

A Tandem of SH3-like Domains Participates in RNA Binding in KIN17, a Human Protein Activated in Response to Genotoxics

Albane le Maire¹, Marc Schiltz², Enrico A. Stura¹
Ghislaine Pinon-Lataillade³, Joël Couprie¹, Mireille Moutiez¹
Muriel Gondry¹, Jaime F. Angulo³ and Sophie Zinn-Justin^{1*}

¹CEA, DSV, DIEP
Gif-sur-Yvette, F-91191
France

²EPFL, Laboratoire de
Cristallographie, Lausanne
CH-1015, Switzerland

³CEA, DSV, DRP
Fontenay-aux-Roses
F-92265, France

The human KIN17 protein is an essential nuclear protein conserved from yeast to human and expressed ubiquitously in mammals. Suppression of *Rts2*, the yeast equivalent of gene *KIN17*, renders the cells unviable, and silencing the human *KIN17* gene slows cell growth dramatically. Moreover, the human gene *KIN17* is up-regulated following exposure to ionizing radiations and UV light, depending on the integrity of the human global genome repair machinery. Its ectopic over-expression blocks S-phase progression by inhibiting DNA synthesis. The C-terminal region of human KIN17 is crucial for this anti-proliferation effect. Its high-resolution structure, presented here, reveals a tandem of SH3-like subdomains. This domain binds to ribonucleotide homopolymers with the same preferences as the whole protein. Analysis of its structure complexed with tungstate shows structural variability within the domain. The interaction with tungstate is mediated by several lysine residues located within a positively charged groove at the interface between the two subdomains. This groove could be the site of interaction with RNA, since mutagenesis of two of these highly conserved lysine residue weakens RNA binding.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: KIN17 protein; X-ray structure; SH3-like fold; RNA binding; tungstate

*Corresponding author

Introduction

Cells are continuously submitted to attack, either from reactive oxygen species resulting from the normal cellular metabolism or by exogenous agents such as radiation and chemicals. Consequent DNA damage triggers specific signalling to activate the repair machinery. The human gene *KIN17*, located in chromosome 10, participates in the general response to genotoxics. It is up-regulated following DNA damage produced by UV-C or by ionizing radiation,^{1,2} and this response to UV light is strictly dependent on the presence of the functional nucleotide excision repair proteins XPA and XPC.^{2,3} The ectopic over-

expression of gene *KIN17* modifies the nuclear morphology and inhibits S-phase progression, thus blocking cell growth.^{4,5}

KIN17 encodes for a nuclear protein conserved from yeast to human. Its ubiquitous expression suggests its participation in a regulatory mechanism common to all cell types. Suppression of *Rts2*, the yeast *KIN17* homolog, renders the cells unviable,⁶ and silencing the human gene *KIN17* impedes cell growth dramatically.⁷ *KIN17* is part of high molecular mass complexes that mediate different types of nucleic acid transactions.^{8,9} The mammalian *KIN17* protein binds to curved DNA found at hot-spots of illegitimate recombination in eukaryotic chromosomes,^{10–12} and partially complements the bacterial transcriptional regulator in *Escherichia coli* called H-NS.¹³ Moreover, proteomic analysis detected protein *KIN17* in several complexes of the human spliceosome, a large RNA–protein assembly consisting of small ribonucleoprotein particles (snRNP) and associated non-small ribo-

Abbreviation used: GST, glutathione-S-transferase.

E-mail address of the corresponding author:
sophie.zinn@cea.fr

nucleoproteins.^{14,15} Interaction between KIN17 and RNA was further characterized by Pinon-Lataillade *et al.*,¹⁶ showing that the mouse KIN17 protein interacts with RNA *in vivo* and that human and mouse KIN17 proteins are able to directly bind different types of RNA homopolymers *in vitro*.

Human KIN17 (hKIN17) is a 45 kDa modular protein composed of a 30 kDa N-terminal region conserved in all species, and a 15 kDa tail found only in higher eukaryotes. The N-terminal region exhibits a C₂H₂ zinc finger motif, which is generally associated with nucleic acid binding function. This domain modulates the interaction between KIN17 and double-stranded DNA as its deletion leads to a reduced DNA binding.¹¹ In contrast, the C-terminal tail does not affect the DNA binding of KIN17.¹¹ It is involved in the anti-proliferative effect of KIN17. Indeed, cell growth is not affected by over-expression of KIN17 protein lacking the C-terminal tail, whereas it is inhibited by over-expression of the whole protein.⁵

The C-terminal domain of hKIN17 comprises a 27 residue KOW motif,¹⁷ shared among three families of ribosomal proteins (microbial L24 and eukaryotic eL26 and eL27), the essential bacterial transcriptional elongation factor NusG, the eukaryotic chromatin elongation factor Spt5p, the T54 proteins and the higher eukaryotic KIN17 proteins.¹⁸ In the three-dimensional structures of the ribosomal proteins¹⁹ and the transcription factor NusG,²⁰ the KOW motif occurs in the first two strands and adjacent loops of a larger src homology 3 (SH₃)-like domain. L24 protein interacts with several RNAs in the ribosome, which is in agreement with the original proposal that KOW is an RNA-binding motif.¹⁷

Here, we report the high-resolution X-ray structure of the C-terminal domain of hKIN17 and show that this domain contributes to the RNA-binding properties of the whole protein. We also present a structure of this domain complexed to tungstate that allowed us to identify two highly conserved lysine residues involved in binding to this ion analogous to phosphate, as well as in RNA binding.

Results

The C-terminal region of hKIN17 exhibits an all- β structure consisting of two tightly packed SH₃-like subdomains

The boundaries of the C-terminal domain of hKIN17 protein (amino acid residues 268–393) were delineated on the basis of residue conservation from sequence alignment of all known KIN17 proteins (Figure 1(a)) and on the basis of hydrophobic amino acid cluster analysis²¹ of the hKIN17 sequence. This domain forms an independently folded unit, as it can be expressed as a soluble and folded protein in solution.

The structure of the isolated C-terminal domain of hKIN17 was determined by means of single

isomorphous replacement with anomalous scattering (SIRAS) phasing using data collected on crystals soaked with iodine as described.²² The structural model, including 211 water molecules, was refined to a resolution of 1.45 Å, leading to final *R* and *R*_{free} values of 13.7% and 17.5%, respectively (Table 1). The geometry can be considered as very good, with 94.2% of the residues in the most preferred ϕ/ψ conformations and only one residue in a disallowed region of the Ramachandran plot. All residues in the X-ray crystal structure are well ordered, except for the first seven residues, (the non-wild-type glycine introduced by the cloning vector and the first six residues) for which there is no electron density, and for which a signal corresponding to random coil is seen in NMR experiments (data not shown). This domain is roughly ellipsoidal, with dimensions of 40 Å × 30 Å × 25 Å and organized into two major subdomains (Figure 1(b)). Subdomain A (spanning residues 284–334 of the hKIN17 sequence) and subdomain B (spanning residues 340–391 of the hKIN17 sequence) both adopt a similar five-stranded β -barrel-like fold and are connected to each other by a short linker of five residues. In each subdomain, the first and smaller β -sheet is formed from strands β -1_{A(or B)}, part of β -2_{A(or B)} and β -5_{A(or B)}, and the second and larger β -sheet is formed from the remainder of β -2_{A(or B)}, β -3_{A(or B)} and β -4_{A(or B)}. The two β -sheets pack against each other at approximately right-angles because of the helical turn between β -4_{A(or B)} and β -5_{A(or B)}. Interestingly, the N-terminal segment 276–283 and the C-terminal fragment 391–393 form the interface between the two subdomains. They interact with both subdomains and the C-terminal fragment interacts with the linker region (Figure 1(b)). The surface area of each subdomain buried in the interface of the tandem structure is approximately 1150 Å². This is a relatively large protein–protein interface with a classical percentage of non-polar residues (58 %).²³

The overall backbone structure of the two subdomains superimposes well, with an rmsd of 1.16 Å on 29 C α atoms, excluding the three variable loops, despite a low level of sequence identity (9% identity; 54% sequence similarity) (Figure 2). A DALI search using the individual A and B subdomain structures shows that these subdomains display structural analogy with known protein domains adopting an SH₃-like fold. Domain III of NusG (PDB code 1M1G) is the best DALI match for both subdomains, and shows 17% and 20% identity with subdomains A and B, respectively. A high DALI score (>6) is obtained also for the dihydrofolate reductase (PDB code 1vie), the Tudor domain of SMN protein (PDB code 1gv5), the SH₃ domain of the protein kinase p56lck (PDB code 1lck-A) and the ribosomal protein L21e (PDB code 1ffk-N). Thus, the C-terminal domain of hKIN17 has structural similarity with various SH₃-like domains, even if it lacks similarities in either primary sequence or charge distribution.

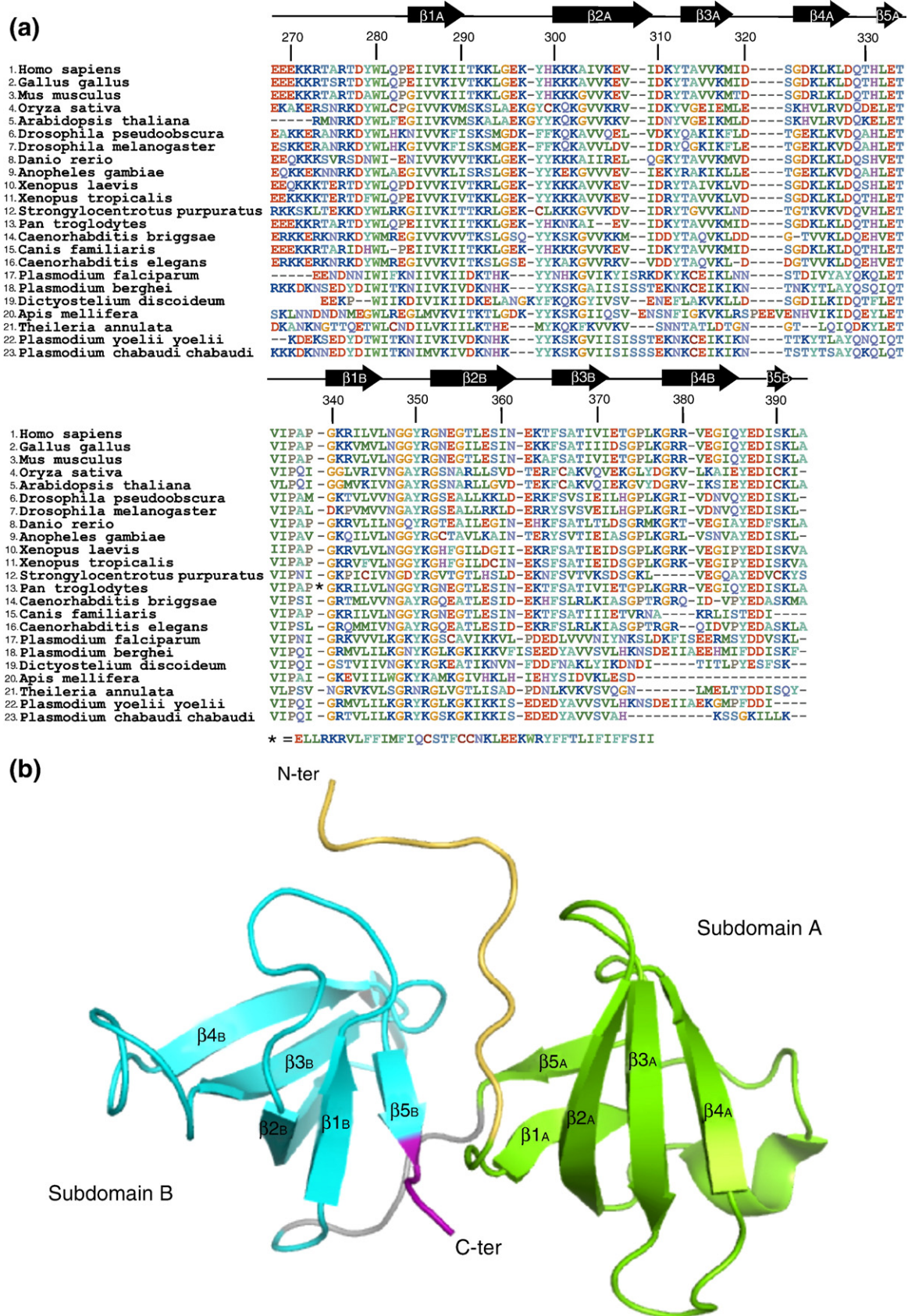


Figure 1 (legend on opposite page)

Table 1. Phasing and refinement statistics

A. Phasing	
FoM (acentric/centric)	0.4961/0.3944
Phasing power iso (acentric/centric)	1.980/1.573
Phasing power anomalous	1.250
B. Refinement	
Space group	P2 ₁ 2 ₁ 2 ₁
Resolution range (Å)	32.62–1.45
Reflections used in refinement	22,249
R (%)	13.7
R _{free} (%)	17.5
Total number of atoms	1181
Number of	
Protein atoms	954
Iodine atoms	12
Water molecules	211
Average B-factor (Å ²)	15.3
r.m.s.d from ideality	
Bond lengths (Å)	0.013
Bond angles (deg.)	1.519
Chirality (deg.)	0.091
Ramachandran plot	
Most-favoured regions (%)	94.2
Additionally allowed regions (%)	4.8
Generously allowed regions (%)	0
Disallowed regions (%)	1 ^a

^a K311 in loop β_{2A}β_{3A} is located in a disallowed region of the Ramachandran plot; however, its backbone conformation was well supported by the electron density.

The C-terminal domain of hKIN17 binds RNA *in vitro*

Since hKIN17 protein is able to recognize RNA homopolymers and can bind RNA *in vivo*,¹⁶ we investigated which region(s) of the protein interact(s) with RNA. Three constructs containing either the N-terminal globular region 1–160 (shown to be folded by NMR analysis; data not shown), the same region without the zinc finger 68–160 (shown to be folded by NMR analysis),²⁴ or the C-terminal domain 268–393 described here were tested for their ability to bind RNA using the same ³²P-labelled RNA probe as that used to characterize the RNA interaction properties of the full-length hKIN17.¹⁶ The North-western analysis showed that the N-terminal domain comprising the zinc finger interacts efficiently with RNA *in vitro*, and that the deletion of the zinc finger from this domain drastically reduces RNA binding (Figure 3). The C-terminal domain showed an intermediate RNA-binding activity. In comparison, a South-western analysis showed that only the

N-terminal domain binds significantly to a double-stranded DNA known to interact with the whole hKIN17 protein (data not shown).

To characterize further the RNA-binding properties of the C-terminal domain, we carried out binding assays with ribonucleotide homopolymers immobilized on agarose beads. This kind of assay has been useful in assessing RNA binding for many other RNA-binding proteins and domains.^{25–27} The C-terminal domain of hKIN17 fused to glutathione-S-transferase (GST) binds to all four RNA homopolymers at 100 mM NaCl, with a clear preference for poly(rG) and poly(rU) (Figure 4(a)). The possibility of GST interference was eliminated by repeating the test with GST alone, since no signal was detected using a specific anti-GST antibody. The cleaved protein was also assayed for RNA binding with results comparable to those obtained with the GST-hKIN17 C-terminal domain fusion (Figure 4(a)). To rule out the possibility of non-specific binding due to electrostatic interactions, the assay was repeated at increasing ionic strength varying from 0.1 M to 1 M NaCl, showing that binding of the C-terminal domain of hKIN17 to poly(rG) and poly(rU) is still significant in concentrations of NaCl up to 0.25 M (Figure 4(b)). Such resistance of the C-terminal domain-poly(rG) or poly(rU) interaction to significant ionic strength suggested that, in addition to ionic contacts, hydrophobic interactions are involved in the RNA-protein interaction. The protein also retains its RNA-binding capacity in the presence of either 1 mg/ml of *E. coli* tRNA or 1 mg/ml of heparin (Figure 4(b)). Finally, the heparin-resistant binding,^{28,29} and the tRNA-resistant binding, demonstrate the stringency and specificity of the binding of the C-terminal domain of hKIN17 to RNA. A North-western assay confirms the interaction of the C-terminal domain of hKIN17 with poly(rG) and poly(rU) (Supplementary Data). On the basis of these experiments, we conclude that the C-terminal domain of hKIN17 is an RNA-binding domain *in vitro*.

Tungstate-binding sites reveal important determinants for the binding of RNA

The binding of a protein to RNA is often mediated by interaction with the phosphate backbone of the nucleic acids, and positively charged amino acids are generally found in contact with the phosphate groups.³⁰ In order to

Figure 1. Sequence and structure of the hKIN17 C-terminal domain. (a) Multiple sequence alignment of the C-terminal region of 23 eukaryotic KIN17 proteins. The numbering scheme refers to the sequence of human KIN17. Positively charged residues are in blue, negatively charged residues are in red and hydrophobic residues are in green. The secondary structure elements corresponding to the hKIN17 C-terminal crystal structure are indicated above the alignment. The length of the hKIN17 C-terminal domain is well conserved, except in the Chimpanzee (*Pan troglodytes*) sequence, which harbours a 42 residue insertion between the two subdomains. (b) A ribbon representation of the 3D structure of the hKIN17 C-terminal domain. It consists of two SH3-like subdomains; the first one (A) is in green and the second one (B) is in blue. The short linker between the two subdomains is in grey (segment 335–339) and the two segments involved in the interface between the subdomains are in orange (N-terminal segment 274–283) and in violet (C-terminal segment 391–393).

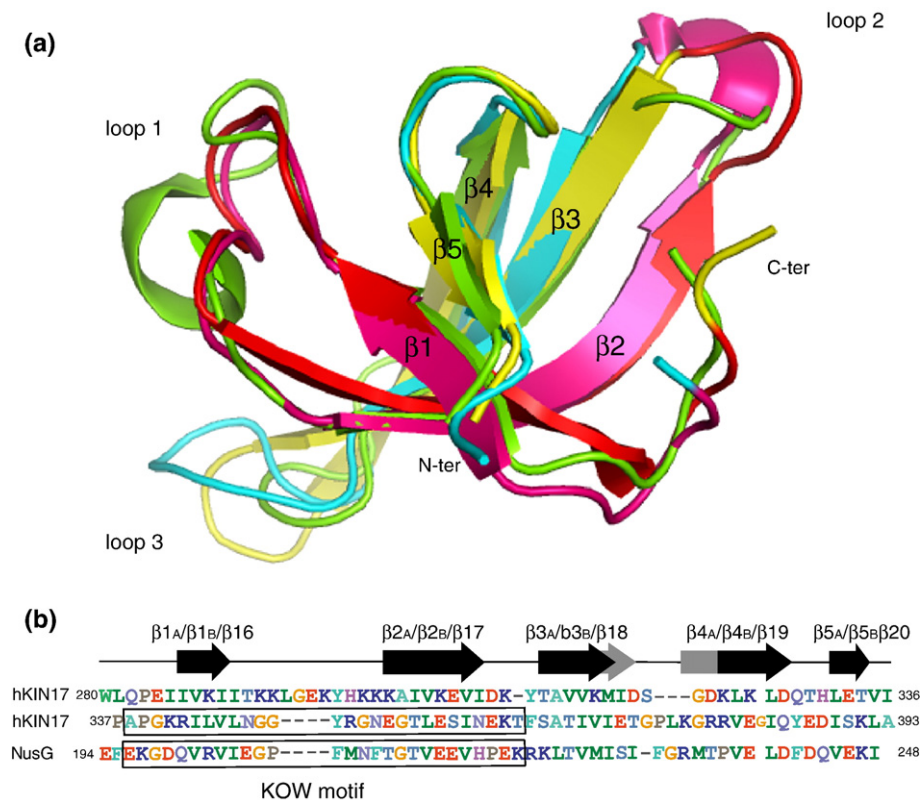


Figure 2. Structural analogues of the two SH₃-like subdomains. (a) Superposition of the two subdomains of hKIN17 C-terminal domain (colours as in Figure 1) and the domain III of *aaeNusG* (PDB code 1M1G; yellow). The rmsd calculated between the backbones of the two subdomains except for the three variable loops reaches 1.16 Å (on 29 C°). The rmsd calculated between the backbones of subdomain B of hKIN17 C-terminal domain and its closest DALI match, the domain III of *aaeNusG* except for the three variable loops reaches 1.20 Å (on 29 C°). The sequences presenting the KOW motif form the first two β-strands and their adjacent loops (in red for NusG and in pink for KIN17). (b) Sequence alignment of the two subdomains of hKIN17 C-terminal domain with the domain III of *aaeNusG*. A β-bulge occurs in the second subdomain causing the insertion of residue G383. Longer β strands (β18 and β19) for the domain III of *aaeNusG* are drawn in grey. The KOW motifs of the second subdomain of hKIN17 C-terminal domain and the domain III of *aaeNusG* are boxed.

identify the positively charged residues of hKIN17 involved in RNA interaction, we screened the surface of the C-terminal domain using tungstate, which can mimic phosphate groups,³¹ and has been useful for identifying phosphate binding sites in other crystal structures.^{32–36} As tungsten is substantially heavier than phosphorus, has more electrons and can give rise to a strong anomalous signal, tungstate ions can be identified easily by X-ray diffraction methods. We collected data at the W *L*_{III} absorption edge from a crystal soaked in Na₂WO₄ solution and a crystal obtained by co-crystallizing with Na₂WO₄ to a resolution of 1.91 Å and 2.30 Å, respectively. Both derivative structures revealed important features of the binding of tungstate anions to hKIN17 C-terminal domain.

The initial 2*F*_o–*F*_c and *F*_o–*F*_c electron density maps obtained from soaked crystals showed a large patch of electron density, defining a cluster of tungsten atoms, confirmed by the anomalous signal. Twelve strong peaks (>10 σ for the anomalous difference) characterise the entire cluster but only six of these have a refined

occupancy greater than 0.1. The tungstate anions are all located around the cationic NH₃⁺ side-chain groups of K302 in β2 of subdomain A, of K341 in β1 of subdomain B and of K391 at the C terminus of β5 in subdomain B of hKIN17 C-terminal domain (Figure 5(a)). A larger cluster of 12 tungstate anions, at almost the same location as in the soaked crystals, is found for the co-crystal (Figure 5(b)). The interaction is characterised by two additional lysine residues (K377' and K297') from molecules related by crystal symmetry and another minor site where one tungsten ion is nestled between the side-chains of R273 and of K317' from a symmetry-related molecule.

Derivative structures reveal structural variability within the C-terminal domain

The structure of the soaked crystal is isomorphous with that of the native domain. On the other hand, the co-crystal of the C-terminal domain of hKIN17 with Na₂WO₄ changes its cell parameters from *a* = 45.75 Å and *b* = 46.31 Å in the native crystal to *a* = 47.67 Å and *b* = 48.72 Å, with concomitant

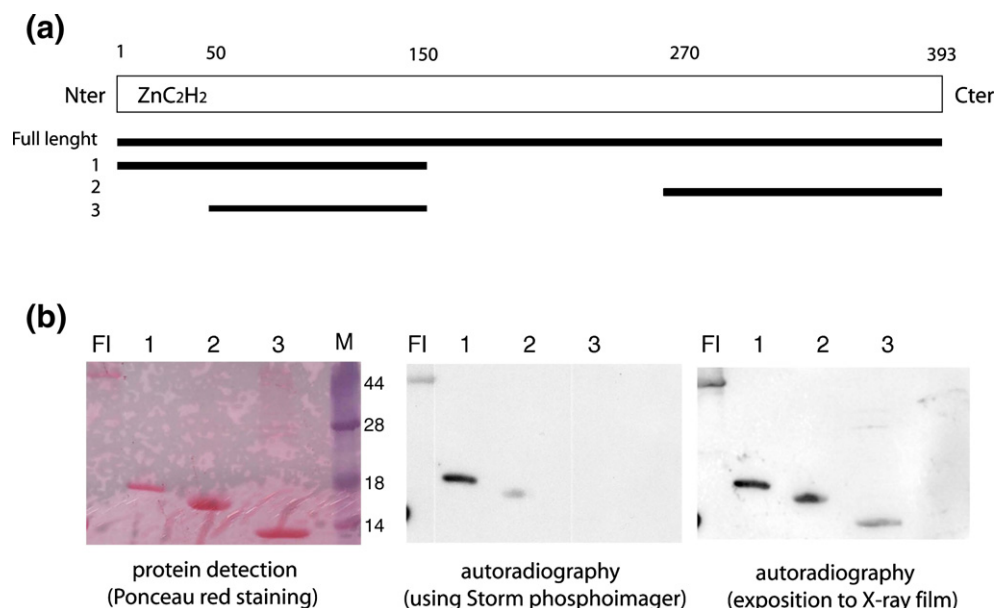


Figure 3. RNA-binding activity of deletions of the hKIN17 protein. (a) A representation of the full-length and deleted hKIN17 proteins. The deleted protein (1) corresponds to the N-terminal tail of hKIN17 and the deleted protein (3) corresponds to the same region without the zinc finger. The deleted protein (2) corresponds to the C-terminal domain of hKIN17. Residues from 150 to 270 of the hKIN17 sequence correspond to a flexible region, without any predicted folded domain. (b) The full-length KIN17 protein (FI) and the deleted proteins were detected by staining with Ponceau red on a nitrocellulose membrane to estimate the amount of input protein. North-western blotting was carried out with the same membrane incubated with a ^{32}P -labelled RNA probe in 0.05 M NaCl. M corresponds to a pre-stained molecular mass marker. The zinc finger and the C-terminal domains are the hKIN17 RNA-binding domains.

differences in crystal packing. To characterise these changes, we compare the models obtained from the crystal soaked in Na_2WO_4 and the Na_2WO_4 co-crystal (Table 2). As shown in Figure 5(a), the hKIN17 C-terminal domain exhibits conformational changes in loops, in particular in the second subdomain, as well as a slight shift in the orientation of the two SH_3 -like subdomains relative to each other. However, the A–B subdomains interface is mostly unchanged.

The C-terminal domain of hKIN17 binds RNA through a surface comprising the highly conserved lysine residues 302 and 391

The tungstate ions interact with several lysine residues from the hKIN17 C-terminal domain, two of which, K302 and K391, are strictly conserved among KIN17 sequences (Figure 1(a)). To confirm that K302 and K391 are also involved in hKIN17 binding to RNA, two mutants were produced where these lysine residues were changed into negatively charged residues: K302E and K391E. Binding of the mutants to RNA was assayed by two independent methods. A binding assay using affinity chromatography to ribonucleotide homopolymers immobilized on agarose beads showed that both mutants K302E and K391E exhibited a reduced RNA-binding ability compared to the wild-type (Figure 6). K302E mutation abolishes the poly(rC) and poly(rA)-

binding activity and clearly reduces the poly(rU) and poly(rG)-binding activity of the C-terminal domain of hKIN17. The K391E mutant has a reduced RNA-binding activity, in particular concerning its poly(rU) binding activity. These results were further confirmed by performing a North-western analysis using ^{32}P end-labelled poly(G) and poly(U) RNAs. This experiment showed that the K302E and K391E mutations abolish the poly(rU)-binding activity and reduce the poly(rG)-binding activity of the C-terminal domain of protein hKIN17 (Supplementary Data). NMR spectroscopy of the mutant proteins shows that the charge reversal mutants have a three-dimensional structure similar to that of the native protein (data not shown), indicating that the surface mutations do not affect their global folding or structural integrity. Thus, we conclude that the two highly conserved lysine residues K302 and K391 are implicated in RNA binding of the C-terminal domain of hKIN17.

Discussion

Contribution of the hKIN17 C-terminal domain to RNA binding

A KOW motif was identified from sequence analysis in the C-terminal domain of hKIN17.

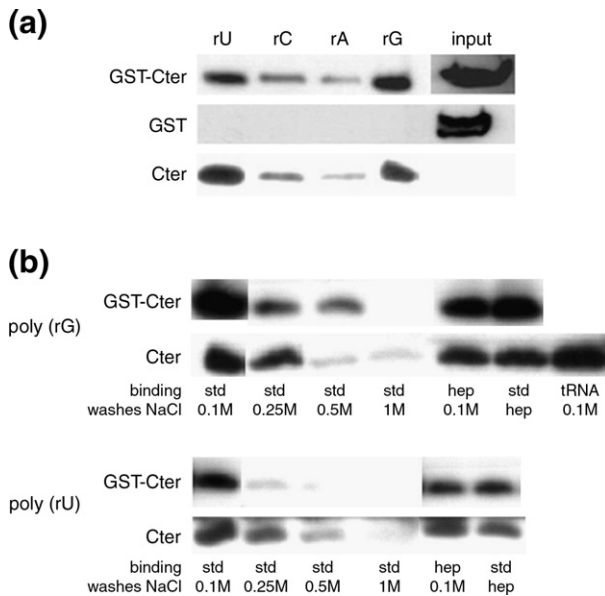


Figure 4. RNA-binding properties of the hKIN17 C-terminal domain. (a) Purified recombinant hKIN17 C-terminal domain (Cter) or in fusion with a GST-tag and purified GST protein as a negative control were incubated with each ribonucleotide homopolymers and washed with 0.1M NaCl. The proteins retained by the immobilized homopolymer RNAs were analysed by immunoblotting. hKIN17 C-terminal domain shows a clear preference for poly(rG) and poly(rU) homopolymers. (b) Characterization of the poly(rG) and poly(rU) RNA interaction with the hKIN17 C-terminal domain alone or in fusion with a GST-tag. Purified recombinant proteins were mixed with immobilized poly(rG) and poly(rU) under various binding conditions, as indicated. Std corresponds to the following binding buffer: 20 mM Tris-HCl (pH 7.4), 2.5 mM MgCl₂, 0.5% (v/v) Triton X-100, 100 mM NaCl. The retained proteins were washed under the conditions described for each lane. Hep corresponds to incubation with 1 mg/ml of heparin. The hKIN17 C-terminal domain binds poly(rG) and poly(rU) in salt-resistant and heparin-resistant conditions.

From the known structure of the large ribosomal subunit,¹⁹ the KOW sequence motif occurs within an SH₃-like fold. It was thus predicted that the hKIN17 C-terminal domain should exhibit an SH₃-like fold. Surprisingly, it contains two structurally similar SH₃-like folds. The predicted KOW sequence element begins in the loop before β_{1B} strand and finishes in the loop $\beta_{2B}\beta_{3B}$. Loop $\beta_{1\beta_2}$ is four residues larger in subdomain A compared to subdomain B, and this insertion prevents the detection of a KOW motif in subdomain A, even if the canonical conserved glycine exists at position 348. On the contrary, loops $\beta_{2\beta_3}$ and $\beta_{3\beta_4}$ are larger in subdomain B. Our RNA interaction studies strengthen the notion that KOW sequence containing domains are involved in RNA binding.

Tungstate derivatives were prepared either by soaking preformed crystals in a solution of Na₂WO₄ or by co-crystallising in the presence of Na₂WO₄.

Methodologically, the two approaches are complementary. In soaked crystals, structural rearrangements can be inhibited by the rigidity of the crystal lattice, while during co-crystallization, artefactual sites can be created at ligand-mediated crystal contacts. In our soaking experiments, we observed no cracking or dissolution of the crystals and a protein conformation identical with that of the free form of KIN17, while the co-crystals showed that conformational changes at the interface between the two subdomains and in loops $\beta_{2\beta_3}$ and $\beta_{3\beta_4}$ occurred as the complex was formed in solution. All these conformational changes that occurred in derivatives were not of sufficient amplitude to overcome the crystal lattice constraints. In the co-crystals, a larger domain rearrangement was possible and unrestrained by crystal interactions, and the loops could adopt a more appropriate conformation. These rearrangements are likely to originate at the tungsten-interacting lysine residues and are amplified at longer distances, culminating in loops $\beta_{2\beta_3}$ and $\beta_{3\beta_4}$ of subdomain B.

Conservatively, we considered meaningful only the tungstate-binding sites common to soaked crystals and co-crystals; namely, K302, K341 and K391. Identification of these lysine residues involved in the interaction with tungstate, a phosphate analog, allowed us to locate putative binding sites of the hKIN17 C-terminal domain with the RNA phosphoribose backbone. Mutagenesis studies confirmed the role of K302 and K391 in RNA binding. This is quite typical, as lysine and arginine residues are often involved in binding to the phosphate backbone of nucleic acids.³⁰ We suggest that the flexibility observed after tungstate binding might allow the protein to adapt to its RNA target.

The RNA-binding activity of hKIN17 seems to be distributed over several domains in the whole protein. At the N terminus, the C₂H₂ zinc-finger domain shows a dual RNA/DNA-binding function *in vitro*, as reported recently for other zinc-finger motifs.³⁷ At the C terminus, the tandem SH₃-like domain binds only RNA *in vitro*. Our methodology that uses tungstate, an electron-dense phosphate mimic, is effective in delimiting in the crystal structure the volume occupied by the RNA phosphate backbone in those stretches where protein-phosphate interactions are dominant, but is unable to define any sequence specificity. The nucleotide specificity suggested by the differential binding of the various RNA homopolymers could arise from base-protein interactions or a preference for a particular phosphate backbone conformation found in its biological RNA target. Given that the hKIN17 C-terminal domain has the same preferences for poly(rG) and poly(rU) as the whole hKIN17 protein, it is likely that this region of the molecule is a major determinant of the *in vivo* RNA binding of the whole hKIN17. From the structural study of several protein-RNA complexes, it was shown that the use of multiple domains allows extended

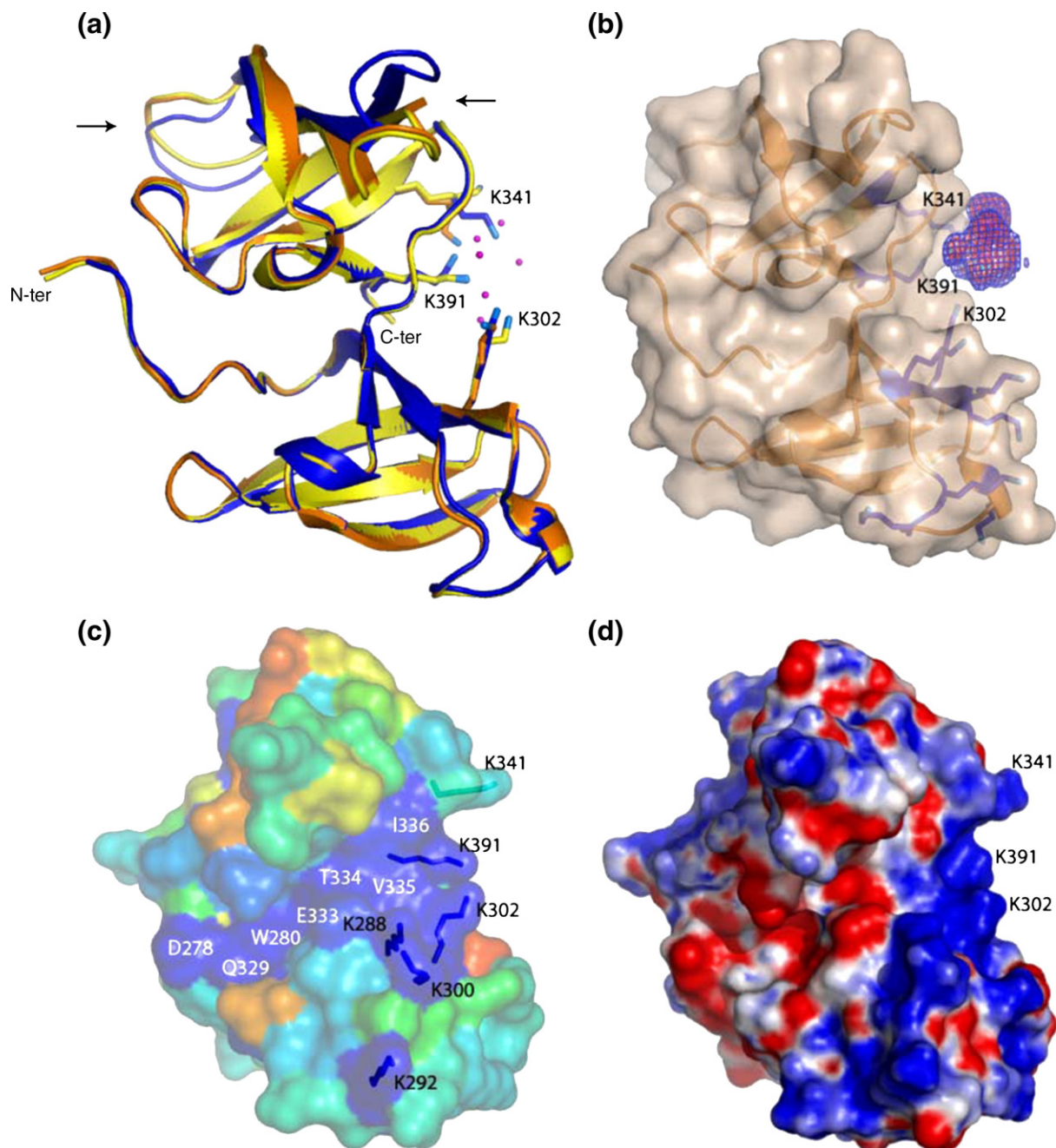


Figure 5. Localisation of the tungstate binding sites in the hKIN17 C-terminal domain. (a) Superimposition on subdomains A of the C-terminal domain of hKIN17 from native (orange), Na_2WO_4 -soaked crystal (yellow) and from the co-crystal (blue). Lysine residues involved in tungstate binding are shown in the corresponding backbone colour. Tungstate anions from the soaked crystal are shown in magenta. Quantification of the conformation changes within the subdomains (in particular, in loops $\beta 2\beta 3$ and $\beta 3\beta 4$ of subdomain B, as indicated by arrows) and of the different positioning of subdomain B is reported in Table 2. (b) Ribbon and surface representation of the hKIN17 C-terminal domain structure. The omit map was calculated using CNS⁵⁸ around the tungstate anions cluster with data collected for the co-crystal. Key lysine residues surrounding the tungstate anions are labelled. (c) A projection of the CONSURF results on the surface of C-terminal domain of hKIN17 indicates the degree of conservation from dark blue (highly conserved) to red (variable). The orientation of the structure has been rotated slightly relative to (b), in order to distinguish the five lysine residues in the positively charged groove. The side-chains of these lysine residues putatively involved in RNA binding are displayed. (d) Surface representation of the C-terminal domain of hKIN17 coloured by electrostatic potential and shown in the same orientation as (c). Blue and red colours indicate positive and negative electrostatic surface potential, respectively.

interactions between proteins and target RNA sequences, thereby contributing to increase binding affinity and specificity. RNA-binding domains

with limited specificity at the individual level can achieve high functional selectivity *via* cooperativity between their modular domains or *via* the

Table 2. Superimposition of the two subdomains of the hKIN17 C-terminal domain

	Native versus soaked crystal	Native versus co-crystal
Superimposition on subdomains A	rmsd _{subdomains A} 0.083	rmsd _{subdomains A} 0.424
	rmsd _{subdomains B} 0.206	rmsd _{subdomains B} 1.801
Superimposition on subdomains B	rmsd _{subdomains A} 0.212	rmsd _{subdomains A} 3.174
	rmsd _{subdomains B} 0.115	rmsd _{subdomains B} 0.781

The rmsd values are calculated on 51 C α atoms for subdomains A (spanning residues 284–334 of the hKIN17 sequence) and 52 C α atoms for subdomains B (spanning residues 340–391 of the hKIN17 sequence).

formation of complexes.³⁸ In the case of hKIN17, the tandem SH₃-like domain could cooperate with the N-terminal zinc-finger domain to associate optimally with specific RNAs.

A highly conserved groove, comprising K302 and K391 and other positively charged residues, is characteristic of the KIN17 family

Analysis of phylogenetic conservation of the C-terminal domain of hKIN17 using PSI-blast revealed that this domain is conserved in KIN17 proteins from metazoans, plants and parasites. Apart from the conserved buried amino acid residues that form part of the protein core, the CONSURF results projected onto the three-dimensional structure of hKIN17 C-terminal domain reveals conserved residues on the surface of the protein that could be structural determinants for interactions with other biomolecules.

Five lysine residues (K288, K292, K300, K302 and K391) are part of a highly conserved and positively charged surface (Figure 5(c)), surrounded by two interfacial exposed and conserved hydrophobic residues (V335 and I336). This surface is essentially located on the $\beta_1\beta_2$ hairpin of the first SH₃-like domain; it also comprises two residues of the linker and one residue at the C terminus of the second SH₃-like domain. On the basis of the surface charge distribution of RNA, we proposed that this positively charged groove may act as an RNA-binding region. This is in agreement with the conclusion that K302 and K391 are involved in RNA binding. A similar situation, where lysine residues from different subdomains located on the groove borders are involved in RNA binding, is found for the *Bacillus stearothermophilus* ribosomal protein L2. In the double- β -barrel protein, one of the subdomains adopts an SH₃-like fold and the other adopts an OB-fold.³⁹ The two subdomains are connected by a four-residue $_310$ helix and are packed tightly approximately perpendicular one to the other. On the basis of mutagenesis data, the authors proposed that protein L2 has a large contact surface for the target RNA in the cleft formed by the interface region of the two subdomains. This RNA-binding

surface of the ribosomal protein L2 comprises positively charged and hydrophobic residues of both subdomains.

At the interface between the two SH₃-like domains of hKIN17 and next to the positively charged surface, another set of highly conserved residues (D278, E333 and E387) forms a mainly negatively charged cleft, and three other conserved residues (W280, Q329 and T334) form the bottom of the cleft. These residues may constitute a surface for other interactions.

The search for analogs of the C-terminal domain of hKIN17 using PSI-blast highlighted significant sequence similarity with the C-terminal region of the metazoan and plant transcription elongation factors SPT5, belonging to the DSIF complex, which exerts both positive and negative effects on RNA polymerase II.^{40,41} In addition, the search identified the poorly studied T54 proteins from metazoa and plants. These sequences are not as well conserved as the KIN17 sequence, in particular concerning the positively charged groove.

Functional significance of the tandem SH₃-like domain

In general, the interaction of the ribosomal SH₃-like domains with RNA is mediated by the variable loops that connect the β -strands.⁴² Variability in the length of the three first loops ($\beta_1\beta_2$, $\beta_2\beta_3$, $\beta_3\beta_4$) is a characteristic of SH₃-like domains and is mainly responsible for the high level of functional versatility of this fold. Consistently, a DALI search using the structures of the hKIN17 SH₃-like subdomains revealed that these are structurally highly similar to SH₃-like domains, exhibiting various functions, from signal transduction (tyrosine kinase) to nucleic acid binding (ribosomal protein L21e).

Only a few three-dimensional structures of tandem SH₃-like domains have been determined. The tandem Tudor domain of human 53BP1,^{43,44} the tandem Tudor domain of FMRP,⁴⁵ the double chromodomains of human CHD1⁴⁶ and the double

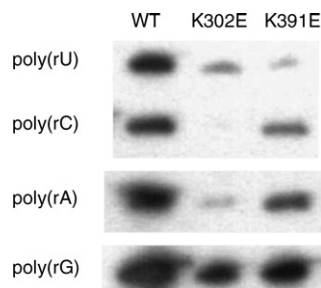


Figure 6. K302E and K391E mutants have a reduced RNA binding activity. The purified wild-type hKIN17 C-terminal domain and the K302E and K391E mutant domains, in fusion with GST, were incubated with each ribonucleotide homopolymers and washed with 0.1 M NaCl. The proteins retained by the immobilized homopolymer RNAs were analysed by immunoblotting. Both mutants show a reduced binding activity for the four homopolymers.

Tudor domain of JMJD2A⁴⁷ were all reported to bind to methylated peptides through a conserved aromatic cage. However, hKIN17 SH₃-like subdomains do not exhibit any solvent-exposed aromatic clusters, suggesting that they do not target similar methylated peptides. Furthermore, the association of a region of 53BP1, containing the tandem Tudor domains, with IR-induced foci is sensitive to treatment with ribonuclease, suggesting that another function of double Tudor of 53BP1 is to bind RNA.⁴⁸ Thus, RNA binding could represent a more general function of tandem SH₃-like domains. In this study, a novel RNA-binding site in a double SH₃-like domain is characterized even if it does not exclude the possibility that this domain is involved also in protein-protein interaction.

Conclusion

RNA-binding proteins are involved in a wide range of cellular processes in the nucleus and in the cytoplasm, including regulation of pre-mRNA splicing, mRNA stability, translation efficiency and transport of RNAs between the nucleus and the cytoplasm. An RNA-binding function for hKIN17 is in accordance with the identification of hKIN17 as a component of the human spliceosome.^{14,15} Furthermore, a Tudor-like domain structurally similar to domain III of NusG and to the subdomains of the hKIN17 C-terminal domain was found recently in the bacterial transcription-repair coupling factor (TRCF), where it mediates interaction with the β subunit of RNA polymerase.⁴⁹ Consistently, the C-terminal domain of hKIN17 as the whole protein co-immunoprecipitate with the RNA polymerase II in HeLa nuclear extracts (G.P.-L., unpublished results). Further investigation will be required to identify which cellular RNA interacts with hKIN17 and to understand how hKIN17 functions in RNA metabolism.

Materials and Methods

Cloning, expression and purification of the domains of hKIN17 protein

The C-terminal domain comprising residues 268–393 of hKIN17 protein was prepared as described.²² Mutant proteins (K302E and K391E) were engineered by site-directed mutagenesis using QuikChange Site-Directed Mutagenesis Kit (Stratagene) and expressed and purified as the wild-type protein. The region comprising residues 1–160 of hKIN17 protein and including the zinc finger (residues 23–50) was prepared as described.²⁴ The region comprising residues 51–160 of hKIN17 protein was prepared following the protocol used for the C-terminal domain.²²

Crystallization, X-ray data collection and refinement

Crystallization and initial phasing of the native structure have been described.²² Briefly, a halide heavy-atom

derivative was prepared by brief-soaking of a native crystal in a solution of KI. SHARP was run to calculate single isomorphous replacement with anomalous scattering phases at 1.45 Å resolution.⁵⁰ After solvent flattening with SOLOMON,⁵¹ the polypeptide chain was traced using a combination of ARP/wARP⁵² and manual rebuilding. Model refinement was carried out with REFMAC5.⁵³

Tungstate derivatives were prepared by both soaking and co-crystallization techniques as described.²² Native crystals were soaked overnight in the crystallisation solution (100 mM sodium acetate (pH 6.3), 10% (v/v) ethylene glycol, 27% (w/v) PEG 6000, 0.5 M LiCl) with 1 mM Na₂WO₄ added. The hKIN17 C-terminal domain was co-crystallized with tungstate in a different crystallisation solution (100 mM sodium acetate (pH 6.3), 25% ethylene glycol, 20% PEG 6000) with 0.5 mM Na₂WO₄ added and co-crystals were then transferred to the cryoprotectant solution (100 mM sodium acetate (pH 6.3), 25% ethylene glycol, 20% PEG 6000) with 100 mM Na₂WO₄ added.

Data were collected using synchrotron radiation at a wavelength of 1.2144 Å, an energy where tungsten scatters X-rays with an anomalous contribution (f'') greater than ten electrons. Molecular replacement was carried out using Molrep and refinement using REFMAC5⁵³ with data collected on a W-soaked crystal. Strong electron density peaks were identified in an anomalous difference Fourier map calculated to locate the tungsten atoms. With the data collected on a co-crystal, AutoSHARP⁵⁴ was used to calculate single wavelength anomalous diffraction (SAD) phases. It was found to be crucial to reduce the default value for the minimum allowed distance between sites for heavy-atom detection in SHELXD as well as for analysis of residual maps in AutoSHARP, since the tungstate anions are best represented as a cluster. Model refinement was then carried out with REFMAC5.⁵³

North-western blot analysis

The 1200 bp RNA probe corresponding to *Kin17* ORF was prepared as described.¹⁶ To allow renaturation of the recombinant proteins after they were separated by SDS-PAGE and transferred onto nitrocellulose membranes, the protein blots were treated for 1–2 h at room temperature in a hybridization buffer (10 mM Tris-HCl (pH 7.4), 50 mM NaCl, 1X Denhardt's solution, 1 mM EDTA) in the presence of 1 μ g/ml of *E. coli* tRNA (Sigma). The blots were then probed at room temperature for 1 h with radiolabelled RNA probe (100,000 cpm for a 2 ml reaction volume) in a hybridization buffer containing 2 μ g/ml of *E. coli* tRNA. Blots were washed twice for 15 min in the hybridization buffer and exposed to X-ray film or to a PhosphorImager screen.

RNA homopolymer binding assay

RNA homopolymer-conjugated agarose beads (poly(rC), poly(rG), poly(rU)) and conjugated Sepharose beads (poly(rA)) were obtained from Sigma. Binding of the purified recombinant protein to RNA homopolymers was performed as described.¹⁶ A portion (20 μ l) of each RNA homopolymer bead was incubated with an equal amount of purified protein (3 μ g), for 30 min at 4 °C with gentle agitation in a total volume of 500 μ l of binding buffer (20 mM Tris-HCl (pH 7.4), 2.5 mM MgCl₂, 0.5% (v/v) Triton X-100 and 100 mM, 250 mM, 500 mM or 1 M NaCl). After incubation, the beads were washed five times with 500 μ l of binding buffer. Some binding reactions or washes

were performed in the presence of 1 mg/ml of heparin or 1 mg/ml of *E. coli* tRNA. Proteins bound to the RNA homopolymer beads were eluted by boiling in SDS loading buffer and loaded onto an SDS/15% (w/v) polyacrylamide electrophoresis gel. After electrophoretic separation, proteins were transferred onto a nitrocellulose filter (Schleicher and Schuell) and detected with anti-GST HRP conjugate antibody (Amersham Biosciences) or anti-KIN17 mAb Igk36 antibody,⁵⁵ depending on the presence or the absence of a GST tag fused to the proteins.

NMR spectroscopy studies

NMR samples (about 1 mM) were prepared in PBS (pH 7.0) containing 10% ²H₂O, 1 mM EDTA, a protease inhibitor cocktail (Sigma), 1 mM NaN₃ and 3-(trimethylsilyl) [2,2,3,3-²H₄] propionate (TSP) was added as a chemical shift reference. The 1D ¹H, 2D ¹H-¹H nuclear Overhauser effect spectroscopy (NOESY) and 2D ¹H-¹H total correlated spectroscopy (TOCSY) experiments were carried out at 297 K on a Bruker DRX-500 spectrometer. All spectra were processed and compared with the program Xwinnmr (Bruker).

Sequence analysis

The amino acid sequences comparisons were optimized using the program MUSCLE,⁵⁶ and were submitted to CONSURF,⁵⁷ in order to calculate for each residue a conservation score based on residue similarity and phylogenetic distance.

Protein Data Bank accession number

The coordinates and structure factors of the X-ray structure have been deposited with the RCSB Protein Data Bank under accession number 2CKK.

Acknowledgements

We thank C. Vonrhein for assistance with auto-SHARP and for helpful discussions. We thank JB. Charbonnier for critical review of the manuscript. We thank B. Gilquin for helpful discussions and MH. Le Du for assistance with crystallography. The laboratory of JF. Angulo work was supported by contract N° 8702 of Electricité de France.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2006.09.033](https://doi.org/10.1016/j.jmb.2006.09.033)

References

1. Biard, D. S., Saintigny, Y., Maratrat, M., Paris, F., Martin, M. & Angulo, J. F. (1997). Enhanced expres-

- sion of the Kin17 protein immediately after low doses of ionizing radiation. *Radiat. Res.* **147**, 442–450.
2. Masson, C., Menaa, F., Pinon-Lataillade, G., Frobert, Y., Chevillard, S., Radicella, J. P. *et al.* (2003). Global genome repair is required to activate KIN17, a UVC-responsive gene involved in DNA replication. *Proc. Natl Acad. Sci. USA*, **100**, 616–621.
3. Kannouche, P., Mauffrey, P., Pinon-Lataillade, G., Mattei, M. G., Sarasin, A., Daya-Grosjean, L. & Angulo, J. F. (2000). Molecular cloning and characterization of the human KIN17 cDNA encoding a component of the UVC response that is conserved among metazoans. *Carcinogenesis*, **21**, 1701–1710.
4. Biard, D. S., Kannouche, P., Lannuzel-Drogou, C., Mauffrey, P., Apiou, F. & Angulo, J. F. (1999). Ectopic expression of (Mm)Kin17 protein inhibits cell proliferation of human tumor-derived cells. *Expt. Cell Res.* **250**, 499–509.
5. Kannouche, P. & Angulo, J. F. (1999). Overexpression of kin17 protein disrupts nuclear morphology and inhibits the growth of mammalian cells. *J. Cell Sci.* **112**, 3215–3224.
6. Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B. *et al.* (1999). Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science*, **285**, 901–906.
7. Biard, D. S., Despras, E., Sarasin, A. & Angulo, J. F. (2005). Development of new EBV-based vectors for stable expression of small interfering RNA to mimic human syndromes: application to NER gene silencing. *Mol. Cancer Res.* **3**, 519–529.
8. Biard, D. S., Miccoli, L., Despras, E., Harper, F., Pichard, E., Creminon, C. & Angulo, J. F. (2003). Participation of kin17 protein in replication factories and in other DNA transactions mediated by high molecular weight nuclear complexes. *Mol. Cancer Res.* **1**, 519–531.
9. Miccoli, L., Frouin, I., Novac, O., Di Paola, D., Harper, F., Zannis-Hadjopoulos, M. *et al.* (2005). The human stress-activated protein kin17 belongs to the multi-protein DNA replication complex and associates in vivo with mammalian replication origins. *Mol. Cell Biol.* **25**, 3814–3830.
10. Mazin, A., Milot, E., Devoret, R. & Chartrand, P. (1994). KIN17, a mouse nuclear protein, binds to bent DNA fragments that are found at illegitimate recombination junctions in mammalian cells. *Mol. Gen. Genet.* **244**, 435–438.
11. Mazin, A., Timchenko, T., Menissier-de Murcia, J., Schreiber, V., Angulo, J. F., de Murcia, G. & Devoret, R. (1994). Kin17, a mouse nuclear zinc finger protein that binds preferentially to curved DNA. *Nucl. Acids Res.* **22**, 4335–4341.
12. Milot, E., Belmaaza, A., Wallenburg, J. C., Gusew, N., Bradley, W. E. & Chartrand, P. (1992). Chromosomal illegitimate recombination in mammalian cells is associated with intrinsically bent DNA elements. *EMBO J.* **11**, 5063–5070.
13. Timchenko, T., Bailone, A. & Devoret, R. (1996). Btcd, a mouse protein that binds to curved DNA, can substitute in *Escherichia coli* for H-NS, a bacterial nucleoid protein. *EMBO J.* **15**, 3986–3992.
14. Makarov, E. M., Makarova, O. V., Urlaub, H., Gentzel, M., Will, C. L., Wilm, M. & Luhrmann, R. (2002). Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science*, **298**, 2205–2208.
15. Rappsilber, J., Ryder, U., Lamond, A. I. & Mann, M.

- (2002). Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **12**, 1231–1245.
16. Pinon-Lataillade, G., Masson, C., Bernardino-Sgherri, J., Henriot, V., Mauffrey, P., Frobert, Y. *et al.* (2004). KIN17 encodes an RNA-binding protein and is expressed during mouse spermatogenesis. *J. Cell Sci.* **117**, 3691–3702.
 17. Kyrpides, N. C., Woese, C. R. & Ouzounis, C. A. (1996). KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem. Sci.* **21**, 425–426.
 18. Ponting, C. P. (2002). Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucl. Acids Res.* **30**, 3643–3652.
 19. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, **289**, 905–920.
 20. Steiner, T., Kaiser, J. T., Marinkovic, S., Huber, R. & Wahl, M. C. (2002). Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities. *EMBO J.* **21**, 4641–4653.
 21. Gaboriaud, C., Bissery, V., Benchetrit, T. & Mornon, J. P. (1987). Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Letters*, **224**, 149–155.
 22. le Maire, A., Schiltz, M., Braud, S., Gondry, M., Charbonnier, J. B., Zinn-Justin, S. & Stura, E. (2006). Crystallization and halide phasing of the C-terminal domain of human KIN17. *Acta Crystallog. sect. F*, **62**, 245–248.
 23. Lo Conte, L., Chothia, C. & Janin, J. (1999). The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198.
 24. Carlier, L., le Maire, A., Braud, S., Masson, C., Gondry, M., Zinn-Justin, S. *et al.* (2006). NMR assignment of region 51–160 of human KIN17, a DNA and RNA-binding protein. *J. Biomol. NMR*. (Electronic publication).
 25. Adinolfi, S., Bagni, C., Musco, G., Gibson, T., Mazzarella, L. & Pastore, A. (1999). Dissecting FMR1, the protein responsible for fragile X syndrome, in its structural and functional domains. *Rna*, **5**, 1248–1258.
 26. Bertrand, S., Burlet, P., Clermont, O., Huber, C., Fondrat, C., Thierry-Mieg, D. *et al.* (1999). The RNA-binding properties of SMN: deletion analysis of the zebrafish orthologue defines domains conserved in evolution. *Hum. Mol. Genet.* **8**, 775–782.
 27. Siomi, H., Siomi, M. C., Nussbaum, R. L. & Dreyfuss, G. (1993). The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell*, **74**, 291–298.
 28. Pinol-Roma, S., Choi, Y. D., Matunis, M. J. & Dreyfuss, G. (1988). Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes Dev.* **2**, 215–227.
 29. Swanson, M. S. & Dreyfuss, G. (1988). Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. *Mol. Cell Biol.* **8**, 2237–2241.
 30. Jones, S., Daley, D. T., Luscombe, N. M., Berman, H. M. & Thornton, J. M. (2001). Protein-RNA interactions: a structural analysis. *Nucl. Acids Res.* **29**, 943–954.
 31. Egloff, M. P., Cohen, P. T., Reinemer, P. & Barford, D. (1995). Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. *J. Mol. Biol.* **254**, 942–959.
 32. Barford, D., Flint, A. J. & Tonks, N. K. (1994). Crystal structure of human protein tyrosine phosphatase 1B. *Science*, **263**, 1397–1404.
 33. Davies, D. R., Interthal, H., Champoux, J. J. & Hol, W. G. (2002). Insights into substrate binding and catalytic mechanism of human tyrosyl-DNA phosphodiesterase (Tdp1) from vanadate and tungstate-inhibited structures. *J. Mol. Biol.* **324**, 917–932.
 34. Gray, C. H., Good, V. M., Tonks, N. K. & Barford, D. (2003). The structure of the cell cycle protein Cdc14 reveals a proline-directed protein phosphatase. *EMBO J.* **22**, 3524–3535.
 35. Longenecker, K. L., Roach, P. J. & Hurley, T. D. (1996). Three-dimensional structure of mammalian casein kinase I: molecular basis for phosphate recognition. *J. Mol. Biol.* **257**, 618–631.
 36. Stuckey, J. A., Schubert, H. L., Fauman, E. B., Zhang, Z. Y., Dixon, J. E. & Saper, M. A. (1994). Crystal structure of Yersinia protein tyrosine phosphatase at 2.5 Å and the complex with tungstate. *Nature*, **370**, 571–575.
 37. Brown, R. S. (2005). Zinc finger proteins: getting a grip on RNA. *Curr. Opin. Struct. Biol.* **15**, 94–98.
 38. Singh, R. & Valcarcel, J. (2005). Building specificity with nonspecific RNA-binding proteins. *Nature Struct. Mol. Biol.* **12**, 645–653.
 39. Nakagawa, A., Nakashima, T., Taniguchi, M., Hosaka, H., Kimura, M. & Tanaka, I. (1999). The three-dimensional structure of the RNA-binding domain of ribosomal protein L2; a protein at the peptidyl transferase center of the ribosome. *EMBO J.* **18**, 1459–1467.
 40. Yamaguchi, Y., Narita, T., Inukai, N., Wada, T. & Handa, H. (2001). SPT genes: key players in the regulation of transcription, chromatin structure and other cellular processes. *J. Biochem. (Tokyo)*, **129**, 185–191.
 41. Zorio, D. A. & Bentley, D. L. (2001). Transcription elongation: the 'Foggy' is lifting ellipsis. *Curr. Biol.* **11**, R144–R146.
 42. Klein, D. J., Moore, P. B. & Steitz, T. A. (2004). The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. *J. Mol. Biol.* **340**, 141–177.
 43. Charier, G., Couprie, J., Alpha-Bazin, B., Meyer, V., Quemeneur, E., Guerois, R. *et al.* (2004). The Tudor tandem of 53BP1: a new structural motif involved in DNA and RG-rich peptide binding. *Structure*, **12**, 1551–1562.
 44. Huyen, Y., Zgheib, O., Ditullio, R. A., Jr, Gorgoulis, V. G., Zacharatos, P., Petty, T. J. *et al.* (2004). Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature*, **432**, 406–411.
 45. Ramos, A., Hollingworth, D., Adinolfi, S., Castets, M., Kelly, G., Frenkiel, T. A. *et al.* (2006). The structure of the N-terminal domain of the fragile X mental retardation protein: a platform for protein-protein interaction. *Structure*, **14**, 21–31.
 46. Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y. *et al.* (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature*, **438**, 1181–1185.
 47. Huang, Y., Fang, J., Bedford, M. T., Zhang, Y. & Xu, R. M. (2006). Recognition of histone H3 lysine-4 methylation by the double Tudor domain of JMJD2A. *Science*, **312**, 748–751.
 48. Pryde, F., Khalili, S., Robertson, K., Selfridge, J., Ritchie, A. M., Melton, D. W. *et al.* (2005). 53BP1 exchanges slowly at the sites of DNA damage and appears to require RNA for its association with chromatin. *J. Cell Sci.* **118**, 2043–2055.

49. Deaconescu, A. M., Chambers, A. L., Smith, A. J., Nickels, B. E., Hochschild, A., Savery, N. J. & Darst, S. A. (2006). Structural basis for bacterial transcription-coupled DNA repair. *Cell*, **124**, 507–520.
50. Bricogne, G., Vornrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallog. sect. D*, **59**, 2023–2030.
51. Abrahams, J. P. & Leslie, A. G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallog. sect. D*, **52**, 30–42.
52. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463.
53. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
54. Vornrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2005). Automated structure solution with auto-SHARP. In *Crystallographic Methods* (Doublé, S., ed), Humana Press, Totowa, NJ, USA.
55. Miccoli, L., Biard, D. S., Creminon, C. & Angulo, J. F. (2002). Human kin17 protein directly interacts with the simian virus 40 large T antigen and inhibits DNA replication. *Cancer Res.* **62**, 5425–5435.
56. Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* **32**, 1792–1797.
57. Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E. & Ben-Tal, N. (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics*, **19**, 163–164.
58. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.

Edited by J. Doudna

(Received 5 July 2006; received in revised form 24 August 2006; accepted 11 September 2006)
Available online 16 September 2006