STRUCTURE OF KI67 FHA DOMAIN 
AND ITS BINDING TO HNIFK

DISSERTATION

Presented in Partial fulfillment of the requirements for 
the Degree Doctor of Philosophy in 
the Graduate School of the Ohio State University

By

Hongyuan Li, M S

*****

The Ohio State University

2003

Dissertation Committee:  
Approved by

Prof. Ming-Daw Tsai, advisor

Prof. Donald Dean

Prof. Mark Foster

Prof. Venkat Gopalan

Biochemistry Department
ABSTRACT

The Ki67 protein is a 359 kDa nuclear protein whose existence is strictly associated with the cell cycle. Consequently, Ki67 has been used extensively as a marker for proliferating cells. The exact role of Ki67 has been elusive despite its great biological significance. The FHA domain has been reported to be a phosphopeptide recognition domain. In many cases, it has been shown to be a signal transduction module. Two proteins, Hklp2 and hNIFK, were reported to bind the FHA domain of Ki67 in a cell cycle- and phosphorylation-dependent manner. The FHA domain in Ki67 serves as a starting point to understand how this protein is involved in the signal transduction network of cell cycle control. In this study, the gene encoding the N-terminus FHA domain of Ki67 was cloned into an expression vector to express in *Escherichia coli* BL21(DE3) CodonPlus strain and the recombinant protein was purified. The binding specificity of this FHA domain was studied extensively using library screening and NMR titration methods. The structure of this domain was solved in collaboration with Dr. In-Ja Byeon using NMR. The phosphorylation site and binding property of its biological target protein hNIFK were studied.

Unlike the FHA1 domain in Rad53, which can bind small phosphopeptides whose consensus sequence can be easily determined by library screening methods, the Ki67
FHA domain only shows binding interaction with phosphopeptides whose length is much greater than the local region near the phosphorylation site. This suggests that higher order structure within the binding target is required for productive binding by Ki67 FHA domain. While non-phosphorylated hNIFK(226-269) fragment can bind to Ki67 FHA domain weakly with a $K_d \sim 10^6 \mu M$, site-specific phosphorylation at T234 increases the affinity by more than 60-fold. HSQC titration analyses showed that this phosphorylation-dependent binding involves an interacting surface composed of a phospho-dependent binding center and additional phospho-independent contacts. The coordination of these two regions may be key to understanding the specific, tight binding of Ki67 FHA domain to hNIFK.

The results presented here provide the first example of an in vitro study of the binding of FHA domain to its biological target. This work shows that while the binding pattern of FHA domain around phospho-interacting center can be similar to those demonstrated by complex structures using small synthetic phosphopeptides from library screening, additional selectivity and/or functionality can be provided through additional contacts. This discovery provides further insights into the binding pattern of FHA domains.
ACKNOWLEDGMENTS

None of the work described here would have been possible without the support from many people. At the time I finish this dissertation, I would like to express my sincere thanks to each of them.

I wish to thank my advisor, Ming-Daw Tsai, for his guidance and continuous support in this study. His enthusiasm and rigorous attitude toward science encourage me to finish each step of my work continuously and carefully.

I would like to thank In-Ja Byeon, who solved the structure of Ki67FHA in collaboration with me and Yong Ju from Tsinghua University, who synthesized the phosphopeptides and phosphopeptide library for me, and Dr. Yoneda of Osaka University for kindly provided me the hNIFK gene and the Hklp2 gene. These are important data and materials to my study.

I am grateful to Chunhua Yuan who taught me how to do HSQC NMR and who helped me in many NMR experiments. Whenever I had a problem on NMR, he was always ready for help. I thank Li Zhao, who helped me on $^{31}$P NMR. I remember her patience and cheerful nature when I did experiment with her. I would like to acknowledge labmates from the FHA projects for helpful discussions and technical
communications. To name a few of them: Hua Liao, Dongyan Qin and Suganya Yongkiettrakul.

I received many advices to get familiar with the lab from Ted Jiang. This certainly helped me carry out my experiments more efficiently.

Finally, I would thank Junan Li and Brandon Lamarche for spending their precious time proofreading my dissertation.
VITA

1970    Born in Tianjin, China
1989-1993   B.S., Biochemistry, Nankai University, Tianjin, China
1993-1996   M.S., Biochemistry, Beijing University, Beijing, China
1996-now   Graduate Teaching and Research Associate
            Department of Biochemistry, The Ohio State University

PUBLICATIONS

1.  Li H, Byeon IJ, Ju Y, Tsai MD. Structure of human Ki67 FHA domain and its
    binding to a phosphoprotein fragment from hNIFK reveal unique recognition sites and
    new views to the structural basis of FHA domain functions, submitted.

2.  Li J, Li H, Tsai MD. Direct Binding of the N-Terminus of HTLV-1 Tax
    Oncoprotein to Cyclin-Dependent Kinase 4 Is a Dominant Path To Stimulate the Kinase

3. Li H, Parthasarathy S, Automatically Deriving Multi-level Protein Structures
    through Data Mining, appeared in the HiPC Workshop on Bioinformatics and
    Computational Biology, 2001

4.  Liao H, Yuan C, Su MI, Yongkiettrakul S, Qin D, Li H, Byeon IJ, Pei D, Tsai
    MD. Structure of the FHA1 domain of yeast Rad53 and identification of binding sites for

FIELDS OF STUDY

Major Field: Biochemistry
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of figures</td>
<td>x</td>
</tr>
<tr>
<td>List of abbreviations</td>
<td>xv</td>
</tr>
<tr>
<td>Chapter 1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Phosphorylation-mediated protein-protein interactions</td>
<td>1</td>
</tr>
<tr>
<td>1.2 The FHA domain</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Background of Ki67 protein</td>
<td>15</td>
</tr>
<tr>
<td>1.4 Studies on Ki67FHA domain’s binding partner</td>
<td>19</td>
</tr>
<tr>
<td>1.5 The goal of this study</td>
<td>25</td>
</tr>
<tr>
<td>Chapter 2 Materials and methods</td>
<td>28</td>
</tr>
<tr>
<td>2.1 Materials</td>
<td>28</td>
</tr>
<tr>
<td>2.2 Gene cloning, expression and purification of Ki67FHA</td>
<td>28</td>
</tr>
<tr>
<td>2.3 Gene cloning, expression and purification of GB1-hNIFK(226-269)</td>
<td>34</td>
</tr>
<tr>
<td>2.4 Gene expression of purification of Hklp2 (1017-1237)</td>
<td>35</td>
</tr>
<tr>
<td>2.5 Library synthesis</td>
<td>37</td>
</tr>
<tr>
<td>2.6 Library screening method</td>
<td>39</td>
</tr>
<tr>
<td>2.7 Solid phase synthesis of phosphorylated peptide corresponding to hNIFK (226-269)</td>
<td>41</td>
</tr>
<tr>
<td>2.8 Binding constant determination using surface plasmon resonance</td>
<td>41</td>
</tr>
<tr>
<td>2.9 Binding study using NMR HSQC titration</td>
<td>42</td>
</tr>
<tr>
<td>2.10 HPLC experiments</td>
<td>42</td>
</tr>
<tr>
<td>2.11 Kinase assay</td>
<td>43</td>
</tr>
<tr>
<td>2.12 Mutagenesis</td>
<td>43</td>
</tr>
<tr>
<td>2.13 Western blotting</td>
<td>44</td>
</tr>
<tr>
<td>Chapter 3 Special binding properties of Ki67FHA</td>
<td>46</td>
</tr>
<tr>
<td>3.1 Library screening cannot generate consensus sequence</td>
<td>46</td>
</tr>
<tr>
<td>3.2 NMR titration experiments showed that Ki67FHA does not bind to short phosphopeptides with different +3 residues</td>
<td>53</td>
</tr>
<tr>
<td>3.3 Summary</td>
<td>55</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Results of pT library screening under less stringent condition</td>
<td>50</td>
</tr>
<tr>
<td>3.2</td>
<td>Results of pT library screening under more stringent condition</td>
<td>51</td>
</tr>
<tr>
<td>3.3</td>
<td>Results of pY Library screening under less stringent condition</td>
<td>52</td>
</tr>
<tr>
<td>3.4</td>
<td>Peptide titration of Ki67FHA</td>
<td>54</td>
</tr>
<tr>
<td>4.1</td>
<td>NMR titration experiments for phospho-independent binding</td>
<td>67</td>
</tr>
<tr>
<td>5.1</td>
<td>Amino acid analysis result of P44</td>
<td>96</td>
</tr>
<tr>
<td>6.1</td>
<td>Special residues R31 and V43</td>
<td>128</td>
</tr>
<tr>
<td>6.2</td>
<td>Binding constants</td>
<td>135</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The α-helical dimeric 14-3-3 phosphopeptide complex (PDB ID=1QJB)</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>WW domain containing fragment in complex with a β-dystroglycan peptide</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Sequence alignment of FHA domains</td>
<td>6</td>
</tr>
<tr>
<td>1.4</td>
<td>Solution structure of the FHA2 domain of Rad53 complexed with a phosphotyrosyl peptide (1FHR)</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>X-Ray structure of the N-terminal FHA domain from S. Cerevisiae Rad53p in complex with a phosphothreonine peptide at 1.6 Å resolution (1G6G)</td>
<td>8</td>
</tr>
<tr>
<td>1.6</td>
<td>FHA domain from hChk2 in complex with a synthetic phosphopeptide (1GXC)</td>
<td>9</td>
</tr>
<tr>
<td>1.7</td>
<td>Crystal structure of the FHA domain of the Chfr mitotic checkpoint protein complexed with tungstate (1LGQ)</td>
<td>9</td>
</tr>
<tr>
<td>1.8</td>
<td>Binding site of the FHA1 domain of Rad53</td>
<td>11</td>
</tr>
<tr>
<td>1.9</td>
<td>Ki67 gene diagram</td>
<td>17</td>
</tr>
<tr>
<td>1.10</td>
<td>Amino acid sequence of Ki67</td>
<td>18</td>
</tr>
<tr>
<td>1.11</td>
<td>Hklp2 gene diagram</td>
<td>19</td>
</tr>
<tr>
<td>1.12</td>
<td>Amino acid sequence of Hklp2</td>
<td>21</td>
</tr>
<tr>
<td>1.13</td>
<td>hNIFK gene diagram</td>
<td>22</td>
</tr>
<tr>
<td>1.14</td>
<td>Amino acid sequence of hNIFK</td>
<td>23</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.6</td>
<td>Competition of tight-binding and weak-binding ingredients</td>
<td>99</td>
</tr>
<tr>
<td>5.7</td>
<td>An expanded region of Figure 5.6</td>
<td>101</td>
</tr>
<tr>
<td>5.8</td>
<td>$^{31}$P NMR titration of Ki67FHA vs. P44</td>
<td>103</td>
</tr>
<tr>
<td>5.9</td>
<td>Partial affinity competition purification using 3 times extra amount of P44</td>
<td>106</td>
</tr>
<tr>
<td>5.10</td>
<td>The tight binding HSQC spectrum after affinity purification</td>
<td>108</td>
</tr>
<tr>
<td>6.1</td>
<td>Structure of Ki67 FHA</td>
<td>115</td>
</tr>
<tr>
<td>6.2</td>
<td>Structure-aided sequence alignment of selected FHA domains</td>
<td>116</td>
</tr>
<tr>
<td>6.3</td>
<td>Diagram of two different kinds of chemical shift changes</td>
<td>119</td>
</tr>
<tr>
<td>6.4</td>
<td>Chemical shift perturbations of the Ki67 FHA backbone amide groups upon binding SP44-234p and hNIFK44</td>
<td>121</td>
</tr>
<tr>
<td>6.5</td>
<td>Phospho-binding center and additional contacts</td>
<td>125</td>
</tr>
<tr>
<td>6.6</td>
<td>HSQC spectrum for R31E</td>
<td>130</td>
</tr>
<tr>
<td>6.7</td>
<td>HSQC spectrum of V43R</td>
<td>133</td>
</tr>
<tr>
<td>6.8</td>
<td>Key binding residues in Ki67FHA</td>
<td>137</td>
</tr>
<tr>
<td>6.9</td>
<td>Possible interactions that determined the binding of Ki67FHA with P44</td>
<td>138</td>
</tr>
<tr>
<td>6.10</td>
<td>The role of H48 in coordinating S45 and N67</td>
<td>142</td>
</tr>
<tr>
<td>A.1</td>
<td>$^{15}$N Ki67FHA titrates with T3-3</td>
<td>156</td>
</tr>
<tr>
<td>A.2</td>
<td>$^{15}$N Ki67FHA titrates with T2-1</td>
<td>158</td>
</tr>
<tr>
<td>A.3</td>
<td>$^{15}$N Ki67FHA titrates with T2-2</td>
<td>160</td>
</tr>
<tr>
<td>A.4</td>
<td>$^{15}$N Ki67FHA titrates with T2-11</td>
<td>162</td>
</tr>
<tr>
<td>A.5</td>
<td>$^{15}$N Ki67FHA titrates with p53 peptide</td>
<td>164</td>
</tr>
</tbody>
</table>
A.6  $^{15}$N Ki67FHA titrates with Hklp2_1021p..............................  166
A.7  $^{15}$N Ki67FHA titrates with Hklp2_1088p..............................  168
A.8  $^{15}$N Ki67FHA titrates with Hklp2_1134p..............................  170
A.9  $^{15}$N Ki67FHA titrates with Hklp2_1144p..............................  172
A.10 $^{15}$N Ki67FHA titrates with Hklp2_1179p..............................  174
A.11 $^{15}$N Ki67FHA titrates with hNIFK-234p..............................  176
A.12 $^{15}$N Ki67FHA titrates with hNIFK_238p..............................  178
B.1   $^{15}$N FHA1 titration with GB1-hNIFK(226-269).......................  181
B.2   $^{15}$N FHA2 titration with GB1-hNIFK(226-269).......................  183
B.3   $^{15}$N GB1-hNIFK(226-269) titration with FHA2.......................  185
B.4   $^{15}$N GB1-hNIFK(226-269) with Ki67FHA..............................  187
B.5   $^{15}$N Ki67FHA titration with GB1 domain..............................  189
C.1   $K_d$ determination for Ki67FHA and GB1-hNIFK(226-269)...........  192
C.2   No binding found between Ki67FHA to itself..............................  183
C.3   Binding between Ki67FHA(R31E) and GB1-hNIFK(226-269)...........  194
C.4   Binding between Ki67FHA(R31E) and GB1-hNIFK(226-269) T238A...........  195
C.5   Binding between Ki67FHA(R31E) and GB1-hNIFK(226-269) T238E...........  196
C.6   $K_d$ determination for Ki67FHA V43R and GB1-hNIFK(226-269) T238E...........  197
C.7   $K_d$ determination for Ki67FHA V43R and GB1-hNIFK(226-269) T238A...........  198
C.8   $K_d$ determination for Ki67FHA V43R and GB1-hNIFK(226-269)...........  199
C.9  $K_d$ determination for Ki67FHA and synthetic phosphopeptide 234p……………………………………………………………………………… 200
C.10  $K_d$ determination for Ki67FHA and synthetic phosphopeptide 238p……………………………………………………………………………… 201
C.11  $K_d$ determination for Ki67FHA and synthetic phosphopeptide 234_238dp…………………………………………………………………… 202
C.12  $K_d$ determination for Ki67FHA V43R and synthetic phosphopeptide 234_238dp…………………………………………………………… 203
C.13  $K_d$ determination for Ki67FHA V43R and synthetic phosphopeptide 238p…………………………………………………………………… 204
C.14  $K_d$ determination for Ki67FHA V43R and synthetic phosphopeptide 234p…………………………………………………………………… 205
C.15  $K_d$ determination for Ki67FHA and GB1-hNIFK(226-269) T234A……………………………………………………………………………… 206
C.16  $K_d$ determination for Ki67FHA and GB1-hNIFK(226-269) T234E……………………………………………………………………………… 207
C.17  $K_d$ determination for Ki67FHA V43D vs. GB1-hNIFK(226-269) T234A…………………………………………………………………………… 208
C.18  $K_d$ determination for Ki67FHA V43D vs. GB1-hNIFK(226-269) T234E…………………………………………………………………………… 209
C.19  $K_d$ determination for Ki67FHA V43D vs. GB1-hNIFK(226-269) T238E…………………………………………………………………………… 210
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amp</td>
<td>Ampicillin</td>
</tr>
<tr>
<td>β-ME</td>
<td>Beta-mercaptoethanol</td>
</tr>
<tr>
<td>Chl</td>
<td>Chloramphenicol</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin dependent kinase</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetracetic acid</td>
</tr>
<tr>
<td>FHA</td>
<td>Forkhead associated</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione-S-transferase</td>
</tr>
<tr>
<td>Hklp2</td>
<td>Human kinesin-like protein 2</td>
</tr>
<tr>
<td>hNIFK</td>
<td>Human nucleolar protein interacting with the FHA domain of Ki67</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear single quantum correlation</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-beta-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>KAPP</td>
<td>Kinase associated protein phosphatase</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo Dalton</td>
</tr>
<tr>
<td>KI domain</td>
<td>Kinase interacting domain</td>
</tr>
<tr>
<td>Ki67FHA</td>
<td>The FHA domain of Ki67 protein (residues 1-120)</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>O/N</td>
<td>Overnight</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PTB</td>
<td>Phosphotyrosine binding</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>SH2</td>
<td>Src-homology 2</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris-(hydroxymethyl) aminomethane</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Phosphorylation-mediated protein-protein interactions

Phosphoprotein binding domains
Protein-protein interactions are an essential part of the machinery of cell signal transduction. Phosphorylation is one important mechanism that mediates the assembly and disassembly of protein complexes in this process. Phosphorylated proteins are usually targets of small modular recognition domains, which are structurally and functionally independent. These domains usually recognize short linear stretches of amino acid residues in their target proteins. Several such domains have been well studied including the SH2 (src-homology 2), PTB (phosphotyrosine binding), 14-3-3 and WW domains.

SH2 and PTB domains can bind to phosphorylated tyrosine motifs in transmembrane receptors and transmit the signal received by the receptor. The 14-3-3
proteins are a large family of acidic proteins of approximately 30 kDa which exist primarily as homo- and hetero- dimers within all eukaryotic cells. They mainly recognize pSer/pThr sequences. The 14-3-3 proteins have diverse roles in signal transduction, such as cell cycle control, mitogen-activated protein kinase activation, apoptosis, gene transcription and chromosome remodeling (Yaffe and Elia 2001). One of the well studied functions of 14-3-3 is the regulation of cdc25. Upon DNA damage, cdc25 is phosphorylated by Chk1 protein to generate a 14-3-3 protein binding site. The 14-3-3 protein binds to cdc25, keeps it in the cytoplasm and blocks its downstream activation of CDC2-cyclin B, thus prevents cellular mitosis. The WW domain is a protein module of around 40 amino acid residues with two highly conserved tryptophan residues. It has been found in a variety of proteins of unrelated functions, including dystrophin (a cytoskeletal protein), mouse NEDD-4 (involved in the embryonic development and differentiation of the central nervous system), and rat FE65 (a transcription-factor activator expressed preferentially in liver). Initially, WW domains were shown to bind to proline rich motifs. But they were later shown to bind to phosphoproteins as well (Lu, Zhou et al. 1999).

**Structures**

The 14-3-3 proteins are α-helical proteins which usually form dimers (Figure 1.1). The WW domain usually contains triple anti-parallel β-sheets (Figure 1.2).
Figure 1.1 The α-helical dimeric 14-3-3 ζ phosphopeptide complex (PDB ID=1QJB).

Figure 1.2 WW domain containing fragment in complex with a β-dystroglycan peptide (PDB ID=1EG4). The WW domain part features a triple anti-parallel β-sheet on the right.
Recognition diversity

Phosphoprotein-binding domains have been shown to recognize different short motifs. The SH2 and PTB domains recognize phosphotyrosine motifs (pY-hydrophobic-hydrophobic-hydrophobic or pY-hydrophobic-X-hydrophobic) (Cohen, Ren et al. 1995; Kavanaugh, Turck et al. 1995). The 14-3-3 family of proteins generally recognize pS/pT sequences such as R(S/Ar)X(pS/pT)XP and RX(Ar/S)X(pS/pT)XP where Ar denotes aromatic residues (Muslin, Tanner et al. 1996). The WW domains have been shown to recognize pThr-Pro/pSer-Pro. But exceptions exist for the rules described above. The 14-3-3 protein can bind sequences that are very different from the consensus sequence described above or even bind non-phosphorylated sequences. The main recognition targets for the WW domains are proline-rich sequences such as PPXY, PPLP or PPR (Sudol and Hunter 2000; Zarrinpar and Lim 2000). Recognition of pThr/pSer is considered to be a functional gain that has only evolved for a few proteins.

Function pattern diversity

There are different function patterns of phosphoprotein-binding domains. For example, the 14-3-3 proteins appear to affect intracellular signaling in one of three ways: (1) direct regulation of the catalytic activity of the bound protein, (2) regulate interactions between the bound protein and other molecules in the cell by sequestration or modification, or (3) control the subcellular localization of the bound ligand.
1.2 The FHA domain

Discovery and existence

The FHA (forkhead associated) domain was found in 1995 by computer-aided sequence alignment (Hofmann and Bucher 1995). A typical FHA core domain comprises 55-75 amino acid residues. These residues are divided into three highly conserved blocks separated by more divergent spacer regions as shown in Figure 1.3. There are only three 100% invariant residues. Later studies showed that the core domain is not long enough to fold independently and form a compact structure. The functional size of FHA domains varies, but it usually contains at least 80-100 amino acid residues (Durocher, Henckel et al. 1999; Liao, Byeon et al. 1999; Hammet, Pike et al. 2000; Sueishi, Takagi et al. 2000).

To date, more than 200 proteins containing FHA domains have been identified with computer-aided sequence analysis. These include DNA damage repair and signaling proteins (such as Chk2, RAD53, Dun1, and NBS1), kinesins (such as KIF1A, KIFB, KIFC in the UNC104 family, and KIAA0042 in the KIF14 family), ring-finger proteins (such as Chfr), forkhead transcription factors (such as Fkh1, Fkh2, Flh1, and Flh2), Ki67 and nuclear inhibitor of protein phosphatase 1 (NIPP-1) (Durocher and Jackson 2002). The diversity of FHA domain containing proteins indicates that FHA domains have a wide range of functions in signal transduction. FHA domains are found in eukaryotes and eubacteria, but not in archaeal.
Figure 1.3 Sequence alignment of FHA domains (Hofmann and Bucher 1995). A predominant β-sheet structure was predicted for FHA domains as indicated by black bars.
**Structure**

The first FHA domain structure, an NMR solution structure of yeast Rad53, was reported in 1999 (Liao, Byeon et al. 1999). Later, more NMR and X-ray diffraction structures from different sources were also reported. Some of them are in complex with synthetic peptides. The currently available structures include: two FHA domains from Rad53 (free and with peptide) (Liao, Byeon et al. 1999; Durocher, Taylor et al. 2000; Liao, Yuan et al. 2000; Wang, Byeon et al. 2000), the FHA domain of Chk2 in complex with a synthetic peptide (Lee and Chung 2001), and a Chfr dimer with tungstate (an analog of phosphate) (Stavridi, Huyen et al. 2002).

FHA domains usually fold into an 11-stranded \( \beta \)-sheet sandwich. There can be some insertions of small \( \alpha \)-helices in the loops between \( \beta \)-strands. The available structures of FHA domains are shown in Figures 1.4 to 1.7.
Figure 1.4 Solution structure of the FHA2 domain of Rad53 complexed with a phosphotyrosyl peptide (1FHR).

Figure 1.5 X-ray structure of the N-terminal FHA domain from Saccharomyces Cerevisiae Rad53p in complex with a phosphothreonine peptide at 1.6 angstrom resolution (1G6G).
Figure 1.6 FHA domain from hChk2 in complex with a synthetic phosphopeptide (1GXC)

Figure 1.7 Crystal structure of the FHA Domain of the Chfr mitotic checkpoint protein complexed with tungstate (1LGQ)
Recognition site

The FHA domain has been demonstrated to be a phosphothreonine peptide-binding domain (Durocher, Henckel et al. 1999; Durocher, Taylor et al. 2000; Liao, Yuan et al. 2000). In the complex structures of the FHA1 domain of Rad53 and the FHA domain of Chk2, the phosphopeptide interacts with the loops connecting $\beta_3/\beta_4$, $\beta_4/\beta_5$ and $\beta_6/\beta_7$. The structure of the FHA1 domain of Rad53 shows that five conserved residues (G69, R70, S85, H88 and N107) are located around the phosphopeptide-binding site. Three of them (R70, S85 and N107) make direct contact with the peptide. R70 and S85 bind directly to pThr of the peptide while N107 binds to the backbone NH of the +3 position. N112 is also a conserved residue, but it is not directly involved in the binding site. It tethers $\beta_{10}$ to the loop between $\beta_7/\beta_8$.

In both the FHA1 domain of Rad53 and the FHA domain of Chk2, it has been shown that the +3 position is critical for the pT peptide selection. In the case of Rad53 FHA1, R83 forms two salt bridges with +3 Asp in the pT peptide through its guanidinium moiety, which is critical for the binding selectivity, while other key interactions at the binding site maintain the conformation of the peptide. In the case of Chk2, the FHA domain prefers an Ile at the +3 position, due to the presence of a shallow pocket on the FHA surface formed by Asn166, Ser192/Leu193, and Thr138.
Figure 1.8 Binding site of the FHA1 domain of Rad53 (Durocher, Taylor et al. 2000). Note that the guadinium group of R83 forms two salt bridges with the +3 position, which is critical for binding selectivity.
FHA domains have been shown to prefer pThr to pSer in several cases. In the case of Chk2 FHA, the γ-methyl group of pThr in the peptide targets interacts with the side chain of a conserved Asn residue (N166) together with the side chain and/or backbone atoms from Thr138, Tyr139, and Ser140 through Van der Waals forces.

Besides the pThr peptide preference of FHA domain, phosphotyrosine peptides were also shown to bind to the FHA2 domain of Rad53 (Wang, Byeon et al. 2000). In this case, a Tyr at the +2 position was preferred. Low affinity to phosphotyrosine and phosphothreonine was observed by NMR titration experiments using a high concentration of ligand. Recently, a crystal structure of Chfr FHA was determined together with tungstate, an analog of phosphate (Stavridi, Huyen et al. 2002).

**Function of FHA**

As mentioned above, FHA domains have been found in a variety of proteins with different functions. There are a few FHA-domain-containing proteins whose biological functions have been shown to be dependent on this domain. The following three proteins have been relatively well studied.

**KAPP**

KAPP (kinase-associated protein phosphatase) was cloned as an interacting partner protein of receptor-like protein kinase RLK5 (Stone, Collinge et al. 1994). Since a 239-
amino acid residue-long fragment (residues 98-336) in the middle of KAPP interacts with phosphorylated RLK5 kinase domain, it was named as KI (Kinase Interacting) domain. Later, an FHA domain was identified in the KI domain and in vitro binding assays using shorter fragments further identified that the minimal fragment of KI domain that can interact with receptor-like protein kinase spans amino acid residues 180 to 298 which covers the FHA domain (Li, Smith et al. 1999). Site-directed mutagenesis verified that the conserved residues of the FHA domain (G211, S226, H229 and N250) play essential roles in interacting with phosphorylated RLKs. Plant receptor-like protein kinases are important signal transduction proteins in hormone reception just like the receptor tyrosine protein kinases. The plant FHA domain in KAPP plays a similar role as the SH2 domain, which binds to phosphorylated receptor tyrosine protein kinases. In vivo experiments further confirmed the role of the FHA domain of KAPP as a negative regulator of signal transduction mediated by RLKs (Williams, Wilson et al. 1997; Stone, Trotochaud et al. 1998).

Rad53

Rad53 is a check point kinase in S. cerevisiae and is unique in that it contains two FHA domains, one at either terminus. In the middle of the protein, there is a Ser/Thr kinase domain. In response to DNA damage, a series of signals are transduced which, consequently, lead to cell cycle arrest at G2/M. Rad53 has been shown to be a downstream regulator of the Mec1 kinase (Sanchez, Desany et al. 1996) and an upstream regulator of Dun1, another kinase which mediates DNA damage response (Zhou and
Rad53 interacts with Rad9, whose phosphorylation is induced by DNA damage under the control of Mec1/Tel1, through the FHA2 domain in a phosphorylation dependent manner (Sun, Hsiao et al. 1998). The biological importance of the FHA2 domain of Rad53 was demonstrated by the fact that mutations in the FHA2 domain abolish check-point cell cycle arrest. Though there is some evidence that the FHA1 domain of Rad53 also plays an important role in the DNA damage response, the details need to be further explored. It is possible that the two FHA domains function at different stages of the cell cycle.

**Chk2**

Human Chk2 protein is considered to be a homolog of yeast Rad53. Like Rad53, it also has a kinase domain. But instead of having two FHA domains, it has only one FHA domain near the N-terminus. There is an SQ/TQ rich region at the very N-terminus. After DNA damage, Chk2 kinase phosphorylates a series of key cell cycle regulator proteins, including BRCA1 (Lee, Collins et al. 2000), p53 (Chehab, Malikzay et al. 2000) and CDC25 (Blasina, de Weyer et al. 1999; Furnari, Blasina et al. 1999). The importance of the FHA domain in transducting cell signals was demonstrated by two oncogenic mutants, I157T and R145W, in the FHA domain. These two mutants were identified in Li-Fraumeni syndrome patients with wild type p53 (Lee, Kim et al. 2001). Since this syndrome is usually associated with mutations in p53, it is possible that the FHA domain mediates the upstream/downstream function of p53. One possible functional pathway of the FHA domain of Chk2 is through upregulation of the kinase activity, which in turn
phosphorylates p53 at S20 and stabilizes it. Upon DNA damage, ATM phosphorylates Chk2 at the SQ/TQ rich region, particularly at T68. T68 phosphorylation further induces autophosphorylation (at T383, T387), which is key for high kinase activity (Lee and Chung 2001). The R145W mutant, which is severely injured in binding phosphopeptides, blocks this autophosphorylation. It is possible that Chk2 forms a dimer through the FHA domain when T68 is phosphorylated and thus mediates the activation process. The I157T mutant, on the other hand, does not show decreased binding affinity to synthetic phosphopeptides, but it has been shown to abolish the binding of several target proteins of Chk2 including BRCA1 and CDC25A. Since I157 is located far from the putative binding center, the data suggest an additional selection mechanism, which has not been elucidated by the complex structures of FHA domains and synthetic phosphopeptide, for binding to biological target proteins.

### 1.3 Background of Ki67 protein

**Discovery and the “missing function”**

The Ki67 protein (Ki-67p, pKi67) was originally identified as an antigen recognized by the monoclonal antibody Ki67 (Gerdes, Schwab et al. 1983). This antigen is present in the nuclei of cells during the G1, S, and G2 phases of cell cycle, as well as in mitosis but not during G0 phase of quiescent cells (Gerdes, Lemke et al. 1984). The stringent existence of Ki67 in proliferating cells has made Ki67 a popular prognostic and
Ki67 is a constituent of compact chromatin (Kreitz, Fackelmayer et al. 2000) and is vital for cell proliferation since down-regulation of Ki67 using antisense RNA prevents cellular proliferation (Schluter, Duchrow et al. 1993). However, knowledge of the exact role of Ki67 protein is limited, due in part to the lack of sequence homology with other proteins/domains of known function (Scholzen and Gerdes 2000).

Ki67 displays many characteristics of a signal transduction protein: a simple primary sequence analysis of Ki67 protein revealed 10 putative nuclear targeting sequences, 143 potential phosphorylation sites for protein kinase C, 89 sites for casein kinase II, 2 tyrosine kinase sites, and 8 consensus sites for CDC2/cyclin B kinase, the key regulator of mitosis transition (Endl and Gerdes 2000). The cellular location of Ki67 is dynamic. At interphase, Ki67 is located mainly in the nucleolus. During cell division, Ki67 is located around the chromatin. These suggest that this protein not only functions in cell proliferation but is also involved in the protein interaction network that drives the cell division cycle. Recent studies on FHA domains have shed light on the possible phosphoprotein interaction role of Ki67. This, together with recent studies on the C-terminus of Ki67, could provide insights into how Ki67 is involved in the whole process of cell cycle control.
**Ki67 gene structure**

The human Ki67 gene locus includes nearly 30,000 base pairs. It has 15 exons which are spliced into two prominent forms (Schluter, Duchrow et al. 1993) with calculated molecular weights of 320 kDa and 359 kDa, respectively. The shorter form of Ki67 lacks amino acids 136-495 (as shown in Figure 1.9).

![Figure 1.9 Ki67 gene diagram](image)

The most salient feature of Ki67 is that it contains 16 approximate repeats of around 122 amino acids. There is a potential ATP binding site at 3034-3041. The FHA domain is located at the N-terminus. While computer alignment showed that the core FHA domain spans residues 27-76, the functional domain is larger.
Figure 1.10 Amino acid sequence of Ki67
1.4 Studies on the binding partner of Ki67FHA domain

To study the potential function of Ki67 FHA domain, Prof. Yoneda’s lab identified two Ki67FHA interacting proteins: Hklp2 (Sueishi, Takagi et al. 2000) and hNIFK (Takagi, Sueishi et al. 2001).

Hklp2

Hklp2 (human kinesin-like protein 2) is a human homolog of Xklp2 (Xenopus kinesin-like protein 2). It has 1388 amino acid residues with positions 32-391 making up a kinesin motor domain. There are two leucine-zippers and one ATP/GTP-binding site (109-116) within the kinesin motor domain (Figure 1.11).

Figure 1.11 Hklp2 gene diagram
Xklp2 is a plus-end directed kinesin-like motor required in centrosome separation and maintenance of spindle bipolarity during mitosis in Xenopus egg extracts (Verheijen, Kuijpers et al. 1989; Verheijen, Kuijpers et al. 1989). Hklp2 shares 53% identity with Xklp2 over the entire sequence and 87.5% identity in the kinesin motor domain. The Ki67 FHA domain (1-99) interacting region was mapped to 1017-1237. Hklp2 can be phosphorylated by mitotic but not interphasic Hela cell extracts. The interaction between Ki67 FHA domain and Hklp2 is more efficient when the latter is treated with mitotic cell extract. Immunofluorescence studies revealed an association between Ki67 and Hklp2 at the periphery of mitotic chromosomes, largely in proximity to the centromeres. Chmadrin, a rat kangaroo protein similar to Ki67, generates severe heterochromatin formation when its c-terminus is overexpressed (Takagi, Matsuoka et al. 1999). The C-terminus of Ki67 can also cause gross changes in chromatin organization (MacCallum and Hall 2000). Therefore, the Ki67 seems to mediate the connection of the motor protein Hklp2 with chromatin through its FHA domain. Recent studies using the C-terminus of Ki67 further support role of Ki67 in regulating higher order chromatin organization (Scholzen, Endl et al. 2002). In this case, Ki67 interacts with members of the heterochromatin protein I family, whose role is targeting proteins to the mitotic spindle (Ainsztein, Kandels-Lewis et al. 1998).
Figure 1.12 Amino acid sequence of Hklp2
hNIFK

hNIFK stands for human nucleolar protein interacting with the FHA domain of Ki67. It contains 293 amino acid residues. There is one RNA-binding domain and one lysine rich domain (Figure 1.13). The Ki67 FHA domain binding motif was mapped to residues 226-269. This motif binds the Ki67 FHA domain efficiently when it is treated with mitotic cell extracts. Phosphatase treatment abolishes this efficient binding. This suggests that this binding is phosphorylation dependent. Since mutation at either T234 or T238 to Ala also abolishes this efficient binding, these two threonines are considered to be candidate phosphorylation sites for phospo-dependent binding to Ki67.

![Figure 1.13 hNIFK gene diagram](image)

hNIFK exists mainly in the nucleolus. Its RNA-binding function was verified by an RNase A treatment, in which great reduction of hNIFK staining was observed in the treated cells. hNIFK co-localizes with Ki67 at the surface of mitotic chromosomes in mitotic cells. It is thought to mediate rRNA metabolism because of its nucleolar location.
During mitosis, rRNA genes are silenced under the control of CDC2-Cyclin B kinase. It is possible that hNIFK mediates rRNA synthesis and is silenced by Ki67 FHA domain when it is phosphorylated.
Figure 1.14 Amino acid sequence of hNIFK
1.5 The goal of this study

To investigate the function of Ki67

As discussed earlier, despite the vital role of Ki67 in cell proliferation, the function of Ki67 remains an enigma since we know very little about Ki67, particularly in the area of signal transduction control. The FHA domain of Ki67 could be a starting point in understanding the signal transduction related functions. The discovery of the two binding partner proteins of Ki67 FHA domain shed some light on the function of Ki67. Since both these two proteins are involved in mitosis, there could be complex signal transduction pathways. Understanding these pathways may contribute to our knowledge not only of Ki67 alone but also of cell cycle control. Though the effect of phosphorylation of the target protein on binding has been qualitatively demonstrated, the detailed binding mechanism remains to be explored. Identifying exact phosphorylation site(s) and elucidating the binding mechanism is the direct goal of this study.

Biological target for FHA domain

To date, the study of FHA domains has been limited to structure determination and binding pattern studies using small phosphopeptides identified in library screening. While there is evidence that small phosphorylated peptides were enough to confer tight binding, this may not completely reflect the biological situation. Ki67 is an ideal protein
to study additional binding interactions of biological target proteins of the FHA domain since the binding of Ki67 to its two target proteins require minimal sizes of fragments in its target proteins (221 amino acid residues for Hklp2 and 44 amino acid residues for hNIFK) (Sueishi, Takagi et al. 2000; Takagi, Sueishi et al. 2001) which are much longer than sequences of phosphothreonine peptides used in complex structure determination and binding study. This longer sequence requirement has been proposed to be an indication of 3-D structure requirement (Li, Lee et al. 2000).

To explore the diverse binding specificity of FHA domain
The functions of several FHA domains have been studied, providing detailed discussions about the roles of key residues. Though these studies provided very good insight, more information needs to be collected because of the diverse binding specificity of FHA domains. The FHA2 domain of Rad53, whose specificity is not so easily defined by phosphopeptide library screening and exhibits dual specificity for pT and pY, may be functionally different to the FHA1 domain of Rad53 whose specificity was clearly determined. Most FHA domains exist as monomers. However, the Chfr FHA domain exists as a dimer and binds to tungstate. All these suggest that FHA domains could be further divided into subfamilies based on their binding specificity. FHA domains adopt a similar structure of an 11-stranded β-sheet sandwich. Small differences at the binding site could affect specificity greatly. Therefore, selectivity determining residues need to be explored in depth. Binding induced structural change is of great interest since it would not only provide further validation of binding but also could provide insights into the
mechanism of regulation of target proteins. This is particularly true for the Ki67 FHA domain. The target proteins of Ki67 FHA domain contain a Thr-Pro motif, which is often recognized by cyclin-dependent protein kinases (CDKs) (Nigg 1995) and MAPKs. Both cyclin-dependent protein kinases and MAPKs are important regulators of cell cycle control. The ability of Ki67 FHA domain to bind this motif could be an important feature that nature has evolved. The proline after threonine can block the activity of certain kinases and the special α-amine group of proline can be an obstacle for peptidases. It will be interesting to see how this proline affects the binding of Ki67 FHA domain to its biological target.

Based on these features in both biological function and selectivity, Ki67 FHA is good choice to study the diversity of FHA domain functional patterns.
CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

Chemicals were purchased from Sigma unless otherwise specified. Restriction enzymes were from New England Biolabs. Sensorchip SA was from BIAcore.

2.2 Gene cloning, expression and purification of Ki67FHA

Gene cloning of Ki67FHA

A cDNA clone (gb262877) harboring a partial ORF corresponding to Ki67FHA was purchased from Incyte Genomics, an official distributor of IMAGE Consortium cDNA clones. Primers were designed to amplify the coding region corresponding to residues 1-120 of Ki67 using polymerase chain reaction (PCR). The PCR product was subcloned into the BamHI and NotI sites of the pEG vector, a derivative of pET29a with the GST
gene downstream of the T7 promoter to generate pEG-KiFHA. The sequence of Ki67FHA gene was verified by DNA sequencing.

**Expression of Ki67FHA**

pEG-KiFHA plasmid was transformed into *E. coli* BL21 (DE3) CodonPlus cells (Stratagene) for overexpression.

*For non-labeled sample:*

A single colony of BL21 (DE3) CodonPlus bacterial strain harboring pEG-KiFHA plasmid was inoculated into 200 ml of LB medium (Kan, Chl) and grown overnight at 37 °C. This seed culture was diluted 1:100 into 10 L of LB medium (Kan, Chl) and grown to OD$_{600}$=0.6-0.8. 1 ml of seed culture was taken for long-term storage in –80 °C freezer with 15% glycerol. 1 mM of IPTG was added to induce expression. After induction, incubation temperature was changed to 25 °C since the expressed protein is more stable at lower temperature. Cells were harvested after incubating 8-10 h and stored at –20 °C.

*For $^{15}$N labeled or $^{15}$N, $^{13}$C double labeled sample:*

A single colony of BL21 (DE3) CodonPlus bacterial strain harboring the pEG-KiFHA plasmid was inoculated into 200 ml of $^{15}$N labeled or $^{15}$N, $^{13}$C double labeled M9 medium (Kan, Chl) with a supplement of thiamine and grown overnight at 37 °C. This seed
culture was diluted 1:100 into 10 L of M9 medium (Kan, Chl) and grown to OD_{600}=0.8-1.0. IPTG was added to a final concentration of 1 mM to induce protein expression. After induction, incubation temperature was changed to 25 °C since the expressed protein is more stable at lower temperature. Cells were harvested after incubating 14 h and stored at –20 °C.

Purification of Ki67FHA

The whole procedure was performed at 4 °C unless otherwise specified.

Cell lysis

The frozen cells were resuspended in lysis buffer (50 mM HEPES, pH 7.5 at RT, 300 mM NaCl, 5 mM DTT, and 2 mM EDTA). Four ml of lysis buffer were used per gram of cell pellet. The suspension was sonicated using a Sonic Dimembrator 550 from Fisher Scientific (1 s on, 3 s off in each cycle for a total on-time of 5 min at power setting 7). The cell lysate was centrifuged using a Beckman JA-20 rotor at 19,000 rpm (48,400 g) for 2 h to clear the supernatant.

Affinity column

The clear cell lysate was loaded on glutathione-agarose beads (Pharmacia) pre-equilibrated with lysis buffer. After washing the column with 10 volumes of lysis buffer,
the glutathione-agarose beads, together with 1 bed volume of buffer, were taken to room temperature. An appropriate amount of thrombin (usually 1 mg) was added to proteolytically remove the GST tag. After incubating at RT for 10 h, the Ki67FHA protein (without the GST tag) was eluted from glutathione-agarose beads using lysis buffer.

**Gel-filtration column**

The eluent was concentrated to less than 10 ml using an Ultrafree 15 centrifugal filter device from Millipore (MWCO=5 kDa), loaded onto an S100 column (Pharmacia) and eluted with S100 buffer (5 mM HEPES pH7.5 at RT, 2 mM DTT, 1 mM EDTA, 150 mM NaCl). Appropriate fractions were collected and concentrated to >0.4 mM and lyophilized for future use. The protein was estimated to be more than 90% pure using SDS-PAGE and silver staining.

**2.3 Gene cloning, expression and purification of GB1-hNIFK(226-269)**

**Gene cloning of hNIFK(226-269)**

A cDNA clone (gi10211895) harboring a partial ORF corresponding to hNIFK was purchased from Incyte Genomics. Primers were designed to amplify the coding region corresponding to residues 226 to 269 of hNIFK using PCR. The PCR product was
subcloned into the BamHI and XhoI sites of pGB vector to generate pGBhNIFK(226-269). The sequence of hNIFK(226-269) was verified by DNA sequencing as shown in Figure 2.1.
Figure 2.1 GB1-hNIFK gene expression product.
**Gene expression of GB1-hNIFK**

Gene expression of GB1-hNIFK followed the same procedure as described in 2.2

**Purification of GB1-hNIFK and GB1 domain**

The procedures were performed at 4 °C unless otherwise specified.

*Cell lysis*

The frozen cells were resuspended in lysis buffer (50 mM HEPES, pH 7.5 at RT, 300 mM NaCl, 15 mM β-ME). Four ml of lysis buffer were used per gram of cell pellet. The suspension was sonicated using a Sonic Dismembrator 550 from Fisher Scientific (1 s on, 3 s off in each cycle for total on-time of 5 min at power setting 7). The cell lysate was centrifuged using a Beckman JA-20 rotor at 19,000 rpm (48,400 g) for 2 h to get clear supernatant.

*Affinity column*

The clear cell lysate was loaded on a nickel-NTA beads (Novagen) pre-equilibrated with lysis buffer. The resin was extensively washed with 10 bed volumes of lysis buffer. The
washed nickel-NTA beads were taken to room temperature and eluted with elute buffer (50 mM HEPES, pH7.5 at RT, 300 mM NaCl, 15 mM β-ME, 500 mM of imidazole).

**Gel-filtration column**

The eluent was concentrated to less than 10 ml using Ultrafree 15 centrifugal filter device from Millipore (MWCO=5 kDa) and then loaded onto S100 column (Pharmacia) and eluted with S100 buffer (5 mM HEPES pH7.5 at RT, 2 mM DTT, 1 mM EDTA, 150 mM NaCl). Corresponding fractions were collected, concentrated to >0.4 mM and lyophilized for future use. The protein was estimated to be more than 90% pure using SDS-PAGE gel and silver staining.

**2.4 Gene expression and purification of Hklp2 (1017-1237)**

The pRSET plasmid harboring gene encoding Hklp2(1017-1237) was a generous gift from Dr. Mashatoshi in Prof. Yoneda’s lab (Osaka University, Japan).

The expression procedure is the same as the expression procedure described in 2.2, except that Ampicillin (200 mg/L) was used instead of kanamycin.

The purification procedure is the same as that described in 2.3.
Figure 2.2 Sequence of expressed Hklp2(1017-1237).
2.5 Library synthesis

The phospho-peptides used for library screening were synthesized by Hua Liao and Yong Ju by the following procedure:

Libraries were constructed with an in-house built peptide synthesis apparatus. The synthesis was performed at room temperature on a 3.0 g scale of TentaGel S-NH2 resin (0.3 mmol/g loading, $2.86 \times 10^6$ beads/g, and 80-100 µm) using Fmoc/HBTu/HOBt chemistry. The linker C terminus-MRBBI/A-N terminus, was first synthesized on the resin, followed by acetyl-AXXX(pT/S/Y)XXX or acetyl-AXXXX(pT)XX depending on different batches. The split synthesis approach was used for randomization of the sequence, resulting in a one bead-one sequence library. Amino acid residues with the same molecular weight were differentiated by capping reagents: 5% acetylglycine was added in the coupling reactions of Lys and norLeu. The complete reaction of each round was monitored by the ninhydrin test. Deprotection was performed in the final stage by reacting for 2 h with a large excess of solution made up of crystalline phenol, trifluoroacetic acid (TFA), anisole, thioanisole, and 1,2-ethanedithiol. The beads were thoroughly washed alternatively with methanol and dichloromethane, dried under vacuum, and kept at -20 °C.

The split synthesis strategy is diagrammed in Figure 2.3.
Figure 2.3 The library synthesis strategy. This strategy ensures that there is one unique sequence on each bead. Note that in each step, 5% peptides on each bead are terminated by AcetylAla and/or acetylGly for subsequent sequence determination by mass spectrometry.
2.6 Library screening method

Preparing biotinylated proteins

Expressed proteins containing FHA domains (Ki67FHA, Dun-1 FHA, Chk2 FHA and Rad53 FHA1) were first dialyzed against PBS (50 mM sodium phosphate pH 8.0, 150 mM NaCl) to remove reagents containing primary amine groups or thiol groups in S100 buffer. The biotinylation reaction was carried out by adding a 1.5-fold excess of NHS-biotin (Sigma) and incubating at RT for 2 h or at 4 °C overnight. Subsequently glycine was added to a final concentration of 0.1 mM to react with unreacted NHS-biotin. Finally, glycine and other small molecules were removed by dialysis against PBS.

Screening positive beads

Library screening was carried out in small batches of 10~20 mg of beads (2800~5600 beads) for each protein. The beads were first washed with 1 ml of DMF 3 times (each round for 15 min) to clean left-over chemicals. Beads were then sequentially washed with 1 ml of HBS (25 mM HEPES pH 7.5, 150 mM or 1 M NaCl) for 10 times, 1 ml of HBST (HBS + 0.1% Tween 20) for 10 times and 1 ml of HBSTB (HBST + 1 mg/ml BSA) for 5 times. After incubating the beads with HBSTB for 1 h to block non-specific binding, 0.1-1 µM of biotinylated protein was added to the beads and incubated for 4 h at RT. Then the beads were washed with 1 ml of HBSTB for 5-10 times to remove extra protein. In 1 ml of HBSTB, 0.5 µg of streptavidin AP conjugate was added to bind to
biotinylated protein and 20 μM of PO$_4^{2-}$ was added to prevent the alkaline phosphatase from cleaving the phosphate group of peptides on the beads. After incubating for 15 min, the beads were sequentially washed with 1 ml of HBSTB for 5-10 times, 1 ml of HBST for 5-10 times, 1 ml of HBS for 5-10 times and 1 ml of TBS (25 mM Tris pH8.0, 150 mM NaCl) for 5 times. BCIP was added to a final concentration of 0.5 mg/ml to develop color. After 1-2 hours, positive beads (blue color) were identified under microscope.

**Bead cleavage**

The positive beads were transferred into separate Eppendorf tubes. 20 μl of CNBr solution was added (20 mg CNBr in 1 ml 70% formic acid) to cleave the peptide off the beads. After reacting for 24 h at RT in dark, the beads were dried with vacuum dryer and dissolved in 5 μl of 0.1% TFA.

**Sequence determination**

1 μl of cleaved sample together with 2 μl matrix (saturated αCCA in 0.1% TFA, 50% acetonitrile (CH$_3$CN)) was spotted on a mass spectrometry metal plate (bruker #25184). The sequence of the peptide was determined using MALDI-TOF.
2.7 Solid phase synthesis of phosphorylated peptide corresponding to hNIFK (226-269)

The phosphorylation peptides initially used in this study (234p, 238p, 234_238dp) were synthesized by Dr. Yong Ju (Tsinghua University, China) from C-terminus to N-terminus by the solid phase peptide synthesis protocols using in-house apparatus. After identification of T234 as the phosphorylation site, a better quality peptide (P44) corresponding to 234p was purchased from Genemed Inc.

2.8 Binding constant determination using surface plasmon resonance

The synthesized phosphopeptides or expressed peptides corresponding to hNIFK(226-269) and BSA(Sigma) were biotinylated with a 1.5-fold excess of NHS-biotin (Sigma) by incubating at 4°C overnight and then dialyzed to remove unreacted excess NHS-biotin. The biotinylation was verified by Western Blotting using streptavidin alkaline phosphatase conjugate (Pierce). Biotinylated peptides were immobilized on a BIAcore sensorchip SA with biotinylated BSA serving as the control channel. Ki67FHA and its mutants in HBSED buffer (5 mM HEPES, 150 mM NaCl, 2 mM DTT and 1 mM EDTA, pH 7.5) were prepared by serial dilution to concentrations ranging from 0.01 to 1000 µM and flowed through the sensorchip. Surface plasmon resonance was detected using BIAcore 3000. The response unit changes vs. concentration plots were used to fit either a double rectangular hyperbolic curve \[ RU = RU_{max1} \frac{c}{c+K_{d1}} + RU_{max2} \frac{c}{c+K_{d2}} \] or a
single rectangular hyperbolic curve \[RU = RU_{max} \times c/(c+K_d)\] to get the dissociation constants depending on whether there was a tight binding. Regeneration buffer (1 M urea + 1 M NaCl) was used to regenerate the sensorchip surface after each round of injection.

2.9 Binding study using NMR HSQC titration

HSQC spectrum of ~0.4 mM $^{15}$N labeled protein (Ki67 FHA, FHA1, FHA2 or GB1-hNIFK(226-269)) was initially acquired as a reference. A series of mixtures containing $^{15}$N-labeled protein and its binding partner (protein: peptide=4:1, 2:1, 1:1 and 1:2) were prepared by adding the binding partner (high concentration, in the same buffer) after each round of HSQC. The HSQC spectra at all ratios were plotted and compared. All experiments were conducted at 293K. 10% D$_2$O was used.

2.10 HPLC experiments

**Ion exchange**

A polymer-based strong cation exchange column was purchased from Vydac. The peptide sample was first treated with 20 mM DTT, 0.1% TFA and then loaded onto the column. The column was washed with 20 ml of low salt eluent (0.1% TFA, 30% acetonitrile) and then eluted with a linear gradient of NaCl.
**Reversed phase**

Semi-preparative reversed phase chromatography columns (C5, C8 and C18) were purchased from Vydac. Pretreated peptide sample was loaded onto the column and washed with 40 ml of 0.1% TFA. Peptide was eluted with a linear gradient of acetonitrile (buffer A 0.1% TFA, buffer B 0.1% TFA + 50% acetonitrile).

**2.11 Kinase assay**

In 50 µl of total volume, 0.1 mM substrate (GB1hNIFK(226-269)) was added to the kinase buffer (0.2 mM ATP, 50 mM HEPES, 10 mM MgCl₂, 2.5 mM EGTA, 0.1 mM Na₃VO₄, 1 mM NaF, 10 mM β-glycerophosphate, 1mM DTT, pH 7.5 at RT). After adding 3-10 units of kinase (3-10 units of CDK4 prepared by our lab or 5 units of cyclin B/CDC 2 from Promega), the reaction was conducted at RT for 10 h. GB1-hNIFK(226-269) 234A and GB1hNIFK(226-269) 238A mutants were used as substrate control.

The reaction product was examined using western blotting with pThr-Pro antibody and followed by ECL detection kit (Pharmacia).

**2.12 Mutagenesis**

Mutagenesis in this study was performed using Stratagene’s Quick Change Kit.

Primers for hNIFK (complementary primers are omitted):
T234E: ctgtggatagccagggccccgaaccagtttgtacaccaacatttttg

T234A: ctgtggatagccagggccccgccccagtttgtacaccaacatttttg

T238E: gggccccacaccagtttgtgaaccaacatttttggagaggcg

T238A: gggccccacaccagtttgtgccccaacatttttggagaggcg

Primers for Ki67FHA (complimentary primers are omitted)

V43R: catccgtatccagcttcctcgtgtgtcaaaacaacattgc

R31E: cagcactctgttttttggagaaggtattgaatgtgacatc

2.13 Western blotting

Protein samples were electrophoresed either using the Phast System (Pharmacia) or Mini-protein II (BioRad).

For the Phast system, nitrocellulose membrane (from BioRad) was applied to the surface of the gel. Protein was transferred to the membrane by heat diffusion at 70 °C for 30 min.

For the BioRad system, protein was transferred to the membrane by electrophoresis transfer in the electrophoresis transfer set (from BioRad) using 60 V for 2
h or 10 V for 10 h in transfer buffer (0.37 g of SDS, 2.9 g of glycine, 4.8 g of Tris base, 200 ml of methanol in 1 liter H2O).

The membrane was blocked using Blotto-Tween (25 mM Tris HCl pH 7.4, 140 mM NaCl, 5% nonfat dried milk, 0.02% NaN3, 0.05% Tween 20, 2.5 mM KCl) for 30 min and then the first antibody was added and incubated for 2 h at RT or 4 °C O/N with shaking. Then the membrane was washed with Western Wash Buffer (WWB, 25 mM Tris HCl pH 7.4, 140 mM NaCl, 0.02% NaN3, 0.05% Tween 20, 2.5 mM KCl) three times with shaking, and soaked with Blotto-Tween. Then 10 µl of secondary antibody was added and incubated for 30 min at RT with shaking. Then the membrane was washed 3 times with WWB, 10 min each time with shaking. Either BCP/TCIP (from Pierce) or ECL (X-film from Pharmacia, chemiluminescence material from Pierce) was used to visualize bands on the membrane depending on the secondary antibody used (HRP conjugate for ECL, AP conjugate for BCP/TCIP). If HRP was used, NaN3 was removed from the blotto-Tween and Western Wash buffer since it is an inhibitor for HRP.
CHAPTER 3

SPECIAL BINDING PROPERTIES OF KI67FHA

3.1 Library screening cannot generate consensus sequence

Library screening strategy

Library screening is a useful method for identifying linear binding sequences of target proteins. The one-bead-one-sequence library synthesis strategy (Lam, Salmon et al. 1991), as described in chapter 2.5 and illustrated in Figure 2.4, makes library screening an easy procedure. The library screening method described in section 2.5 and diagrammed in Figure 3.1 (150 mM of NaCl in all the buffers) has been successfully used to identify tight binding phosphothreonine peptides for the FHA1 domain and phosphotyrosine peptides for the FHA2 domain of Rad53 (Liao, Yuan et al. 2000; Wang, Byeon et al. 2000). So the same method was used to screen the FHA domain of Ki67.
Figure 3.1 Diagram of library screening procedure
**Library sequence**

Three different kinds of libraries (pT, pS and pY) were used in this study. Depending on different batches of synthesis, the library varied slightly in linker and in sequence. Initial libraries used MRBBA as the linker (M for cleavage by CNBr, R for positive ionization in MALDI-TOF, β-Alanine, B for flexibility) and the library sequence has a sequence of bead-C-terminus-MRBBAX₁X₂X₃(pT/S/Y)X₄X₅X₆AA-N-terminus. Later, libraries used MRBBI as the linker and the library had a sequence of bead-C-terminus-MRBBIX₁X₂X₃X₄(pT/S/Y)X₅X₆AA-N-terminus. The latter one focused more on C-terminus sequence after pT/S/Y under the assumption that FHA domains mainly select C-terminal sequence around the +3 position. While random amino acids were in each position, X₆ should not be considered as fully random because of the scale of synthesis.

**Library screening condition**

The library screening condition is controlled by three factors: the amount of protein added, the wash time in each step and the salt concentration. Protein concentrations of 0.1 µM and 1 µM were mainly used. When protein concentration of lower than 0.1 µM was used, very few positive beads were identified. Based on empirical estimation, the washing condition was arbitrarily classified as stringent (around 10 times each step) and less stringent (around 5 times each step). NaCl concentrations of 150 mM (physiological) and 1 M (high salt) were used in different screenings.

During library screening, absence of protein was used as negative control and FHA1 was used as positive control.
Screening result

A total amount of 1 g of beads was used for different batches of library screening. All library screenings failed to give good consensus sequences for Ki67 FHA while the positive control of FHA1 consistently showed a preference of Asp at +3 position. Under less stringent condition (1 µM of protein, less wash and low salt), the positive beads tended to have sequences with arginine(s) at various position no matter which library was used (Table 3.1). The presence of a large number of arginines at various positions was also observed for library screenings of other FHA domains and was proven to be non-specific. Under the more stringent condition (0.1 µM of protein, extended wash and low salt), positive beads tend to have sequences with tyrosine(s) at various positions but focused on the positions around pT/pS (Table 3.2). Like the preference of Arg, the preference of Tyr was deemed as non-specific. Because of the frequent presence of Tyr around pT/pS, the possible preference of pY library was tested. However, pY library screening failed to generate good consensus sequences as well (Table 3.3). Under the high salt condition (1M NaCl), virtually no positive beads were generated, suggesting preference of Arg or Tyr is due to nonspecific ionic interactions.
Table 3.1 Results of pT library screening under less stringent condition. The peptide library has a sequence as: N--terminus-AX₁X₂(pT)X₃X₄X₅X₆IBBRM-COO--C terminus. Less stringent conditions were used. Results showed a nonspecific preference of R.
| A | 0 | 0% | 0 | 0% | 0 | 0% | 2 | 7.70% | 1 | 3.80% |
| B | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% |
| C | 1 | 50% | 0 | 0% | 1 | 6.70% | 0 | 0% | 0 | 0% |
| D | 0 | 0% | 1 | 14.30% | 2 | 13.30% | 2 | 7.70% | 0 | 0% |
| E | 0 | 0% | 0 | 0% | 0 | 0% | 1 | 3.80% | 1 | 3.80% |
| F | 0 | 0% | 0 | 0% | 0 | 0% | 16.70% | 1 | 3.80% | 5 | 19.20% |
| G | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 3 | 11.50% |
| H | 0 | 0% | 0 | 0% | 1 | 6.70% | 1 | 3.80% | 5 | 19.20% |
| I | 0 | 0% | 2 | 28.60% | 0 | 0% | 4 | 15.40% | 3 | 11.50% |
| J | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% |
| K | 0 | 0% | 0 | 0% | 0 | 0% | 1 | 3.80% | 0 | 0% |
| L | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% |
| M | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 1 | 3.80% |
| N | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 4 | 15.40% |
| O | 0 | 0% | 0 | 0% | 0 | 0% | 2 | 7.70% | 1 | 3.80% |
| P | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% |
| Q | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 | 0% |
| R | 1 | 50% | 1 | 14.30% | 3 | 20% | 2 | 7.70% | 4 | 15.40% |
| S | 0 | 0% | 0 | 0% | 0 | 0% | 1 | 3.80% | 0 | 0% |
| T | 0 | 0% | 0 | 0% | 1 | 6.70% | 1 | 3.80% | 1 | 3.80% |
| U | 0 | 0% | 0 | 0% | 2 | 13.30% | 1 | 3.80% | 1 | 3.80% |
| V | 0 | 0% | 0 | 0% | 1 | 14.30% | 0 | 0% | 0 | 0% |
| W | 0 | 0% | 1 | 14.30% | 0 | 0% | 0 | 0% | 0 | 0% |
| X | 0 | 0% | 2 | 28.60% | 5 | 33.30% | 13 | 50% | 6 | 23.10% |
| Y | 0 | 0% | 2 | 28.60% | 5 | 33.30% | 13 | 50% | 6 | 23.10% |

Table 3.2 Results of pT library screening under more stringent condition. The peptide library has a sequence as: N-terminus-AX₁X₂X₃(pT)X₄X₅X₆ABBRM-COO--C terminus. Stringent conditions were used.
Table 3.3 Results of pY Library screening under less stringent condition. The peptide library has a sequence as: N-terminus-AX₁X₂(pY)X₃X₄X₅X₆IBBRM-COO--C terminus. Less stringent conditions were used.
3.2 NMR titration experiments showed that Ki67FHA does not bind to short phosphopeptides with different +3 residues

Based on the library screening results of FHA domains in Rad53 and Chk2, the FHA domain seems to prefer to recognize the +3 position in the peptide while other residues around pT have limited effects on the binding. Hence, peptides with different residues at the +3 position were chosen to test binding using NMR HSQC experiments. Peak shifting can be observed even for weak binding provided the ligand concentration is high enough. This is the ideal choice to test the weak binding between protein and peptide ligands even when the binding is not optimal. Weak binding between FHA1 domain and pTyr and pThr was observed by NMR HSQC when the ligand concentration was high.

12 peptides from different sources were tested using $^{15}$N Ki67FHA in NMR HSQC experiments. These include peptides with aromatic, hydrophobic, acidic and basic residues at the +3 position as well as phosphopeptides from binding partner proteins (Table 3.4). None of the peptides caused any visible chemical shift changes in the HSQC spectrum (see Appendix A for all titration spectra).
<table>
<thead>
<tr>
<th>Peptide name</th>
<th>Sequence</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>PepT 3_3</td>
<td>ARTL(pT)ARY</td>
<td>Y at +3</td>
</tr>
<tr>
<td>PEPRadT2_11</td>
<td>KEKD(pT)FNID</td>
<td>I at +3, no R</td>
</tr>
<tr>
<td>P53_peptide</td>
<td>SGSYSQE(pT)FSDLG</td>
<td>D at +3</td>
</tr>
<tr>
<td>PepT2_1</td>
<td>AMHS(pT)HRR</td>
<td>R at +3</td>
</tr>
<tr>
<td>PEP 1</td>
<td>SLEV(pT)EADATFVQ</td>
<td>D at +3</td>
</tr>
<tr>
<td>Hklp2_1021</td>
<td>NCKYN(pS)ALVDR</td>
<td>Binding partner</td>
</tr>
<tr>
<td>Hklp2_1088</td>
<td>SGLLQ(pS)AQEEL</td>
<td>Binding partner</td>
</tr>
<tr>
<td>Hklp2_1134</td>
<td>EHVM(pS)AEDP</td>
<td>Binding partner</td>
</tr>
<tr>
<td>Hklp2_1144</td>
<td>PQSPK(pT)PPHFQ</td>
<td>Binding partner</td>
</tr>
<tr>
<td>Hklp2_1169</td>
<td>EDGRA(pS)KTSLE</td>
<td>Binding partner</td>
</tr>
<tr>
<td>hNIFK_234</td>
<td>VDSQGP(pT)PVCTPT</td>
<td>Binding partner</td>
</tr>
<tr>
<td>hNIFK_238</td>
<td>GTPFVC(pT)PTFLER</td>
<td>Binding partner</td>
</tr>
</tbody>
</table>

Table 3.4 Peptide titration of Ki67FHA
3.3 Summary

Lack of a consensus sequence from library screening strongly suggests that Ki67FHA cannot recognize short peptides with conserved residues around pT/pS/pY. This is in agreement with the yeast 2-hybrid experiment which shows that the binding motif of Ki67FHA cannot be mapped to a smaller fragment in either Hklp2 or hNIFK. Apparently, this binding pattern is very different from that of the FHA1 domain of Rad53, in which case a conserved Asp at +3 position was observed in more than 90% of all library screening results and tight binding can be observed by NMR. Absence of any chemical shift changes in NMR HSQC experiments with different phosphothreonine and phosphoserine peptides further supported the library screening results. The fact that five 11-amino acid phosphopeptides from Hklp2 and the two 13-amino acid phosphopeptides from hNIFK also failed to bind to Ki67FHA suggests a tertiary structure requirement and/or size requirement for Ki67FHA binding. It is possible that the binding of Ki67FHA requires a special conformation which can only be well-formed by a larger fragment of ligand protein. Small peptides, which have relative freedom in conformational change, could require a relatively high energy for this special conformation. Larger ligand fragments could also provide additional contacts which contribute to the binding. Further experiments are needed to explore the Ki67 FHA binding mechanism and why tertiary structure is required.
CHAPTER 4

PHOSPHO-INDEPENDENT BINDING

4.1 Introduction

Ki67 FHA domain has been shown to bind to non-phosphorylated target proteins (Hklp2 and hNIFK) weakly (Sueishi, Takagi et al. 2000; Takagi, Sueishi et al. 2001). Non-phosphorylated hNIFK can be pulled down by GST-Ki67 FHA domain though phosphorylated hNIFK can be pulled down much more efficiently. The phosphorylation of hNIFK is mitotic cell dependent; however, a small uncharacterized population of hNIFK also interacts with Ki67FHA domain in other cell cycle phases. This minor population of hNIFK is likely in the non-phosphorylated form. Though the biological significance is elusive, phoso-independent binding of Ki67FHA to non-phosphorylated hNIFK serves as an intermediate step for studying the phosphorylation-dependent tight binding due to the ease in sample preparation. It also serves as a necessary control experiment for later study.
4.2 Theoretical basis: HSQC NMR spectrum and binding

HSQC NMR is a convenient method to study the binding between a labeled protein and its ligand. A few key points from Biomolecular NMR Spectroscopy (Evans, 1995) and NMR of Macromolecules (Roberts, 1993) are included here for easy analysis of data.

The simplest binding situation is a second-order exchange reaction:

\[
E + L \xrightleftharpoons[k_{-1}]{k_{+1}} EL
\]  \[1\]

Where E is enzyme (in our case FHA domain), L is the ligand (the binding partner of FHA domain).

The dissociation constant,

\[
K_d = \frac{k_{-1}}{k_{+1}} = \frac{[E][L]}{[EL]} \]  \[2\]

The fraction of ligand,

\[
p_L = \frac{[L]}{[L] + [EL]} \]  \[3\]

The fraction of enzyme bound with ligand,

\[
p_{EL} = \frac{[EL]}{[L] + [EL]} \]  \[4\]

The life time in bound state,

\[
\tau_{EL} = \frac{1}{k_{-1}} \]  \[5\]

The life time for free ligand state,
\[ \tau_L = 1/k_{+1}[E] \] \[ \tau_E = 1/(k_{+1}[L]) = p_E/(k_{-1}p_{EL}) \]

The total life time of second-order exchange,

\[ 1/\tau = 1/\tau_{EL} + 1/\tau_L = k_{-1}(1+p_{EL}/p_L) \]

According to Biomolecular NMR Spectroscopy chapter 1.3 (Evans, 1995), the NMR timescale is defined by the nuclear precession frequency, and kinetic processes or molecular motions that occur on this time scale are reflected directly in the chemical shift (\( \delta \)), scalar coupling (\( J \)), and relaxation parameters (\( T_1 \) and \( T_2 \)). See Figure 4.1. If the measured variable is chemical shift, with exchange occurring between two environments characterized by shifts \( \delta_A \) and \( \delta_B \), slow exchange occurs if the exchange rate \( k << |\delta_A - \delta_B| \), intermediate exchange if \( k \approx |\delta_A - \delta_B| \), and fast exchange if \( k >> |\delta_A - \delta_B| \). Whether a system is in the fast, intermediate, or slow exchange range with respect to chemical shift can be judged from the appearance of the spectrum.
Figure 4.1 NMR time scale. Reproduced from (Roberts 1993).

<table>
<thead>
<tr>
<th>Time scale</th>
<th>Slow</th>
<th>Intermediate</th>
<th>Fast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical shifts, $\delta^a$</td>
<td>$k \ll \delta_A - \delta_B$</td>
<td>$k = \delta_A - \delta_B$</td>
<td>$k \gg \delta_A - \delta_B$</td>
</tr>
<tr>
<td>Coupling constant, $J^b$</td>
<td>$k \ll J_A - J_B$</td>
<td>$k = J_A - J_B$</td>
<td>$k \gg J_A - J_B$</td>
</tr>
<tr>
<td>$T_2$ relaxation $^c$</td>
<td>$k \ll \frac{1}{T_{2,A}} - \frac{1}{T_{2,B}}$</td>
<td>$k = \frac{1}{T_{2,A}} - \frac{1}{T_{2,B}}$</td>
<td>$k \gg \frac{1}{T_{2,A}} - \frac{1}{T_{2,B}}$</td>
</tr>
</tbody>
</table>

$^a (\delta_A - \delta_B)$ is typically in the order of hundreds of hertz.

$^b J$ is typically in the order of 1–10 Hz.

$^c \frac{1}{T_{2,A}} - \frac{1}{T_{2,B}}$ is typically in the order of 1–20 Hz for protons; the linewidth at half-height, $\Delta = 1/\pi T_2$. 
Figure 4.2 Change in chemical shifts and linewidths in the presence of chemical exchange between two equally populated environments. Reproduced from (Roberts 1993).
When exchange between the two states is slow, separate resonance lines are observed for each state (Figure 4.2). The exchange rate can be readily measured from the line widths of the resonances. These yield the apparent spin-spin relaxation rates, $1/T_{2i,\text{obs}}$, which are given by the following equations:

\[
1/T_{2L,\text{obs}}=1/T_{2L}+1/\tau_L=1/T_{2L}+k_{-1}p_{EL}/p_L \quad \cdots \cdots \cdot [9]
\]

\[
1/T_{2E,\text{obs}}=1/T_{2E}+1/\tau_E=1/T_{2E}+k_{-1}p_{EL}/p_E \quad \cdots \cdots \cdot [10]
\]

and

\[
1/T_{2EL,\text{obs}}=1/\tau_{EL}=1/T_{2EL}+k_{-1} \quad \cdots \cdots \cdot [11]
\]

For fast exchange a single resonance line is observed, with a chemical shift that is a weighted average of the individual species:

\[
\delta_{\text{obs}}=\alpha \delta_A+(1-\alpha)\delta_B \quad \cdots \cdots \cdot [12]
\]

in the case of enzyme-ligand binding,

\[
\delta_{\text{obs}}=\delta_L p_L+\delta_{EL} p_{EL} \quad \cdots \cdots \cdot [13]
\]

Where $\alpha$ is the fractional population of species A.

The $K_d$ determination for very fast exchange according to [2], [4], [13] can use the following equation:

\[
\delta_{\text{obs}}-\delta_E=(\delta_{EL}-\delta_E)((E_T+L_T+K_d)-((E_T+L_T+K_d)^2-4E_T L_T))^{1/2}/2E_T \quad \cdots \cdots [14]
\]
Enzyme and ligand concentrations ($E_T$ and $L_T$) are known. $\delta_L$ can be measured accurately. This equation can be used to fit the data obtained from HSQC to obtain $K_d$.

4.3 HSQC titration experiments

$^{15}$N Ki67FHA and GB1-hNIFK(226-269) were produced as described in Materials and Methods. hNIFK(226-269) was produced as a fusion protein with a 56 amino acid GB1 domain since the GB1 domain has been shown to be a solubility enhancement tag (SET) for small proteins/domains (Zhou, Lugovskoy et al. 2001) and can greatly increase protein stability. GB1-hNIFK(226-269) has a high yield and can be concentrated to greater than 3 mM without any precipitation. GB1-hNIFK(226-269) was eluted as a monomer on S100 and the purity of GB1-hNIFK(226-269) was estimated to be more than 90% by SDS-PAGE followed by silver staining. Ligand binding experiment was performed by recording a series of 2-D $^{15}$N-HSQC spectra on uniformly $^{15}$N labeled protein samples (Ki67FHA) with different concentrations of ligands (GB1-hNIFK(226-269)).

Titration HSQC spectrum showed a typical fast/intermediate exchange weak binding

In the HSQC titration spectrum of $^{15}$N Ki67FHA vs. GB1-hNIFK(226-269) (Figure 4.3), more than 50 peaks were shifted or missing altogether. All the shifted peaks showed gradual shifts, the shift distances of which are positively correlated to the amount of
ligand added. Most of the shifted peaks also showed different degrees of signal weakening which was the result of linewidth broadening. The disappeared or weakened peaks are an indication of intermediate binding or the existence of multiple states. There could be several different exchanging conformations upon the interaction with ligand in that particular region of Ki67FHA, which cause different chemical shifts of nearby residues. The existence of multiple states at the binding site eventually weakened the signals out.
Figure 4.3 HSQC spectrum of $^{15}$N Ki67 FHA binding to GB1-hNIFK(226-269)

Black: $^{15}$N Ki67FHA
Red: $^{15}$N Ki67FHA: GB1-hNIFK(226-269)=1:2
**HSQC titration control experiments**

Control experiments were designed to test FHA domain specificity of the phospho-independent binding and the effect of the GB1 domain. The control experiment design and results are listed in Table 4.1. Titration spectra are listed in Appendix B.

The results showed:

1. This binding is specific to Ki67FHA. There is no binding between FHA1 and GB1-hNIFK or FHA2 and GB1-hNIFK.

2. The hNIFK fragment may not be well structured since the NMR spectrum for GB1-hNIFK only displays well-dispersed peaks for the GB1 domain.

3. The GB1 fusion does not interfere with the binding either directly or indirectly since neither GB1 domain caused any chemical shifts for Ki67FHA nor Ki67FHA caused any chemical shifts in the GB1 domain when GB1 is in fused form with hNIFK(226-269).
<table>
<thead>
<tr>
<th>$^{15}$N Protein</th>
<th>Ligand</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ki67FHA</td>
<td>GB1-hNIFK*</td>
<td>More than 30 peaks shift more than 0.1 ppm or disappeared</td>
</tr>
<tr>
<td>Rad53 FHA2</td>
<td>GB1-hNIFK*</td>
<td>No peak shifts</td>
</tr>
<tr>
<td>Rad53 FHA1</td>
<td>GB1-hNIFK*</td>
<td>No peak shifts</td>
</tr>
<tr>
<td>GB1-hNIFK*</td>
<td>FHA2</td>
<td>No peak shifts</td>
</tr>
<tr>
<td>GB1-hNIFK*</td>
<td>Ki67FHA</td>
<td>No peak shifts for GB1 domain</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peak shifts for hNIFK* part observed</td>
</tr>
<tr>
<td>GB1-hNIFK*</td>
<td>None</td>
<td>Use as reference</td>
</tr>
<tr>
<td>Ki67FHA</td>
<td>GB1</td>
<td>No peak shifts</td>
</tr>
</tbody>
</table>

Table 4.1 NMR titration experiments for phospho-independent binding

*: hNIFK fragment 226-269
4.4 Roles of T234 and T238 in the target protein in phospho-independent binding

Two threonines (T234 and T238) in hNIFK have been identified to be important for tight binding in vivo (Takagi, Sueishi et al. 2001). T234A and T238A mutants were made to test whether these residues play an important role in phospho-independent binding. HSQC titration spectra for GB1-hNIFK(226-269)T234A and GB1-hNIFK(226-269)T238A showed little difference from wild type as shown in Figure 4.4 and Figure 4.5, indicating that these two threonines do not play an important role in phospho-independent binding. Glutamic acid is a commonly used analog for phosphothreonine or phosphoserine, for example, it was successfully used to study the role of autophosphorylation of protein kinase C (Feng and Hannun 1998). GB1-hNIFK(226-269) T234E and T238E mutants were constructed to investigate whether the phosphorylation in Ki67FHA target protein can be mimicked by these mutations. HSQC titration experiments were carried out under the same conditions as that for wild type. This result showed that phosphorylated Thr cannot be mimicked by Glu though the latter also provides negative charges and has a similar structure. The HSQC titration spectra (Figure 4.6, Figure 4.7) for these two mutants were basically the same as that for wild type GB1-hNIFK(226-269).
Figure 4.4 HSQC titration of GB1-hNIFK(226-269)T234A
Black: free protein
Red: protein: peptide=1:0.5
Green: protein: peptide=1:1
Blue: protein: peptide=1:4
Figure 4.5 HSQC titration of GB1-hNIFK(226-269)T238A
Black: free protein
Red: protein: peptide=1:0.25
Green: protein: peptide=1:0.5
Blue: protein: peptide=1:1
Turquoise: protein: peptide=1:2
Purple: protein: peptide=1:4
Figure 4.6 HSQC titrations of GB1-hNIFK(226-269)234E
Black: free protein
Red: protein: peptide=1:0.25
Green: protein: peptide=1:0.5
Blue: protein: peptide=1:1
Turquoise: protein: peptide=1:2
Purple: protein: peptide=1:4
Figure 4.7 HSQC titration of GB1hNIFK(226-269)238E
Black: free protein
Red: protein: peptide=1:0.25
Green: protein: peptide=1:0.5
Blue: protein: peptide=1:1
Turquoise: protein: peptide=1:2
Purple: protein: peptide=1:4
4.5 $K_d$ determination using SPR

The dissociation constant ($K_d$) of Ki67FHA and GB1-hNIFK(226-269) was determined using surface plasmon resonance (SPR) as described in Materials and Methods. Biotinylated BSA was used as the reference channel. GB1-hNIFK(226-269) was immobilized on the BIAcore chip as the stationary phase. Ki67FHA was used as mobile phase to flow over the BIAcore chip. Response unit change of each concentration was collected at platform region of the sensorgrams (Figure 4.8). After fitting response unit change vs. concentration to a hyperbolic curve (Figure 4.9), $K_d$ was determined to be 108.6 µM. Ki67FHA was immobilized on a separate channel to test whether Ki67FHA forms oligomers. Ki67FHA showed no tendency for oligomerization as there was no observable response unit change. Therefore this binding constant reflects the binding between Ki67FHA and GB1-hNIFK(226-269) only.

Similarly, the dissociation constants for T234A, T238A, T234E and T238E were determined to be 131 µM, 100 µM, 130 µM and 115 µM respectively (detailed curves are listed in Appendix C), which are comparable to 108.6 µM observed with wild type GB1-hNIFK(226-269). This further supports that T234 and T238 do not play an important role in the phospho-independent binding.
Figure 4.8 Sensorgram overlay for the binding of Ki67FHA and GB1-hNIFK(226-269). Data points were taken at the platform region (shown by the arrow in the curve) of the curves and fit to a hyperbolic curve to determine $K_d$ (Figure 4.9). Background was subtracted.
Figure 4.9 $K_d$ determination for Ki67 FHA and GB1-hNIFK(226-269) using SPR

$R = R_{\text{max}} \frac{c}{K_d + c}$

$R_{\text{max}} = 531.9187$ RU

$K_d = 108.5887$ uM

correlation coefficient = 0.99679787
4.6 K<sub>d</sub> determination using chemical shift changes upon binding

As a verification of the K<sub>d</sub> determined using SPR, the dissociation constant of the phospho-independent binding between Ki67FHA and GB1-hNIFK(226-269) was alternatively determined using chemical shift changes observed in HSQC spectra. Since the titration spectra of wild type GB1-hNIFK and mutants basically overlap each other but the titration of GB1-hNIFK(226-269)T234E has more data points, it is used for the determination. The chemical shift changes of E34 and F95 were used because of their clear signals in HSQC spectra. Ki67FHA concentration, ligand concentration and measured chemical shift changes were used to fit the 3-variable equation 14 (Figure 4.10) to get the dissociation constant (K<sub>d</sub>) and maximal chemical shift change (δ<sub>obs</sub>−δ<sub>E</sub>]<sub>max</sub>). The binding constants were determined to be 53.8 and 47.0 µM which are a little lower than but in the same order of magnitude to that determined using SPR.
Figure 4.10 $K_d$ determination using chemical shift changes. Total Ki67FHA concentration ($E_T$, free and bound), total GB1-hNIFK concentration ($L_T$) and chemical shift changes ($\delta_{\text{obs}} - \delta_E$) were used to fit equation: 

$$
\delta_{\text{obs}} - \delta_E = (\delta_{\text{EL}} - \delta_E)((E_T + L_T + K_d) - ((E_T + L_T + K_d)^2 - 4E_T L_T))^{1/2}/2E_T.
$$

Upper: chemical shift changes of E34 were used. Lower: chemical shift changes of F95 were used.
\[ \Delta \delta_{\text{max}} = 0.2391 \text{ ppm} \]
\[ K_d = 0.0528 \text{ mM} \]
\[ R = 0.99503959 \]

\[ \Delta \delta_{\text{max}} = 0.2391 \text{ ppm} \]
\[ K_d = 0.0470 \text{ mM} \]
\[ R = 0.9967 \]
4.7 Summary

Consistent with the previously reported pull-down assays (Takagi, Sueishi et al. 2001), HSQC titration and SPR experiments verified that Ki67FHA can bind to non-phosphorylated form of hNIFK weakly. HSQC titration experiments showed that this binding involves a large number of the residues (more than 50 residues are perturbed) in Ki67FHA and does not involve T234 and T238, the candidate phosphorylation sites, in target protein hNIFK. The data here serve as a basis to characterize phospho-dependent binding. A detailed discussion of the structural basis of phospho-dependent and phospho-independent binding is included in Chapter 6.
5.1 Phosphopeptides used for phospho-dependent binding

Although the binding between Ki67FHA and non-phosphorylated hNIFK(226-269) is weak, it supports the idea that Ki67FHA binds to hNIFK through a longer sequence. Phospho-dependent binding of Ki67FHA to hNIFK in this study was based on this sequence. Among all possible threonine and serine residues in this sequence, only T234 and T238 were reported to be required for tight binding \textit{in vivo} (Takagi, Sueishi et al. 2001). Three phosphopeptides corresponding to hNIFK(226-269) with phosphorylation at T234, T238, and both T234 and T238 (Figure 5.1) were synthesized to test the phosphorylation effect on binding.
Peptide-A (234_238dp)
KTVDS QGP(pT)P VC(pT)PT FLERR KSQVA ELNDD DKDDE IVFKQ PISCM*

Peptide- B (238p)
KTVDS QGP(T)P VC(pT)PT FLERR KSQVA ELNDD DKDDE IVFKQ PISCM*

Peptide-C (234p)
KTVDS QGP(pT)P VC(T)PT FLERR KSQVA ELNDD DKDDE IVFKQ PISCM*

P44
KTVDS QGP(pT)P VC(T)PT FLERR KSQVA ELNDD DKDDE IVFKQ PISC
M* Homoserine lactone

Figure 5.1 The sequences of phosphopeptides. The first 3 peptides were synthesized using an in-house apparatus as described in materials and methods (Thanks to Dr. Yong Ju). P44 was synthesized by Genemed Synthesis Inc. For the peptides synthesized using solid phase synthesis, an extra homoserine lactone was appended to the C-terminus as the result of CNBr cleavage.
5.2 Identification of T234 as the phosphorylation site

The peptides (234p, 238p and 234_238dp) were immobilized on the same BIAcore sensorchip SA on different channels. BSA was immobilized on a separate channel as a control. Ki67FHA at different concentrations was used as the flowing phase. Sensorgrams were monitored (Figure 5.2 A, B and C). Response unit changes were collected and plotted against Ki67FHA concentrations (Figure 5.3A). The binding curve clearly showed that there are two distinct types of binding (one tight binding and one weak binding) for 234p and 234_238dp while there is only one weak binding for 238p. The difference between binding curves is obvious at low concentrations (Figure 5.3B). In this range, the binding curves for 234p and 234_238dp display a rapid burst of response unit change forming a hyperbola-like shape while the binding curve for 238p only shows a slow, steady increase of response unit change which is the prelude of weak binding. The \( K_d \)s of 234p and 234_238dp were determined by fitting the data to a double rectangular hyperbolic curve. The \( K_d \) of 238p was determined by fitting the data to a single rectangular hyperbolic curve.

The tight binding of 234p and 234_238dp (\( K_d = 1.6 \, \mu M \) and 1.4 \( \mu M \) respectively) is more than 60 times stronger than phospho-independent binding (\( K_d = 108.6 \, \mu M \), see chapter 4.5). The results suggest that phosphorylation of T234 is essential for the tight binding. Weak binding around 100-200 \( \mu M \) was observed for all three peptides. This suggests that the weak binding is not dependent on site-specific phosphorylation. Apparently, the weak binding constants are in the same range of phospho-independent binding.
After the identification of T234 as the phosphorylation site, a peptide of the same sequence was obtained through a commercial source (Genemed Synthesis Inc.) for a better quality. This new phosphopeptide, named P44, was essentially the same as 234p except it does not contain a homoserine lactone, a byproduct of solid phase synthesis, at the end of sequence. P44 showed similar binding property to 234p. The dissociation constants ($K_d$) were determined to be 2.1 µM and 91.2 µM for the tight binding and the weak binding respectively (Figure 5.4). The binding curves of both P44 and of 234p showed clearly that there are two different types of binding in phosphorylated hNIFK: a tight-binding and a weak-binding. The percentage of each ingredient can be estimated by the maximal response unit changes. In the SPR sensorgram, the response unit change is proportional to the amount of protein immobilized by the ligand. The maximal response unit change is proportional to the total amount of bound ligand. In a one-to-one binding, given that the tight binding competes with weak binding, which will be discussed later in this chapter, the relative amount of peptides that can mediate tight binding can be calculated. For P44, the maximal response unit changes for tight binding and weak binding are 122 and 732 respectively. Therefore, around 14% of the total peptide is tight-binding ingredient.
Figure 5.2 Sensorgram overlay for the binding of Ki67FHA and synthetic phosphopeptides. (A) 234p. (B) 238p. (C) 234_238dp.
Figure 5.3 Binding constant determinations of 234p, 238p and 234-238dp using SPR. (A) curves at full range of concentrations (B) curves at low concentrations. For 234p and 234_238dp, the curves were fit to a double rectangular hyperbola ($\Delta RU = RU_{max1} \times c / (K_{d1} + c) + RU_{max2} \times c / (K_{d2} + c)$). For 238p, the curve was fit to a single rectangular hyperbola ($\Delta RU = RU_{max} \times c / (K_d + c)$) where c is the concentration in µM. For 234p, R=0.99984252, $RU_{max1}=188.8$, $K_{d1}=1.6$ µM, $RU_{max2}=814.9$, $K_{d2}=146.2$ µM. For 234_238dp, R=0.99993463, $RU_{max1}=68.3$, $K_{d1}=1.4$ µM, $RU_{max2}=563.2$, $K_{d2}=135.1$ µM. For 238p, R=0.99967888, $RU_{max}=989.1$, $K_d=101.2$ µM.
Figure 5.4 The binding constant determination of P44
5.3 Purity analysis of P44

The purity of P44 was analyzed using mass spectrometry (Figure 5.5) and amino acid analysis (Table 5.1). From the mass spectrometry, the full-length P44 (MW 5032 Da) is the component of highest percentage. However, there were also other peaks (4932 Da and 5119 Da) around 5032 Da whose identities cannot be determined. Amino acid analysis showed that the calculated amino acid content roughly matches the peptide sequence though there are some variations. Amino acid analysis has an average error range of around 15%. The error ranges for some residues are much larger than others for different reasons: Trp and phospho amino acids (destroyed), Cys, Ser and Thr (partially destroyed), hydrophobic residues such as Val, Leu and Ile (hard to hydrolyze). Therefore, amino acid analysis data could not be used to accurately determine the purity of the peptide, however it provides information on the percentage of peptide content. 72% (3.6 µg) of total injection amount (5 µg) was detected by amino acid analysis, indicating around 72% of the lyophilized powder is the peptide.

The phosphorus content of P44 was analyzed using $^{31}$P NMR (Figure 5.7A). Two references were used: 0.1 mM phosphate ($\delta=1.847$) and 0.2 mM glucose-1-phosphate ($\delta=2.092$). The concentration of peptide with a phosphate group was calculated by the peak integration of P44 ($\delta=3.035, 3.087, 3.539$ and 3.661 respectively) divided by that of the references and multiplied by concentration of references. This calculated amount of P44 (with a phosphate group) was compared with total P44 used in the sample (3 mM). Result showed that 60% of the phospho-peptide used in the sample has a phosphate group.
There was no peak for P44 with missing phosphate on the mass spectrometry. The peak at 4932 Da on the mass spectrometry may be a combination of peptides with a missing phosphate and a missing amino acid.
Figure 5.5 Mass spectrometry of P44. Upper: P44 upon arrival. Lower: P44 after affinity purification and HPLC separation.
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>MW</th>
<th>Picomole</th>
<th>microgram</th>
<th>mole %</th>
<th>calc AA</th>
<th>AA count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp+Asn</td>
<td>114.62</td>
<td>5901.79</td>
<td>0.6765</td>
<td>18.91</td>
<td>7.7531</td>
<td>7</td>
</tr>
<tr>
<td>Glu+Gln</td>
<td>128.64</td>
<td>4842.7</td>
<td>0.623</td>
<td>15.52</td>
<td>6.3632</td>
<td>6</td>
</tr>
<tr>
<td>Trp</td>
<td>204.23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ser</td>
<td>87.09</td>
<td>2024.76</td>
<td>0.1763</td>
<td>6.49</td>
<td>2.6609</td>
<td>3</td>
</tr>
<tr>
<td>Gly</td>
<td>57.05</td>
<td>583.3</td>
<td>0.0333</td>
<td>1.87</td>
<td>0.7667</td>
<td>1</td>
</tr>
<tr>
<td>His</td>
<td>137.16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Arg</td>
<td>156.2</td>
<td>1459.68</td>
<td>0.228</td>
<td>4.68</td>
<td>1.9188</td>
<td>2</td>
</tr>
<tr>
<td>Thr</td>
<td>101.12</td>
<td>1985.85</td>
<td>0.2008</td>
<td>6.36</td>
<td>2.6076</td>
<td>3</td>
</tr>
<tr>
<td>Ala</td>
<td>71.09</td>
<td>780.85</td>
<td>0.0555</td>
<td>2.5</td>
<td>1.025</td>
<td>1</td>
</tr>
<tr>
<td>Pro</td>
<td>97.13</td>
<td>2770.57</td>
<td>0.2691</td>
<td>8.88</td>
<td>3.6408</td>
<td>4</td>
</tr>
<tr>
<td>Tyr</td>
<td>163.19</td>
<td>9.68</td>
<td>0.0016</td>
<td>0.03</td>
<td>0.0123</td>
<td>0</td>
</tr>
<tr>
<td>Val</td>
<td>99.15</td>
<td>2313.86</td>
<td>0.2294</td>
<td>7.41</td>
<td>3.0381</td>
<td>4</td>
</tr>
<tr>
<td>Met</td>
<td>131.21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cys2</td>
<td>204.28</td>
<td>297.31</td>
<td>0.0607</td>
<td>0.95</td>
<td>0.3895</td>
<td>2</td>
</tr>
<tr>
<td>Ile</td>
<td>113.18</td>
<td>1336.52</td>
<td>0.1513</td>
<td>4.28</td>
<td>1.7548</td>
<td>2</td>
</tr>
<tr>
<td>Leu</td>
<td>113.18</td>
<td>1580.28</td>
<td>0.1789</td>
<td>5.06</td>
<td>2.0746</td>
<td>2</td>
</tr>
<tr>
<td>Phe</td>
<td>147.19</td>
<td>1704.86</td>
<td>0.2509</td>
<td>5.46</td>
<td>2.2386</td>
<td>2</td>
</tr>
<tr>
<td>Lys</td>
<td>128.17</td>
<td>3613.82</td>
<td>0.4632</td>
<td>11.58</td>
<td>4.7478</td>
<td>4</td>
</tr>
</tbody>
</table>

Total 31205.83 3.6

Table 5.1 Amino acid analysis result of P44. 5 µg of P44 was injected for analysis. Cysteines and phosphothreonines were destroyed.
5.4 The tight-binding ingredient competes with the weak-binding ingredient

The relationship of tight binding and weak binding was examined using HSQC titration experiments. HSQC Titration using protein to peptide ratio less than 1:1 results in a spectrum similar to that of the non-phosphorylated peptide, GB1-hNIFK. The tight binding signals were barely visible due to the low percentage of tight binding ingredient. Titration using an increased amount of peptide results in strengthened tight-binding signals. Figure 5.6A showed a titration spectrum (protein: peptide=1: 1.75) of two sets of signals (tight binding and weak binding). The tight binding signal is still much weaker than the weak binding signal. An expanded region of Figure 5.6 was shown in Figure 5.7. As more and more P44 was added, the weak binding signal became weaker and weaker until at last all the weak binding signals were gone (protein: peptide=1: 7) (Figure 5.6B).

5.5 $^{31}$P titration experiment

$^{31}$P NMR of P44 showed four major peaks (Figure 5.8A). Presumably these 4 peaks represent 4 major components (either different conformations or impurities in the peptide). The relationship of peaks and binding was examined by $^{31}$P titration experiment (Figure 5.8). Adding Ki67FHA to P44 caused Peak 2 and Peak 4 shifted to a new position at ~ 4.3 ppm, while Peak 1 and Peak 3 did not shift. The new peak is much broader than either Peak 2 or Peak 4, suggesting it is in bound form. Peak broadening was also observed for Peak 1 and Peak 3 when protein to peptide ratio was high, probably
due to the weak binding between Ki67FHA and P44 as discussed in section 5.4. At the end of titration, when Peak 2 and Peak 4 completely disappeared, the new peak constituted 30% of total integration area, suggesting around one-third of the total P44 with phosphate group can mediate tight binding. Since 60% of P44 is phosphorylated, this result indicates around 18% of total peptide can mediate tight binding, similar to that determined by SPR experiment (14%).
Figure 5.6 Competition of tight-binding and weak-binding ingredients. (A) protein: peptide=1:1.75. (B) protein: peptide=1:7
Figure 5.7  An expanded region of Figure 5.6. t: tight binding peaks. w: weak-binding peaks. Notice that the weak-binding peaks shift the same way as those in GB1-hNIFK titration spectrum.
Figure 5.8 $^{31}$P NMR titration of Ki67FHA vs. P44. (A) P44 only (B) protein: peptide=1:3.4 (C) protein: peptide=1:1.7 (D) protein: peptide=1.2:1. Two reference peaks were used: phosphate peak at 1.847 and glucose-1-phosphate peak at 2.092.
5.6 Affinity competition purification of P44

Since the tight-binding ingredient can effectively compete off the weak-binding ingredient in P44, affinity purification was applied to get the tight-binding complex. Basically, phosphopeptide P44 was pretreated with 20 mM DTT to make sure that its cysteines were in reduced form. $^{15}$N labeled Ki67FHA protein was mixed with several fold amount of P44, adjusted to pH 7.5 with 1 M HEPES and incubated 1 hour at 4 °C. The mixture was loaded onto a S100 column and eluted with S100 buffer (5 mM HEPES pH 7.5, 150 mM NaCl, 2 mM DTT, and 1 mM EDTA). Fractions were collected and examined by 20% SDS PAGE. The bound portion of peptide was eluted early together with Ki67FHA in a complex. The composition of the complex was examined NMR HSQC titration experiments. When 3-fold of peptide was used, the HSQC of Ki67FHA-P44 complex showed a mixture of tight binding and weak binding. As shown in Figure 6.9, the peaks of tight binding did not shift with changed concentrations of protein while the peaks of weak binding shifted. When 7 fold of peptide was used, the HSQC spectrum of Ki67FHA-P44 complex showed a clear set of peaks corresponding to the tight binding one (Figure 5.10). This complex was later used to map out the tight binding surface (Chapter 6). The HSQC experiment here and a 3D edited-NOESY experiment using the same affinity purified complex clearly identified most of the residues involved in the tight binding.

The tight binding peptide in the complex can be separated using an HPLC C8 column by a linear gradient of acetonitrile. Mass spectrometry showed that the purified peptide contains full-length P44 (figure 5.5).
Figure 5.9 Partial affinity competition purification using 3 times extra amount of P44. Similar to the competition experiment, the spectrum showed two sets of peaks.
Figure 5.10 The tight binding HSQC spectrum after affinity competition purification. 7 times of extra amount of P44 were used. All the weak-binding peaks disappeared in this spectrum, indicating the affinity competition purification has successfully removed the weak binding ingredient.
5.7 Discussion

Peptide purity and results in this study

Synthetic peptides of more than 40 amino acid residues often have a purity problem because of the accumulated synthetic error of each step. The major impurity possibly comes from peptides missing one or a few residues. The impure peptide may have similar molecular weight, charge and hydrophobicity and is extremely hard to remove by HPLC. However, the full-length peptide is still the ingredient of highest percentage as shown by mass spectrometry. This ingredient can be detected by the SPR method. Because of the purity of the phosphopeptide and limitations of the SPR method (solid state, spatial effect, etc.), the $K_d$ determined for the phosphopeptide here may not completely reflect the true binding affinity. However, it should reflect a maximal value for the true $K_d$. The affinity purification takes the advantage that the phosphopeptide with correct sequence binds to Ki67FHA most tightly. An excessive amount of phosphopeptide competed for limited amount of Ki67FHA and only the tight-binding ingredient was kept. This procedure may not remove all the impurities such as peptides missing amino acid residues which are not involved in the binding. However, it does not interfere with the mapping of the binding surface, which is composed of amino acids that are involved in the binding. The uniformity of interactions is indicated by the HSQC spectrum of tight binding, which shows one clear set of peaks. Except for three residues, all the shifted peaks were clearly identified. This further supports that all the interactions are uniform. Of course, 100% peptide of full-length is always desirable; however,
limitations in synthetic peptide technique, purification technique and cost of time and money have prevented further efforts on it.

**The conformation of the pThrPro motif in the peptide may contribute to the binding**

The phosphorus in the peptide has 4 major peaks in $^{31}$P NMR spectrum. There are three possibilities: (1) P44 has 4 major conformations. (2) P44 has 3 major incorrect synthetic products in addition to the correct one. (3) The combination of conformation and incorrect synthetic products. Possibility (2) is less likely since an incorrect product usually does not affect the chemical shift of phosphorus seriously unless it involves a direct interaction or a close indirect interaction with the phosphorus and no major incorrect synthetic products of P44 were detected using HPLC that are sufficient to contribute to major peaks. The pT in P44 is followed by a proline, which usually has two conformations. The chemical shift differences between Peak 1 and Peak 3, and Peak 2 and Peak 4 are in the same range of chemical shift difference observed for cis- and trans-conformations of small phosphopeptides containing a pThrPro motif (Schutkowski, Bernhardt et al. 1998). The areas of Peak 3 and Peak 4 are 7% of those of Peak 1 and Peak 2 respectively, which is in the range of cis/trans ratio in a pThrPro motif observed for small peptides (5%-15%). Therefore, Peak 1 and Peak 2 possibly represent the trans conformation while Peak 3 and Peak 4 represent the cis conformation. In the $^{31}$P titration experiment, disappearance of both Peak 2 and Peak 4 was observed. Assuming only one correct conformation can fit into the binding site, this result seems to support a
conversion of cis and trans conformations upon binding. Since Peak 1 and Peak 3 do not disappear upon adding Ki67FHA, it is possible that these peaks represent a cumulative effect of incorrect sequences of P44 which do not bind to Ki67FHA tightly. However, since pThr is preceded by a proline which can have two conformations, the possibility that Peak 1 and Peak 3 are caused by the cis and trans conformations of this proline can not be excluded.

**The weak binding ingredient**

The existence of weak binding ingredient is detected by both SPR and NMR titrations (\(^{15}\text{N} \text{HSQC} \text{titration} \text{and} \ 31\text{P} \text{titration})\). However, the exact reason for the weak binding can not be determined. Since the HSQC titration spectrum of weak binding is similar to the non-phosphorylated hNIFK, it is possible that the weak binding is caused by dephosphorylated form of P44. However, some incorrect synthetic products of P44, which affect only the tight binding, and incorrect conformations of P44 can also contribute to the weak binding and have a similar HSQC spectrum as non-phosphorylated hNIFK. The weak binding detected by the above methods is likely to be a combination of all three.
6.1 Structure of Ki67FHA

Sample preparation
Initially, the portion of the Ki67 gene covering residues 14-88 was cloned and expressed. The construct covers the core domain of Ki67 FHA as determined by computer program with around 12 extra amino acid residues flanking each side. While this construct could be overexpressed very well, the expressed product was extremely unstable and could not be purified. This was in accordance with later published results that a functional FHA domain is longer than the core domain as aligned by computer analysis. A new construct covering Ki67 residues 1-120 was constructed and used for all later studies. Ki67FHA refers to Ki67 residues 1-120 unless otherwise specified.

Ki67 FHA expressed and purified as described in chapter 2.1 has a yield of 13 mg per liter of LB and 5.2 mg per liter of M9 minimal media. The purified Ki67FHA showed one band on SDS PAGE followed by silver staining. The purity was estimated to be more than 90%.
Ki67FHA sample at NMR working concentration (~0.4 mM) was stable for at least 2 weeks at pH 8.4 (5 mM Tris, 2 mM DTT, 1 mM EDTA, pH 8.4) without salt at 293K. pH lower than this value made Ki67FHA unstable. This condition was used for structure determination. Later a lower pH buffer with medium amount of salt (5 mM HEPES pH 7.5, 2 mM DTT, 1 mM EDTA, 150 mM NaCl) was used for NMR experiments including most HSQC titrations. The protein stability is slightly better in medium salt buffer though the pH is lower. The two conditions gave similar spectra while lower pH with medium amount of salt gave better HSQC signals.

**Structure determination**

The structure of Ki67FHA was determined by Dr. In-Ja Byeon, OSU.

**Structure description**

The structure of Ki67 FHA (residues 1-100) is composed of an 11-β stranded sandwich (as shown in Figure 6.1). Residues (101-120) are excluded from the structure since they form a random coil. When compared with other FHA domains, Ki67FHA showed a compact structure since it does not contain insertions of α-helices between the β-sheets, which are away from the putative binding site, as compared with those in the FHA1 and Chk2FHA (Figure 6.2). This compact structure is also demonstrated by the number of residues in the loops, particularly at loop formed by β10 and β11, where there is only one residue connecting the two β-strands.
Figure 6.1 Structure of Ki67 FHA
Figure 6.2 Structure-aided sequence alignment of selected FHA domains. β-strands are marked in red. α-helices are marked in turquoise. Highly conserved residues are indicated by “*”.
<table>
<thead>
<tr>
<th></th>
<th>β1</th>
<th>β2</th>
<th>β3</th>
<th>β4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ki67</td>
<td>4</td>
<td>TRRLVTIKRSGVDPHFPLSL---L----------STCLFGRGIECDIR</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>FHA1</td>
<td>32</td>
<td>TVCRVICTTGQ---PIRDLSDISSLQVLYKEKRSIKKVKVWRFGRNPACDYH</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>FHA2</td>
<td>576</td>
<td>RFLTLKPLPDSSIESLEIQQGV----------NPFFIGRSEDCNCK</td>
<td>612</td>
<td></td>
</tr>
<tr>
<td>CHK2</td>
<td>93</td>
<td>WALWALQDG---ANLCEV---N----------DNYWGRDKSCYEYC</td>
<td>124</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>5</th>
<th>β5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ki67</td>
<td>39</td>
<td>IQ----------LPVVSQHKCKIEIEH----------QEAILHNF</td>
<td>63</td>
</tr>
<tr>
<td>FHA1</td>
<td>78</td>
<td>LGN----------ISRLSNHKFQILLGE----------DGNLLLNDI</td>
<td>104</td>
</tr>
<tr>
<td>FHA2</td>
<td>613</td>
<td>I-E----------DNRLSRVHCFIFKRRHAVGKSMYESPAGLDDIYCHT</td>
<td>652</td>
</tr>
<tr>
<td>CHK2</td>
<td>125</td>
<td>FDEPLLKRDKYRTYSSKHFRIFREVGPKNS----------YIAYIYDH</td>
<td>163</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>β7</th>
<th>β8</th>
<th>β9</th>
<th>β10</th>
<th>β11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ki67</td>
<td>64</td>
<td>SSTNPQVNSVVID---EPVRLKHDGVITII-------DRSFRYENE</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FHA1</td>
<td>105</td>
<td>ST-NGTWNLQKVSKNSSLQSDIETGVGVESD-ILSLVIFIN</td>
<td>148</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FHA2</td>
<td>653</td>
<td>GNVSYLNNNNRMISQGKTLQQDGDEIKIIDWKKFKVIGFKVEIN</td>
<td>697</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHK2</td>
<td>164</td>
<td>SG-NGTFTVNLVGGGKRRPLNNNSIALSLSSR------NKVFVFDDL</td>
<td>204</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6.2 The mapping of binding surface by HSQC NMR

The binding of Ki67FHA to hNIFK(226-269) differs from previous binding studies of FHA domains to small phosphopeptide by involving a large number of residues. These residues form a binding interface and were studied using HSQC titration spectra. HSQC spectra of Ki67FHA titrated with GB1-hNIFK(226-269) were used to map out residues which are involved in the phospho-independent binding. Peaks were identified by the moving trends as increased amount of ligand was added. Maximal chemical shift changes were calculated using equation 14 in Chapter 4. HSQC spectrum of the Ki67FHA-P44 tight-binding complex (purified using affinity method as described in Chapter 5) was used to map out residues that are involved in the phospho-dependent binding. Peaks were assigned by comparing their NOE patterns with those of free Ki67 FHA using 3D $^{15}$N-edited NOESY experiments.

The normalized chemical shift change was calculated from chemical shift changes along $^1$H and $^{15}$N axes using empirical formula $\delta = [(\delta_H)^2 + (0.17*\delta_N)^2]^{1/2}$ (Farmer, Constantine et al. 1996). Residues were classified as having either same-trend chemical shift change (the peaks shift to a similar destination) or different-trend chemical shift change (the peaks obviously shift to different destinations) based on their chemical shift trends on the phospho-independent binding HSQC spectrum and phospho-dependent binding HSQC spectrum (Figure 6.3). The chemical shift changes data were plotted against residue number (Figure 6.4).
Figure 6.3 Diagram of two different kinds of chemical shift changes. Upper: Partial HSQC spectrum from Ki67FHA titrated with affinity purified P44. Lower: Partial HSQC spectrum from Ki67FHA titrated with GB1-hNIFK(226-269). L28 shows same-trend chemical shift change. S45 and H48 show different-trend chemical shift change.
Figure 6.4 Chemical shift perturbations of the Ki67 FHA backbone amide groups upon binding SP44-234p (dark green) and hNIFK44 (light green). The combined $^1$H and $^{15}$N chemical shift differences, calculated using the equation $\delta = [(\delta_\text{H})^2 + (0.17*\delta_\text{N})^2]^{1/2}$ used to normalize the $^1$H and $^{15}$N chemical shifts (Farmer, Constantine et al. 1996; Yuan, Yongkiettrakul et al. 2001), were plotted against residue numbers. Negative bars show residues of different chemical shift change trends in phospho-binding and phospho-independent binding. Residues with chemical shift difference <0.1 ppm are omitted. No data are shown if the residues are prolines (no HN protons) or their assignments could not be made unambiguously.
6.3 The binding surface anatomy

Phospho-dependent and phospho-independent bindings use the same binding surface

Figure 6.4 showed that the shifted/weakened peaks in the phospho-independent binding spectrum are the same peaks as those shifted in the phospho-dependent binding spectrum with a few exceptions. Most peaks that do not shift in phospho-independent binding titration spectrum do not shift in the phospho-dependent binding titration spectrum either. In other words, phospho-dependent binding and phospho-independent binding happen on the same binding surface.

Phospho-binding center and additional contacts

Based on the chemical shift trends, the binding surface can be roughly divided into two parts: phospho-binding center and additional contacts (Figure 6.5).

The residues with different-trend chemical shift changes suggest that their roles are affected by the phosphate group. These residues, mostly located around the β3-β4, β4-β5 and β6-β7 region, form the phospho-binding center. Among all the residues in the phospho-binding center, V43 (1.09 ppm), S45 (0.63 ppm) K46 (0.45 ppm) showed significantly larger chemical shift changes. These residues are likely to be strongly involved in binding. Structural comparison with other FHA domains in complex with
synthetic phosphopeptides suggests that S45 and K46 interact with the phosphate group with the help of N67 while V43 may interact with the residues at +1 ~ +3 positions.

The residues with same-trend chemical shift changes suggest that their roles are similar in the phospho-dependent and phospho-independent bindings. These residues, mostly located around the β1-β2 and β10-β11 region, form the additional contacts since their roles are not directly related to the phosphate recognition. Hydrophobic residues dominate the residues with relatively larger chemical shift changes in this region (I10, G14, V15, H19, S23, L28, I90, I91 and F95), suggesting that these additional contacts are mainly hydrophobic interactions.

The peaks of I39, Q40, L41, D92, R93 and S94 disappeared on the phospho-independent binding, therefore whether they play similar roles in phospho-dependent binding and phospho-independent binding cannot be determined. The peaks of I39, Q40, L41 and D92 can be identified on the phospho-dependent binding spectrum with good signal, indicating their roles are much better defined in the phospho-dependent binding. These residues showed large chemical shift changes (I39, 0.76 ppm; Q40, 0.40 ppm; L41, 0.74 ppm; and D92, 0.40 ppm) and are located close to V43. These residues (shown in orange in Figure 6.5) may be important in residue selection in target peptide around +1~+3 region. Since the peaks of R93 and S94 cannot be uniquely identified on the phospho-dependent binding spectrum, they are not shown in Figure 6.5 though they might also play a role in either phospho-dependent binding or phospho-independent binding.
Figure 6.5 Phospho-binding center and additional contacts. The residues in phospho-binding center are represented by red balls. The residues in the additional contacts region are represented by green balls. Left: front view. Right: side view.
6.4 Mutant study

R31 and V43 display special characteristics (Table 6.1). R31 is a conserved residue whose chemical shift changes in phospho-dependent binding and phospho-independent binding are of the same trend but of different distance. It is possible that this residue play a similar role in both types of binding but with different interaction strength. V43 is a non-conserved residue, however, its huge chemical shift (>1 ppm) upon phospho-dependent binding suggests its key role in this binding. It is possible that this residue contributes greatly to residue selection in target peptide around +1~+3 positions. Consequently, these residues are chosen for further study.
<table>
<thead>
<tr>
<th></th>
<th>R31</th>
<th>V43</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservation</td>
<td>Conserved</td>
<td>Not conserved</td>
</tr>
<tr>
<td>Environment</td>
<td>Middle, preceded by the most conserved G30</td>
<td>Tip of a big loop</td>
</tr>
<tr>
<td>Chemical shift change</td>
<td>Same trend, different degree</td>
<td>Huge change (with phospho-peptide)</td>
</tr>
<tr>
<td>Proposed roles</td>
<td>1. Direct interaction with pThr through side chain</td>
<td>Key auxiliary phospho-binding determinant</td>
</tr>
<tr>
<td></td>
<td>2. Conformational change key to accommodating the peptide.</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1 Special residues R31 and V43
**R31E mutant abolishes phospho-dependent binding and affects phospho-independent binding dramatically**

The Ki67FHA R31E mutant was constructed, expressed and purified to more than 90% pure. The structural integrity was examined by HSQC experiment (Figure 6.6). The HSQC spectrum showed that the global structure of R31E is retained. The perturbed residues include the putative phosphate-interacting residues S45, K46, suggesting E31 interacts with phosphate group through its side chain. Binding constants between Ki67FHA R31E mutant and phospho-peptide P44, GB1-hNIFK(226-269) and mutants were measured using SPR. The results (Table 6.2) showed that R31E affects both phospho-dependent binding and phospho-independent binding dramatically. R31E destroyed phospho-dependent binding completely and no binding was detected for P44 at micromolar range. While this mutation did not completely destroy phospho-independent binding, the binding affinity decreased for more than 5 fold. R31E can bind both wild type hNIFK(226-269) and threonine mutants similarly. But the binding constants could not be determined accurately since they exceed the concentration limit where Ki67FHA is stable.
Figure 6.6 HSQC spectrum for R31E
V43R and V43D mutants abolish the phospho-dependent binding only

Similar approaches were applied to Ki67FHA V43D and V43R mutants. HSQC spectrum of V43R showed no global structural perturbation (Figure 6.7). The perturbed residues include a few residues centered at D92, suggesting the side chain of R43 in V43R reaches the vicinity of D92. SPR experiments showed different effects of V43 mutants on the phospho-dependent and phospho-independent bindings. While either V43 mutant abolishes the phospho-dependent binding completely, neither of them has significant impact on the phospho-independent binding. V43D showed slightly weaker binding to the non-phosphorylated hNIFK fragment than wild type Ki67FHA, probably due to a charge effect around the local region. V43R can bind to non-phosphorylated hNIFK fragment as strongly as wild type Ki67FHA. The binding of V43R to both the phosphopeptides (234p and 238p) is about 2 times stronger than to non-phosphorylated hNIFK fragment, which may result from a charge effect between the phosphate group and Arg side chain. The V43 mutants have limited effects on the binding of the threonine mutants as well. The result showed that V43 plays an important role in phospho-dependent binding though it does not interact with the phosphate group directly, and that it does not affect the phospho-independent binding.
Figure 6.7 HSQC spectrum of V43R
Table 6.2 Binding constants (µM). Binding constants were determined using SPR. Detailed curve fitting see Appendix 3.
*: K_{d1} means tight binding (<10 µM). K_{d2} means weak binding (≥10 µM)
N.D.: not detected
N.A.: not available (data not measured)
See Appendix C for detailed binding constant determination data.
6.5 Discussion

**Key binding residues and P44 orientation**

The structural analysis of the binding of Ki67FHA using HSQC spectra and structural comparison with FHA structures in complex with synthetic peptides revealed key residues (Figure 6.8) in Ki67FHA and possible interactions with P44 (Figure 6.9). These interactions help map out the orientation of P44. Basically, the pT of P44 is situated in the crevice between the β4/β5 and β6/β7 loops and interacts with S45, K46 and N67 around this region and also interacts with the guanidinium group on the long side chain of R31. Then P44 goes through the crevice formed by β4/β5 and β10/β11. The three residues (PVC) that follow pT may interact with I39, L41 and V43 on the β4/β5 loop and probably with L91 and D92 on the β10/β11 loop. Then about 15 residues that follow the +1 ~ +3 region go through the crevice formed by β1/β2 and β10/β11 loops. Secondary structure prediction showed that these residues have a slight tendency to form an α-helix structure. It is possible that the hydrophobic residues in this region are on one side of the α-helix-like structure to interact with the hydrophobic residues in Ki67FHA. The following residues are very hydrophilic in P44, indicating they are out of the binding interface.
Figure 6.8 Key binding residues in Ki67FHA. Residues near phospho group are displayed in ball and stick style. Residues near +1~+3 of P44 are displayed in line style. Residues in additional contacts are displayed in stick style. Two flexible residues detected by NMR are displayed by red balls.
Figure 6.9 Possible interactions that determined the binding of Ki67FHA with P44
Phospho-dependent binding is a better-fit binding

In phospho-independent binding, there are large numbers of residues whose peaks disappeared or greatly weakened in the HSQC titration spectrum upon adding ligand peptide. A possible explanation for this observation is the existence of multiple conformations. The binding ligand may bind on and off frequently with different interactions prevailing at different part at one particular moment. In the phospho-dependent binding HSQC spectrum, however, there are few missing peaks, indicating the phospho-dependent binding is a two-state slow exchange binding. As indicated by the chemical shift changes, except for R31, S45 and K46 which interact with the phosphate group, phospho-dependent and phospho-independent binding use approximately the same set of residues. This means that besides the addition of phosphate interactions, phospho-dependent binding is just a better-fit version of the phospho-independent binding. One possible mechanism to coordinate the structural fitting is through R31. This residue is situated in the β3/β4 loop near the β3 strand, however, it interacts with phosphate group in P44, which is situated quite a distance away between β4/β5 and β6/β7 loops, through its long side chain. This residue is preceded by G30, a highly conserved residue which could provide non-substitutable flexibility. The fact that R31E not only destroys the phospho-dependent binding but also affects the phospho-independent binding supports the proposed role of this residue.
The auxiliary phospho-binding determinants

In the FHA1 domain of Rad53, the side chain of non-conserved residue R83 interacts with the side chain of Asp at the +3 position and thereby influences its binding to phosphopeptide. This residue is deemed as key residue to phosphopeptide recognition. V43 in Ki67 FHA aligns with R83 in Rad53. A similar role of this residue is not only supported by its huge chemical shift change upon phospho-dependent binding but also by the V43R and V43D mutants that specifically destroy phospho-dependent binding. Although the whole binding surface may contribute to the specificity of binding, this residue seems to play a key role in it. It is still not clear how this residue is involved in the binding, but hydrophobic interaction seems to be important since valine contains two hydrophobic methyl groups. In the phospho-dependent binding HSQC, large chemical shift changes were also observed for I39, Q40, L41 and D92. These residues are all located near the +1 to +3 region of P44 and their roles are much better defined in the phospho-dependent binding. It is likely that these residues all contribute to the phospho-binding.

A new model of FHA domain binding mechanism

Previous studies of FHA domains and their binding to small synthetic peptide indicate that FHA domains recognize primarily the pT residue and the +3 residue (D, I, or L at the +3 position). The results in this study clearly indicate that Ki67 FHA does not follow this simple model. On the basis of the analysis described above, the interaction of Ki67FHA with its target protein includes three factors:
(1) Direct phospho-binding determinants. This involves a few conserved residues in Ki67FHA such as R31, S45, K46 and N67. Some other conserved residues may bear the structural importance to keep the interactions in right position, such as H48, which could coordinate S45 and N67 through its imidazole group (Figure 6.10).

(2) Auxiliary phospho-binding determinants. This involves residues that interact with the +1 to +3 region in the peptide. Most of them are located in the β4/β5 loop. Some residues in the β10/β11 loop may also be involved. This part may also contribute to the phospho-independent binding though it contributes more in the phospho-dependent binding.

(3) Additional contacts. This involves residues in β1/β2 and β10/β11 region, most of which are hydrophobic. Additional contacts are distinct from auxiliary phospho-binding interactions in that it functions relatively independently and their roles are conserved in both phospho-dependent and phospho-independent bindings.
Figure 6.10 The role of H48 in coordinating S45 and N67.
Other issues about the binding surface

Accessory hydrophobic surface has been proposed for the Chk2 FHA domain based on the binding studies with the Chk2FHA I157T mutant (Li, Williams et al. 2002). The authors argued that I157 could act to stabilize an overall interaction. However, since the binding of short phospho-peptide to Chk2 FHA is already strong, it is more likely that mutation at this residue serves as a negative regulator rather than a positive regulator itself. I157 aligns with L60 in Ki67FHA, which is located far from the binding surface discussed in this study. The significance of the hydrophobic surface proposed for Chk2 FHA remains to be demonstrated.
CONCLUSIONS AND PERSPECTIVES

In this study, we have demonstrated the unique binding property of Ki67FHA, characterized the phospho-independent binding of hNIFK, identified the phosphorylation site in hNIFK and determined the effect of this site-specific phosphorylation on binding. We also analyzed the structural basis and provided a new model for the binding of FHA domains. As the first example of in vitro characterization of the binding of an FHA domain and its biological target protein, this study provides some new information on the function of Ki67 and the binding mechanism of FHA domains.

7.1 A connection between Ki67 and cell cycle control

The specific recognition of Ki67FHA toward its biological target protein

Ki67FHA exhibits a high specificity toward its biological target protein, phosphorylated hNIFK. This high specificity was indicated by several pieces of evidence. First of all, the binding of Ki67FHA is distinct from other FHA domains since it excludes the binding
of small phosphopeptides even if they are from biological target proteins. This unique property was ascribed to a tertiary structure requirement in the target protein. This tertiary structure requirement was supported by the $^{31}$P NMR experiments, which indicated that Ki67FHA probably prefers a specific proline conformation in the vicinity of pThr. High selectivity of Ki67FHA was also supported by the compact structure of this domain, particularly the compact loop formed by $\beta_{10}$ and $\beta_{11}$, where the flexibility of D92, one of the potential auxiliary phospho-binding residues, is greatly confined. And finally, high specificity is indicated by the fact that a large number of residues, which form a coordinating binding surface, are involved in the binding.

**Ki67’s relationship with cell cycle-related kinases**

The fact that the existence of Ki67 is strongly associated with proliferating cells suggests its close relationship with the cell cycle. One of the key cell cycle regulation mechanisms is through cyclin dependent kinases (CDKs). In this study, the phosphorylation site was determined to be T234. The target peptide of Ki67FHA contains pThr-Pro motif which is usually the phosphorylation product of either CDKs or MAPKs (mitogen-activated protein kinases). It is possible that Ki67 interacts with its partner proteins, which are phosphorylated by CDKs or MAPKs, through the FHA domain which recognizes the pThr-Pro motif. Preliminary experiment of kinase assay (as described in Materials and Methods) was unsuccessful in identifying CDC2/cyclinB and CDK4/cyclinD as the kinase for GB1-hNIFK(226-269). However, since T238 was reported to be required for phosphorylation-dependent tight binding *in vivo*, it is possible that the phosphorylation of
T234 involves a complex process which cannot be mimicked by \textit{in vitro} kinase assay. It is interesting to explore the relationship between the recognition site of phospho-dependent binding and the kinase recognition site. If the two sites overlap with each other, it suggests a continuous regulation. Otherwise, it suggests a cross regulation in the cell cycle regulation network.

**Ki67’s functional pattern**

Previous studies of hNIFK and Hklp2 have shown that Ki67 co-localizes with these proteins during the cell cycle. Relocation could be an important regulation mechanism for the activities of Ki67 and/or these two proteins. However, this does not exclude the possibility that the activity of Ki67 or its partner protein is regulated through a conformational change induced by the FHA domain. The existence of pThrPro in the binding site is reminiscent of the sequence-specific and phospho-dependent proline isomerization role of Pin1 in cell cycle regulation (Yaffé, Schutkowski et al. 1997). Recently tumor suppressor p53 has been shown to be under the control of Pin1, which recognizes pThrPro, through a conformational change (Zacchi, Gostissa et al. 2002; Zheng, You et al. 2002). Since our preliminary data suggest that Ki67FHA may prefer one specific conformation related to the prolines flanking pThr, Ki67 could potentially regulate the activity of its target protein through the stabilization of one particular conformation.
**Other proteins that interact with Ki67FHA**

Besides hNIFK, another identified biological target protein of Ki67FHA is Hklp2. The Ki67-interacting motif of Hklp2 has been mapped to 1017-1237. When comparing this fragment of Hklp2 with its homologue in Xenopus, Xklp2, there are only 5 threonines/serines that are conserved (S1021, S1088, S1134, T1144 and S1169). Since FHA domains have been shown to prefer pThr to pSer, T1144 seems to be the most possible phosphorylation site for tight binding. Comparison of the sequences around T1144 in Hklp2 and T234 in hNIFK revealed a similar pattern of pThr-Pro(+1)-hydrophobic(+2)-proton donor(+3)-α-helix forming region. It would be interesting to see whether this is the recognition pattern of Ki67FHA. Once identified, it will help discover more proteins that interact with Ki67FHA.

**7.2 New concepts on FHA domains**

**The biological boundary of FHA domains**

It has been demonstrated that functional FHA domains should contain a much longer region than the core domain identified by computer alignment. This is supported by the structural instability of truncated forms of FHA domains. Structures of FHA domains also demonstrate that FHA domains usually contain 11 β-strands forming a sandwich structure. In this study, we identified additional contacts that are actually located within the first and last loops, forming part of the binding surface. This supports that the 11-β-strand sandwich structure is functionally required as well.
**FHA domains may have additional specificity determinants**

Previous studies on FHA domains indicated a low specificity for target peptides. The specificity of FHA domains was proposed to be determined by pT and the +3 residue following it. For example, in the study of Rad53 FHA1 domain, a short phosphopeptide with an Asp at the +3 position is enough to confer tight binding. This pattern, which coincides with the casein kinase II-recognition site, can be found in numerous places in proteins. This has made the identification of phosphorylation site(s) in biological target proteins using library screening result very difficult. FHA domains were also shown to bind small phosphopeptides without obvious biological relevance. For example, in the study of Rad53, FHA2 domain was shown to bind a pTyr peptide, the biological relevance of which is elusive. It is reasonable to expect additional determinants that affect the binding specificity of an FHA domain to its target protein. These additional determinants are observed in this study. Based on the new binding model of the binding of FHA domain, the recognition of target peptides involves three factors including phosphate interacting residues, auxiliary phospho-binding determinants and additional contacts that are remote from the phospho-binding center. This binding involves much more residues (estimated to be around 20 amino acid residues), which could form certain secondary structure, in the binding target. The method used in this study to discover the additional contacts in Ki67FHA provides a framework to explore the additional determinants in other FHA domains.
Limitations of library screening method in the study of FHA domain

The new model proposed for the binding of FHA domains provides a challenge to the library screening method, which was extensively used in previous studies, since it suggests that much more residues at the C-terminus side from pT are involved in binding. *In vivo* identification seems to be necessary for a reliable conclusion.
REFERENCES


APPENDIX A

PEPTIDE TITRATION SPECTRA OF KI67FHA WITH SMALL PHOSPHOPEPTIDES
Figure A.1 $^{15}$N Ki67FHA titrates with T3-3
Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.2 $^{15}$N Ki67FHA titrates with T2-1

Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.3 $^{15}$N Ki67FHA titrates with T2-2

Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.4 $^{15}$N Ki67FHA titrates with T2-11

Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.5 $^{15}$N Ki67FHA titrates with p53_peptide

Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.6 $^{15}$N Ki67FHA titrates with Hklp2_1021p
Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.7 $^{15}$N Ki67FHA titrates with Hklp2_1088p
Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.8 $^{15}$N Ki67FHA titrates with Hklp2_1134p
Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.9 $^{15}$N Ki67FHA titrates with Hklp2_1144p

Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
N15 Ki67FHA vs HKP2-1144P
Figure A.10 $^{15}$N Ki67FHA titrates with Hklp2_1179p

Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:4
Figure A.11 $^{15}$N Ki67FHA titrates with hNIFK-234p

Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:0.25
Green: protein: peptide=1:1
Blue: protein: peptide=1:4
Figure A.12 $^{15}$N Ki67FHA titrates with hNIFK_238p

Black: before adding peptide
Red: after adding peptide. Protein: peptide=1:0.25
Green: protein: peptide=1:0.5
Blue: protein: peptide=1:1
Turquoise: protein: peptide=1:4
APPENDIX B

CONTROL EXPERIMENTS FOR PHOSPHO-INDEPENDENT BINDING
Figure B.1 $^{15}$N FHA1 titration with GB1-hNIFK(226-269)

Black: $^{15}$N FHA1 only
Red: $^{15}$N FHA1:GB1-hNIFK=1:0.5
Green: $^{15}$N FHA1:GB1-hNIFK=1:1
Blue: $^{15}$N FHA1:GB1-hNIFK=1:2
Figure B.2 $^{15}$N FHA2 titration with GB1-hNIFK(226-269)

Black: $^{15}$N FHA2 only  
Red: $^{15}$N FHA2:GB1-hNIFK=1:0.5  
Green: $^{15}$N FHA2:GB1-hNIFK=1:2
Figure B3 $^{15}$N GB1-hNIFK(226-269) titration with FHA2

Black: $^{15}$N GB1-hNIFK(226-269) only
Red: $^{15}$N GB1-hNIFK(226-269): FHA2=1:0.5
Green: $^{15}$N GB1-hNIFK(226-269):FHA2=1:1
Figure B.4 $^{15}$N GB1-hNIFK(226-269) with Ki67FHA
Black: $^{15}$N GB1-hNIFK(226-269) only
Red: $^{15}$N GB1-hNIFK(226-269): Ki67FHA=1:0.25
Green: $^{15}$N GB1-hNIFK(226-269): Ki67FHA=1:0.5
Blue: $^{15}$N GB1-hNIFK(226-269): Ki67FHA=1:1
Turquoise: $^{15}$N GB1-hNIFK(226-269):Ki67FHA=1:2
Figure B.5 $^{15}$N Ki67FHA titration with GB1 domain
Black: $^{15}$N Ki67FHA only
Red: $^{15}$N Ki67FHA: GB1=1:0.5
Green: $^{15}$N Ki67FHA: GB1=1:1
APPENDIX C

$K_D$ DETERMINATION USING SPR
Figure C.1 $K_d$ determination for Ki67FHA and GB1-hNIFK(226-269)
Figure C.2 No binding found between Ki67FHA to itself
Figure C.3 binding between Ki67FHA(R31E) and hNIFK(226-269). The binding constant is beyond accurate determination range but is estimated to be >500µM.
Figure C.4 binding between Ki67FHA(R31E) and GB1-hNIFK(226-269) T238A. The binding constant is beyond accurate determination range but is estimated to be >500µM.
Figure C.5 binding between Ki67FHA(R31E) and hNIFK(226-269) T238E
The binding constant is beyond accurate determination range but is estimated to be >500µM
Figure C.6 $K_d$ determination for Ki67FHA V43R and GB1-hNIFK(226-269) T238E
Figure C.7 $K_d$ determination for Ki67FHA V43R and GB1-hNIFK(226-269) T238A
Figure C.8 $K_d$ determination for Ki67FHA V43R and GB1-hNIFK(226-269)

- $R_{\text{Umax}} = 446.1962$
- $K_d = 101.1671 \mu\text{M}$
- $R = 0.99930316$
Figure C.9 $K_d$ determination for Ki67FHA and synthetic phosphopeptide 234p

\[
f = \frac{a \cdot x}{b + x} + \frac{c \cdot x}{d + x}
\]

$R = 0.99984252$

$a = 188.8185$

$b = 1.6261$

$c = 814.9263$

$d = 146.1757$
Figure C.10 $K_d$ determination for Ki67FHA and synthetic phosphopeptide 238p
Figure C.11 $K_d$ determination for Ki67FHA and synthetic phosphopeptide 234_238dp

\[ f = a \times \frac{1}{x} + c \times \frac{x}{d + x} \]

\[ R = 0.99993463 \]

\[ a = 68.2762 \]

\[ b = 1.3870 \]

\[ c = 563.2336 \]

\[ d = 1.361103 \]
Figure C.12 K_d determination for Ki67FHA V43R and synthetic phosphopeptide 234_238dp

\[ f = \frac{a \times x}{b + x} \]

\[ R = 0.99940018 \]

\[ a = 596.0245 \]

\[ b = 56.8385 \]
Figure C.13 $K_d$ determination for Ki67FHA V43R and synthetic phosphopeptide 238p

\[ f = \frac{a \times x}{b + x} \]

$R = 0.99965019$

$a = 928.8776$

$b = 55.6000$
Figure C.14 $K_d$ determination for Ki67FHA V43R and synthetic phosphopeptide 234p.

The equation used is:

$$f = \frac{ax}{b+x}$$

with:

- $R = 0.99969441$
- $a = 879.5455$
- $b = 56.0904$
Figure C.15 $K_d$ determination for Ki67FHA and GB1-hNIFK(226-269) T234A

\[ R = R_{\text{max}} \frac{c}{(K_d + c)} \]

$R_{\text{max}} = 639.9060 \text{ RU}$

$K_d = 131.2061 \text{ uM}$

correlation coefficient = $0.99966079$
Figure C.16 K\textsubscript{d} determination for Ki67FHA and GB1-hNIFK(226-269) T234E

\[ R = R_{\text{max}} \cdot \frac{c}{(K_d + c)} \]

- \( R_{\text{max}} = 790.8534 \text{ RU} \)
- \( K_d = 129.9422 \text{ uM} \)

Correlation coefficient = 0.99970630
Figure C.17 $K_d$ determination for Ki67FHA V43D vs. GB1-hNIFK(226-269) T234A

$R = 0.99517522$

$RU_{max} = 717.6227$

$K_d = 181.3215$
Figure C.18 $K_d$ determination for Ki67FHA V43D vs. GB1-hNIFK(226-269) T234E
Figure C.19 Kd determination for Ki67FHA V43D vs. GB1-hNIFK(226-269) T238E

\[ R = 0.99484017 \]
\[ R_{\text{max}} = 810.9403 \]
\[ K_d = 158.3931 \]