Crystal structure of a tankyrase 1–telomere repeat factor 1 complex

Bo Li,^a Ruihong Qiao,^a Zhizhi Wang,^b Weihong Zhou,^a Xin Li,^a Wenqing Xu^{b*} and Zihe Rao^{a*}

^aCollege of Life Sciences, Nankai University, Weijin Road #94, Tianjin 300071, People's Republic of China, and

^bDepartment of Biological Structure, University of Washington, 1959 NE Pacific Street, Seattle, Washington 98195, USA.

*Correspondence e-mail: wxu@u.washington.edu, raozh@nankai.edu.cn, raozh@nankai.edu.cn

Received 11 February 2016

Accepted 11 March 2016

Edited by L. J. Beamer, University of Missouri, USA

Keywords: TNKS; tankyrase; telomere; TRF1; telomere repeat factor 1; PARYlation; complex structure.

PDB reference: TNKS1–TRF1 complex, 5hkp

Supporting information: this article has supporting information at journals.iucr.org/f

Telomere repeat factor 1 (TRF1) is a subunit of shelterin (also known as the telosome) and plays a critical role in inhibiting telomere elongation by telomerase. Tankyrase 1 (TNKS1) is a poly(ADP-ribose) polymerase that regulates the activity of TRF1 through poly(ADP-ribosyl)ation (PARYlation). PARYlation of TRF1 by TNKS1 leads to the release of TRF1 from telomeres and allows telomerase to access telomeres. The interaction between TRF1 and TNKS1 is thus important for telomere stability and the mitotic cell cycle. Here, the crystal structure of a complex between the N-terminal acidic domain of TRF1 (residues 1–55) and a fragment of TNKS1 covering the second and third ankyrin-repeat clusters (ARC2-3) is presented at 2.2 Å resolution. The TNKS1–TRF1 complex crystals were optimized using an ‘oriented rescreeing’ strategy, in which the initial crystallization condition was used as a guide for a second round of large-scale sparse-matrix screening. This crystallographic and biochemical analysis provides a better understanding of the TRF1–TNKS1 interaction and the three-dimensional structure of the ankyrin-repeat domain of TNKS.

1. Introduction

Telomeres function to protect the ends of chromosomes and are critical for cell proliferation and survival. Owing to their unique structure, a six-subunit complex called shelterin is required to maintain telomeres, in particular during DNA replication and chromosome separation at mitosis (Martínez & Blasco, 2011; Diotti & Loayza, 2011; Longhese, 2012). The TRF1 protein encoded by the *TERF1* gene is one of the subunits of shelterin. It interacts directly with the telomere DNA and inhibits telomere elongation by telomerase (de Lange, 2005; Walker & Zhu, 2012). A key mechanism for regulating shelterin/TRF1 activity is PARYlation of TRF1 by tankyrase, which promotes the release of TRF1 from telomeres and allows telomere elongation by telomerase (de Lange, 2005; Diotti & Loayza, 2011; Hsiao & Smith, 2008; De Boeck *et al.*, 2009). High levels of TRF1 have been associated with colon, kidney, gastric and lung cancers, and TRF1 abundance correlates with telomere length in tumor cells, suggesting that TRF1 may play an important role in tumorigenesis (Miyachi *et al.*, 2002; Garcia-Aranda *et al.*, 2006; Valls-Bautista *et al.*, 2012; Pal *et al.*, 2015; Nakanishi *et al.*, 2003). The interaction between TNKS and TRF1 is critical for TRF1 PARYlation and telomere regulation.

There are two tankyrases, tankyrase 1 and 2 (TNKS1 and TNKS2), which have similar domain structures and are

functionally redundant in many cases. Both TNKS1 and TNKS2 contain 20 ankyrin repeats clustered into five ankyrin-repeat clusters (ARCs; Fig. 1*a*), which are responsible for TNKS localization and substrate recruitment (Hsiao & Smith, 2008; Haikarainen *et al.*, 2014). C-terminal to the TNKS ankyrin-repeat domain are the homology to sterile alpha motif (SAM) oligomerization and poly(ADP-ribose) polymerase (PARP) catalytic domains. Our current understanding of the structural basis of TNKS–ligand interaction has been derived from crystallographic analyses of the TNKS2(ARC4)–3BP2 and TNKS1(ARC2–3)–Axin complexes (Morrone *et al.*, 2012; Guettler *et al.*, 2011). These structural studies, along with previous biochemical analyses, showed that TNKS ARC1, ARC2, ARC4 and ARC5 can individually recognize a TNKS-binding motif (TBM) with a consensus sequence **RxxG/P/AxG**, with the highlighted Arg and Gly being strictly conserved (Morrone *et al.*, 2012; Seimiya & Smith, 2002; Guettler *et al.*, 2011; Sbodio & Chi, 2002; Seimiya *et al.*, 2004). ARC3 of TNKS1/2 does not interact with TBM, but may form a dimeric structure (Morrone *et al.*, 2012).

The human TRF1 protein contains an N-terminal acidic domain (residues 1–58) which is responsible for TNKS binding (Fig. 1*a*). A core TBM has been identified within the acidic domain (**RGCADG**; residues 13–18 in human TRF1). Previous work showed that the entire acidic domain binds to TNKS1 much more tightly than this core TBM (Sbodio & Chi, 2002). The crystal structure of the six-residue **RGCADG** sequence embedded in a 16-residue 3BP2 protein sequence in complex with ARC4 of TNKS2 (PDB entry 3tw5) has been reported (Guettler *et al.*, 2011). However, it remains unclear how the TRF1 residues flanking the core **RGCADG** motif may interact with TNKS1/2, in particular with TNKS ARCs other than ARC4, and whether the TRF1 flanking sequences also affect the binding conformation of its core motif. Here, we report a 2.2 Å resolution crystal structure of a complex between the human TRF1 fragment 1–55 and a mouse TNKS1(ARC2–3) fragment (residues 308–655), which is 99% identical to the corresponding human TNKS1 sequence. In addition, we performed GST-pulldown and isothermal titration calorimetric (ITC) analyses of the interactions between

Figure 1

(*a*) Domain structures of TNKS1 and TRF1. Schematic representation of the domain organizations of mTNKS1 and hTRF1. ARC, ankyrin-repeat cluster; SAM, homology to sterile alpha motif; PARP, poly(ADP-ribose) polymerase catalytic domain; Acidic, acidic domain of TRF1; TRFH, TRF1 homology domain; Myb, Myb-type DNA-binding motif. Red boxes highlight the TNKS1 and TRF1 regions which are included in the crystal structure. (*b*) Sequence alignment of the acidic domain of TRF1. The sequence used for crystallization is boxed with green dashed lines. Residues with clear electron density in the crystal structure are boxed with green solid lines. The two most critical TRF1 residues (Arg13 and Gly18) involved in interaction with TNKS1 are highlighted with green stars. (*c*) Overall structure of the TNKS1–TRF1 complex in two orthogonal orientations. In each asymmetric unit there are two TNKS1(ARC2–3) molecules (colored in green and cyan, respectively) and two TRF1(1–55) molecules (colored in purple and orange, respectively). The six blue arrows indicate the six inter-ARC linker helices, three in each TNKS1(ARC2–3) molecule.

Table 1
Macromolecule-production information.

	mTNKS1	hTRF1
Source organism	<i>Mus musculus</i>	<i>Homo sapiens</i>
DNA source	Mouse cDNA library	Human cDNA library
Forward primer†	GGAATTCGAAAACCTGTATTTTCAGGGCGGGAAATCAGCTCTGGAC	GGAATTCATATGGCGGAGGAAGTTTCCTCAG
Reverse primer†	ATAAGAATGCGGCCGCTAGTCACCAGCTTTCGATGC	ATAAGAATGCGGCCGCTCACTCGGGGGCCC
Cloning vector	pGEX-4T-1	pCool
Expression vector	pGEX-4T-1	pCool
Expression host	<i>E. coli</i> Transetta (DE3)	<i>E. coli</i> Transetta (DE3)
Complete amino-acid sequence of the construct produced	GKSALDLADPSAKAVLTGEYKKDELLEAARSGNEEKLMLLTPLVNCH- ASDGRKSTPLHLAAGYNRVIRIVQLLLQHGDVHAKDKGGLVPLHNAC- SYGHYEVTELLKHGACVNMAMDWQFTPLHEAASKNRVEVCSLLLSH- GADPTLVNCHGKSADVMAPTPELRERLTYEFGHSLQAAREADLAK- VKKTLALEIINFKQPQSHETALHCAVASLHPKRKQVAELLLRKGANV- NEKNKDFMTPLHVAERAHNDVMEVLHKGAKMNALDSLQGTALHRA- ALAGHLQTCRLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPM- RTSDVDYRLLEASKAGD	MAEEVSSAAPSPPRCADGRDADPTTEQMAETERNDEEQFECQELLECV- QVGAPE

† The restriction sites are underlined; the TEV cleavage site is indicated in italics.

TRF1 and TNKS1 (both wild type and mutants), which provide further insights into the TRF1–TNKS interaction.

2. Materials and methods

2.1. Expression and purification of mTNK1(ARC2-3) and hTRF1(1–55)

Mouse tankyrase 1 (mTNKS1; residues 308–655) was cloned into pGEX-4T-1 vector (GE Healthcare) with a TEV protease-cleavage site following the GST tag (Morrone *et al.*, 2012). Human TRF1 (hTRF1) residues 1–55 were cloned into pCool vector, again with a TEV protease-cleavage site following the GST tag (Addgene). The two constructs were transformed individually into *Escherichia coli* Transetta (DE3) competent cells (TransGen Biotech). Cultured cells were grown in the presence of 100 µg ml^{−1} ampicillin at 310 K with vigorous shaking until an OD₆₀₀ of 0.8 was reached and were then chilled to 289 K and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16–18 h. The cell pellets were harvested by centrifugation at 4000g for 15 min and stored at 253 K. Cells were lysed by sonication on ice in buffer A [consisting of 20 mM Tris–HCl pH 8.0, 200 mM NaCl, 5% glycerol, 5 mM dithiothreitol (DTT)] and centrifuged to remove cell debris. The cell supernatant was injected onto a Glutathione Sepharose 4B column (GE Healthcare) pre-equilibrated with buffer A. Following extensive washing with buffer B (consisting of 20 mM Tris–HCl pH 8.0, 400 mM NaCl, 5% glycerol, 5 mM DTT) to remove nonspecifically bound proteins, the target proteins were cleaved on the column with TEV protease to remove the GST tag at 277 K overnight. After the affinity-chromatography step, mTNKS1 was further purified on a HiTrap Q HP 5 ml column (GE Healthcare) and eluted with a gradient of NaCl (in 20 mM Tris–HCl pH 8.0 buffer, 5% glycerol, 5 mM DTT). The main peak fractions from the HiTrap Q HP column were collected, concentrated and further purified using a Superdex 200 10/300 GL column (GE Healthcare). The purified mTNKS1 was concentrated to approximately 10 mg ml^{−1} in 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 5% glycerol, 5 mM DTT and stored as aliquots

at 193 K. Human TRF1(1–55) was purified using a very similar procedure, except for the last size-exclusion chromatography (SEC) step, in which a Superdex 75 column was used instead of a Superdex 200 column. Every purification step above was analyzed by SDS–PAGE to check the purities of the target proteins. Macromolecule-production information is summarized in Table 1.

2.2. Crystallization of the TNK1(ARC2-3)–TRF1(1–55) complex

Mouse TNKS1(ARC2-3) and excess human TRF1(1–55) were mixed and incubated on ice for 1 h to form the complex. Crystallization conditions for the TNKS1(ARC2-3)–TRF1(1–55) complex were screened by the sitting-drop vapor-diffusion method at 293 K (1 µl protein solution plus 1 µl reservoir solution) using multiple commercially available crystallization screening kits. The initial crystals were obtained using reagent No. 3 [0.05 M calcium chloride dihydrate, 0.1 M MES monohydrate pH 6.0, 45%(v/v) polyethylene glycol 200] from the PEGRx 2 crystallization kit (Hampton Research). Based on this initial crystallization condition, we performed extensive rescreening using an ‘oriented rescreening’ strategy. In this approach, each of the rescreening reservoir solution contained 70% (by volume) of initial crystallization condition solution and 30% (by volume) of a screening kit solution (from multiple commercially available crystallization screening kits). This approach allowed us to obtain several different crystal forms. The crystal that diffracted to the best resolution was obtained from 35 mM calcium chloride dehydrate, 70 mM MES monohydrate pH 6.0, 30 mM Tris pH 8.5, 31.5% PEG 200, 6% ethanol. The crystals were flash-cooled in this ‘cryo-ready’ solution and stored in liquid nitrogen for data collection. Crystallization information is summarized in Table 2.

2.3. X-ray data collection and structure determination

A data set was collected to 2.2 Å resolution on the BL18U beamline at Shanghai Synchrotron Radiation Facility, People’s Republic of China. The data set was processed with *HKL-2000* (Otwinowski & Minor, 1997). The crystal belonged to space

Table 2
Crystallization.

Method	Sitting-drop vapor diffusion
Plate type	48-well plastic plate
Temperature (K)	293
Protein concentration (mg ml ⁻¹)	10
Buffer composition of protein solution	20 mM Tris-HCl pH 8.0, 150 mM NaCl, 5% glycerol, 5 mM DTT
Composition of reservoir solution	35 mM calcium chloride dehydrate, 70 mM MES monohydrate pH 6.0, 30 mM Tris pH 8.5, 31.5% PEG 200, 6% ethanol
Volume and ratio of drop	1 µl protein solution plus 1 µl reservoir solution
Volume of reservoir (µl)	100

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL18U, SSRF
Wavelength (Å)	0.978
Temperature (K)	100
Detector	Pilatus 6M
Crystal-to-detector distance (mm)	350.0
Rotation range per image (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	1
Space group	C2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	135.00, 100.07, 75.95
α , β , γ (°)	90, 107.54, 90
Mosaicity (°)	0.555
Resolution range (Å)	50.00–2.20 (2.26–2.20)
Total No. of reflections	149561
No. of unique reflections	48730
Completeness (%)	98.7 (97.7)
Multiplicity	3.1 (3.1)
$\langle I/\sigma(I) \rangle$	8.8 (2.7)
$R_{\text{r.i.m.}}$ [†]	0.133 (0.477)
$R_{\text{p.i.m.}}$ [‡]	0.073 (0.262)
Overall <i>B</i> factor from Wilson plot (Å ²)	56.20

[†] $R_{\text{r.i.m.}} = \sum_{hkl} \{ [N(hkl)/[N(hkl) - 1]]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl) \}$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity of multiple observations of symmetry-related reflections. [‡] $R_{\text{p.i.m.}} = \sum_{hkl} \{ 1/[N(hkl) - 1] \}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl) \}$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity of multiple observations from symmetry-related reflections.

group C2 and there are two TNKS1 molecules and two TRF1 peptides in the asymmetric unit (Table 3). The ARC2 and ARC3 domains from the TNKS1–Axin complex (PDB entry 3utm; Morrone *et al.*, 2012) were used as the initial search models for molecular replacement in *Phaser* (McCoy *et al.*, 2007). The 2.2 Å resolution diffraction data set was used for refinement (Table 4). The complex model was improved using iterative cycles of manual model building in *Coot* (Emsley *et al.*, 2010) and refinement with *REFMAC5* in *CCP4* (Winn *et al.*, 2011). TLS was used throughout refinement (Winn *et al.*, 2003). The diffraction data and final structural coordinates of the TNK1(ARC2-3)–TRF1(1-55) complex have been deposited in the PDB with accession code 5hkp.

2.4. Site-directed mutagenesis and purification of mutant proteins

Seven missense mutants of mTNKS1 and hTRF1(1-35) were generated using the Fast Mutagenesis System (TransGen

Table 4
Structure solution and refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	79,000–2.200 (2.253–2.196)
Completeness (%)	97.7 (85.4)
σ Cutoff	$F > 0\sigma(F)$
No. of reflections, working set	45650 (2946)
No. of reflections, test set	2425 (150)
Final R_{cryst}	0.198 (0.280)
Final R_{free}	0.229 (0.277)
Cruickshank DPI [†]	0.1989
No. of non-H atoms	
Protein	4943
Ligand	0
Solvent	106
Total	5049
R.m.s. deviations	
Bonds (Å)	0.009
Angles (°)	1.229
Average <i>B</i> factors (Å ²)	
Protein	53.43
Ramachandran plot [‡]	
Most favored (%)	99.4
Allowed (%)	0.6

Biotech). All sequence-verified mutant mTNKS1 and hTRF1 constructs were transformed into *E. coli* Transetta (DE3) competent cells (TransGen Biotech). The expression and purification of mutant mTNKS1 and TRF1 proteins were identical to these of the wild-type (WT) proteins, as described above.

2.5. Isothermal titration calorimetric (ITC) assays

To examine the interactions between WT or mutant mTNKS1 proteins and the hTRF1 protein, isothermal titration calorimetric (ITC) assays were performed at 289 K using a MicroCal iTC200 (Malvern Instruments Ltd). All of the WT and mutant mTNKS1 protein samples were diluted to 20 µM using a protein stock solution consisting of 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 5% glycerol, 5 mM DTT. TRF1 protein sample was diluted to 200 µM using the same buffer. TNKS1 protein samples were placed into the sample cell and TRF1 was titrated into the sample cell in one injection of a 0.4 µl aliquot and another 19 injections of 2 µl aliquots at 2 min intervals. Data analysis was performed using the *Origin* software, which gave the stoichiometry (*n*), dissociation constant (*K_d*) and enthalpy (ΔH) of the interaction.

2.6. In vitro GST pull-down assay

50 µg of purified GST or GST-tagged hTRF1 was incubated at 277 K for 1 h with 50 µl Glutathione Sepharose 4B resin (GE Healthcare). 200 µg of purified WT or mutant mTNKS1(ARC2-3) protein was then added to each 50 µl of resin bound to GST or GST-tagged hTRF1, respectively. The mixtures were incubated for a further 1 h at 277 K, washed three times in 1 ml 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 5% glycerol, 5 mM DTT and finally resuspended in 50 µl SDS-PAGE protein sample buffer. The samples were analyzed by SDS-PAGE and stained with Coomassie Blue.

3. Results and discussion

3.1. Overall structure of the TNKS1(ARC2-3)–TRF1(1–55) complex

To better understand the TNKS–TRF1 interaction, we chose to use hTRF1(1–55) for our crystallographic analysis, which contains all of the conserved regions in the acidic domain including the core TBM (Figs. 1*a* and 1*b*). We used mouse TNKS1(ARC2-3) in our crystallographic analysis, which is 99% identical to human TNKS1(ARC2-3) and is 100% identical within ARC2, which is responsible for the recognition of the TBM. For convenience, we use mTNKS1 residue numbering for structural description and discussions. The crystal structure was determined using molecular replacement and was refined at 2.2 Å resolution (Table 4).

In our structure, each mTNKS1 ARC consists of four ankyrin repeats, and each ankyrin repeat contains two helices with a β -hairpin in between (Fig. 1). TNKS1(ARC2-3) has a tendency to form a homodimer (Morrone *et al.*, 2012). In the crystal lattice there are two ARC2-3 molecules in each asymmetric unit, which form a swapped dimer, with the swap occurring behind the first helix of ARC3 (Fig. 1*c*). Two TRF1

molecules were found to bind to each of the ARC2 domains in the TNKS1(ARC2-3) homodimer.

Although the hTRF1(1–55) fragment was used in our crystallization and the fragment remained intact in our crystallographic lattice (Supplementary Fig. S1), only hTRF1 residues 11–22 were clearly visible in the electron-density map. This indicates that the TRF1 acidic domain is structurally flexible (intrinsically disordered) in the absence of the binding partner. To evaluate the potential contribution of the C-terminal region of TRF1(1–55), we performed ITC analysis of TNKS1(ARC2-3) with both TRF1(1–55) and TRF1(1–35). Both TRF1 fragments bind to TNKS1(ARC2-3) with comparable affinities ($K_d = 0.31 \mu\text{M}$ versus $0.46 \mu\text{M}$; Table 5), confirming that the N-terminal half of the TRF1 acidic domain is predominantly responsible for TNKS1 binding.

3.2. The TNKS1–TRF1 interface

Human TRF1 residues 11–22 (SPRGCADGRDAD) bind to a groove formed by the second helices and the β -hairpins of TNKS ARC2 in a manner similar to other TBM ligands. Within this TRF1 sequence, three side chains form hydrogen

Figure 2

(*a*) Details of the interface in stereoview. TNKS1(ARC2) is shown as green cartoons, with key TRF1-interacting residue side chains shown as sticks. The TRF1(11–22) fragment is shown as sticks, with residues labelled in black. (*b*) Surface illustration of the TNKS1–TRF1 interaction. One copy of ARC2 of mTNKS1 is shown as a white solid surface and residues involved in interaction are labeled, with C, O and N atoms in green, red and blue, respectively. A corresponding copy of hTRF1 is shown as a stick model in yellow and residues involved in the interaction with TRF1 are labeled in blue.

Table 5

Summary of isothermal titration calorimetric (ITC) analysis.

All experiments were performed at $T = 289$ K. n.s., no detectable signal. n.a., not available.

(a) Interactions between hTRF1(1–55) and mTNKS1 (WT or mutants).

	ΔH (kcal mol ⁻¹)	$T\Delta S$ (kcal mol ⁻¹)	n	K_d (μM)
mTNKS1	-29.4 ± 0.5	-1.15	0.94 ± 0.01	0.31 ± 0.05
mTNKS1 L396A	-25.1 ± 4.6	-1.05	1.00†	31.2 ± 12.0
mTNKS1 N401A	-17.5 ± 5.1	-0.62	0.99 ± 0.18	22.6 ± 8.6
mTNKS1 Y405A	n.s.	n.s.	n.a.	n.a.
mTNKS1 H407A	-24.7 ± 5.2	-1.09	1.00†	178.6 ± 46.6
mTNKS1 D425A	-14.3 ± 1.6	-0.45	1.00†	25.5 ± 6.5
mTNKS1 E434A	-25.3 ± 1.6	-1.10	1.00†	67.7 ± 6.4
mTNKS1 R440A	-24.4 ± 0.2	-0.90	0.85 ± 0.01	0.60 ± 0.05

(b) Interactions between hTRF1(1–35) and WT mTNKS1.

	ΔH (kcal mol ⁻¹)	$T\Delta S$ (kcal mol ⁻¹)	n	K_d (μM)
mTNKS1	-26.8 ± 0.2	-1.02	1.00 ± 0.01	0.47 ± 0.03

† n was fixed to 1.0 for these calculations, since the K_d values are higher than the protein concentrations.

bonds or charge–charge interactions with TNKS1 and thus contribute to the TRF1–TNKS1 binding specificity (Fig. 2). The side chain of Arg13 of TRF1 forms two charge–charge interactions with Glu434 and Asp425 of TNKS1, respectively; the Asp17 side chain forms a hydrogen bond to the side chain of Ser363 of TNKS1, whereas the side chain of Asp20 of TRF1 interacts with the side chain of Arg440 of TNKS1. There are also a number of interactions provided by TRF1 main-chain groups that stabilize the interaction and align TRF1 in the TNKS1 groove, including hydrogen bonds between the main-chain carbonyl groups of Ser11 and Ala16 of TRF1 and the side chains of Trp427 and Tyr405 of TNKS1, respectively.

Gly18, an absolutely conserved residue in the TBM, is recognized by a narrow ‘gate’ formed by the parallel Tyr372 and Tyr405 side chains, which are separated by only 7.4 Å. The C α atom of Gly18 is located in the mid-bottom of this gate, with its N-terminal side fixed by a hydrogen bond between the main-chain carbonyl group of Asp17 and the side chain of Asn401 of TNKS1, and the C-terminal side fixed by an interaction between the main-chain carbonyl group of Arg19 and the side chain of His407 of TNKS1 (Figs. 2*a* and 2*b*). Overall, our structure explains the sequence specificity of the TRF1 TBM well, and reveals additional interactions from the TRF1 flanking residues beyond the core TBM.

3.3. Mutagenetic analysis of TNKS1 mutants

The sequence selectivity of short TBM peptides and TNKS2(ARC4) has been studied with 3BP2–TBM and TNKS2 mutants using a fluorescence polarization assay (Guettler *et al.*, 2011). To evaluate how flanking sequences may affect the sensitivity of key TNKS1 and TRF1 interface residues to mutagenesis, we have analyzed the interactions between TRF1(1–55) and both WT TNKS1(ARC2–3) and mutants using both GST-pulldown and ITC assays.

In our ITC assay (Table 5 and Supplementary Fig. S2), mutations of either of the two TNKS1 residues that interact with Arg13 of TRF1, D425A and E434A, dramatically increased the K_d (*i.e.* reduced the binding affinity) from 0.31 μM (for WT) to 25 and 68 μM , respectively. Mutations of the two TNKS1 residues that interact with the TRF1 main-chain region between Arg13 and Gly18, L396A and N401A, also reduced the binding affinity to 31 and 23 μM , respectively. Importantly, individual mutations of two residues interacting with the main chain of Gly18 and its neighboring Arg19, Y405A and H407A, completely abolished the TRF1–TNKS1 interaction. In contrast, mutation of Arg440 of TNKS1, which interacts with Asp20 of TRF1, which is C-terminal to the core TBM, only reduced the binding affinity by twofold. Our GST-pulldown results are consistent with our ITC results in that the Y405A and H407A mutations have the most dramatic effect on the TRF1–TNKS1 interaction (Fig. 3*a* and Supplementary Fig. S3). These results indicate that the Gly-recognition ‘tyrosine gate’ works as a Gly-specificity filter and contributes to the TNKS1–TRF1 binding affinity.

3.4. Structural comparison with a TNKS2(ARC4)–3BP2–TRF1 chimera complex

TNKS1/2 ARC2 and ARC4 are highly homologous (Fig. 3*b*), and their three-dimensional structures are also very similar (Fig. 3*c*). Most of the TRF1-interacting TNKS1(ARC2) residues that we observed (with the exception of Arg440) are also found in TNKS2(ARC4). A crystal structure of TNKS2(ARC4) in complex with a 3BP2–TRF1 chimera (PDB entry 3tws), which has the TBM core from TRF1 (RGADG) but a 3BP2 flanking sequence (Fig. 3*c*), has been determined with the same crystal packing as the TNKS2(ARC4)–3BP2 peptide complex (Guettler *et al.*, 2011). Owing to a 3_{10} -helix in the N-terminus of the 3BP2 peptide, the position of the Arg corresponding to Arg13 in TRF1 slightly deviates from that of Arg13 in native TRF1 (Fig. 3*c*). However, with minor local structural adjustment in both TRF1 and TNKS1, two very similar charge–charge interactions are formed between the side chain of Arg13 of TRF1 and Glu434 and Asp425 of TNKS1, respectively. This suggests that the interface between the conserved Arg residue of TBM peptides (here Arg13) and the negatively charged TNKS1 pocket are quite promiscuous and can accommodate different conformations. In fact, in the TNKS1–Axin complex structure the second TNKS1 binding motif of Axin (residues 60–79) uses an arginine (hAxin Arg62) seven residues away from the normal position to make similar charge–charge interactions by a looping-back conformation (Morrone *et al.*, 2012). In contrast, despite having different C-terminal flanking sequences, both TRF1 and TNKS1 proteins have almost identical main-chain and side-chain conformations around the conserved Gly (Gly18), suggesting a fixed recognition interface in this region.

3.5. Comparison of TNKS(ARC2–3) structures and implication for the overall structure of TNKS

Both TNKS1 and TNKS2 contain five ARC domains/clusters. The four linkers between ARCs are also conserved

(Sbodio *et al.*, 2002). It remains unclear how these five ARCs relate to one another. The TNKS1(ARC2-3) fragment that we crystallized contains three of the four linkers. Our previous and current work demonstrates that each of the conserved inter-ARC linkers forms a single long helix and packs on the ends of neighboring ARCs (Fig. 1*b*; Morrone *et al.*, 2012). This may also be a conserved feature of the linker between ARC4 and ARC5, which is absent in our structure. To understand whether and how these linkers allow inter-ARC flexibility, we performed a superposition analysis between our current

TNKS1–TRF1 complex and the TNKS1–Axin complex (PDB entry 3utn). Despite being obtained from distinct crystallization conditions, the crystal packing of the TNKS1(ARC2-3)–TRF1(1–55) complex is very similar to that of the TNKS1(ARC2-3)–Axin(1–80) complex; both belong to space group *C*2, with similar unit-cell parameters (*a* = 131.78, *b* = 106.59, *c* = 73.48 Å, β = 105.76° for the TNKS1–Axin crystal and *a* = 135.00, *b* = 100.07, *c* = 75.95 Å, β = 107.54° for the TNKS1–TRF1 crystal). This is owing to the similar overall ARC2-3 structures

Figure 3
(*a*) Mutagenetic analysis of the TNKS1(ARC2-3) residues involved in TRF1 binding. GST-tagged TRF1(1–55) was used to pull down WT and mutant TNKS1(ARC2-3) proteins. (*b*) Sequence alignment of ARC2 and ARC4 of mTNKS1/2 and hTNKS1/2. Individual ankyrin repeats are boxed with blue dashed lines and labeled. Critical residues involved in interaction are highlighted with black stars; the two critical tyrosines that form the 'tyrosine gate' are highlighted with red stars. (*c*) Superposition of the TNKS1(ARC2-3)–TRF1(1–55) complex structure with the TNKS2(ARC4)–3BP2-TRF1 chimera complex. The TNKS1 ARC2 domain in the TNKS1–TRF1 complex is superimposed with the TNKS2 ARC4 domain in the TNKS2–3BP2-TRF1 complex. Corresponding sequences of hTRF1 and the 3BP2-TRF1 chimeric peptide are aligned. The C α position of Gly18 is indicated with a purple sphere.

Figure 4

Superposition of the four TNKS1(ARC2-3)-containing structures indicates inter-ARC structural flexibility. Cartoon models are shown for all four TNKS1(ARC2-3) structures: subunit *A* of TNKS1-TRF1 in green, subunit *B* of TNKS1-TRF1 in cyan, subunit *A* of TNKS1-Axin in yellow and subunit *B* of TNKS1-Axin in orange. All four models are superposed on the ankyrin repeats of ARC2 (residues 352–470).

in both of the crystal structures. Interestingly, when the TNKS1(ARC2) subdomains in these structures were superposed, the positions of TNKS1(ARC3) are quite different (Fig. 4). This indicates that although neighboring TNKS ARCs are tethered together by interacting with the shared inter-ARC helix, the relative orientations of the neighboring ARCs are not rigidly fixed. Instead, there is significant inter-ARC flexibility, at least for the ARC2-ARC3 interface.

4. Conclusions

The TRF1-TNKS1 interaction plays an important role in telomere regulation. In this work, we provide a structural basis for understanding how the acidic domain of TRF1 interacts with the ankyrin-repeat domain of TNKS1. We show that while the C-terminal half of the acidic domain of TRF1 is structurally flexible and non-essential to TNKS binding, the residues flanking the RxxP/G/AxG motif, including Asp20, contribute to the interaction. While the interface between Arg13 of TRF1 and the acidic pocket of TNKS1/2 may be quite flexible, the Gly-recognition 'tyrosine gate' is crucial as a specificity filter and as a binding-affinity hotspot. Our work also provides new insights into the structural dynamics of the ankyrin-repeat domain of TNKS1/2, suggesting that the inter-ARC linkers provide structural flexibility among the five rigid ARC domains. These insights help us to better understand the ligand recognition and regulation of TNKS1 and TNKS2, which play key roles in many biological systems.

Acknowledgements

We gratefully acknowledge the assistance of the staff of beamline 18U at Shanghai Synchrotron with the X-ray diffraction data collection. We also thank the staff of the Core Facility at Nankai University. This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant Nos. XDB08020200 and XDB08010303).

References

- Cruickshank, D. W. J. (1999). *Acta Cryst.* **D55**, 583–601.
- De Boeck, G., Forsyth, R. G., Praet, M. & Hogendoorn, P. C. (2009). *J. Pathol.* **217**, 327–344.
- Diotti, R. & Loayza, D. (2011). *Nucleus*, **2**, 119–135.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Garcia-Aranda, C., de Juan, C., Diaz-Lopez, A., Sanchez-Pernaute, A., Torres, A. J., Diaz-Rubio, E., Balibrea, J. L., Benito, M. & Iniesta, P. (2006). *Cancer*, **106**, 541–551.
- Guettler, S., LaRose, J., Petsalaki, E., Gish, G., Scotter, A., Pawson, T., Rottapel, R. & Sicheri, F. (2011). *Cell*, **147**, 1340–1354.
- Haikarainen, T., Krauss, S. & Lehtio, L. (2014). *Curr. Pharm. Des.* **20**, 6472–6488.
- Hsiao, S. J. & Smith, S. (2008). *Biochimie*, **90**, 83–92.
- Lange, T. de (2005). *Genes Dev.* **19**, 2100–2110.
- Longhese, M. P. (2012). *Front. Biosci.* **17**, 1715–1728.
- Martínez, P. & Blasco, M. A. (2011). *Nature Rev. Cancer*, **11**, 161–176.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Miyachi, K., Fujita, M., Tanaka, N., Sasaki, K. & Sunagawa, M. (2002). *J. Exp. Clin. Cancer Res.* **21**, 269–275.
- Morrone, S., Cheng, Z., Moon, R. T., Cong, F. & Xu, W. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 1500–1505.
- Nakanishi, K., Kawai, T., Kumaki, F., Hiroi, S., Mukai, M., Ikeda, E., Koering, C. E. & Gilson, E. (2003). *Clin. Cancer Res.* **9**, 1105–1111.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pal, D., Sharma, U., Singh, S. K., Kakkar, N. & Prasad, R. (2015). *PLoS One*, **10**, e0115651.
- Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95–99.
- Sbodio, J. I. & Chi, N.-W. (2002). *J. Biol. Chem.* **277**, 31887–31892.
- Sbodio, J. I., Lodish, H. F. & Chi, N.-W. (2002). *Biochem. J.* **361**, 451–459.
- Seimiya, H., Muramatsu, Y., Smith, S. & Tsuruo, T. (2004). *Mol. Cell Biol.* **24**, 1944–1955.
- Seimiya, H. & Smith, S. (2002). *J. Biol. Chem.* **277**, 14116–14126.
- Valls-Bautista, C., Piñol-Felis, C., Reñé-Espinet, J. M., Buenestado-García, J. & Viñas-Salas, J. (2012). *Rev. Esp. Enferm. Dig.* **104**, 530–536.
- Walker, J. R. & Zhu, X.-D. (2012). *Mech. Ageing Dev.* **133**, 421–434.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Winn, M. D., Murshudov, G. N. & Papiz, M. Z. (2003). *Methods Enzymol.* **374**, 300–321.