doi:10.1016/j.jmb.2003.10.032 J. Mol. Biol. (2004) 335, 371–381

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

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

Hongyuan Li, In-Ja L. Byeon*, Yong Ju and Ming-Daw Tsai*

Departments of Biochemistry and Chemistry, Campus Chemical Instrument Center The Ohio State University Columbus, OH 43210, USA

**Corresponding authors*

Recent studies by use of short phosphopeptides showed that forkhead-associated (FHA) domains recognize pTXX(D/I/L) motifs. Solution structures and crystal structures of several different FHA domains and their complexes with short phosphopeptides have been reported by several groups. We now report the solution structure of the FHA domain of human Ki67, a large nuclear protein associated with the cell-cycle. Using fragments of its binding partner hNIFK, we show that Ki67-hNIFK binding involves ca 44 residues without a pTXX(D/I/L) motif. The pThr site of hNIFK recognized by Ki67 FHA is pThr234-Pro235, a motif also recognized by the proline isomerase Pin1. Heteronuclear single quantum coherence (HSQC) NMR was then used to map out the binding surface, and structural analyses were used to identify key binding residues of Ki67 FHA. The results represent the first structural characterization of the complex of an FHA domain with a biologically relevant target protein fragment. Detailed analyses of the results led us to propose that three major factors control the interaction of FHA with its target protein: the pT residue, +1 to +3 residues, and an extended binding surface, and that variation in the three factors is the likely cause of the great diversity in the function and specificity of FHA domains from different sources.

Keywords: FHA; NMR; protein structure; binding surface; phosphoprotein

Introduction

The forkhead-associated (FHA) domain has been demonstrated to be a modular phosphopeptide-binding domain.[1–5] It has been identified in more than 200 proteins with great functional diversity, including DNA damage repair and signaling proteins, kinesins, ring-finger proteins, forkhead transcription factors, nuclear inhibitor of protein phosphatase 1 (NIPP-1), and Ki67.[6] Results from studies with short pT-containing peptides and peptide libraries indicate that FHA recognize pTXX(D/I/L) motifs. The structures of several FHA domains and their complexes with short phosphopeptides have been reported.[1,3,4,7–11] However, the biological target proteins of FHA domains have been proposed or identified in only a few cases, including RLK5 (KAPP FHA domain),[12] Rad9 (two FHA domains of Rad53),[13] Chk2 (Chk2 FHA domain),[14] Pan3 (Dun1 FHA domain),[15] and Hklp2[16] and hNIFK[17] (Ki67FHA domain). Some of the evidence is indirect and remains to be verified further. Most of these suggested target proteins contain multiple pTXX(D/I/L) motifs. For example, Rad9 contains five TXXD motifs (potential binding motif for the FHA1 domain of Rad53 binding) and 22 TXX(I/L) motifs (potential binding motif for the FHA2 domain of Rad53). Specific binding sites have been suggested in some cases on the basis of the results from short phosphopeptides.[4] However,

Present address: I.-J. L. Byeon, Building 5, Room B2-41, NIH, Bethesda, MD 20892-0520, USA.

Y.J. was on sabbatical leave from Tsinghua University, Beijing, People's Republic of China.

Abbreviations used: NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; HSQC, heteronuclear single quantum coherence; r.m.s.d., root-mean-square deviation; FHA, forkhead-associated; SPR, surface plasmon resonance; CDK, cyclin-dependent kinase.

E-mail addresses of the corresponding authors: tsai.7@osu.edu; in-jab@intra.niddk.nih.gov

in no case has the actual pT residue recognized by an FHA domain been demonstrated by chemical or structural analysis of the complex between FHA and a protein-binding partner.

The Ki67 protein was originally identified as an antigen which was recognized by a monoclonal antibody Ki67.[18] This antigen is present in the nuclei of cells in the G_1, S, and G_2 phases of dividing cells as well as in mitosis but not in the G_0 phase of quiescent cells.[19] This stringent feature has made Ki67 a popular prognostic and diagnostic tool.[20] The location of Ki67 is dynamically related to cell cycle. It is a constituent of compact chromatin[21] and is vital for cell proliferation, since removal of Ki67 protein using antisense nucleotides prevents cell proliferation.[22] Recent studies showed that the C terminus of Ki67 interacts with heterochromatin protein 1 family and potentially plays a role in higher-order chromatin organization.[23] All these suggest that Ki67 protein is involved in the protein interaction network that drives cell division cycles. Since Ki67 contains an FHA domain, recent biological studies suggested that Ki67 protein could interact with phosphorylated proteins through its FHA domain. Two Ki67FHA domain interacting proteins, Hklp2[16] and hNIFK,[17] have been identified. Both are phosphorylated in mitotic cells and could be involved in complex signal transduction pathways. Understanding the structural basis of the interactions between Ki67 and its binding partners will contribute to our knowledge not only on the functions of Ki67 but also on the control of cell-cycles.

Here, we present the structure of the Ki67FHA domain. The specific phosphoamino acid recognized by Ki67FHA was identified by use of various synthetic phosphopeptide fragments of hNIFK. The binding motif was shown to be different from all other FHA domains studied with short phosphopeptides. The detailed binding interaction was characterized by a 44 amino acid residue phosphopeptide corresponding to the residues 226–269 from the biological target protein hNIFK. A new view to the mechanism of FHA recognition is proposed.

Results and Discussion

Solution structure of Ki67FHA and comparison with other FHA domains

The construct for the Ki67FHA domain contains residues 1–120 of Ki67 and eight additional linker residues (GSPEFPGG) attached to the N terminus of Ki67FHA. The fragment is unstable and has a tendency to aggregate, thus NMR experiments were performed at lower temperature (17 °C) and lower concentration (< 0.5 mM). While N and C-terminal residues exhibit characteristics of flexible random coiled conformation, residues 4–100 display well-dispersed HSQC peaks. Of this core section, all but six (R12, S13, E55, S65, T66, N67) backbone amide groups and more than 95% of all NMR resonances were assigned. The unassigned amides were not detected in $^{15}N-^{1}H$ heteronuclear single quantum coherence (HSQC) most likely due to exchange broadening.

The structures of Ki67FHA were calculated using a simulated annealing protocol using the NIH version[24] of the program X-PLOR.[25] The total number of distance and torsion angle restraints that were applied to the calculation is 1921, and thus ca 20 restraints per residue. The best-fit superposition of the final 22 ensemble structures is shown in Figure 1 and structural statistics are provided in Table 1. The structures are well defined, satisfy the experimental restraints, and show very good covalent geometry. All residues are observed in the allowed Ramachandran space. They also show very good structural convergence with r.m.s.d. values of 0.37(±0.05) Å for the backbone atoms and 0.84(±0.05) Å for the heavy atoms excluding the disordered loop between β strands 1 and 2 (residues 11–16). The disorder of the loop is likely a result of intermediate loop motions in chemical shift time-scale because the two backbone NH groups in the loop (R12 and S13) are not detected in $^{15}N-^{1}H$ HSQC and the loop exhibits little medium or long-range NOEs.

Residues 4–100 of Ki67 show a β-sandwich scaffold essentially identical with those of the FHA domains from Rad53,[1,3,4] Chk2[11] and Chfr.[7] Search of structures in the Protein Data Bank using the program DALI[26] confirmed that Ki67FHA shows close structural similarity (Z-score > 10 to Ki67FHA) to all FHA domains in the data bank: Rad53 FHA1 (Z-score = 11.5 for free FHA1, 1J4O; 12.8 and 11.3 for the FHA1/pT peptide complexes, 1G6G and 1J4P, respectively), Rad53 FHA2 (Z-score = 10.6 for free FHA2, 1FHQ; 10.8 for the FHA2/pY peptide complex, 1J4K), and Chk2 FHA (Z-score = 12.3 for the FHA/pT peptide complex, 1GXC). Sequence alignments of Ki67FHA to other FHA domains using the structural homology are shown in Figure 2A. Note that the loop length and structures vary substantially although all the 11 β-strands are highly conserved in all FHA domains. It is interesting to note that Ki67FHA has the smallest number of residues to fold into the FHA characteristic β-sandwich because their loops are shorter than the corresponding loops in the other FHA domains. Figure 2B shows the comparison of the ribbon diagrams between Ki67FHA and Rad53 FHA1. It is clear that these two FHA domains are well superimposed: the backbone atomic r.m.s.d. value is 1.6 Å when the extended loops found in FHA1 between β2 and β3 and β9 and β10 are excluded. Figure 2B also displays the side-chains that have been identified to interact with the phosphate group of the phosphorylated peptides in FHA1, FHA2, and Chk2. It is interesting to observe that residues 31(Arg), 45(Ser) and 46(Lys) of Ki67FHA are located in a three-dimensional space very similar to the corresponding ones of the other three FHA domains. Their interactions with

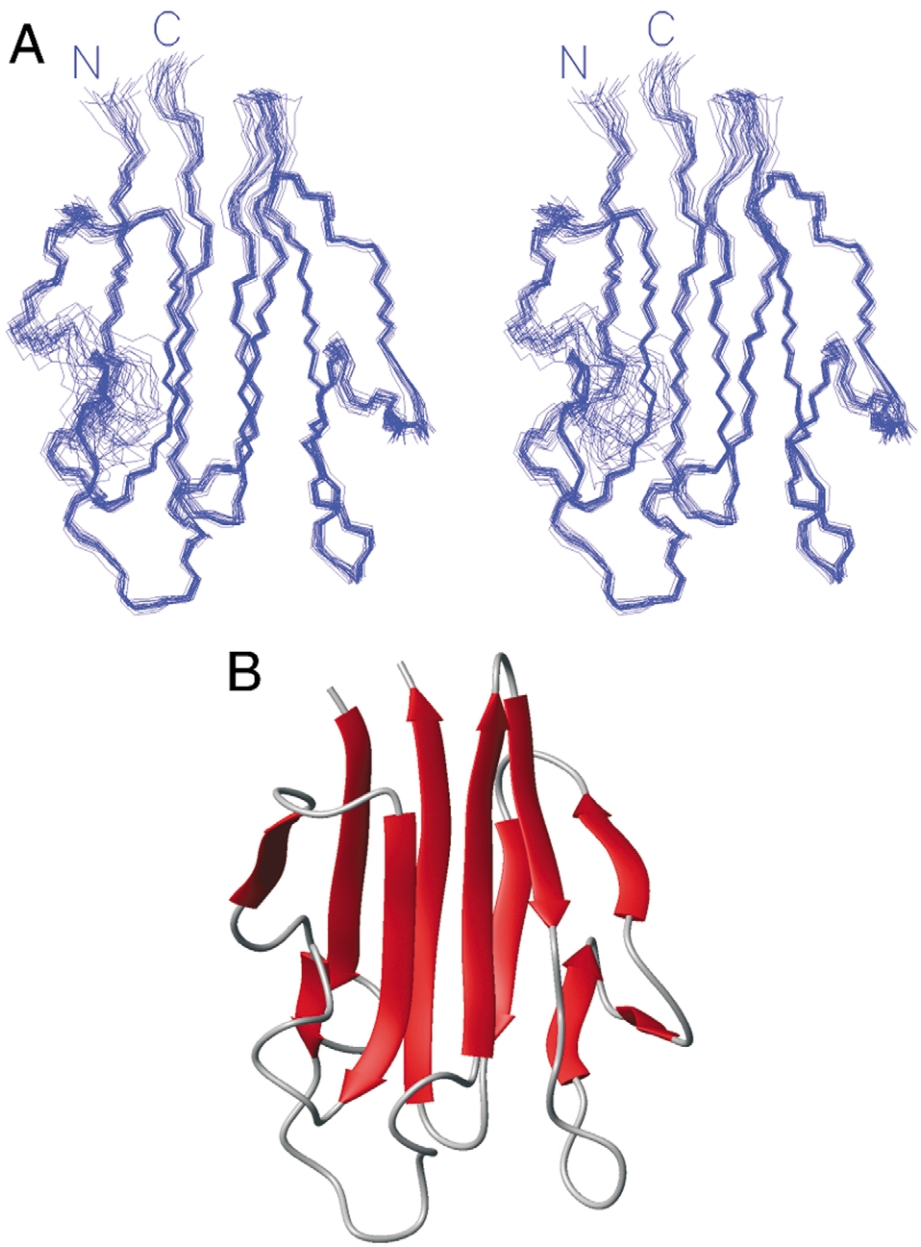

Figure 1. A, Stereo view showing superposition of the backbone traces (N, C^{α}, C′) of the final 22 ensemble structures of Ki67FHA (residues 4–100). B, The ribbon diagram.

the phosphate group have been observed in all the structures of FHA-phosphopeptide complexes determined previously. On the other hand, S65 of Ki67 FHA is found in a rather distant position from that of the corresponding residue of Rad53 FHA1 (T106, which was shown to interact with the phosphate group).[3] It remains to be established whether S65 of Ki67 is not involved in the interaction with the phosphate group, or the loop structure around residue 65 is not well defined due to the lack of its assignment.

Ki67FHA does not bind short phosphopeptides containing pTXX(D/I/L) motifs

Previously, the binding of FHA domains has been shown to be highly dependent on the linear sequence near pT, particularly at the +3 position.[2–4] However, the three approaches that have been used to identify tight-binding peptides for other FHA domains have not been successful with the Ki67 FHA domain. (a) The library screening method which has successfully identified the tight binding phosphopeptides for FHA1 domain and FHA2 domain of Rad53,[4,8] was used to screen Ki67FHA. The three libraries used in this study are $X_1X_2X_3(pT/pS/pY)X_4X_5X_6$. In all experiments, FHA1 was used as a positive control. Under the conditions when FHA1 gave more than 90% of Asp at the +3 position, Ki67FHA did not show a clear selection at the +3 position or other positions. (b) As a complementary approach, synthetic

Table 1. Structural statistics for the FHA domain of Ki67

	22 Conformer ensemble	Regularized mean structure
A. *r.m.s.d. from experimental restraints*[a]		
Distance restraints (Å)[b]		
All (1756)	0.035 ± 0.002	0.032
Sequential ($\|i-j\|=1$) (392)	0.030 ± 0.003	0.029
Medium ($\|i-j\|<5$) (141)	0.034 ± 0.008	0.038
Long ($\|i-j\|\geq 5$) (808)	0.043 ± 0.002	0.038
Intra-residue (353)	0.013 ± 0.010	0.012
Hydrogen bonds (62)[c]	0.028 ± 0.003	0.019
Dihedral angle restraints (°) (165)[d]	0.485 ± 0.114	0.493
B. *r.m.s.d. from idealized covalent geometry*		
Bonds (Å)	0.0028 ± 0.0003	0.0028
Angles (deg.)	0.547 ± 0.018	0.593
Impropers (deg.)	0.431 ± 0.021	0.540
C. *Measures of structural quality*		
Ramachandran plot[e]		
Most favored regions (%)	81.4 ± 1.5	84.4
Additionally allowed regions (%)	18.4 ± 1.8	15.6
Generously allowed regions (%)	0.2 ± 0.5	0.0
Disallowed regions (%)	0.0 ± 0.0	0.0
E_{L-J} (kcal/mol)[f]	−337 ± 9	−338
Coordinate precision (Å)[g]		
Backbone atoms	0.37 ± 0.05	–
All heavy atoms	0.84 ± 0.05	–

[a] The number of restraints is given in parentheses. The restraints were exclusively from the core residues 4–100.
[b] None of the ensemble structures exhibited distance violations greater than 0.5 Å.
[c] All hydrogen-bonding restraints involve slowly exchanging NH protons.
[d] Dihedral angle restraints comprise 72 ϕ, 71 ψ and 22 χ_1 angles. None of the structures exhibit dihedral angle violations greater than 6°.
[e] Calculated by PROCHECK. Poorly defined regions (residues 1–4, 1–16 and 100–120) were excluded in the statistics.
[f] E_{L-J} is the Lennard–Jones van der Waals energy calculated with the CHARMM empirical energy function and is not included in the target function for simulated annealing or restrained minimization.
[g] Coordinate precision for the ensemble is defined as the average atomic r.m.s.d. value from the mean structure obtained by averaging the coordinates of the 22 ensemble structures best fit to each other excluding poorly defined regions (residues 1–4, 11–16 and 100–120).

pT-peptides with different types of residues at the +3 position were used to titrate ^{15}N Ki67FHA in HSQC NMR experiments. These include peptides with aromatic, hydrophobic, acidic and basic residues at position +3. Addition of up to 2 mM of any of these peptides to a 0.4 mM Ki67FHA solution causes no observable chemical shift changes on the HSQC spectrum of Ki67FHA. (c) We next tested the binding of short phosphorylated peptides from the biological targets of Ki67FHA. Seven peptides (11–13 residues) corresponding to possible phosphorylation sites within the identified binding motif in hNIFK and Hklp2 proteins were synthesized and used for HSQC titration. These include five peptides from Hklp2 centered at conserved serine and threonine residues (S1021, S1088, S1134, T1144 and S1169) and two peptides from hNIFK centered at T234 and T238 whose mutation has been reported to disrupt tight binding *in vivo*.[17] None of the peptides caused any chemical shift change in the ^{15}N HSQC spectrum of Ki67 FHA, indicating lack of binding. Overall, the results are very different from those of other FHA domains studied to date, which are able to bind more than one type of short phosphopeptide. It is possible that only longer peptide fragments from the biological partner are able to bind to the FHA domain of Ki67.

Identification of a tight-binding 44 residue fragment consisting of a (pThr)Pro motif from the biological binding partner hNIFK

On the basis of yeast two-hybrid experiments, Ki67FHA has been suggested to interact with a 44 residue fragment from hNIFK covering residues 226–269. Furthermore, pull-down assays using mitotic cell extract-treated hNIFK and mutants suggested that phosphorylation at T234 and/or T238 are required for tight binding *in vivo*.[17] We therefore overexpressed this 44 residue fragment of hNIFK (designated as hNIFK44) in *Escherichia coli* as a fusion protein with a solubility enhancement tag GB1 domain.[27] This tag greatly increased the yield and stability of hNIFK but did not interact with Ki67FHA domain directly or indirectly as tested by NMR. HSQC NMR experiments demonstrated that Ki67FHA indeed binds to hNIFK44, albeit weakly; details of NMR results are described in the next section. The K_d value was determined using surface plasmon resonance (SPR) to be around 100 μM.

The results described above suggested that a

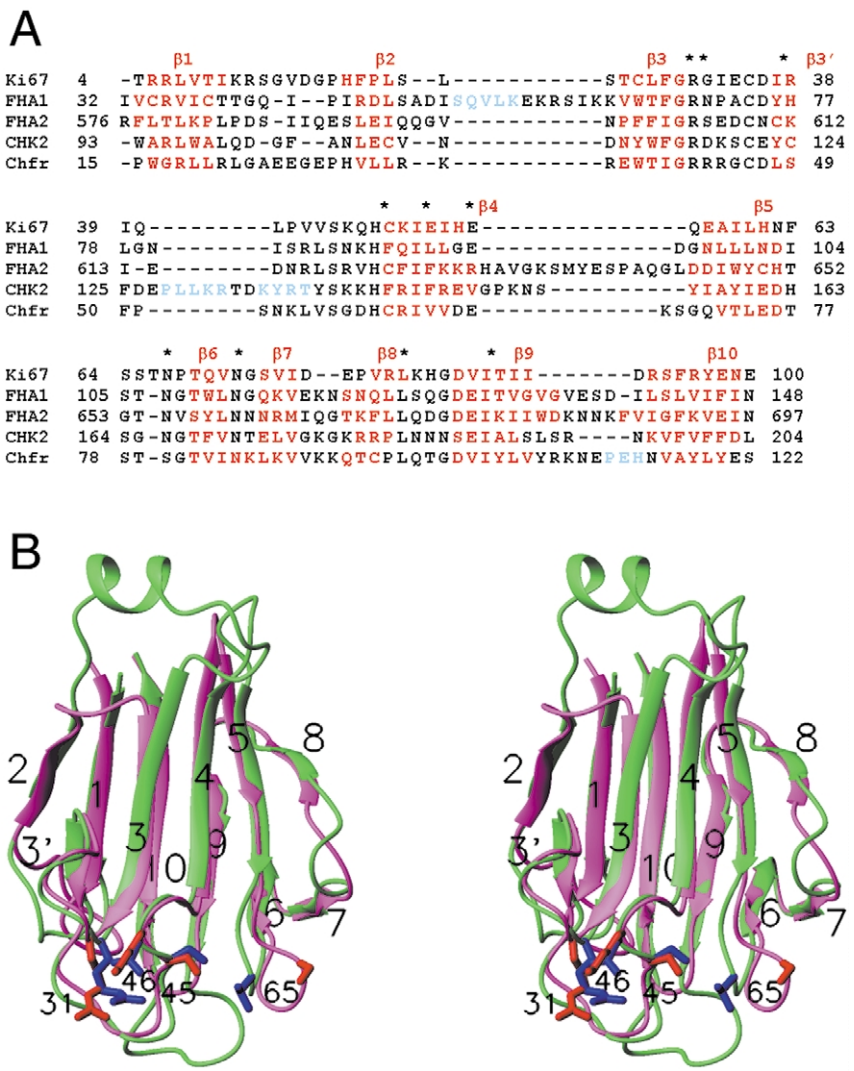

Figure 2. A, Sequential and structural alignments of Ki67 FHA to FHA1 and FHA2 of yeast Rad53, the human Chk2 FHA domain, and the FHA domain of the Chfr mitotic checkpoint protein. The β-strands are shown in red, and helices found in loops are shown in cyan. Identical residues are indicated by stars. B, Stereo view of the superposition of the ribbon diagrams of Ki67FHA (magenta) and Rad53 FHA1 (green, 1G6G). The 11 β-strands are numbered sequentially as in FHA2,[8] namely 1–10 for antiparallel strands and 3′ for the parallel β-strand. The side-chains of 31, 45, 46 and 65 in Ki67 are colored in red, and the corresponding residues in FHA1 where they interact directly with the phosphate group of a pT peptide are in blue.

tight-binding fragment could be obtained upon phosphorylation of the Thr residues in hNIFK44. Attempts to achieve Thr phosphorylation by commercially available protein kinases were unsuccessful. As an alternative approach, three phosphopeptides corresponding to hNIFK44 with phosphorylation at T234, T238 or both (designated as SP44-234p, SP44-238p, and SP44-234-238dp, respectively) were chemically synthesized and tested for binding to Ki67 FHA by SPR. Tight binding was observed for SP44-234p and SP44-234_238dp ($K_d = 1.63$ μM and 1.38 μM, respectively) but not for SP44-238p, indicating that phosphorylation at T234 but not T238 is essential to mediate the tight binding. The synthetic peptide SP44-234p, which has the following sequence, was used for further structural analysis as described in the next section.

^{226}KTVDSQGP(pT)PVCTPTFLERRKSQVAELNDDDKDDEIVFKQPISC.

While both hNIFK44 and SP44-234p are able to bind to Ki67 FHA as described above, the 1D proton NMR experiment indicated that neither displays clearly folded tertiary structures in the free form. Structure needed for specific binding is likely to be induced upon binding with the FHA domain.

There are two significant aspects in the results described above: a long phosphopeptide is required for binding to Ki67 FHA, and a proline residue follows immediately the phosphothreonine site. In the other known binding partner protein of Ki67FHA, Hklp2, the only conserved threonine (T1144) in the mapped binding region is also followed by a proline residue.[16] Since the function of Ki67 is closely related to cell proliferation and the cell cycle is controlled by specific kinases such as

cyclin-dependent kinases (CDKs), it is tempting to associate Ki67FHA phosphoprotein binding activity with the kinase activities. CDKs and MAPKs (mitogen-activated protein kinases) have been reported to recognize Thr-Pro motif.[28] While previous studies with other FHA domains show that small phosphopeptides are sufficient to convey tight binding, little is known about whether these FHA domains bind to their biological target proteins in a similar way. On the other hand, in a few cases including Ki67FHA, the region of FHA domains involved in binding to their biological target has been mapped to a region of a few hundred residues using *in vivo* methods, but the exact phosphorylation site has not been demonstrated by *in vitro* experiments. The results in this study serve to bridge this important gap.

NMR analyses of the interaction of Ki67 FHA with phosphorylated and non-phosphorylated fragments

Tight binding of Ki67FHA with SP44-234p was characterized by slow-exchange chemical shift changes in both ^{31}P NMR and ^{15}N-HSQC experiments. Figure 3 shows the overlay of the HSQC spectra of the Ki67FHA domain in the free form and complexed to SP44-234p. Assignment of the ^{15}N-^{1}H HSQC resonances of Ki67FHA complex with SP44-234p was conducted by comparing their NOE patterns with those of free Ki67FHA using 3D ^{15}N-edited NOESY. This approach allowed us to assign the Ki67FHA ^{15}N-^{1}H resonances in the complex for all but 11 residues, and thus to identify the residues affected by the binding. R38, R93 and S94 could not be assigned because of insufficient NOE evidence. However, these residues are likely involved in binding because their HN peaks were not found within 0.3 ppm compared to the peptide-free position. R12, S13, E55, S65, T66 and N67 could not be assigned because these residues were not assigned in the free Ki67 FHA. I33 and F63 are located in a region too crowded to assign. The summary of the chemical shift perturbation introduced by binding of SP44-234p is shown in Figure 4A. The greatest perturbations are observed in the loop between β3′ and β4; all residues in this loop are labeled as red or orange, as shown in Figure 4B.

Binding of the non-phosphorylated fragment hNIFK44 was also examined by HSQC NMR. The binding is clearly weak as the peaks shift

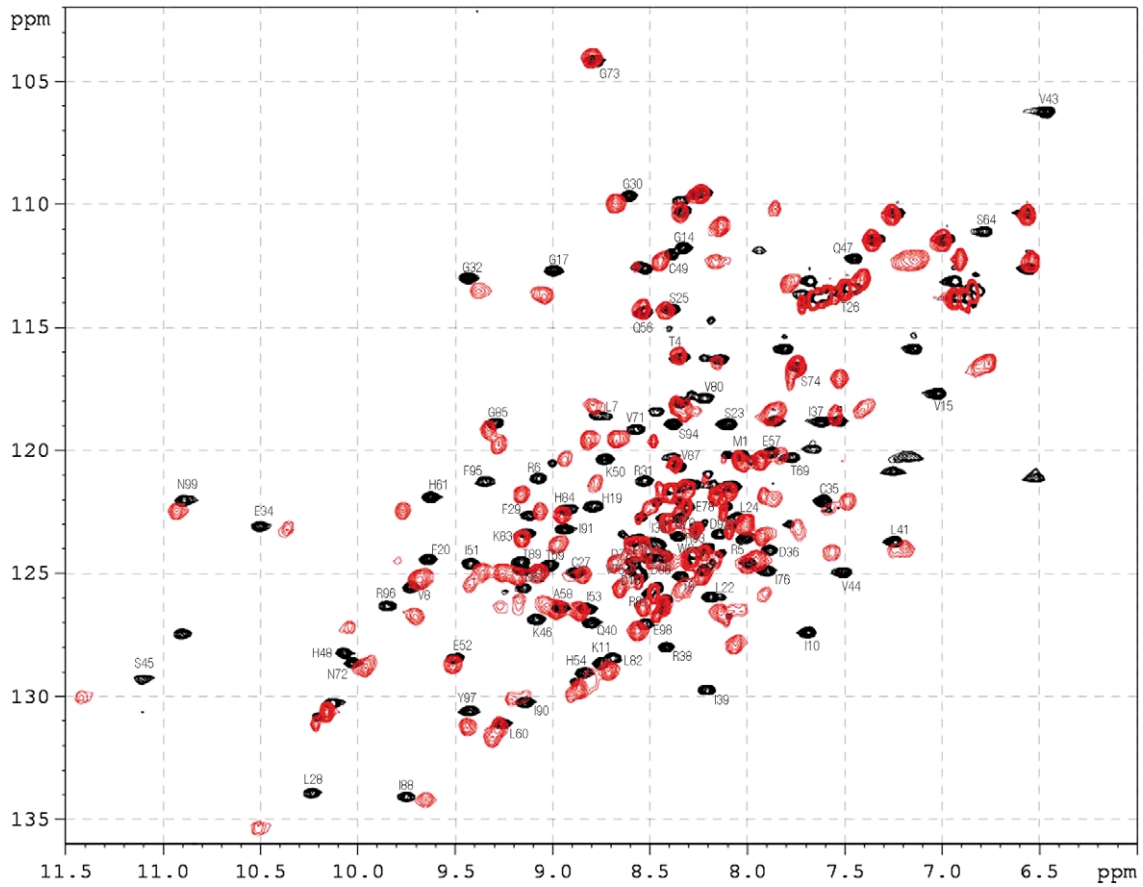

Figure 3. Overlay of the ^{15}N HSQC spectra of Ki67FHA in the free form (black) and with the 44 amino acid phosphopeptide SP44-234p bound (red). Only the assignments of the free form are shown in order to avoid overcrowding. The assignments of the complex were performed as described in the text.

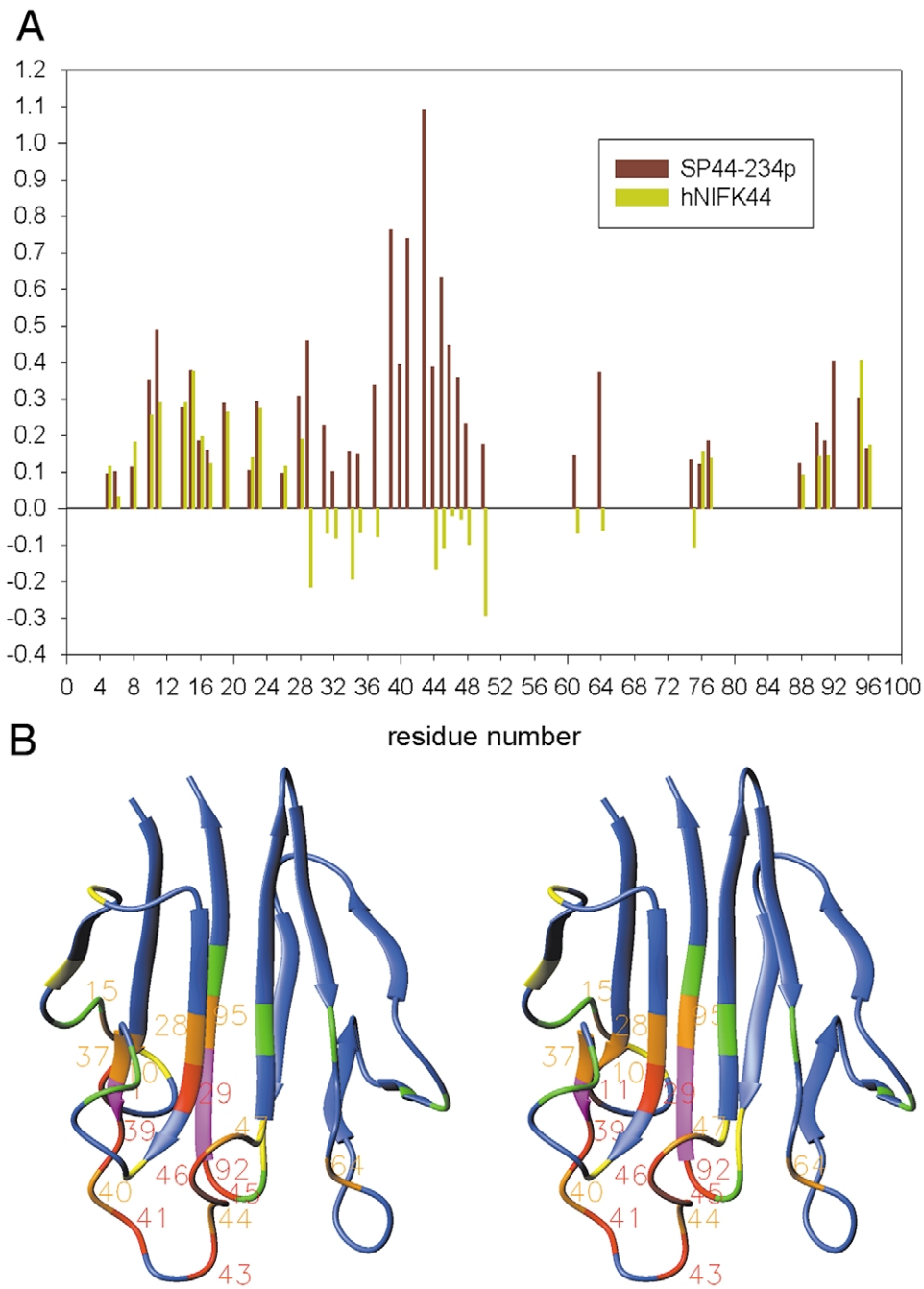

Figure 4. A, Chemical shift perturbations of the Ki67FHA backbone amide groups upon binding SP44-234p (brown) and hNIFK44 (green). The combined ^{1}H and ^{15}N chemical shift differences, calculated using the equation $\Delta ppm = [(\Delta\delta_{HN})^2 + (\Delta\delta_N \times \alpha_N)^2]^{1/2}$, where α_N is the scaling factor (0.17) used to normalize the ^{1}H and ^{15}N chemical shifts,[9,39] were plotted against residue numbers. The residues with no data shown are those with $\Delta ppm < 0.1$ ppm, those unassigned, or proline residues (no HN protons). The residues with comparable brown and green bars are mostly the residues in the extended binding surface. The residues with high brown bars but no green bars or negative green bars (chemical changes in different directions) are mostly the residues involved either directly or indirectly in the binding of the (pT)PVC residues as described in the text. B, Mapping of the SP44-234p-binding site on the Ki67 FHA structure. The residues were grouped into five groups using rainbow colors: red, $\Delta \geq 0.4$ ppm; orange, 0.3–0.4 ppm; yellow, 0.2–0.3 ppm; green, 0.13–0.2 ppm; and blue, <0.13 ppm. In addition, three residues, R38, R93 and S94 are defined as possible binding residues and are labeled in magenta because their NH peaks disappeared and were not found within ± 0.3 ppm compared to the peptide-free HSQC. Residues numbers are given for red and orange residues.

gradually, but it is specific, since it was not observed for other FHA domains such as FHA1 and FHA2 of Rad53 used as controls. The chemical shift changes are also summarized in Figure 4A. The results suggest that the FHA residues involved in the binding of hNIFK44 are generally the same as those involved in the binding of SP44-234p, with the exception of the residues involved in the binding of pThr. In consistence with this interpretation, mutations of hNIFK44 T234 to Ala or Glu caused little perturbation in binding as suggested by HSQC experiments. The K_d values of Ki67FHA binding to T234A and T234E mutants determined by SPR were nearly identical with that of wild-type hNIFK44 reported above, which support the interpretation of NMR results.

New view: FHA domain recognition involves three main factors

Previous results indicate that FHA domains recognize primarily the pT residue and the pT + 3 residue (D, I, or L). Our present results clearly indicate that Ki67 FHA does not follow the "+3 rule". On the basis of the results described above, we propose that three major factors control the interaction of FHA with its target protein: the pT residue, residues +1 to +3 (rather than residue +3 alone), and an extended binding surface, and that variation in the three factors likely leads to the great diversity in the function and specificity of FHA domains from different sources. The pT specificity is likely to be universal for all FHA domains biologically, though we have shown that Rad53 FHA2 can bind to pY peptides.[8] The residues involved in pT binding are likely to be conserved among all FHA domains. On the other hand, the other two factors are likely to vary greatly among different FHA domains. We further analyze and suggest possible interactions between Ki67 FHA and the SP44-234p peptide as follows.

Possible FHA residues interacting with pT

The sequential/structural alignments given in Figure 2 show that Ki67 FHA contains all the essential residues to interact with the pT phosphate group of peptides: R31, S45, K46 (or N, R in other FHA) and S65 (or T). Together with the chemical shift change results (Figure 4), we suggest that Ki67FHA interacts with the pT phosphate group using these residues.

Possible FHA residues interacting with residues +1 to +3

Previously it was proposed that the binding specificity of pTXXD in FHA1 is largely determined by R83 *via* a salt-bridge,[3,9] whereas that of pTXXI/L in FHA2 is by R617 and I681 *via* hydrophobic contacts.[10] Ki67 FHA has Val (V43) and Ile (I91) as the corresponding residues, suggesting that Ki67 FHA would prefer hydrophobic residues at the pT + 3 position. However, SP44-234p has a Cys at position pT + 3 although the neighboring residues are hydrophobic. It is interesting to mention that many hydrophobic residues are defined as the hot-binding ones (red and orange in Figure 4B): I10, V15, L28, F29, I37, I39, L41, V43, V44 and F95. This suggests that hydrophobic interactions may play a significant role in the binding specificity in Ki67 FHA.

Like the known structures of FHA complexes with short phosphopeptides, we think that the pTPVC residues exist in an extended conformation, with proline in the *trans* form. It is possible that FHA domains have evolved to bind the pTXXX motif in an extended conformation through backbone interactions, in addition to side-chain recognitions. A paired H-bonding between the backbone CO of pT + 1 and NH of pT + 3 and the side-chain $N^{\delta}H$ and CO^{δ} of a conserved Asn residue (N107 of FHA1,[3] N655 of FHA2,[10] N166 of Chk2,[11] and N67 of Ki67) are observed in all structures of pT peptide-FHA complexes, which suggests the H-bonding may play a role in keeping the extended backbone structure of the pT peptide.

Identification of an extended binding surface

The binding of Ki67 FHA with the non-phosphorylated fragment hNIFK44 and its phosphorylated counterpart SP44-234p share some common features, which can be considered to represent the extended binding surface. This is indicated by ca 20 peaks showing similar chemical shift changes in both complexes. Based on such analyses, this new interface includes the loop residues between β1 and β2, namely I10, K11, G14 and V15, and many residues of β10, a strand neighboring β1 and β2.

Such an extended binding surface, however, has not been observed for the Rad53 FHA domains, since their binding sites were mapped by use of short phosphopeptides. It remains to be established whether the extended binding surface also plays important roles for the Rad53 FHA domains when they bind to longer phosphopeptides. Analyses of the differences in the structures of free FHA domains indicate potentially significant differences in the length and structure of the loop between β9 and β10. In Ki67 FHA, β9 and β10 are connected by a single residue forming a tight β turn, whereas in FHA1 and FHA2, the β strands are longer and they use more (four to five) residues to form a loop, resulting in a protruding and lengthy loop. It is possible that the short and concise β9/10 loop structure may readily allow the pTXXX binding site to be continued to the back of the structure containing β1–2 and β10 to form a tertiary and extended binding surface. On the other hand, FHA1 and 2 are not expected to provide such a continuous binding surface due to the

protruding β9 and β10 structure, allowing only a short and localized binding surface. Therefore, in FHA1 and FHA2, the specificity of the +3 position may be more important than in Ki67 FHA to ensure tight binding for short pT peptides (ca 1 μM K_d). Note that Chk2 FHA also has a long β9/10 loop and also shows the +3 specificity, similar to FHA1 and FHA2.

An accessory binding surface was suggested by the study of Chk2 FHA[11] based on the deleterious effect on its binding to BRCA1, CDC25A[29] and p53[30] when I157, a distant residue from the pT recognition residues, is mutated to Thr. The extended binding surface observed in this study is different from the accessory binding surface in that it is an extended part of the pT recognition residues and provides additional affinity and specificity.

Conclusions and future perspectives

Complete solution structure of the Ki67FHA/SP44-234p complex is not yet available because it requires preparation of ^{13}C,^{15}N-labeled synthetic phosphopeptide SP44-234p. However, the structure of free Ki67FHA, along with extensive functional and structural analyses, has provided significant insight into the mechanism and function of FHA domains as described here. Furthermore, our results also raise the following important points which could lead to a broad impact in understanding how FHA domains mediate signal transduction processes in living cells. (a) Previous results indicate that the major mechanism of recognition by FHA domains is the pTXX(D/I/L) motif. This is apparently not the case with Ki67FHA. The result indicates great diversity in FHA specificity and calls for further studies of different FHA domains. (b) The results of FHA binding studies obtained by use of short phosphopeptides (10–15 residues) should be treated with caution. It is possible that the recognition site elucidated by use of short phosphopeptides is only chemically meaningful. This could explain why the results with short phosphopeptides have not led to unequivocal identification of the actual biological binding site of most other FHA domains. Our results suggest that longer peptide fragments or the whole binding protein should be used in future structure–function analyses of FHA domains. (c) Whether the proline in the (pT)P motif is *cis* or *trans* is a very interesting question, particularly since phosphorylation-dependent proline isomerization is an important mechanism to regulate the cell-cycle.[31,32] While we favor the *trans* form in the structural analysis described above, it remains to be verified by complete structural determination of the complex. Recently tumor suppressor p53 has been shown to be under the control of Pin1 through a conformational change.[33,34] The hNIFK was proposed to be involved in rRNA metabolism.[17] A potential regulation mechanism is through a conformational change at the (pT)P motif. Whether prolyl isomerases are involved in this process and how they are related to the recognition by Ki67FHA is important in understanding cell-cycle controls.

Materials and Methods

Gene cloning, expression and purification of Ki67FHA and hNIFK44

The cDNA clone AA262877 harboring a partial open reading frame corresponding to Ki67FHA, and BE790697 harboring a partial open reading frame for hNIFK, were purchased from Incyte Genomics. Primers were designed to amplify the coding region corresponding to residues 1–120 of Ki67 and residues 226–269 of hNIFK using polymerase chain reaction (PCR). The PCR product of FHA was subcloned into the *Bam*HI and *Not*I sites of a pEG vector, a derivative of pET29a with GST gene downstream T7 promoter to generate pEG-KiFHA. The pEG-KiFHA plasmid was transformed into *E. coli* BL21 (DE3) CodonPlus cell (from Stratagene) for overexpression. GST-Ki67FHA protein was purified by a glutathione-agarose column (Pharmacia). After GST was removed by thrombin digestion, Ki67 FHA was purified using S100 column (Pharmacia). The PCR product of hNIFK was subcloned into the *Bam*HI and *Xho*I sites of a pGB vector downstream from the GB1 gene to generate pGBhNIFK(226–269). The pGBhNIFK(226–269) plasmid was transformed into *E. coli* BL21 (DE3) CodonPlus cell (from Stratagene) for overexpression. The fragment (designated as hNIFK44) was purified through a nickel-NTA column (QIAgen) followed by an S100 column (Pharmacia).

Synthesis of phosphopeptides

The three phosphopeptides corresponding to hNIFK(226–269) with phosphorylation at 234, 238, and both positions were synthesized using in-house standard solid-phase synthesis techniques. The synthetic peptide SP44-234p was then obtained from Genemed. For the peptides synthesized using solid-phase synthesis, an extra homoserine lactone was appended at the C terminus as the result of CNBr cleavage.

The commercially obtained SP44-234p was affinity-purified: ^{15}N labeled Ki67FHA protein was mixed with ten times the amount of SP44-234p, adjusted to pH 7.5 with 1 M Hepes and incubated for one 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, 1 mM EDTA). Corresponding fractions were collected and examined with SDS/20% PAGE.

NMR spectroscopy and structure determination

All spectra were recorded at 17 °C on NMR samples containing ca 0.3–0.5 mM protein, 10 mM Tris–HCl buffer (pH 8.4 at 4 °C), 2 mM DTT and 1 mM EDTA using Bruker DRX800, DRX600 and DMX600 spectrometers. Later the buffer condition was changed to 5 mM Hepes (pH 7.5), 2 mM DTT, 1 mM EDTA, 150 mM NaCl for optimal resolution. ^{1}H, ^{13}C and ^{15}N sequential assignments were achieved using heteronuclear multi-dimensional experiments.[35,36] Inter-proton distance restraints were derived from 2D NOESY and 3D ^{13}C-edited NOE and ^{15}N-edited NOE experiments, which were recorded

using mixing times of 80–100 ms. Approximate inter-proton distance restraints were grouped into four distance ranges: 1.8–2.7 Å, 1.8–3.3 Å, 1.8–5.0 Å and 1.8–6.0 Å, corresponding to strong, medium, weak and very weak NOEs, respectively. In addition, 0.5 Å was added to the upper limit of inter-proton distance restraints involving methyl groups. For hydrogen bonds, distance restraints of 1.5–2.5 Å (H–O) and 2.4–3.5 Å (N–O) were employed. Slowly exchanging amide protons were detected from the ^{1}H-^{15}N HSQC spectrum of lyophilized protein freshly dissolved in $^{2}H_2O$ at 10 °C. The backbone torsion angle (ϕ and ψ) restraints were obtained using TALOS.[37] The $\chi 1$ torsion angle restraints were derived by quantitative *J*-correlation spectroscopy.[38] The experimentally determined distance and torsion angle restraints (Table 1) were applied in a simulated annealing protocol using the NIH version[24] of the program X-PLOR.[25] An ensemble of 100 structures was generated and the final 22, the major structural family with lowest X-PLOR target function values, were selected. The regularized mean structure was calculated by restrained energy minimization of the mean structure of the final 22 structures.

Library screening, HSQC NMR and surface plasmon resonance

HSQC NMR was performed as described,[9,10] as were library screening and SPR.[4]

Coordinates

Coordinates have been deposited at the RCSB Protein Data Bank, with accession numbers PDB ID 1R21 and RCSB ID RCSB020341. Chemical shifts have been deposited at the BioMagResBank (BMRB), with accession number 5959.

Acknowledgements

We thank Dr Yoneda of Osaka University for kindly providing us the hNIFK gene, the Hklp2 gene and Hklp2 mutant plasmids; Dr Angela Gronenborn of NIH for allowing the final stage of structural determination to be performed in her laboratory by I.J.B.; and Suganya Yongkiettrakul of our laboratory for useful discussion. This work was supported by NIH grant CA87031.

References

1. Liao, H., Byeon, I. J. & Tsai, M. D. (1999). Structure and function of a new phosphopeptide-binding domain containing the FHA2 of Rad53. *J. Mol. Biol.* **294**, 1041–1049.
2. Durocher, D., Henckel, J., Fersht, A. R. & Jackson, S. P. (1999). The FHA domain is a modular phosphopeptide recognition motif. *Mol. Cell*, **4**, 387–394.
3. Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P. *et al.* (2000). The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. *Mol. Cell*, **6**, 1169–1182.
4. Liao, H., Yuan, C., Su, M. I., Yongkiettrakul, S., Qin, D., Li, H. *et al.* (2000). Structure of the FHA1 domain of yeast Rad53 and identification of binding sites for both FHA1 and its target protein Rad9. *J. Mol. Biol.* **304**, 941–951.
5. Li, J., Smith, G. P. & Walker, J. C. (1999). Kinase interaction domain of kinase-associated protein phosphatase, a phosphoprotein-binding domain. *Proc. Natl Acad. Sci. USA*, **96**, 7821–7826.
6. Durocher, D. & Jackson, S. P. (2002). The FAH domain. *FEBS Letters*, **513**, 58–66.
7. Stavridi, E. S., Huyen, Y., Loreto, I. R., Scolnick, D. M., Halazonetis, T. D., Pavletich, N. P. & Jeffrey, P. D. (2002). Crystal structure of the FHA domain of the Chfr mitotic checkpoint protein and its complex with tungstate. *Structure*, **10**, 891–899.
8. Wang, P., Byeon, I. J., Liao, H., Beebe, K. D., Yongkiettrakul, S., Pei, D. *et al.* (2000). II. Structure and specificity of the interaction between the FHA2 domain of Rad53 and phosphotyrosyl peptides. *J. Mol. Biol.* **302**, 927–940.
9. Yuan, C., Yongkiettrakul, S., Byeon, I. J., Zhou, S. & Tsai, M. D. (2001). Solution structures of two FHA1-phosphothreonine peptide complexes provide insight into the structural basis of the ligand specificity of FHA1 from yeast Rad53. *J. Mol. Biol.* **314**, 563–575.
10. Byeon, I. J., Yongkiettrakul, S. & Tsai, M. D. (2001). Solution structure of the yeast Rad53 FHA2 complexed with a phosphothreonine peptide pTXXL: comparison with the structures of FHA2-pYXL and FHA1-pTXXD complexes. *J. Mol. Biol.* **314**, 577–588.
11. Li, J., Williams, B. L., Haire, L. F., Goldberg, M., Wilker, E., Durocher, D. *et al.* (2002). Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2. *Mol. Cell*, **9**, 1045–1054.
12. Stone, J. M., Collinge, M. A., Smith, R. D., Horn, M. A. & Walker, J. C. (1994). Interaction of a protein phosphatase with an *Arabidopsis* serine–threonine receptor kinase. *Science*, **266**, 793–795.
13. Schwartz, M. F., Duong, J. K., Sun, Z., Morrow, J. S., Pradhan, D. & Stern, D. F. (2002). Rad9 phosphorylation sites couple Rad53 to the *Saccharomyces cerevisiae* DNA damage checkpoint. *Mol. Cell*, **9**, 1055–1065.
14. Ahn, J. Y., Li, X., Davis, H. L. & Canman, C. E. (2002). Phosphorylation of threonine 68 promotes oligomerization and autophosphorylation of the Chk2 protein kinase *via* the forkhead-associated domain. *J. Biol. Chem.* **277**, 19389–19395.
15. Hammet, A., Pike, B. L. & Heierhorst, J. (2002). Post-transcriptional regulation of the RAD5 DNA repair gene by the Dun1 kinase and the Pan2–Pan3 poly(A)-nuclease complex contributes to survival of replication blocks. *J. Biol. Chem.* **277**, 22469–22474.
16. Sueishi, M., Takagi, M. & Yoneda, Y. (2000). The forkhead-associated domain of Ki-67 antigen interacts with the novel kinesin-like protein Hklp2. *J. Biol. Chem.* **275**, 28888–28892.
17. Takagi, M., Sueishi, M., Saiwaki, T., Kametaka, A. & Yoneda, Y. (2001). A novel nucleolar protein, NIFK, interacts with the forkhead associated domain of Ki-67 antigen in mitosis. *J. Biol. Chem.* **276**, 25386–25391.
18. Gerdes, J., Schwab, U., Lemke, H. & Stein, H. (1983). Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. *Int. J. Cancer*, **31**, 13–20.
19. Gerdes, J., Lemke, H., Baisch, H., Wacker, H. H., Schwab, U. & Stein, H. (1984). Cell cycle analysis of

a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *J. Immunol.* **133**, 1710–1715.
20. Brown, D. C. & Gatter, K. C. (2002). Ki67 protein: the immaculate deception? *Histopathology*, **40**, 2–11.
21. Kreitz, S., Fackelmayer, F. O., Gerdes, J. & Knippers, R. (2000). The proliferation-specific human Ki-67 protein is a constituent of compact chromatin. *Expt. Cell Res.* **261**, 284–292.
22. Schluter, C., Duchrow, M., Wohlenberg, C., Becker, M. H., Key, G., Flad, H. D. & Gerdes, J. (1993). The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins. *J. Cell Biol.* **123**, 513–522.
23. Scholzen, T., Endl, E., Wohlenberg, C., van der Sar, S., Cowell, I. G., Gerdes, J. & Singh, P. B. (2002). The Ki-67 protein interacts with members of the heterochromatin protein 1 (HP1) family: a potential role in the regulation of higher-order chromatin structure. *J. Pathol.* **196**, 135–144.
24. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Marius Clore, G. (2003). The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73.
25. Brünger, A. T. (1992). *X-PLOR: Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT
26. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.
27. Zhou, P., Lugovskoy, A. A. & Wagner, G. (2001). A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins. *J. Biomol. NMR*, **20**, 11–14.
28. Nigg, E. A. (1995). Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. *Bioessays*, **17**, 471–480.
29. Falck, J., Mailand, N., Syljuasen, R. G., Bartek, J. & Lukas, J. (2001). The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis. *Nature*, **410**, 842–847.
30. Falck, J., Lukas, C., Protopopova, M., Lukas, J., Selivanova, G. & Bartek, J. (2001). Functional impact of concomitant *versus* alternative defects in the Chk2-p53 tumour suppressor pathway. *Oncogene*, **20**, 5503–5510.
31. Lu, K. P. (2000). Phosphorylation-dependent prolyl isomerization: a novel cell cycle regulatory mechanism. *Prog. Cell Cycle Res.* **4**, 83–96.
32. Zhou, X. Z., Lu, P. J., Wulf, G. & Lu, K. P. (1999). Phosphorylation-dependent prolyl isomerization: a novel signaling regulatory mechanism. *Cell. Mol. Life Sci.* **56**, 788–806.
33. Zacchi, P., Gostissa, M., Uchida, T., Salvagno, C., Avolio, F., Volinia, S. *et al.* (2002). The prolyl isomerase Pin1 reveals a mechanism to control p53 functions after genotoxic insults. *Nature*, **419**, 853–857.
34. Zheng, H., You, H., Zhou, X. Z., Murray, S. A., Uchida, T., Wulf, G. *et al.* (2002). The prolyl isomerase Pin1 is a regulator of p53 in genotoxic response. *Nature*, **419**, 849–853.
35. Clore, G. M. & Gronenborn, A. M. (1998). Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol.* **16**, 22–34.
36. Clore, G. M. & Gronenborn, A. M. (1991). Two-, three-, and four-dimensional NMR methods for obtaining larger and more precise three-dimensional structures of proteins in solution. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 29–63.
37. Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.
38. Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R. & Zhu, G. (1994). Measurement of homo- and heteronuclear *J* couplings from quantitative *J* correlation. *Methods Enzymol.* **239**, 79–105.
39. Farmer, B. T., II, Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchunas, J., Jr *et al.* (1996). Localizing the NADP + binding site on the MurB enzyme by NMR. *Nature Struct. Biol.* **3**, 995–997.

Edited by P. Wright

(Received 16 July 2003; received in revised form 8 October 2003; accepted 9 October 2003)