682
TYRO3	0.1933	YES	0.1636	SRPK1	0.1543
MNK2	0.1925	BLK	0.1623	DYRK1A	0.1329


D
Haspin (DSF)(r =0.03)
Aurora-B (1uM) (r = 0.17)
EGFR (1uM) (r = 0.55)
YES (1uM) (r = 0.29)
SRC (1uM) (r = 0.23)
CDK1/cyclinB (1uM) (r = 0.04)
10
5
-5
100
50
-50
-50
50
100

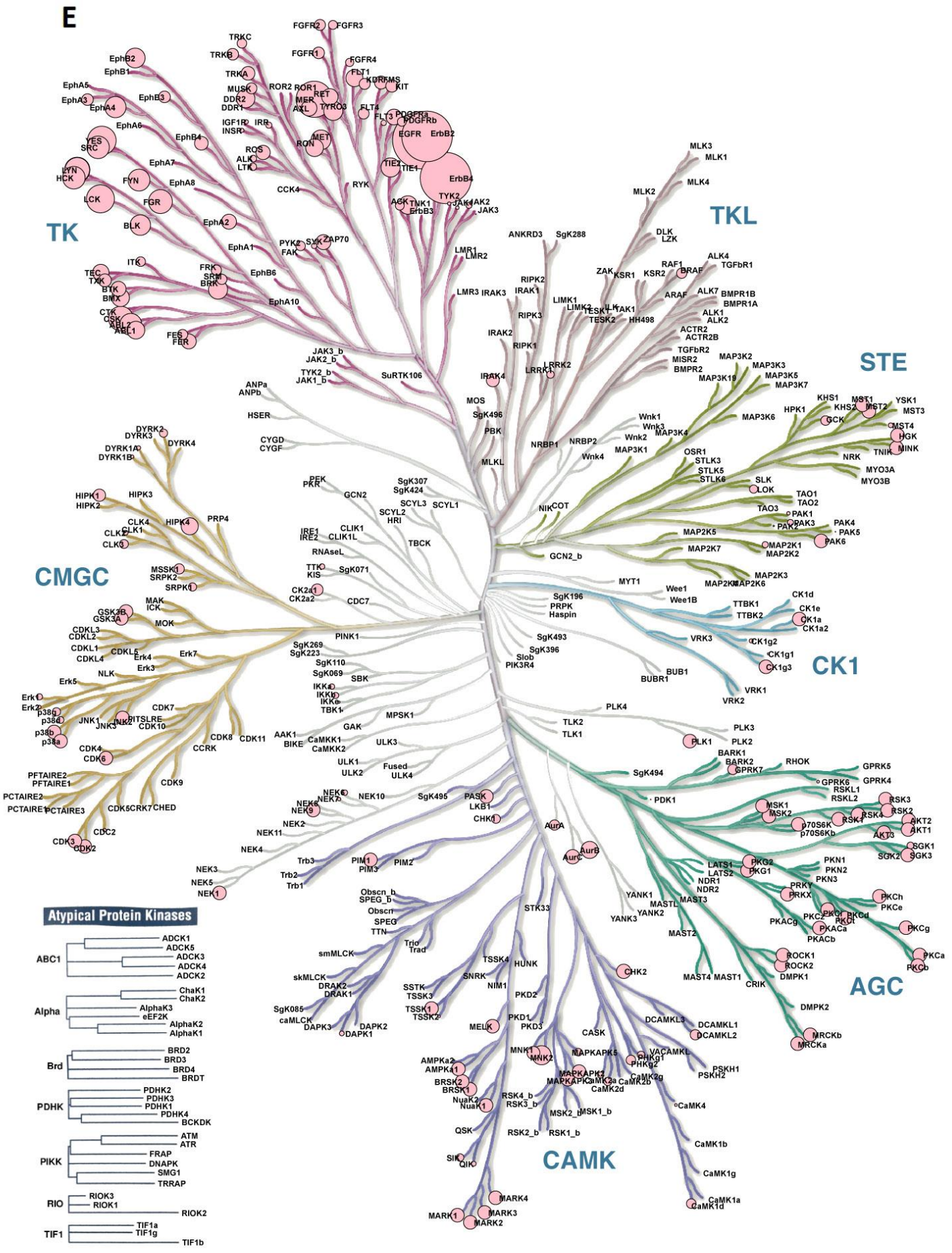

E
TK
TKL
STE
CMGC
CK1
AGC
CAMK
Atypical Protein Kinases
ABC1
Alpha
Brd
PDHK
PIKK
RIO
TIF1
"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"

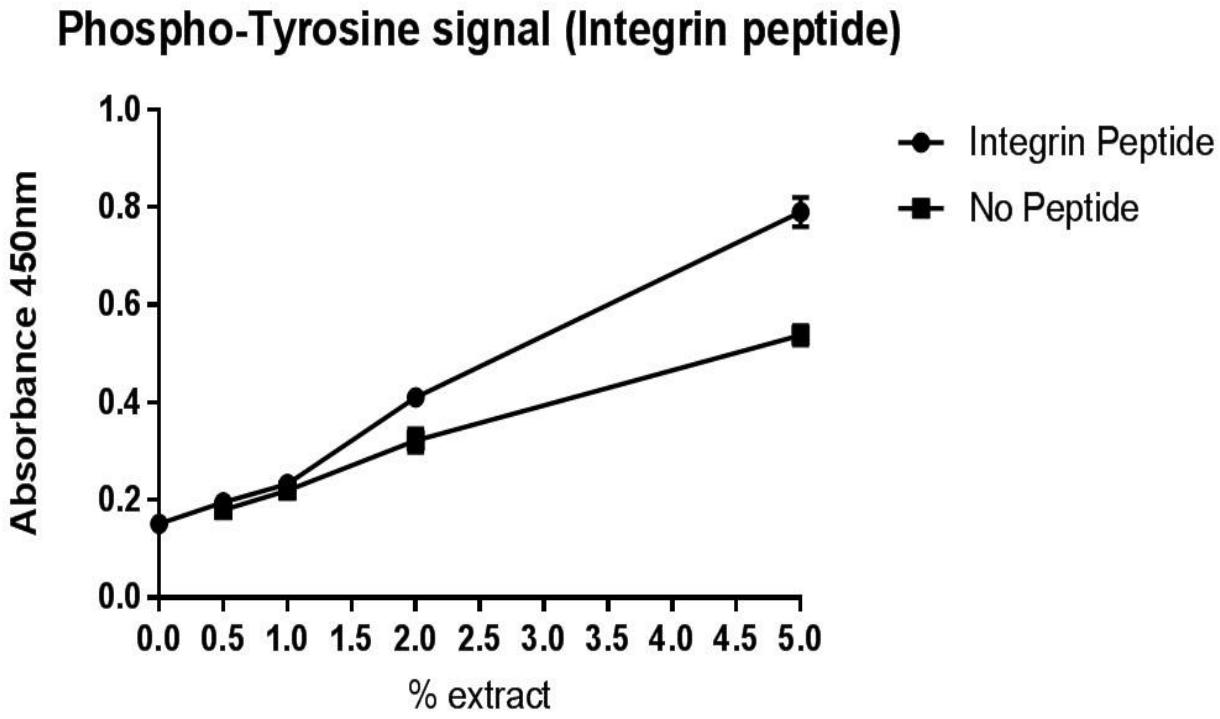
Figure 4.7: **EGFR Y1016ph KiPIK screen with PKIS1 compounds** : **A** KiPIK calibration of peptides corresponding to residues 1013-1021 of EGFR using EGF stimulated (50 ng/ml 5 mins) A431 cell extract. anti-phosphotyrosine antibody used for detection **B** % inhibition of signal by PKIS1 compounds following synchronous kinase reactions **C** Pearson correlation of recombinant kinase *in vitro* inhibition profiles with anti-phosphotyrosine lysate signal inhibition profile (PKIS1 compounds), (right) top 20 correlating kinases from 3 datasets of inhibition profiles **D** Scatter plots of PKIS1 compounds. (x-axis) anti-phosphotyrosine signal inhibition in lysate reactions, against (y-axis) inhibition of recombinant kinase *in vitro* (from inhibitor profiling dataset: Nanosyn 1 µM, except Haspin from DSF 1 µM dataset) **E** Pearson correlation coefficients (as in **C**), for kinases from Nanosyn 1 µM dataset, mapped onto phylogenetic tree of the human kinome. Circle size corresponds to correlation coefficient score.

4.2.10 KiPIK screening identifies SRC kinases as Integrin β1A Y795 kinase

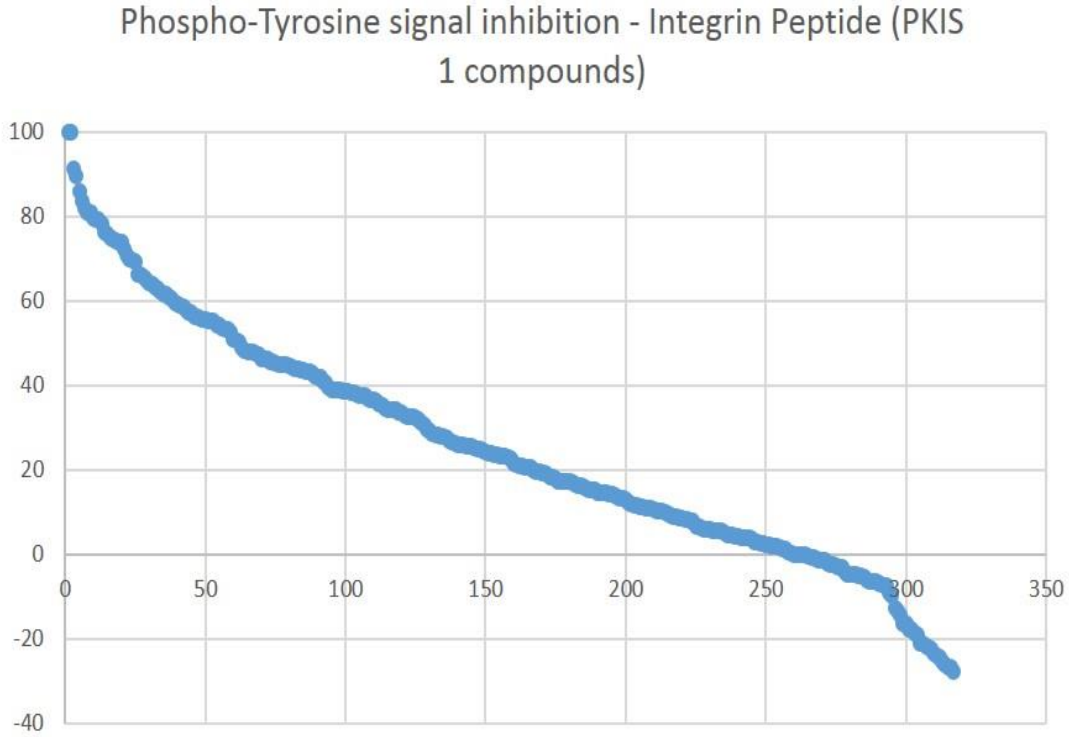
To further test our screening approach, and verify the viability of non-specific phospho-antibodies, we next tried KiPIK screening on a peptide encompassing an Integrin β1A phosphosite at Y795, thought to be a substrate of Src family kinases (Calderwood et al., 2013, Sakai et al., 2001).

The best correlating kinases were found in the Nanonsyn 1 µM library. Of those scoring a Pearson's r value >0.5, 6 out of 7 were Src family kinases. Interestingly, although the screen clearly identified Src family kinases as expected, EGFR was also among the top correlating kinases with r = 0.5235 (note EGF stimulated A431 cells had been used as a source of cell extract) (Figure 4.8C). (Note SRC kinases are not present in the DSF assay panel)

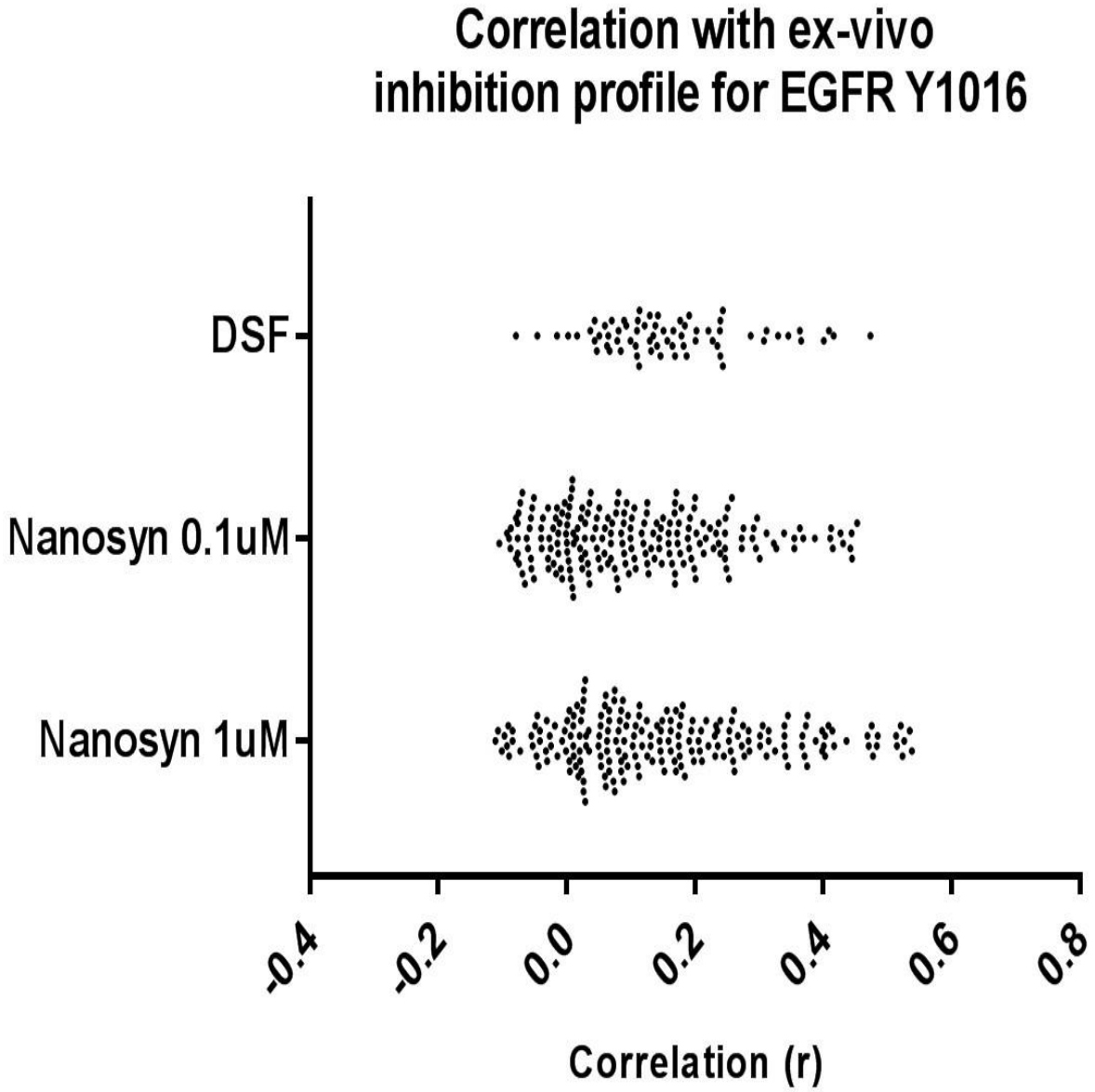
A


Phospho-Tyrosine signal (Integrin peptide)
Absorbance 450nm
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
% extract
Integrin Peptide
No Peptide

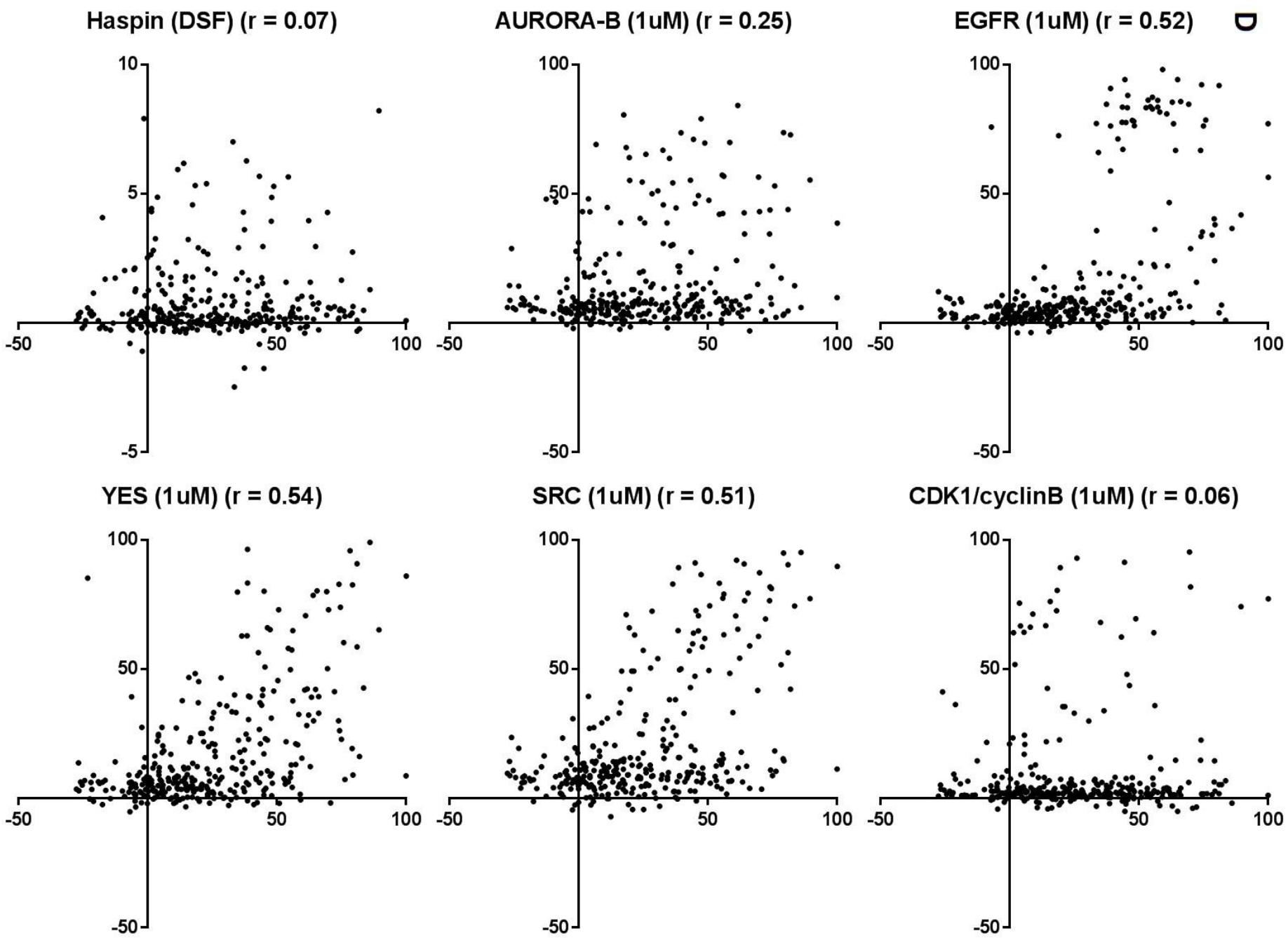
B

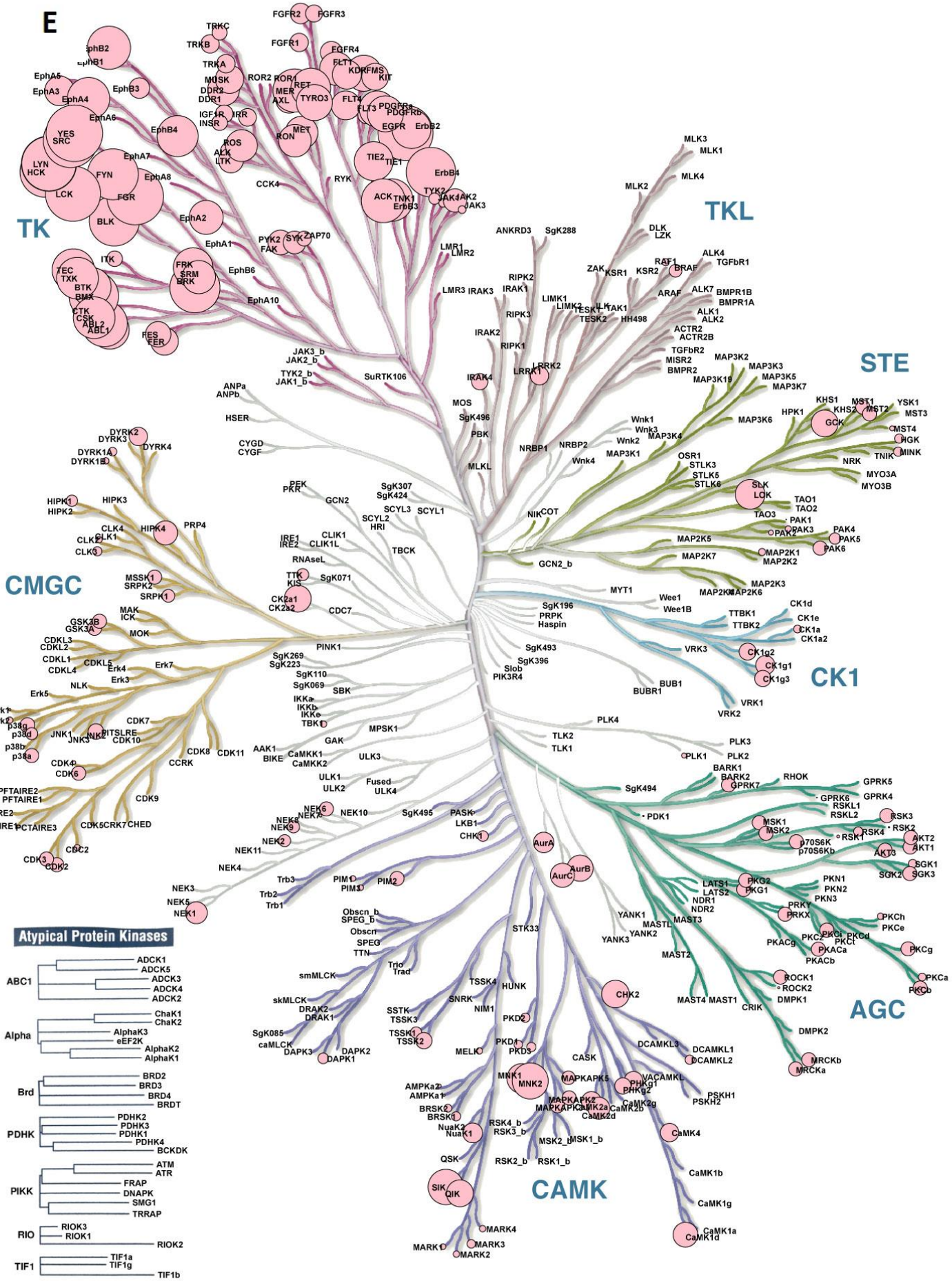

Phospho-Tyrosine signal inhibition - Integrin Peptide (PKIS 1 compounds)
100
80
60
40
20
0
-20
-40
0
50
100
150
200
250
300
350

C



Nanosyn 1uM		Nanosyn 0.1uM		DSF	
YES	0.5375	LCK	0.4525	BMX	0.4736
FGR	0.535	SRC	0.4445	CHEK2	0.4163
LCK	0.5262	YES	0.4442	ABL2	0.4092
EGFR	0.5235	FGR	0.4422	NEK2	0.4007
LYNA	0.5213	LYNA	0.4313	STK10	0.3659
SRC	0.5146	LYNB	0.4267	FES	0.3643
LYNB	0.5138	ARG	0.415	PLK4	0.3461
HCK	0.4851	ABL1	0.4119	GAK	0.33
FYN	0.4838	EGFR	0.4112	TYK2	0.3118
BRK	0.4778	HCK	0.3873	SLK	0.3078
BLK	0.4761	ERBB4	0.3693	STK33	0.2871
BMX	0.4714	FYN	0.3637	PAK4	0.2442
ERBB4	0.4688	PDGFRb	0.3551	MST4	0.244
ARG	0.4362	TNK1	0.3548	MAPK8	0.2403
TNK2	0.4165	BRK	0.3387	STK39	0.2401
EPHB2	0.414	BLK	0.3268	TEC	0.2385
MER	0.4081	RET	0.3233	CAMKK2	0.2353
TNK1	0.4029	PDGFRa	0.3117	GRK5	0.2282
TEC	0.4027	KDR	0.3005	CAMK2A	0.221
EPHA4	0.4008	MNK2	0.2965	PHKG2	0.2017


D
Haspin (DSF) (r = 0.07)
AURORA-B (1uM) (r = 0.25)
EGFR (1uM) (r = 0.52)
YES (1uM) (r = 0.54)
SRC (1uM) (r = 0.51)
CDK1/cyclinB (1uM) (r = 0.06)
10
5
-5
100
50
-50



"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"

Figure 4.8: **Integrin β1A Y795ph KiPIK screen with PKIS1 compounds** : **A** KiPIK calibration of peptides corresponding to residues 784-798 of Integrin β1A using EGF stimulated (50 ng/ml 5 mins) A431 cell extract. anti-phosphotyrosine antibody used for detection **B** % inhibition of signal by PKIS1 compounds following synchronous kinase reactions **C** Pearson correlation of recombinant kinase *in vitro* inhibition profiles with anti-phosphotyrosine lysate signal inhibition profile (PKIS1 compounds), (right) top 20 correlating kinases from 3 datasets of inhibition profiles **D** Scatter plots of PKIS1 compounds. (x-axis) anti-phosphotyrosine signal inhibition in lysate reactions, against (y-axis) inhibition of recombinant kinase *in vitro* (from inhibitor profiling dataset: Nanosyn 1 μM, except Haspin from DSF 1 μM dataset) **E** Pearson correlation coefficients (as in **C**), for kinases from Nanosyn 1 μM dataset, mapped onto phylogenetic tree of the human kinome. Circle size corresponds to correlation coefficient score.

4.2.11 In silico inhibitor subsampling can aid KiPIK screen analysis

The high correlation score of EGFR in the Integrin β1A Y795 screen was unexpected. We reasoned that it was either a consequence of similarity in the inhibition profiles of our inhibitor set between SRC and EGFR kinases (in which case EGFR high correlation was probably an artefact of high SRC correlations); or could represent a genuine and distinct phosphorylation signal occurring in our lysate.

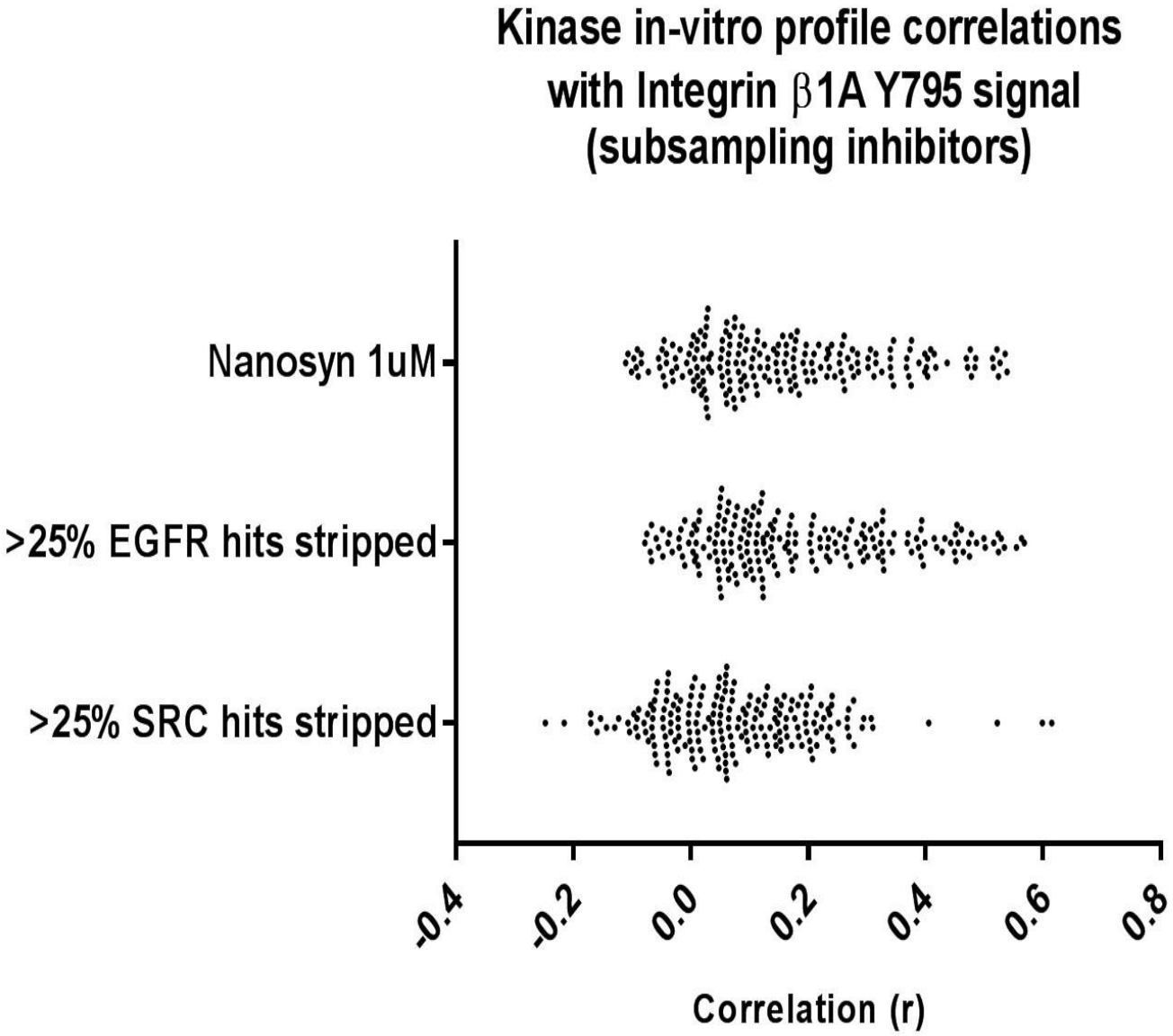
To probe whether the EGFR correlation could be separated from SRC kinases we tried sub-sampling *in silico* the inhibitors we included in our correlation analysis. First we tried stripping out all inhibitors that inhibited EGFR >25% in the Nanosyn screening panel at 1 μM concentration (55 compounds removed of 317 total). This resulted in a significant increase in the correlation score of SRC (r = 0.51 increases to r = 0.57) and LYN kinases. Interestingly when the removed inhibitors were highlighted on the correlation plot for SRC (Figure 4.9B), it was clear that the bulk of these sat around 50% on the X-axis (KiPIK screen inhibitory %) and very low on the Y-axis (Nanosyn profiling % inhibition). The result of stripping these inhibitors out was therefore to ‘clean up’ the correlation. When the complementary procedure was applied (removing compounds that inhibit SRC > 25%; 76 compounds removed of

317 total), the effect on EGFR correlation was similarly striking (r = 0.52 increases to r = 0.61). SRC inhibitor stripping had the effect of dramatically ‘cleaning’ the EGFR correlation plot of compounds that were weakening the correlation due to high KiPIK inhibitory scores and low Nanosyn profiling scores (Figure 4.9B).

It therefore seems likely that distinct activity from both SRC kinases and EGFR were contributing to the phosphorylation signal we were measuring. As we used a generic phospho-tyrosine antibody to detect phosphorylation in this assay, and observed a strong peptide independent signal in our calibration, it is possible that this peptide independent background is the source of EGFR dependent signal. A site-specific antibody to the Integrin peptide phosphosite could be used to confirm this. Notably, the extract used was from A431 cells stimulated with EGF, so EGFR was likely hyperactivated in these extracts. EGFR has also been reported to activate SRC family kinases in some cell types so it is also possible that hyperactivated EGF is contributing to SRC activation in our lysates (Furcht et al., 2015, Bromann et al., 2004).

It also cannot be ruled out that EGFR was contributing to phosphorylation of the Integrin peptide.

A



Nanosyn 1uM		>25% EGFR hits stripped		>25% SRC hits stripped	
YES	0.5375	SRC	0.5674	EGFR	0.6142
FGR	0.535	LYNA	0.5632	ERBB4	0.5982
LCK	0.5262	LYNB	0.5541	ERBB2	0.522
EGFR	0.5235	LCK	0.5339	MER	0.4053
LYNA	0.5213	HCK	0.5286	CAMK1D	0.3094
SRC	0.5146	ARG	0.5235	CHEK2	0.3067
LYNB	0.5138	FGR	0.5227	BRK	0.2941
HCK	0.4851	BMX	0.5159	TNK2	0.2914
FYN	0.4838	FYN	0.5013	AXL	0.278
BRK	0.4778	EPHB2	0.4991	PTK5	0.2768
BLK	0.4761	EPHA4	0.4861	CAMK4	0.2689
BMX	0.4714	TXK	0.478	YES	0.2686
ERBB4	0.4688	BLK	0.4745	SYK	0.2663
ARG	0.4362	YES	0.4727	TSSK2	0.2661
TNK2	0.4165	TNK2	0.4635	LCK	0.2513
EPHB2	0.414	TEC	0.4602	FES	0.2423
MER	0.4081	EPHB4	0.4552	IRR	0.2404
TNK1	0.4029	PDGFRb	0.4528	PRAK	0.2402
TEC	0.4027	TNK1	0.4513	CAMK2A	0.2396
EPHA4	0.4008	BRK	0.4503	CAMK2D	0.2368

B

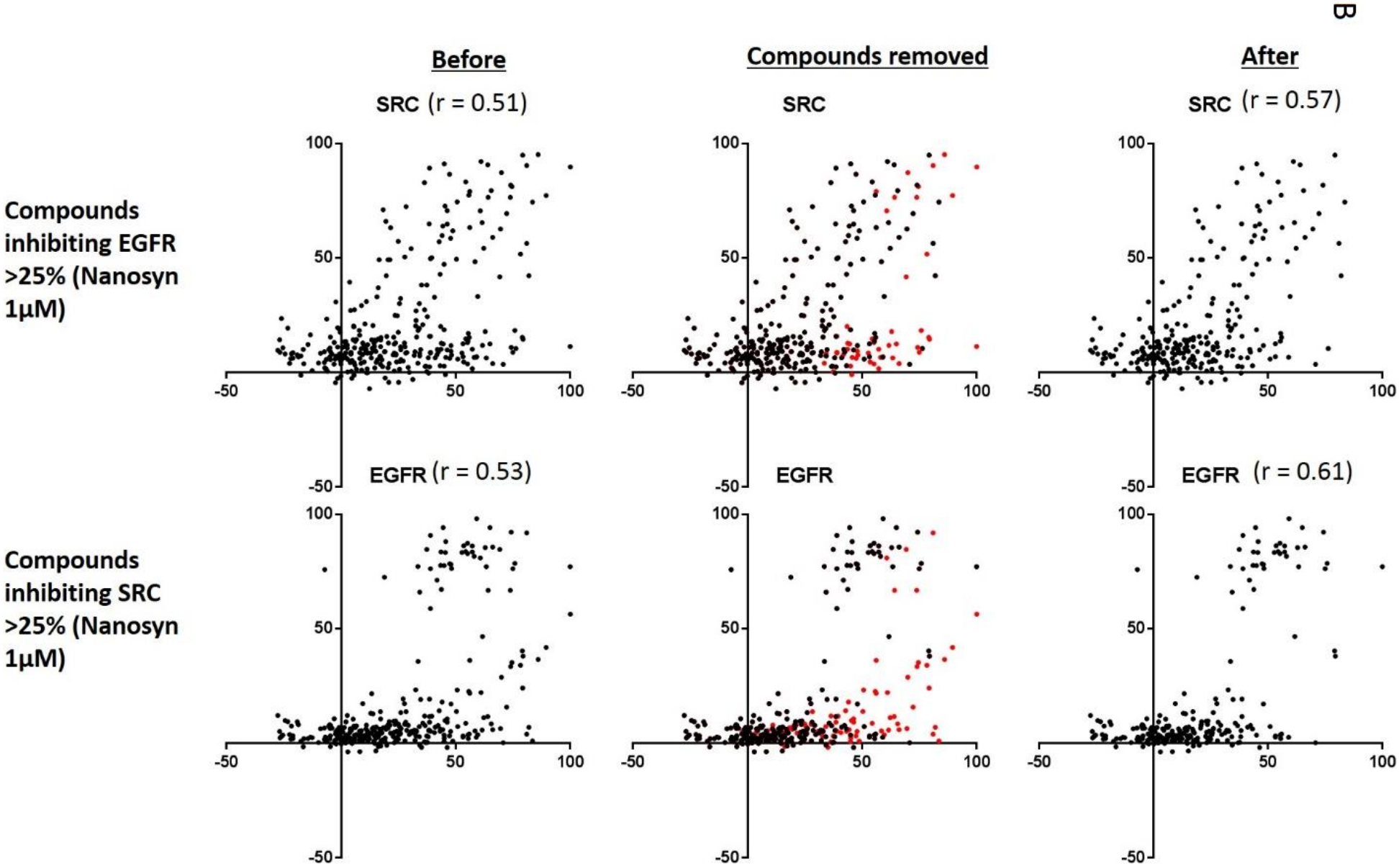

Before
Compounds removed
After
SRC (r = 0.51)
SRC
SRC (r = 0.57)
EGFR (r = 0.53)
EGFR
EGFR (r = 0.61)
Compounds inhibiting EGFR >25% (Nanosyn 1µM)
Compounds inhibiting SRC >25% (Nanosyn 1µM)
100
50
-50
-50
50
100

Figure 4.9: **Effect of *in silico* subsampling PKIS1 compounds on Integrin β1A Y795ph KiPIK screen correlations: A E**ffect on kinase correlations (from Figure 4.8) of removing all compounds which inhibit EGFR >25% (55 compounds of 317) or SRC >25% (76 compounds of 317) (All correlations with Nanosyn 1 µM). **B S**catter plots of PKIS1 compounds before and after stripping indicated compounds (x-axis) anti-phosphotyrosine signal inhibition in lysate reactions, against (y-axis) inhibition of recombinant kinase *in vitro* (from Nanosyn 1 µM inhibitor profiling dataset)

4.2.12 Investigating H3T11ph with KiPIK

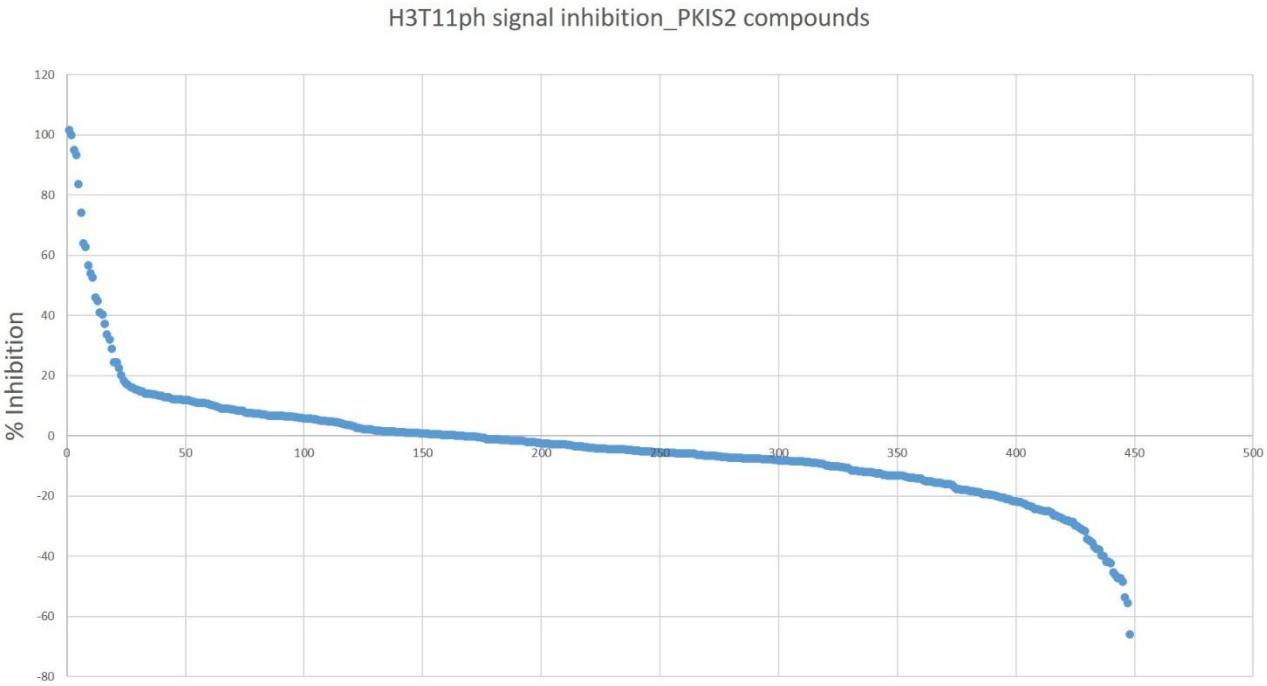
We also carried out KiPIK screening on H3T11ph using mitotic extracts. First, we tried KiPIK screening with the PKIS1 library. However we found that strikingly few compounds were capable of inhibiting the H3T11ph signal. Of the library of PKIS1 compounds tested, only two were capable of inhibiting the phosphorylation >50%. Although non-inhibition can also be informative when correlating kinases by KiPIK, none of the kinases in the PKIS1 profiled kinase had strong correlations with our H3T11ph signal. To gain more insight, we decided to KiPIK screen H3T11 phosphorylation with a larger inhibitor library which had been profiled on more kinases. We obtained PKIS2, which consists of 645 inhibitors which have been profiled on 392 unique protein kinases (all at 1 µM on the DiscoverX assay, a competitive binding assay).

We completed a partial screen of PKIS2 on H3T11ph (447 compounds). As we had seen with PKIS1, compounds which inhibited H3T11ph >50% were rare (11 total) (Figure 4.10A). Resulting correlations with PKIS2 profiled kinases were also low (Figure 4.10B) and the range of correlations across the kinome was very narrow in comparison to previous screens. This was particularly evident when plotting the results of the screen onto the kinome tree. Unlike in previous screens, the highest hitting kinases were not clustered to a particular region of the tree (which would indicate phylogenetic similarity) (Figure 4.10C).

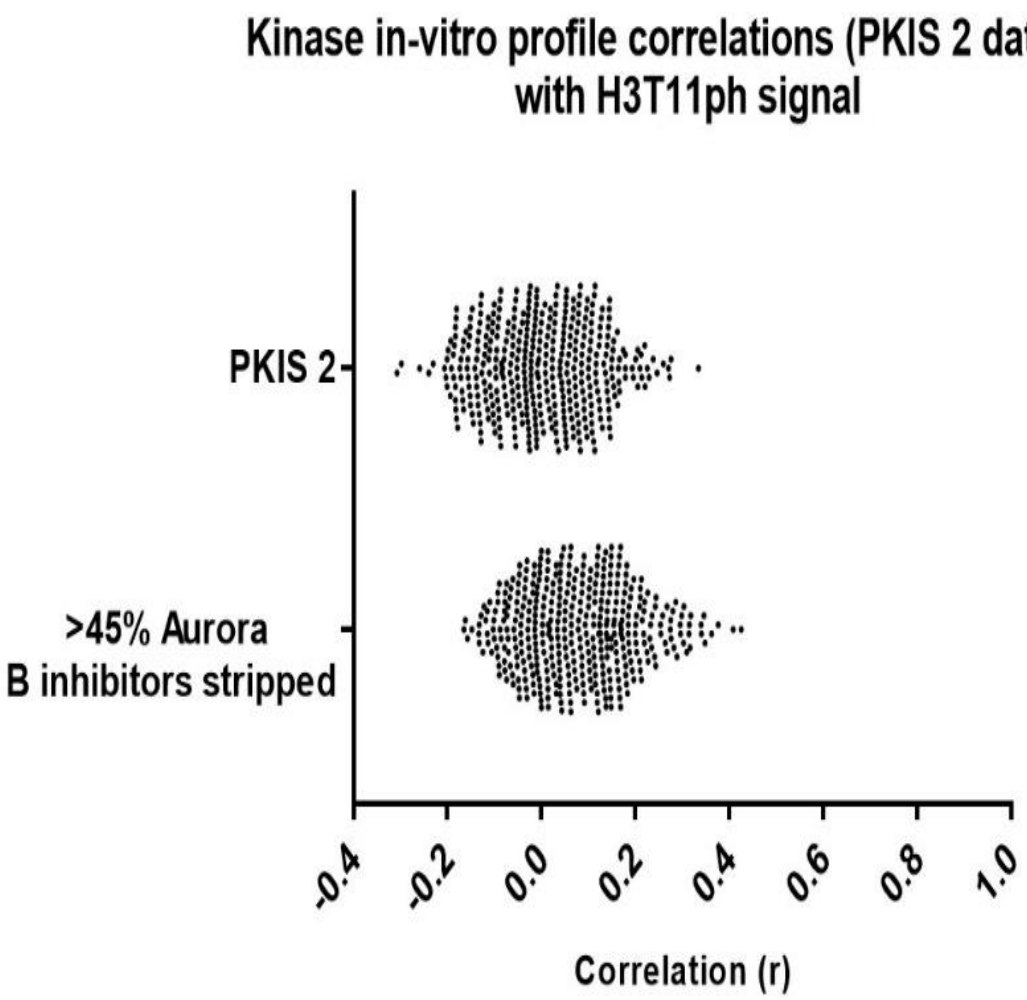
NEK10 had the highest correlation but it was clear from the compound scatter plot that a significant portion of compounds which inhibited NEK10 >50% in the

Nanosyn assay were not reducing our H3T11ph signal; many were even increasing H3T11ph. As discussed below, we determined that Aurora kinase inhibition was increasing our H3T11ph signal. This effect would therefore shift all compounds which significantly inhibited Aurora kinases towards the left side of the x-axis (lysate reaction inhibition) in our correlation scatter plots – distorting correlations based on H3T11ph signal inhibition. To prevent this distortion of our correlation results we removed, *in silico,* compounds which inhibited Aurora B >45% from our correlation calculations. This improved the correlation of NEK10 and several other kinases. It also improved a visual assessment of their correlation by removing compounds which increased the H3T11ph signal in our assay (those with a negative inhibitory score), yet inhibited the kinases (NEK10 and PHKG2) in the DiscoverX assay (Figure 4.10D). After this adjustment the top 2 candidates were PHKG2 and NEK10 with correlation scores of 0.42 and 0.41 respectively. Follow up of these candidates in cells is ongoing.

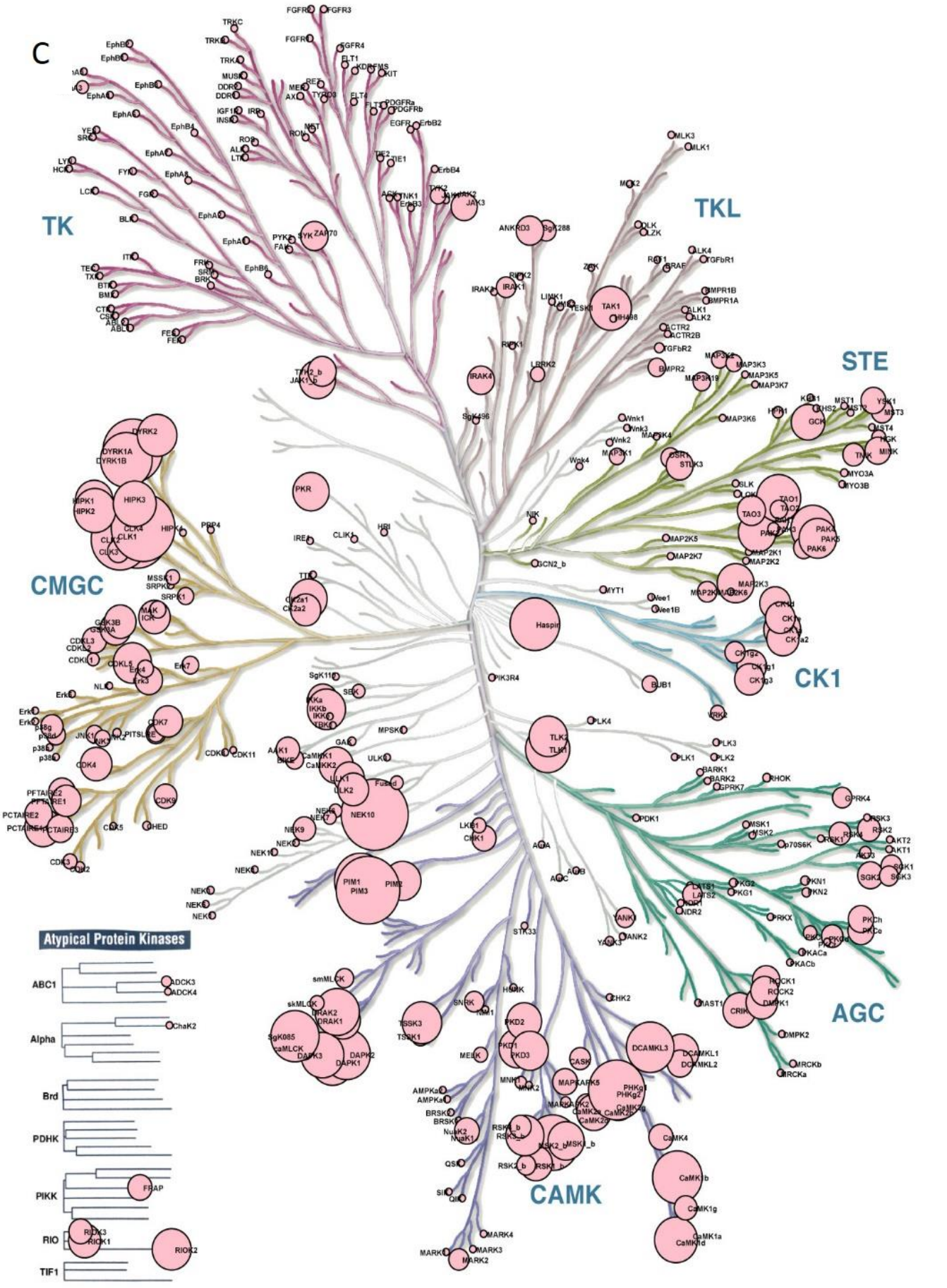
A


H3T11ph signal inhibition_PKIS2 compounds
% Inhibition
120
100
80
60
40
20
0
-20
-40
-60
-80
0
50
100
150
200
250
300
350
400
450
500

Bottom correlations (PKIS2)	
ABL1-nonphosphorylated	-0.182
EPHA1	-0.182
CSK	-0.183
SRMS	-0.183
ABL2	-0.185
FLT4	-0.187
EPHA6	-0.187
VEGFR2	-0.191
RIPK2	-0.193
TXK	-0.193
LOK	-0.198
TNK1	-0.201
TIE1	-0.201
TNK2	-0.203
BTK	-0.206
EPHB6	-0.230
HCK	-0.239
AURKA	-0.259
BRK	-0.298
AURKB	-0.307



PKIS2		>45% Aurora B inhibitors stripped	
NEK10	0.335	PHKG2	0.426
CLK4	0.276	NEK10	0.408
PIM3	0.273	PIM3	0.377
CLK1	0.271	CLK1	0.362
DYRK1A	0.260	CLK4	0.349
.CLK2	0.247	MYLK4	0.346
PHKG2	0.238	SNARK	0.341
TLK1	0.229	TAOK1	0.340
DCAMKL3	0.226	MST3	0.318
RPS6KA4(Kin.Dom.2-C-terminal)	0.221	MAP4K2	0.317
PIM1	0.220	TLK1	0.315
CAMK1B	0.214	CAMK1D	0.314
DYRK1B	0.212	DAPK2	0.310
TAOK1	0.211	DYRK1A	0.303
VPS34	0.206	PIM1	0.302
HIPK3	0.205	MEK3	0.299
HASPIN	0.201	HIPK3	0.297
MYLK4	0.196	PAK7	0.296
DAPK2	0.194	CLK2	0.289
DRAK2	0.179	JAK3(JH1domain-catalytic)	0.287



"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"

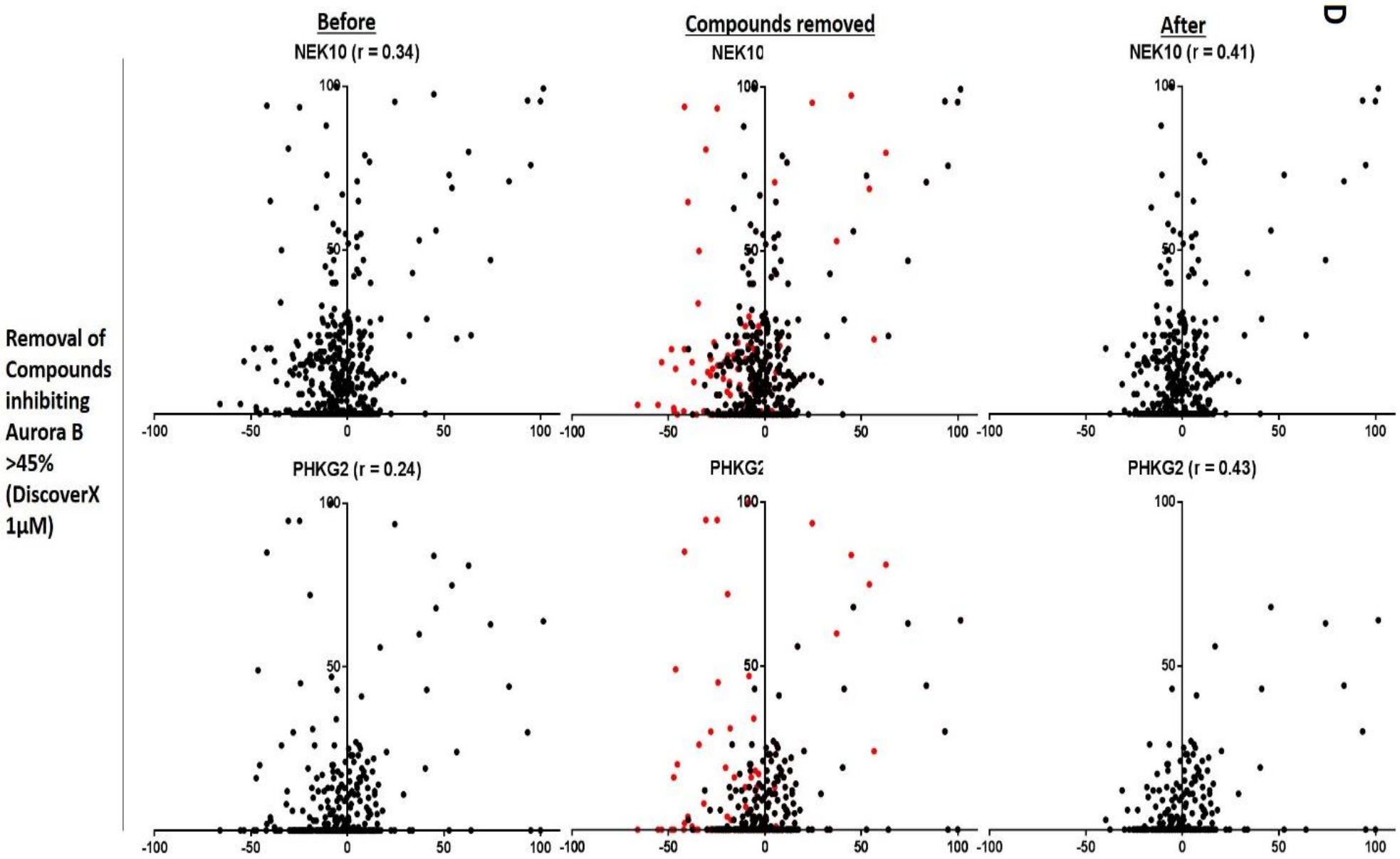

D
Before
Compounds removed
After
NEK10 (r = 0.34)
NEK10
NEK10 (r = 0.41)
PHKG2 (r = 0.24)
PHKG2
PHKG2 (r = 0.43)
Removal of Compounds inhibiting Aurora B >45% (DiscoverX 1µM)
100
50
-100
-50
0
50
100

Figure 4.10: **H3T11ph KiPIK screen with PKIS2 compounds** : **A** % inhibition of H3T11ph signal by PKIS2 compounds following synchronous kinase reactions **B** Pearson correlation of recombinant kinase *in vitro* inhibition profiles with H3T11ph lysate signal inhibition profile (PKIS2 compounds), (left) bottom 20 correlating kinases (right) top 20 correlating kinases. Also with compounds inhibiting Aurora B >45% stripped from dataset **C** Pearson correlation coefficients (as in **B,** PKIS2), for kinases from DiscoverX 1 µM dataset, mapped onto phylogenetic tree of the human kinome. Circle size corresponds to correlation coefficient score. **D** Effect of removing >45% Aurora B inhibiting compounds (in DiscoverX) on **s**catter plots of PKIS2 compounds (x-axis) H3T11ph signal inhibition in lysate reactions, against (y-axis) inhibition of recombinant kinase *in vitro* (from DiscoverX 1 µM inhibitor profiling dataset)

Inverse correlation with Aurora B

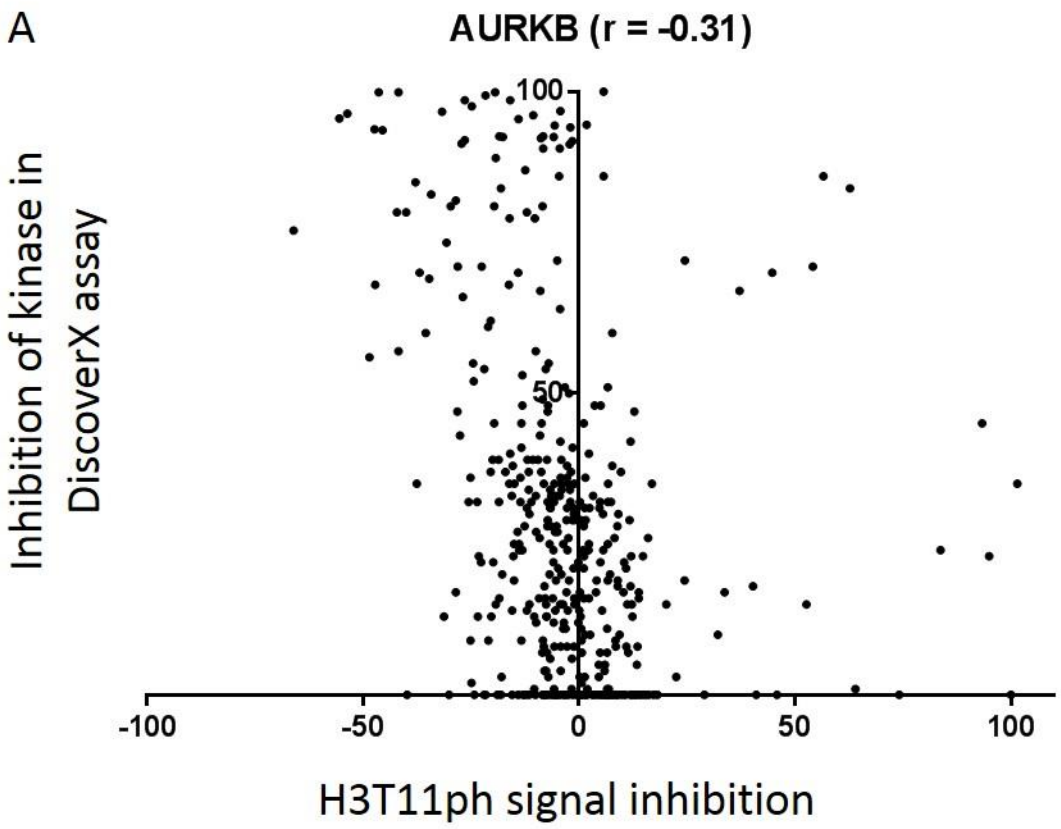
Interestingly the screen indicated an inverse correlation with both Aurora A and B. This negative correlation was clearly evident upon visual inspection of the compound scatter plot of Aurora B (vs H3T11 lysate signal) (Figure 4.11A), indicating that compounds which inhibited Aurora B increased the signal of H3T11ph in our lysate reactions. As a major substrate of Aurora B is the neighbouring residue of H3S10, we hypothesized that phosphorylation of H3S10 was inhibitory to phosphorylation of H3T11. To test this hypothesis, we tested the ability of mitotic lysate to phosphorylate another H3 peptide on which H3S10 was pre phosphorylated (Figure 4.11B). Strikingly, in contrast to an unmodified H3 peptide, we detected no H3T11 phosphorylation, suggesting that pre-existing H3S10 phosphorylation prevents phosphorylation of H3T11. Importantly in figure 3.5 (chapter 3) we established that our H3T11 antibody was only modestly reduced in binding affinity for H3T11ph when neighbouring H3S10 was also phosphorylated. This suggests that the unknown H3T11 kinase competes with Aurora kinases for H3 peptides in our lysates, thereby accounting for the inverse correlation with Aurora A and B in our KiPIK screen (reflecting an increase in H3T11 phosphorylation in lysates with Aurora inhibitors).

We were interested in whether this phenomenon could be preserved *in vivo*. H3S10ph is strong throughout mitosis so we reasoned that if we depleted H3S10ph by inhibiting Aurora B we might see an increase in H3T11ph. To test this, we probed the effect of Aurora B inhibition (with ZM-447439) on H3T11ph levels in

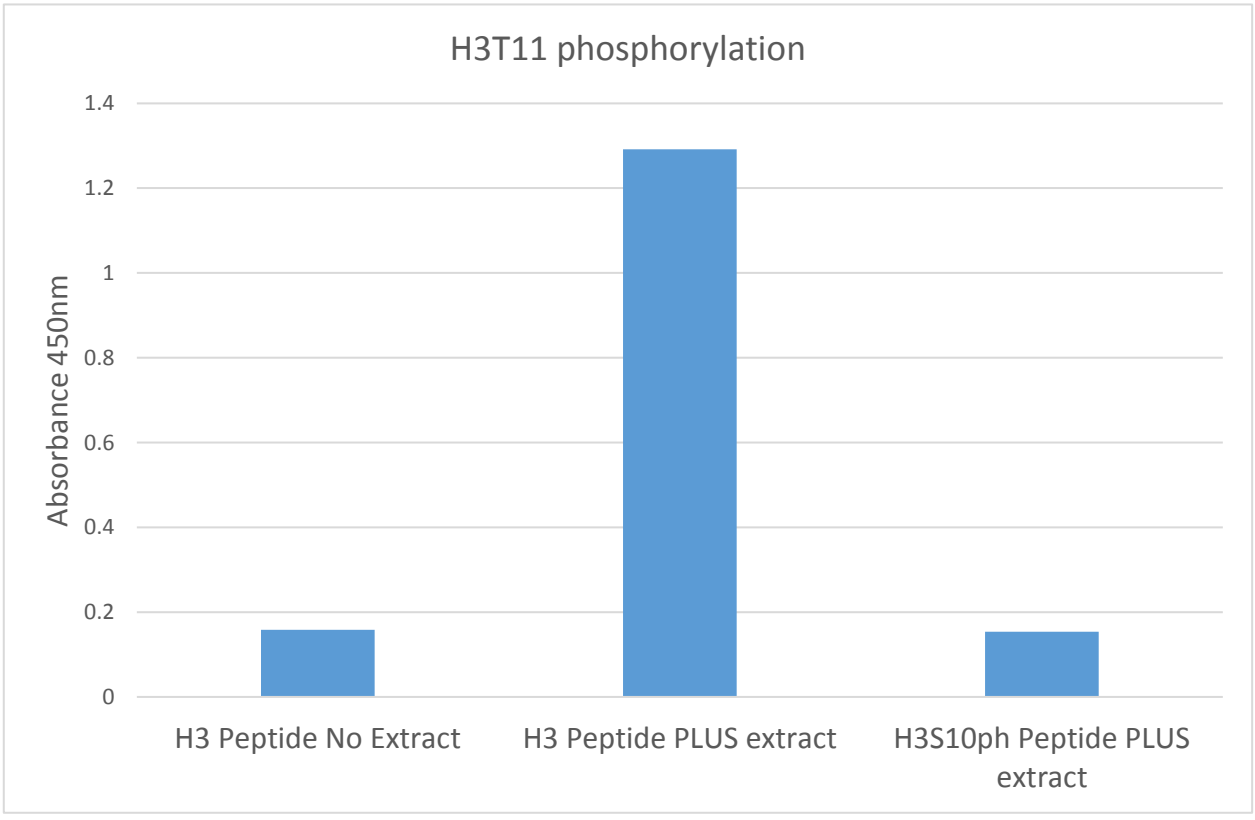
Hela cells. Cells were treated with either Nocodazole alone (6 hours), Nocodazole and ZM-447439 (6 hours), or Nocodazole for 5 hours followed by addition of ZM-447439 for 1 hour.

ZM-447439 addition consistently resulted in a spreading of the H3T11ph immunofluorescence signal from a somewhat punctate signal, strongest around the centromeres, to a uniform signal covering the entire chromatin. Strikingly this signal was significantly brighter when ZM-447439 was added for one hour at the end of a Nocodazole arrest rather than at the beginning (Figure 4.11C) (discussed further in 4.3.2).

A


AURKB (r = -0.31)
Inhibition of kinase in DiscoverX assay
100
50
-100
-50
0
50
100
H3T11ph signal inhibition

B


H3T11 phosphorylation
Absorbance 450nm
1.4
1.2
1
0.8
0.6
0.4
0.2
0
H3 Peptide No Extract
H3 Peptide PLUS extract
H3S10ph Peptide PLUS extract

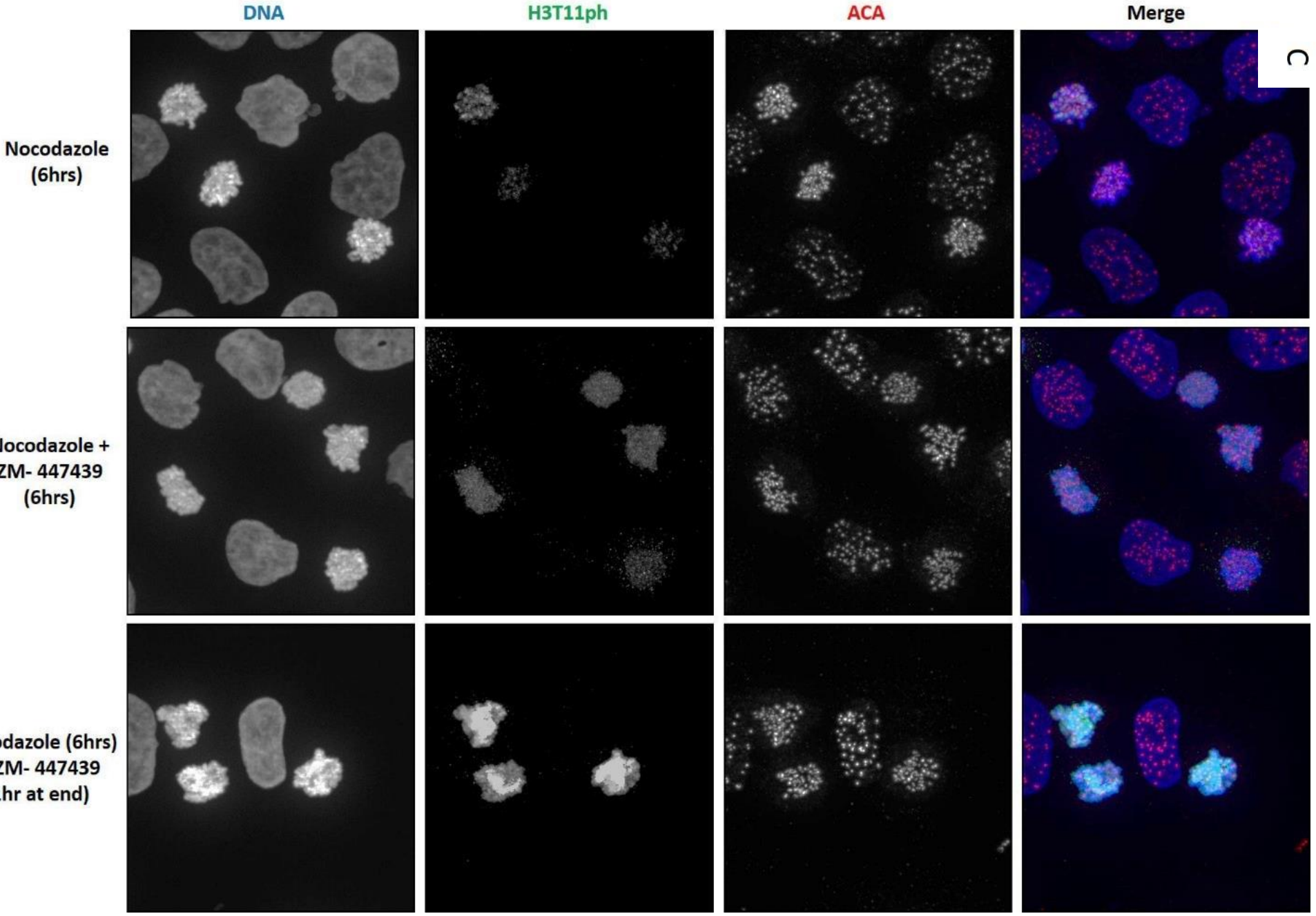

DNA
H3T11ph
ACA
Merge
C
Nocodazole (6hrs)
Nocodazole + ZM- 447439 (6hrs)
Nocodazole (6hrs) + ZM- 447439 (1hr at end)

Figure 4.11: **Antagonism of H3T11ph by Aurora B in lysates and in cells** **A** Scatter plot of PKIS2 compounds from KiPIK screen performed in 4.10, (x-axis) H3T11ph signal inhibition in lysate reactions, against (y-axis) inhibition of recombinant kinase *in vitro* (from DiscoverX 1 µM inhibitor profiling dataset). **B** H3 1-21 and H3(S10ph) 1-21 peptides were incubated with or without Mitotic extract (2%) in KiPIK buffer for 30 minutes and probed for H3T11ph by ELISA. **C** HeLa cells were treated as indicated, stained with DAPI and probed for H3T11ph and ACA. (Nocodazole at 200 nmol, ZM447439 at 10 µM)

4.3 Discussion

4.3.1 KiPIK screening is effective for identifying direct kinases of phosphorylation sites of interest

KiPIK screening was able to unambiguously identify the established direct kinases for 4 diverse phosphorylation sites tested, indicating that we have developed a generally applicable method for identifying the direct kinases of phosphorylation sites of interest.

The KiPIK screen measuring H3T3 phosphorylation clearly identified Haspin as the phosphorylating kinase, in line with the literature (Dai et al., 2005). Other high scoring kinases included DYRK and PIM kinases. A conserved feature among many Haspin inhibiting compounds is activity against the DYRK kinases, presumably reflective of similarities in the structure of their active sites (Cuny et al., 2010, Cuny et al., 2012). Similarly, when we carried out hierarchical clustering of kinases based on their inhibition profiles across profiled inhibitor datasets (such as PKIS2) the PIMs (along with DYRKs) consistently cluster with Haspin, reflecting that they share similar inhibition profiles (See Appendix A for an example hierarchical clustering; of PKIS1 Nanosyn 1 µM). This feature of KiPIK screening means that superior correlation scores will often be found for kinases with pharmacological similarity to the genuine substrate phosphorylating kinase. For this reason, examination of phylogenetic trees (as illustrated multiply in this chapter), or a hierarchical clustering tree of the inhibitor panel (which are often strikingly similar to phylogenetic trees), can be useful when considering the results of a KiPIK screen.

The KiPIK screen for the H3S28 phosphorylating kinase highlighted Aurora kinases as responsible for this phosphorylation event, again in line with the literature (Crosio et al., 2002). The correlation scores were considerably lower than for the H3T3ph screen, likely a consequence of having to discard a considerable portion of the data due to a technical error (4.3.7). Aurora kinases have highly conserved catalytic domains and subcellular localisation is thought to facilitate differences in the substrates they target in cells (Carmena et al., 2009, Li et al., 2015). It is therefore possible that all expressed Aurora kinases phosphorylated our H3 peptide in the e*x vivo* KiPIK reactions. Despite the similarity of Aurora kinases, several of the compounds in the PKIS1 set differ quite significantly in their inhibition of each Aurora kinase; this could potentially contribute to reduced correlation scores for each Aurora kinase if, in the lysate reaction, they are all contributing to the phosphorylation of H3S28.

KiPIK screening of EGFR Y1016ph, an autophosphorylation site (Walton et al., 1990, Rotin et al., 1992), clearly highlighted EGFR as the phosphorylating kinase. A generic phospho-tyrosine antibody was used for signal detection in this assay indicating that a site specific antibody is not required for KiPIK screening (when using ELISA as the detection method) (see 6.6 for discussion of alternative detection methods). The background signal detected with the phospho-tyrosine was higher than for the screens using phosphorylation site specific antibodies. Despite this, the correlation scores for EGFR were very high.

The Integrin β1A Y795 KiPIK screen also used a generic phospho-tyrosine antibody. This phosphorylation site is thought to be the target of SRC family kinases (Calderwood et al., 2013, Sakai et al., 2001). Consistent with this, 6 of the top 7 (and 10 of the top 11) correlating kinases in our screen were from the Src family. The receptor tyrosine kinase EGFR was also among the top correlating kinases, but is rather distinct from the src kinases phylogenetically. Because of this (and as the extract had been stimulated with EGF) we suspected that the EGFR signal was a distinct signal, rather than it having high correlation as a consequence to SRC family kinase similarity. To explore this possibility, we investigated whether it was possible to separate the two signals by sub sampling our dataset. *In silico* sub

sampling to remove EGFR hitting inhibitors significantly increased the correlation of some Src family kinases (SRC and LYN), while reducing the correlation of others (YES). It also seemed to clean the SRC correlation plot specifically of inhibitors that were reducing correlation. Performing the inverse sub sampling had an even greater effect on EGFR, strengthening the correlation from 0.52 to 0.61, and 'cleaning' the correlation plot dramatically. While this approach was far from unbiased, the observation that removing each set of inhibitors seemed to 'clean' the correlation plots, combined with the increases in correlation scores, supports the notion that the EGFR signal was distinct (possibly a consequence of using the generic phospho-tyrosine antibody and imperfect plate washing; or the interplay of EGFR and Src kinase signalling in cells). Rerunning the KiPIK screen with a site specific phospho antibody could help to determine this.

4.3.2 KiPIK screening of H3T11 phosphorylation in mitosis

KiPIK screening results

We also KiPIK screened H3T11ph, for which the mitotic kinase is unknown. H3T11ph was potently inhibited by only 2 compounds in the PKIS1 inhibitor set, which led to very poor correlations with all the kinases profiled on PKIS1 (data not shown). The PKIS1 compounds were assembled from published GlaxoSmithKline small molecule inhibitors. Although the authors avoided overpopulating PKIS1 with compounds from particular chemotypes there are biases in the inhibitory coverage of the set (likely over representing kinases that have received more drug-development attention). Of the 224 kinases profiled on the PKIS1 set 18 kinases were not inhibited >50% by any of the compounds (at 1 µM) (Elkins et al., 2016). This suggests that the H3T11ph kinase is not within the kinases profiled on PKIS1, and that it could be an unusual or understudied kinase.

We then tried KiPIK screening H3T11ph using the PKIS2 inhibitor set (which is larger and more diverse than PKIS1) (Drewry et al., 2017). 11 compounds inhibited H3T11ph >50%. After removing compounds which inhibited Aurora B >45% (discussed below and in 4.2.12), PHKG2 and NEK10 correlated most strongly with our H3T11ph signal. Investigation of these kinases *in vivo* is ongoing.

KiPIK screening reveals an antagonistic role for Aurora B on H3T11 phosphorylation in mitosis

The PKIS2 KiPIK screen of H3T11ph also revealed an inverse correlation with Aurora kinases. As the neighbouring residue H3S10 is a prominent Aurora B substrate we suspected that competing phosphorylation of this residue could be the source of the inhibitory role of Aurora kinase activity on H3T11ph. In agreement with this we found that a pre-phosphorylated H3 peptide (H3S10ph) could not be phosphorylated at H3T11 in a mitotic lysate kinase reaction. Additionally, we observed that Aurora B inhibition *in vivo* led to a spreading of the H3T11ph signal on mitotic chromosomes following a 6 hour nocodazole arrest (from a usual centromerically enriched signal to an evenly spread signal across the chromosome arms). Interestingly, when Aurora B was inhibited during mitotic arrest, the H3T11ph signal was both spread and much more intense. It therefore seems likely that H310ph antagonises or prevents phosphorylation of H3T11 by the H3T11 kinase.

Interestingly, when Aurora B was inhibited for 1 hour at the end of a mitotic arrest rather than at the beginning, the increase and spreading of the H3T11ph signal was greater. The reason for this is unclear. It could be related to Aurora B's role in inhibiting the mitotic phosphatase PP1 (Nasa et al., 2018) which is thought to have an indirect role in dephosphorylating H3T11ph in mitosis (Qian et al., 2011). Another possible explanation is that an Aurora B activity at the beginning of mitosis also facilitates H3T11ph by some means. One possible (though speculative) model of this is that HP1 bound to the adjacent H3K9 residue (when di or tri methylated) also antagonises phosphorylation of H3T11. As discussed in the introduction (1.5.6) H3S10ph is required for the removal of HP1 from di or tri methylated H3K9 at the beginning of mitosis. An important difference in our two Aurora inhibitor treatments could be that they have significantly different HP1 occupancy levels on H3K9. If cells enter mitosis without Aurora B activity, HP1 is not displaced and remains on the chromatin (Fischle et al., 2005). Whereas, if Aurora B is inhibited for 1 hour after cells have already been in mitosis for some time, H3S10 undergoes dephosphorylation (we observed loss of H3S10ph, data not shown), but HP1 has already been displaced – potentially allowing greater H3T11 phosphorylation.

Chapter 5. Using KiPIK screening to identify the kinase responsible for an unassigned phosphorylation site on INCENP

5.1 Introduction

5.1.1 Identifying Kinases for an unassigned phosphorylation site

We were interested to see if we could use KiPIK screening to identify kinases for unassigned phosphorylation sites.

In a phosphoproteomic study of mitotically enriched cells, Dephoure *et al.* (2008) categorised the identified phosphosites by motif analysis and concluded that the majority could be assigned to the consensus motifs of known mitotic kinases. Interestingly, however, they identified two unique motifs and suggested that there were undiscovered mitotic kinases (Dephoure et al., 2008). Among proteins which were phosphorylated on these unique motifs, we identified an appealing candidate: INCENP (residue S446).

Interestingly, this phosphorylation site was also detected in a study by Hegemann *et al.* (2011). These authors carried out phosphoproteomic analysis on mitotic cells that they had treated with small molecule kinase inhibitors and identified INCENP S446ph as among the phosphorylation sites that were sensitive to Hesperadin. On the basis of Hesperadin sensitivity they propose that INCENP S446ph is regulated by Aurora B and potentially a direct substrate (Hegemann et al., 2011). Although Hesperadin is known to have strong inhibitory effect against Aurora kinases it is a very promiscuous inhibitor (Bamborough et al., 2008); the sequence around INCENP S446 is also not similar to known Aurora phosphosites. We therefore decided to screen INCENP S446ph with our KiPIK methodology to see if a direct kinase could be determined.

Importantly, a phospho-specific antibody to INCENP S446ph was available, which would make *in vivo* validation and functional analysis more straightforward.

5.1.2 Aims

1. Use KiPIK screening to identify the kinase responsible for INCENP S446 phosphorylation

5.2 Results

5.2.1 INCENP S446 phosphorylation does not require Aurora B kinase activity

As INCENP S446ph in cells had been reported to be sensitive to 100 nM Hesperadin (a potent inhibitor of Aurora B), it has been proposed that it is dependent on Aurora B activity and is potentially a direct substrate (Hegemann et al., 2011). To explore this possibility, we acquired a phospho-specific antibody against INCENP S446ph (a gift from Jan-Michael Peters, IMP, Vienna) and examined whether it was possible to generate INCENP S446ph in the absence of Aurora B activity *in vivo*.

Staining Hela cells with the antibody revealed an exclusively mitotic signal and clear co-localisation with Aurora B (Figure 5.1). To assess whether the signal was dependent on Aurora B we treated asynchronous cells for 20 minutes with ZM447439 (a potent Aurora B inhibitor, which has been profiled on numerous protein kinases and has good selectivity (LINCS 2018) (Ditchfield et al., 2003). Cells were costained for H3S10ph (a well-known Aurora B substrate), Aurora B and INCENP S446ph. We saw no evidence of a reduced INCENP S446ph signal, but, as expected, there was a substantial reduction in H3S10ph intensity in mitotic cells. Strikingly, we were able to observe multiple prophase cells with unimpaired INCENP S446ph signal but a complete absence of detectable H3S10ph (Figure 5.1B). These cells had almost certainly entered mitosis in the presence of inhibited Aurora B, yet had phosphorylated INCENP S446. We therefore concluded that Aurora B was not required for INCENP S446ph.

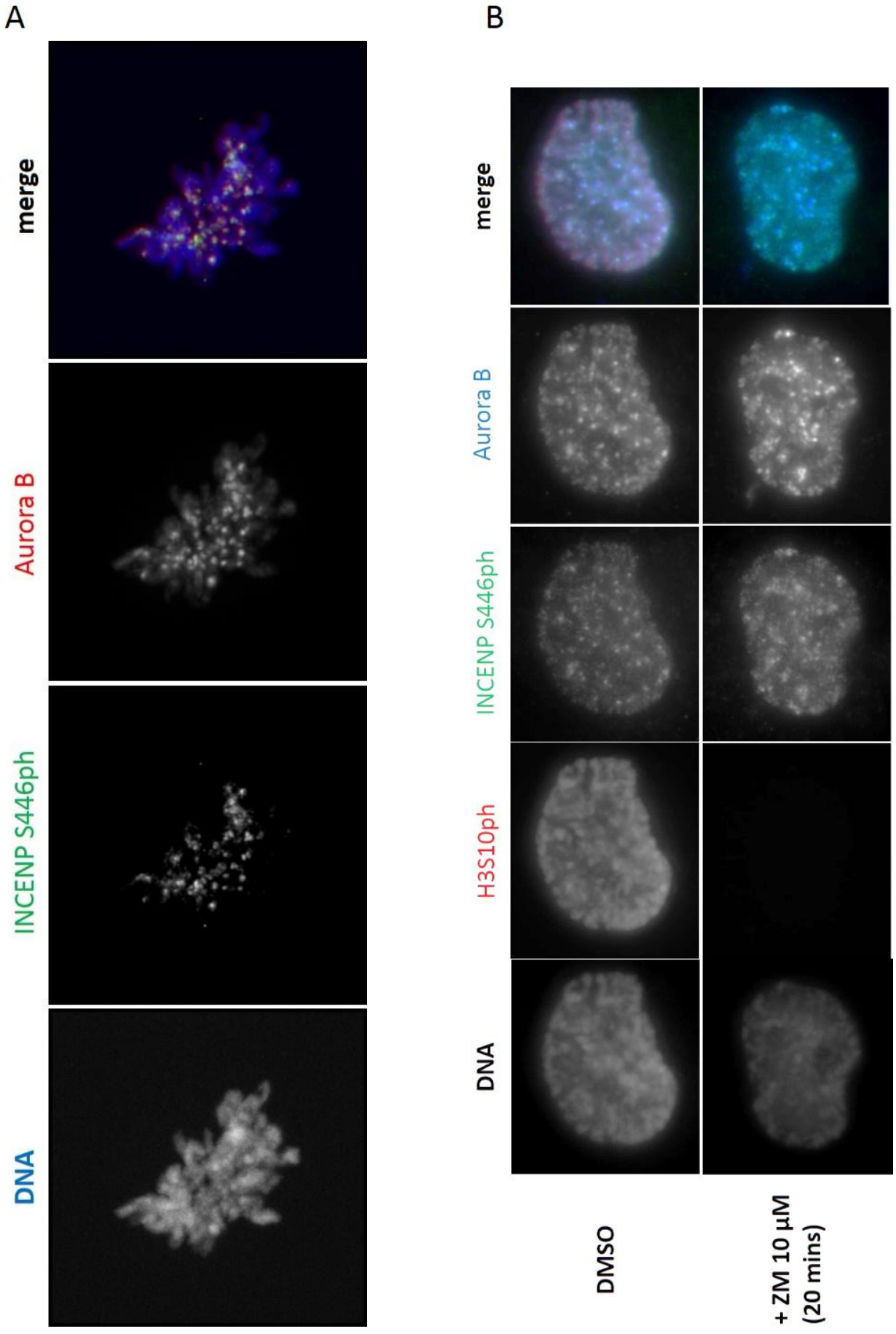

A
merge
Aurora B
INCENP S446ph
DNA
B
merge
Aurora B
INCENP S446ph
H3S10ph
DNA
DMSO
+ ZM 10 µM
(20 mins)

Figure 5.1: **INCENP S446ph does not require Aurora B kinase activity. A** Representative image of a mitotic HeLa cell DAPI stained and immunofluorescently stained with INCENP S446ph antibody and anti-Aurora B. **B** HeLa cells were untreated or treated with ZM447439 for 20 minutes (10 µM). Representative prophase cells from each treatment DAPI stained and immunofluorescently stained with INCENP S446ph antibody, anti-Aurora B and H3S10ph antibodies

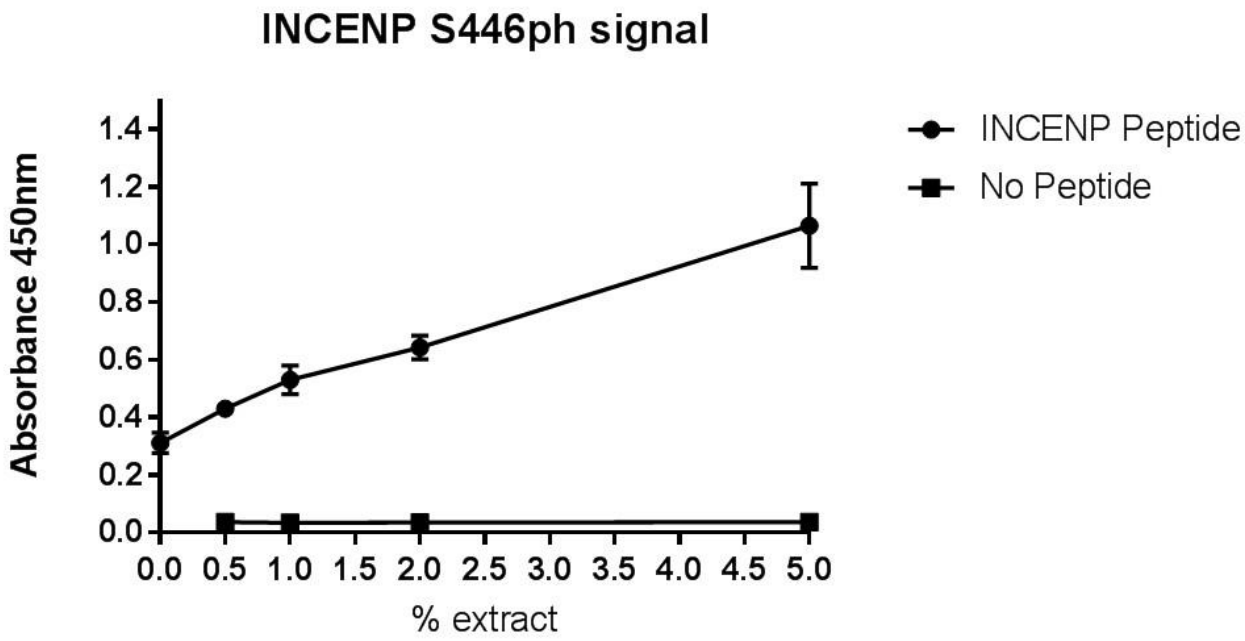
5.2.2 KiPIK screening identifies Cyclin B/CDK1 as the INCENP S446ph kinase

In order to identify the kinase responsible for INCENP S446 phosphorylation, we performed a KiPIK screen on a short unphosphorylated peptide corresponding to residues 438-453 of INCENP.

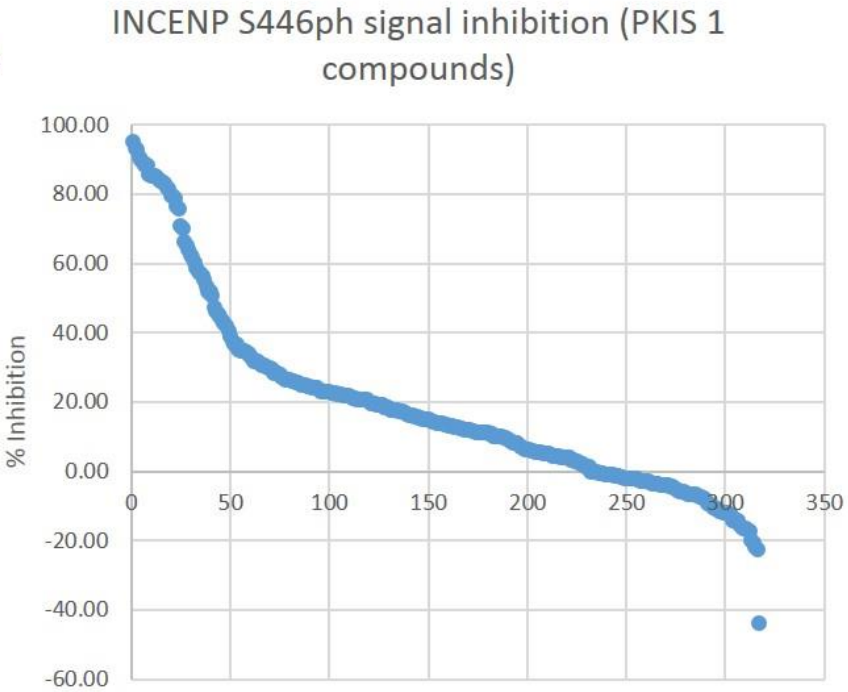
We used mitotic cell extract as a source of kinase activity and the INCENP S446ph antibody for the detection of phosphorylation by ELISA. An initial kinase reaction calibration was carried out by titrating increasing concentrations of mitotic HeLa cell extract in the presence or absence of the peptide (note antibody was also titrated, the concentration with best signal-to-noise is presented). We observed a strong peptide dependent signal that increased with increasing extract, indicating extensive de novo phosphorylation of the peptide.

After establishing that we were able to generate and detect *de novo* phosphorylation of the peptide, we proceeded with a KiPIK screen using the PKIS1 inhibitor set. A subset of the inhibitors was able to inhibit the phosphorylation potently (Figure 5.2B). Pearson's correlation of our INCENP S446ph signal inhibition profile with the inhibition profiles of kinases in the Nanosyn and DSF datasets indicated a very strong correlation with CDKs, with CDK1/cyclin B scoring the strongest correlation at r = 0.805 (Figure 5.2C).

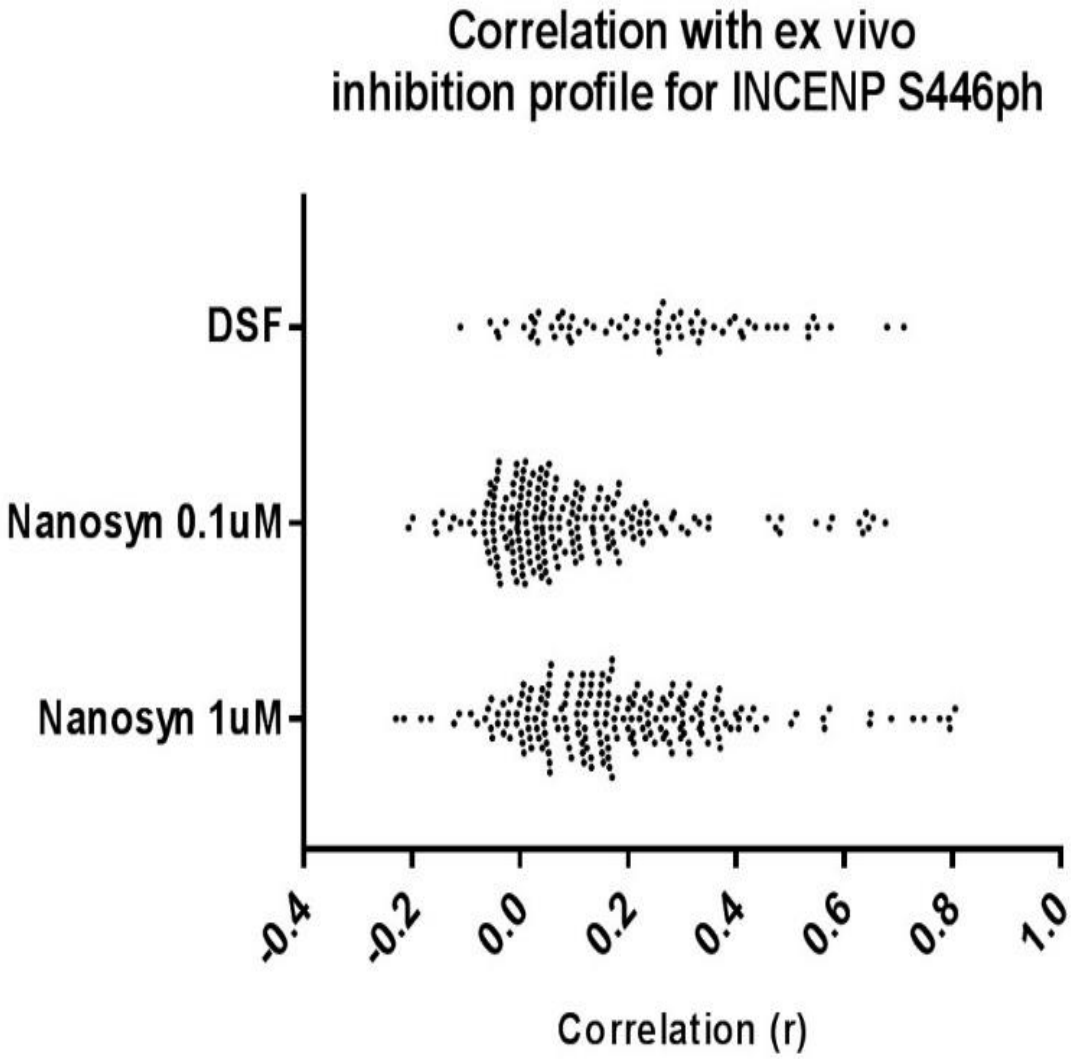
A


INCENP S446ph signal
Absorbance 450nm
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
% extract
INCENP Peptide
No Peptide

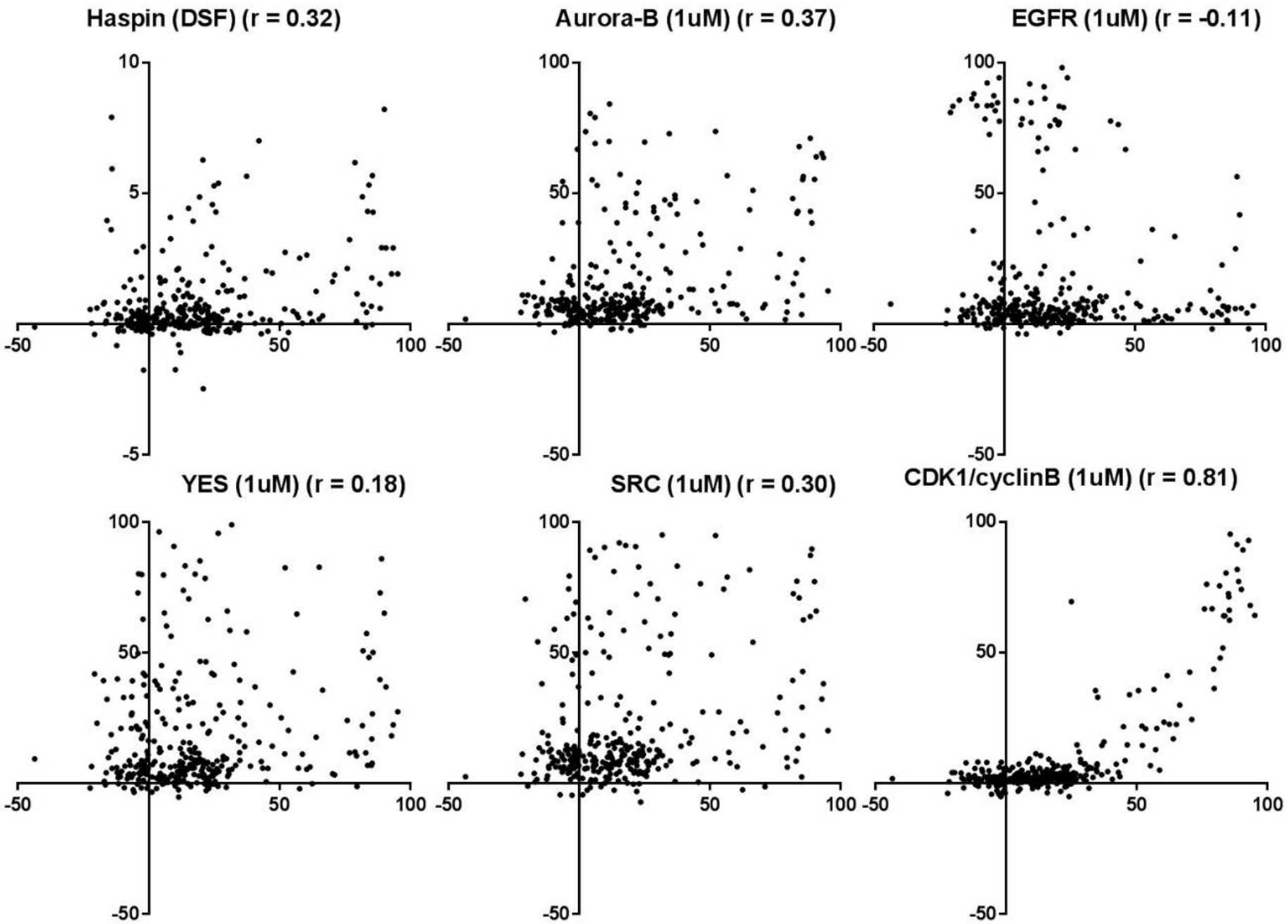
B


INCENP S446ph signal inhibition (PKIS 1 compounds)
% Inhibition
100.00
80.00
60.00
40.00
20.00
0.00
-20.00
-40.00
-60.00
0
50
100
150
200
250
300
350

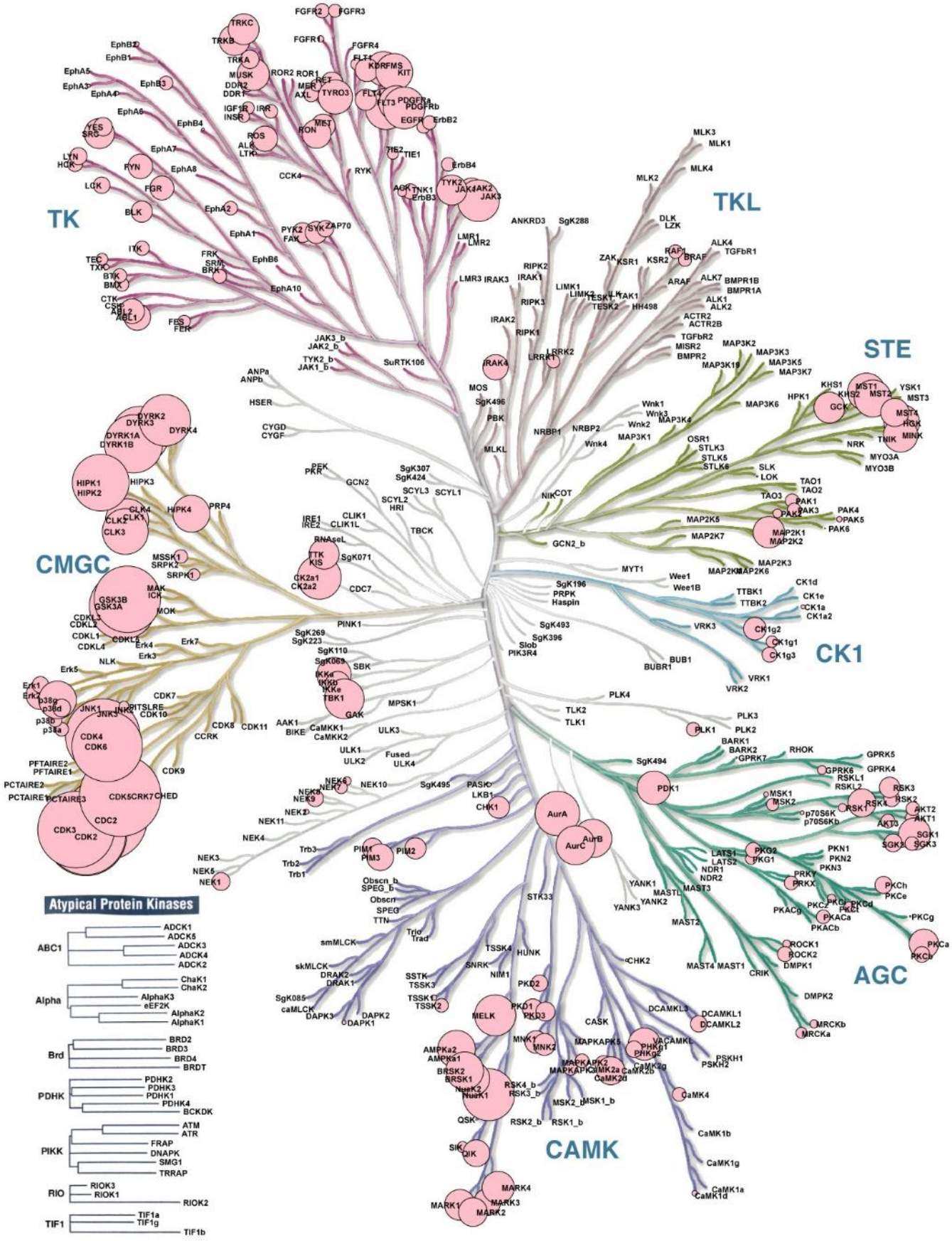
C



Nanosyn 1uM		Nanosyn 0.1uM		DSF	
CDK1/cyclinB	0.805	CDK2/cyclinA	0.676	CDK9	0.711
CDK2/cyclinA	0.795	CDK1/cyclinB	0.653	CDK2	0.679
CDK5/p35	0.793	CDK5/p35	0.645	STK16	0.575
CDK2/cyclinE	0.777	GSK3B	0.640	STK17A	0.550
CDK3/cyclinE	0.748	CDK2/cyclinE	0.634	PCTK2	0.543
CDK4/cyclinD	0.727	GSK3A	0.628	CLK4	0.533
CDK6/cyclinD3	0.687	CDK3/cyclinE	0.577	CDKL1	0.533
GSK3A	0.649	CDK4/cyclinD	0.572	STK17B	0.492
GSK3B	0.647	CDK6/cyclinD3	0.548	PRKAA2	0.474
DYRK1B	0.573	DYRK1A	0.483	AAK1	0.458
DYRK1A	0.563	DYRK1B	0.481	DYRK2	0.435
HIPK1	0.561	HIPK1	0.473	TLK1	0.423
ARK5	0.511	ARK5	0.460	TYK2	0.412
DYRK2	0.502	PDGFRb	0.349	PRKAA1	0.408
CLK3	0.455	CLK2	0.349	DYRK1A	0.398
AURORA-A	0.437	MST1	0.331	CLK3	0.388
CK2	0.432	DYRK2	0.323	MST4	0.375
CLK2	0.424	MELK	0.311	PDPK1	0.359
JAK3	0.409	KIT	0.299	CAMK2A	0.340
FLT3	0.405	MST2	0.285	CAMKK2	0.336


Haspin (DSF) (r = 0.32)
Aurora-B (1uM) (r = 0.37)
EGFR (1uM) (r = -0.11)
YES (1uM) (r = 0.18)
SRC (1uM) (r = 0.30)
CDK1/cyclinB (1uM) (r = 0.81)
10
5
-5
100
50
-50
-50
50
100

E



"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"

Figure 5.2: **INCENP S446ph KiPIK screen with PKIS1 compounds** : **A** KiPIK calibration of peptides corresponding to residues 438-453 of INCENP using mitotic cell extract. INCENP S446ph antibody used for detection **B** % inhibition of signal by PKIS1 compounds following synchronous kinase reactions **C** Pearson correlation of recombinant kinase *in vitro* inhibition profiles with INCENP S446ph lysate signal inhibition profile (PKIS1 compounds), (right) top 20 correlating kinases from 3 datasets of inhibition profiles **D** Scatter plots of PKIS1 compounds. (x-axis) INCENP S446ph signal inhibition in lysate reactions, against (y-axis) inhibition of recombinant kinase *in vitro* (from inhibitor profiling dataset: Nanosyn 1 µM, except Haspin from DSF 1 µM dataset) **E** Pearson correlation coefficients (as in **C**), for kinases from Nanosyn 1 µM dataset, mapped onto phylogenetic tree of the human kinome. Circle size corresponds to correlation coefficient score.

5.2.3 CDK1/Cyclin B phosphorylates INCENP S446 in-vitro

To verify the findings of the KiPIK screen we next tested whether CDK1/Cyclin B could phosphorylate INCENP S446 *in vitro*.

CDK1/Cyclin B and Aurora B were both tested for ability to phosphorylate INCENP S446. *In vitro* kinase assays were conducted on a mixture of 3 fragments of recombinant INCENP (see figure legend), only one of which covered the region containing S446 (at 68 kDa). The results of the reactions were detected by western blotting with the INCENP S446ph antibody and revealed clear phosphorylation of the S446 containing fragment by CDK1/Cyclin B. Aurora B, in contrast, appeared to result in little if any phosphorylation of S446 *in vitro* (Figure 5.3).

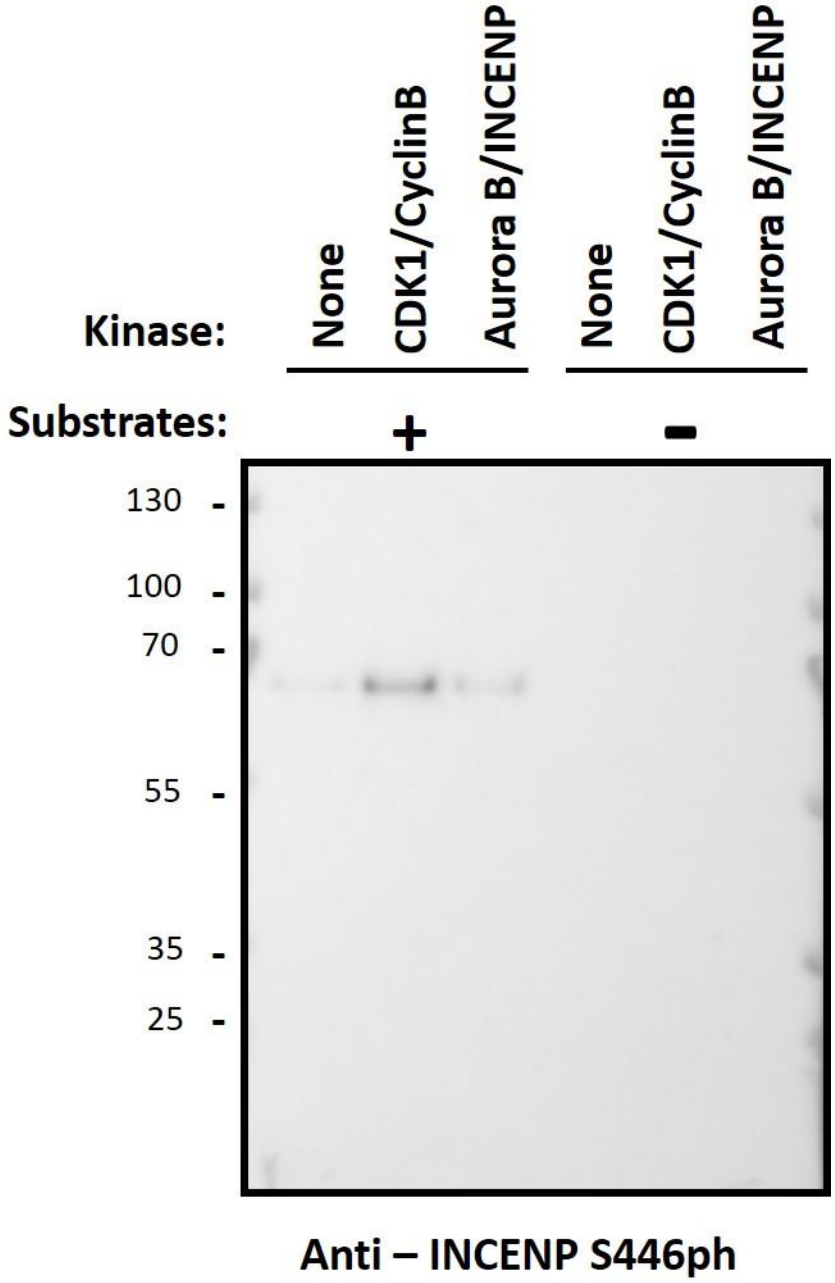


Figure 5.3: **CDK1/Cyclin B phosphorylates INCENP S446 *in vitro*.** *In vitro* kinase reactions were carried out using recombinant CDK1/Cyclin B or Aurora B/INCENP (0.05µg kinase) on a mixture of recombinant INCENP fragments (0.25µg each substrate). INCENP 369-583 (68 kDa), INCENP 1-405 (45 kDa), INCENP 826-919 (35 kDa). Reactions were stopped after 30 minutes and signal detected by immunoblotting with INCENP S446ph antibody

5.2.4 INCENP S446ph is lost within minutes of acute CDK1/Cyclin B inhibition in mitotically arrested cells

We next sought to confirm that INCENP S446ph is phosphorylated by CDK1/Cyclin B in cells. CDK1 substrates are difficult to demonstrate *in vivo* because disrupting CDK1 activity entirely both prevents cells from entering mitosis and leads to rapid exit from mitosis (Vassilev et al., 2006). For this reason, identifying direct CDK1/Cyclin B substrates from substrates that depend on CDK1/Cyclin B indirectly (i.e. essentially all phosphorylation events that depend on cells being in mitosis) is challenging.

To limit the problem of mitotic exit occurring upon CDK1/Cyclin B inhibition, we examined whether INCENP S446ph was susceptible to acute inhibition of CDK1/Cyclin B in cells that were already in mitosis. As an additional control, we also examined phosphorylation levels of the C-terminal TSS motif of INCENP concurrently to further ensure any response to cyclin B/ CDK1 inhibition was specific to S446ph, and not an indirect effect on INCENP phosphorylation levels broadly. (INCENP TSSph antibody was a gift from Michael Lampson, as described in (Salimian et al., 2011), the two serines of this motif are Aurora B substrates).

Briefly, cells grown in 96 well plates were treated with MG132 for 1 hour. This was followed by inhibitor treatment for either 0, 3, 6 or 9 minutes, after which cells were fixed in formaldehyde and immunofluorescently stained. For detection we used high-content imaging, collecting fluorescence intensity measurements for INCENP S446ph / INCENP TSSph, Aurora B and H3S10ph in >250 mitotic cells for each time point and condition. Mitotic exit that might be induced by the inhibitors was further controlled by using positive H3S10ph (commonly used as a mitotic marker) as a filter to ensure the cells we were selecting for intensity measurement were still mitotic.

INCENP S446ph was acutely inhibited by RO-3306, a potent CDK1/Cyclin B inhibitor, and reduced to less than 50% of initial intensity levels after 6 minutes. Roscovitine, which also inhibits CDK1/Cyclin B, had a less potent effect but S446ph

levels were reduced to 68% after 9 minutes. ZM447439 appeared to cause an increase in S446ph intensity although this was accompanied by a small increase in Aurora B intensity and is therefore possibly the result of a spreading of the Aurora B immunofluorescence signal rather than increased phosphorylation (in the absence of Aurora B activity the CPC has a diffuse localisation across all of the chromatin rather than a predominantly centromeric localisation (Wang et al., 2011)).

INCENP TSSph responded quite differently, exhibiting no loss in intensity due to either CDK1/Cyclin B inhibitor (RO-3306 or Roscovitine). In contrast Aurora B inhibition by ZM447439 caused a dramatic drop to around 25% of initial intensity after 3 minutes treatment (in line with the reported direct role of Aurora B in phosphorylating this site (Bishop and Schumacher, 2002)).

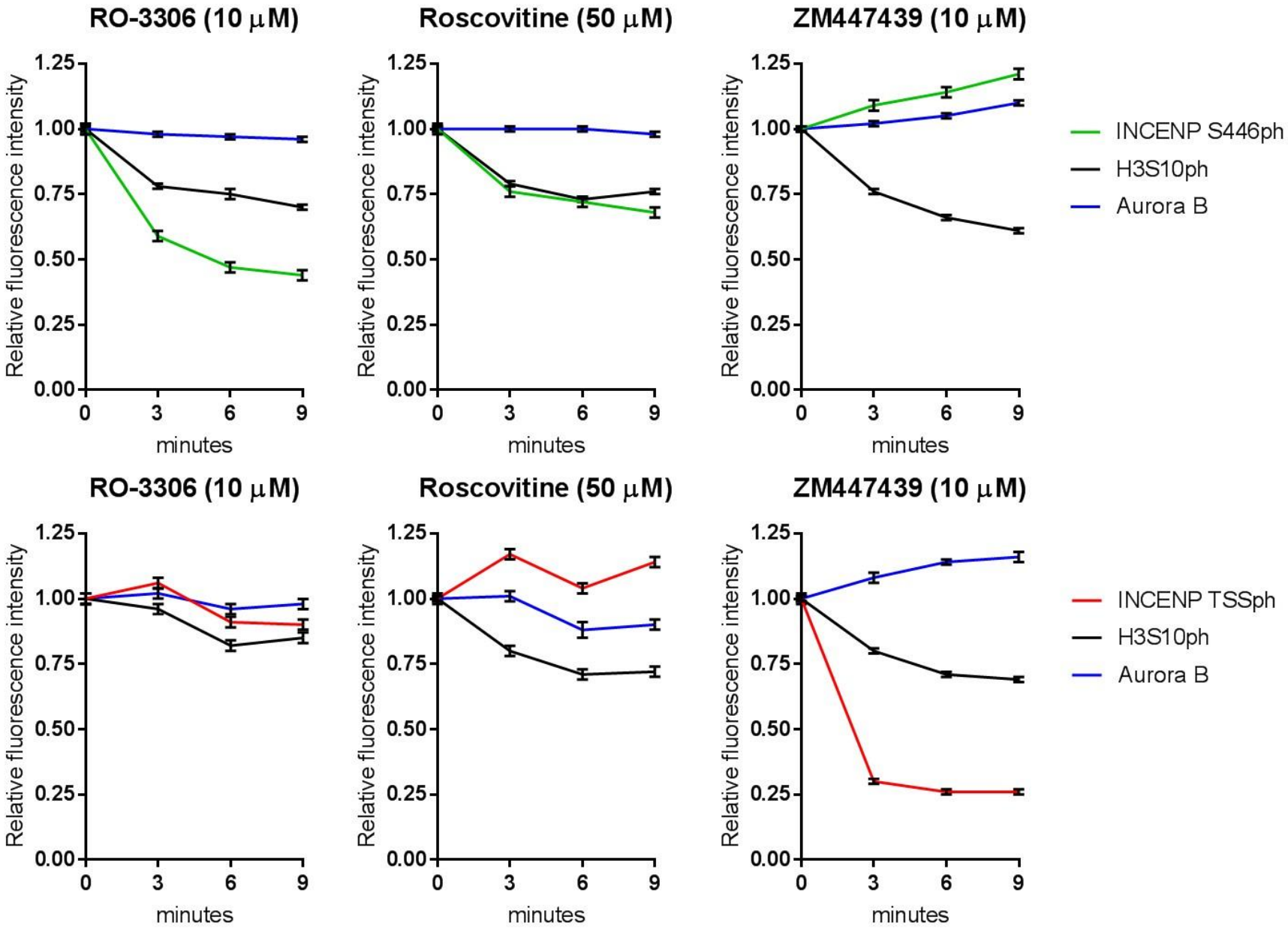

RO-3306 (10 μM)
Roscovitine (50 μM)
ZM447439 (10 μM)
Relative fluorescence intensity
1.25
1.00
0.75
0.50
0.25
0.00
0
3
6
9
minutes
INCENP S446ph
H3S10ph
Aurora B
RO-3306 (10 μM)
Roscovitine (50 μM)
ZM447439 (10 μM)
Relative fluorescence intensity
minutes
INCENP TSSph
H3S10ph
Aurora B

Figure 5.4: **Acute treatment with CDK1 inhibitors reduces INCENP S446 phosphorylation in cells.** HeLa cells were arrested in MG132 for 1 hour. Then RO-3306, Roscovitine, or ZM447439 was added for 0 (not added), 3, 6, or 9 minutes. Cells were fixed and immunofluorescently stained for Aurora B, H3S10ph and INCENP S446ph or INCENP TSSph. Mean fluorescent intensities of H3S10ph positive mitotic cells were collected for approx 250 cells per treatment. Error bars: standard error of mean

5.3 Discussion

5.3.1 Identification of CDK1/Cyclin B as the direct kinase for INCENP S446 by KiPIK screening

Here we report using KiPIK screening to identify the direct kinase responsible for INCENP S446ph in cells. The correlation scores following KiPIK screening with the PKIS1 inhibitor library clearly identified the CDK family of kinases as responsible for phosphorylation in our lysate reactions. CDK1/cyclin B had the strongest correlation, in agreement with the observation that INCENP S446ph is enriched in mitosis.

To confirm the results of the KiPIK screen we assessed *in vitro* and in cells whether CDK1/cyclin B could be responsible for INCENP S446 phosphorylation. In agreement with the screen, INCENP S446 was readily phosphorylated *in vitro* with purified CDK1/cyclin B. Furthermore, we demonstrated in cells that pre-established INCENP S446ph was acutely sensitive to CDK1 inhibitors, diminishing by more than 50% after 6 minutes of RO-3306 treatment.

In further support of a direct role for CDK1/cyclin B in INCENP S446 phosphorylation, it was recently shown that CDK1/cyclin B phosphorylates a number of non-canonical (non-Ser/Thr-Pro) motifs (Suzuki et al., 2015). The peptide sequence around S446 corresponds exactly to this alternative CDK1/cyclin B motif: P-x-S-x-R/K-R/K-R\K-R\K. Using Arg/Lys scanning peptide libraries, Suzuki et al. (2015) report that this motif is even more efficiently phosphorylated than a x-S-P-x containing peptide mixture (Suzuki et al., 2015).

We have therefore demonstrated the effectiveness of KiPIK screening in identifying the kinase responsible for an ‘orphan’ phosphorylation site.

The identification of this non-canonical CDK1/cyclin B mediated phosphorylation event in mitosis underlines the power of KiPIK screening as an unbiased approach for uncovering novel biology. We subsequently tested 8 *in silico* prediction methods to assess their ability to predict CDK1/cyclin B mediated regulation of this event. Only Scansite (which can only make predictions for 31 kinases) and NetworKIN 3.0 (which integrates information from the STRING database) indicated CDK1 as the most likely kinase (see Appendix D).

Chapter 6. Discussion

6.1 Purpose of the study

Our motivation for carrying out this research stems from the current methodological challenge in identifying the kinase directly responsible for a phosphorylation event of interest. Indeed, we are not aware of a method that has found widespread use that allows researchers to selectively identify direct upstream kinases for any substrate. As regulation by phosphorylation plays a fundamental role in cell biology, a methodological advance addressing this difficulty would be of great benefit.

In exploring this problem, we began with a study assessing the utility of siRNA kinome screening for identifying upstream kinases (detailed in chapter 3). While we had some success with this approach, limitations associated with indirect effects were also apparent (discussed further in section 6.2).

In chapters 4 and 5, we detail the development of a novel methodology for identifying direct kinases for specific phosphorylation sites. We validated this method on diverse known kinase-substrate pairs and used it to identify the kinase for an as yet unassigned phosphorylation site on INCENP. The approach, which we have termed KiPIK screening, represents a technically straightforward and broadly applicable strategy to identify kinases responsible for phosphorylation events of interest. In addressing the methodological insufficiency in this fundamental area it has the potential to benefit research widely.

In this chapter we begin with a review of the findings from chapter 3 and discuss how these relate to the strengths and limitations inherent in genetic screens in intact cells (for upstream kinases; section 6.2).

We then go on to discuss the results of chapters 4 and 5 and what these can tell us about the KiPIK screening method (section 6.3), This is followed by a conceptual discussion of the factors that contribute to the effectiveness of KiPIK screening in light of the determinants of kinase–substrate specificity in a cell (section 6.4). The

strengths and limitations are discussed (section 6.5), followed by propositions for the further development of KiPIK screening (section 6.6).

6.2 Kinome-wide siRNA screening as a method for identifying mitotic histone kinases

6.2.1 Overview

In chapter 3, we carried out kinome wide siRNA screens to identify upstream kinases for 4 mitotic histone phosphorylation marks, H3T3ph, H3S10ph, H2BS6ph and H3T11ph.

Screen results

The H3T3ph screen performed well, clearly identifying Haspin (GSG2) as the kinase of greatest importance in determining H3T3 phosphorylation levels, in line with its role as the direct H3T3ph kinase in mitosis (Dai et al., 2005). The screen also provided useful information on the indirect regulation of H3T3ph in mitosis. Aurora B was highlighted, which has important roles in activating Haspin and preserving H3T3ph from dephosphorylation (Qian et al., 2013). The screen also indicated an increase in H3T3ph levels when BUBR1 levels were reduced. We propose that this could be due to a proportionate increase in newer mitotic cells in our experimental setup, preventing the proteasome dependent decline in H3T3ph we observed over an extended mitosis (which probably occurs as a result of Cyclin B destruction (Brito and Rieder, 2006)). Alternatively, it could be related to BUBR1's role in the recruitment of mitotic phosphatases (Suijkerbuijk et al., 2012) (as discussed in 3.3.1).

In contrast, the effect of DDR2 on H3T3ph was an unexpected. Follow up experiments confirmed a modest effect of DDR2 knockdown on H3T3ph levels in mitosis. Interestingly the relative effect size of DDR2 knockdown appeared enhanced when cells were grown on collagen, the DDR2 receptor ligand. However, we also observed a similar decrease in H3S10ph in this condition, possibly suggesting that the enhanced effect of DDR2 knockdown on both these mitotic marks could be due to a broader role for DDR2 in regulating proliferation generally (Labrador et al., 2001, Olaso et al., 2002, Marquez and Olaso, 2014). It would be

interesting to carry out further immunofluorescence experiments to examine the impact of collagen and DDR2 knockdown on H3T3ph levels in individual mitotic cells.

The H2BS6ph screen prominently highlighted Aurora B as necessary for normal phosphorylation levels. This led us to establish that Aurora B has a crucial but indirect role in facilitating H2BS6ph through the inhibition of PP1 (Seibert et al. 2019). BUBR1 was also highlighted. As in the H3T3ph and H3T11ph screens, we think this may be attributable to a shortening of the arrest time in mitosis of the cells we measured (see section 3.3.2). Further work from our collaborators identified CDK1/Cyclin B as the likely direct kinase for H2BS6ph. In our screen, CDK1 (CDC2) was ranked #14 in kinases whose reduction caused diminished H2BS6ph; the indirect effects mediated by Aurora B and BUBR1 were significantly more prominent (discussed further in section 6.2.2.).

Our screen for H3S10ph regulating kinases was less informative. Knockdown of Aurora B, well established as the direct kinase for H3S10ph in mitosis (Crosio et al., 2002, Fischle et al., 2005), caused only a 1.4 standard deviation reduction in H3S10ph, and the effect size of all kinases was minimal. H3S10 is a very robust substrate of Aurora B and has proved difficult to reduce by Aurora B siRNA in our hands and others (Girdler et al., 2006). Work from Xu et al. (2010) indicates that different Aurora B substrates require different levels of Aurora B activity for robust phosphorylation to occur (Xu et al., 2010), potentially accounting for our ability to detect more pronounced effects from Aurora B knockdown for H2BS6 and H3T3. However, it also seems likely that the methodological alteration in mitotic arrest agent we made between screen 1 (H3T3ph and H2BS6ph, nocodazole) and screen 2 (H3T11ph and H3S10ph, taxol) was a mistake and contributed to the small effect on H3S10ph we observed from Aurora B knockdown. Ditchfield et al. (2003) demonstrate that potent inhibition of Aurora B (with ZM447439) quickly compromises a taxol mediated mitotic arrest by preventing error correction and allowing the SAC to become satisfied, leading to exit from mitosis (a nocodazole arrest is much more resistant) (Ditchfield et al., 2003). Therefore, more potently Aurora B depleted cells in our siRNA screen would have quickly exited mitosis in our taxol arrest (in screen 2). Consequently the cells we measured H3S10ph and

H3T11ph intensities for were biased to those with less efficient Aurora B depletion, likely significantly accounting for the apparent low effect of Aurora B knockdown on H3S10ph. Interestingly, the number of mitotic cells we recorded in Aurora B knocked down wells was 97% of the screen mean, suggesting that perhaps only the most potently Aurora B depleted cells escaped our taxol arrest, and that mitotic exit caused by Aurora B knockdown was not widespread in our screen.

The results of our H3T11ph screen appear dominated by kinases with indirect roles in regulating the phosphorylation site. Indeed, the most prominent appear attributable to the unusual dependence of H3T11ph levels on time spent in mitotic arrest (or perhaps simply in mitosis). In support of this, BUBR1 and MPS1's effects on H3T11ph were rescued when mitotic arrest was independent of the SAC (which their knockdown would likely compromise). Similar to BUBR1 and MPS1, CDK1 reduction through siRNA could also be expected to reduce the length of a taxol dependent mitotic arrest. Rattani et al. (2014) report that CDK1 catalytic activity is necessary for SAC functioning (through an as of yet undetermined mechanism) (Rattani et al., 2014). In agreement with an indirect role for CDK1 in H3T11ph, we found that CDK1 was not responsible for H3T11 phosphorylation in lysates.

Kinases whose knockdown caused a prominent increase in H3T11ph can also be understood through this explanatory model of H3T11 phosphorylation. It has been established that an extended mitosis causes a gradual increase in apoptotic cells (increasing with time spent in mitosis), due to the gradual degradation of anti-apoptotic protein MCL-1 (Haschka et al., 2015, Topham and Taylor, 2013). In line with this, two of the most prominent kinases whose knockdown caused an increase in average H3T11ph intensity are regulators of apoptosis (ASK and STK17A). Their knockdown could likely suppress this apoptotic process allowing cells to survive a longer extended mitosis (and therefore accumulate greater H3T11ph levels).

Similarly, PLK1 knockdown induced upregulation of H3T11ph. This could be due to the prometaphase arrest that incomplete PLK1 inhibition reportedly causes (Petronczki et al., 2008). This Taxol-independent arrest would result in an extension of the mitotic arrest in our experiment (as arrest would begin prior to Taxol addition and therefore be in effect for a longer time period).

Identifying more direct regulators of H3T11ph from those that simply shorten or extend a Taxol induced mitotic arrest might be possible with a different experimental design (see 6.2.2).

General conclusions

siRNA screening for kinases involved in specific mitotic histone phosphorylation marks provided significant insight into their regulation (except perhaps for H3S10ph).

Although the H3T3ph screen clearly identified Haspin (the direct kinase) as centrally important, for the other screens performed indirect effects were dominant. Information as to the indirect regulation of a phosphorylation event can be useful as it can provide greater understanding of the roles and regulation of a specific phosphorylation event (as we saw with H2BS6ph). However, if indirect effects are too dominant they can prevent or obscure identification of direct regulators.

Although indirect regulation is an inherent feature of screens in intact cells, the relative strength of indirect (compared to direct) regulatory factors revealed by a genetic screen will clearly vary between specific phosphorylation events. It will be determined not just by the type of indirect regulation naturally governing a particular phosphorylation event in a cell, but also by how and which of these regulatory factors intersect with or are elicited by the experimental design of the screen.

With sufficient knowledge of the indirect regulatory factors impacting a particular phosphorylation event it might be possible in some cases to mitigate/control for these factors by changes in experimental design, with the aim increasing the strength and detectability of direct regulators.

6.2.2 Screen design alterations to mitigate indirect effects

Prominent indirect effects and were evident in our H2BS6ph and H3T11ph screens. Here we discuss potential reasons for this and possible adjustments that could be made to experimental design in light of our discoveries about the nature of their indirect regulation.

In the case of H2BS6ph, subsequent work from our collaborators leads us to believe that CDK1/Cyclin B is the direct kinase (Seibert et al. 2019). Because CDK1/Cyclin B is necessary for cells to enter and maintain a mitotic state, siRNA screening (or any genetic method) are poor tools for identifying mitotic phosphorylation events directly regulated by CDK1/CyclinB activity. In all of our siRNA screens, the CDK1 siRNA knocked down cells we measured were necessarily those that incurred an incomplete knockdown of CDK1. Cells in which CDK1 was knocked down very effectively would have been unable to enter or maintain mitosis, and therefore were not measured.

In the H3T11ph screen we were unable to identify the direct kinase, and its identity remains unknown. Unlike for H2BS6ph, it is unclear why the direct kinase was not prominent in our screen results.

H3T11ph evidently develops relatively (and unusually) slowly in an extended mitosis (over a number of hours). The most straightforward explanation for this is that the activity level of the direct H3T11 kinase (on H3T11) is very low; either because it is in very low abundance, is very inactive, or that H3T11 is a very bad substrate for it. However, in either of these cases one would expect knockdown of the phosphorylating kinase to dramatically reduce H3T11ph levels.

Another explanation for this slow and linear development over time is almost perfectly balanced activity levels of the phosphorylating kinase and dephosphorylating phosphatase (slightly favouring the kinase). However, one would also expect in this case that knockdown of the phosphorylating kinase to dramatically reduce H3T11ph levels, especially as an unbalancing of these hypothetical kinase – phosphatase levels would be compounded by time (and our arrest time was 6 hours), which was not evident in our screen.

If these possibilities are rejected, another explanation is that the rate of H3T11 phosphorylation depends on another process (which occurs slowly over time) to a much greater extent than it does H3T11 direct kinase activity levels. Indeed, if H3T11 develops very slowly over time at a rate limited by an external other than H3T11 direct kinase activity rate then the H3T11ph kinase might only require very low levels of activity or abundance to keep up with this other rate limiting feature – it's siRNA knockdown might therefore not reduce the direct kinase levels below a level that allows it to 'keep up with' the other rate limiting factor.

We hypothesise that dephosphorylation of the neighbouring reside H3S10, could be this rate limiting factor. In our KiPIK investigations of H3T11ph, we determined that its phosphorylation by mitotic lysate was inversely correlated with Aurora kinase activity (Aurora B is the H3S10ph kinase). Moreover, pre-existing H3S10ph on a H3 peptide prevented H3T11 phosphorylation with mitotic lysate. Additionally, we observed in cells an increase in intensity and clear spreading along the chromosome arms of H3T11ph when Aurora B activity was inhibited with the inhibitor ZM-447439.

In view of the features of H3T11ph exposed by our investigations (development over time in extended mitosis, inverse correlation with Aurora in lysates and probable inhibition by H3S10ph), a further siRNA screen could be designed to mitigate the contribution of indirect effects and potentially expose the role of the direct H3T11ph kinase.

Because H3T11ph develops over an extended mitosis it is indirectly susceptible to siRNA inhibition that cuts short the length of mitotic arrest. The arrest induced by taxol is dependent on the SAC, so can be cut short by indirect effects which compromise SAC signalling. For this reason, taxol was a poor choice for mitotic arrest inducer in our screen.

To overcome this, we could use MG132 to induce mitotic arrest, as it does so in a SAC independent manner. Although MG132 would reduce the number of mitotic cells we accumulated (it prevents mitotic entry) it would also ensure that the cells

measured will have been in mitosis for a similar time period; potentially advantageous in normalizing H3T11ph levels in the cells measured.

The other proposed cause of indirect effects could potentially be mitigated by preventing or removing H3S10 phosphorylation, and therefore preventing the proposed dependence of H3T11 phosphorylation on H3S10 dephosphorylation. As siRNA seemed ineffective in knocking Aurora B down sufficiently to significantly reduce H3S10ph levels the Aurora B inhibitor ZM447439 could be used instead.

A screen could be carried out where, after a 48 hour siRNA kinome library treatment, ZM447439 is added in addition to MG132 (optimum timings to be determined experimentally). This would remove H3T11 phosphorylation dependence on the SAC and potentially expose the activity of the H3T11 direct kinase as (hopefully the most) prominent factor in the H3T11ph levels recorded.

6.3 The KiPIK screening method

Screen results

In chapters 4 and 5 we detailed the development of KiPIK screening and used it to identify direct kinases responsible for the phosphorylation of diverse sites in cells. We were able to identify the established direct kinases for H3T3ph, H3S28ph, EGFR Y1016ph and Integrin β1A Y795ph. We also tested 2 'orphan' phosphorylation sites, H3T11ph and INCENP S446ph. H3T11ph was inhibited by remarkably few small molecule inhibitors, possibly suggesting it is phosphorylated by an unusual or understudied kinase. Our KiPIK screen with the larger PKIS2 inhibitor set identified interesting candidate kinases which we are currently investigating for a role in H3T11ph. In contrast our KiPIK screen of INCENP S446ph very clearly identified CDK1/Cyclin B as the phosphorylating kinase in our extract reaction. Our follow up investigations in cells support the notion that CDK1/Cyclin B in the direct kinase for INCENP S446ph *in vivo*.

Resistance to indirect effects

We have designed KiPIK screening to diminish indirect effects and promote identification of the direct kinase responsible for phosphorylating the site of interest.

KiPIK screening incorporates several features which contribute to this aim. Firstly the use of small-molecule inhibitors allows the timeframe of inhibition to be kept short (unlike in genetic screens where long timeframes potentiate indirect effects). In addition, KiPIK screening makes use of lysates rather than intact cells. The loss of cellular context and increasing in dilution is very likely to disrupt and dampen cellular signalling networks and pathways which would give rise to indirect effects in cells. Furthermore lysates are prepared from cells where the phosphorylation activity of interest is robust, and therefore the kinase responsible is active. This active state is preserved from indirect effects associated with dephosphorylation (and then the requirement for reactivation by upstream kinases for example) by including phosphatase inhibitors at every stage of KiPIK screening, thereby "freezing" kinases in their active state. Moreover an inherent feature of our KiPIK screen correlation coefficients is that their resistance to indirect effects is determined by the relative effect size between signals, an indirect effect contributing to 20% of a phosphorylation signal is unlikely to 'throw off' a robust correlation score.

In the H3T3ph KiPIK screen, Aurora B, which is known to be an important upstream regulator in cells (Wang et al., 2011), was not highlighted (no Aurora kinase in the top 40 correlations) consistent with the notion that upstream regulatory effects are diminished and direct kinase identification favoured.

It was also notable in the INCENP S446ph screen that we were able to identify CDK1/Cyclin B with great prominence by KiPIK screening. As discussed CDK1/Cyclin B is essential to entry and maintenance of mitosis (Vassilev et al., 2006). This makes identification and separation of its effects on individual substrates from its essential roles in mitosis difficult by genetic screening methods.

A very specific type of indirect regulation which KiPIK is susceptible to was highlighted in the H3T11ph screen. It was clear from this screen that Aurora B activity on the neighbouring residue (H3S10) was indirectly impacting H3T11 phosphorylation. It is therefore worth considering the possible impact of post-translational modifications on adjacent residues when selecting substrates to probe with KiPIK screening.

6.4 Why does KiPIK screening work?

In this section we discuss important component parts of KiPIK screening and how these may contribute to the effectiveness of the technique.

For KiPIK screening to successfully identify the *in vivo* kinase for a specific phosphorylation site, two things are essential. Firstly, it must be able to successfully identify the kinase phosphorylating the substrate residue in question within the lysate reaction. Secondly, the kinase phosphorylating the substrate residue within the lysate reaction must be the same kinase responsible for phosphorylation of this site in cells.

It is clear, and perhaps unsurprising, that the inhibitor specificities collected by profiling on *in vitro* recombinant kinases are highly conserved in our *ex vivo* KiPIK reactions (which constitute whole cell lysate diluted in kinase buffer), for all known kinases tested; Haspin, Aurora B, EGFR and Src kinases. Consequently, as KiPIK screening demonstrates, correlating the inhibitory effect of all inhibitors in a panel functions as a powerful diagnostic approach for identifying a dominant kinase signal acting on a probed substrate.

Comparability both within and between the profiled dataset and our KiPIK reactions is essential for tight correlations to be possible. Great care has been taken by the developers of commercial recombinant kinase screening panels to ensure comparability in the published inhibitory datasets**.** In our KiPIK reactions tight comparability between reactions is achieved by using the same freshly thawed cell lysate and a liquid handling robot to carry out all reactions in parallel.

For the kinases identified by KiPIK to be the *in vivo* kinases the lysate reactions need to preserve enough of the factors which determine *in vivo* kinase-substrate specificities.

In the introductory chapter, we discussed the linear sequence around a phosphorylated residue and the regulation of local concentrations of kinases and

substrates as key determinants of kinase-substrate specificity in a cell. The peptide substrates we have been using thus far for KiPIK screening preserve the linear sequence of the substrate, but it is doubtful that the factors which regulate local concentrations of particular kinases and substrates in a cell are operating effectively in the homogenised diluted lysate of our *ex vivo* reactions. Furthermore, our short peptides are unlikely to incorporate all the features of the cognate protein which facilitate its protein-protein interactions in a cell.

However, while the potency of local concentration regulating specificity determinants is likely diminished, in providing the substrate peptide well in excess of physiological concentrations, it also seems likely that the necessity for them is diminished; the rate of interaction between kinases and the substrate will be high via diffusion.

Importantly, in addition to preserving the contribution of linear substrate sequence, KiPIK screening also maintains activation states of kinases characteristic of the cells from which the lysates were prepared. As discussed earlier, an important component of a KiPIK screen is preparing lysate from cells which have high levels of the phosphorylation being investigated. For example preparing mitotic cell lysate when investigation a mitotic phosphorylation event; or receptor stimulated lysate when investigating a phosphorylation event known to be downstream of a particular cell surface receptor.

In essence, the lysate reactions are essentially a competition of phosphorylation efficiency between all kinases in the lysate. If a kinase stands out above the rest in terms of efficiency with which it phosphorylates the query substrate it will be responsible for the bulk of the phosphorylation signal detected.

For this reason, one can also imagine KiPIK screening being resistant to small amounts of erroneous substrate phosphorylation by kinases which would not usually interact with the query substrate in a cell. Successful kinase identification by KiPIK screening may only require that the rate of phosphorylation being carried out on the substrate by the genuine *in vivo* kinase is greater than that of other kinases by a significant factor. To be identified, the genuine kinase would likely just have to

be responsible for a more significant bulk of substrate phosphorylation than the next most efficient kinases.

In conclusion, KiPIK in its present form essentially answers the question:
in this cell lysate, which kinase most efficiently phosphorylates a substrate with this linear sequence? Although this might not always be the same as the kinase responsible for phosphorylating a substrate in an intact cell, on the basis of the screens performed thus far it often is.

6.5 KiPIK screening in context

6.5.1 Advantages of KiPIK screening over currently available techniques

In the introductory chapter, we discussed the techniques that are currently available for the identification of upstream kinases for specific phosphorylation sites. Here we compare KiPIK screening to these techniques and discuss the advantages of this new methodology.

As discussed in the introduction (1.4.3), *in silico* approaches for kinase prediction base their predictions upon the short linear amino acid sequences surrounding phosphosites and therefore rest on the assumption that recognition of these sequences by the substrate binding site of kinases is the primary determinant of substrate specificity (as discussed in 1.3.2). KiPIK screening is similar in this regard as it also queries the ability of kinases to phosphorylate a particular substrate based primarily on the short linear sequence around a phosphorylation site (although as discussed in 6.6 there is no reason that recombinant proteins could not be used as substrates for KiPIK screening, thereby potentially integrating more in cell interaction determining factors). However, while most *in silico* approaches utilise primary sequence alone, KiPIK screening has the advantage that it also incorporates the contextual information of relative kinase activity levels of the cells from which the cell lysate is derived.

The major weakness of *in silico* approaches for kinase prediction is that they are highly dependent on the quality and quantity of training data available. The current lack of experimentally verified substrates for many kinases leads to major limitations

in terms of kinome coverage and accuracy for current computational approaches. A recent paper by Song et al. (2017) illustrates this point: for the training of their phosphorylation prediction algorithm they limited their coverage to kinases (or kinase families) for which at least 50 experimentally verified phosphorylation sites were available (in the PhohphoELM database), applying this filter allowed them to train models for only 9 kinases and 3 kinase families (Song et al., 2017). KiPIK screening in contrast is not reliant on the presence and quality of past data and can in principle identify any kinase for which sufficient profiling data is available (discussed in 6.5.2). When compared with a spectrum of *in silico* methods, including those that attempt to incorporate information on biological context, KiPIK was the most reliable approach for identifying the kinases responsible for the five substrates examined in here (see appendix D).

Screens in intact cells (such as siRNA screens, CRISPR/Cas9 or overexpression screens) allow complete kinome coverage and retain the regulatory features governing phosphorylation in the cell. However, the combination of the long treatment timeframes of genetic methods and intact cellular pathways and networks predisposes these approaches to indirect effects. Another concern, particularly with overexpression screens, is the potential for both direct and indirect phosphorylation artefacts resulting from artificially high (or low) levels of examined genes. Artificially high concentrations of a kinase are likely to cause less stringent substrate selection as well as over activation (or inhibition) of downstream kinases or processes. KiPIK screening conversely limits the possibility for indirect effects by acute inhibition timescales and disrupted cellular pathways and networks.

A further limitation of screens in intact cells that rely on genetic disruption of normal kinase expression levels is the requirement for efficient transfectability of the cells or tissues under study. KiPIK screening has the advantage that any cells or tissue can be used as a source of lysate.

The available biochemical methods we discussed in the introductory chapter (1.4) can be divided into 3 types: *in vitro* recombinant screens, co-purification strategies, and purification of catalytic activity by chromatography.

In vitro recombinant screens with purified components are not susceptible to indirect effects and in principle, like KiPIK screening, allow the efficiency of phosphorylation of a substrate to be compared between kinases. However abnormal substrate promiscuity is common in situations in which purified kinases are exposed to purified substrates (Peck, 2006, Cheng et al., 1993). Importantly, KiPIK screening has the major advantage that lysate kinases are present in abundancies and activation states reflective of the cells from which they are derived. It is therefore far more likely to reflect the kinase-substrate specificities found in a cell. The cost of preparing and screening panels of recombinant kinases can also be prohibitive, while KiPIK screening is very inexpensive.

Several methods have also been developed that attempt to stabilize transient kinase–substrate interactions to allow co-purification of substrates with upstream kinases (Dedigama-Arachchige and Pflum, 2016, Maly et al., 2004, Zeng et al., 2017) (Statsuk et al., 2008). Strategies involving crosslinker addition to ATP or to a substrate of interest have both been developed with the aim of capturing kinase-substrate interactions (Dedigama-Arachchige and Pflum, 2016, Maly et al., 2004, Statsuk et al., 2008). However, while these approaches had some success *in vitro,* in cells their specificity proved too low for viable non-biased kinase identification. Another approach to kinase-substrate co-purification utilises bimolecular fluorescence complementation (BiFC) (Zeng et al., 2017). However this strategy also has specificity concerns as it detects interactions per se rather than specifically those that result in the phosphorylation of a particular residue. It also requires overexpression of kinase and substrate (potentially causing artefacts), is protein- rather than phosphosite-specific, and requires subsequent mass spectrometry. KiPIK screening in contrast informs of catalytic events occurring on specific phosphorylation sites. Detection is straightforward and relative kinase abundancies are intact.

The purification of catalytic activity via chromatography has historically, and more recently, also been reported as a strategy for identifying the kinase of a specific phosphorylation event(Rubin and Rosen, 1975, Downward et al., 1984, Kubota et al., 2009, Ji et al., 2010). Similar to KiPIK screening, this approach uses cell lysate as a source of catalytic activity. However, this methodology requires an enormous

quantity of cells as starting material (7-8 x10^{10} cells), followed by multiple rounds of chromatographic purification, kinases assays, and finally mass spectrometric identification of isolated kinases. It is therefore poorly suited as a generic approach for kinase identification.

In contrast, KiPIK screening allows identification of a specific kinase activity in cell lysate without a requirement for purification. Typical screens can be performed with lysate from 7-8 x10^{7} or fewer cells and require one round of kinase reactions (rather than multiple). Consequently, it is far less labour intensive, and much more practical as a generic approach for kinase identification.

6.5.2 Limitations of KiPIK screening

The coverage and resolution across the kinome of a KiPIK screen is limited by the characteristics of the inhibitor set used and how much of the kinome this set has been profiled upon.

The extent of inhibitor profiling is limited by the number of kinases for which activity assays have been developed, a recent report identified commercially available assays for 436 kinases (Drewry et al., 2017). Several large libraries of commercially available compounds have been profiled on a large proportion of the kinome. The largest of these include PKIS1 (Elkins et al., 2016), in which 367 inhibitors were profiled on 196 unique protein kinases, and PKIS2, in which 645 inhibitors were profiled on 392 unique protein kinases (Drewry et al., 2017).Consequently with currently available compounds and profiling datasets KiPIK screening has coverage for nearly 80% of the human kinome.

Effective resolvability of kinases by KiPIK requires that the library of inhibitors being used has sufficient diversity such that there are distinguishable differences in the inhibition profile (or pattern) for each kinase across the set of inhibitors. Kinases from different families will generally have many small-molecule inhibitors that differentiate them. Potential issues with resolvability arise with kinases that are very phylogenetically similar, as there will generally be fewer inhibitors which inhibit them

to substantially varying degrees. In practice it is very unusual to find kinases that are not significantly resolved by some of the inhibitors in reasonably diverse sets of 100 or more inhibitors. One way of describing the resolution of an inhibitor set is by calculating the Euclidian distance between kinases across all inhibitors in a set. See appendix A for the Euclidian distances of kinases in PKIS1 (arranged by hierarchical clustering based on Euclidian distances).

Another potential weakness of KiPIK screening is that it measures reactions taking place in cell lysate rather than intact cells. The results we have generated so far indicate that KiPIK screening integrates sufficient specificity determining factors to reflect in-cell specificities (discussed in 6.4). However, there are several closely related kinases which have evolved differences primarily in their regulatory rather than catalytic domains. This likely limits the resolvability of these kinases by both inhibitor sets and intrinsically in lysate reactions due to loss of the determinants of sub cellular localisation which separate their functions in cells. As discussed in the introduction, Aurora A and B share highly conserved catalytic domains and divergent regulatory domains. The divergent regulatory domains mediate dramatic differences in subcellular localisation in a cell, which is thought to account for their distinct substrate selection (Carmena et al., 2009, Li et al., 2015). Although a technical error likely contributed to the poorly resolved correlation scores in this particular experiment, our H3S28ph screen results in chapter 4 reflect the difficulty of differentiating between the three Aurora kinases (); Aurora A, B and C all were top hits in this experiment. Nevertheless, the ability to unambiguously identify member(s) of the Aurora kinase family as responsible for H3S28ph is a significant step when there are more than 500 kinases in the human genome.

Although not intrinsic to the KiPIK method, antibody-based detection imparts further limitations to the methodology described here. Antibodies to specific phosphorylation sites are not always available, although as we demonstrated in chapter 4, generic (phospho-tyrosine) antibodies can be used. If generic antibodies are required it potentially limits the scope of the phosphorylation sites that can be screened. Flanking residues of the same type around a query phosphorylation site could result in mixed signals if these flanking sites are also phosphorylated. If

phosphorylation site specific antibodies are not available it also complicates in cell verification of the results of a screen.

6.6 Future development of KiPIK screening

There are several avenues which could be explored to increase the power and versatility of KiPIK screening.

Firstly, gaps in coverage and resolution could be filled with an expanded and optimised inhibitor set to allow complete kinome coverage by KiPIK screening. The recently described PKIS2 set goes a long way towards this goal. Moreover, the authors of this paper are in the process of compiling a 1000-1500 compound "comprehensive kinase chemogenomic set" (KCGS) which they also intend to profile on nearly 400 kinases. They intend to make these compounds freely available to the research community, suggesting near complete kinome coverage by KiPIK screening will be achievable in the near future (Drewry et al., 2017)

Another way in which KiPIK screening could be made more versatile would be through integrating and optimising mass spectrometry based detection into the methodology. This could allow any phosphorylation site to be screened by KiPIK, regardless of antibody availability or surrounding sequence suitability. One strategy for this would be to make use of the "absolute quantification" (AQUA) strategy described by Kirkpatrick et al (2005). This approach allows precise quantitation of phosphorylated peptides by mass spectrometry via spiking a peptide mixture (after the kinase reaction in the case of KiPIK) with a known concentration of a synthetic internal standard peptide (Kirkpatrick et al., 2005).

Another potential development to KiPIK screening would be to utilise recombinant whole proteins as substrates rather than short peptides. We have experimented with this recently and were able to conduct a KiPIK screen with a recombinant MEK1 protein for phosphorylation on S221. The assay is still being developed, and the concentration of recombinant MEK1 we used was clearly sub-optimal, but our preliminary screen identified 3 Raf proteins in the top 6 results, in line with the literature (Park, 2014) (data not shown). Utilising recombinant proteins rather than

peptides might improve the specificity of some KiPIK screens as whole proteins potentially incorporate more in-cell specificity determining factors.

If mass spectrometric detection and recombinant proteins are both integrated into the KiPIK methodology it may be possible to screen multiple phosphorylation sites simultaneously on a single recombinant protein.

KiPIK screening could also be further developed by experimenting with different types of cell extract. Stock batches of extract from different cell lines or tissues and/or treated with different compounds that arrest cells at particular stages or stimulate particular pathways could be prepared and tested on query substrates of interest prior to KiPIK screening. Lysates which produce the phosphorylation of interest with the best signal to noise could then be taken forward for screening. Similarly, sub-cellular fractionation could also be explored in order to separate different parts of the cell and potentially mimic the in-cell kinase-substrate accessibility determined by these factors.

If inhibition times were kept short (to minimize indirect effects), it might also be possible to perform KiPIK screens in intact cells using changes in endogenous protein phosphorylation as a readout. Immunofluorescent detection could be utilised as in the siRNA screens in chapter 3. Alternatively, quantitation by mass spectrometry might be possible. At the extremes of technical ambition, it is conceivable that a great number of endogenous phosphorylation sites could be screened simultaneously (with mass spectrometric detection). Calibrating the minimum acute inhibition time sufficient to cause robust changes in direct substrate phosphorylation with inhibitors across multiple representative phosphorylation sites would be a good starting point in this endeavour. However, any such acute in-cell approach would rely on endogenous phosphatase activity that acted more swiftly than indirect effects.

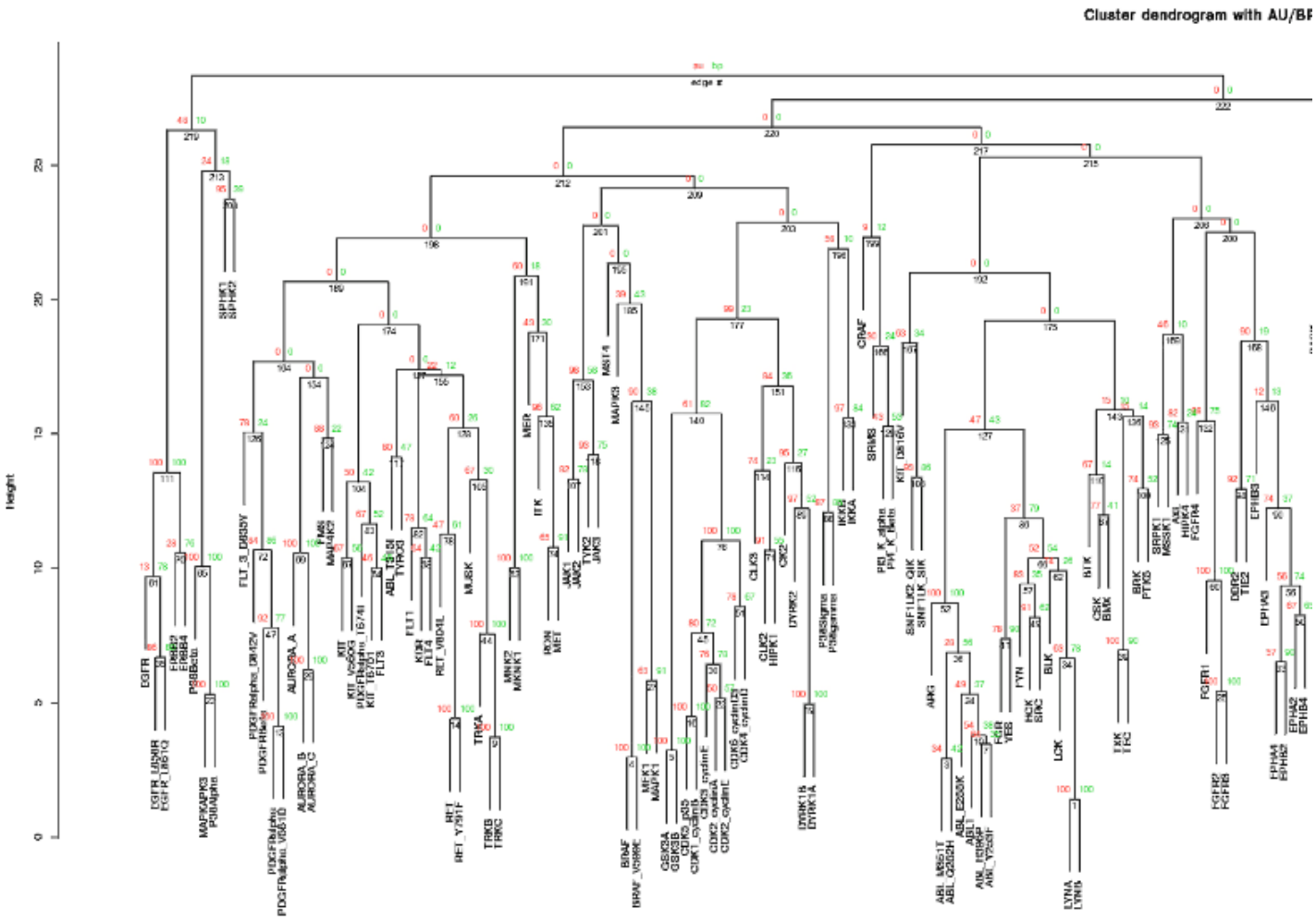
If in-cell KiPIK were attempted, and indirect effects were preventing clear signals being detected, some kind of *in silico* deconvolution might be feasible to separate mixed signals. A basic approach to probe for separable signals could be the systematic application of the inhibitor stripping described in 4.2.11.

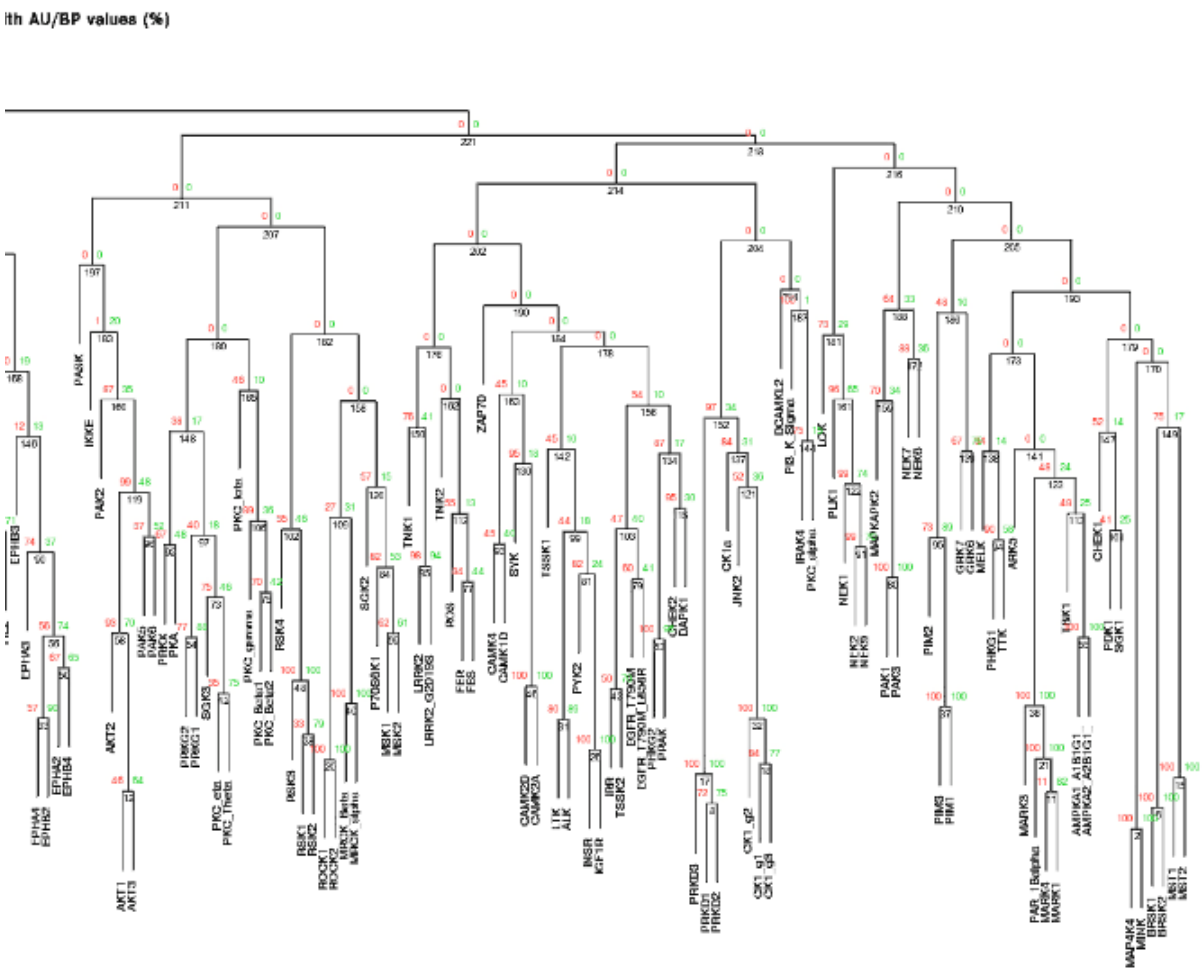
6.7 Further potential applications for KiPIK screening

An important feature of the KiPIK screening methodology is that the lysate used as a source of kinases can be derived from any cell line (or tissue). A report by Huttlin et al (2010) describes extensive tissue specific phosphorylation occurring in different mouse tissue types (Huttlin et al., 2010). KiPIK screening could have extensive application in exploring kinases responsible for important tissue specific phosphorylation events. There could be similar utility for KiPIK screening in developmental biology. Importantly, the catalytic domains of mouse kinases are 97% conserved with humans on average (Caenepeel et al., 2004) so it seems likely that a lot of inhibitor profiling data would be accurate for kinases across similar species.

Appendices

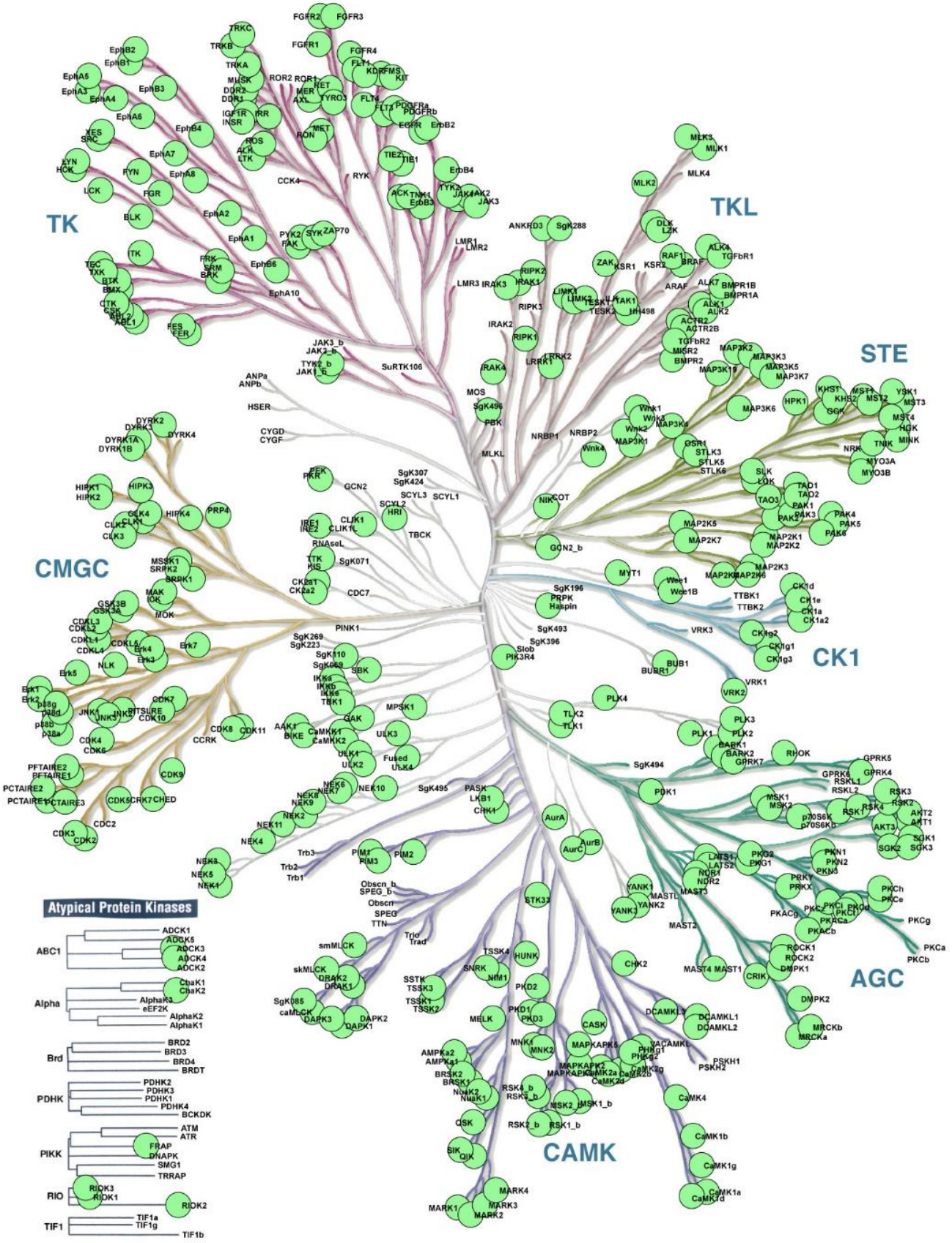
A





Appendix A: Hierarchical clustering of kinases using the Nanosyn 1 µM inhibition profiles for PKIS1 compounds. Green and red value, are bootstrap values signifying % of hierarchical clusters that cluster each branch as depicted. Black values are Euclidian distance between kinases (sum of difference of all PKIS1 compound inhibitions)

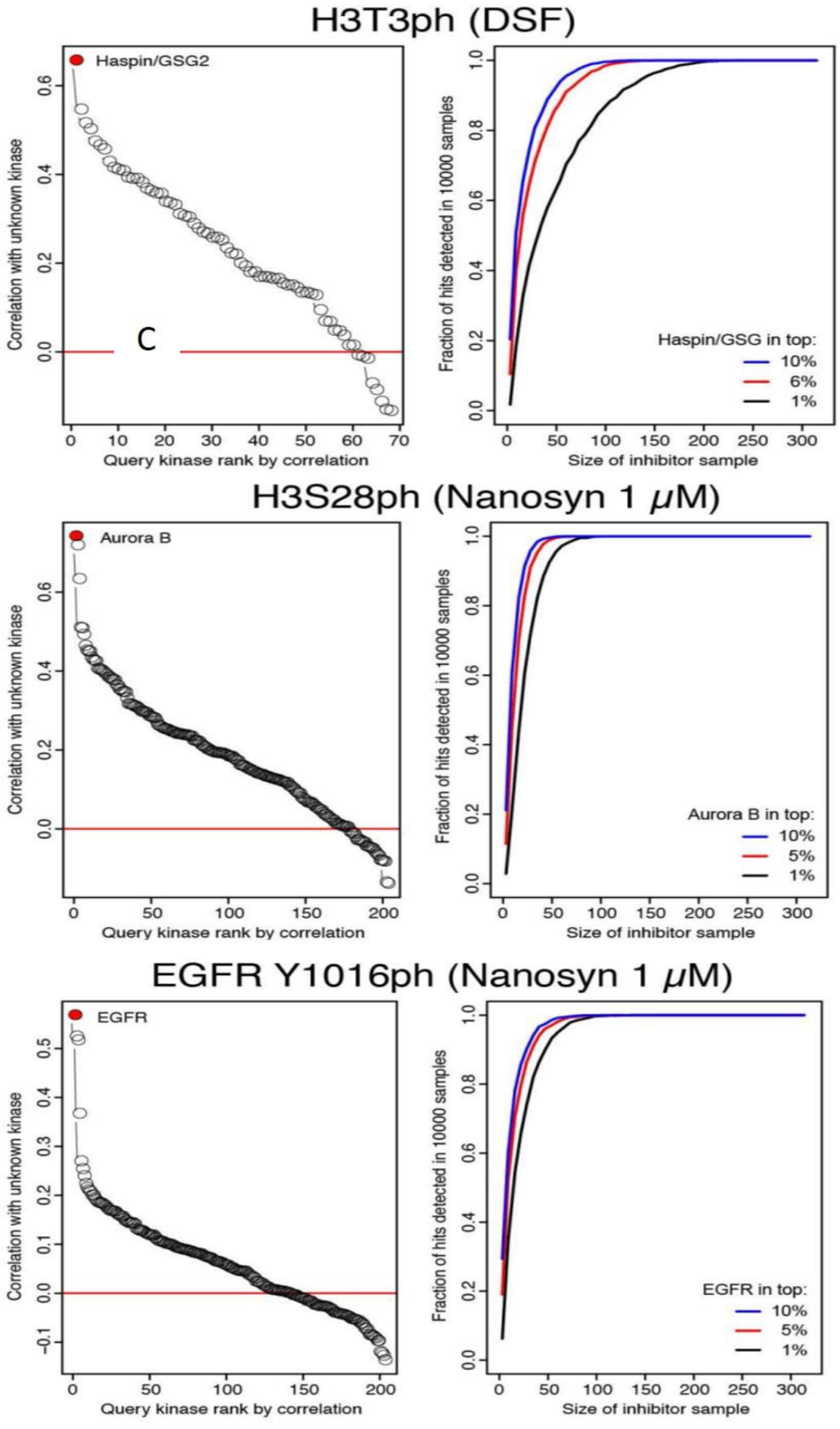
B



"Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)"

Appendix B: Kinome coverage of DiscoverX assay PKIS2 was profiled on

C


H3T3ph (DSF)
Haspin/GSG2
Correlation with unknown kinase
Query kinase rank by correlation
Fraction of hits detected in 10000 samples
Size of inhibitor sample
Haspin/GSG in top:
10%
6%
1%
H3S28ph (Nanosyn 1 μM)
Aurora B
Aurora B in top:
10%
5%
1%
EGFR Y1016ph (Nanosyn 1 μM)
EGFR
EGFR in top:
10%
5%
1%

Appendix C: **Computational downsampling of inhibitor numbers reveals robustness of top hit results for H3T3ph, H3S28ph (repeat screen), and EGFR Y1016ph KiPIK screens**. The righthand plot of each pair shows the fraction of times the expected kinase was found in the top 1%, 5% or 10% of all kinases in the profiling dataset, as a function of inhibitor library size. Lefthand plots show the ranked correlation scored for the unknown kinase with each query kinase in the library. Analysis performed by Conor Lawless

D

H3T3ph

	GPS 3.0		KinasePhos 2.0		NetPhos 3.1		NetPhorest 2.1		NetworKIN 3.0		PHOSIDA		PhosphoNET		ScanSite 4.0	
1	CK1:VRK:VRK2	26.75	No prediction		CDC2/CDK1	0.463	GRK group	0.12	TTK	1.85	No prediction		DNAPK/PRKDC	51	No prediction	
2	CK1 group	20.222			GSK3	0.437	DAPK group	0.11	DAPK3	1			ATR	42		
3	Other:Haspin/GSG2	15			CaM-II/CAMK2	0.436	PKC/PRKC group	0.1	PKCbeta/PRCKB	0.94			MEK7/MAP2K7	36		
4	STE:STE20:MST:MST1/STK4	13.935			PKG	0.400	DMPK group	0.09	PKCalpha/PRKCA	0.6			MEK4/MAP2K4	25		
5	AGC:DMPK:ROCK group	13.242			CKI/CSNK1	0.391	MAP2K group	0.07	PKCzeta/PRKCZ	0.4			H/PK4	21		
6	CK1:VRK group	10.4			p38MAPK	0.361	CLK group	0.06	PKCdelta/PRKCD	0.33			ULK2	6		
7	CAMK:CAMKL:QIK:SIK1	9.625			DNAPK/PRKDC	0.341	DNAPK/PRKDC	0.06	GRK1	0.3			H/PK2	4		
8	CK1:VRK:VRK1	7.7			CKII/CSNK2	0.293	TTK	0.05	GRK3	0.3			H/PK3	3		
9	CK1:CK1:CK1-D:CK1e/CSNK1E	6.941			CDK5	0.236	SLK group	0.04	GRK5	0.3			Trb3/TRIB3	2		
10	STE:STE11:MEKK1/MAP3K1	5.857			ATM	0.233	MST group	0.04	GRK4	0.3			H/PK1	2		

H3S28ph

	GPS 3.0		KinasePhos 2.0		NetPhos 3.1		NetPhorest 2.1		NetworKIN 3.0		PHOSIDA		PhosphoNET		ScanSite 4.0	
1	AGC:DMPK:ROCK group	47.919	Aurora group	0.976391	PKA	0.684	PKA group	0.24	PKAalpha/PRKACA	1.38	PKA	N/A	MTOR/FRAP	409	PKC delta/PRKCD	0.466
2	AGC:GRK:BARK:BARK1/GRK2	38.214	ATM	0.923114	PKG	0.574	PAK group	0.18	PAK1	1.36			PIM1	392	PKA/PRKACG	0.650
3	AGC:PKC:PKCa:PRKCA	20.859	MAPK group	0.809572	PKC	0.477	RCK/ROCK group	0.16	PKCbeta/PRCKB	1.24			PIM3	388	AURORA B	0.716
4	Other:AUR:Aurora C	18.833	IKK/IKBKB	0.697333	GSK3	0.455	CLK group	0.14	PKCalpha/PRKCA	0.99			MSK1/RPS6KA5	387		
5	AGC:RSK:MSK:RPS6KA5	16.741	PKB/AKT group	0.666591	CaM-II/CAMK2	0.425	PIM1/3 group	0.13	PKAbeta/PRKACB	0.49			PKCb/PRKCB1	384		
6	AGC:PKC:PKCa:PRKCB	16	CK1/CSNK1	0.549269	CKI/CSNK1	0.415	PKC group	0.13	PKAgamma/PRKACG	0.49			PKCt/PRKCQ	378		
7	Other:AUR:Aurora B	12.185	RSK group	0.535955	CDC2/CDK1	0.406	DMPK group	0.12	PKCzeta/PRKCZ	0.46			PKCe/PRKCE	375		
8	AGC:GRK:BARK group	11.395	STK4	0.51238	RSK	0.379	ACTR2/2B TGFBR2 group	0.08	PKCdelta/PRKCD	0.36			PKCh/PRKCH	374		
9	Other:AUR:Aurora B:AURKB	11.346	CHK1	0.51238	DNAPK/PRKDC	0.357	PKD group	0.07	PAK3	0.35			PKCd/PRKCD	372		
10	AGC:RSK:MSK group	10.37	AKT1	0.5	ATM	0.319	PIM2	0.07	PAK2	0.35			PIM2	371		

Integrin β1A Y795ph

	GPS 3.0		KinasePhos 2.0		NetPhos 3.1		NetPhorest 2.1		NetworKIN 3.0		PHOSIDA		PhosphoNET		ScanSite 4.0	
1	TK:Eph group	12.927	No prediction		EGFR	0.456	Trk group	0.06	KDR	6.84	EGFR	N/A	ITK	391	No prediction	
2	TK:Csk:CSK	12.815			INSR	0.403	Eph group	0.05	FLT1	0.37			TEC	382		
3	TK:Src:SrcB:HCK	10.2			SRC	0.397	KDR FLT1 group	0.04	TRKA/NTRK1	0.22			ERBB2	377		
4	TK:Jak:JAK2	10.136					MAP2K group	0.04	TRKC/NTRK3	0.19			SRM/SRMS	375		
5	TK:Csk:CTK/MATK	10							TRKB/NTRK2	0.00			BMX	375		
6	TK:Csk group	9.889											BTK	373		
7	TK:Src:SrcA:FGR	8.9											EGFR	373		
8	TK:EGFR:ERBB2	8.6											ERBB4	370		
9	TK:EGFR group	8.494											ALK	367		
10	TK:Trk group	7.4											RET	367		

EGFR Y1016ph

	GPS 3.0		KinasePhos 2.0		NetPhos 3.1		NetPhorest 2.1		NetworKIN 3.0		PHOSIDA		PhosphoNET		ScanSite 4.0	
1	TK:Src:SrcA:FGR	10.9	EGFR	0.942926	No prediction		EGFR group	0.08	ERBB2	23.85	EGFR	N/A	ZAP70	573	EGFR	0.459
2	TK:Abl group	10.562	ZAP70	0.528317			Eph group	0.06	EGFR	14.31			SYK	572	FGR	0.470
3	TK:Eph:EPHB1	10	PDGFR	0.527931									TEC	556	PDGFRB	0.562
4	TK:Syk:SYK	9.521	JAK2	0.526844									ARG/ABL2	555	ABL1	0.673
5	TK:EGFR group	9.253	IGF1R	0.523152									ALK	549		
6	TK:Src:SrcB:LYN	8.861	BTK	0.522771									TXK	548		
7	TK:Syk:ZAP70	8.774	FES	0.521712									AXL	548		
8	TK:Eph:EPHA2	8	RET	0.518117									BLK	547		
9	TK:Trk group	7.867	EPH group	0.513706									KIT	547		
10	TK:Tec:TEC	7.5	IR/INSR	0.507936									FGFR4	546		

INCENP S446ph

	GPS 3.0		KinasePhos 2.0		NetPhos 3.1		NetPhorest 2.1		NetworKIN 3.0		PHOSIDA		PhosphoNET		ScanSite 4.0	
1	AGC:PKC:PKCi group	14.479	ATM	0.886475	PKC	0.692	PKC group	0.39	CDK1	18.96	No prediction		PIM3	544	CDC2/CDK1	0.560
2	AGC:PKC:PKCa:PRKCB	12.552	Aurora group	0.834161	GSK3	0.480	RCK/ROCK group	0.24					PIM1	538	PKC alpha/beta/gamma	0.569
3	AGC:AKT:AKT2	9.929	CK1/CSNK1	0.797518	p38MAPK	0.436	CDK1/2/3/5 group	0.24					PIM2	536	CDK5	0.572
4	AGC:PKC:PKCa group	9.043	MAPK group	0.555368	CaM-II/CAMK2	0.419	PKB/AKT group	0.09					PKCb/PRKCB1	478	GSK3 group	0.602
5	CK1:VRK:VRK2	7.75	RSK group	0.540009	CKI/CSNK1	0.365							PKCt/PRKCQ	474	CDK1 motif 1	0.955
6	AGC:PKC:PKCh group	7.513	STK4	0.513056	DNAPK/PRKDC	0.347							MSK1/RPS6KA5	468	CDK1 motif 2	0.996
7	TKL:IRAK:IRAK1	6.5	CHK1	0.513056	CDC2/CDK1	0.317							ASK1/MAP3K5	466		
8	CAMK:CASK:CASK	6	AKT1	0.510758	RSK	0.307							PKCa/PRKCA	462		
9	CAMK:RAD53 group	5.755	PKG	0.5	ATM	0.286							PKCh/PRKCH	462		
10	CAMK:CAMKL:NuaK:NUAK1	4.625			CDK5	0.283							MAP3K15	459		

Appendix D. Predicted kinases for specific phosphorylation sites using *in silico* approaches.

We used 8 different methods for which web-based tools are available to predict the most likely kinases for each of the phosphorylation sites used in KiPIK screening. Expected kinases are shown in red, and closely related kinases in green. None of the approaches was consistently able to predict the correct kinase. For example, although many methods identified EGFR as the likely kinase for EGFR Y1016ph, three of the methods failed to place this kinase in the top three (GPS 3.0, NetPhos 3.1, and PhosphoNET). In contrast, GPS 3.0 was the only method that was able to place Haspin/GSG2 in the top three places for H3T3ph. For each phosphosite, the number of predictions that placed the correct kinase(s) as the top candidate were: H3T3ph, 0/8; H3S28ph, 1/8; Integrin 1A Y795ph, 0/8; EGFR Y1016ph, 5/8; INCENP S446ph, 2/8. This failure rate occurs despite the fact that the expected kinase is available to the prediction algorithms at high frequency: Haspin, 2/8; Aurora family, 7/8; Src family, 8/8; EGFR, 8/8; Cdk1, 8/8.

GPS 3.0 (http://gps.biocuckoo.org/index.php) uses similarity to experimentally determined kinase target sequences from the phospho.ELM database. Predictions can be made that cover 464 human protein kinases although, because this is a group-based prediction system, many kinases are in subgroups and there are approximately 280 human kinase-specific predictions. The method rests on assumption that kinases that are on a similar branch of a sequence-based dendrogram have similar substrates. The higher the score, the more likely the residue is phosphorylated.

KinasePhos 2.0 (http://kinasephos2.mbc.nctu.edu.tw/index.html) uses sequence-based amino acid coupling pattern analysis and solvent accessibility in a SVM (support vector machine). Validated sites from the phospho.ELM and Swiss-Prot databases were used for training. Predictions can be made for 58 human protein kinases. The higher the score, the more likely the residue is phosphorylated.

NetPhos 3.1 (http://www.cbs.dtu.dk/services/NetPhos/) predicts phosphorylation sites using neural networks based on phosphorylation sites in the phospho.ELM database. Predictions are made for 17 kinases: ATM, CKI/CSNK1, CKII/CSNK2, CaM-II/CAMK2, DNAPK/PRKDC, EGFR, GSK3, INSR, PKA, PKB/AKT, PKC, PKG, RSK, SRC, CDC2/CDK1, CDK5 and p38MAPK. Output scores are in the range 0 to 1; scores above 0.5 indicate “positive predictions”.

NetPhorest 2.1 (http://www.netphorest.info/index.shtml) uses a collection of probabilistic classifiers based on position-specific scoring matrices (PSSM) or neural networks to classify phosphorylation sites according to the likely kinase responsible. Like GPS 3.0, it assumes that kinases that are on a similar branch of a sequence-based dendrogram have similar substrates. It makes predictions for approximately 70 subgroups covering 222 kinases. Scores are from 0 to 1, with high scores being more confident predictions.

NetworKIN 3.0 (http://www.networkin.info/index.shtml) builds on NetPhorest by including contextual information on kinases and substrates from the STRING database, such as binding interactions or substrates that have other residues that are already known targets of a kinase. The theoretically neutral score is 1, and the higher the score for a given kinase, the higher the likelihood it is indeed the kinase. Kinome coverage is as for NetPhorest 2.1. Note that CDK1 is a strong “hit” for INCENP S446ph in part because the STRING database recognises that other INCENP residues are known targets for this kinase.

PHOSIDA (http://141.61.102.18/phosida/index.aspx) includes the simple Motif Matcher tool that searches for sequence matches with annotated kinase recognition patterns for 33 motifs covering 25 kinases. No score is provided7.

PhosphoNET (http://www.phosphonet.ca) includes Kinase Predictor V2 that makes predictions based on determinants of specificity (DoS) within the primary amino acid sequences of the catalytic domains of 488 human protein kinases. The higher the score, the better the prospect that a kinase will phosphorylate a given site, with a maximum possible score of 1000 .

ScanSite 4.0 (https://scansite4.mit.edu/4.0/#home). Phosphorylation sites for particular kinases are predicted using PSSMs determined from oriented peptide library techniques. Thirty-three motifs covering 31 human kinases are included. Scores are on a scale of 0 to infinity, where 0 means a protein sequence perfectly matches the optimal binding pattern, and larger numbers indicate progressively poorer matches to the optimal consensus sequence.

References

ALEXANDER, J., LIM, D., JOUGHIN, B. A., HEGEMANN, B., HUTCHINS, J. R., EHRENBERGER, T., IVINS, F., SESSA, F., HUDECZ, O., NIGG, E. A., FRY, A. M., MUSACCHIO, A., STUKENBERG, P. T., MECHTLER, K., PETERS, J. M., SMERDON, S. J. & YAFFE, M. B. 2011. Spatial exclusivity combined with positive and negative selection of phosphorylation motifs is the basis for context-dependent mitotic signaling. *Sci Signal,* 4**,** ra42.

ALLEN, J. J., LI, M., BRINKWORTH, C. S., PAULSON, J. L., WANG, D., HUBNER, A., CHOU, W. H., DAVIS, R. J., BURLINGAME, A. L., MESSING, R. O., KATAYAMA, C. D., HEDRICK, S. M. & SHOKAT, K. M. 2007. A semisynthetic epitope for kinase substrates. *Nat Methods,* 4**,** 511-6.

AMANCHY, R., PERIASWAMY, B., MATHIVANAN, S., REDDY, R., TATTIKOTA, S. G. & PANDEY, A. 2007. A curated compendium of phosphorylation motifs. *Nat Biotechnol,* 25**,** 285-6.

AZORSA, D. O., ROBESON, R. H., FROST, D., MEEC HOOVET, B., BRAUTIGAM, G. R., DICKEY, C., BEAUDRY, C., BASU, G. D., HOLZ, D. R., HERNANDEZ, J. A., BISANZ, K. M., GWINN, L., GROVER, A., ROGERS, J., REIMAN, E. M., HUTTON, M., STEPHAN, D. A., MOUSSES, S. & DUNCKLEY, T. 2010. High-content siRNA screening of the kinome identifies kinases involved in Alzheimer's disease-related tau hyperphosphorylation. *BMC Genomics,* 11**,** 25.

BAILEY, A. O., PANCHENKO, T., SATHYAN, K. M., PETKOWSKI, J. J., PAI, P. J., BAI, D. L., RUSSELL, D. H., MACARA, I. G., SHABANOWITZ, J., HUNT, D. F., BLACK, B. E. & FOLTZ, D. R. 2013. Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. *Proc Natl Acad Sci U S A,* 110**,** 11827-32.

BAIN, J., PLATER, L., ELLIOTT, M., SHPIRO, N., HASTIE, C. J., MCLAUCHLAN, H., KLEVERNIC, I., ARTHUR, J. S., ALESSI, D. R. & COHEN, P. 2007. The selectivity of protein kinase inhibitors: a further update. *Biochem J,* 408**,** 297-315.

BAMBOROUGH, P., DREWRY, D., HARPER, G., SMITH, G. K. & SCHNEIDER, K. 2008. Assessment of chemical coverage of kinome space and its implications for kinase drug discovery. *J Med Chem,* 51**,** 7898-914.

BANNISTER, A. J., ZEGERMAN, P., PARTRIDGE, J. F., MISKA, E. A., THOMAS, J. O., ALLSHIRE, R. C. & KOUZARIDES, T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature,* 410**,** 120-4.

BISHOP, A. C., UBERSAX, J. A., PETSCH, D. T., MATHEOS, D. P., GRAY, N. S., BLETHROW, J., SHIMIZU, E., TSIEN, J. Z., SCHULTZ, P. G., ROSE, M. D., WOOD, J. L., MORGAN, D. O. & SHOKAT, K. M. 2000. A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature,* 407**,** 395-401.

BISHOP, J. D. & SCHUMACHER, J. M. 2002. Phosphorylation of the carboxyl terminus of inner centromere protein (INCENP) by the Aurora B Kinase stimulates Aurora B kinase activity. *J Biol Chem,* 277**,** 27577-80.

BLETHROW, J. D., GLAVY, J. S., MORGAN, D. O. & SHOKAT, K. M. 2008. Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proc Natl Acad Sci U S A,* 105**,** 1442-7.

BOUTROS, M., HEIGWER, F. & LAUFER, C. 2015. Microscopy-Based High-Content Screening. *Cell,* 163**,** 1314-25.

BREEN, M. E. & SOELLNER, M. B. 2015. Small molecule substrate phosphorylation site inhibitors of protein kinases: approaches and challenges. *ACS Chem Biol,* 10**,** 175-89.

BRINKWORTH, R. I., BREINL, R. A. & KOBE, B. 2003. Structural basis and prediction of substrate specificity in protein serine/threonine kinases. *Proc Natl Acad Sci U S A,* 100**,** 74-9.

BRITO, D. A. & RIEDER, C. L. 2006. Mitotic checkpoint slippage in humans occurs via cyclin B destruction in the presence of an active checkpoint. *Curr Biol,* 16**,** 1194-200.
BRITTLE, A. L., NANBA, Y., ITO, T. & OHKURA, H. 2007. Concerted action of Aurora B, Polo and NHK-1 kinases in centromere-specific histone 2A phosphorylation. *Exp Cell Res,* 313**,** 2780-5.
BROMANN, P. A., KORKAYA, H. & COURTNEIDGE, S. A. 2004. The interplay between Src family kinases and receptor tyrosine kinases. *Oncogene,* 23**,** 7957-68.
BROWN, N. R., NOBLE, M. E., ENDICOTT, J. A. & JOHNSON, L. N. 1999. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat Cell Biol,* 1**,** 438-43.
CAENEPEEL, S., CHARYDCZAK, G., SUDARSANAM, S., HUNTER, T. & MANNING, G. 2004. The mouse kinome: discovery and comparative genomics of all mouse protein kinases. *Proc Natl Acad Sci U S A,* 101**,** 11707-12.
CALDERWOOD, D. A., CAMPBELL, I. D. & CRITCHLEY, D. R. 2013. Talins and kindlins: partners in integrin-mediated adhesion. *Nat Rev Mol Cell Biol,* 14**,** 503-17.
CARMENA, M., RUCHAUD, S. & EARNSHAW, W. C. 2009. Making the Auroras glow: regulation of Aurora A and B kinase function by interacting proteins. *Curr Opin Cell Biol,* 21**,** 796-805.
CARMENA, M., WHEELOCK, M., FUNABIKI, H. & EARNSHAW, W. C. 2012. The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nat Rev Mol Cell Biol,* 13**,** 789-803.
CHANG, F. T., CHAN, F. L., JD, R. M., UDUGAMA, M., MAYNE, L., COLLAS, P., MANN, J. R. & WONG, L. H. 2015. CHK1-driven histone H3.3 serine 31 phosphorylation is important for chromatin maintenance and cell survival in human ALT cancer cells. *Nucleic Acids Res,* 43**,** 2603-14.
CHENG, H. C., MATSUURA, I. & WANG, J. H. 1993. In vitro substrate specificity of protein tyrosine kinases. *Mol Cell Biochem,* 127-128**,** 103-12.
CHILD, E. S. & MANN, D. J. 2006. The intricacies of p21 phosphorylation: protein/protein interactions, subcellular localization and stability. *Cell Cycle,* 5**,** 1313-9.
CROSIO, C., FIMIA, G. M., LOURY, R., KIMURA, M., OKANO, Y., ZHOU, H., SEN, S., ALLIS, C. D. & SASSONE-CORSI, P. 2002. Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. *Mol Cell Biol,* 22**,** 874-85.
CUNY, G. D., ROBIN, M., ULYANOVA, N. P., PATNAIK, D., PIQUE, V., CASANO, G., LIU, J. F., LIN, X., XIAN, J., GLICKSMAN, M. A., STEIN, R. L. & HIGGINS, J. M. 2010. Structure-activity relationship study of acridine analogs as haspin and DYRK2 kinase inhibitors. *Bioorg Med Chem Lett,* 20**,** 3491-4.
CUNY, G. D., ULYANOVA, N. P., PATNAIK, D., LIU, J. F., LIN, X., AUERBACH, K., RAY, S. S., XIAN, J., GLICKSMAN, M. A., STEIN, R. L. & HIGGINS, J. M. 2012. Structure-activity relationship study of beta-carboline derivatives as haspin kinase inhibitors. *Bioorg Med Chem Lett,* 22**,** 2015-9.
DAI, J., SULTAN, S., TAYLOR, S. S. & HIGGINS, J. M. G. 2005. The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes Dev.,* 19**,** 472-488.
DAR, A. C. & SHOKAT, K. M. 2011. The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling. *Annu Rev Biochem,* 80**,** 769-95.
DAVIS, M. I., HUNT, J. P., HERRGARD, S., CICERI, P., WODICKA, L. M., PALLARES, G., HOCKER, M., TREIBER, D. K. & ZARRINKAR, P. P. 2011. Comprehensive analysis of kinase inhibitor selectivity. *Nat Biotechnol,* 29**,** 1046-51.

DE ANTONI, A., MAFFINI, S., KNAPP, S., MUSACCHIO, A. & SANTAGUIDA, S. 2012. A small molecule inhibitor of Haspin alters kinetochore functions of Aurora B. *J Cell Biol,* 199**,** 269-84.

DE OLIVEIRA, P. S., FERRAZ, F. A., PENA, D. A., PRAMIO, D. T., MORAIS, F. A. & SCHECHTMAN, D. 2016. Revisiting protein kinase-substrate interactions: Toward therapeutic development. *Sci Signal,* 9**,** re3.

DEDIGAMA-ARACHCHIGE, P. M. & PFLUM, M. K. 2016. K-CLASP: A Tool to Identify Phosphosite Specific Kinases and Interacting Proteins. *ACS Chem Biol,* 11**,** 3251-3255.

DEIBLER, R. W. & KIRSCHNER, M. W. 2010. Quantitative reconstitution of mitotic CDK1 activation in somatic cell extracts. *Mol Cell,* 37**,** 753-67.

DELUCA, K. F., LENS, S. M. & DELUCA, J. G. 2011. Temporal changes in Hec1 phosphorylation control kinetochore-microtubule attachment stability during mitosis. *J Cell Sci,* 124**,** 622-34.

DEPHOURE, N., ZHOU, C., VILLEN, J., BEAUSOLEIL, S. A., BAKALARSKI, C. E., ELLEDGE, S. J. & GYGI, S. P. 2008. A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci U S A,* 105**,** 10762-7.

DITCHFIELD, C., JOHNSON, V. L., TIGHE, A., ELLSTON, R., HAWORTH, C., JOHNSON, T., MORTLOCK, A., KEEN, N. & TAYLOR, S. S. 2003. Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores. *J. Cell Biol.,* 161**,** 267-80.

DOWNWARD, J., YARDEN, Y., MAYES, E., SCRACE, G., TOTTY, N., STOCKWELL, P., ULLRICH, A., SCHLESSINGER, J. & WATERFIELD, M. D. 1984. Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. *Nature,* 307**,** 521-7.

DREWRY, D. H., WELLS, C. I., ANDREWS, D. M., ANGELL, R., AL-ALI, H., AXTMAN, A. D., CAPUZZI, S. J., ELKINS, J. M., ETTMAYER, P., FREDERIKSEN, M., GILEADI, O., GRAY, N., HOOPER, A., KNAPP, S., LAUFER, S., LUECKING, U., MICHAELIDES, M., MULLER, S., MURATOV, E., DENNY, R. A., SAIKATENDU, K. S., TREIBER, D. K., ZUERCHER, W. J. & WILLSON, T. M. 2017. Progress towards a public chemogenomic set for protein kinases and a call for contributions. *PLoS One,* 12**,** e0181585.

EGLI, D., BIRKHOFF, G. & EGGAN, K. 2008. Mediators of reprogramming: transcription factors and transitions through mitosis. *Nat Rev Mol Cell Biol,* 9**,** 505-16.

ELKINS, J. M., FEDELE, V., SZKLARZ, M., ABDUL AZEEZ, K. R., SALAH, E., MIKOLAJCZYK, J., ROMANOV, S., SEPETOV, N., HUANG, X. P., ROTH, B. L., AL HAJ ZEN, A., FOURCHES, D., MURATOV, E., TROPSHA, A., MORRIS, J., TEICHER, B. A., KUNKEL, M., POLLEY, E., LACKEY, K. E., ATKINSON, F. L., OVERINGTON, J. P., BAMBOROUGH, P., MULLER, S., PRICE, D. J., WILLSON, T. M., DREWRY, D. H., KNAPP, S. & ZUERCHER, W. J. 2016. Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat Biotechnol,* 34**,** 95-103.

ELLIS, J. J. & KOBE, B. 2011. Predicting protein kinase specificity: Predikin update and performance in the DREAM4 challenge. *PLoS One,* 6**,** e21169.

ESWARAN, J., PATNAIK, D., FILIPPAKOPOULOS, P., WANG, F., STEIN, R. L., MURRAY, J. W., HIGGINS, J. M. & KNAPP, S. 2009. Structure and functional characterization of the atypical human kinase haspin. *Proc Natl Acad Sci U S A,* 106**,** 20198-203.

FIELDS, A. P. & THOMPSON, L. J. 1995. The regulation of mitotic nuclear envelope breakdown: a role for multiple lamin kinases. *Prog Cell Cycle Res,* 1**,** 271-86.

FISCHLE, W., TSENG, B. S., DORMANN, H. L., UEBERHEIDE, B. M., GARCIA, B. A., SHABANOWITZ, J., HUNT, D. F., FUNABIKI, H. & ALLIS, C. D. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature,* 438**,** 1116-22.

FONSECA, J. P., STEFFEN, P. A., MULLER, S., LU, J., SAWICKA, A., SEISER, C. & RINGROSE, L. 2012. In vivo Polycomb kinetics and mitotic chromatin binding distinguish stem cells from differentiated cells. *Genes Dev,* 26**,** 857-71.

FRIEDMAN, A. & PERRIMON, N. 2006. A functional RNAi screen for regulators of receptor tyrosine kinase and ERK signalling. *Nature,* 444**,** 230-4.

FURCHT, C. M., BUONATO, J. M. & LAZZARA, M. J. 2015. EGFR-activated Src family kinases maintain GAB1-SHP2 complexes distal from EGFR. *Sci Signal,* 8**,** ra46.

GEHANI, S. S., AGRAWAL-SINGH, S., DIETRICH, N., CHRISTOPHERSEN, N. S., HELIN, K. & HANSEN, K. 2010. Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation. *Mol Cell,* 39**,** 886-900.

GHEGHIANI, L., LOEW, D., LOMBARD, B., MANSFELD, J. & GAVET, O. 2017. PLK1 Activation in Late G2 Sets Up Commitment to Mitosis. *Cell Rep,* 19**,** 2060-2073.

GHENOIU, C., WHEELOCK, M. S. & FUNABIKI, H. 2013. Autoinhibition and Polo-dependent multisite phosphorylation restrict activity of the histone H3 kinase Haspin to mitosis. *Mol Cell,* 52**,** 734-45.

GIRDLER, F., GASCOIGNE, K. E., EYERS, P. A., HARTMUTH, S., CRAFTER, C., FOOTE, K. M., KEEN, N. J. & TAYLOR, S. S. 2006. Validating Aurora B as an anti-cancer drug target. *J Cell Sci,* 119**,** 3664-75.

GOOD, M. C., ZALATAN, J. G. & LIM, W. A. 2011. Scaffold proteins: hubs for controlling the flow of cellular information. *Science,* 332**,** 680-6.

GOTO, H., YASUI, Y., NIGG, E. A. & INAGAKI, M. 2002. Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cells,* 7**,** 11-7.

GOVIN, J., DORSEY, J., GAUCHER, J., ROUSSEAUX, S., KHOCHBIN, S. & BERGER, S. L. 2010. Systematic screen reveals new functional dynamics of histones H3 and H4 during gametogenesis. *Genes Dev,* 24**,** 1772-86.

GRUNEBERG, U., NEEF, R., HONDA, R., NIGG, E. A. & BARR, F. A. 2004. Relocation of Aurora B from centromeres to the central spindle at the metaphase to anaphase transition requires MKlp2. *J Cell Biol,* 166**,** 167-72.

HAKE, S. B., GARCIA, B. A., KAUER, M., BAKER, S. P., SHABANOWITZ, J., HUNT, D. F. & ALLIS, C. D. 2005. Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc Natl Acad Sci U S A,* 102**,** 6344-9.

HALE, T. K., CONTRERAS, A., MORRISON, A. J. & HERRERA, R. E. 2006. Phosphorylation of the linker histone H1 by CDK regulates its binding to HP1alpha. *Mol Cell,* 22**,** 693-9.

HASCHKA, M. D., SORATROI, C., KIRSCHNEK, S., HACKER, G., HILBE, R., GELEY, S., VILLUNGER, A. & FAVA, L. L. 2015. The NOXA-MCL1-BIM axis defines lifespan on extended mitotic arrest. *Nat Commun,* 6**,** 6891.

HEGEMANN, B., HUTCHINS, J. R., HUDECZ, O., NOVATCHKOVA, M., RAMESEDER, J., SYKORA, M. M., LIU, S., MAZANEK, M., LENART, P., HERICHE, J. K., POSER, I., KRAUT, N., HYMAN, A. A., YAFFE, M. B., MECHTLER, K. & PETERS, J. M. 2011. Systematic phosphorylation analysis of human mitotic protein complexes. *Sci Signal,* 4**,** rs12.

HERTZ, N. T., WANG, B. T., ALLEN, J. J., ZHANG, C., DAR, A. C., BURLINGAME, A. L. & SHOKAT, K. M. 2010. Chemical genetic approach for kinase-substrate mapping by covalent capture of thiophosphopeptides and analysis by mass spectrometry. *Curr Protoc Chem Biol,* 2**,** 15-36.
HINDRIKSEN, S., LENS, S. M. A. & HADDERS, M. A. 2017. The Ins and Outs of Aurora B Inner Centromere Localization. *Front Cell Dev Biol,* 5**,** 112.
HIROTA, T., LIPP, J. J., TOH, B. H. & PETERS, J. M. 2005. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature,* 438**,** 1176-80.
HORN, H., SCHOOF, E. M., KIM, J., ROBIN, X., MILLER, M. L., DIELLA, F., PALMA, A., CESARENI, G., JENSEN, L. J. & LINDING, R. 2014. KinomeXplorer: an integrated platform for kinome biology studies. *Nat Methods,* 11**,** 603-4.
HORNBERG, J. J., BRUGGEMAN, F. J., BINDER, B., GEEST, C. R., DE VAATE, A. J., LANKELMA, J., HEINRICH, R. & WESTERHOFF, H. V. 2005. Principles behind the multifarious control of signal transduction. ERK phosphorylation and kinase/phosphatase control. *FEBS J,* 272**,** 244-58.
HSU, J. Y., SUN, Z. W., LI, X., REUBEN, M., TATCHELL, K., BISHOP, D. K., GRUSHCOW, J. M., BRAME, C. J., CALDWELL, J. A., HUNT, D. F., LIN, R., SMITH, M. M. & ALLIS, C. D. 2000. Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell,* 102**,** 279-91.
HUTTI, J. E., JARRELL, E. T., CHANG, J. D., ABBOTT, D. W., STORZ, P., TOKER, A., CANTLEY, L. C. & TURK, B. E. 2004. A rapid method for determining protein kinase phosphorylation specificity. *Nat Methods,* 1**,** 27-9.
HUTTLIN, E. L., JEDRYCHOWSKI, M. P., ELIAS, J. E., GOSWAMI, T., RAD, R., BEAUSOLEIL, S. A., VILLEN, J., HAAS, W., SOWA, M. E. & GYGI, S. P. 2010. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell,* 143**,** 1174-89.
INBAL, B., SHANI, G., COHEN, O., KISSIL, J. L. & KIMCHI, A. 2000. Death-associated protein kinase-related protein 1, a novel serine/threonine kinase involved in apoptosis. *Mol Cell Biol,* 20**,** 1044-54.
JANSSON, D., NG, A. C., FU, A., DEPATIE, C., AL AZZABI, M. & SCREATON, R. A. 2008. Glucose controls CREB activity in islet cells via regulated phosphorylation of TORC2. *Proc Natl Acad Sci U S A,* 105**,** 10161-6.
JI, J. H., HWANG, H. I., LEE, H. J., HYUN, S. Y., KANG, H. J. & JANG, Y. J. 2010. Purification and proteomic identification of putative upstream regulators of polo-like kinase-1 from mitotic cell extracts. *FEBS Lett,* 584**,** 4299-305.
JOHNSON, L. N. 2011. Substrates of mitotic kinases. *Sci Signal,* 4**,** pe31.
JOHNSON, L. N. & LEWIS, R. J. 2001. Structural basis for control by phosphorylation. *Chem Rev,* 101**,** 2209-42.
JOUGHIN, B. A., LIU, C., LAUFFENBURGER, D. A., HOGUE, C. W. & YAFFE, M. B. 2012. Protein kinases display minimal interpositional dependence on substrate sequence: potential implications for the evolution of signalling networks. *Philos Trans R Soc Lond B Biol Sci,* 367**,** 2574-83.
KAWASHIMA, S. A., YAMAGISHI, Y., HONDA, T., ISHIGURO, K. & WATANABE, Y. 2010. Phosphorylation of H2A by Bub1 prevents chromosomal instability through localizing shugoshin. *Science,* 327**,** 172-7.
KELLY, A. E., GHENOIU, C., XUE, J. Z., ZIERHUT, C., KIMURA, H. & FUNABIKI, H. 2010. Survivin reads phosphorylated histone H3 threonine 3 to activate the mitotic kinase Aurora B. *Science,* 330**,** 235-9.

KIRKPATRICK, D. S., GERBER, S. A. & GYGI, S. P. 2005. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods,* 35**,** 265-73.

KORNEV, A. P., HASTE, N. M., TAYLOR, S. S. & EYCK, L. F. 2006. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A,* 103**,** 17783-8.

KUBOTA, K., ANJUM, R., YU, Y., KUNZ, R. C., ANDERSEN, J. N., KRAUS, M., KEILHACK, H., NAGASHIMA, K., KRAUSS, S., PAWELETZ, C., HENDRICKSON, R. C., FELDMAN, A. S., WU, C. L., RUSH, J., VILLEN, J. & GYGI, S. P. 2009. Sensitive multiplexed analysis of kinase activities and activity-based kinase identification. *Nat Biotechnol,* 27**,** 933-40.

KUFER, T. A., SILLJE, H. H., KORNER, R., GRUSS, O. J., MERALDI, P. & NIGG, E. A. 2002. Human TPX2 is required for targeting Aurora-A kinase to the spindle. *J Cell Biol,* 158**,** 617-23.

KUNITOKU, N., SASAYAMA, T., MARUMOTO, T., ZHANG, D., HONDA, S., KOBAYASHI, O., HATAKEYAMA, K., USHIO, Y., SAYA, H. & HIROTA, T. 2003. CENP-A phosphorylation by Aurora-A in prophase is required for enrichment of Aurora-B at inner centromeres and for kinetochore function. *Dev Cell,* 5**,** 853-64.

LABRADOR, J. P., AZCOITIA, V., TUCKERMANN, J., LIN, C., OLASO, E., MANES, S., BRUCKNER, K., GOERGEN, J. L., LEMKE, G., YANCOPOULOS, G., ANGEL, P., MARTINEZ, C. & KLEIN, R. 2001. The collagen receptor DDR2 regulates proliferation and its elimination leads to dwarfism. *EMBO Rep,* 2**,** 446-52.

LAU, P. N. & CHEUNG, P. 2011. Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing. *Proc Natl Acad Sci U S A,* 108**,** 2801-6.

LEITINGER, B. 2011. Transmembrane collagen receptors. *Annu Rev Cell Dev Biol,* 27**,** 265-90.

LI, S., DENG, Z., FU, J., XU, C., XIN, G., WU, Z., LUO, J., WANG, G., ZHANG, S., ZHANG, B., ZOU, F., JIANG, Q. & ZHANG, C. 2015. Spatial Compartmentalization Specializes the Function of Aurora A and Aurora B. *J Biol Chem,* 290**,** 17546-58.

LIU, X., SONG, Z., HUO, Y., ZHANG, J., ZHU, T., WANG, J., ZHAO, X., AIKHIONBARE, F., ZHANG, J., DUAN, H., WU, J., DOU, Z., SHI, Y. & YAO, X. 2014. Chromatin protein HP1 interacts with the mitotic regulator borealin protein and specifies the centromere localization of the chromosomal passenger complex. *J Biol Chem,* 289**,** 20638-49.

MALY, D. J., ALLEN, J. A. & SHOKAT, K. M. 2004. A mechanism-based cross-linker for the identification of kinase-substrate pairs. *J Am Chem Soc,* 126**,** 9160-1.

MANNING, G., WHYTE, D. B., MARTINEZ, R., HUNTER, T. & SUDARSANAM, S. 2002. The protein kinase complement of the human genome. *Science,* 298**,** 1912-34.

MARQUEZ, J. & OLASO, E. 2014. Role of discoidin domain receptor 2 in wound healing. *Histol Histopathol,* 29**,** 1355-64.

MATEESCU, B., ENGLAND, P., HALGAND, F., YANIV, M. & MUCHARDT, C. 2004. Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. *EMBO Rep,* 5**,** 490-6.

METZGER, E., YIN, N., WISSMANN, M., KUNOWSKA, N., FISCHER, K., FRIEDRICHS, N., PATNAIK, D., HIGGINS, J. M., POTIER, N., SCHEIDTMANN, K. H., BUETTNER, R. & SCHULE, R. 2008. Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation. *Nat Cell Biol,* 10**,** 53-60.

MILLER, C. J. & TURK, B. E. 2018. Homing in: Mechanisms of Substrate Targeting by Protein Kinases. *Trends Biochem Sci,* 43**,** 380-394.
MILLER, M. L., JENSEN, L. J., DIELLA, F., JORGENSEN, C., TINTI, M., LI, L., HSIUNG, M., PARKER, S. A., BORDEAUX, J., SICHERITZ-PONTEN, T., OLHOVSKY, M., PASCULESCU, A., ALEXANDER, J., KNAPP, S., BLOM, N., BORK, P., LI, S., CESARENI, G., PAWSON, T., TURK, B. E., YAFFE, M. B., BRUNAK, S. & LINDING, R. 2008. Linear motif atlas for phosphorylation-dependent signaling. *Sci Signal,* 1**,** ra2.
MOFFAT, J., GRUENEBERG, D. A., YANG, X., KIM, S. Y., KLOEPFER, A. M., HINKLE, G., PIQANI, B., EISENHAURE, T. M., LUO, B., GRENIER, J. K., CARPENTER, A. E., FOO, S. Y., STEWART, S. A., STOCKWELL, B. R., HACOHEN, N., HAHN, W. C., LANDER, E. S., SABATINI, D. M. & ROOT, D. E. 2006. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell,* 124**,** 1283-98.
MOK, J., KIM, P. M., LAM, H. Y., PICCIRILLO, S., ZHOU, X., JESCHKE, G. R., SHERIDAN, D. L., PARKER, S. A., DESAI, V., JWA, M., CAMERONI, E., NIU, H., GOOD, M., REMENYI, A., MA, J. L., SHEU, Y. J., SASSI, H. E., SOPKO, R., CHAN, C. S., DE VIRGILIO, C., HOLLINGSWORTH, N. M., LIM, W. A., STERN, D. F., STILLMAN, B., ANDREWS, B. J., GERSTEIN, M. B., SNYDER, M. & TURK, B. E. 2010. Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Sci Signal,* 3**,** ra12.
NAKAYAMA, J., RICE, J. C., STRAHL, B. D., ALLIS, C. D. & GREWAL, S. I. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science,* 292**,** 110-3.
NASA, I., RUSIN, S. F., KETTENBACH, A. N. & MOORHEAD, G. B. 2018. Aurora B opposes PP1 function in mitosis by phosphorylating the conserved PP1-binding RVxF motif in PP1 regulatory proteins. *Sci Signal,* 11.
NATALYA NADY, J. M., MICHAEL KARETA, FREDERIC CHEDIN AND CHERYL ARROWSMITH 2008. A SPOT on the chromatin landscape? Histone peptide arrays as a tool for epigenetic research. *Cell*.
NOZAWA, R. S., NAGAO, K., MASUDA, H. T., IWASAKI, O., HIROTA, T., NOZAKI, N., KIMURA, H. & OBUSE, C. 2010. Human POGZ modulates dissociation of HP1alpha from mitotic chromosome arms through Aurora B activation. *Nat Cell Biol,* 12**,** 719-27.
OLASO, E., LABRADOR, J. P., WANG, L., IKEDA, K., ENG, F. J., KLEIN, R., LOVETT, D. H., LIN, H. C. & FRIEDMAN, S. L. 2002. Discoidin domain receptor 2 regulates fibroblast proliferation and migration through the extracellular matrix in association with transcriptional activation of matrix metalloproteinase-2. *J Biol Chem,* 277**,** 3606-13.
OUE, Y., MURAKAMI, S., ISSHIKI, K., TSUJI, A. & YUASA, K. 2018. Intracellular localization and binding partners of death associated protein kinase-related apoptosis-inducing protein kinase 1. *Biochem Biophys Res Commun,* 496**,** 1222-1228.
PAPAGEORGIOU, A., RAPLEY, J., MESIROV, J. P., TAMAYO, P. & AVRUCH, J. 2015. A genome-wide siRNA screen in mammalian cells for regulators of S6 phosphorylation. *PLoS One,* 10**,** e0116096.
PARK, J. I. 2014. Growth arrest signaling of the Raf/MEK/ERK pathway in cancer. *Front Biol (Beijing),* 9**,** 95-103.
PECK, S. C. 2006. Analysis of protein phosphorylation: methods and strategies for studying kinases and substrates. *Plant J,* 45**,** 512-22.

PETI, W. & PAGE, R. 2013. Molecular basis of MAP kinase regulation. *Protein Sci,* 22**,** 1698-710.
PETRONCZKI, M., LENART, P. & PETERS, J. M. 2008. Polo on the Rise-from Mitotic Entry to Cytokinesis with Plk1. *Dev Cell,* 14**,** 646-59.
PREUSS, U., LANDSBERG, G. & SCHEIDTMANN, K. H. 2003. Novel mitosis-specific phosphorylation of histone H3 at Thr11 mediated by Dlk/ZIP kinase. *Nucleic Acids Res,* 31**,** 878-85.
PUNTERVOLL, P., LINDING, R., GEMUND, C., CHABANIS-DAVIDSON, S., MATTINGSDAL, M., CAMERON, S., MARTIN, D. M., AUSIELLO, G., BRANNETTI, B., COSTANTINI, A., FERRE, F., MASELLI, V., VIA, A., CESARENI, G., DIELLA, F., SUPERTI-FURGA, G., WYRWICZ, L., RAMU, C., MCGUIGAN, C., GUDAVALLI, R., LETUNIC, I., BORK, P., RYCHLEWSKI, L., KUSTER, B., HELMER-CITTERICH, M., HUNTER, W. N., AASLAND, R. & GIBSON, T. J. 2003. ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res,* 31**,** 3625-30.
QIAN, J., BEULLENS, M., HUANG, J., DE MUNTER, S., LESAGE, B. & BOLLEN, M. 2015. Cdk1 orders mitotic events through coordination of a chromosome-associated phosphatase switch. *Nat Commun,* 6**,** 10215.
QIAN, J., BEULLENS, M., LESAGE, B. & BOLLEN, M. 2013. Aurora B defines its own chromosomal targeting by opposing the recruitment of the phosphatase scaffold Repo-Man. *Curr Biol,* 23**,** 1136-43.
QIAN, J., LESAGE, B., BEULLENS, M., VAN EYNDE, A. & BOLLEN, M. 2011. PP1/Repo-man dephosphorylates mitotic histone H3 at T3 and regulates chromosomal aurora B targeting. *Curr Biol,* 21**,** 766-73.
RAMAKRISHNAN, R. & RICE, A. P. 2012. Cdk9 T-loop phosphorylation is regulated by the calcium signaling pathway. *J Cell Physiol,* 227**,** 609-17.
RATTANI, A., VINOD, P. K., GODWIN, J., TACHIBANA-KONWALSKI, K., WOLNA, M., MALUMBRES, M., NOVAK, B. & NASMYTH, K. 2014. Dependency of the spindle assembly checkpoint on Cdk1 renders the anaphase transition irreversible. *Curr Biol,* 24**,** 630-7.
ROTIN, D., MARGOLIS, B., MOHAMMADI, M., DALY, R. J., DAUM, G., LI, N., FISCHER, E. H., BURGESS, W. H., ULLRICH, A. & SCHLESSINGER, J. 1992. SH2 domains prevent tyrosine dephosphorylation of the EGF receptor: identification of Tyr992 as the high-affinity binding site for SH2 domains of phospholipase C gamma. *EMBO J,* 11**,** 559-67.
RUBIN, C. S. & ROSEN, O. M. 1975. Protein phosphorylation. *Annu Rev Biochem,* 44**,** 831-87.
RUPPERT, J. G., SAMEJIMA, K., PLATANI, M., MOLINA, O., KIMURA, H., JEYAPRAKASH, A. A., OHTA, S. & EARNSHAW, W. C. 2018. HP1alpha targets the chromosomal passenger complex for activation at heterochromatin before mitotic entry. *EMBO J,* 37.
SAKAI, T., JOVE, R., FASSLER, R. & MOSHER, D. F. 2001. Role of the cytoplasmic tyrosines of beta 1A integrins in transformation by v-src. *Proc Natl Acad Sci U S A,* 98**,** 3808-13.
SALIMIAN, K. J., BALLISTER, E. R., SMOAK, E. M., WOOD, S., PANCHENKO, T., LAMPSON, M. A. & BLACK, B. E. 2011. Feedback control in sensing chromosome biorientation by the Aurora B kinase. *Curr Biol,* 21**,** 1158-65.
SCHLESSINGER, J. & LEMMON, M. A. 2003. SH2 and PTB domains in tyrosine kinase signaling. *Sci STKE,* 2003**,** RE12.

SCOTT, J. D. & PAWSON, T. 2009. Cell signaling in space and time: where proteins come together and when they're apart. *Science,* 326**,** 1220-4.
SHAH, K., LIU, Y., DEIRMENGIAN, C. & SHOKAT, K. M. 1997. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc Natl Acad Sci U S A,* 94**,** 3565-70.
SHARMA, K., D'SOUZA, R. C., TYANOVA, S., SCHAAB, C., WISNIEWSKI, J. R., COX, J. & MANN, M. 2014. Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Rep,* 8**,** 1583-94.
SHI, Y. 2009. Serine/threonine phosphatases: mechanism through structure. *Cell,* 139**,** 468-84.
SHIMADA, M., NIIDA, H., ZINELDEEN, D. H., TAGAMI, H., TANAKA, M., SAITO, H. & NAKANISHI, M. 2008. Chk1 is a histone H3 threonine 11 kinase that regulates DNA damage-induced transcriptional repression. *Cell,* 132**,** 221-32.
SONG, J., WANG, H., WANG, J., LEIER, A., MARQUEZ-LAGO, T., YANG, B., ZHANG, Z., AKUTSU, T., WEBB, G. I. & DALY, R. J. 2017. PhosphoPredict: A bioinformatics tool for prediction of human kinase-specific phosphorylation substrates and sites by integrating heterogeneous feature selection. *Sci Rep,* 7**,** 6862.
SONGYANG, Z., BLECHNER, S., HOAGLAND, N., HOEKSTRA, M. F., PIWNICA-WORMS, H. & CANTLEY, L. C. 1994. Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol,* 4**,** 973-82.
STATSUK, A. V., MALY, D. J., SEELIGER, M. A., FABIAN, M. A., BIGGS, W. H., 3RD, LOCKHART, D. J., ZARRINKAR, P. P., KURIYAN, J. & SHOKAT, K. M. 2008. Tuning a three-component reaction for trapping kinase substrate complexes. *J Am Chem Soc,* 130**,** 17568-74.
STELLFOX, M. E., BAILEY, A. O. & FOLTZ, D. R. 2012. Putting CENP-A in its place. *Cell Mol Life Sci*.
STUCKE, V. M., SILLJE, H. H., ARNAUD, L. & NIGG, E. A. 2002. Human Mps1 kinase is required for the spindle assembly checkpoint but not for centrosome duplication. *EMBO J,* 21**,** 1723-32.
SUDAKIN, V., CHAN, G. K. & YEN, T. J. 2001. Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2. *J Cell Biol,* 154**,** 925-36.
SUIJKERBUIJK, S. J., VLEUGEL, M., TEIXEIRA, A. & KOPS, G. J. 2012. Integration of kinase and phosphatase activities by BUBR1 ensures formation of stable kinetochore-microtubule attachments. *Dev Cell,* 23**,** 745-55.
SUZUKI, K., SAKO, K., AKIYAMA, K., ISODA, M., SENOO, C., NAKAJO, N. & SAGATA, N. 2015. Identification of non-Ser/Thr-Pro consensus motifs for Cdk1 and their roles in mitotic regulation of C2H2 zinc finger proteins and Ect2. *Sci Rep,* 5**,** 7929.
TANOUE, T., ADACHI, M., MORIGUCHI, T. & NISHIDA, E. 2000. A conserved docking motif in MAP kinases common to substrates, activators and regulators. *Nat Cell Biol,* 2**,** 110-6.
TATEBE, H., MURAYAMA, S., YONEKURA, T., HATANO, T., RICHTER, D., FURUYA, T., KATAOKA, S., FURUITA, K., KOJIMA, C. & SHIOZAKI, K. 2017. Substrate specificity of TOR complex 2 is determined by a ubiquitin-fold domain of the Sin1 subunit. *Elife,* 6.
TAYLOR, S. S. & KORNEV, A. P. 2011. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci,* 36**,** 65-77.
TERADA, Y. 2006. Aurora-B/AIM-1 regulates the dynamic behavior of HP1alpha at the G2-M transition. *Mol Biol Cell,* 17**,** 3232-41.

TOPHAM, C. H. & TAYLOR, S. S. 2013. Mitosis and apoptosis: how is the balance set? *Curr Opin Cell Biol,* 25**,** 780-5.
TSUKAHARA, T., TANNO, Y. & WATANABE, Y. 2010. Phosphorylation of the CPC by Cdk1 promotes chromosome bi-orientation. *Nature,* 467**,** 719-23.
UBERSAX, J. A. & FERRELL, J. E., JR. 2007. Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol,* 8**,** 530-41.
VAN DE WEERDT, B. C., LITTLER, D. R., KLOMPMAKER, R., HUSEINOVIC, A., FISH, A., PERRAKIS, A. & MEDEMA, R. H. 2008. Polo-box domains confer target specificity to the Polo-like kinase family. *Biochim Biophys Acta,* 1783**,** 1015-22.
VAN DER HORST, A., VROMANS, M. J., BOUWMAN, K., VAN DER WAAL, M. S., HADDERS, M. A. & LENS, S. M. 2015. Inter-domain Cooperation in INCENP Promotes Aurora B Relocation from Centromeres to Microtubules. *Cell Rep,* 12**,** 380-7.
VARIER, R. A., OUTCHKOUROV, N. S., DE GRAAF, P., VAN SCHAIK, F. M., ENSING, H. J., WANG, F., HIGGINS, J. M., KOPS, G. J. & TIMMERS, H. T. 2010. A phospho/methyl switch at histone H3 regulates TFIID association with mitotic chromosomes. *EMBO J,* 29**,** 3967-78.
VASSILEV, L. T., TOVAR, C., CHEN, S., KNEZEVIC, D., ZHAO, X., SUN, H., HEIMBROOK, D. C. & CHEN, L. 2006. Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. *Proc Natl Acad Sci U S A,* 103**,** 10660-5.
VERMEULEN, M., EBERL, H. C., MATARESE, F., MARKS, H., DENISSOV, S., BUTTER, F., LEE, K. K., OLSEN, J. V., HYMAN, A. A., STUNNENBERG, H. G. & MANN, M. 2010. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell,* 142**,** 967-80.
VOGEL, W., GISH, G. D., ALVES, F. & PAWSON, T. 1997. The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol Cell,* 1**,** 13-23.
VON STECHOW, L., FRANCAVILLA, C. & OLSEN, J. V. 2015. Recent findings and technological advances in phosphoproteomics for cells and tissues. *Expert Rev Proteomics,* 12**,** 469-87.
WALTON, G. M., CHEN, W. S., ROSENFELD, M. G. & GILL, G. N. 1990. Analysis of deletions of the carboxyl terminus of the epidermal growth factor receptor reveals self-phosphorylation at tyrosine 992 and enhanced in vivo tyrosine phosphorylation of cell substrates. *J Biol Chem,* 265**,** 1750-4.
WANG, F., DAI, J., DAUM, J. R., NIEDZIALKOWSKA, E., BANERJEE, B., STUKENBERG, P. T., GORBSKY, G. J. & HIGGINS, J. M. G. 2010. Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. *Science,* 330**,** 231-235.
WANG, F. & HIGGINS, J. M. 2013. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol,* 23**,** 175-84.
WANG, F., ULYANOVA, N. P., DAUM, J. R., PATNAIK, D., KATENEVA, A. V., GORBSKY, G. J. & HIGGINS, J. M. G. 2012. Haspin inhibitors reveal centromeric functions of Aurora B in chromosome segregation. *J Cell Biol,* 199**,** 251-68.
WANG, F., ULYANOVA, N. P., VAN DER WAAL, M. S., PATNAIK, D., LENS, S. M. A. & HIGGINS, J. M. G. 2011. A positive feedback loop involving Haspin and Aurora B promotes CPC accumulation at centromeres in mitosis. *Curr. Biol.,* 21**,** 1061-9.
WANG, T. H., POPP, D. M., WANG, H. S., SAITOH, M., MURAL, J. G., HENLEY, D. C., ICHIJO, H. & WIMALASENA, J. 1999. Microtubule dysfunction induced by paclitaxel initiates apoptosis through both c-Jun N-terminal kinase (JNK)-dependent and -independent pathways in ovarian cancer cells. *J Biol Chem,* 274**,** 8208-16.

WULLSCHLEGER, S., LOEWITH, R. & HALL, M. N. 2006. TOR signaling in growth and metabolism. *Cell,* 124**,** 471-84.
XU, Z., VAGNARELLI, P., OGAWA, H., SAMEJIMA, K. & EARNSHAW, W. C. 2010. Gradient of increasing Aurora B kinase activity is required for cells to execute mitosis. *J Biol Chem,* 285**,** 40163-70.
XUE, L., GEAHLEN, R. L. & TAO, W. A. 2013. Identification of direct tyrosine kinase substrates based on protein kinase assay-linked phosphoproteomics. *Mol Cell Proteomics,* 12**,** 2969-80.
XUE, Y., LI, A., WANG, L., FENG, H. & YAO, X. 2006. PPSP: prediction of PK-specific phosphorylation site with Bayesian decision theory. *BMC Bioinformatics,* 7**,** 163.
YAFFE, M. B. & ELIA, A. E. 2001. Phosphoserine/threonine-binding domains. *Curr Opin Cell Biol,* 13**,** 131-8.
YAMAGISHI, Y., HONDA, T., TANNO, Y. & WATANABE, Y. 2010. Two histone marks establish the inner centromere and chromosome bi-orientation. *Science,* 330**,** 239-43.
YANG, W., XIA, Y., HAWKE, D., LI, X., LIANG, J., XING, D., ALDAPE, K., HUNTER, T., ALFRED YUNG, W. K. & LU, Z. 2012. PKM2 phosphorylates histone H3 and promotes gene transcription and tumorigenesis. *Cell,* 150**,** 685-96.
YU, Y., ANJUM, R., KUBOTA, K., RUSH, J., VILLEN, J. & GYGI, S. P. 2009. A site-specific, multiplexed kinase activity assay using stable-isotope dilution and high-resolution mass spectrometry. *Proc Natl Acad Sci U S A,* 106**,** 11606-11.
YU, Z., ZHOU, X., WANG, W., DENG, W., FANG, J., HU, H., WANG, Z., LI, S., CUI, L., SHEN, J., ZHAI, L., PENG, S., WONG, J., DONG, S., YUAN, Z., OU, G., ZHANG, X., XU, P., LOU, J., YANG, N., CHEN, P., XU, R. M. & LI, G. 2015. Dynamic phosphorylation of CENP-A at Ser68 orchestrates its cell-cycle-dependent deposition at centromeres. *Dev Cell,* 32**,** 68-81.
ZANZONI, A., CARBAJO, D., DIELLA, F., GHERARDINI, P. F., TRAMONTANO, A., HELMER-CITTERICH, M. & VIA, A. 2011. Phospho3D 2.0: an enhanced database of three-dimensional structures of phosphorylation sites. *Nucleic Acids Res,* 39**,** D268-71.
ZEITLIN, S. G., SHELBY, R. D. & SULLIVAN, K. F. 2001. CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. *J Cell Biol,* 155**,** 1147-57.
ZENG, L., WANG, W. H., ARRINGTON, J., SHAO, G., GEAHLEN, R. L., HU, C. D. & TAO, W. A. 2017. Identification of Upstream Kinases by Fluorescence Complementation Mass Spectrometry. *ACS Cent Sci,* 3**,** 1078-1085.
ZHOU, L., TIAN, X., ZHU, C., WANG, F. & HIGGINS, J. M. 2014. Polo-like kinase-1 triggers histone phosphorylation by Haspin in mitosis. *EMBO Rep,* 15**,** 273-81.

In review:

SEIBERT, M., KREUGER, M., WATSON, N.A., SEN, O., DAUM, J., SLOTMAN.J, BRAUN. T, HOUTSMULLER, A., GORBSKY, G., JACOB, R., KRACHT, M., HIGGINS, J.M.G., SCHMITZ, L. 2019. CDK1-mediated phosphorylation at the novel histone modification site H2B serine 6 is required for mitotic chromosome segregation. *Journal of Cell Biology*.

Websites:

LINCS 2018, Harvard University, accessed 14 October 2018,
< http://lincs.hms.harvard.edu/db/datasets/20202/results?page=8>
LINCS 2018, Harvard University, accessed 16 October 2018,
< http://lincs.hms.harvard.edu/db/datasets/20047/results>