The Tudor Tandem of 53BP1: A New Structural Motif Involved in DNA and RG-Rich Peptide Binding

Gaëlle Charier,¹ Joël Couprie,¹
Béatrice Alpha-Bazin,¹ Vincent Meyer,²
Eric Quéméneur,² Raphael Guérois,²
Isabelle Calelmbaut,¹ Bernard Gilquin,¹
and Sophie Zinn-Justin,¹,*
¹Département d'Ingénierie et d'Etudes des Protéines
²Département de Biologie Joliot-Curie CEA SACLAY
91191 Gif-sur-Yvette France
³Département d'Ingénierie et d'Etudes des Protéines
CEA VALRHRO
30207 Bagnols-sur-Ceze France
⁴Département de Biologie Structurale
LMCP, CNRS UMR 7590
4 Place Jussieu
75252 Paris Cedex 05 France

Summary

53BP1 is a key transducer of the DNA damage checkpoint signal, which is required for phosphorylation of a subset of ATM substrates and p53 accumulation. After cell irradiation, the 53BP1 N-terminal region is phosphorylated. Its two C-terminal BRCT motifs interact with p53. Its central region is required and sufficient for 53BP1 foci formation at DNA strand breaks and for 53BP1 binding to the kinetochore. It contains an RG-rich segment and interacts with DNA in vitro. Here we show that the major globular domain of the 53BP1 central region adopts a new structural motif composed of two tightly packed Tudor domains and a C-terminal α helix. A unique surface essentially located on the first Tudor domain is involved in the binding to 53BP1 RG-rich sequence and to DNA, suggesting that the Tudor tandem can act as an adaptor mediating intramolecular as well as intermolecular protein-protein interactions and protein-nucleic acid associations.

Introduction

The DNA double-strand break (DSB) is one of the most serious damages for cells. Such an event may result in rearrangement or loss of genetic information, and lead to cell death or carcinogenesis. DSBS arise during normal endogenous processes of cells (DNA replication, meiosis, V(D)J recombination), but can also be induced by ionizing radiations (IR). In general, cells do not enter S or M phase before the DNA lesions are properly repaired due to the action of the DNA damage checkpoint (Hartwell and Weinert, 1989). The sensitivity of cancer cells to DNA-damaging agents is explained by the fact that cancer cells have often lost some aspects of their checkpoint functions, thus acquiring a higher rate of genomic evolution and a growth advantage (Hartwell and Kastan, 1994).

The mammalian protein p53 Binding Protein 1 (53BP1) was originally identified in a yeast two-hybrid screen as a protein that interacts with p53 DNA binding domain through its two C-terminal BRCT motifs (Iwabuchi et al., 1994; Joo et al., 2002; Derbyshire et al., 2002). BRCT domains are 100–150 residue motifs found in a large number of proteins involved in the cellular responses to DNA damages (Bork et al., 1997; Callebaut and Mornon, 1997a; Clapperton et al., 2004; Williams et al., 2004). Consistently, upon exposure to IR, 53BP1 was shown to rapidly form foci at the sites of DSBs and to be phosphorylated via ATM, a central signaling kinase of the response to DSBs (Schultz et al., 2000; Rappold et al., 2001; Anderson et al., 2001). Further experiments using small-interfering RNA or gene targeting to knockdown 53BP1 expression have shown that 53BP1 is required for the accumulation of p53, for the intra-S-phase and G2-M checkpoints, and for the phosphorylation of a subset of ATM substrates such as Chk2, BRCA1, and SMC1 in response to IR damage (Wang et al., 2002). These results indicate that 53BP1 is a central mediator of the DNA damage checkpoint.

Radiation-induced phosphorylation of 53BP1 N terminus by ATM kinase is not essential for 53BP1 foci formation. However, the region 1052–1639 of 53BP1, comprised between the N-terminal phosphorylated region and the BRCT domains, is required and sufficient for the recruitment of 53BP1 to DNA strand breaks. In vitro pull-down assays revealed that an overlapping region, comprising residues 956–1354, binds to phosphorylated but not unphosphorylated histone H2AX (Ward et al., 2003). Moreover, several experiments suggested a direct interaction between 53BP1 and phosphorylated H2AX in vivo. First, after irradiation 53BP1 colocalizes with phosphorylated H2AX in megabase regions surrounding the sites of DNA breaks (Schultz et al., 2000; Rappold et al., 2001). Second, phosphorylation of H2AX at serine 140 is critical for efficient 53BP1 foci formation (Ward et al., 2003). Third, H2AX-deficient cells lack normal 53BP1 foci formation and lack 53BP1-deficient cells manifest a G2-M checkpoint defect after exposure to low doses of IR (Fernandez-Capetillo et al., 2002). Thus, recruitment of the central region of 53BP1 to phosphorylated H2AX foci seems to be a crucial step for the initiation and maintenance of 53BP1 foci formation.

53BP1 directly interacts with DNA in vitro. The region 1052–1709 of 53BP1, largely overlapping the region involved in foci formation, binds to a double-strand 10 bp oligonucleotide in blot overlay assays. Furthermore, electrophoresis mobility shift assay experiments suggested that 53BP1 possesses at least two DNA binding domains; both regions 1319–1480 and 1480–1616 bind to linear double- and single-strand DNA substrates. The region 1480–1616 also promotes DNA end joining by the
Table 1. Structural Statistics for the Mouse 53BP1 (1469-1591) Fragment

<table>
<thead>
<tr>
<th>Number of violations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>nOe distance restraints &gt; 0.5 Å</td>
<td>0</td>
</tr>
<tr>
<td>Dihedral restraints &gt; 10°</td>
<td>0</td>
</tr>
<tr>
<td>Experimental restraints</td>
<td></td>
</tr>
<tr>
<td>Distance restraints (Å)</td>
<td>2397 (rmsd: 0.04 ± 0.001)</td>
</tr>
<tr>
<td>Unambiguous</td>
<td>2081</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>256</td>
</tr>
<tr>
<td>Hydrogen bonds</td>
<td>60</td>
</tr>
<tr>
<td>Dihedral restraints (°)</td>
<td>206 (rmsd: 1.2 ± 0.06)</td>
</tr>
<tr>
<td>Rms deviation from idealized covalent geometry</td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.016 ± 0.0003</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>3.6 ± 0.04</td>
</tr>
<tr>
<td>Improvers (°)</td>
<td>3.0 ± 0.2</td>
</tr>
<tr>
<td>Energy (kcal/mol)*</td>
<td></td>
</tr>
<tr>
<td>van der Waals</td>
<td>171 ± 11</td>
</tr>
<tr>
<td>Electrostatic</td>
<td>-408 ± 22</td>
</tr>
<tr>
<td>Ramachandran plot (%)</td>
<td></td>
</tr>
<tr>
<td>Most favored regions</td>
<td>79.8</td>
</tr>
<tr>
<td>Additionally allowed regions</td>
<td>18.4</td>
</tr>
<tr>
<td>Coordinate precision (residues 7–129)</td>
<td></td>
</tr>
<tr>
<td>Backbone atoms</td>
<td>0.82 ± 0.2</td>
</tr>
<tr>
<td>Heavy atoms</td>
<td>1.3 ± 0.16</td>
</tr>
</tbody>
</table>

*The van der Waals energy is calculated with a Lennard-Jones potential. The electrostatic energy is calculated with no net charge on side chain atoms and a distance-gated dielectric constant. CHARMM22 parameters were used.

DNA ligase IV/Xrcc4 complex, which is involved in the nonhomologous end joining (NHEJ) pathway of DSB repair in mammalian cells. Thus, the central region of 53BP1 might directly participate to the repair of DNA DSBs (Iwabuchi et al., 2003).

Finally, 53BP1 localizes to kinetochores and is hyper-phosphorylated during mitosis under conditions where the spindle checkpoint is activated. The minimal 53BP1 kinetochore binding domain resides again in a region located between the N terminus phosphorylated in an ATM-dependent manner and the BRCT domains. It corresponds to residues 1220–1601 in mouse 53BP1 and 1235–1616 in human 53BP1. Thus, the central region of 53BP1 is involved both in the DNA damage response and in the signaling at the kinetochore during mitosis (Jullien et al., 2002).

We have analyzed the sequence of 53BP1 between residues 956 and 1709. It contains several low complexity regions. However, a segment showing globular domain characteristics (about one-third of hydrophobic residues) is found between residues 1480 and 1616. This segment is predicted to contain a Tudor domain, a conserved motif of 50 amino acids found in several RNA-associated proteins (Ponting, 1997; Maurer-Stroh et al., 2003; Callebaut and Mornon, 1997b). A Tudor domain was also found in the SMN protein, a protein linked to spinal muscular atrophy. Its 3D structure was solved by NMR (Selenko et al., 2001) and by X-ray crystallography (Sprangers et al., 2003). The Tudor domain of SMN binds to symmetrically dimethylated arginines (sDMA) of RG-rich sequences found in Sm protein C-terminal tails (Selenko et al., 2001). Interestingly, an RG-rich motif of 8 residues (RGRGRRGR) is highly conserved within the three known 53BP1 sequences (human, mouse, Xenopus). This motif is located 80 residues before the predicted Tudor domain in mouse 53BP1.

To gain molecular insight into the functional role of the region of 53BP1 comprised between the N-terminal phosphorylated region and the BRCT domains, we have solved the three-dimensional structure of the segment 1463–1617 of mouse 53BP1 (53BP1TT) using NMR. Here we show that this segment, which is 99% identical to region 1478–1632 of human 53BP1, surprisingly comprises not only one but two Tudor domains. The first Tudor domain corresponds to the predicted one and presents a cavity analogous to the sDMA binding region of SMN. At this point, NMR measurements were used to characterize the interaction of 53BP1TT with an RGRGRRGR peptide symmetrically dimethylated or nonmethylated on the arginine residues. The DNA binding region of 53BP1TT was also mapped by NMR. These different interactions were analyzed, in order to understand their specificity and their potential mode of regulation. Finally, searching for similar functional sites in the other detected tandem Tudor sequences was carried out in order to understand the functional role of this new structural family.

Results

53BP1TT Folds into Three Structural Motifs

The solution structure of 53BP1TT was determined by heteronuclear double and triple resonance NMR spectroscopy (Table 1). Coordinates and NMR restraints were deposited at the Protein Data Bank (entry code: 1SSF, http://www.rcsb.org/pdb/). In the following, the supplementary glycine resulting from the biomolecular construction is numbered 0 and the last residue 155. 53BP1TT is folded between residues 8 and 129. Indeed, positive 1H-15N nOe were measured between residues 8 and 129 (data not shown), and the backbone root-mean-square deviation (rmsd) calculated on this fragment with respect to the mean coordinates yields 0.85 Å. The 3D structure of region 8–129 is constituted of three structural motifs (Figures 1A and 1B). The first and second motifs correspond to residues 12–57 and 64–112, respectively. They both adopt a β-barrel-like fold and are connected by a 6 residue linker. The third motif essen-
Figure 1. Structure of the Mouse 53BP1 (1463-1617) Region

(A) Stereoview of the backbone atoms (N, Cα, C') for residues 7–129 (corresponding to residues 1469–1591 of mouse 53BP1) of a set of ten superimposed structures. Secondary structures are colored in blue for β strands of Tudor motif 1, green for β strands of Tudor motif 2, orange for 3_10 helices of motifs 1 and 2, and red for the C-terminal helix.

(B) Ribbon representation of residues 7–129 colored as in (A).

(C) Superimposition of Tudor motifs 1 and 2 on backbone atoms (N, Cα, C'). Dark blue and green correspond to fitted residues (12–17 on 65–70, 24–31 on 76–83, 37–41 on 92–96, and 43–57 on 97–111). Light blue and green correspond to the other residues.

This suggests that the minor conformation of 53BP1_TT corresponds to a cis conformer of Pro127. Our structural analysis focuses on the major form of the 53BP1_TT fragment.

Motifs 1 and 2 Adopt Similar β-Barrel-like Structures
As shown in Figure 1C, motifs 1 and 2 are mostly superimposable, except for two structurally variable loops located between the first and the second strands and between the second and the third strands of each motif.
Figure 2. Comparison with the Tudor Domain of SMN Protein

(A) Sequence alignment of the Tudor domain of SMN with the two Tudor domains of 53BP1. Secondary structure elements are indicated in blue for β strands and in red for the 3_10 helix. "Consensus tudor" indicates a consensus sequence obtained from the SMART database: “s” means small and medium size residues at 80%; “h” stands for hydrophobic residues at 80%; “l” means aliphatic residues at 80%; “@” means aromatic residues at 50% or hydrophobic at 65%; and “—” stands for Asp at 65%.


(C) Comparison of the aromatic clusters of SMN and 53BP1. SMN Tudor domain is colored in red, and 53BP1 Tudor 1 in blue. Residues named in the text are labeled.

The rmsd calculated between the backbones of the two motifs except for the variable loops yields 1.9 Å. Each of these motifs is constituted of two strongly bent antiparallel β sheets: β sheets β1, β2, β5 and β2’, β3, β4 form motif 1; β sheets β6, β7, β10, and β7’, β8, β9 form motif 2. The two large β strands β2 and β7 are bent by a 90° angle due to the presence of classical β-bulges at positions (30,31; 39) and (82,83; 94) respectively. β2’ and β7’ represent these strands after their respective β-bulges. Furthermore, β4 is connected to β5 and β9 to β10 by 3_10 helices of three residues at positions 52–54 and 106–108, respectively, such that the angle between the two consecutive strands is about 90° in both cases. Thus, the two antiparallel β sheets are roughly perpendicular in both motifs. The β-barrel-like structures are closed by interactions of β1 with β5 and β6 with β10, respectively.

Structural alignment between the two motifs yields 18% of sequence identity (Figure 2A). The conserved residues are either small and medium size amino acids
Motifs 1 and 2 Correspond to a New Tudor Tandem Fold
A DALI search using motif 1 gives NusG (PDB code: 1M1G) and SMN (PDB code: 1G5V) as the two closest matches. These structures are also found in second and fourth positions when motif 2 is proposed to DALI. Clearly, the β sheet arrangement of motifs 1 and 2 is similar to that found in the Tudor fold of SMN (Selenko et al., 2001) and the Tudor-like fold of NusG (Steiner et al., 2002). If the structures of motifs 1 and 2 are superimposed onto the structure of the SMN Tudor domain, the backbone rmsd calculated on the five β strands yields 2.0 and 1.5 Å, respectively (Figure 2B). All the small and medium size amino acids conserved between motifs 1 and 2 of 53BP1 are also small and medium size amino acids in the SMN Tudor domain (Figure 2A). However, the aromatic residues conserved between motifs 1 and 2 are not found in the SMN Tudor domain, suggesting that these residues are probably not crucial for the Tudor fold.

Motif 2 was not predicted as a Tudor domain by SMART (Letunic et al., 2004). The sequence alignment of Figure 2A shows which of the 16 best-conserved positions characterizing a Tudor domain are conserved in SMN and 53BP1 motifs 1 and 2. Clearly, if motif 1 presents most of the markers of a Tudor domain, motif 2 only presents two-thirds of them, and those essentially correspond to buried small or hydrophobic residues. In particular, the solvent-exposed aromatic patch found in SMN and motif 1 is absent in motif 2.

The DALI search using 53BP1 Tudor motifs further revealed a structural analogy with PWWP and MBT domains, evolutionary related to the Tudor domains within the “Royal family” described by Ponting and coworkers (Maurer-Stroh et al., 2003). Finally, the DALI search pointed out the structural analogy between the two first 53BP1 motifs and other SH3-like barrels of various biological functions.

53BP1TT Interacts with the 53BP1 Arg-Gly-Rich Sequence
Alignment of the three known 53BP1 sequences shows that a RGRGRRGR stretch is highly conserved 80 amino acids upstream from the Tudor tandem (Figure 3A). Mouse 53BP1 is methylated in vivo probably through this RG-rich stretch (Y. Adachi, personal communication). To investigate the potential interaction of nonmethylated or symmetrically dimethylated RGRGRRGR peptides with 53BP1TT, we performed NMR titrations. We added each peptide to the 15N-labeled 53BP1TT sample and followed the chemical shift perturbations of 53BP1TT residues by recording 1H–15N HSQC experiments. Figure 3B presents an overlay of the 1H–15N HSQC of 53BP1TT free and saturated with the nonmethylated RG-rich peptide. A similar overlay was obtained after saturation with the methylated peptide (data not shown).

Clearly, 53BP1TT binds to the RGRGRRGR peptides, whether the arginines are nonmethylated or symmetrically dimethylated. The exchange rate between free and bound protein conformations is fast, suggesting that the affinity between 53BP1TT and the peptides is relatively low. Fitting the variation of weighted chemical shift displacements against the peptide concentration yielded a Kd value of 3.9 ± 2.2 mM and 5.8 ± 2.7 mM for the methylated and the nonmethylated peptides, respectively. The estimation of the affinity for the nonmethylated peptide was confirmed by fluorescence measurements (data not show).

Chemical shift changes due to peptide addition are shown in Supplemental Figures S1B and S1C. Residues significantly involved in the interaction with the nonmethylated and the methylated peptides are described in Figures 3C and 3D, respectively. The two mapped interaction surfaces are highly similar. The central peptide binding region (defined by weighted chemical shift displacements higher than 0.05 ppm) contains Trp18, Ser19, Asn21, Gly22, Tyr23 (in loop β1β2), Asp44, Tyr46 (in loop β3β4), Glu47, Cys48 (in strand β4), and Ala69 (in strand β6). Thus, both peptides bind in a cavity essentially composed of three aromatic residues and two negatively charged residues, located between loops β1β2, β3β4 and strand β4 in motif 1. Other residues surrounding this cavity also show substantial chemical perturbations (comprised between 0.025 and 0.05 ppm; Figures 3C and 3D). In particular, several residues of motif 2, located in loop β6β7 and strand β10, are affected by the addition of RG-rich peptides, suggesting that both motifs 1 and 2 are involved in peptide binding.
Figure 3. Interaction of 53BP1 with Two RG-Rich Peptides

(A) Sequence alignment of human, mouse, and Xenopus 53BP1 in the RG-rich region (mouse numbering).

(B) Overlay of the $^{1}H$-$^{15}N$ HSQC spectra obtained for protein free at 0.6 mM (red) and for protein saturated with the nonmethylated RG-rich peptide (blue).

(C) Interaction with the nonmethylated RG-rich peptide. Left: ribbon representation, with Tudor 1 in blue and Tudor 2 in green. The side chains of residues whose NH-group chemical shift perturbation ($|\Delta \delta(1H)| + 0.1 \times |\Delta \delta(^{15}N)|$) is higher than 0.05 ppm (2 standard deviations) are colored in red and labeled. The backbone is colored in orange if the corresponding chemical shift perturbation is higher than 0.025 ppm. Right: the interaction surface is shown in the same orientation, with the same colors for perturbed amino acids.

(D) Interaction with the symmetrically dimethylated RG-rich peptide. Color codes and orientations are the same as in (C).
In order to characterize the specificity of these interactions, we tested the binding of 53BP1 \(^{TT}\) to L-arginine alone and to a KG-rich peptide where all the arginines of the initial peptide were substituted by lysines. No substantial chemical shift (>0.02 ppm) variation of 53BP1 \(^{TT}\) residues was observed after addition of a 50- and 54-fold molar excess of L-arginine and KG-rich peptide, respectively (data not shown). Furthermore, in order to evaluate the potential role of the residues flanking the RG-rich segment in the binding of 53BP1 RG-rich region to 53BP1 \(^{TT}\), a 24 residue peptide comprising the RG-rich sequence and relatively well conserved in the three 53BP1 sequences (GGKAPVTPGRGRGRRPPSR TTGT, nonmethylated, cf. Figure 3A) was added to an NMR sample of 53BP1 \(^{TT}\). The longer RG-rich peptide also binds to 53BP1 \(^{TT}\) with an affinity in the millimolar range: the observed dissociation constant yields 4.7 ± 2.0 mM; the corresponding binding region (weighted chemical shift displacements higher than 0.05 ppm) includes Ala16 (in \(\beta_1\)), Trp18, Ser19, Gly22, Tyr23 (in loop \(\beta_1\beta_2\)), Phe24, Ser26 (in \(\beta_2\)), Phe42 (in \(\beta_3\)), Gly45, Tyr46 (in \(\beta_3\beta_4\)), Glu47, Cys48 (in \(\beta_4\)), Asp59 (in the linker), and Ser71 and Asp73 (in \(\beta_6\beta_7\)) (Supplemental Figure S1D).

53BP1 \(^{TT}\) Is Also Involved in DNA Binding

Doherty and coworkers recently showed that the human 53BP1 (1052-1709) fragment binds to a 10 bp oligonucleotide (5'-AACTCAGGT-3') (Iwabuchi et al., 2003). This fragment overlaps the mouse 53BP1 (1463-1617) fragment. The binding of 53BP1 \(^{TT}\) to this 10 bp oligonucleotide was investigated using NMR. As shown on the \(^{1}H\)-\(^{15}N\) HSQC spectra overlay of Figure 4A, addition of DNA induces substantial chemical shift displacements of a set of NH signals, indicating that the 53BP1 \(^{TT}\) is indeed able to bind to DNA. The dissociation constant estimated for this interaction yields 0.46 ± 0.03 mM.

Figure 4B shows the residues whose NMR signals are strongly perturbed after addition of a 3-fold molar excess of DNA. All chemical shift variations are shown in Supplemental Figure S1A. The central DNA binding region (defined by weighted chemical shift displacements higher than 0.3 ppm) comprises Trp18, Asn21, Gly22, Tyr23, Asp44, and Glu47. Thus DNA binds to 53BP1 \(^{TT}\) through a surface formed by loops \(\beta_1\beta_2\) and \(\beta_3\beta_4\) in motif 1. Other residues surrounding this central region also show substantial chemical shift displacements (comprised between 0.05 and 0.30 ppm). A larger DNA binding surface can thus be defined involving both Tudor motifs (loops \(\beta_1\beta_2\), \(\beta_3\beta_4\), and \(\beta_6\beta_7\) and strands \(\beta_4\), \(\beta_6\), and \(\beta_{10}\)). Figure 4B shows this large DNA binding surface. Surprisingly, it is similar to the region involved in the binding to RG-rich peptides.

Discussion

The RG-Rich Peptide Binding Site of 53BP1 \(^{TT}\) Is Essentially Composed of Variable Loops of an SH3-like Fold and Is Largely Superimposable to the Functional Site of SMN

Tudor domains belong to the SH3-like superfamily. SH3-like domains are often part of modular proteins showing a high functional versatility from signal transduction to nucleic acid binding. Their topology is characterized by a five \(\beta\) strand motif and four connecting loops. Large variations in sequence, length, and flexibility of the three first loops are responsible for the functional specificity of the SH3-like domains. In contrast, the fourth loop, which comprises a short \(\beta_4\) helix, is structurally very conserved and plays a key role in preserving the SH3 fold (Dalgarno et al., 1997). In the case of SH3 domains of protein kinases, loop \(\beta_1\beta_2\) is particularly large and is involved with loop \(\beta_2\beta_3\), in proline-rich peptide binding. In the case of 53BP1 motif 1, loop \(\beta_1\beta_2\) is also involved in the interaction with the targeted RG-rich peptides, but together with the contiguous loop \(\beta_3\beta_4\). In both cases, an aromatic patch is involved in the binding, but this patch is close to loop \(\beta_2\beta_3\) in SH3 domains and close to loop \(\beta_3\beta_4\) in 53BP1.

In contrast, the RG-rich peptide binding site of 53BP1 motif 1 is similar to the binding site of the SMN Tudor domain to Sm protein RG-rich tails. Indeed, using NMR titrations, Sattler and coworkers showed that the amide groups strongly involved in the SMN/poly RG interaction belong to a cluster of conserved aromatic residues: Trp102 (in \(\beta_1\beta_2\)), Tyr109 (in \(\beta_2\)), Tyr127 (in \(\beta_3\)), and Tyr130 (in \(\beta_3\beta_4\)). A negatively charged amino acid, Glu134, is also important since the binding to Sm proteins is abolished when this amino acid is mutated to a lysine (Selenko et al., 2001). In the alignment of 53BP1 motif 1 with SMN (Figure 2A), these aromatic residues correspond to Trp18, Tyr25, Phe42, and Asp44, respectively. Glu134 in SMN is aligned with Cys48 in 53BP1. Three and four of these five 53BP1 residues are perturbed by the addition of the unmethylated and the methylated RG-rich peptides, respectively. Tyr25 of 53BP1 is not involved in the binding, but the close Tyr23 is affected by addition of the peptides. Thus, the RG-rich peptide binding site of 53BP1 is mostly superimposable to the SMN functional site, but it is larger because located both in motifs 1 and 2 and involving the more distant Tyr23 (Figure 2C).

Arginine Methylation of the RG-Rich Peptide or Addition of 8 Residues from the 53BP1 Sequence on Both Sides Do Not Influence the Binding to 53BP1 \(^{TT}\)

SMN and 53BP1 use similar aromatic pockets to bind with a relatively low affinity to nonmethylated and arginine symmetrically dimethylated RG-rich peptides. Sattler and coworkers showed that arginine methylation increases the affinity of the SMN Tudor domain for RG repeats contained in the C terminus of Sm proteins (Spranger et al., 2003). However, SMN efficiently binds some other RG-rich containing substrates, such as fibrillarin, nucleolin GAR1, Sm core proteins hnRNP Q, R, and U, RNA helicase A, and p80 coilin (Young et al., 2003). In their paper, Lorson and coworkers reported that SMN Tudor domain binds to the nonmethylated RG-rich peptide from the Ewing’s sarcoma protein (EWS) with a dissociation constant of 3 mM and to the symmetrically dimethylated peptide with a dissociation constant of 5 mM (Young et al., 2003). We report similar values for the binding of 53BP1 \(^{TT}\) to nonmethylated and symmetrically dimethylated RG-rich peptides (affinity
Figure 4. Interaction of 53BP1 TT with DNA

(A) Overlay of the $^1$H-$^{15}$N HSQC spectra obtained for protein free at 0.3 mM (red) and for protein saturated with a 10 bp oligonucleotide. Residues whose chemical shift perturbation ($|\Delta \delta(\text{H})| + 0.1 \times |\Delta \delta(\text{N})|$) is higher than 0.15 ppm are labeled in orange and in red if it is superior to 0.3 ppm.

(B) Left: ribbon representation, with Tudor 1 in blue and Tudor 2 in green. The side chains of residues whose NH-group chemical shift perturbation is higher than 0.3 ppm (2 standard deviations) are colored in red and labeled. The backbone is colored in orange if the chemical shift perturbation is higher than 0.15 ppm and in yellow if it is higher than 0.05 ppm. Orientation is the same as in Figure 2. Right: the interaction surface is shown in the same orientation, with the same colors for perturbed amino acids.

values: 3.9 and 5.8 mM, respectively). Thus, in both cases, arginine symmetrical dimethylation is not critical for the interaction. Nevertheless, the presence of arginine in the peptide is critical, as shown by the absence of measurable affinity of the KG-rich peptide for 53BP1 TT, and the peptide sequence is important because arginine alone is also not capable to bind to 53BP1 TT with a measurable affinity.

We have also tested the affinity of 53BP1 TT for a 24 residue peptide of 53BP1 containing the already tested 8 residue RG-rich core, and relatively well conserved in the three 53BP1 sequences. This peptide binds to
53BP1TT on the same surface and with a similar affinity ($K_d \approx 4.7 \pm 2.0 \text{ mM}$), as compared to the initial peptide ($K_d \approx 5.8 \pm 2.7 \text{ nM}$). This suggests that the flanking residues are not involved in the observed interaction.

**Both Tudor Motifs of 53BP1 Are Necessary for DNA Binding**

Tandem-arranged protein-interaction modules with restrained orientations have already been observed, in which each domain alone retains its structural integrity and peptide ligand binding activity (Hatada et al., 1995; Hof et al., 1998; Ottinger et al., 1998; Jacobson et al., 2000). By contrast, in the case of mouse 53BP1TT, the DNA and peptides binding sites are centered on an aromatic cavity located in motif 1 (Trp18, Tyr23, Tyr46). These aromatic residues are not conserved in motif 2, justifying the absence of an equivalent binding site in motif 2.

Binding of DNA to 53BP1 was extensively studied by Doherty and coworkers (Iwabuchi et al. 2003). They showed that the region 1480–1616 of human 53BP1 (corresponding to our mouse 53BP1TT) binds to single-strand and double-strand DNA in vitro. On the contrary, they found no interaction between these DNA substrates and the fragments 1480–1540 and 1540–1616 that correspond to isolated Tudor domains 1 and 2 respectively. Moreover, they reported that both Tudor motifs are needed to form foci after X irradiation in cells. As shown by our NMR titration, several residues of motif 2 close to the aromatic cavity of motif 1 are also affected by the addition of DNA (Figure 4B), suggesting that motif 2 is also involved in DNA binding. Our results are thus in agreement with Doherty and coworkers’ biological data suggesting that both Tudor domains are needed for interaction with DNA.

Very recently, Zhang and coworkers have reported the three-dimensional structure of tandem PDZ in the rat GRIP1 (Feng et al., 2003). The tandem adopts a compact and stable structure, while the first repeat alone is less stable and its 3D structure is distorted. Mutual stabilization of the two repeats allows ligand binding by the second repeat. Similarly, the tandem BRCT of BRCA1 behave as a single stable fragment in limited proteolysis and X-ray crystallographic studies (Williams et al., 2001), and both repeats are needed for phospho-specific peptide interaction (Manke et al., 2003). Our Tudor tandem structure might be another example where tandem-arranged modules represent functional supramodules with distinct structures and biological functions with respect to individual domains.

**A Similar Region of 53BP1TT for DNA and RG-Rich Peptide Binding**

Surprisingly, the DNA binding surface is similar to the region involved in the binding to RG-rich peptides. However, DNA is mostly negatively charged and the peptide is largely positively charged. Two remarks can be done to understand this result.

First, among the residues involved in DNA binding, no positively charged residue (Arg, Lys) is found, as is usually the case in protein-DNA interfaces. We compared our results with the analysis of 129 protein-DNA complex structures made by Thornton and coworkers; only three residues (Asn21, Asp44, and Glu47) correspond to typical amino acids forming hydrogen bonds at protein-DNA interfaces (Luscombe et al. 2001). Furthermore, in the same study, the authors noticed that cysteines have a high propensity to contact DNA backbone; in our case, the NMR signal of Cys48 is significantly perturbed by the binding to DNA. Finally, two phenylalanines (Phe 24, Phe42) are also perturbed by the addition of DNA; this amino acid has a high affinity for many DNA base types, which can be explained by its ability to produce extensive ring-stacking interactions (Luscombe et al. 2001). These observations suggest that the 53BP1TT-DNA interface essentially involves hydrogen bonding to nonpositively charged residues and base stacking to aromatic residues.

Second, the RG-rich peptide binding surface is mainly composed of hydrophobic residues (Trp18, Tyr23, Phe24, Phe42, Tyr46); these are able to develop cation-π interactions with the arginines of the peptide (Zacharias and Dougherty, 2002). The Tudor domain of SMN possesses a similar aromatic cavity. However, in the case of SMN, an arginine probably interacts with the negatively charged residue Glu134 of the cavity. In 53BP1TT, no negatively charged side chain is found in the binding site.

To sum up, both DNA and RG-rich peptides interact with 53BP1TT mainly via hydrophobic contacts, cation-π interactions, ring stacking, and hydrogen bonding. No positively charged residue of 53BP1TT is clearly involved in DNA binding and no negatively charged residue is observed in the RG-rich peptide binding site. This may explain why both interactions involve the same region of 53BP1TT.

**Tudor Tandem Is Associated to Gene Transcriptional Regulation, Chromatin Remodeling, and DNA Repair**

The three-dimensional structure described in this paper is the first solved structure of a tandem of Tudor domains. Many Tudor proteins contain multiple Tudor repeats, and Tudor domains are often arranged in closely linked pairs (Letunic et al., 2004). For example, the human ESET protein possesses two Tudor domains connected by a 38 residue linker, while the human GASC1 protein possesses two such domains connected by an 11 residue linker. Alignment of the sequences of these Tudor tandem domains shows that 28 amino acids of 53BP1TT are conserved in more than 70% of the sequences (Figure 5). Twelve residues are buried and constitute the hydrophobic core of motif 1. Eight residues play a similar structural role in motif 2. Seven residues are localized at the interface between motif 1 and motif 2: Gly11, Arg13, Gly22, Phe24, Tyr25, Asp43, and Arg106. Such a distribution of conserved residues suggests that the role of these residues is mainly to stabilize the Tudor tandem fold.

The Tudor tandem is mostly associated with SET, MBD, PHD, and Zinc Finger C2H2 or BRCT domains, which suggests that their biological function is always linked to gene transcriptional regulation, chromatin remodeling, and DNA repair (Letunic et al., 2004). However, most residues of 53BP1TT involved in RG-rich pep-
Figure 5. Sequence Alignment of Tudor Tandem Containing Proteins

Conserved residues are colored (dark green for hydrophobic, red for negatively charged, blue for positively charged, cyan for aromatic, and yellow for small residues). Secondary structures of 53BP1TT are indicated (arrows for β strand and rectangle for α helix); mouse 53BP1 residues colored in gray have a relative accessible surface below 20%. Red asterisks show the residues whose NMR signal is affected by the binding to both RG-rich peptides and DNA.

tide or DNA binding are not conserved within other Tudor tandem sequences. Thus, by now, no evidence indicates that Tudor tandems show common target binding properties.

Conclusion

The Tudor domains of SMN and 53BP1 are able to bind to proteins containing RG-rich sequences. Moreover, the ribosomal protein RL24 of H. marismortui has a Tudor-like fold and interacts with RNA (Kyrpides et al., 1996; Steiner et al., 2002; Ban et al., 2000). By comparing the RNA binding site of RL24 and the RG-rich peptide binding site of SMN, Wahl and coworkers (Steiner et al., 2002) proposed that the Tudor-like domain of the microbial transcription modulator NusG could bind concomitantly to proteins and nucleic acids via different surfaces. On the opposite, our analysis shows that the Tudor tandem of 53BP1 uses the same surface to bind both peptides and DNA. The presence of this multipartner binding surface suggests that 53BP1TT acts as an adaptor mediating both protein-protein and protein-DNA interactions.

Furthermore, since the tested RG-rich sequence belongs to 53BP1 itself, 53BP1TT may be involved in intramolecular or intermolecular associations. An intramolecular interaction may allow a regulation of the accessibility to other partners of the 53BP1TT functional site. An intermolecular association could entail the accumulation of 53BP1 molecules and, therefore, be a key step in the mechanism of nuclear foci formation observed after cell irradiation. Further studies should be performed to search for other partners of 53BP1TT in the nucleus and to investigate the role of methylation in the interactions between 53BP1TT and RG-rich proteins.

Experimental Procedures

NMR Spectroscopy

53BP1TT was expressed and purified as described previously (Carrier et al., 2004). NMR samples were prepared in Tris-HCl 50 mM buffer (pH 7.2) containing 150 mM NaCl in either 90% H2O/10%
Structure and Interactions of 53BP1 Tudor Tandem

D2O or in 100% D2O. 1 mM EDTA, a protease inhibitor cocktail (SIGMA), 1 mM NaNO3, and 1 mM 3-trimethylsilyl[2,2,3,3-2H4] propionate (TSP) were added. All assignment experiments were performed at 27°C on Bruker DRX-500, DRX-600 equipped with triple-resonance probes according to the previously reported procedure (Charrier et al., 2004). The NOE crosspeak volumes used for structure calculation were measured on four NOESY experiments (two 1H-N-HMQC-NOESY recorded at 750 MHz at the European Large Scale Facilities in Utrecht, Netherlands, and a 1H-15N-HSOC-NOESY in D2O, a 1H-15N-HSOC-NOESY in H2O, and a 13C-1H-HSOC-NOESY in the 13C aromatic region all three recorded on a local 600 MHz spectrometer equipped with a triple resonance TXI cryoprobe). φ torsion angle values were deduced from the analysis of the Hn-H and the HMCG-Q-J experiments (Vuister and Bax, 1993; Vuister et al., 1994). Hydrogen bond restraints were derived from slowly exchanging amide protons, identified after exchange of H2O to D2O followed on 1H-15N HSQC spectra recorded at different times. All spectra were processed with the programs XwinNmr (Bruker) or NMRPipe (Delaglio et al., 1995) and analyzed using Felix (Molecular Simulations).

Structure Determination
The solution structure of region 7–129 was solved on the basis of 2337 interproton distances deduced from the NOE data. These distances were estimated from 3368 integrated peak volumes obtained from the four NOESY experiments (935 on the 1H-N-HSOC-NOESY, 473 on the 1H-15N-HSOC-NOESY in D2O, 1685 on the 1H-15N-NOE spectra in H2O, and 275 on the 13C-1H-HSOC-NOESY in the 13C aromatic region). A semiautomated iterative assignment procedure was applied for the assignment and the construction of the 3D structures (Savarin et al., 2001). A force field adapted to NMR structure calculation (file parallelhlg.pro in CNS 1.0 [Brunger et al., 1998]) was used. 206 torsion angle (φ or ψ) values were deduced from the analysis of the Hn-H and the HMCG-Q-J experiments and from the backbone H, N, and C chemical shifts using the program TALOS (Cornilescu et al., 1999). Finally, 30 hydrogen bonds were imposed during the structure calculation. At the last step, 1000 structures were calculated and the 10 best structures were selected and refined with a standard energy function (CHARMM22), including an electrostatic energy term. This term is calculated with no net charge on the side chain atoms and with a distance-gated dielectric constant.

Titration with Peptides and DNA
NMR titrations were carried out by recording 1H-15N HSQC spectra at concentrations of 0.3–0.6 mM. The four peptides: RGRGRRGG, RGRGRRGRG (R = symmetrically dimethylated R), KGKKGKGK, and KGKAPVTPGRGRRGRPPRSRTTGGT, from Peptide Speciality Laboratories GmbH (Heidelberg), were added up to 9-, 6-, 5-, and 34-fold molar excess to the protein samples, respectively. The 10 bp oligonucleotide 5’-AAAGGTTTTGTTAAC-3’ (PROLIGO, Paris) was annealed prior to NMR experiments and added up to a 3-fold molar excess to the protein sample. Assuming that the exchange rate is fast and that these interactions show low affinities, we can consider that the concentration of the free ligand is approximately the same as the concentration of the added ligand. Thus dissociation constants were estimated by fitting the titration curves obtained for strongly perturbed residues with Kaleidograph software, using the approximate equation y = Δvmax x Kx / (Kx + x), where y is the weighted chemical shift displacements [Δv(1H)] + 0.1 x [Δv(15N)] (Δv = Δmax – ΔoFree is the observed chemical shift, Δi is the initial chemical shift before adding the ligand), x is the ligand concentration, Δvmax is the maximum variation of the weighted chemical shift displacements, and Kx is the estimated dissociation constant.

Bio-Informatics
An initial multiple alignment was built from Tudor tandem sequences found in the SMART database (Letunic et al., 2004) (http://smart.templbl-heidelberg.de/smart/). The sequence of 53BP1™ was structurally aligned to this set of sequences. Using HMMER2.3 (Durbin et al., 1998) (http://hmmer.wustl.edu/), a HMM profile was built and used to scan the Non Redundant database (ftp://ftp.ncbi.nih.gov/blast/db/). New sequences were detected (E-value < 1e-25) and added to the initial alignment. Sequences with high identity were excluded. The final alignment was manually optimized using the secondary structure of 53BP1™.

Supplemental Data
Histograms representing the absolute value of chemical shift displacements of the whole residues of 53BP1™ after saturation with DNA or different RG-rich peptides are provided as Supplemental Data and can be found at http://www.structure.org/cgi/content/full/12/9/1551/DC1.

Acknowledgments
The 750 MHz spectrum was recorded at the SON NMR Large Scale Facility in Utrecht, which is funded by the “Access to Research Infrastructures” program of the European Union (HPRI-CT-2001-00172). We are grateful to Philippe Savarin and Flavio Toma who kindly lent us their 600 MHz spectrometer. We thank D. Jullien for providing mouse 53BP1 cDNA and G. Stier for TEV protease cDNA. Vincent Meyer is supported by grants from Association Française contre les Myopathies (AFM, grant Decrypton 2003 #9457).

References


Accession Numbers
The NMR restraints and protein coordinates have been deposited in the PDB under entry code 1SSF. The chemical shifts have been deposited at the BMRP under accession number 5878.