Roles for Motifs of Cell Cycle Regulating Kinases
Beyond Substrate Selection of Individual Kinases

by

Jes Alexander

B.S., Biology and Chemistry
Stanford University, 2000

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biology
at the
Massachusetts Institute of Technology

May 2008

© 2008 Jes Alexander. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly
paper and electronic copies of this thesis document in whole or in part.

Signature of Author

Department of Biology
May 14, 2008

Certified by

Michael B. Yaffe
Associate Professor of Biology
Thesis Supervisor

Accepted by

Stephen P. Bell
Professor of Biology
Chairman, Graduate Student Committee
Roles for Motifs of Cell Cycle Regulating Kinases
Beyond Substrate Selection of Individual Kinases

By
Jes Alexander

Submitted to the Department of Biology on May 14, 2008 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in Biology

ABSTRACT

Errors in the cell cycle, particularly during mitosis, have recently been implicated in
tumorigenesis and cancer formation. Several protein kinases, including the major mitotic
kinases Cdk1, Aurora A, Aurora B, Nek2, and Plk1, and the Polo-like kinase (Plk)
family, Plk1, Plk2, Plk3, and Plk4, are known to play important roles in the regulation of
mitosis and other phases of the cell cycle to prevent such errors. However, the
mechanisms underlying many of the roles of these kinases are unknown because the
appropriate substrates have not been identified. To aid in substrate identification and to
gain insight into the role of substrate specificity in the regulation of these kinases, we
have determined the motifs of the major mitotic kinases and the Plk family of kinases by
Positional Scanning Oriented Peptide Library Screening (PS-OPLS), which gives
complete, unbiased motifs. The PS-OPLS motif of Plk1 revealed a previously unreported
specificity of Plk1 leading to the identification of new candidate substrates, including
p31/Comet, which is involved in the silencing of the spindle assembly checkpoint for
anaphase onset. Additionally, the motifs and the localizations of the major mitotic
kinases lead us to put forward a model potentially explaining how the major mitotic
kinases, despite having overlapping localizations and access to the same sites on
substrates, phosphorylate distinct sites. This model, if true, would suggest that the motifs
and localizations of all of the major mitotic kinases coordinately direct the group of
kinases to the appropriate substrates. The motifs of the Plk family reveal that Plk2 and
Plk3 phosphorylate sites previously primed by phosphorylation by a tyrosine kinase,
suggesting that these kinases may integrate tyrosine kinase and serine/threonine kinase
signaling in a novel way. Finally, we find that Plk3 has a phosphorylation dependent
toggle switch which changes the motif of the kinase. The result suggests that the motif of
a kinase is not necessarily static, as was previously thought, and that the motif of certain
kinases may be modulated for different roles. These results give us deeper insight into
the function of kinases at the level of motifs, substrates, and the regulation of groups of
kinases.

Thesis Supervisor: Michael B. Yaffe
Title: Associate Professor of Biology
Table of Contents

Abstract
Acknowledgments

Chapter One
The Structural Basis for Kinase Substrate Specificity and the Known Roles of the Major Mitotic Kinases and Polo-like Kinase Family in the Cell Cycle
Kinase Substrate Specificity
The Major Mitotic Kinases
The Polo-like Kinase Family
Summary and Preview
References

Chapter Two
The Determination of the Motifs of the Major Mitotic Kinases: Motifs and Localizations of the Major Mitotic Kinases May Cooperatively Regulate the Kinases as a Group
Abstract
Introduction
Experimental Procedures
Results
Discussion
References

Chapter Three
Determination of the Substrate Specificities of the Polo-like Kinase Family Reveals that Plk3 has a Motif that Changes Depending on the Phosphorylation Status of the Activation Loop
Abstract
Introduction
Experimental Procedures
Results
Discussion
References

Chapter Four
Conclusions and Future Directions
Analysis of Motifs of Kinases Related by Function Versus those Related by Sequence
Summary and Future Directions
References
Acknowledgments

I would like to thank all of the members of the Yaffe lab for their technical and intellectual support, for generally making the lab a good place to work, and for putting up with me over the years. I would like to particularly thank the other graduate students in the lab, past and present, Sarah Bissonette, Drew Lowery, Duaa Mohammad, and Andrea Tentner, who I have had the good fortune to get to know as we suffered through graduate school together. Their technical, intellectual, and emotional support and friendship have been immensely helpful—Sarah Bissonette for doing and saying what needed to be done and said, Drew Lowery for knowing every detail published about the cell cycle and for sharing that information with me during discussions about my work, and Duaa for being a good friend, discussing my work with me, providing technical help, and for entertaining me with stories about her life. I would also like to thank all the postdocs in the lab, past and present, Chris Ellson, Scott Floyd, Alexandra Gardino, Jeong-Ho Hong, Kazu Kishi, Dan Lim, Rune Linding, Coky Nguyen, Gerry Ostheimer, Michael Pacold, Christian Reinhardt, Erik Wilker, and Marcel van Vugt, for their time and effort in teaching me and creating an environment conducive to learning— particularly, Dan Lim for intellectual and technical support and putting up with my ranting, Christian Reinhardt for working with me on the phospho-dependent binding domain screen and his extra patience and effort in teaching me, Michael Pacold and Scott Floyd for career advice, Chris Ellson for being a good friend, discussing my work with me, telling the truth when it needed to be said, giving technical assistance (also known as cloning), and being generally supportive. Special thanks goes to Mary Stewart who keeps the lab running and without whom I would never have been able to complete a single experiment. Finally, I would like to thank Mike for hiring some good people, procuring money to fund my experiments, and for teaching me what to do and what not to do if I am ever as fortunate as he has been and have the opportunity to head my own lab.

Outside the lab, I would like to thank Ben Turk for helping us get the PS-OPLS technique running in the Yaffe lab, Karl Clauser in Steve Carr’s lab for his valiant efforts in attempting to find Plk3 substrates, and Eric Weiss for a fruitful collaboration. I would also like to thank my prelim and thesis committee including Steve Bell, Paul Chang, Ray Erikson, Frank Gertler, Terry Orr-Weaver, and Peter Sorger for their time and effort, helpful suggestions, and for pushing me to become a better scientist through their insightful and needed critique. Additionally, my undergraduate research advisor Doug Brutlag and my research advisor at the NCI, John Weinstein, were so encouraging. I would like to thank them for their role in getting me to this point.

Most importantly, I would like to thank all the people outside of the biology community who have been so supportive and have made me happy despite all my painful experiences in graduate school. I would like to thank the University of Kansas men’s basketball team who won the NCAA National Championship this year (2008). Following them for the last twenty years through the failures and disappointments to their run through the tournament and to the championship this year has not only made me happy, but has been a huge inspiration to me. More importantly, my parents and sister have been extremely understanding and have supported me in every way they could, as they have in all my endeavors. And last, but far from least, my biggest debt of gratitude is to my wife, Alexis, who, despite being a medical resident, has found time to support me through the difficult times in graduate school and in life.
Chapter One

The Structural Basis for Kinase Substrate Specificity and the Known Roles of the Major Mitotic Kinases and Polo-like Kinase Family in the Cell Cycle

Contributions:
I wrote all sections. I created Figures 1.1, 1.6, and 1.7. Other figures were adapted from the sources cited.
The goal of the eukaryotic cell cycle is the replication of the genetic material and division of the genetic material and other contents of the cell to two daughter cells. The cell cycle is divided into four phases, G1, S, G2, and M. M phase includes two separate, but coordinated processes, mitosis and cytokinesis. M-phase is particularly complicated as a series of events during mitosis, including centrosome separation and maturation, spindle assembly, chromosome segregation, and spindle disassembly, and during cytokinesis must occur in rapid succession without any errors. Errors in M-phase result in genetic instability and chromosomal aberrations, which are hallmarks of cancer and may be initiating events in tumorigenesis (1, 2).

To prevent errors, the events within each phase of the cell cycle are regulated by, among other things, the carefully timed expression and destruction of proteins, formation of specific protein-protein interactions, and transduction of signals. Protein phosphorylation, performed by protein kinases, is the most common mechanism employed by the cell for signal transduction and the cell cycle is not an exception. There are an estimated 700,000 potentially phosphorylatable sites at any given time in a eukaryotic cell (3). Each protein kinase chooses specific sites on specific substrates. While the actual number of sites may be fewer than the estimated 700,000 since some of these sites may not be surface accessible, each kinase must select which sites to phosphorylate from a large number of inappropriate sites. Clearly, the substrate specificity of each kinase is important for the fidelity of processes such as the cell cycle.

Understanding the regulation and substrate specificity of kinases involved in the cell cycle and, in particular mitosis, should give insight into the process of tumorigenesis and may lead to modes of prevention or treatments of cancer (1, 2). This thesis examines the substrates specificities of two sets of kinases (Figure 1.1). The major mitotic kinases, Cdk1, Aurora A, Aurora B, Nek2, and Plk1, are a functional family with all members playing significant roles regulating the processes of mitosis (4). The other set of kinases, the Polo-like kinase (Plk) family, which includes Plk1, Plk2, Plk3, and Plk4, are related by sequence, but have functions associated with different phases of the cell cycle (5). Knowledge of the substrate specificity of these kinases will aid in the discovery of their substrates, allowing a deeper understanding of the roles of these kinases in the cell cycle. Additionally, the substrate specificities of these sets of kinases may give us other insights
Figure 1.1: Roles of the Major Mitotic Kinases and Polo-like Kinases in the cell cycle. Each kinase is represented by a colored dot as indicated on the right side of the figure. Schematics of a cell at each stage of the cell cycle are represented and the phases are color coded, orange for M-phase and red for G1-, S-, and G2-phases. Spindle Assembly Checkpoint is abbreviated SAC and microtubule is abbreviated MT.

into mechanisms of substrate specificity involved in the regulation of the processes in which these kinases function and how substrate specificity differs between kinases related by function and those related by sequence.

**Kinase Substrate Specificity**

There are 518 protein kinases in the human genome (6). These kinases can be divided into seven major groups based on sequence similarity with many kinases not falling into any of these groups. The major mitotic kinases, which cooperate functionally
Figure 1.2: Phylogeny of human kinases based on sequence similarity of the kinase domains with the seven major groups of kinases labeled in small lettering. A. the Major Mitotic Kinases, Cdk1, Aurora A, Aurora B, Nek2, and Plk1 are labeled in large lettering and indicated by red circles and B. The Polo-like kinase family, Plk1, Plk2, Plk3, and Plk4, are indicated by the red oval. This figure was adapted from (6).
to regulate mitosis, show significant sequence diversity and do not cluster to a single phylogenetic class. Cdk1 is a member of the CMGC kinases, Aurora A and Aurora B are very similar to AGC kinases, Nek2 is most similar to the CAMK group, and Plk1 is most similar to the AGC kinases (Figure 1.2A). In contrast to the major mitotic kinases, members of the Polo-like family of kinases, which seem to have different roles in distinct parts of the cell cycle, have very high sequence similarity and cluster in the phylogeny near the AGC kinases (Figure 1.2B). Despite the clustering of kinases into different groups based on sequence, all protein kinases catalyze the same reaction: transfer of the \( \gamma \)-phosphate from a molecule of ATP to the hydroxyl group of a serine, threonine, or tyrosine residue of the substrate protein. For phosphorylation of a substrate to occur, the site to be phosphorylated on the substrate must bind in the active site cleft of the kinase while the kinase is active. For most kinases, the phosphorylation reaction is regulated at a minimum of three levels: kinase activation, recognition of the sequence surrounding the phosphorylatable residue by the active site cleft, and co-localization of the kinase with the substrate. These aspects are discussed herein.

**Kinase Domain Structure and the Structural Basis of Kinase Activation**

The first X-ray crystal structure of a kinase domain was that of the AGC kinase group member cAMP-Dependent Protein Kinase also known as Protein Kinase A (PKA) (7). Subsequent crystal structures of other kinase domains have shown that the protein kinase domain fold is highly conserved (7-9). Kinase domains are bi-lobed with a smaller NH\(_2\)-terminal lobe (N-lobe) and a larger COOH-terminal lobe (C-lobe) (Figure 1.3). The N-lobe is primarily responsible for positioning ATP for \( \gamma \)-phosphate transfer and consists of a five-stranded \( \beta \)-sheet (composed of strands \( \beta 1 \) to \( \beta 5 \)) and two \( \alpha \)-helices including \( \alpha C \). The loop between \( \beta 1 \) and \( \beta 2 \) is the phosphate binding loop (P-loop). The C-lobe is primarily responsible for peptide substrate binding and phospho-transfer. It is mostly \( \alpha \)-helical containing six \( \alpha \)-helices including \( \alpha EF \) and \( \alpha F \) and four small \( \beta \)-strands including \( \beta 6 \) through \( \beta 9 \). The loop between \( \beta 6 \) and \( \beta 7 \) is the catalytic loop directly involved in phosphate transfer and the loop between \( \beta 8 \) and \( \beta 9 \) is the \( \text{Mg}^{2+} \) binding loop. Between the N-lobe and C-lobe is the active site cleft, where peptide and ATP are bound and catalysis occurs.
Several residues are invariant because of their important roles in catalysis (Figure 1.3 and 1.4). Using the amino acid numbering scheme for PKA, the ATP phosphates are positioned by a glycine rich region in the P-loop; a Mg\(^{2+}\) ion, which bridges the \(\beta\)- and \(\gamma\)-phosphates and is stabilized by D184 in the Mg\(^{2+}\) loop; and K72 in \(\beta3\), which forms an hydrogen bond with \(\beta\)- and \(\gamma\)-phosphates and is itself localized by a salt-bridge interaction with E91 on \(\alpha\)C. The Mg\(^{2+}\) ion interacts with the \(\beta\)- and \(\gamma\)-phosphates of the bound ATP molecule and is required for phospho-transfer. D166 is the residue directly involved in phosphate transfer. D166 was originally thought to act as a base to remove the proton from the hydroxyl group on the substrate serine or threonine or from the phenol group on the substrate tyrosine allowing the resulting alkoxide or phenoxide anion to perform a nucleophilic attack on the ATP \(\gamma\)-phosphate resulting in phospho-transfer (10, 11). However, the current thinking is that D166 is not only a base catalyst, but
Figure 1.4: Schematic representation of the invariant catalytic residues of PKA, also showing the RD arginine R165. This figure is from (11).

Instead may play a major role in positioning the substrate serine or threonine hydroxyl or tyrosine phenol for attack on the ATP γ-phosphate (12). D166 itself is positioned by hydrogen bonding interactions from N171 also from the catalytic loop.

Most kinases can attain at least two conformational states, a minimally active state and a maximally active state. Conformational changes in the activation segment cause the kinase to switch between these states by positioning the catalytic residues
appropriately for phospho-transfer (7, 13, 14). The activation segment runs from a highly conserved tripeptide, 184-DFG-186, in the middle of the Mg$^{2+}$ binding loop to another conserved tripeptide, 206-APE-208, which is at the junction of the N-terminus of $\alpha$EF and the loop preceding it, the P+1 loop (Figure 1.5). Within the activation segment, in order from N-terminal to C-terminal, are the Mg$^{2+}$ binding loop, $\beta$9, the activation loop, and the P+1 loop, named for its interaction with the amino acid one position C-terminal to the phospho-acceptor residue, although it interacts with other positions as well.

In many kinases, activation requires phosphorylation on serine, threonine or tyrosine residues on the activation loop (15). This phosphorylation can occur through autophosphorylation or by phosphorylation by a different kinase. In the unphosphorylated state, the activation segment often collapses into the active site cleft blocking access to ATP and the peptide substrate. Phosphorylation of the activation loop can occur at one or more residues depending on the kinase, but in most cases one of the sites plays a dominant role (T197 in PKA) in the conformational change. Phosphorylation on the activation loop causes refolding of the activation segment by formation of interactions between the phospho residue and a group of amino acids composing a basic pocket (7, 13, 14). In PKA, pT197 interacts with H87 on $\alpha$C, R165
Figure 1.6: Structure of Protein Kinase A (PKA) (PDB ID: 1ATP) showing the interaction between the phosphoresidue on the activation segment and the basic pocket. Basic pocket residues (green) and the activation segment phosphothreonine (yellow) are shown in stick representation. The activation segment (red) and the rest of the protein (white) are shown in the backbone representation. In this figure, the N-lobe as at the top and the C-lobe is at the bottom and the structure, as compared to Figure 1.3, has been rotated 180° around an axis running from N-lobe to C-lobe.
on the catalytic loop, and K189 on β9 (Figure 1.6). All kinases that require phosphorylation for activation have an arginine (R165) just preceding the invariant catalytic aspartate (D166) (Figure 1.3, 1.4, 1.6) (11). Kinases with this RD sequence are known as RD kinases. Interaction between pT197 and R165 seems to have little effect on the positioning of the catalytic loop, but is likely to be important for positioning of the activation segment and in particular the DFG tripeptide (15). Interaction with K189 also seems to be involved in positioning the DFG tripeptide (7, 13, 14). Interaction with H87 plays a role in positioning the N- and C-lobes with respect to each other through positioning of the αC. In the inactive state of the kinase, electrostatic repulsion between the amino acids of the basic pocket is thought to prevent the activation segment from achieving the active state conformation (15). Neutralization of the basic pocket by interaction with the dianionic pT197 allows the activation segment to take its active state conformation and facilitates activation of the kinase.

Upon phosphorylation, the conformational changes associated with the interaction of the phosphoresidue with the basic pocket are propagated through the kinase resulting in the activation segment moving out of the active site cleft, repositioning of the DFG tripeptide sequence, and positioning of the catalytic residues allowing efficient phospho-transfer. Repositioning of the 184-DFG-186 tripeptide localizes D184 for chelation of the Mg$^{2+}$ ion which is required for positioning of the phosphates of ATP for catalysis (7, 11, 13, 14, 16). Additionally, F185 forms hydrophobic interactions with amino acids on αC, which are reinforced by interactions between the pT197 and H87 on αC. These interactions with αC are important for positioning of the invariant catalytic E91, which is on αC and forms an ionic interaction with the invariant catalytic K72. As mentioned above this interaction is required to localize K72 for its role in positioning the phosphates of ATP. In serine/threonine kinases, phosphorylation on the activation loop also results in a hydrogen bonding interaction between the P+1 loop conserved serine or threonine (T165) and residues in the catalytic loop. Though not part of the activation loop, the loop between αEF and αF seems to play a role in stabilization of the conformation of the activation loop in the active state of the kinase. Together these conformational changes activate the kinase.
Residues other than the RD arginine of the basic pocket are not as highly conserved, even in kinases that require phosphorylation for activation (15). For example, in Aurora A, another kinase activated by phosphorylation on the activation loop, the equivalent residue to K189 is not basic (17). Instead, a lysine on the activation loop interacts with the activation loop phosphorylation sites. The basic pocket itself, however, is highly conserved, even in kinases not requiring phosphorylation for activation. Glycogen Synthase Kinase 3 beta (Gsk3β) and Casein Kinase 1 (CK1) are examples (18-20). The amino acid in the equivalent position to T197 is a valine in Gsk3β and an asparagine in CK1, making phosphorylation on the equivalent site impossible. GSK3β has a basic pocket composed of the RD arginine, a residue from αC, and a residue from β9; CK1 has a basic pocket composed of the RD arginine and a residue from β9. The crystal structures of both kinases have been solved and in both cases the basic pocket was found to be occupied by an anion from the crystallization solution. Crystal structures of Gsk3β have also been solved in which the basic pocket is empty (21). In these structures, both the activation segment of Gsk3β and the basic pocket are in the active conformation confirming that Gsk3β is active without phosphorylation on the T197 equivalent site or an anion in the basic pocket. The structure also implies that the basic pocket is open to be filled by an anion. Both of these kinases phosphorylate substrates that have previously been phosphorylated (primed) by another kinase on a nearby residue. Gsk3β phosphorylates a serine or threonine four amino acids N-terminal to a phosphoserine (22). CK1 phosphorylates a serine or threonine three amino acids C-terminal to a glutamate, aspartate, phosphoserine or phosphothreonine (23). Taken together, these data suggest that a phoshoresidue on the substrate can fit into the basic pocket, implying that the basic pocket is involved in substrate selection rather than kinase activation in these kinases. It should be noted that no kinase has been described for which the required priming phosphorylation is catalyzed by a tyrosine kinase.

Structural Basis for Kinase Substrate Specificity

The original studies of kinase substrate specificity were performed on PKA in the mid-to-late 1970's and showed that primary structure of the substrate was primarily responsible for the choice of sites phosphorylated (24). At each position near the
phosphoacceptor residue, certain amino acids increase the likelihood that a kinase will phosphorylate the site while others reduce the likelihood. The consensus sequence is a way of representing which sites a kinase is likely to phosphorylate. There are two types of consensus sequences, minimal consensus sequences and optimal consensus sequences. The minimal consensus sequence delineates which amino acids in which positions are required for phosphorylation. The optimal consensus sequence delineates which amino acids in which positions result in a site that is highly favored by the kinase. In either case, each position of the sequence is represented as a list of amino acids separated by slashes ("/"). The amino acids in each list are considered to have similar effects on the likelihood that the kinase will phosphorylate the site. An X is used when there is no significant difference among the amino acids in their effect on the likelihood of phosphorylation. A motif is a broader concept including consensus sequence, but expanding it to include all residues in each position affecting the likelihood of phosphorylation, both positively and negatively. The PKA optimal consensus sequence was determined to be R-R-X-S/T-ϕ, where X is any amino acid and ϕ is a hydrophobic amino acid (24). By convention, the phosphoacceptor site (S/T) is position 0, positions N-terminal to this are numbered with more negative numbers going away from position 0 (e.g. X is -1) and positions C-terminal to the phosphoacceptor are numbered with more positive numbers going away from position 0 (e.g. ϕ is +1).

An understanding for how specificity is mediated at the molecular level was not gained until kinetic studies on PKA alanine scanning mutants were performed and the X-ray crystal structure of PKA in complex with the PKI inhibitor peptide was solved in the early 1990's (25-27). This structure has served as a common reference point for the analysis of all subsequent kinase structures. The sequence of the inhibitor in the region of interaction with the active site cleft of the kinase, with the exception of an alanine instead of a serine or threonine at what would be the phosphoacceptor site, is the same as the optimal sequence. The crystal structure of the complex shows that the substrate binds in an extended conformation in the active site cleft, explaining the importance of primary structure (Figure 1.2, Figure 1.7), rather than secondary or tertiary structure, in specificity. Additionally, there are clear interactions between amino acids composing the surface of the active site cleft and the amino acids of the inhibitor peptide, particularly the
Figure 1.7: Structure of PKA in complex with the PKI peptide (PDB ID: 1ATP) showing the amino acids in the PKA active site cleft that interact with the amino acids in the bound pseudo-substrate peptide. PKA is shown in the space filling representation. Amino acids E127, E170, and E230, which interact with the arginines in the -3 and -2 positions on the peptide are shown in green, as are L198 and L205, which interact with the isoleucine in the +1 position of the peptide. The rest of the kinase is shown in white. The peptide (orange) is shown in the backbone representation, except for the two arginines and the isoleucine mentioned above, which are shown in stick representation.
arginines at positions -3 and -2 and the hydrophobic amino acid at position +1. The R in the -3 position forms electrostatic interactions with E127 and E170 on the surface of the active cleft formed by the N-lobe and the R in the -2 position forms an electrostatic interaction with E230 in αF in the C-lobe. The hydrophobic amino acid in the +1 interacts with a hydrophobic pocket formed by L198 and L205 in the P+1 loop of the activation segment.

Interestingly, Cyclin Dependent Kinases (CDKs) and Mitogen Activated Protein Kinases (MAPKs), members of the CMGC group, have a very different specificity in the +1 position. These kinases only phosphorylate sequences with a proline in the +1 position. This is in contrast to kinases of the AGC and CAMK groups, Aurora A kinase, as well as probably others, which will not phosphorylate any sequence with proline in the +1 position. The crystal structure of activated Cdk2 in complex with Cyclin A3 and a substrate peptide has been solved and seems to explain the rare specificity of proline-directed kinases (8). Proline is unique in that its backbone nitrogen cannot hydrogen bond due to the cyclic nature of the amino acid. The cyclic nature of the amino acid also results in kink in any peptide chain containing proline. The authors suggested that any other amino acid in the +1 position would result in an unsatisfied hydrogen bond due to a lack of a hydrogen bonding partner from the amino acids of the kinase in the vicinity of the +1 position. Such an unsatisfied hydrogen bond would make any amino acid other than proline much less favorable in the +1 position. However, this selectivity may also be due to the kink in the peptide chain caused uniquely by proline.

Subsequent crystal structures of other kinases with selectivity in the -3, -2, and +1 positions have demonstrated the role of residues homologous to the ones described above in the selectivities of these kinases (28). The structures have also shown that for most kinases, only residues within four or five positions on either side of the phosphoacceptor site on the substrate interact with the active site cleft, which explains why the residues involved in kinase-substrate recognition are commonly found within four of five positions of the phosphoacceptor (3, 8, 27). However, selectivity at positions farther out can occur through interactions with the outer surface of the kinase domain (27). Other determinants of specificity also exist for specific kinases. As mentioned previously, priming phosphorylation can interact with the basic pocket in the active site cleft of
certain kinases. In general, the active site cleft is complementary to the substrate sequence allowing interactions based on charge, hydrogen bonding, or hydrophobicity, and this complementarity defines the kinase specificity.

Techniques for Kinase Motif Determination

The original approach used to determine the specificity of PKA employed kinetics studies on a series of peptides based on a site found to be phosphorylated by PKA in pyruvate kinase (24, 29). The peptides were designed such that each had a different single amino acid mutation. This technique employs the comparison of affinity and reaction velocity for each of the different peptides, from which the effect on specificity of different amino acids in each position can be inferred. This technique is particularly useful for determining a minimal consensus sequence, as this just requires alanine scanning. It is less useful if a substrate site is not known as a site must be determined prior to employing this technique. It is also less useful if the goal is to determine an optimal consensus sequence or a motif, as to do this correctly requires kinetic studies on peptides with single mutations where the mutations scan through each amino acid at each position of the sequence. This can amount to hundreds or thousands of experiments. If all amino acids are not tested in each position, the resulting optimal consensus sequence or motif may be incomplete and, therefore, biased.

To overcome these problems the technique of oriented peptide library screening (OPLS) was developed in the mid 1990’s (30). In this technique, an in vitro kinase assay is performed on a degenerate library of peptides instead of a peptide with a single sequence. The library consists of peptides of the same length each, with a fixed phosphoacceptor residue in the center, which orients the library. All other positions in the library are composed of a mixture of amino acids. There are usually four to seven such degenerate positions on each side of the phosphoacceptor and each usually includes an equal mixture of all 20 amino acids except for serine, threonine, and tyrosine, as these would disorient the library, and cysteine, as this might result in disulfide cross-links between peptides. After the reaction with the kinase, those peptides that were phosphorylated are separated from the unphosphorylated peptides and sequenced by Edman Degradation. The percentage of each amino acid at each position of the
phosphorylated peptides is determined from these results. The ratio of the percentage of a given amino acid in a given position to the percentage of that amino acid in that position from the original library gives the relative enrichment of that amino acid and is a measure of the selectivity of the kinase for that amino acid in that position. To further examine specificity, the same technique can be repeated with a library containing, in addition to the fixed phosphoacceptor residue, a second fixed position corresponding to the highest selectivity determined with the previous library. This amplifies the selectivity ratios of less highly selected positions to better distinguish between signal and noise.

OPLS advantages include not requiring a known substrate site, requiring much less work to achieve an unbiased motif, and determination of selectivity at each position independently of all other positions. This technique has been successfully employed on many kinases (30-36). However, it has a few important drawbacks. The process of sequencing is error prone because of preview and lag, which are caused by excessive or incomplete peptide cleavage during Edman Degradation, as compared with the bulk of the peptides. When this occurs around a position of high selectivity for a given amino acid, the positions just prior to (preview) and just after (lag) will often show high selectivity for the same amino acid. As a result some subjectivity is required to decide which positions are showing real selectivity and which are showing artifacts of the sequencing reaction. This problem extends to positions in which certain amino acids are strongly selected against. Preview and lag will have the effect of covering up the extent of negative selection, making interpretation of residues mediating negative selection difficult and somewhat unreliable. Finally, this technique does not allow for testing the role of priming phosphorylation in specificity. To do this would require the addition of phosphoamino acids to the peptide library. However, phosphoamino acids would make it impossible to distinguish between the peptides phosphorylated by the kinase and those not phosphorylated by the kinase since all peptides would have phosphoresidues.

A more advanced technique based on OPLS, positional scanning oriented peptide library screening (PS-OPLS), which was published in 2004, addresses the deficiencies in OPLS (37). In this technique instead of a single in vitro kinase assay on single oriented peptide library, in vitro kinase assays on 198 separate oriented peptide libraries are performed in parallel (Figure 2.2). The stock peptide libraries are arranged in an array
with 9 rows and 22 columns. Each peptide library is biotinylated and has two fixed positions, the phosphoacceptor position and a position for testing kinase specificity. Scanning across the columns puts all 20 amino acids plus phosphotyrosine and phosphothreonine in the second fixed position. Scanning down the rows moves the second fixed position from -5, relative to the phosphoacceptor S/T, to +4. At the intersection of a column and a row is a library with the column’s designated amino acid fixed at the row’s designated position and with all other positions degenerate except the S/T phosphoacceptor site. The degenerate positions contain a mixture of all 20 amino acids except cysteine. The kinase reactions are also arranged in 9 rows and 22 columns and the libraries are added to the reactions in parallel maintaining their positions in the array. The kinase assays are performed in the presence of $\gamma$-32P-ATP. After the assays are performed, a portion of each reaction is transferred in parallel to a streptavidin membrane to which the peptides bind. The membrane is washed to remove the unreacted $\gamma$-32P-ATP and then visualized by autoradiogram or phosphorimager. The darker the spot at the intersection of a given column and given row, the more the substrate binding cleft selects for the amino acid designated by the column at the position designated by the row. The resulting data can be quantitated by densitometry.

PS-OPLS has been used successfully on several kinases (31, 37-44). A drawback of the technique is that it requires significantly more kinase protein than is required for OPLS since 198 reactions are being performed instead of just one. However, PS-OPLS does not have the problems associated with the sequencing artifacts of OPLS. Therefore, no subjectivity is required in determining which amino acids in which positions are selected, and, importantly, this technique also allows the reliable determination of negatively selected amino acids. In addition, since peptides phosphorylated in the kinase reactions do not need to be separated from the unphosphorylated peptides, phosphoamino acids can be included in the libraries. This allows testing for a role of priming phosphorylation in specificity. Based on these advantages, we chose to use PS-OPLS for the motifs determined in this work.
Other Determinants of Substrate Specificity

Whereas the motif of a kinase decides which sites on a substrate can be phosphorylated and, therefore, makes the final determination as to whether a site will be phosphorylated, phosphorylation requires that the kinase co-localize with the substrate while the kinase is active. This co-localization requirement is particularly important since the binding of peptides in the kinase active site cleft has dissociation constants often in the hundreds of micromolar to millimolar range (45). To compensate for low affinity, either high local concentrations of the kinase in the vicinity of the substrate or secondary binding sites are necessary to enable phosphorylation. To increase local concentrations, determinants of substrate specificity other than motifs are often employed.

Different kinases use different determinants. Some kinases recognize substrate docking motifs, which are separate from the site to be phosphorylated, through kinase domain interaction surfaces, which are separate from the active site cleft. MAPKs recognizing D domains on their substrates are examples of this (46). Some kinases have targeting domains. Non-receptor tyrosine kinases often contain Src Homology 2 (SH2) and/or Src Homology 3 (SH3) domains, which bind phosphotyrosine containing sequences and proline-rich sequences in substrates, respectively (47). Polo-like kinase 1 (Plk1) has a phospho-dependent binding domain, the Polo-Box domain (PBD), that binds S-pS/pT motifs (48). Two models have been proposed for the role of the PBD in Plk1 substrate interactions (49). In the Processive model, the PBD directly binds Plk1 substrates; in the Distributive model, the PBD binds a protein which situates Plk1 in close proximity to other proteins which are substrates. Evidence for both modes of substrate recognition exist (49). Some kinases have targeting subunits. The targeting subunits of Cdns are Cyclins, with which they form stable heterodimers, as their name suggests. Cyclin A directs Cdk2 to many of its substrates through interactions between a hydrophobic patch on the Cyclin and an RXL motif on these substrates (50). Some kinases are localized to particular subcellular compartments or structures. Interaction through an RXL motif seems to be dispensable for substrate recognition by Cdk1/Cyclin B (51). However, Cyclins are also responsible for Cyclin/Cdk complex localization. Cdk1/Cyclin B1 is excluded from the nucleus prior to mitosis by a nuclear export sequence (NES) and cytoplasmic retention sequence (CRS) on Cyclin B1 (52).
early mitosis, phosphorylation on the CRS inactivates it causing rapid translocation of Cdk1/Cyclin B1 into the nucleus where its substrates are (53-55). In all cases, the local concentration of kinase in the vicinity of substrate is increased, satisfying the spatial co-localization requirement and elevating the likelihood of phosphorylation (3).

**The Major Mitotic Kinases**

Mitosis is the part of the cell cycle in which the nuclear contents, particularly the DNA which was replicated in S-phase, is divided between the emerging daughter cells. This process relies on two major structures, the centrosome and the mitotic spindle. The centrosome is the microtubule organizing complex (MTOC) and during the cell cycle it undergoes replication in preparation for its role in mitosis. Prior to mitosis, cells contain a duplicated centrosome consisting of two mother-daughter pairs of centrioles tethered together and surrounded by pericentriolar material (PCM), which nucleates microtubules. At the G2/M transition, the tether is severed and the centrosomes begin to move apart to opposite ends of the cell. During this process, the centrosomes mature, meaning that proteins such as γ-tubulin and other regulators of microtubule assembly are recruited to the PCM to enable the formation of the mitotic spindle. In the first phase of mitosis, prophase, the chromosomes condense. Prometaphase begins with breakdown of the nuclear envelope allowing attachment of the chromosomes to the newly formed mitotic spindle. Microtubules of the mitotic spindle attach to kinetochores, protein complexes which bind to the centromeric DNA regions on the chromosomes. The kinetochore on each chromosome must attach to microtubules from each centrosome to form bivalent or amphitelic attachments. Until this occurs, the spindle assembly checkpoint prevents chromosome segregation. When all kinetochores have formed attachments to microtubules from each centrosome, the chromosomes align on the metaphase plate and the cell has reached metaphase. At this point, the E3 ubiquitin ligase anaphase promoting complex/cyclosome (APC/C) ubiquitinates Securin, which holds the sister chromatids together, and Cyclin B causing their destruction and allowing anaphase onset. During anaphase, the sister chromatids move toward the centrosomes as the centrosomes move apart. In telophase, the chromosomes reach the opposite ends of the cell and decondense and the nuclear envelope reassembles forming the two new nuclei of the daughter cells.
The major mitotic kinases, Cdk1, Aurora A, Aurora B, Nek2, and Plk1 regulate these events through carefully timed phosphorylation of specific sites on substrates. Each of these events is regulated by more than one of these kinases and these kinases have overlapping localizations, often even interacting with and phosphorylating the same substrate proteins on different residues. What follows are brief descriptions of what is known about each of these kinases including how they are activated, where they localize, what their substrate specificity is, and what their known roles are with emphasis on known substrates.

Cdk1

Cyclin-Dependent Kinases (Cdks) are dependent on cyclins for their activation and targeting, as their name suggests. Cyclin-Cdk complexes are the master regulators of the cell cycle with different cyclin-Cdk complexes controlling different parts of the cell cycle. This regulation is based on carefully timed expression, activation, and destruction of the cyclins and other regulators of these complexes. Vertebrates have the G1-phase cyclin D, the G1/S-transition cyclin E, the S-phase cyclin A, and the M-phase cyclin B. Cdk1, 2, 4, and 6 have the same kinase motif, S/T-P-X-K/R, but K/R in the +2 seems not to be necessary, particularly for Cdk1 (8, 30, 36). Unlike other Cyclin-Cdk complexes that also rely on an interaction between a hydrophobic patch on the cyclin and an RXL motif on the substrate, Cdk1/Cyclin B seems not to require this for substrate selection (51). The Cdk1/Cyclin B complex is the master regulator of mitosis and seems to play a role in almost every aspect of mitosis.

Activation of Cdk1/Cyclin B, which occurs at the centrosome, marks entry into mitosis (56). Cyclin B expression begins in S-phase and protein levels peak at the beginning of mitosis (57). Cyclin B binds to Cdk1 on the aC helix which contains the PSTAIRE sequence recognized by cyclins for binding. This binding results in a conformational change to a partially active state. Phosphorylation on the activation loop at T160 by the Cdk activating kinase (CAK) is also required for full activation of the kinase (8, 58). However, prior to mitosis, Cdk1 is inhibited by a second set of phosphorylations on T14 and Y15 of the P-loop generated by Myt1 and Wee1 kinases (59, 60). These phosphorylations inhibit the kinase by blocking access to the ATP
binding pocket through electrostatic repulsion with the phosphates of ATP. At the onset of mitosis, these phosphorylations are removed by the phosphatases Cdc25B and Cdc25C (61). A small amount of active Cdkl/Cyclin B phosphorylates Wee1, Cdc25B and Cdc25C decreasing the activity of Wee1, increasing the activities of Cdc25B and Cdc25C, and generating a positive feedback loop which rapidly activates the rest of the pool of Cdkl/Cyclin B (62, 63). As discussed later, Plk1 also contributes to mitotic entry through a role in generating this positive feedback loop.

Once active, Cdkl/Cyclin B activity plays a role in centrosome separation and the resultant bipolar spindle. At least one substrate of Cdkl/Cyclin B involved in this process is known. HsEg5 is a kinesin-related motor protein that associates with the centosome in mitosis, but not in interphase. Its localization to the centrosome at mitosis is required for centrosome separation. Disruption of its localization results in monastral spindles. Cdkl/Cyclin B triggers localization to the centrosome by phosphorylation on a conserved site in the C-terminus (64).

Relocalization of Cdkl/Cyclin B allows it to regulate other mitotic processes. There are three B-type cyclins. Cyclin B1 and B2 are ubiquitously expressed, whereas Cyclin B3 seems to be restricted to the testis (65, 66). Cdkl binds Cyclin B1 and Cyclin B2 to form two separate complexes that perform different roles in mitosis based on their different localizations (67). Prior to mitosis, both Cdkl/Cyclin BI and Cdkl/Cyclin B2 are sequestered in the cytoplasm by a cytoplasmic retention sequence (CRS) and a nuclear export sequence (NES) (52). Cdkl/Cyclin B2 remains outside the nucleus first localizing to the golgi where it phosphorylates GM130 for mitotic golgi fragmentation (67, 68). Later in prophase, Cdkl/Cyclin B2 localizes diffusely in the cytoplasm (69). Cdkl/Cyclin B1 becomes phosphorylated on its NES and CRS causing abrogation of nuclear export and enhanced nuclear import resulting in translocation of Cdkl/Cyclin B into the nucleus (53-55). Once inside the nucleus, Cdkl/Cyclin B1 has several different roles. It phosphorylates lamins to cause nuclear envelope breakdown (70, 71). It also localizes to chromatin, spindle microtubules, and unattached kinetochores (72). Cdkl/Cyclin B1 is thought to play a role in chromatin condensation (73). It phosphorylates histone H1, which was thought to mediate its role in chromatin condensation, but the true significance of this phosphorylation has turned out to be less
Cdk1/Cyclin B1 phosphorylates MAP4 and Stathmin/Opl8 to regulate spindle microtubule dynamics (74-76). Additional substrates likely exist at the kinetochore, but their identities are not known; Cdk1/Cyclin B1 localization specifically to unattached kinetochores, however, suggests it may have a role in amphitelic attachment of chromosomes to the spindle (72).

Another role of Cdk1/Cyclin B is to prepare cells for anaphase and exit from mitosis through its own destruction. The E3 ubiquitin ligase, Anaphase Promoting complex/Cyclosome (APC/C), ubiquitinates mitotic proteins for destruction by the proteasome. The APC/C can bind two WD40-repeat containing proteins that change its substrate specificity, Cdc20 and Cdh1 (77). Cdc20 targets the APC/C to proteins containing a sequence known as the destruction box (D-box). Cdh1 targets the APC/C to proteins containing a D-box or proteins containing a sequence known as the KEN box. Cyclin-Cdk phosphorylation of Cdh1 prevents it from interacting with the APC/C (78). At the beginning of mitosis, Cdk1/Cyclin B is fully active and functions with Plk1 to phosphorylate subunits of APC/C to activate it (79). Degradation of D-box containing proteins then begins with Cyclin A shortly after nuclear envelope breakdown in prophase (80). Destruction of Cyclin B and Securin, which also contain D-boxes, does not occur until after all chromosomes are aligned on the metaphase plate with amphitelic attachments and the spindle assembly checkpoint has been silenced (81, 82). Securin is an inhibitor of sister chromatid separation. Destruction of Cyclin B and Securin allows chromosome segregation in anaphase. The destruction of Cyclin B inactivates Cdk1, allowing Cdh1 to bind the APC/C, thus, changing the APC/C substrate specificity to allow other mitotic proteins to be destroyed. Geminin, an inhibitor of DNA replication, is ubiquitinated by the APC/C$^{\text{Cdh1}}$ to allow relicensing of origins of replication (83). Inactivation of Cdk1 is required for mitotic exit, cytokinesis, chromatin decondensation, and reformation of the nuclear envelope (4, 84).

**Aurora A**

Aurora Kinases are conserved from budding yeast to humans. Budding yeast have one Aurora, Ipl1, whereas mammals have three, Aurora A, B, and C. Aurora B is
the ortholog of Ipl1 and will be discussed below. Aurora C is normally expressed only in the testis, playing a role in meiosis (85).

Aurora A is ubiquitously expressed beginning in S-phase, through G2-phase, and into mitosis with protein levels and activity peaking shortly after mitotic entry (86). It localizes to the centrosome in S- and G2- phases and to both the centrosome and to a portion of the spindle microtubules proximal to the centrosome in mitosis (87). There is rapid exchange of localized Aurora A at both centrosomes and microtubules with a cytoplasmic pool of the kinase (88). Aurora A is partially activated by autophosphorylation on its activation loop at T288 (equivalent to PKA T197) (17). However, pT288 does not completely interact with the basic pocket resulting in the activation segment still partially blocking the active site cleft. Additionally, since pT288 is not protected by the basic pocket, it is subject to dephosphorylation by Protein Phosphatase 1 (PP1), which regulates Aurora A activity (89). TPX2, which binds Aurora A after nuclear envelope breakdown, fully activates Aurora A. The binding of TPX2 to Aurora A causes the activation segment to move out of the active site cleft and buries pT288 resulting in full activation of the kinase and preventing its dephosphorylation (17). The motif of Aurora A has been determined to be R-X-S/T-ϕ, where ϕ is any hydrophobic amino acid (90, 91). Based on loss of function studies, Aurora A plays roles in centrosome maturation, mitotic entry, and centrosome separation (92-94).

The centrosome is the microtubule organizing center in metazoans and consists of a pair of centrioles surrounded by pericentriolar material (PCM). The pericentriolar material is the site of microtubule nucleation. Centrosome maturation, which occurs in G2 and early mitosis, involves the recruitment of proteins to the pericentriolar material such as γ-tubulin ring complex and γ-tubulin, which are involved in microtubule nucleation. Aurora A is localized to centrosomes through Plk1 activity, and Aurora A activity at the centrosome is required for centrosome maturation including the recruitment of γ-tubulin, although how this is mediated is unclear (95). One role of Aurora A in the process of maturation is the phosphorylation of NDEL1, which is required for TACC3 recruitment to the centrosome (93). TACC3, which is also a substrate of Aurora A, interacts with XMAP215, which stabilizes microtubules (96, 97). PAK, Ajuba, and TPX2 interact individually with Aurora A at the centrosome and
activate it. The kinase PAK plays a role in focal adhesion dynamics as does Hefl, another protein that interacts with Aurora A at the centrosome, suggesting Aurora A may coordinate mitosis with substratum adhesion (98, 99). Another known substrate of Aurora A at the centrosome is the kinase Lats2. Phosphorylation of Lats2 by Aurora A may play a role in localizing Lats2 to the centrosome where it is required for centrosome maturation (100, 101). TPX2, which fully activates Aurora A kinase and is phosphorylated by it, is responsible for re-localization of part of the Aurora A pool from the centrosome to the spindle microtubules (102). The localization of TPX2 and, therefore, that of Aurora A on the spindle microtubules is also regulated by Plk1 (95). The role of Aurora A on microtubules is unknown.

Aurora A plays at least two roles in mitotic entry. It is involved in the localization of Cdk1/Cyclin B to the centrosome where the activation of Cdk1/Cyclin B begins (94). Aurora A also phosphorylates Cdc25B to activate it (103). As discussed previously, Cdc25B activates Cdk1/Cyclin B by removing inhibiting phosphorylations, which results in a positive feedback loop resulting in activation of the pool of Cdk1/Cyclin B and mitotic entry. Cdk1/Cyclin B also phosphorylates PP1, reducing its activity. This results in increased Aurora A activation and furthering the positive feedback loop for Cdk1/Cyclin B activation and generating a positive feedback loop for Aurora A activation (104).

In Drosophila, the mutation of the Aurora A homolog results in monopolar spindles suggesting a role for Aurora A in centrosome separation (105). This phenotype has also been seen in human cell culture treated with Aurora A inhibitors or siRNA (92, 106). However, the role of Aurora A in the separation or the maintenance of separation of centrosomes is far from clear. The Xenopus ortholog of Aurora A has been shown to phosphorylate the Xenopus ortholog of Eg5, which as discussed previously, is phosphorylated by Cdk1/Cyclin B to localize it to the centrosome for its role in centrosome separation (107). However, the role of the phosphorylation by Aurora A is unknown.

Aurora A is destroyed by the proteasome beginning in anaphase, but its levels do not become undetectable until G1-phase (84, 108). Aurora A contains a D-box which is only active in concert with the D-box activating domain (DAD) in Aurora A (109).
Phosphorylation in the DAD protects Aurora A from ubiquitination and subsequent degradation (110). Although the D-box should be recognized by both Cdc20 and Cdh1, ubiquitination of Aurora A is mediated by the Cdh1 form of the APC/C (108, 111).

**Aurora B**

Aurora B is expressed from S- to M-phase with protein levels peaking in mitosis (112). It is part of a complex including the non-enzymatic proteins INCENP, survivin, and borealin, known as the chromosomal passenger complex (CPC) (113). As part of this complex, Aurora B travels to several localizations during mitosis (114). In prophase, the complex can be found diffusely over chromosomes, on both the arms and on the centromeres. By metaphase, the complex is concentrated on the centromeres, and in anaphase and telophase, the complex moves to the spindle midzone and midbody, respectively. There is exchange of the complex with a cytoplasmic pool at the centromeres through metaphase, but turnover is greatly reduced in anaphase at the spindle midzone (113). The binding of INCENP to Aurora B activates Aurora B. Initial binding of INCENP changes the position of the αC helix causing a partially active conformation of the activation segment (115). Phosphorylation of INCENP by Aurora B and autophosphorylation on its activation loop moves the αC and the activation segment to the fully active conformation. In addition to its role in activating Aurora B, INCENP is thought to be involved in localization of the complex, as are survivin and borealin (116-118). The Ipl1 motif, R/K-X-S/T-I/L/V, determined based on comparison of substrate sites has been assumed to be the motif of Aurora B (110, 119).

Aurora B may plays roles in chromosome condensation, spindle assembly, kinetochore-microtubule attachments, and spindle assembly checkpoint. Aurora B phosphorylates Histone H3 on S10 and S28, which correlates with chromosome condensation during prophase, but the causal link, if one exists, is not known (120).

At least two modes of spindle assembly exist: a mode driven by the centrosomes, in which centrosomes nucleate microtubules and kinetochore capture stabilizes them and a mode in which microtubules that assemble in the cytoplasm are stabilized in the vicinity of the chromatin. Aurora B seems to play a role in the latter form of spindle assembly at least in *Xenopus* egg extracts as inhibition of Aurora B or depletion of the CPC inhibits
spindle assembly. This may occur through phosphorylation and inhibition of Stathmin/Op18, which destabilizes microtubules or through phosphorylation and re-localization of MCAK, which also destabilizes microtubules (121, 122).

Aurora B affects kinetochore-microtubule attachments through a role in formation of the trilaminar kinetochore (123). It also plays a role in fixing incorrect kinetochore-microtubule attachments in response to a lack of tension on the kinetochore. Release of kinetochore-microtubule attachments seems to involve phosphorylation of Hec1, which is part of a complex required for the kinetochore to bind microtubules (124). In budding yeast, lack of tension seems to be sensed by the INCENP-survivin complex that binds at the kinetochore-microtubule interface (125). This complex may recruit Aurora B when there is a lack of tension, but not when there is appropriate tension. Aurora B seems to recruit MCAK at merotelic attachments in a kinase activity dependent manner, where MCAK can depolymerize inappropriately attached microtubules (126). This suggests a mechanism in which Aurora B causes the detachment of inappropriately attached microtubules, either syntelic or merotelic, by phosphorylating Hec1 and can then cause depolymerization of the previously merotelically attached microtubules through recruitment of MCAK. Aurora B is also involved in maintaining activation of the spindle checkpoint until all chromosomes have aligned with appropriate microtubule attachments on the metaphase plate (127). Phosphorylation by Cdk1/Cyclin B and then Plk1 on INCENP regulates Aurora B in anaphase (128). Finally, Aurora B activity is required for cytokinesis (113).

Nek2

Nek2 is a member of a family of kinases consisting of eleven members in human. The kinase domain of Nek2 is the most similar of this family to the founding member of the family in Aspergillus nidulans, Never In Mitosis A (NIMA) (129). NIMA was discovered in a screen for mutants that do not undergo mitosis, as its name suggests (129). Nek2 protein levels and activity peak in S- and G2-phases (130). Nek2 contains a leucine zipper motif C-terminal to its kinase domain which it uses for homodimerization, which allows autophosphorylation for activation of the kinase (131). The activation of Nek2 is also regulated by PP1 which dephosphorylates Nek2 to deactivate it (132). PP1
is inhibited by Nek2 phosphorylation, as well as Cdk1/Cyclin B phosphorylation and the binding of Inhibitor-2, a protein that specifically inhibits PP1 (132, 133). This suggests a positive feedback loop in activation of Nek2. The optimal phosphorylation motif of Nek2 on substrates is not known.

In S-phase and G2, Nek2 localizes to the core of the centrosome at the proximal ends of the centrioles between the two pairs of centrioles, where it is both structural and required for centrosome separation (134). The proteins Rootletin and Centrosomal Nek2-associated protein 1 (C-Nap1) are also at the proximal ends of the centrioles interacting with each other to tether the two pairs of centrioles together (134, 135). Nek2 phosphorylates both Rootletin and C-Nap1 causing them to be displaced from the centrosome at the G2/M transition. Additionally, β-Catenin interacts with Rootletin and when Rootletin is displaced by Nek2, β-Catenin relocates to other parts of the centrosome (136). The displacement of C-Nap1, Rootletin and β-Catenin allow centrosome separation. β-Catenin was also shown to be a substrate of Nek2, but the significance of this is not clear.

Nek2 has also been implicated in regulating microtubule organization at the centrosome. Ninein-like protein (Nlp) interacts with gamma-tubulin ring complex at the centrosome in interphase cells, but does not localize to the centrosome in mitosis suggesting that it has a role in microtubule nucleation in interphase, but not mitosis. Nek2 and Plk1 phosphorylate Nlp to cause its displacement from the centrosome (137). How this process is important for mitosis is unclear.

There are at least three splice variants of Nek2: Nek2A, Nek2B, and Nek2C. Nek2A and Nek2C both have a KEN-box and a D-box, but Nek2B has neither. Although Cdc20 is required for Nek2A and Nek2C ubiquitination and degradation, APC/C recognizes Nek2A and Nek2C directly, not through Cdc20 (138, 139). As a result, Nek2A and Nek2C are degraded shortly after nuclear envelope breakdown in a spindle assembly checkpoint independent manner. Nek2B levels remain constant through mitosis, but then decrease in G1-phase (140). How this decrease in protein levels occurs is unknown.
Plk1

Polo-like kinase 1 (Plk1) is a member of the Polo-like kinase family, which is named after the founding member discovered in *Drosophila*, Polo (141). Polo was named for its mutant phenotype of abnormal spindle poles (141). Like *Drosophila*, *Saccharomyces* only has a single Polo family member, Cdc5. However, humans have 4 Polo kinases, Plk1, Plk2, Plk3, and Plk4. All four members play roles in the cell cycle, but the roles of Plk2, Plk3, and Plk4 are less well understood than that of Plk1, which is the ortholog of Cdc5 (142). Plk2, Plk3, and Plk4 will be discussed below.

Plk1 is expressed from late S-phase into M-phase with protein levels peaking in mitosis (143). It localizes to the centrosome in S- and G2-phases, to the centrosome and kinetochores in prophase, to the centrosome, kinetochores, and spindle microtubules in metaphase, and to the centrosomes, kinetochores, and spindle midzone in anaphase (143). Plk1 localized to the centrosomes and kinetochores exchanges rapidly with a cytoplasmic pool of the protein (Kazuhiro Kishi, Personal Communication). In addition to an N-terminal kinase domain, Plk1 has two C-terminal Polo-box motifs, which have recently been found to fold together to form a Polo-box domain (PBD) (144). The Polo-box domain of Plk1 was found to be a phospho-dependent binding domain that binds sequences with the consensus sequence S-pS/pT (48). Kinase activation occurs concurrently with expression and involves at least two possibly related mechanisms. Kinase activity is inhibited by an interaction between the kinase domain and the PBD. This autoinhibition is relieved by the binding of the PBD to phospho-peptides (144). It has also been reported phosphomimic mutation of the activation loop reduces the interaction between the kinase domain and PBD, suggesting that phosphorylation on the activation loop, aside from activating the kinase by changing the activation loop conformation to the active state, may also relieve the autoinhibition (145). The upstream kinase that phosphorylates and Plk1 for activation is not known. The consensus sequence for Plk1 phosphorylation has been determined by kinetics on peptides based on a known site in Cdc25C. The resulting motif, D/E-X-S/T-φ-X-D/E, where φ is any hydrophobic amino acid, is biased (146).

Plk1 is involved in centrosome separation and centrosome maturation. The original study of Polo mutants demonstrated monopolar spindles and subsequent work
has confirmed this (141, 147, 148). However, the role of Plk1 in centrosome separation is still unknown, possibly because its roles in separation and maturation are not clearly separable. Several roles for Plk1 in centrosome maturation are at least partially known. Plk1 binds and recruits γ-tubulin to the centrosome for microtubule nucleation (148, 149). Plk1 plays a role in inhibiting the microtubule destabilizing protein Stathmin/Opl8 (150). As previously mentioned, it also regulates Nlp in conjunction with Nek2 (137). Plk1 phosphorylates TCTP likely regulating its association with microtubules as it binds microtubules to stabilize them until metaphase (151). Plk1 also phosphorylates Asp which is involved in microtubule organization at the centrosome (152).

Plk1 plays two roles in mitotic entry. Plk1 phosphorylates Cdc25C as part of the activation of Cdc25C (153). As mentioned previously Cdc25C removes phosphorylations by Wee1 and Myt1 that inhibit Cdk1/Cyclin B. Further contributing to Cdk1/Cyclin B activation, Plk1 phosphorylates Wee1 and Myt1 resulting in their inactivation (154, 155). Cdk1/Cyclin B generates PBD binding sites on Cdc25C and Myt1 that enable Plk1 phosphorylation of these proteins (153, 155).

Plk1 may play at least two roles associated with the kinetochore. It is required for the association of Hecl, which is required for kinetochore-microtubule attachments, Mad2, which is involved in the spindle checkpoint, and other proteins involved in kinetochore structure (156). It may also play a direct role in the spindle assembly checkpoint by generating the 3F3/2 phosphoepitope on kinetochores lacking tension (156, 157).

Plk1 also regulates the metaphase-anaphase transition. Plk1 activates the APC/C in two ways. At the beginning of mitosis, Plk1 and Cdk1/Cyclin B phosphorylate the APC/C to activate it (79). Early mitotic inhibitor 1 (Emi1) inhibits the APC/C in S-phase, but is degraded in prometaphase allowing activation of the APC/C. Plk1, stimulated by Cdk1/Cyclin B, phosphorylates Emi1 resulting in its destruction (158, 159). The APC/C degrades Securin, which inhibits the protease Separase from degrading cohesin and allowing sister chromatid separation (160). In budding yeast, Cdc5 phosphorylates the cohesin Scc1 to increase its affinity for separase, thus further contributing the metaphase-anaphase transition. Although Plk1 has roles in cytokinesis,
it begins to be ubiquitinated by the APC/C and destroyed in anaphase and is gone by G1-phase (84, 161).

**The Polo-like Kinase Family**

As mentioned above, the Polo-like kinase family in humans has four members, Plk1, Plk2, Plk3, and Plk4. This family of kinases is defined by two conserved features, an N-terminal kinase domain and at least one Polo-box motif in the C-terminus. The kinase domains are highly conserved with the highest sequence similarity between Plk2 and Plk3. Plk1 is the next most similar and Plk4 is the least. Plk1, Plk2, and Plk3 have two Polo-box motifs, which fold together to form a Polo-box Domain (PBD) (162). The Polo-box domain of Plk1 was shown to be a phospho-dependent binding domain and the binding consensus sequences for the PBDs of Plk1, Plk2, and Plk3 were published to be essentially the same, S-pS/pT. Plk4 only has a single Polo-box, the function of which seems to be different from that of the other Plks, as described below. All members of this family seem to play non-redundant roles in the cell cycle as the localizations and patterns of expression are not the same. However, the roles of Plk2, Plk3, and Plk4 are poorly understood and almost no substrates for these kinases are known.

**Plk2**

Plk2 was discovered as a serum inducible immediate early gene in NIH 3T3 cells (163). It is expressed in G1 and protein levels and activity peak and fall off in G1 (163, 164). Phosphomimic mutation on the activation loop was found to increase the activity of the kinase 10-fold suggesting that the kinase may need to be phosphorylated on the activation loop to be fully activated (164).

Plk2 is a non-essential gene in mouse, but embryos showed slight growth retardation and MEFs from these mice showed delayed S-phase entry, suggesting a role for Plk2 in the cell cycle (165). At least one cell cycle role may be in centrosome duplication as Plk2 localizes to the centrosome and both overexpression and knockdown results in abnormal numbers of centrioles (166). In addition to being expressed in response to serum, it is expressed after X-ray and UV irradiation and other DNA damaging agents due to p53 transactivation, implying a role in response to cellular stresses (167, 168). Plk2 is also expressed in postmitotic cells, where it has been reported
to interact with calcium and integrin binding protein (CIB), which inhibits its activity and spine associated RapGAP (SPAR) which is involved in remodeling dendritic spines (164, 165, 169, 170). Association with both of these proteins is through its PBD. However, the phospho-dependency of the binding has not been checked. SPAR seems to be a substrate of Plk2, but the site of phosphorylation was not identified (170). Indeed, no phosphorylation sites have been mapped and the kinase motif has not been determined.

Plk3

Much of data on Plk3 is conflicting and/or surprisingly similar to that of Plk1 or Plk2. Like Plk2, Plk3 was originally discovered as an immediate early gene transcribed in response to fibroblast growth factor (FGF) treatment of NIH 3T3 cells (171). Like Plk2, Plk3 mRNA is expressed only in G1 (171, 172). Reports of protein expression have been inconsistent. Plk3 protein levels were originally reported to remain constant throughout the cell cycle (173, 174). However, recently Plk3 protein levels have been reported to track with the mRNA levels: increasing, peaking, and falling off in G1 (175). The localization of the kinase is also controversial. The kinase was originally described to localize to the cellular cortex (176, 177). Later it was suggested to localize to centrosomes, mitotic spindle, and the midbody during mitosis, which is very similar to the distribution of Plk1 (178, 179). Subsequent work suggests that the kinase does not localize to any of these structures, but instead to the nucleolus (175). The mechanism of activation of the kinase has not been examined at a biochemical level and the motif has not been determined.

Plk3 seems to have the same function as Plk2 in neurons as it also interacts with CIB and SPAR (169, 170). However, its role in the cell cycle seems to be very different from that of Plk2. Surprisingly, Plk3 has been reported to both inhibit as well as promote cell cycle progression.

Plk3 seems to become transcribed and activated after oxidative stress and DNA damage to mediate a stress response (174, 180, 181). This activation is at least partially mediated by phosphorylation by Chk2 (181). Plk3 then phosphorylates Chk2, resulting in further activation of Chk2 (182). It has also been reported to phosphorylate S20 of p53 (181). Phosphorylation at S20 of p53 is important for the cell cycle arrest mediated by
p53 and was previously reported to be phosphorylated by Chk1 and Chk2 after DNA damage (183, 184). Chk1 and Chk2 have motifs requiring an R in the -3 position (185). The S20 site of p53 has an E in the -3 position making it a more likely target of Plk3 than Chk1 or Chk2, based on the sequence similarity between Plk1 and Plk3 kinase domains and what is known about the Plk1 motif (146). Plk3 has recently been reported to generate a priming phosphorylation for Glycogen Synthase Kinase 3 beta (Gsk3β) on Cdc25A resulting in phosphorylation of Cdc25A by Gsk3β and the degradation of Cdc25A (186). Since Cdc25A promotes cell cycle progression, Plk3 is implicated in inhibition of the cell cycle by two different pathways, one through stabilization of p53 and one through destruction of Cdc25A.

Plk3 has also been reported to phosphorylate Cdc25C resulting in the translocation of Cdc25C into the nucleus as part of mitotic entry (187). It also may be involved in promoting golgi fragmentation which is required for proper mitosis (188). Recently, Plk3 has been shown to be required for entry into S-phase (175). Finally overexpression of Plk3 results in cell cycle arrest and apoptosis, which was explained at least partially as the outcome of defects in cytokinesis (177). Based on these reports, Plk3 seems to also have a complex role in promoting the cell cycle. The role of Plk3 in the cell cycle remains mysterious, but determination of the motif and additional substrates should help to clear up the controversy.

Plk4

Plk4, like the other Plk family members, has an N-terminal kinase domain, but unlike the other family members, it only has a single polo-box, which does not function as a phospho-dependent binding domain. Instead, the polo-box seems to function in homodimerization (189). A second region that mediates homodimerization independently of the polo-box also exists between the kinase domain and polo-box (189). Plk4 is expressed from late G1-phase through mitosis with mRNA levels falling in early G1 (190). The cell cycle regulation of protein and kinase activity levels has not been determined. Plk4 localizes to the nucleolus in G2 and to the centrosome in G2 and M (191). The centrosome localization seems to be mediated by the polo-box (189). The motif of Plk4 has been reported to be χ-X-S/T-ϕ-ϕ-χ/P, where
\( \chi \) is a charged residue and \( \phi \) is a large hydrophobic residue (192).

Plk4 seems to be required for mitosis as the Plk4 knockout mouse is embryonic lethal with the embryo dying with large numbers of cells in mitosis with high Cyclin B levels (191). MEFs from mice heterozygous for Plk4 showed centrosome amplification, multipolar spindles, and aneuploidy (193). Plk4 has been shown to be required for centriole duplication, likely explaining the phenotypes seen in the Plk4 heterozygote MEFs (194, 195). Substrates of the kinase have yet to be determined.

**Summary and Preview**

Protein kinases play key roles in regulation of the cell cycle. Activation of a kinase frequently involves phosphorylation on the activation loop. The phosphoresidue on the activation loop interacts with a basic pocket in the active site cleft causing a conformational change that aligns the catalytic residues for optimal phospho-transfer. The motif of a kinase is determined by the residues in the active site cleft which select sequences that the kinase can phosphorylate based on their ability to make favorable electrostatic, hydrogen-bonding, and/or hydrophobic interactions. In the case of kinases that do not require phosphorylation for activation, the basic pocket can contribute to the specificity of the kinase by binding sequences that have been previously phosphorylated upstream or downstream of the site to be phosphorylated. Determinants of localization of a kinase play an important role in substrate specificity by increasing the local concentration of kinase in the area of the substrate. Activation, motifs, and localizations play key roles in the substrate specificity of the major mitotic kinases Cdk1, Aurora A, Aurora B, Nck2, and Plk1. Each of the events in mitosis is regulated by more than one of these kinases and these kinases have overlapping localizations often even interacting with and phosphorylating the same substrate proteins on different residues. The motifs of this functional family have not been completely defined. Very little is known about the activation, motifs, and localizations of the rest of the Polo-like kinase family, Plk2, Plk3, and Plk4, or their roles in the cell cycle. Very few substrates of these kinases have been identified.

In this thesis, we set out to determine complete, unbiased motifs of two sets of kinases involved in the regulation of the cell cycle. The motifs of the functionally
related, but sequence dissimilar major mitotic kinases were determined (Chapter Two) as well of those of the sequence related, but functionally distinct Polo-like kinase family (Chapter Three). The motif of Plk1 led to the discovery of new substrates. Additionally, the motifs of the major mitotic kinases along with their localizations lead us to put forward a hypothesis to explain how despite overlapping localizations and overlapping motifs, these kinases are able to phosphorylate distinct substrates sites, suggesting that motifs may be involved in the coordinated regulation of these kinases as a group not just the kinases individually. The motifs of the Polo-like kinase family result in the exciting novel motif of Plk3 involving specificity for phosphotyrosine and the even more exciting result that the motif of a kinase is not necessarily ‘static’ as was previously thought. In Chapter 4, the motifs of these two sets of kinases are compared and contrasted with respect to sequence and resulting function and follow-up experiments are proposed.
References


126. Knowlton AL, Lan W, Stukenberg PT. Aurora B is enriched at merotelic attachment sites, where it regulates MCAK. Curr Biol. 2006 Sep 5;16(17):1705-10.


164. Ma S, Liu MA, Yuan YL, Erikson RL. The serum-inducible protein kinase snk is a G1 phase polo-like kinase that is inhibited by the calcium- and integrin-binding protein CIB. Mol Cancer Res. 2003 Mar;1(5):376-84.


182. Bahassi el M, Myer DL, McKenney RJ, Hennigan RF, Stambrook PJ. Priming phosphorylation of Chk2 by polo-like kinase 3 (Plk3) mediates its full activation by ATM and a downstream checkpoint in response to DNA damage. Mutat Res. 2006 Apr 11;596(1-2):166-76.


Chapter Two

The Determination of the Motifs of the Major Mitotic Kinases: Motifs and Localizations of the Major Mitotic Kinases May Cooperatively Regulate the Kinases as a Group

Jes Alexander, Dan Lim, Brian Joughin, Fabio Sessa, Frank Ivins, Andrew Fry, P. Todd Stuckenberg, Andrea Mussachio, Steve Smerdon, Jan-Michael Peters, Michael Yaffe

Contributions:
Dan and I determined the motif of Plk1 together (Figure 2.2A). Bioinformatics localization enrichment data (Figure 2.6A) and methods are from Brian. Data in Figure 2.3C is from Jan-Michael. Figure 2.3D and 2.4D were created by Dan. Figure 2.5D was created by Steve. Aurora A protein was from Todd, Aurora B protein was from Fabio and Andrea, Nek2 protein was from Frank, Andrew, and Steve, and Plk1 protein was from Dan. I performed all the other experiments, wrote the rest of the sections, and created the other figures.
Abstract

Mitosis is highly regulated, as errors in this process result in aneuploidy and genetic instability. Regulation of these processes relies heavily on precise protein phosphorylation of specific substrate sites and the major mitotic kinases, Cdk1/Cyclin B, Aurora A, Aurora B, Nek2, and Plk1, play key roles in this regulation. Some of the substrates of these kinases are known, but many remain to be identified as many of the roles of these kinases in mitosis have yet to be understood at the molecular level. These kinases have overlapping localizations and, therefore, have access to the same substrates and sites on substrates. Yet, each kinase phosphorylates distinct sites from those phosphorylated by each of the other kinases. How this specificity is maintained is not clear. To gain insight into this specificity and to aid in the identification of new substrates of these kinases, we have determined the motifs of the major mitotic kinases using Positional Scanning Oriented Peptide Library Screening (PS-OPLS) which gives complete and unbiased motifs. We verified the motifs using kinetic studies on peptides with single amino acid changes from the optimal peptide determined by PS-OPLS. The motifs match known substrates of these kinases. Using structural models of the kinase domains with the optimal peptides docked in, we were able to rationalize the substrate specificity of these kinases. Plk1 was found to have selectivity for N in the -2 position, which was not known previously, and this led to the identification of new phosphorylation sites in Scc1, Bub1, and p31/Comet. Based on the motifs and the previously known localizations we put forward a hypothesis that may explain how these kinases maintain distinct phosphorylation sites and, if true, would suggest that the individual motifs and localizations of the major mitotic kinases cooperatively regulate the substrate site selection of these kinases in a coordinated manner.
Introduction

Mitosis is the process by which the contents of the nucleus are partitioned to daughter cells during cell division. This process includes centrosome separation and maturation, chromatin condensation, nuclear envelope breakdown, spindle formation, amphitelic chromosome attachment, and chromosome segregation. The timing and mechanics of each of these processes is carefully regulated as problems in these processes result in aneuploidy, genetic instability, and cancer (1, 2). Regulation of these processes relies heavily on precise protein phosphorylation of specific substrate sites and the major mitotic kinases, Cdk1/Cyclin B, Aurora A, Aurora B, Nek2, and Plk1, play key roles in this regulation (3). Although some of the substrates of these kinases are known, many have yet to be identified (3). Additionally, how the specificity of these kinases is established and maintained remains to be understood.

Cdk1/Cyclin B is the master regulator of mitosis and its activation marks entry into mitosis. It has roles in chromatin condensation, nuclear envelope breakdown, and spindle formation (4-10). Aurora A is involved in centrosome separation and maturation, while Aurora B is involved in chromosome condensation and assuring amphitelic chromosome attachment (11-16). Nek2 is required for centrosome separation and plays a role in centrosome maturation (17, 18). Plk1 has been shown to regulate Cdk1/Cyclin B activation, centrosome maturation, spindle formation, and amphitelic chromosome attachment (19-24).

Not surprisingly, since these kinases cooperate to regulate processes within mitosis, they have overlapping localizations (Figure 2.1). At and just prior to the G2/M transition, Aurora A and Plk1 are localized to the pericentriolar material, while Nek2 localizes to the proximal centriole at the core of the centrosome and Aurora B localizes to the centromeres of the decondensed, but duplicated chromosomes (17, 25-27). At this time, Cdk1/CyclinB activity is low but increasing. At metaphase, Cdk1/Cyclin B is maximally active and diffusely localized throughout the cell, but intense activity appears to be present at the centrosomes, spindle microtubules, and kinetochores (10, 28-30). Aurora A and Plk1 co-localize with Cdk1/Cyclin B to the pericentriolar material and the
The major mitotic kinases have overlapping localizations in mitosis. Localizations of the major mitotic kinases, Cdk1/Cyclin B (red), Aurora A (green), Aurora B (purple), Plk1 (blue), and Nek2 (yellow), at the G2/M transition (left) and at metaphase (right) are shown.

spindle microtubules (25, 31). Plk1 also co-localizes with Aurora B and Cdk1/Cyclin B at the kinetochores (10, 25, 27).

Although mitosis has been studied extensively, there is still much to understand about the process at the molecular level. Some substrates of these kinases are known, but there seem to be more substrates to be identified, as the known substrates do not to explain all the phenotypic changes these kinases are thought to be involved in. Additionally, despite overlapping subcellular localizations, and therefore, access to overlapping sets of potential phosphorylation sites on substrates, these kinases have distinct and apparently mutually exclusive sets of phosphorylation sites \textit{in vivo}, as seems to be required for the proper regulation of mitosis. How the activities of, and substrate targeting by, these kinases are coordinated so that each phosphorylates its own sites, but not those of the other kinases is not clear. Since consensus phosphorylation motifs of individual kinases play important roles in substrate selection for many well-studied kinases, we suspected that determining the optimal motifs for these mitotic kinases might shed additional light on the problem of substrate selection and co-localization, as well as help in the identification of new substrates of these kinases.
We, therefore, have determined the motifs of Cdk1/Cyclin B, Aurora A, Aurora B, Nek2, and Plk1, using Positional Scanning Oriented Peptide Library Screening (Figure 2.2A), which gives an unbiased and complete motif. In particular, this method allows each amino acid in each position in a sequence motif to be tested independently of the rest of the sequence, and defines not only which amino acids are positively selected in each position, but also which are discriminated against. Subsequent kinetic studies verified the motifs determined by PS-OPLS, and the motifs closely matched several previously mapped sites on known substrates. Additionally, we were able to rationalize the basis for motif selection using models of the optimal peptide docked into the X-ray crystal structures of these kinases. In addition, the new expanded motif of Plk1 allowed the identification of new substrates for Plk1. Finally, based on the motifs of these kinases and their localizations, we propose the hypothesis that in mitosis, major mitotic kinases with overlapping localization do not have overlapping motifs and major mitotic kinases with overlapping motifs do not have overlapping localizations. Such regulation would imply that all the individual motifs and localizations of the major mitotic kinases together cooperatively regulate the substrate site selection of these kinases in a coordinated manner, giving us the beginnings of a systems-level understanding of the regulation of mitosis by protein phosphorylation.
Experimental Procedures

Kinase Protein Production and Purification

Recombinant full-length human wildtype C-terminally hexa-His-tagged Aurora A, X. Laevis wildtype Aurora B(60-361):INCENP(790-847) complex, and full-length human T175A Nek2A proteins were produced as described previously (32-34). To produce T210D Plk1 kinase domain, nucleotide sequence encoding the kinase domain of human Plk1 (aa. 38-346) was cloned into NdeI and XhoI restriction sites in the pET28a bacterial expression vector (Novagen). The resulting fusion protein contains an N-terminal hexa-His-tag and a thrombin protease cleavage site between hexa-His-tag and the kinase domain. A Thr-210 mutation to Asp was introduced using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and verified by sequencing. Protein was expressed in E. coli Rosetta (Novagen) cells. A starter culture was grown overnight in LB. 6L of TB culture was initiated by addition of the starter culture at a 1:100 dilution. After 4 hours of growth, the 6L of TB culture were cooled to 4°C on ice, IPTG was added to a final concentration of 1 mM and the cultures were incubated at room temperature for 16 hours. Cells were pelleted by centrifugation at 4000 X g for 10 min at 4°C, the pellet from each liter of culture was resuspended in ~40 mL of lysis buffer (50 mM Tris, 500 mM NaCl, 14 mM β-mercaptoethanol), transferred to separate 50 mL Falcon tubes, and lysed by sonication using a Branford sonicator with a probe tip (power level 8, duty cycle 50%, 2:33 min:sec, 2 times) in an ethanol-ice bath. The lysate was clarified by ultracentrifugation at 40,000 RPM in a Beckman 45Ti rotor for 30 min at 4°C. Initial purification utilized Ni-NTA affinity chromatography. Following ultracentrifugation, the supernatant was mixed with 25 mL of packed Ni-NTA beads by end-over-end rotation at 4°C for 2 hours. The bead slurry was loaded onto a column for FPLC with the aid of a peristaltic pump. The column was transferred to the FPLC, washed with 280 mL of wash buffer (500 mM NaCl, 14 mM β-mercaptoethanol, 10 mM imidazole, 10 mM Phosphate, pH 7), and the protein was eluted using buffer of the same composition except containing 300 mM imidazole. Fractions were analyzed by on-line UV absorbance and SDS/PAGE. Fractions containing Plk1 kinase domain protein were pooled, 100 units of thrombin
protease were added to cleave off the hexa-His-MBP tag, and the mixture was incubated at 4°C overnight. To remove the cleaved tag, the protein was re-passaged through Ni-NTA using the same conditions as described above. The resulting ~10 mL of material was applied to a Superose 12 column (16 mm X 30 cm) and eluted using gel filtration buffer (500 mM NaCl, 2 mM DTT, 10 mM Tris, pH 8). Fractions (2 mL) were analyzed for Plk1 content by on-line UV absorbance and SDS-PAGE.

**Phosphorylation Motif Determination by Peptide Library Array Screening**

Positional Scanning Oriented Peptide Library Screening (PS-OPLS) was performed following Hutti et al. (35). Briefly, solution-phase kinase reactions were performed in parallel on 198 separate biotinylated, partially degenerate oriented peptide libraries (Anaspec, Inc) arrayed in a 384 well microtiter plate in a 22 row X 9 column format. Each peptide library contains an N-terminal biotin tag, a 50:50 mix of serine and threonine at the orienting phospho-acceptor residue, a single second fixed amino acid located between the -5 and +4 position, and a mixture of amino acids at all other positions. Individual libraries contain any of the 20 natural amino acids as well as phosphothreonine and phosphotyrosine in the second fixed position, corresponding to the 22 rows. Scanning down the rows in the array moves the position of the fixed amino acid from -5 to +4 relative to the fixed phospho-acceptor residue, while scanning across the columns changes the identity of the second fixed amino acid. As an example, the peptide library with Lys fixed in the -4 position has the following sequence: Y-A-X-K-X-X-S/T-X-X-X-A-G-K-K-biotin, where amino acids are represented in 1-letter code, and X is an equal mixture of all 17 natural amino acids excluding Cys, Ser, and Thr to prevent oxidation effects and eliminate secondary phosphorylation events. S/T denotes a 50:50 mix of Ser and Thr. Kinase reactions were performed at 30°C in a total volume of 16 μl containing 31.25 μM peptide library, 100 μM ATP, and 200 μCi of [32P]-γ-ATP, in 150 mM NaCl (500 mM NaCl for Nek2 kinase reactions), 10 mM MgCl2, 1 mM DTT, 0.1% Tween 20, and 50 mM Tris, pH 7.5. For Aurora A and Aurora B, reactions were performed for 6 hours with 0.25 and 0.792 μg of protein per reaction, respectively. For Nek2, reactions were performed for 8 hours with 2.4 μg of protein per reaction. Reactions for the Plk1 motif were performed for 3 hours with 0.5 μg of protein per
reaction. For Cdk1/Cyclin B, reactions were performed for 4 hours with 0.08 μg of protein complex (Millipore) per reaction. Following incubation, 2 μl of each reaction were simultaneously transferred to a streptavidin-coated membrane (Promega SAM² biotin capture membrane) using a 384 slot pin replicator (VP Scientific). The membrane was washed three times with 140 mM NaCl, 0.1% SDS, 10 mM Tris, pH 7.4, three times with 2 M NaCl, twice with 2 M NaCl containing 1% H₃PO₄, and once with water. The extent of peptide library phosphorylation was determined by imaging the membrane with a phosphorimager (Molecular Dynamics).

**In Vitro Kinase Assays**

Kinase assays for kinetic parameter determination were performed at 30°C in 90 μL of kinase reaction buffer (50 mM Tris, pH 7.5, 150 mM NaCl (500 mM NaCl for Nek2 assays), 10 mM MgCl₂, 100 μM ATP, 9 μC [³²P]-γ-ATP, and 1 mM DTT). Each reaction contained 0.003 μg of Aurora B:INCENP complex, 3.6 μg of T175A Nek2 protein, or 0.054 μg of T210D Plk1 kinase domain. The sequences of the optimal peptides for Aurora B, Nek2 and Plk1 were ARRHSMGWAYKKKK, WFRMSIRGGYKKKK, GHDTSFYWAAYKKKK, respectively. These optimal peptides were determined by taking the most highly selected amino acid from each position, -4 to +3, of the PS-OPLS blot and using S as the phospho-acceptor residue, 0. A C-terminal tyrosine was added to each peptide to allow determination of concentration of peptide solutions by UV spectrometry, and four C-terminal lysines were appended to increase solubility and electrostatic interaction with phosphocellulose paper. Additional peptides with single amino acid changes from the optimal peptides were as indicated. Concentrations of peptides were as indicated. 5 μL of each reaction were spotted on phosphocellulose at 0, 3, 6, 9, 12, and 15 minutes in duplicate. The phosphocellulose paper was washed 4 times with 0.5% phosphoric acid, transferred vials containing 3 mL of scintillation cocktail, and scintillation counted. From these kinase assays, $K_m$, $V_{max}$, and $V_{max}/K_m$ values were determined by curve fitting assuming classic Michaelis-Menten kinetics. For each concentration of peptide, care was taken to insure that less than 5% of the total substrate was phosphorylated and that the reaction rate was linear with respect to time.
Kinase assays for comparison of activity of each kinase against the optimal peptides of each kinase were performed at 30°C in 25 μL of kinase reaction buffer. Each reaction contained 0.008 μg of Aurora B:INCENP complex, 1 μg of T175A Nek2 protein, or 0.046 μg of T210D Plk1 kinase domain. Amounts of each kinase were chosen so that reactions of each kinase with its optimal peptide would yield approximately equal amounts of phosphorylated peptide. Concentrations of optimal peptide were set to the $K_m$ of the kinase in the reaction. Reactions were performed in triplicate and 5 μL of each reaction was spotted on phosphocellulose at 0 minutes and 1 hour. The phosphocellulose paper was washed 4 times with 0.5% phosphoric acid, added to vials containing scintillation fluid, and counted.

**Structural Modeling**

The X-ray crystal structures of Plk1 (PDB ID: 2OWB), Aurora B (PDB ID: 2BFX), and Nek2 (PDB ID: 2JAV) were used as base models. Peptides were manually docked into the substrate binding cleft based on threading the structures of the kinases onto the structure of the PKA:PKI complex (38). Molecular surface representations of the Aurora B, Nek2, and Plk1 active sites were created with GRASP (33, 34, 36, 37) and shaded by electrostatic potential using surface projections of charge calculated with DelPhi (37).

**Centrosome and Spindle Enrichment Informatics**

A list of proteins associated with the centrosome was obtained from mass spectrometry data of Andersen et al. (39). A list of spindle proteins was obtained from mass spectrometry data of Sauer et al. (40). Sequences for these proteins were downloaded directly from NCBI Entrez Protein using NCBI Entrez Utilities (http://www.ncbi.nlm.nih.gov/entrez/query/static/eutils_help.html). The human proteome was downloaded as version 3.23 of the file ipi.HUMAN.fasta from the International Protein Index (http://www.ebi.ac.uk/IPI/), and a list was generated of the amino acid sequences surrounding each serine or threonine amino acid residue within the sequences of each protein in each of these three data sources. The human proteome yielded 1850231 sites from 66619 proteins. The centrosome data and spindle data yielded 35425
sites from 524 proteins and 34909 sites from 277 proteins, respectively. For each motif of interest, statistical significance of that motif in the centrosomal or spindle proteomes was calculated using the hypergeometric distribution:

\[ P = \sum_{i=k}^{n} \frac{\binom{m}{i} \binom{N-m}{n-i}}{\binom{N}{n}} \]

where \( N \) is the number of S/T sites in the human proteome, \( n \) is the number of S/T sites in the centrosomal or spindle proteome, \( m \) is the number of motif sites in the human proteome, and \( k \) is the number of motif sites in the centrosomal or spindle proteome. This corresponds to the probability of seeing as many instances or more of the motif as are seen in the centrosomal or spindle proteome by chance if drawing a dataset the same size as the centrosomal or spindle proteome at random from the human proteome.
Results

Determination of the Optimal Consensus Motif of Plk1

To address our question of how major mitotic kinases phosphorylate distinct sites and to aid in substrate identification, we determined the optimal consensus motif of Plk1 by Positional Scanning Oriented Peptide Library Screening (Figure 2.2A). Plk1 showed strong selection for D, N, or E in the -2 position relative to the S/T phospho-acceptor residue, as evidenced by darker spots on the autoradiogram of the spotted streptavidin membrane. Plk1 also showed strong selection for hydrophobic amino acids in the +1 position, particularly F, Y, I, and M. Interestingly, P in the +1 was strongly discriminated against as indicated by the arrowhead in the right panel of Figure 2.2A.

To validate the PS-OPLS results, we verified the motif revealed by this technique by determining kinetic parameters for Plk1-dependent phosphorylation of the optimal peptide (GHDTSFYWAAYKKKK) as well as peptides containing single amino acid mutations in optimal peptide (Figure 2.2B). $V_{\text{max}}/K_m$ ratios of these variant peptides followed the same trends as displayed in the PS-OPLS. The optimal peptide had neither the lowest $K_m$ nor the highest $V_{\text{max}}$, but had the highest $V_{\text{max}}/K_m$ ratio which is consistent with the peptide library screening approach determining optimal substrate sequences based on maximum $V_{\text{max}}/K_m$ ratio (i.e. maximal turnover). All variations of the optimal peptide tested resulted in decreased $V_{\text{max}}/K_m$ ratios suggesting that the peptide chosen as the optimal peptide is correct and verifying that the optimal peptide can be determined by PS-OPLS by choosing the most highly selected amino acid in each position. Particularly, changing the D to an A in the -2 resulted in a greater than 20-fold drop in the $V_{\text{max}}/K_m$ ratio and activity of Plk1 against the peptide with P in the +1 was so low that it was not possible to fit the data to a Michaelis-Menten curve. The lack of activity against this peptide recapitulated the lack of activity against P in the +1 in the PS-OPLS (Figure 2.2A). Based on these data the optimal consensus sequence is D/N/E-X-S/T-Φ, where X is any amino acid and Φ is hydrophobic amino acids except P. (Figure 2.2A).

To investigate the structural basis for the optimal motif we performed molecular modeling studies using a previously published structure of the Plk1 kinase domain and a
Figure 2.2: Plk1 Selects D, N, or E in the -2 relative to the phospho-acceptor S/T and hydrophobic amino acids in the +1, but strongly discriminates against P in the +1 in opposition to Cdk1. A. Experimental flow of Positional Scanning Oriented Peptide Library Screening (PS-OPLS) technique for determination of kinase substrate specificity (left) and PS-OPLS blot of Plk1 T210D kinase domain (a.a. 38-346). B. Determination of kinetic parameters ($K_m$, $V_{max}$, and $V_{max}/K_m$ ratio) for reactions of Plk1 T210D kinase domain with the optimal substrate peptide and peptides with single amino acid changes from the optimal. Graph of reaction data fitted to the Michael-Menten equation for selected peptides as indicated (left) and table of kinetic parameters (right). C. PS-OPLS blot of Cdk1/Cyclin B.
model of the PKA:substrate complex as base models (Figure 2.3C) (36). The optimal peptide sequence, HDTSFYWA, was modeled into the active site in an extended conformation. Selection for D and E in the -2 seems to be based on an electrostatic on electrostatic and/or hydrogen-bonding interaction with K178. This conclusion is supported by the observation that mutation of the -2 residue to one incapable of forming either type of interaction (i.e. A) reduced the $K_m$ by 6-fold, while replacement with a positively charged residue (i.e. R) reduced the $K_m$ nearly 10-fold. F and other hydrophobic amino acids selected in the +1 appear to fit into a hydrophobic pocket formed by L211, P215, and I218 in the activation loop. Substitution of small or negatively charged amino acids in this position reduced the $K_m$ by a factor of 20 or greater.

The motif we determined is consistent with mapped Plk1 phosphorylation sites, of which several are known (Figure 2.3A). Most Plk1 phosphorylation sites contain a D or E in the -2 and a hydrophobic amino acid in the +1 position. This motif is also consistent with a previously published motif for Plk1: D/E-X-S/T-φ-X-D/E, where φ is any hydrophobic amino acid (41). However, the new motif differs from the previous motif since it expands the -2 position selection to include N, which has a similar level of selectivity to that of D or E in that position and demonstrates no selection for D or E in the +3 position (Figure 2.2A).

This expanded motif suggested that there may be novel substrates containing Plk1 phosphorylation sites with N in the -2 relative to the phospho-acceptor. To identify such substrates we performed a phospho-proteomic screen (Figure 2.3B). Plk1 phosphorylation was inhibited using the Plk1 specific inhibitor, BI 2536, in nocodazole arrested HeLa cells. Phosphorylations on specific mitotic complexes were mapped by mass spectrometry. In parallel, phosphorylations were mapped on these same mitotic complexes from nocodazole arrested, but non-Plk1-inhibited HeLa cells. Phosphorylations found in the non-Plk1-inhibited data set, but not in the inhibited data set indicate Plk1 dependent phosphorylation. Many Plk1 phosphorylation sites were determined (data not shown) including 3 with N in the -2 position relative to the phosphorylation (Figure 2.3C). Two of these proteins, Bub1 and p31/Comet, are involved in spindle assembly checkpoint and the third, Scc1/Rad21, is involved in sister
Figure 2.3: The substrate specificity motif of Plk1 matches previously mapped Plk1 phosphorylation sites and aids in the identification of new N in the -2 sites. A. Previously published mapped Plk1 phosphorylation sites contain D or E in the -2 and hydrophobic amino acids in the +1. B. Experimental strategy for the identification of new N in the -2 Plk1 substrate sites. C. Table of mapped N in the -2 Plk1 phosphorylation sites identified as in B. D. Molecular surface representation of the Plk1 active site with the optimal substrate peptide, HDTSFYWA, shown in stick representation. The electrostatic potentials are colored red (negative) and blue (positive).
chromatid cohesion in mitosis (7, 42, 43). Regulation of the spindle assembly checkpoint and sister chromatid cohesion are known roles of Plk1 (44, 45).

The strong discrimination of Plk1 against phosphorylating sites with P in the +1 suggests a possible mechanism by which Cdk1/Cyclin B and Plk1 could phosphorylate mutually exclusive sites. Cdk1 is a proline-directed kinase with the consensus phosphorylation motif, S/T-P-X-K/R (46). A P in the +1 is essential for Cdk1 to phosphorylate a site, which is indicated by the lack of selection of other amino acids in the +1 position of the PS-OPLS blot (Figure 2.2C.) Since P in the +1 is essential Cdk1-mediated for phosphorylation and Plk1 seems incapable of phosphorylating peptides with P in the +1, Cdk1 and Plk1 should not be able to phosphorylate the same sites on proteins. Aurora A has also been shown previously to strongly discriminate against sites with P in the +1 suggesting that Aurora A and Cdk1 also phosphorylate mutual exclusive sites (47).

**Determination of the Optimal Consensus Motif of Aurora A and Aurora B**

To verify the discrimination against P in the +1 by Aurora A and to determine if this is a more general mode of regulation between Cdk1 and the major mitotic kinases, we determined the motifs of Aurora A and Aurora B by PS-OPLS (Figure 2.4A). Both Aurora A and Aurora B showed extremely strong selection for R in the -2 and strong discrimination against all other amino acids in the -2. Surprisingly, even K in the -2 is a poor substitute for R. Additionally, there is a much smaller selectivity for R in the -3 position. In the +1 position, although both kinases selected hydrophobic amino acids, the specific set of amino acids selected differed slightly between Aurora A and B. Aurora B selected I and M most strongly, while Aurora A selected F and I in addition to L and M. Importantly, both Aurora A and Aurora B showed strong discrimination against P in the +1. The similarity in overall phosphorylation motifs between Aurora A and B is not surprising as these kinases have a high level of sequence similarity.

This result verified the previous report of Aurora A being largely incapable of phosphorylating sites with P in the +1. To independently verify the PS-OPLS results for Aurora B, we determined kinetic parameters for Aurora B-dependent phosphorylation of the optimal peptide (ARRHSMGWAYKKK) and peptides with single amino acid
**A.**

Peptide Km

\[ V_m, (pmol/min/pg) \]

\[ V_{max} / K_m \]

**Optimal**

<table>
<thead>
<tr>
<th>Peptide</th>
<th>( K_m ) (pM)</th>
<th>( V_{max} ) (pmol/min/pg)</th>
<th>( V_{max} / K_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARRHSMGWAYKKKK</td>
<td>247</td>
<td>5425</td>
<td>22</td>
</tr>
<tr>
<td>R-2D tide</td>
<td>1238</td>
<td>2085</td>
<td>1.7</td>
</tr>
<tr>
<td>M+1P tide</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

**Optimal: ARRHSMGWAYKKKK**

**B.**

**C.**

<table>
<thead>
<tr>
<th>Substrate Protein</th>
<th>Mapped Aurora A Phosphorylation Site(s)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>AurkA</td>
<td>RTT(^{TM})L</td>
<td>(47)</td>
</tr>
<tr>
<td>Cdc25B</td>
<td>RRS(^{TM})V</td>
<td>(67)</td>
</tr>
<tr>
<td>HURP</td>
<td>RMS(^{TM})L</td>
<td>(68)</td>
</tr>
<tr>
<td>Lats2</td>
<td>RYS(^{TM})L</td>
<td>(69)</td>
</tr>
<tr>
<td>MBD3</td>
<td>RRS(^{TM})G</td>
<td>(70)</td>
</tr>
<tr>
<td>p53</td>
<td>RHS(^{TM})V</td>
<td>(71)</td>
</tr>
<tr>
<td>RafA</td>
<td>RKS(^{TM})L</td>
<td>(72)</td>
</tr>
</tbody>
</table>

**D.**

<table>
<thead>
<tr>
<th>Substrate Protein</th>
<th>Mapped Aurora B Phosphorylation Site(s)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>AurKB</td>
<td>RKT(^{TM})M</td>
<td>(73)</td>
</tr>
<tr>
<td>CENP-A</td>
<td>RRS(^{TM})R</td>
<td>(74)</td>
</tr>
<tr>
<td>Desmin</td>
<td>RTS(^{TM})G</td>
<td>(75)</td>
</tr>
<tr>
<td>Histone H3</td>
<td>RKS(^{TM})T, RKS(^{TM})A</td>
<td>(13, 76)</td>
</tr>
<tr>
<td>INCENP</td>
<td>RTS(^{TM})S</td>
<td>(50, 77)</td>
</tr>
<tr>
<td>MCAK</td>
<td>RKS(^{TM})C</td>
<td>(16, 78)</td>
</tr>
<tr>
<td>MgcRacGAP</td>
<td>RLS(^{TM})T, RRS(^{TM})T, RST(^{TM})S, RIS(^{TM})G</td>
<td>(79, 80)</td>
</tr>
<tr>
<td>Vimentin</td>
<td>RSS(^{TM})V</td>
<td>(73, 81)</td>
</tr>
</tbody>
</table>
Figure 2.4: Aurora A and Aurora B strongly select R in the -2 and hydrophobic amino acids in the +1, but strongly discriminate against P in the +1. A. PS-OPLS blots of Aurora A (left) and Aurora B:INCENP complex (right). B. Determination of kinetic parameters ($K_m$, $V_{max}$, and $V_{max}/K_m$ ratio) for reactions of Aurora B:INCENP with the optimal substrate peptide and peptides with single amino acid changes from the optimal. Graph of reaction data fitted to the Michaelis-Menten equation for peptides as indicated (left) and table of kinetic parameters (right). C. Previously published mapped Aurora A (left) and Aurora B (right) phosphorylation sites contain R in the -2 and hydrophobic amino acids in the +1. D. Molecular surface representation of the Aurora B active site with the optimal substrate peptide, RRHSMGW, shown in stick representation. The electrostatic potentials are colored red (negative) and blue (positive).

Changes in which the R in the -2 was changed to a D and the M in the +1 was changed to a P (Figure 2.4B). Changing R to D resulted in a 13-fold drop in $V_{max}/K_m$ as compared to the optimal peptide. As with Plk1, the activity of Aurora B against the P in the +1 peptide so low that it was not possible to fit the data to a Michaelis-Menten equation. Based on these data the minimal optimal consensus sequence for both Aurora A and Aurora B is R-X-S/T-Φ, where X is any amino acid and Φ is any hydrophobic amino acid except P.

To investigate the structural basis for the optimal motif we performed molecular modeling studies using the previously published structure of the Aurora B kinase domain (Figure 2.4C) (33). The optimal peptide sequence, RRHSMGW, was modeled into the active site in an extended conformation. Selection for R in the -3 seems to be based on an electrostatic interaction with E177. Selection for R in the -2 is likely due to an electrostatic interaction with E220 and E281. Consistent with this model, replacement of the -2 R by a D resulted in a 5-fold decrease in $K_m$. M and other hydrophobic amino acids in the +1 position appear to fit into a hydrophobic pocket formed by W237 and M249 on the activation loop and L256 of the active site cleft.

The motifs of both Aurora A and Aurora B are consistent with mapped Aurora A and Aurora B phosphorylation sites (Figure 2.4C) as most Aurora phosphorylation sites
have an R in the -2 and a hydrophobic in the +1. The Aurora A motif presented here is consistent with previously published consensus motifs (47, 48). The Aurora B motif, R-X-S/T-Φ, corrects a previously deduced motif, R/K-R/K-X-S/T- φ (49, 50), which claimed that K is an acceptable substitute for R in the -2 and allowed P in the +1. These data suggest that Aurora A and Cdk1 phosphorylate mutually exclusive sites, as do Aurora B and Cdk1.

Determination of the Optimal Consensus Motif of Nek2

To determine if the other major mitotic kinase, Nek2, also follows the emerging pattern of not phosphorylating sites with P in the +1, we performed PS-OPLS Nek2 (Figure 2.5A). Nek2 showed strong discrimination against non-hydrophobic amino acids in the -3 position with particularly strong positive selection for F, L, and M. In the -2 position, Nek2 showed no strong preference, although R was slightly preferred to other amino acids and P was disfavored. In the +1 position, Nek2 favored hydrophobic amino acids and showed extremely strong discrimination against P in the +1. Nek2 also showed discrimination against D and E in the +1.

To verify the PS-OPLS data we measured kinetic parameters for Nek2-dependent phosphorylation of its optimal peptide (WFRMSIRGYKKK) and peptides containing single amino acid changes from the optimal. Changing the F in the -3 of the optimal to a V resulted in an approximately 8-fold decrease in the $V_{\text{max}}/K_m$ as compared to the optimal peptide, primarily due to a reduction in $K_m$. V was chosen because it was neither favored nor disfavored in the -3 position. Non-hydrophobic amino acids, which were discriminated against in the -3 position, would be expected to result in even larger decreases $V_{\text{max}}/K_m$. As with Plk1 and Aurora B, the activity of Nek2 against the P in the +1 peptide was so low that it was not possible to fit the data to a Michaelis-Menten curve. Based on these data the minimal optimal consensus motif for Nek2 is $\varphi$-R/X-X-S/T-Φ-R/H, where $\varphi$ is any hydrophobic amino acid, X is any amino acid, and Φ is any hydrophobic amino acid except P.

To investigate the structural basis for the optimal motif we performed molecular modeling studies using the previously published structure of the Nek2 kinase domain, and since in this structure the activation loop is disordered, a model of the PKA:substrate
Figure 2.5: Nek2 strongly selects hydrophobic amino acids in the -3 and the +1, but strongly discriminates against P in the +1. A. PS-OPLS blot of Nek2. B. Determination of kinetic parameters ($K_m$, $V_{max}$, and $V_{max}/K_m$ ratio) for reactions of Nek2 with its optimal substrate peptide and peptides with single amino acid changes from the optimal. Graph of reaction data fitted to the Michael-Menten equation for peptides as indicated (left) and table of kinetic parameters (right). C. Previously published mapped Nek2 phosphorylation sites contain hydrophobic amino acids in the -3 and the +1. D. Molecular surface representation of the Nek2 active site with the optimal substrate peptide, FRASIR, shown in stick representation. The electrostatic potentials are colored red (negative) and blue (positive).
complex as base models (Figure 2.5C) (34). The optimal peptide sequence, FRASIR, was modeled into the active site in an extended conformation. Selection for F and other hydrophobic amino acids in the -3 seems to be based on a surface formed by two alanines that are F and E, respectively, in PKA. Selection for R in the -2 may be due to an electrostatic interaction with E208. I and other hydrophobic amino acids in the +1 of the peptide likely fit into a hydrophobic pocket formed by F176 in the activation loop, and P180 and M183 of the active site. The R and H in the +2 position of the peptide may electrostatically interact with E48, which could also form hydrogen-bonds to hydroxyl groups of S and T.

There are only two mapped Nek2 phosphorylation sites including the activation loop autophosphorylation site, T175 (Figure 2.5D). The motif of Nek2 is consistent with these mapped sites, which both contain F in the -3 position. No motif for Nek2 has been published previously. Importantly, this motif confirms our hypothesis that Nek2, just like Plk1, Aurora A, and Aurora B will not phosphorylate sites targeted by Cdk1 and vice versa.

**Examination of the Role of Localization and Motifs on the Regulation of the Major Mitotic Kinases**

The motif data suggests a mode of regulation of mitotic kinase signaling based on the presence or absence of a P in the +1 as a kind of switch allowing phosphorylation of a site by either Plk1, Aurora A, Aurora B, or Nek2 if a residue other than proline is present in the +1 or Cdk1 if a proline is present in the +1. This regulation prevents Cdk1 from phosphorylating Plk1, Aurora A, Aurora B, or Nek2 sites and Plk1, Aurora A, Aurora B and Nek2 from phosphorylating Cdk1 sites. Previously, this proline specificity has been suggested as a mode of regulation between Cdk1 and Aurora A and between the CMGC proline-directed kinase group, of which Cdk1 is a member, and the basophilic CAMK and AGC kinase groups, to which Aurora A and Aurora B are closely related but are not members (47, 51). The data presented here expands this concept from just Aurora A to include the other major mitotic kinases, Aurora B, Nek2, and Plk1.

To attempt to understand how the major mitotic kinases aside from Cdk1 phosphorylate mutually exclusive sites despite having overlapping localizations, we re-
examined the role of kinase localization on substrate selection performing an unbiased bioinformatics analysis of potential sites for Plk1, Nek2, and Aurora kinases at the centrosome and on the spindle apparatus compared to the proteome as a whole (Figure 2.6A). The Plk1 motif was found to be enriched at the centrosome and spindle apparatus. The Nek2 motif was enriched at the centrosome, but not on the spindle as might be expected for a centrosomal protein. The Aurora motif was not significantly enriched at the centrosome or the spindle. The lack of enrichment of Aurora kinase sites is, perhaps, not surprising as there are many basophilic kinases that control processes outside of mitosis. The enrichment of potential substrates at subcellular structures where Plk1 and Nek2 localize supports the intuitive idea that these kinases are likely to phosphorylate proteins that co-localize to the areas of the cell with the largest abundance of these kinases rather than to areas of low abundance of the kinase, and suggests that proteins at these structures may have co-evolved to make use of the appropriate kinase for phospho-dependent regulation. This, in turn, suggests that localization may play a role in maintaining mutually exclusivity of major mitotic kinase phosphorylation sites.

To further explore this emerging potential relationship between localization and motifs in the regulation of major mitotic kinases we more closely examined the ability of these kinases to phosphorylate the same sequences. Although we have already determined kinetic parameters for Plk1 phosphorylating a D-2R variant of its optimal peptide (Figure 2.2B) and Aurora B and Nek2 phosphorylating R-2D variants of their optimal peptides (Figure 2.4B, 2.5B) and seen that these changes away from the optimal resulted in significant reductions in the $V_{max}/K_m$ ratio compared to the optimal, we wanted to directly compare these kinases on a common set of peptides. The optimal peptide of each kinase, Aurora B, Nek2, and Plk1, was reacted with each kinase in this set (Figure 2.6B). As expected, each kinase phosphorylated its optimal peptide most strongly. The amount of phosphorylation of the Plk1 optimal peptide by Nek2 and Aurora B was about 8% and less than 1% of the phosphorylation by Plk1, respectively. The amount of Aurora B and Plk1 phosphorylation of the Nek2 optimal peptide was about 23% and less than 1% of the phosphorylation by Nek2, respectively. The extent of Nek2 and Plk1 phosphorylation of the Aurora B optimal peptide was about 1% of that of Aurora B. These data suggest that kinases with non-overlapping localizations such
Figure 2.6: The motif of Nek2 overlaps with that of the Aurora kinases and Plk1, but the motifs of Plk1 and the Aurora kinases are non-overlapping. A. Table of significance scores for the enrichment of consensus motifs in centrosome and mitotic spindle proteins. B. Comparison of ability of each kinase (Plk1 in black, Nek2 in gray, and Aurora B in white) to phosphorylate each of the other kinases’ optimal substrate peptides as a percentage of the amount of phosphorylation of its own optimal substrate peptide. C. Pseudo-color PS-OPLS blots of Aurora A (yellow), Aurora B (purple), Nek2 (cyan), and Plk1 (red) were superimposed to show similarities and differences between the motifs. Equal selectivity in the Aurora A-Aurora B overlay is cyan, Aurora A-Nek2 overlay is red, Aurora A-Plk1 overlay is purple, Nek2-Plk1 overlay is maroon, and Aurora B-Plk1 overlay is green.
as Plk1 or Aurora kinases and Nck2 may be able to phosphorylate the same sequences, whereas kinases with overlapping subcellular localizations, such as Plk1 and Aurora kinases, are not able to phosphorylate the same sequences as a strict consequence of non-overlapping phosphorylation motifs.
Discussion

Mitosis is an intricate and highly regulated process. Several kinases including the master regulator Cdk1 and the other major mitotic kinases, Aurora A, Aurora B, Nek2, and Plk1, phosphorylate specific sites on specific proteins to orchestrate the process of mitosis. To better understand the mechanisms underlying the roles these kinases play in mitosis, we have determined unbiased, complete motifs of these kinases. The motif of Cdk1 has been published previously to be S/T-P-X-R/K, and we have verified this motif (46). The motif of Nek2 is novel. The motifs of Plk1, Aurora A, and Aurora B have been determined by PS-OPLS, with results that differed significantly from previously published, but biased and incomplete motifs (41, 49, 50). We further verified the PS-OPLS motifs by kinetic studies on individual peptides, testing the effects of altering amino acids away from the optimal for individual positions. The motifs match known substrates of these kinases and we were able to rationalize the structural basis for the motifs using structural models of the kinase domains with the optimal peptide docked in.

As part of this study we have identified new Plk1 substrates. Motifs are particularly useful when combined with phosphoproteomic mass spectrometry screens, where they can help to identify the kinases that created the discovered phosphosites. The previously reported motif of Plk1 suggested that Plk1 only phosphorylated sites with D or E in the -2 position (41). PS-OPLS revealed the complete motif showing additional specificity for N in the -2. The new motif was applied to a phosphoproteomic mass spectrometry screen in which Plk1 phosphorylation sites were determined by comparing nocodazole arrested mitotic HeLa cells treated with the specific Plk1 inhibitor BI 2536 to HeLa cells just arrested in mitosis with nocodazole (52). Sites in Sccl/Rad21, Bub1, and p31/Comet had N in the -2 matching the new Plk1 consensus sequence and are Plk1 activity dependent suggesting that they are novel Plk1 sites. Sccl is a subunit of the cohesin complex that prevents chromosome segregation until it is cleaved by separase at the metaphase-anaphase transition (53). Sccl is known to be phosphorylated at multiple sites by Plk1 enhancing the cleavage of Sccl by separase, although all of the sites have not been mapped (53). The location of the phosphorylation site discovered in this study is near one of the separase cleavage sites suggesting it may also function in increasing the
cleavability of Scc1. Bub1 has previously been shown to be phosphorylated by Cdk1 recruiting Plk1 to kinetochores, and Plk1 has been shown to phosphorylate Bub1 although the site had not been determined (54). P31/Comet has recently been shown to bind Mad2, causing Mad2 to release Cdc20 (55). The Mad2-Cdc20 interaction is required for maintenance of the spindle checkpoint until all kinetochores have obtained amphitelic attachments (56). Plk1 is known to be required for the recruitment of Mad2 to the kinetochore for the spindle assembly checkpoint, but this phosphorylation in p31/Comet suggests it may have other roles also (23).

Of the major mitotic kinases, Nek2 stands alone in its localization to the proximal ends of the centrioles at the core of the centrosome. The other major mitotic kinases have overlapping localizations with Aurora A, Plk1, and Cdk1 localizing to the pericentriolar material and spindle poles and Aurora B, Plk1, and Cdk1 localizing to the kinetochores and centromeres. As these kinases are all Serine/Threonine kinases, these overlapping localizations suggest that these kinases will have access to the same proteins and, therefore, the same potential phosphorylation sites. Since the mapped phosphorylation sites of each kinase do not contain very many, if any, common sites with any of the other kinases in the set, we wanted to understand how these kinases maintain their specificity.

The motifs in this study shed light on this problem. Figure 2.6C shows each motif in pseudo-color and pairs of motifs superimposed. The overlay of the Aurora A motif in green and Aurora B motif in purple shows how similar the motifs are, as the majority of spots are neither purple nor green but bluish to cyan -- the R in the -2, in particular, which dominates the motifs. As mentioned above, Aurora A and Aurora B have mutually exclusive localizations during mitosis, which suggests that these kinases, despite having essentially identical motifs, also phosphorylate mutually exclusive sites as they do not have access to the same sites unless the protein being phosphorylated has a localization overlapping with the localizations of both Aurora A and B. This is supported by the known substrates for these kinases (Figure 2.4C). Additional proteins known to form complexes with Aurora-A and -B, such as Bora and Survivin and INCENP, respectively, may also play a role in substrate selection.

The Nek2-Plk1 and Aurora A-Plk1 motif overlays show the selectivity in the -2 of the Plk1 and Aurora A motifs and the selectivity in the -3 of the Nek2 motif. The optimal
peptide of Nek2 contains an R in the -2 which suggests that optimal Nek2 sequences might be phosphorylatable by the Aurora kinases, which we have verified on individual peptides (Figure 2.6B). Nek2 may also be able to phosphorylate the same sequences as Plk1 if such sequences had hydrophobic amino acids in the -3 to satisfy the Nek2 motif, which we have verified on individual peptides (Figure 2.6B). However, Nek2 has been shown to localize to the proximal centrioles at the core of the centrosome, whereas Aurora A localizes to the periphery of the pericentriolar material, where Plk1 is also thought to localize (26). Nek2 and Aurora A or Plk1 should, therefore, not be able to phosphorylate the same sites unless the substrate localizes to both the core of the centrosome and to the periphery of the pericentriolar material.

In the Aurora A-Plk1 motif overlay, with Aurora A in green and Plk1 in cyan, the R in the -2 is green and the D, N, and E in the -2 are cyan. Similarly, in the Aurora-B-Plk1 motif overlay with Aurora B in purple and Plk1 in cyan, the R in the -2 is purple and the D, N, and E in the -2 are cyan. This difference in selection in the -2 suggests that Plk1 and Aurora kinases, despite having overlapping localizations, will not phosphorylate the same sequences, which we verified using peptides (Figure 2.6B).

Integrating all of this data suggested to us a mode of regulation to enforce specificity of phosphorylation by these kinases. Plk1 and the Aurora kinases do not phosphorylate the same sites despite overlapping localizations because their motifs are non-overlapping. Nek2 and Plk1 and the Aurora kinases do not phosphorylate the same sites despite overlapping motifs because of non-overlapping localizations. Thus we hypothesize that the major mitotic kinases with overlapping localizations do not have overlapping motifs and major mitotic kinases with overlapping motifs do not have overlapping localizations.

Another way to visualize this hypothesis is shown in the Venn diagrams depicted in Figure 2.7. Kinases exist in two functionally orthogonal spaces: 'localization space' and 'motif space'. 'Localization space' reflects all of the subcellular locales where a protein can reside. 'Motif space' contains the optimal sequence motifs of all S/T kinases. In 'localization space', the red circle represents the localization of Cdk1, which overlaps
Figure 7: An alternate view of the proposed hypothesis in which kinase functionality is represented by Venn diagrams of ‘localization space’ and ‘motif space’. In ‘localization space’, each circle represents the total subcellular locales available to the kinase. In ‘motif space’, each circle represents the sequences that can be phosphorylated by the kinase. In this representation, a major mitotic kinase can overlap with every other kinase in at most one of these two spaces.

with all of the other major mitotic kinase. Similarly, Plk1, whose localization is represented by the blue circle, overlaps with the localizations of Aurora A (green) and Aurora B (purple) localizations, but not Nek2 (yellow). The ‘motif space’ panel is interpreted similarly. In the context of these two spaces an equivalent statement of the hypothesis becomes evident: Each major mitotic kinase can overlap every other kinase in, at most, one of localization or motif space. As examples, Cdk1 overlaps all of the other kinases in ‘localization space’, but not in ‘motif space’ and Nek2 overlaps Plk1 and Aurora Kinases in ‘motif space’, but not in localization spaces’. This representation begs the question of how temporal aspects of ‘regulation fit into this model. Although, prior events undoubtedly play roles in which substrates and which sites in these substrates are phosphorylated by the major mitotic kinases, this aspect of regulation is beyond the scope of this study.
This hypothesis, if true, implies a coordinated regulation of these kinases. It has been known for some time that motifs and localization direct individual kinases to their substrates. However, what has been suggested by our data is that the motifs and localizations of individual kinases may be cooperatively directing a group of kinases to their substrates and that the major mitotic kinases may be under a coordinated selective pressure to maintain their motifs and localizations, suggesting a systems-level regulation and evolution of the major mitotic kinases and their substrates.
References


Chapter Three

Determination of the Substrate Specificities of the Polo-like Kinase Family Reveals that Plk3 has a Motif that Changes Depending on the Phosphorylation Status of the Activation Loop

Jes Alexander, Dan Lim, Michael Yaffe

Contributions:
Dan and I determined the motif of Plk1 T210D together (Figure 3.4 and 3.5). I performed all the other experiments. I wrote all the sections and created all the figures.
Abstract

The Polo-like kinase (Plk) family in humans consists of four members, Plk1, Plk2, Plk3, and Plk4. These kinases are non-redundant regulators of the cell cycle. Although the role of Plk1 in mitosis has been extensively explored, the roles of the other Plk family members are just beginning to be understood and few substrates have been identified. The complete, unbiased motifs of Plk2, Plk3, and Plk4 have not been reported. Using Positional Scanning Oriented Peptide Library Screening, we have determined the motifs of all four wildtype Plk family members as well as the motifs of activation loop phosphomimic mutants of these kinases. The four kinases have similar, but distinct acidophilic motifs. Therefore, these kinases are likely to phosphorylate some of the same sites, but will also phosphorylate different sites on proteins. The motifs of Plk2 and Plk3 include specific selectivity for phosphotyrosine, suggesting that phosphorylation of potential substrates by a tyrosine kinase may prime these substrates for phosphorylation by Plk2 or Plk3. The motif of wildtype unphosphorylated Plk3 is dramatically different from the motif of the Plk3 T219D phosphomimic mutant, suggesting that phosphorylation on the activation loop of Plk3 may act as a switch to change the motif of the kinase. \textit{In vitro} phosphorylation of the kinase verified the phosphorylation-based motif switch. This is the first report of kinase that has a motif that is non-static. Finally, we examined the ability of the Plk3 Polo-box domain to bind phosphopeptides and find that it binds unphosphorylated peptides more strongly than phosphorylated peptides, suggesting that it does not function as a phospho-dependent binding domain in substrate selection.
Introduction

The Polo-like kinase (Plk) family in humans consists of four members, Plk1, Plk2, Plk3, and Plk4. These kinases are non-redundant regulators of the cell cycle. Each member of the Plk family has an N-terminal kinase domain and at least one Polo-box motif in the C-terminus. The family is named after the founding member in Drosophila, Polo. Polo, itself, was named after the phenotype of abnormal spindles poles, caused by mutation of the gene (1).

The ortholog of Polo in higher eukaryotes is Plk1, the sequence of which is conserved from yeast (where the kinase is named Cdc5) to humans. Plk1 is a key regulator of M-phase, involved in centrosome separation and maturation, mitotic entry, spindle assembly, the spindle assembly checkpoint, and cytokinesis (2). An incomplete motif has been published as D/E-X-S/T-ϕ-D/E, where ϕ is any hydrophobic amino acid (3). We have determined an unbiased, complete motif of Plk1 previously to be D/N/E-X-S/T-Φ, where Φ is any hydrophobic amino acid except proline (Figure 2.2). Many substrates of Plk1 are known (Chapter 1, Figure 2.3). The two polo-box motifs in Plk1 were found to fold together into a phospho-dependent binding domain, the Polo-box domain (PBD), which binds the motif S-pS/pT (4). In contrast to Plk1, very little is known about the other Plk family members.

Complete, unbiased kinase motifs of Plk2 and Plk3 have not been reported, but the binding motifs of the Plk2 and Plk3 PBDs, determined by phosphopeptide library screening, were reported to be the same as that of the Plk1 PBD, S-pS/pT (5). Plk2 and Plk3 are known to have post-mitotic roles in neurons where Plk2 is known to phosphorylate spine associated RapGAP (SPAR), but the phosphorylation site(s) were not identified (6-8). Within the cell cycle, Plk2 seems to play a role in centriole duplication, but no substrates have been identified (9).

Plk3’s role in the cell cycle is controversial. Plk3 has been reported to play a role in stress response pathways activated by oxidative stress and DNA damage (10, 11). As part of its role, it has been reported to phosphorylate p53 S20 and Cdc25C S216 (11, 12). Both sites are involved in arresting the cell cycle, although the amino acid sequences flanking the phosphorylated residues are quite different. Surprisingly, Plk3 has also been
reported to phosphorylate Cdc25C at S191 to enable mitotic entry (13). Plk3 has also been reported to have other pro-cell cycle roles being required for entry into S-phase and playing roles in golgi fragmentation in mitosis, and in cytokinesis (14-16).

Plk4 may be required for mitosis, as the Plk4 knockout mouse is inviable and the embryo dies with large numbers of cells in mitosis with high Cyclin B levels (17). Plk4 has also been shown to be required for centriole duplication explaining the presence of abnormal spindles, which are seen in both overexpression and loss of function studies of Plk4 (18-20). No substrates of Plk4 have been identified. Unlike the other Plk family members, Plk4 only has a single polo-box motif, which may be involved in homodimerization rather than phospho-dependent binding of substrates (21).

Identification of substrates should elucidate the roles of Plk2 and Plk4 in centriole duplication and clarify the conflicting literature on Plk3. Importantly, many kinases show strong selection for residues 3 amino acids N-terminal to the phospho-acceptor. For the Ser-20 phospho-acceptor in p53, this residue is a Glu, while for the Ser-216 site in Cdc25C, this residue is an Arg, making it unlikely that both sites are truly phosphorylated directly by Plk3. To further clarify optimal phosphorylation motifs and aid in substrate identification, we have determined the motifs of wildtype Plk1, Plk2, Plk3, and Plk4 and the motifs of activation loop phosphomimic mutants of these kinases using Positional Scanning Oriented Library Screening (PS-OPLS). Plk family members have similar, but distinct acidophilic motifs. The motif of Plk3 is particularly interesting. Specific selection for phosphotyrosine suggests that phosphorylation of the substrate by a tyrosine kinase may prime for Plk3 phosphorylation of the substrate. Additionally, Phosphorylation on the activation loop of Plk3 acts as a switch, dramatically changing the motif of the kinase. This is the first report of a kinase with a motif that changes as a function of activation loop phosphorylation.
Experimental Procedures

Kinase Protein Production and Purification

To produce wildtype and T239D Plk2 kinase domain protein, wildtype, K106A, R184A, R209A, K217A, T219D, and T219A Plk3 kinase domain protein, and wildtype and T170D Plk4 kinase domain protein, nucleotide sequences encoding the kinase domains of human Plk1 (aa. 38-346), Plk2 (aa.53-359), Plk3 (aa.49-345), and Plk4 (aa.1-290) were cloned separately into the NdeI and XhoI restriction sites in a modified pET28a expression vector (Novagen) (pET28A-His-MBP, courtesy of Dan Lim). Mutations were introduced using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and verified by sequencing. The resulting fusion proteins contain an N-terminal hexa-His-tag followed by MBP and a Tev protease cleavage site between MBP and the kinase domain. Protein was expressed in E. coli Rosetta (Novagen) cells.

For Wildtype and T239D Plk2 kinase domain, wildtype, K106A, R184A, R209A, K217A, T219D, and T219A Plk3 kinase domain, and T170D Plk4 kinase domain fusion proteins, starter cultures were grown overnight in LB. 50 mL of TB culture was initiated by addition of the starter culture at a 1:100 dilution. After 4 hours of growth, the 50 mL of TB culture was cooled to 4°C on ice, IPTG was added to a final concentration of 1 mM, and the cultures were incubated at room temperature for 16 hours. Cells were pelleted by centrifugation at 4000 X g for 10 min at 4°C, resuspended in 5 mL of lysis buffer (50 mM Tris, 500 mM NaCl, 1 mM DTT), and lysed by sonication using a Branford sonicator with a probe tip (power level 4, duty cycle 50%, 2:33 min:sec) in an ethanol-ice bath. The lysate was clarified by centrifugation at 18,000 RPM in a Sorvall SL-50T rotor in a Sorvall Super T21 table top centrifuge for 30 min at 4°C. The supernatant was added to 0.5 mL of packed amylose beads and end-over-end rotated a 4°C overnight. Beads were washed 5 times rapidly with lysis buffer and protein was eluted by addition of 1 mL of elution buffer (50 mM Tris, 500 mM NaCl, 1 mM DTT, 100 mM maltose) and end-over-end rotation for 2 hours at 4°C.

To produce wildtype and T219A PKA phosphorylated Plk3 kinase domain protein, the wildtype and T219A Plk3 kinase domain plasmid constructs described above were used. Nucleotide sequence encoding the catalytic domain of PKA (aa.11-352) was
cloned into the pCDFDuet-1 expression vector (Novagen). Protein was expressed in *E. coli* Rosetta (Novagen) cells containing either the Plk3 wildtype or Plk3 T219A and the PKA expression vectors. Expression and purification was as described for other Plk3 constructs above.

Wildtype Plk4 kinase domain protein was expressed in *E. coli* Rosetta (Novagen) cells. A starter culture was grown overnight in LB. 2L of TB culture was initiated by addition of the starter culture at a 1:100 dilution. After 4 hours of growth, the 2L of TB culture were cooled to 4°C on ice, IPTG was added to a final concentration of 1 mM and the cultures were incubated at room temperature for 16 hours. Cells were pelleted by centrifugation at 4000 X g for 10 min at 4°C, the pellet from each liter of culture was resuspended in ~40 mL of buffer (50 mM Tris, 500 mM NaCl, 14 mM β-mercaptoethanol), transferred to separate 50 mL Falcon tubes, and lysed by sonication using a Branford sonicator with a probe tip (power level 8, duty cycle 50%, 2:33 min:sec, 2 times) in an ethanol-ice bath. The lysate was clarified by ultracentrifugation at 40,000 RPM in a Beckman 45Ti rotor for 30 min at 4°C. Initial purification utilized Ni-NTA affinity chromatography. Following ultracentrifugation, the supernatant was mixed with 25 mL of packed Ni-NTA beads by end-over-end rotation at 4°C for 2 hours. The bead slurry was loaded onto a column for FPLC with the aid of a peristaltic pump. The column was transferred to the FPLC, washed with 280 mL of wash buffer A (500 mM NaCl, 14 mM β-mercaptoethanol, 10 mM imidazole, 50 mM Phosphate, pH 7), and the protein was eluted in buffer with the same composition except containing 300 mM imidazole. The kinase domain was further purified by affinity chromatography on amylose beads. Fractions containing Plk4 protein from the Ni-NTA column, as determined by on-line UV absorbance and SDS-PAGE, were loaded onto a column containing 25 mL of packed amylose beads, washed with 100 mL of wash buffer B (50 mM Tris, 500 mM NaCl, 14 mM β-mercaptoethanol, 10 mM maltose), and the protein was eluted in buffer with the same composition except containing 100 mM imidazole. Fractions containing Plk4 protein were pooled, 5 mg of hexa-his tagged Tev protease was added to cleave off the hexa-His-MBP tag, and the mixture was incubated at 4°C overnight. To remove the cleaved tag and Tev protease, the protein was re-passaged through Ni-NTA using the same conditions as described above. The recovered material
was applied to a Superose 12 column (16 mm X 30 cm) and eluted using gel filtration buffer (500 mM NaCl, 2 mM DTT, 10 mM Tris, pH 8). Fractions (2 mL) were analyzed for Plk4 protein by on-line UV absorbance and SDS-PAGE.

To produce wildtype and T210D Plk1 kinase domain, nucleotide sequence encoding the kinase domain of human Plk1 (aa. 38-346) was cloned into NdeI and XhoI restriction sites in the pET28a bacterial expression vector (Novagen). The resulting fusion protein contains an N-terminal hexa-His-tag and a thrombin protease cleavage site between hexa-His-tag and the kinase domain. A Thr-210 mutation to Asp was introduced using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and verified by sequencing. Protein was expressed in *E. coli* Rosetta (Novagen) cells. A starter culture was grown overnight in LB. 6L of TB culture was initiated by addition of the starter culture at a 1:100 dilution. After 4 hours of growth, the 6L of TB culture were cooled to 4°C on ice, IPTG was added to a final concentration of 1 mM and the cultures were incubated at room temperature for 16 hours. Cells were pelleted by centrifugation at 4000 X g for 10 min at 4°C, the pellet from each liter of culture was resuspended in ~40 mL of lysis buffer (50 mM Tris, 500 mM NaCl, 14 mM β-mercaptoethanol), transferred to separate 50 mL Falcon tubes, and lysed by sonication using a Branford sonicator with a probe tip (power level 8, duty cycle 50%, 2:33 min:sec, 2 times) in an ethanol-ice bath. The lysate was clarified by ultracentrifugation at 40,000 RPM in a Beckman 45Ti rotor for 30 min at 4°C. Initial purification utilized Ni-NTA affinity chromatography. Following ultracentrifugation, the supernatant was mixed with 25 mL of packed Ni-NTA beads by end-over-end rotation at 4°C for 2 hours. The bead slurry was loaded onto a column for FPLC with the aid of a peristaltic pump. The column was transferred to the FPLC, washed with 280 mL of wash buffer (500 mM NaCl, 14 mM β-mercaptoethanol, 10 mM imidazole, 10 mM Phosphate, pH 7), and the protein was eluted using buffer of the same composition except containing 300 mM imidazole. Fractions were analyzed by on-line UV absorbance and SDS/PAGE. Fractions containing Plk1 kinase domain protein were pooled, 100 units of thrombin protease were added to cleave off the hexa-His-MBP tag, and the mixture was incubated at 4°C overnight. To remove the cleaved tag, the protein was re-passaged through Ni-NTA using the same conditions as described above. The resulting ~10 mL of material was applied to a Superose 12 column (16 mm X
30 cm) and eluted using gel filtration buffer (500 mM NaCl, 2 mM DTT, 10 mM Tris, pH 8). Fractions (2 mL) were analyzed for Plk1 content by on-line UV absorbance and SDS-PAGE.

**Phosphorylation Motif Determination by Peptide Library Array Screening**

Positional Scanning Oriented Peptide Library Screening (PS-OPLS) was performed following Hutti et al. (22). Briefly, solution-phase kinase reactions were performed in parallel on 198 separate biotinylated, partially degenerate oriented peptide libraries (Anaspec, Inc) arrayed in a 384 well microtiter plate in a 22 row X 9 column format. Each peptide library contains an N-terminal biotin tag, a 50:50 mix of serine and threonine at the orienting phospho-acceptor residue, a single second fixed amino acid located between the -5 and +4 position, and a mixture of amino acids at all other positions. Individual libraries contain any of the 20 natural amino acids as well as phosphothreonine and phosphotyrosine in the second fixed position, corresponding to the 22 rows. Scanning down the rows in the array moves the position of the fixed amino acid from -5 to +4 relative to the fixed phospho-acceptor residue, while scanning across the columns changes the identity of the second fixed amino acid. As an example, the peptide library with Lys fixed in the -4 position has the following sequence: Y-A-X-K-X-X-X-S/T-X-X-X-X-A-G-K-K-biotin, where amino acids are represented in 1-letter code, and X is an equal mixture of all 17 natural amino acids excluding Cys, Ser, and Thr to prevent oxidation effects and eliminate secondary phosphorylation events, respectively. S/T denotes a 50:50 mix of Ser and Thr. Kinase reactions were performed at 30°C in a total volume of 16 µl containing 31.25 µM peptide library, 100 µM ATP, and 200 µCi of [³²P]-γ-ATP, in 150 mM NaCl, 10 mM MgCl₂, 1 mM DTT, 0.1% Tween 20, and 50 mM Tris, pH 7.5. Reactions for the Plk1 T210D and wildtype motifs were performed for 3 hours with 0.5 µg of protein per reaction and 3 hours with 5 µg of protein per reaction, respectively. Reactions for the Plk2 T239D motif and Plk4 wildtype motif were performed for 6 hours with 0.5 µg of protein per reaction and 2 hours with 0.1 µg of protein per reaction. Reactions for Plk3 wildtype and T219D were performed for 6 hours with 1 and 1 µg of protein per reaction, respectively. Reactions for Plk2 wildtype motif and Plk4 T170D motif were performed for 6 hours with 0.5 µg of protein per reaction and
2 hours with 10 μg of protein per reaction, respectively. Following incubation, 2 μl of each reaction were simultaneously transferred to a streptavidin-coated membrane (Promega SAM² biotin capture membrane) using a 384 slot pin replicator (VP Scientific). The membrane was washed three times with 140 mM NaCl, 0.1% SDS, 10 mM Tris, pH 7.4, three times with 2 M NaCl, twice with 2 M NaCl containing 1% H₃PO₄, and once with water. The extent of peptide library phosphorylation was determined by imaging the membrane with a phosphorimager (Molecular Dynamics).

**In vitro Kinase Assays**

All kinase assays were performed at 30°C for 45 minutes in a total volume of 40 μl containing kinase reaction buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 10 mM MgCl₂, 100 μM ATP, 9 μCi [³²P]-γ-ATP, and 1 mM DTT) and the amounts of kinase domain protein and peptide substrate indicated below. For Plk3, 0.6 μg of the indicated wildtype or mutant kinase domain protein were reacted against 0.5 mM pY+2tide, with sequence GEDTSFpYWAAYKKKK, or Y+2tide, with sequence GEDTSFYWAAYKKKK. For Plk1, 0.006 μg of wildtype or T210D kinase domain protein were reacted with 0.5 mM Plktide, with sequence GHDTSFYWAAYKKKK. For Plk2, 0.12 μg of wildtype or T219D kinase domain protein were reacted with 0.5 mM Y+2tide or pY+2tide. For Plk4, 0.019 μg of wildtype or T170D kinase domain protein were reacted with 0.5 mM Plk4tide, with sequence GHETSFYWAAXKKK. Reactions were performed in triplicate and 5 μL of each reaction was spotted on phosphocellulose at 0 minutes and 45 minutes. The phosphocellulose paper was washed 4 times with 0.5% phosphoric acid, added to scintillation fluid, and scintillation counted.

**IVT pulldown assays of Plk3 PBD**

Six degenerate peptide libraries were used, which each had either a fixed phosphoresidue or the corresponding non-phosphoresidue, which was flanked on either side with four positions containing degenerate mixtures of amino acids. Additionally, each library was biotinylated at its N-terminus. The phosphoresidue was one of pT, pS, or pY. As an example, the pT peptide library had the following sequences: biotin-Z-G-Z-G-X-X-X-pT-X-X-X-X-Ala-K-K, where Z indicates aminohexanoic acid, and X denotes all
amino acids except C. Streptavidin beads (Pierce, 75pmol/μL gel) were incubated with a five-fold molar excess of each biotinylated library in 20 mM Tris/HCl (pH 7.5), 125 mM NaCl, 0.5% NP-40, 1 mM EDTA and washed four times with the same buffer to remove unbound peptide. The bead immobilized libraries (30μL gel) were added to 6 μL of an in vitro translated 35S-labeled phospho-binding domain construct in 200 μL binding buffer (20 mM Tris/HCl (pH 7.5), 125 mM NaCl, 0.5% NP-40, 1 mM EDTA, 1 mM DTT, 4 μg/mL pepstatin, 4 μg/mL aprotinin, 4 μg/mL leupeptin, 200 μM Na3VO4, 50 mM NaF). Plk3 PBD, Plk1 PBD, and Src SH2 domain constructs included the amino acids indicated from nucleotides sequences cloned into pCDNA3.1+. After incubation at 4°C for 2-3 hours, the beads were rapidly washed 3 times with binding buffer prior to SDS-PAGE (12%) and autoradiography. Determination of binding was based on comparison of the resulting bands from the pulldown with phosphopeptide library and the non-phosphopeptide library.
Results

Kinase Motifs of wildtype Polo-like Kinase family members

Activation loop phosphorylation is a common mechanism of activation of most kinases. However, most kinases also have a residual level of activity even without phosphorylation on the activation loop. It has previously been shown that Plk1 is phosphorylated at T210 on the activation loop and that this phosphorylation results in activation (23, 24). The phosphomimic T210D mutation has been shown to at least partially replicate this activation. This T is conserved in the other three human Polo-like kinase family members. An activation loop phosphomimic mutation has been published to activate Plk2 also (7). For Plk3 or Plk4, neither the phosphorylation status of the activation loop T nor the effect of activation loop phosphorylation or phosphomimic mutations on kinase activity has been reported. To the best of our knowledge there have been no reports of a difference in the specificity of a kinase when in the less active non-phosphorylated state as compared to the more active phosphorylated state. As we had no reason to suspect that the wildtype, unphosphorylated Polo-like kinase family members would behave any differently than other known kinases, we first determined the kinase motifs of the wildtype kinases by Positional Scanning Oriented Peptide Library Screening (Figure 3.1).

In the Plk1 PS-OPLS blot, the darkest spots, indicating strongly selected amino acids in the optimal phosphorylation motif, correspond to D, N, and E in the -2 and hydrophobic amino acids, particularly F, in the +1 (Figure 3.1). The lightest spot on the blot, indicating a residue that is specifically discriminated against in the phosphorylation motif, corresponds to P in the +1. The resulting optimal consensus motif of Plk1 is D/N/E-X-S/T-Φ, where X is any amino acid, S/T is the phosphoacceptor residue and Φ is any hydrophobic amino acid except P. This motif is the same as the Plk1 T210D motif (Figure 2.2 and Figure 3.1). The optimal consensus motif of Plk4 T170D is D/N/E-X-S/T-φ-Y/W, where φ is any hydrophobic amino acid. This motif is similar to but not exactly the same as a motif of Plk4 that was published while this study was underway. That approach employed phosphorylation of arrays of peptides synthesized on a membrane, in which each position of an individual reference peptide was replaced with
Figure 3.1: Positional Scanning Oriented Peptide Library Screening (PS-OPLS) blots of wildtype Polo-like kinase (Plk) family kinase domains. The blot shows relative amounts of radioactive phosphate incorporation in *in vitro* kinase assays into degenerate peptide libraries with the indicated amino acid fixed at the indicated position relative to the phosphoacceptor S/T. The darker the spot, the more highly selected the amino acid indicated by the columns is at the position indicated by the row.

Each of the 20 amino acids while keeping the other amino acids of the peptide fixed (25). In contrast to this approach, PS-OPLS systematically tests each amino acid in each position independently of all others positions resulting in unbiased motifs and each kinase reaction in this technique occurs in solution which results in fewer artifacts than methods involving substrate attached to solid supports (22). The motifs of Plk2 and Plk3 are very...
similar, but not the same. In both motif blots, the pY in the +2 is the strongest spot. However, for Plk2, the E in the -3 and the F in the +1 are almost as strong as the pY in the +2. For Plk3, the selectivity for pY in the +2 dominates the blots as no other spot is nearly as dark. The optimal consensus motif of Plk2 is E-X-X-S/T-F-pY and of Plk3 is S/T-X-pY, where pY is phosphotyrosine. This is an intriguing result as no other kinase is known to select specifically for pY in its motif, and it suggests that tyrosine kinase phosphorylation primes substrates for Plk3 phosphorylation and, perhaps for Plk2 phosphorylation as well.

To verify this unexpected motif, we compared the activity of Plk3 against an optimal peptide containing pY in the +2 (pY+2tide) with its activity against a peptide with the same sequence except with Y substituted for pY in the +2 (Y+2tide) (Figure 3.2A, 3.2B). We chose to focus on Plk3 since the pY in the +2 was dominant in the Plk3 motif, but less so in the Plk2 motif. The activity of Plk3 against the pY+2tide was approximately three times the activity against the Y+2tide confirming the selectivity for pY in the Plk3 motif revealed by PS-OPLS.

The pY in the motif of Plk3 interacts with the basic pocket that in most kinases interacts with the activation loop phosphorylation

The motif of Plk3 is reminiscent of the unusual Gsk3β motif, which is S/T-X-X-X-pS/pT, where pS is phosphoserine and pT is phosphothreonine. As is indicated by its motif, a priming S or T phosphorylation 4 amino acids C-terminal to the phosphoacceptor S or T on the substrate is required to allow Gsk3β to phosphorylate the phosphoacceptor residue on that substrate (26). Gsk3β is also unusual in that it does not require phosphorylation on its activation loop to attain an active confirmation (27). In most kinases including MAP kinases, Cyclin-dependent kinases, PKA, Akt, and Aurora kinases, phosphorylation on the activation loop allows the kinase to attain an active state through conformational changes caused by interaction of the phosphate with a basic pocket composed of positively-charged residues on the catalytic loop and N-terminal lobe of the kinase (27, 28). In contrast, the X-ray crystal structure of Gsk3β suggests that the pS/pT priming phosphorylation on the substrate fits into this basic pocket instead, which accounts for the pS/pT-containing motif of Gsk3β (27).
Figure 3.2: *In vitro* kinase assays comparing activity of Plk3 against a peptide with pY in the +2 position (pY+2tide) with a peptide with the same sequence except with Y in the +2 (Y+2tide) show that A. overall Plk3 activity is unaffected by activation loop phosphomimic mutation (T219D), but that the activity of the kinase against a peptide with pY in the +2 is reduced and B. that this reduction in activity against the pY+2tide is recapitulated by phosphorylation at T219 and may be due to an interaction between K106 and the pY on the substrate peptide. Data in B is shown as the logarithm base two of the ratio of the counts per minute (CPM) of the pY+2tide to the Y+2tide for ease of visualization.
Since the motifs of Plk3 and Gsk3β are similar, we hypothesized that like in the case of Gsk3β, the priming phosphorylation on the substrate might be interacting with the basic pocket in Plk3. To test this hypothesis we measured the activity of mutants, in which residues potentially forming the basic pocket were mutated to alanines, against pY+2tide and Y+2tide (Figure 3.2B.). As the structure of Plk3 has not been solved, but the Polo-like kinase family is most similar to AGC kinases, we chose potential Plk3 basic pocket residues based on sequence alignment of Plk3 with PKA and Aurora A, for which the basic pocket is known based on solved X-ray crystal structures of these kinases. The amino acids in PKA that form the basic pocket are H87 on αC, R165 which is the RD arginine on the catalytic loop, and K189 on β9. The equivalent residues in Plk3 are K106, R184, and A208, respectively. Since A208 is not a basic residue, but R209 is, we decided to include this residue as a potential basic pocket residue instead of A208. Aurora A also does not have a basic residue equivalent to K189. Instead it has a lysine two amino acids N-terminal to the phosphoacceptor T on the activation loop that is part of the basic pocket. We included the equivalent residue, K217 in Plk3, in the group of potential basic pocket residues. For easier visual comparison, the ratios of activity against the pY+2tide to the activity against the Y+2tide in Figure 3.2B are shown as log ratios. The ratio of the activity of the unphosphorylated, wildtype kinase against the pY+2tide to the activity against the Y+2tide, as mentioned previously, was approximately 3 resulting in a log ratio of 1.6. The log ratios of the activities of the R209A and K217A mutants were nearly the same as that of the wildtype at 1.5, but the log ratio for the K106A mutant was approximately -0.75, meaning that the activity of the K106A mutant against the Y+2tide was greater than against the pY+2tide. This suggests that K106 interacts with the pY on the substrate. The R184A mutation eliminated kinase activity making assessment of its role in binding the pY on the substrate impossible.

Another approach to interrogating the role of the basic pocket in pY selectivity is to test the activity of Plk3 constructs that have an intact basic pocket, but that prevent the pY on the substrate from interacting with the pocket. Since in most kinases, the activating phosphorylation on the activation loop fits into the basic pocket as do activation loop phosphomimic mutations, we would expect a phosphorylation or phosphomimic mutation on the activation loop of Plk3 to fit into the pocket and exclude
the pY on the substrate. We examined the effect of phosphorylation on T219 and the
effect of the T219D mutation on activity against the pY+2tide and Y+2tide (Figure 3.2B).
Plk3 was co-expressed in bacteria with PKA, which has a motif, R-X-X-S/T- φ, that
matches the sequence surrounding T219 on Plk3, 216-RKKT²¹⁹I-220. Plk3 protein was
then purified away from PKA protein by affinity chromatography using the His-MBP tag.
To control for phosphorylation of sites other than T219, we also co-expressed Plk3
T219A with PKA. The activities of PKA phosphorylated wildtype and PKA
phosphorylated T219A Plk3 were determined. The log ratio of PKA phosphorylated
T219A Plk3 was slightly higher than unphosphorylated wildtype, while the log ratio of
PKA phosphorylated wildtype was near zero, suggesting that phosphorylation on the
activation loop dramatically reduced the specificity of the kinase for the pY in the +2.
Recapitulating the results with PKA phosphorylated Plk3 and the basic pocket mutations,
activity of T219D Plk3 against the pY+2tide was brought down to the same level as the
activity against the Y+2tide resulting in a log ratio near zero.

Since negative charge on the activation loop could electrostatically repulse the pY
on the substrate irrespective of any effect on the basic pocket, we also tested the activities
of T219H, T219K, and T219R mutants, which will have increased positive charge on the
activation loop. The positive charge would be expected to electrostatically attract the pY
on the substrate and, consequently, result in an increased pY+2tide to Y+2tide activity
ratio as compared to the wild-type, unphosphorylated form, if the mechanism for the
reduced ratio in T219D and pT219 Plk3 is only based on electrostatic repulsion, rather
than via the basic pocket. The log ratios of these three mutants were all less than that of
the wildtype, unphosphorylated form. It therefore, seems likely that the pY on the
substrate fits into the basic pocket, which would explain the remarkable pY-directed
motif.

**Activation Loop Phosphomimic mutation does not increase the activity of all
members of the Polo-like kinase family**

The activities of wildtype, unphosphorylated and T219D Plk3 against the Y+2tide
were approximately the same, despite higher activity of the wildtype, unphosphorylated
form of Plk3 against the pY+2tide as compared to T219D Plk3 (Figure 3.2A, 3.2B). This
suggests that the overall activity of the kinase does not change with the activation loop phosphomimic mutation. This is unexpected as the sequence homology between Plk3 and Plk1 is quite high and Plk1 activity has been shown to increase with the activation loop phosphomimic mutation (23).
This result led us to wonder whether Plk4 was activated by an activation loop phosphomimic mutation like Plk1 or unaffected like Plk3. In addition, we wanted to know the effect of the phosphomimic mutation on the activity of Plk2 against the pY+2 tide as compared to the Y+2 tide. We measured the activity of Plk1 wildtype and T210D against the optimal peptide based on the Plk1 motif, Plk2 wildtype and T239D against the pY+2 tide and Y+2 tide, and Plk4 wildtype and T170D against a peptide based on the Plk4 motif (Figure 3.1). The phosphomimic mutation increased the activity of Plk1 approximately 10-fold, agreeing with the previously published data (Figure 3.3A) (24). The activity of both wildtype and T239D Plk2 against the pY+2 tide was greater than against the Y+2 tide. The activity of Plk2 T239D against the pY+2 tide was approximately 7-fold higher than the wildtype against the pY+2 tide and activity of Plk2 T239D against the Y+2 tide was more than 40-fold greater than the wildtype against the same peptide (Figure 3.3B). Thus unlike Plk3, the phosphomimic mutation in Plk2 significantly increases its activity and this increase in activity was not specific to the sequence tested. The activity of Plk4 wildtype was more than 200-fold higher than Plk4 T170D (Figure 3.3C).

**Kinase motifs of activation loop phosphomimic mutants of Polo-like Kinase family members**

That activation loop phosphorylation or phosphomimic mutation does not change the overall activity of Plk3, but does reduce the activity of this kinases against pY in the +2 containing sequences, suggests that these modifications to the kinase may simply be altering the specificity of this kinase from that of the wildtype, unphosphorylated form. To examine this possibility, we determined the kinase motif of Plk3 T219D as well as the motifs of Plk1 T210D, Plk2 T239D, and Plk4 T170D as controls to determine if this is a general property of Polo-like kinase family members or specific to Plk3 (Figure 3.4). The optimal consensus motif of Plk1 T210D and Plk4 T170D are D/N/E-X-S/T-Φ and D/N/E-X-S/T-Φ-Y/W, respectively (Figure 3.4). These motifs are the same as the motifs of the wildtype, unphosphorylated kinases (Figure 3.4). The optimal consensus motif of Plk2 T239D is E-D/E-pY-S/T-Φ-pY/Y/W. The motif of Plk3 T219D is E-D/E-pY/D/E-S/T-Φ. This motif is distinctly different from the motif of the wildtype,
Figure 3.4: Comparison of the motifs of the wildtype Plk family members with the motifs of the activation loop phosphomimic mutants. In the red-green overlay in the third column, the motifs of the wildtype kinases are green and the motifs of the phosphomimic mutants are red. Equal contribution from both motifs results in a yellow spot.

unphosphorylated Plk3, S/T-X-pY (Figure 3.4). Clearly, the phosphomimic mutation changes the motif. This is a striking result. Phosphorylation on the activation loop of Plk3 acts as a toggle to switch the kinase from a pY-directed kinase to an acidophilic kinase with a motif of E-D/E-pY/D/E-S/T-Φ. To our knowledge, this is the first report of a kinase that has a motif that switches.

Comparison of the motifs of the most active form of each of the Polo-like kinase family members

Although most closely related to the AGC group of kinases, the Polo-like kinase family does not belong to any of the seven major groups of kinases (TK, TKL, STE, CK,
Figure 3.5: Pseudo-color PS-OPLS blots of the more active form (wildtype or activation loop phosphomimic mutant) of Plk1 (purple), Plk2 (red), Plk3 (yellow), and Plk4 (cyan) were superimposed to show similarities and differences between the motifs. Equal selectivity in the Plk1-Plk2 overlay is dark blue, Plk1-Plk3 overlay is blue, Plk1-Plk4 overlay is yellow, Plk2-Plk3 overlay is green, and Plk3-Plk4 overlay is magenta.

AGC, CAMK, and CMGC). The Polo-like kinase family may be thought of as a group in and of itself and, therefore, studying the motifs of this family may give us insight into the diversity of motifs within any of the major groups of kinases without having to determine the motifs of tens of kinases. We compared the motifs of Plk1 T210D, Plk2 T239D, wildtype, unphosphorylated Plk3, and wildtype, unphosphorylated Plk4 as these motifs are of the more active form. To more easily visualize the similarities and differences between the motifs, Figure 3.5. shows pairs of motifs in pseudocolor superimposed. The motifs of Plk2 T239D and Plk3 are very similar. Selection for pY in the +2 is clear, but other weak selection in Plk3 is evident in the Plk2 motif. This similarity is not surprising as their sequence homology is higher than that of any other pair of Polo-like kinase family members. The E in the -3 position distinguishes the motifs of these two kinases from that of Plk1 T210D or Plk4. The D/N/E in the -2 is selected by all members of the family, with weaker selection by Plk2 T239D, Plk3, and Plk4, but strong selection by Plk1 T210D. In the +1 position, all members of the family, except Plk3, show strong
selection for F as well as a general selection for hydrophobic amino acids, which are weak selections for Plk3. We have previously shown that Plk1 will not phosphorylate sequences with P in the +1 and the motif of Plk2 T239D shows similar discrimination against P in the +1 as that of Plk1 (Figure 2.2). This is in contrast to the motif of Plk4 which tolerates P in the +1. Selection for Y and W in the +2 is strongest in the Plk4 motif and to a lesser extent that of Plk1 T210D, but is also present in the motif of Plk2 T239D. Overall, the members of the family have kinase motifs that are similar, but the motifs of Plk1 T210D, Plk2 T239D, Plk3, and Plk4 all contain aspects that make them distinct from each other. This suggests that these kinases may phosphorylate some of the same sites, but will also phosphorylate different sites on proteins.

The Polo-Box Domain of Plk3, unlike that of Plk1, is not a phospho-dependent binding domain

To further explore the biochemistry of Plk3, we re-examined the ability of the Plk3 PBD to function as a phospho-dependent binding domain. It was previously published that the Plk1 PBD is a phospho-dependent binding domain that binds sequences matching the motif, S-pS/pT-X (4). Based on the high sequence homology including the critical residues, H538 and K540 in Plk1 and H590 and K592 in Plk3, that interact with the phosphate it was suggested that Plk3 PBD was also a phospho-dependent binding domain (5). Three constructs were tested for their ability to bind to pS, pT, or pY oriented peptide libraries as compared to non-phosphorylated control peptide libraries. The three constructs were full-length Plk3 (aa. 1-646), a construct including a C-terminal portion of the kinase domain through the PBD (aa. 312-646), and a construct starting just C-terminal to the kinase domain and including the PBD (aa. 335-646) (Figure 3.6). The construct composed of amino acids 312-646 has been published to localize to the centrosome and the construct including amino acids 335-646 is the equivalent of the Plk1 326-603 sequence that was originally discovered to be a phospho-dependent binding domain (4, 29). All three constructs showed stronger binding to the non-phosphorylated control peptide libraries than the phospho-peptide libraries, whereas the Plk1 PBD bound more strongly to the pS and pT oriented peptide libraries than the non-phosphorylated controls and the Src SH2 domain bound more strongly to the pY
Figure 3.6: *In vitro* transcription/translation pulldown assays of Plk3 Polo-box domain (PBD) constructs with phosphorylated or unphosphorylated peptide libraries shows that the Plk3 PBD binds more strongly to non-phosphopeptides than to phosphopeptides suggesting it is not a phospho-dependent binding domain. Three constructs of the Plk3 PBD were tested with two controls, the Plk1 PBD and the Src SH2 domain.

Oriented library than the non-phosphorylated control (Figure 3.6). These data suggest that the Plk3 PBD is not a phospho-dependent binding domain, but may actually bind non-phosphorylated sequences. Unfortunately, it is not possible to determine from this technique what sequences are preferred by this domain.
Discussion

In this study, we determined the motifs of the Polo-like kinase family members, Plk1, Plk2, Plk3, and Plk4, as both the wildtype and with the T210D, T239D, T219D, and T170D phosphomimic mutations, respectively, on the activation loops. The phosphomimic mutation motifs of Plk1 and Plk4 were essentially the same as the wildtype motifs of these kinases. The Plk1 motif is D/N/E-X-S/T-Φ and the Plk4 motif is D/N/E-X-S/T-ϕ-Y/W, where Φ is any hydrophobic amino acid except proline and ϕ is any hydrophobic amino acid. Unexpectedly, the wildtype motifs of Plk2 and Plk3 were different from the activation loop phosphomimic mutation motifs of these kinases. The wildtype motif of E-X-X-S/T-F-pY and of Plk3 is S/T-X-pY. The Plk2 T239D motif is E-D/E-pY-S/T-Φ-pY/Y/W and the Plk3 T219D motif is E-D/E-pY/D/E-S/T-Φ. The motifs of Plk3 make it clear that Plk3 is an acidophilic kinase making it unlikely that if phosphorylates the S216 site in Cdc25C.

Phosphomimic mutation on the activation loop seems to have different effects on different members of the Polo-like kinase family, suggesting that phosphorylation on the activation loops would have similarly diverse effects. We reconfirmed that phosphomimic mutation on the activation loop of Plk1 stimulates kinase activity. Phosphomimic mutation on the activation loop of Plk2 and Plk3 seems to switch the motif. The motif switch for Plk2 is less dramatic than that of Plk3 as the strongest determinants of specificity in both the wildtype, unphosphorylated and Plk2 T239D motifs are the same. Additionally, the phosphomimic mutation increased the activity of the kinase by more than 40-fold compared to the wildtype, suggesting that the kinase may need to be phosphorylated in vivo for activity. In Plk3, the phosphomimic mutation does not seem to affect catalytic competency of the kinase, but seems only to switch the motif. The motif switch is quite dramatic, changing the motif from being dominated by selectivity for pY in the +2 to having specificity for acidic residues in the positions N-terminal to the phosphoacceptor and almost reducing the pY in the +2 specificity to background levels. The motif switching is not an artifact of the phosphomimic mutation since phosphorylation in the activation loop also switches the motif. This is an extremely important result as it shows that kinases do not necessarily have static motifs, as was
previously assumed in the field. The pY in the +2 specificity of the wildtype kinase is at least in part mediated by K106, which may be part of a basic pocket involved in both phospho-dependent kinase activation as well as substrate selection in the absence of activation loop phosphorylation. Since the structure of Plk3 has not been solved, it is impossible to know for sure. The phosphomimic mutation of Plk4 resulted in a greater than 200-fold drop in activity of the kinase. Although, based on sequence alignment, T170 of Plk4 seems to be the equivalent residue to Plk1 T210, which has been shown to be phosphorylated for activation of Plk1, this can only be verified with crystal structures. It may be that Y169 in Plk4 is actually the residue that is phosphorylated for activation of Plk4. Nevertheless, the results lead to the hypothesis that phosphorylation on the activation loop may function as a way of shutting off Plk4 kinase activity. Although clearly possible, whether or not phosphorylation is actually employed in vivo to swap Plk3 motifs or shut off kinase activity of Plk4 remains to be tested.

Intriguingly, the Plk2 and Plk3 motifs show specific selection for pY in the +2 position. In other kinase motifs when pY is selected, either other negatively charged residues or large hydrophobic residues are also selected in the same position, as in the case of CKII and PAK kinases, suggesting that what is important for these kinases is charge or hydrophobicity, not the presence of phosphotyrosine itself (22, 30). Two kinases are known to require priming phosphorylations within a few amino acids of the site to be phosphorylated, GSK3β, as mentioned previously, and CKI, which has the motif pS/pT/D/E-X-X-S/T (31). In both cases, the priming phosphorylation is on a pS or pT residue. No kinase has previously been described that requires priming phosphorylation on a pY residue. This is particularly interesting in a serine/threonine kinase as it suggests that Plk2 and Plk3 may integrate tyrosine kinase signaling with serine/threonine kinase signaling in a novel way.

One other serine/threonine kinase integrates tyrosine kinase signaling and serine/threonine kinase signaling in a similar, but distinct manner. Protein Kinase C delta (PKCδ) contains a Conserved Domain 2 (C2) domain that has been found to be a phosphotyrosine binding domain (32). Thus, the C2 domain can recognize a phosphotyrosine motif on a substrate and the PKC kinase domain could then phosphorylate the substrate at another site. This is similar to the PBD of Plk1, which
binds a phosphoserine/phosphothreonine motif on a substrate and allows Plk1 to phosphorylate that protein or other proteins in the vicinity (33). We re-examined the ability of the Plk3 PBD to function as a phospho-dependent binding domain and found that it binds more strongly to non-phosphorylated sequences suggesting that it is not a phospho-dependent binding domain. Both Plk1 and Plk3 seem to function as molecular logic gates, integrating the activity of another kinase with its own and outputting a phosphorylation on the substrate. Plk1 functions as an AND gate. Its two inputs are phosphorylation status of the activation loop of the kinase domain, which decides if the kinase is active or not, and phosphorylation status of the substrate or a protein nearby, which is recognized by the PBD. Both inputs should be in the phosphorylated state for output phosphorylation of the substrate by Plk1. PKCδ functions similarly. Plk3 functions as an Exclusive OR (XOR) gate. Its two inputs are phosphorylation status of the activation loop, which decides which motif the kinase will phosphorylate and the phosphorylation status of the substrate, which is recognized by the active site cleft. For Plk3 exactly one input should be in the phosphorylated state for output phosphorylation on the substrate. The phosphopeptide binding function of the PBD, which is conserved in orthologs of Plk1 from yeast to humans, seems in Plk3 to have been subsumed by the Plk3 kinase domain, leaving the Plk3 PBD able to handle other as yet undefined functions (5). Since the kinase domain is responsible for recognition of the priming phosphorylation, phosphorylation by Plk3 occurs on more carefully chosen sites than in the case of Plk1 because the priming phosphorylation must be in the same protein exactly two amino acids away from the residue to be phosphorylated.

Kinase motifs become more important and useful as the number of phosphoproteomic mass spectrometry screens performed increases. The motifs combined with phosphoproteomic data and bioinformatics approaches can help to identify the kinases creating the phosphorylation sites and thereby predict sites phosphorylated by these kinases (34). The substrates of Plk2, Plk3, and Plk4 are almost unknown and their roles are only beginning to be understood. In this study we have determined the motifs of these kinases, which should aid in the identification of substrates and help to increase our standing of the roles of these kinases in regulation of both the cell cycle and other post-mitotic functions.
References


7. Ma S, Liu MA, Yuan YL, Erikson RL. The serum-inducible protein kinase snk is a G1 phase polo-like kinase that is inhibited by the calcium- and integrin-binding protein CIB. Mol Cancer Res. 2003 Mar;1(5):376-84.


Chapter Four

Conclusions and Future Directions

Contributions:
I wrote all sections.
Analysis of Motifs of Kinases Related by Function Versus those Related by Sequence

The Polo-like kinase family has high sequence similarity, but each member of the family seems to have a different role in cell cycle regulation. The motifs of the Polo-like kinase family were found to be similar, but distinct. In particular, Plk3 T219D and the more active forms of Plk1, Plk2, and Plk4 all showed selection for D and N in the -2 position and a hydrophobic amino acid in the +1 position. This similarity suggests that given access to the same proteins, these kinases might phosphorylate the same sites on these proteins. In addition to these common specificities, each kinase selected for other amino acids in other positions that differentiated their motifs, like selection for E in the -3 by Plk2 and the strong selection for W and Y in the +2 by Plk4, suggesting that each member of the family will also phosphorylate distinct sites on substrates. As noted in Chapter 1, these kinases have common localizations and could have common substrates, particularly at the centrosome, where each has been reported to localize.

The expression patterns of the kinases likely contribute to their distinct roles, which are likely mediated by phosphorylation of distinct substrates. Plk4 protein is expressed from late G1 through the cell cycle to early in the next G1, when it becomes undetectable (1). Plk2 is expressed in G1 with protein and activity levels peaking in mid-G1 and then falling off before S-phase (2). Plk1 protein is expressed from late S-phase until it is destroyed in late M-phase and early G1 (3). Plk2 may be at the centrioles during the period in which Plk4 protein is undetectable preparing centrioles for duplication prior to Plk4’s role in centriole duplication (4). Plk2 protein levels probably fall off at approximately the same time as Plk4 levels are rising again. Plk1 and Plk4 may be at the centrosome at the same time in mitosis, but Plk1 seems to be associated with the pericentriolar material whereas Plk4 is associated specifically with the centrioles (5), suggesting that they may have non-overlapping localizations at the centrosome possibly preventing phosphorylation of the same substrates. How Plk3 fits into this picture is unclear, since when Plk3 protein is expressed during the cell cycle or where it localizes is controversial. However, the motif of wildtype, unphosphorylated Plk3 is distinct from the motifs of the other kinases and so it likely has distinct substrates. It
seems likely that timing of expression and localization play an important role in substrate specificity of the Plk family.

The substrate specificity of the Protein Kinase C (PKC) family, which has high sequence similarity and highly similar motifs, is regulated similarly to the above proposed regulation of the Plk family. However, the PKC family is different from the Plk family, since PKC kinases have some redundant roles (6). All nine members of the family select for an R in the -3 and hydrophobic amino acids in the +1 position (7). Each has slightly different specificities at other positions of the motif. Regulation of PKC substrate specificity seems to involve differential expression as well as localization, like the Plk family (8). The PKCs are different from the Plks in that instead of temporal regulation of expression, PKCs are expressed in different cell types (9). Like the Plk family, the PKCs have different localizations. Localization is mediated by receptors for activated C-kinase (RACKs) (8). RACKs likely limit the substrates with which each kinase can interact in addition to allosterically activating the bound kinase. The similarities in regulation of substrate specificity of the PKCs to the proposed regulation of the Plk family lend some credence to the proposed model of regulation.

Unlike the Polo-like kinases, which function in different aspects of the cell cycle and are expressed during different phases of the cell cycle, the major mitotic kinases must all be expressed during mitosis. The major mitotic kinases, which, except for Aurora A and Aurora B, do not have very high sequence similarity, but do cooperate in function, have very different motifs. Cdk1 will not phosphorylate any sites phosphorylated by any of the other major mitotic kinases and vice versa despite overlapping localizations; Plk1 will not phosphorylate any sites phosphorylated by the Aurora kinases and vice versa despite overlapping localizations. However, the motifs of Aurora A and Aurora B are almost identical. We have proposed that the localization of these kinases is a key factor in generating their specificity as Aurora A localizes to the centrosome and spindle microtubules whereas Aurora B localizes to the DNA and kinetochores (10, 11). It might be that for kinases that function together, like the major mitotic kinases, localization and motifs play larger roles in generating specificity than for kinases which do not, like the PKC or Plk family, which can rely on differential expression to generate some of the specificity.
Summary and Future Directions

Although motifs have been known for some time to direct individual kinases to their substrates, the motifs presented in this study suggest that motifs can play larger roles. In Chapter 2, we put forward a hypothesis that, if true, would imply that the motifs and localizations of individual kinases cooperatively direct a group of kinases to their substrates in a coordinated manner. Although motifs function at the level of individual kinases they may also function at the level of groups of functionally related kinases. In Chapter 3, we showed that the motifs of Plk2 and Plk3 select for phosphotyrosine, suggesting that phosphorylation by a tyrosine kinase on the substrate primes for phosphorylation of the substrate by Plk2 and Plk3. Motifs are not just involved in passing along signals, but can actually integrate signals. In the case of Plk2 and Plk3, motifs may integrate tyrosine kinase signaling with serine/threonine kinase signaling in a novel way. Also in Chapter 3, we showed that Plk3 has a phosphorylation based toggle switch that changes the motif of the kinase. This result suggests that the motif of a kinase is not necessarily static and for some kinases it might be modulated to cause the kinase to phosphorylate different sites under different conditions. Thus, the motif of certain kinases may actually change the functional roles of those kinases.

Coordinated Regulation of the Major Mitotic Kinases by Motifs and Localizations

In Chapter 2, based on the results of biochemical characterization of the motifs of the major mitotic kinases we put forward a hypothesis potentially explaining how the major mitotic kinases are regulated to phosphorylate distinct sites in proteins despite having access to the same sites: In mitosis, major mitotic kinases with overlapping motifs do not have overlapping localizations and major mitotic kinases with overlapping localizations do not have overlapping motifs. If this hypothesis accurately describes the regulation theses kinase are under, then mislocalization of a kinase A so that it co-localizes with a kinase B with an overlapping motif should allow the sites phosphorylated by the kinase B to be phosphorylated by kinase A. Additionally, changing the motif of a kinase C that is co-localized with a kinase D so that both kinases have the same motif should allow the kinase C to phosphorylate sites that had originally only been
phosphorylated by the kinase D. Testing either of these predictions of the model directly is difficult because all of the major mitotic kinases are pleiotropic and modified function by either mislocalization or changing of the motif would result in overlaid, difficult to interpret phenotypes. In particular, the localizations of substrates and kinases would be altered. A better way to test the hypothesis would be to manipulate a substrate rather than a kinase itself. Either mislocalization of a substrate from the localization of one kinase to that of another kinase with an overlapping motif or mutation of a site known to be phosphorylated by one kinase to match the motif of second kinase with the same localization, but a different motif would make interpretation of the results of the experiment easier, since an appropriately chosen substrate will not be pleiotropic.

One such convenient substrate is the budding yeast, Scc1. Scc1 is a component of the cohesin complex which holds sister chromatids together during mitosis to prevent premature chromosome segregation (12). When all chromosome have obtained amphitelic attachments and the spindle checkpoint has been silenced, the APC/C ubiquitinates securin resulting in the destruction of securin (13). Securin inhibits the protease separase from degrading Scc1 (14). In budding yeast, degradation of Scc1 is essential for anaphase onset (15). Separase cleaves Scc1 at two sites, of which the C-terminal site is preferred by the endopeptidase. A few amino acids N-terminal to each of the cleavage sites is a site matching the Plk1 consensus motif (173-DTSL-176 and 261-DNSV-264). Alexandru et al. showed that the yeast ortholog of Plk1, Cdc5, phosphorylates Scc1 at these sites to make it a better substrate for separase in a securin deletion strain (16). Mutation of the serines to non-phosphorylatable residues resulted in delay of Scc1 degradation by about 30 minutes and increased the number of cells with undivided nuclei undergoing cytokinesis, known as the “cut” (chromosomes untimely torn) phenotype. Phosphorylation of Scc1 was assayed by electrophoretic mobility shift.

Mutation of the sites phosphorylated by Cdc5 to match the motif of Ipl1 (D173R and D261R), the budding yeast ortholog of Aurora B, or Cdk1 (L176P and V264P) would allow the hypothesis to be tested as both Ipl1 and Cdk1 have appropriate localizations (Angelika Amon, Personal Communication). The yeast system would allow this to be done with relative ease, and wildtype and non-phosphorylatable Scc1 could be used as controls. The same assays used by Alexandru et al. could be used. Phosphorylation and
degradation of Sccl can be monitored over a time-course by electrophoretic mobility shift and the appearance of cleavage products, respectively. The number of yeast displaying the “cut” phenotype would also be monitored. If the hypothesis is true, the motif mutant Sccl should be phosphorylated and degraded with kinetics more similar to the wildtype Sccl than the non-phosphorylatable mutant. Also there should be fewer cells displaying the “cut” phenotype for the motif mutant than for the non-phosphorylatable mutant. The proposed experiments are currently in progress.

It is also possible that the motif mutant Sccl will result in a different phenotype. Both Ipl1 and Cdk1 are under different regulation than Cdc5. The motif mutant Sccl might be phosphorylated earlier or later by Ipl1 or Cdk1 than it would be by Cdc5, resulting in premature or delayed separase cleavage of the motif mutant Sccl. While a delayed separase cleavage may just result in a delayed, but otherwise normal, anaphase, a premature cleavage prior to all kinetochores attaining amphitelic attachments would cause a lack of tension, an extended spindle checkpoint, and, probably, premature cytokinesis resulting in aneuploid cells (12).

A phenotype for the Sccl motif mutant different than the wildtype would have important implications for the regulation of mitosis. It would suggest that it matters which kinase phosphorylates a given site. In essence, each site must not only be phosphorylated, but must be phosphorylated in the appropriate setting, which would be arranged by the differential upstream regulation of the appropriate major mitotic kinase. Such a result would explain why the regulation suggested by the hypothesis would be required and why more than one kinase is necessary for proper regulation of mitosis. As mitosis is an intricate process, this differential upstream regulation must be important in fine tuning when events occur in mitosis. Although, temporal regulation is not explicitly stated in the hypothesis, it would implicitly be part of the hypothesis. Whether or not there is a different phenotype for the Sccl motif mutant as compared to the wildtype, if the hypothesis is true, there is a coordinated regulation of the major mitotic kinases based on the individual kinases’ motifs and localizations.
Regulation of p31/Comet Function by Plk1

p31/Comet was shown to be a candidate substrate of Plk1 in nocodazole arrested cells. Previously, p31/Comet was shown to be involved in silencing the spindle checkpoint by binding Mad2 (17). Mad2 has at least two conformations, open-Mad2 and closed-Mad2 (18). In the "Mad2 Template" Model, closed-Mad2 binds Mad1 at the kinetochores for its localization to the kinetochore (19). Open-Mad2 binds closed-Mad2 on kinetochores without amphitelic attachments where it can also bind Cdc20. Upon binding of Cdc20, open-Mad2 changes conformation to closed-Mad2 and the Mad2-Cdc20 complex dissociates from the kinetochore. The interaction of Mad2 with Cdc20 prevents interaction of Cdc20 with the APC/C preventing anaphase onset. After a kinetochore has obtained an amphitelic attachment, p31/Comet binds the closed-Mad2 on the kinetochore preventing open-Mad2 from binding closed-Mad2 and preventing the formation of Mad2-Cdc20 complexes (20).

How p31/Comet is prevented from binding Mad2 on kinetochores still lacking amphitelic attachments has not been determined, although we hypothesize that phosphorylation by Plk1 may play a role. The role of Plk1 in the spindle assembly checkpoint is unclear. Plk1 is required for recruitment of proteins involved in the spindle assembly checkpoint such as Mad1 and Mad2, and it also has been reported to generate the 3F3/2 phosphoepitope which appears on kinetochores lacking amphitelic attachments to the spindle (21-23). The X-ray crystal structure of the Mad2-p31/Comet complex has been solved (24). However, the candidate Plk1 phosphorylation site on P31/Comet is not near the interface suggesting it does not directly inhibit interaction of the Mad2-p31/Comet complex. However, the phosphorylation may still be a signal resulting in the prevention of p31/Comet from binding Mad2 prior to amphitelic attachment. To test this hypothesis, a non-phosphorylatable mutant p31/Comet could be expressed in p31/Comet knockdown cells. Appropriate controls would be expression of wildtype p31/Comet and a phosphomimic mutant in the p31/Comet knockdown cells. If the hypothesis is true, the spindle checkpoint would be silenced early resulting in premature chromosome segregation and increased aneuploidy as compared to the wildtype control. The phosphomimic might show delayed anaphase onset.
Promotion and Inhibition of the Cell Cycle by Plk3

As discussed previously, Plk3 has been reported to have roles both in promoting the cell cycle and promoting cell cycle arrest (25-31). There are at least two possible explanations for how Plk3 could mediate this conflicting behavior, one involving the PBD and the other involving the kinase domain.

In Chapter 3, the PBD of Plk3 was shown to not be a phospho-dependent binding domain, and this has been corroborated by an independent study published recently (32). However, the PBD has previously been shown to mediate protein-protein interactions with CIB and SPAR as part of its non-cell cycle related roles in neurons (33, 34). The interaction with CIB seems to occur in a phospho-independent manner as only one of the Polo-box motifs of the pair of Polo-box motifs comprising the full Polo-box domain is required. GFP-tagged Plk3 PBD has been reported to localize to centrosomes also suggesting a binding function (35). When cell lysates from untreated cells or cells subjected to cellular stress were fractionated on a size-exclusion column, Plk3 was reported to associate with complexes of different sizes (27). The later data is somewhat suspect as it relies on an antibody, which we have found to immunoprecipitate the heat shock protein, Hsp70, rather than Plk3 based on mass spectrometry (data not shown) and recognizes a band by Western that is not fainter in lysates from Plk3 knockdown cells as compared to untreated cells (36). Nevertheless, the data suggest the hypothesis that the Plk3 PBD may function as context dependent binding domain forming complexes with different proteins after cellular stress than during a normal cell cycle. The switch deciding which proteins it interacts with may be on the PBD itself or may be a property of the interacting proteins.

More excitingly, the kinase domain may contain the switch which decides whether Plk3 contributes to cell cycle progression or cell cycle arrest. In Chapter 3, phosphorylation or phosphomimic mutation of Plk3 at T219, which is on the activation loop, was found to act as a toggle switch to swap between a motif almost entirely consisting of specificity for pY in the +2 position and the generally acidophilic motif, E-D/E-pY/D/E-S/T-Φ, where Φ, is any hydrophobic amino acid. Although this result suggests that motif swapping is possible for kinases, which is an important result in itself, it does not address whether or not this toggle is employed in vivo on Plk3 and to what
end. Phosphorylation by Plk3 on p53 and Cdc25A has been reported to occur on acidophilic sites (ETFS^{20}D in p53 and EST^{80}D in Cdc25A) to cause cell cycle arrest and Plk3 has been reported to be phosphorylated by Chk2 following cellular stress, although the site was not identified (26, 27, 37). The activation loop T219 is part of a sequence (EQRKK{T}^{219}) that matches a minimal Chk2 consensus sequence of R-X-X-S/T (38). Together these data suggest the hypothesis that Plk3 may be phosphorylated on T219 of the activation loop to swap its motif from being dominated by pY specificity for the roles of Plk3 in promoting the cell cycle to being acidophilic for the roles of Plk3 in inhibiting the cell cycle.

Although there is an antibody that recognizes the phosphorylated activation loop of Plk1, no such antibody exists for Plk3 (39). The sequences of Plk1 and Plk3 on the activation loop are sufficiently similar, however, that the antibody might recognize Plk3 phosphorylated on the activation loop. To test the antibody, recombinant phosphorylated and unphosphorylated Plk3 could be used in a Western blot. If the antibody does recognize T219 phosphorylated Plk3, an easy experiment presents itself that could test this hypothesis using cell lines. Plk3 could be immunoprecipitated from lysates from cells exposed to ionizing radiation to cause DNA damage and the phosphospecific antibody used in a Western blot to determine if Plk3 is phosphorylated. An equal amount of lysate from undamaged cells would be used as a control. If the antibody does not recognize T219 phosphorylated Plk3, a mass spectrometric approach would be required.

One possible experiment to clarify the relevance of motif switching might be to make Plk3 knock-down cells, or analogue-specific inhibited forms of Plk3, and ask whether T219A or T219D mutants of Plk3 were equivalent to wild-type Plk3 in their ability to reverse a knock-down or chemically inhibited phenotype. In our preliminary experiments, RNAi-mediated knock-down of Plk3 appeared to induce cell lethality, although the absence of a good antibody against Plk3 limits our ability to make strict conclusions.

Modulation of Src Homology 2 (SH2) Domain Binding by Plk3

Tyrosine kinase signaling is involved in, among other processes, cellular substratum adhesion (40). We and others have shown that GFP-tagged Plk3 localizes to
the cellular cortex where tyrosine kinase signaling regulating substratum adhesion is concentrated (data not shown), (40, 41). Additionaly, overexpression of Plk3 results in cellular rounding and release from the substratum, suggesting Plk3 may play a role in regulating substratum adhesion (data not shown), (41). The relevant substrates are unknown. SH2 domains, which are found in many proteins involved in tyrosine kinase signaling, bind tyrosine phosphorylated sequences in proteins (42, 43). The motifs of many SH2 domains have been determined and were found to select tyrosine phosphorylated sequences based on the amino acids C-terminal to the phosphotyrosine residue (44, 45). However, recently a more extensive study of SH2 domain binding motifs has revealed that many SH2 domains also have specificity for amino acids one and two positions N-terminal to the phosphotyrosine residue (46). Specificity for phosphoserine or phosphothreonine in any position was not examined in any of these studies (44-46). The motif of wildtype, unphosphorylated Plk3 is S/T-X-pY, suggesting that Plk3 selects sequences for phosphorylation that are similar to sequences bound by SH2 domains. This similarity in specificity suggests the hypothesis that Plk3 could phosphorylate phosphotyrosine containing sequences to either increase or decrease their affinity for SH2 domains. Phosphorylation by Plk3 would place two negative charges on the amino acid two positions N-terminal to the phosphotyrosine, which could affect SH2 binding as some domains specifically select negatively charged amino acids in this position while some others select serine or threonine (46). Such modulation of SH2 domain binding and consequently tyrosine kinase signaling may mediate Plk3’s role in substratum adhesion.

**Shutting off Plk4 Activity**

We showed that the wildtype Plk4 kinase domain was approximately 200-fold more active the activation loop phosphomimic mutant, Plk4 T170D. This suggests that phosphorylation on the activation loop of Plk4 might act as a toggle to turn on or off kinase activity when it is no longer needed either locally or globally. Plk4 is involved in centriole biogenesis and overexpression results in multiple centrioles (5). Since Plk4 is expressed from G1 to the next G1, but centriole duplication takes place in S-phase, it may
be that Plk4 kinase activity at the centrioles needs to be inhibited, but activity elsewhere is still required since Plk4 also seems to be required for mitosis (1, 47).

No kinase has been reported for which phosphorylation on the activation loop has resulted in inactivation. So, this would be a novel mode of regulation of kinase activity. To determine if Plk4 is phosphorylated on the activation loop will require a mass spectrometric approach as an antibody is not currently available. One way to test if phosphorylation on the activation loop plays a role in regulating Plk4 activity with respect to centriole duplication would be to express a non-phosphorylatable Plk4 mutant at near endogenous levels in HeLa cells in the setting of a Plk4 knockdown and an S-phase arrest by aphidicolin and look for centriole overduplication by immunofluorescence using C-Napl as a marker (5). Expression of wildtype Plk4 and catalytically inactive Plk4 in the same setting would be appropriate controls. If inactivation of Plk4 by phosphorylation on the activation loop is required, then centrioles would overduplicate in cells expressing the non-phosphorylatable mutant, but not in cells expressing the wildtype or catalytically inactive forms.

Enhancing the PS-OPLS Technique

Although phosphorylation of proteins is the most widespread post-translational modification employed by the cell for signal transduction, ubiquitination, sumoylation, acetylation, and methylation, among others, are also known to play roles in signal transduction (8, 48). Currently, it is not known if post-translational modifications other than phosphorylation are part of the motif of any kinases. Ubiquitin and SUMO are too large to fit in the active site cleft of a kinase, but acetylations and methylations are small modifications that could play a role in kinase specificity. On proteins such as histones and p53, phosphorylation, acetylation, and methylation occur on residues in close enough proximity to each other that one post-translationally modified residue could act as a specificity determinant for the enzymes creating subsequent post-translational modifications (49, 50). Although, there are no clear examples of acetylated or methylated residues acting as kinase specificity determinants, phosphorylation of p53 at S371 inhibits methylation on the adjacent K372 suggesting that the methyltransferase affinity for the sequence is reduced by the phosphoresidue (50). If phosphoresidues can
affect methyltransferase affinity, it does not seem unreasonable that methylated or acetylated residues could be determinants of kinase specificity. Interestingly, Aurora B phosphorylates Histone H3 next to trimethylated K9 at S10 during mitosis (51, 52). Although the unmethylated site matches the motif of Aurora B, the effect of methyl-lysine on the specificity of the kinase for the site is unknown. It may be that Aurora B has a higher affinity for a site with methyl-lysine in the -1 position. The peptide libraries used in the current PS-OPLS technique allow for the testing of specificity towards phosphothreonine and phosphotyrosine in addition to the twenty natural amino acids, but the libraries could be expanded to include other post-translational modifications such as methyl-lysine, dimethyl-lysine, trimethyl-lysine, methyl-arginine, and acetyl-lysine to allow us to determine if these post-translational modifications affect the specificity of kinases.

Additionally, this technique could also be modified to enable the determination of motifs of other enzymes that post-translationally modify proteins. A peptide spot array has been used to determine the specificity of one methyltransferase (53). In this approach, peptides containing single amino acids changes from the sequence of a known site were synthesized on a membrane in separate spots. Each amino acid in each position was represented in a different spot. The PS-OPLS technique has the advantages that the reaction occurs in solution, which reduces artifacts associated with the solid support, and that each amino acid is tested in each position independently of all others resulting in an unbiased motif. Methyltransferases and acetyltransferases might be amenable to this approach. As acetyltransferases and methyltransferase modify not only histones, but transcription factors, and other proteins, the motifs of these enzymes might help to determine substrates, which would help to define the roles of these enzymes (50, 54). The peptide libraries used for the determination of the motifs of these enzymes would have to be oriented on lysines. S-adenosyl-methionine with a radioactively labeled methyl group would be need for the methyltransferases; Acetyl CoA with a radioactively labeled acetyl group would be needed for the acetyltransferases. The rest of the experimental set up, however, could be the same.
References


