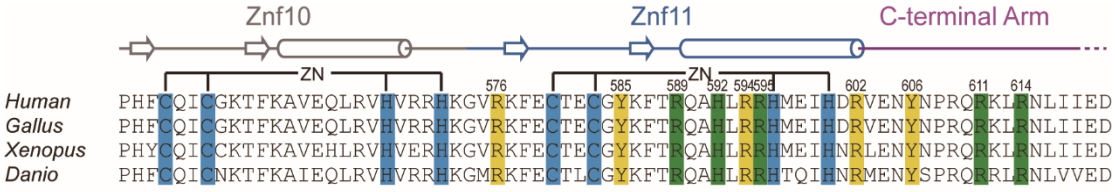
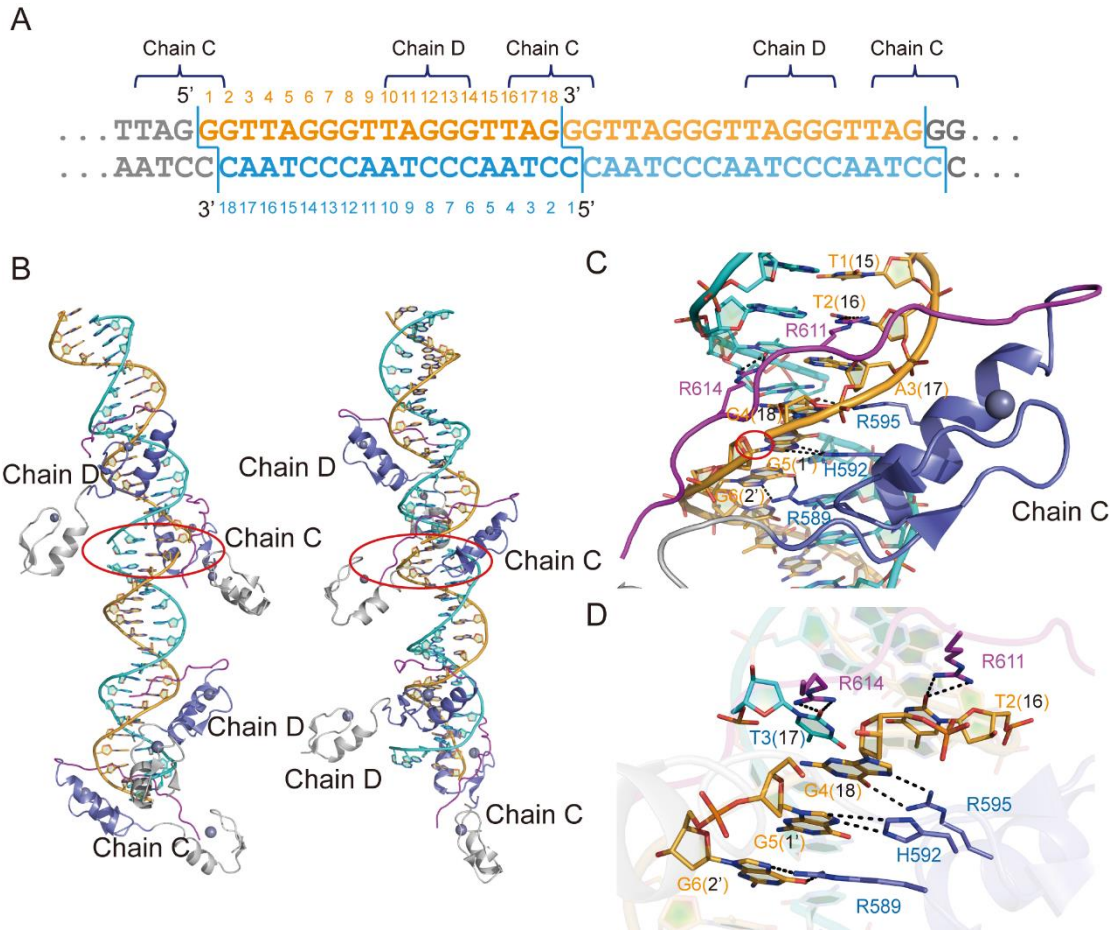
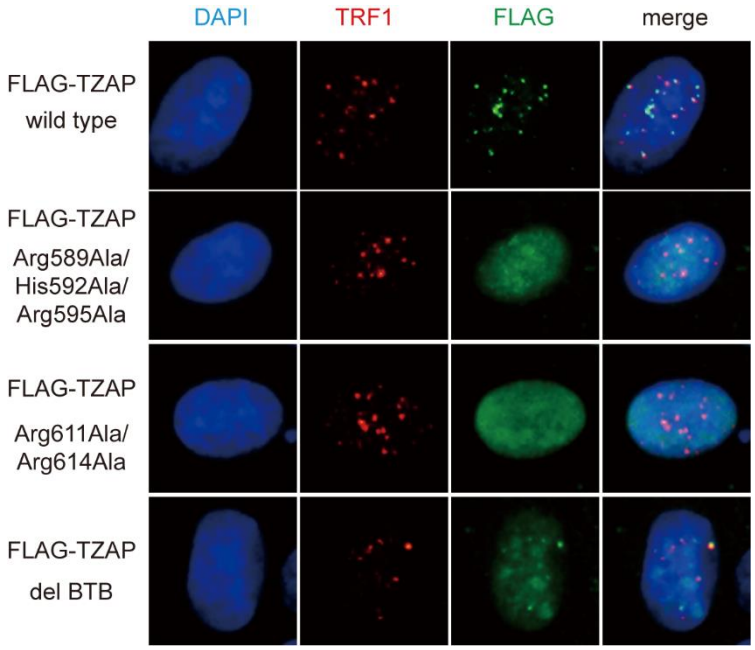


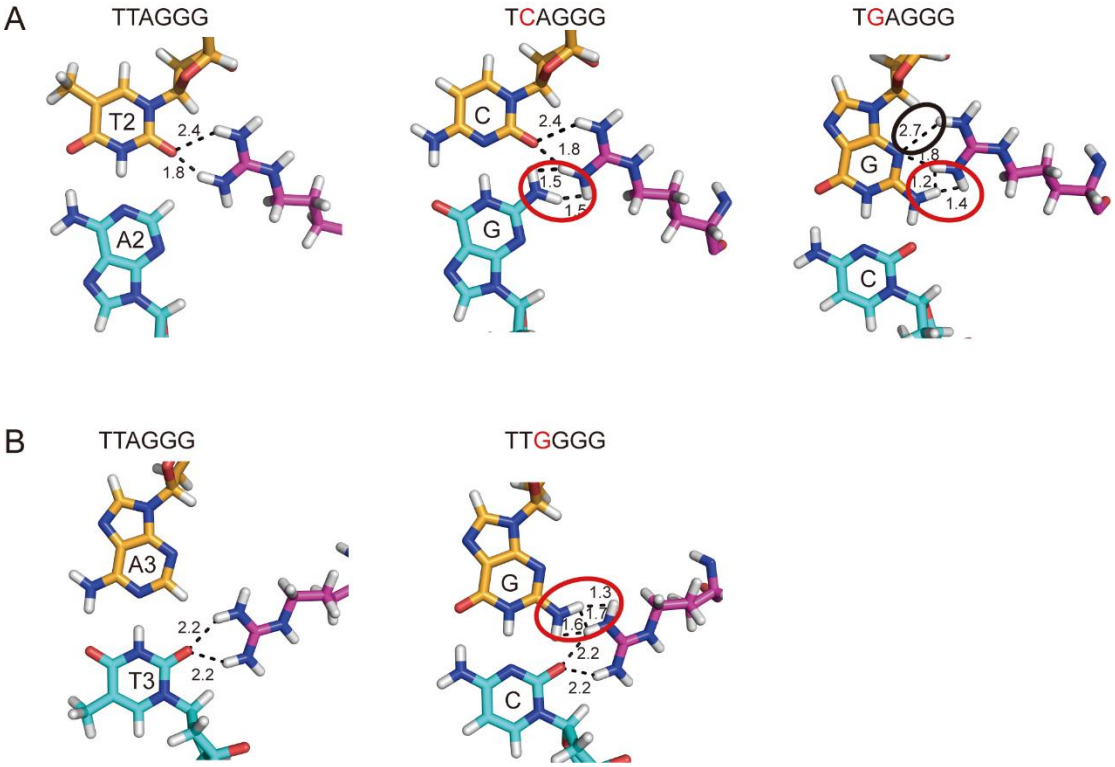
Supplementary information, Figure S1 Sequence alignment of the Znf10-11-C region of human TZAP and its homologues in *Gallus*, *Xenopus laevis*, and *Danio rerio*. Conserved residues are highlighted: cyan (for zinc-coordinating); yellow (for phosphate backbone interactions); green (for base-specific interactions).



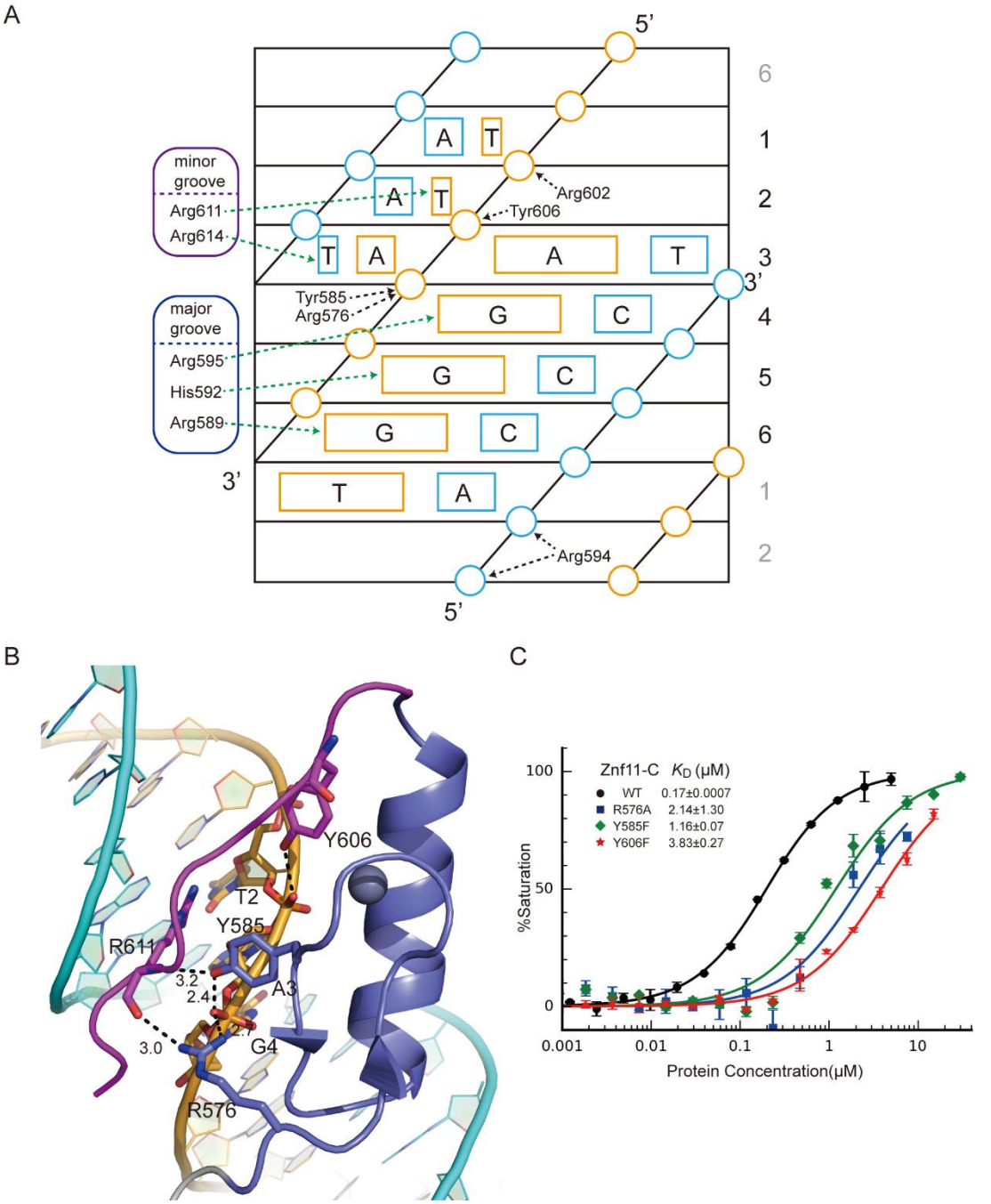
Supplementary information, Figure S2 (A) The DNA sequence used for crystallization is shown. The overhanging G and C of two neighboring DNA molecules pair to form pseudo-continuous TTAGGG duplex. The top and bottom numbers indicate the nucleotide positions of the G-strand and C-strand in the pdb file, respectively. Chain D binds TTAGGG site in the DNA duplex, while Chain C binds the end to end stacking TTAGGG site. (B) Cartoon representation of how Chain C binds the end to end stacking TTAGGG site of two neighboring DNA molecules. (C-D) Detailed intermolecular interactions between the DNA bases and Chain C. The hydrogen bonds are depicted as black dashed lines. Chain C and Chain D bind DNA in a similar way.



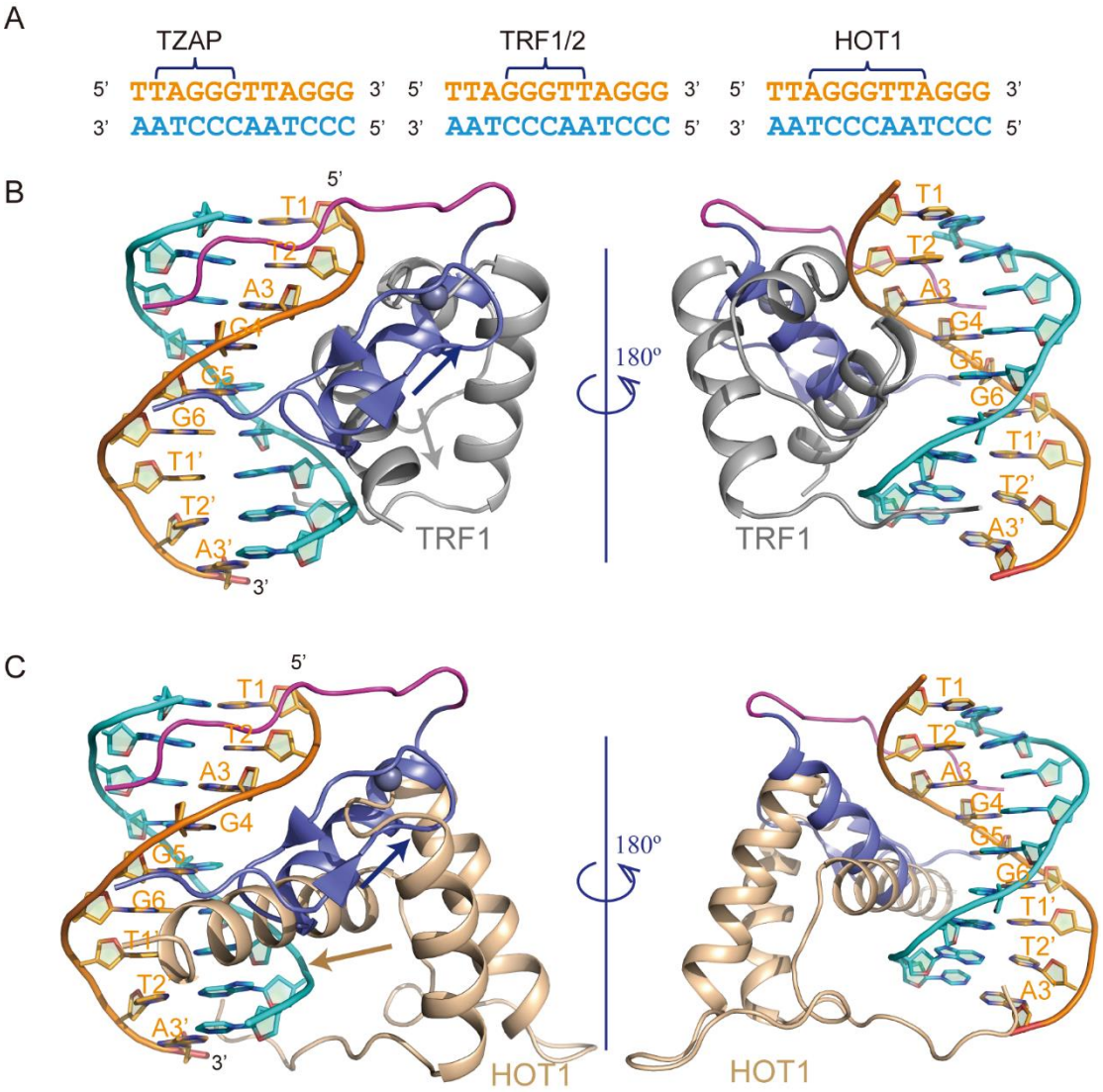
Supplementary information, Figure S3 Immunofluorescence (IF) stainings for exogenous FLAG-TZAP wild type, triple-point mutant for Znf11 (Arg589Ala/His592Ala/Arg595Ala), double-point mutant for C-terminal arm (Arg611Ala/Arg614Ala), and BTB deletion (del BTB) mutant in U2OS cells.



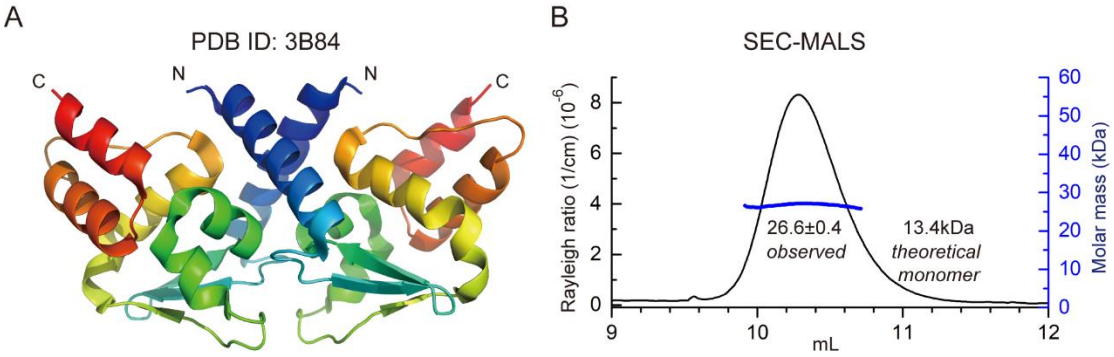
Supplementary information, Figure S4 (A) Models for the interactions between subtelomeric DNA sequences TCAGGG and TGAGGG with Arg611. The steric clash is marked with red ovals, while the loss of hydrogen bonds is highlighted with black ovals. The models were generated from Chain C of the pdb file. The hydrogen atoms were added by Pymol. (B) Models for the interactions between subtelomeric DNA sequence TTGGGG with Arg614. The steric clash is marked with red ovals.



Supplementary information, Figure S5 (A) Summary of direct protein-DNA contacts in the TZAP-DNA complex. The DNA is represented as a cylindrical projection. Phosphates are represented with circles. Green dash lines with arrowheads indicate base-specific recognition by forming hydrogen bonds. Black dashed lines with arrowheads indicate phosphate contacts. (B) Arg576 and Y585 interact with the phosphate backbone and the main chain of Arg611. Y606 within the C-terminal arm interacts with the phosphate backbone. (C) The effects of R576A, Y585F and Y606F mutants of Znf11-C on TTAGGG binding.



Supplementary information, Figure S6 (A) TZAP recognizes TAGGG directly; TRF1/2 recognize GGGTT; HOT1 directly recognizes AGGGTTA. (B) Alignment and superimposition of the DNA moieties (TTAGGGTTA) of the TZAP and TRF1 crystal structures. Only the DNA moieties of TZAP complex structure are shown. (C) Alignment and superimposition of the DNA moieties (TTAGGGTTA) of the TZAP and HOT1 crystal structures. Only the DNA moieties of TZAP complex structure are shown.



Supplementary information, Figure S7 The BTB domain of human TZAP acts as a dimer in crystal structure and in solution. (A) The BTB dimer is show in cartoon representation. (B) MALS (black line, left axis) and calculated mass (blue line, right axis) determined from SEC peak.

Supplementary information, Materials and Methods

Protein expression and purification

ZnF9-11 (residues 516-605), ZnF9-11-C (residues 516-620) and ZnF11-C (residues 573-620) were amplified by PCR from a human brain cDNA library and cloned into a modified pGEX-4T-1 vector with a TEV cleavage site. All the mutants were generated using MutanBEST kit (Takara) and confirmed by DNA sequencing. All the proteins were expressed in *Escherichia coli* BL21 (Gold) cells. Cells were grown in LB medium supplemented with 100 μM $ZnCl_2$ at 37℃ until OD_{600} reached 0.8. Then the proteins was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 5 hours. The proteins were purified by glutathione sepharose (GE healthcare) in a buffer containing 20 mM Tris-HCl (pH 7.5), 1 M NaCl, followed by TEV cleavage and size-exclusion chromatography on a Hiload 16/60 Superdex 75 column (GE healthcare). The purified proteins were dialyzed with Buffer A (20 mM Tris-HCl (pH7.5), 150 mM NaCl), and concentrated for subsequent analysis.

The BTB domain of TZAP (residues 1-119) was cloned into a modified PET28a vector with a sumo tag. The protein was expressed in *Escherichia coli* BL21 (Gold) cells and induced with 0.1 mM for 24 hours at 16℃. The protein was purified by Ni-chelating column (Qiagen), followed by Ulp1 cleavage, and size-exclusion chromatography on a Hiload 16/60 Superdex 75 column (GE healthcare).

Crystallography

Purified Znf9-11-C was mixed with the 18bp double-stranded oligo (Table S2) at a 2:1 ratio in 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, followed by dialyzing against Buffer A to form the protein-DNA complex. The complex was further concentrated up to about 1 mM. The crystals were grown using the hanging drop vapor diffusion method at 20℃ by mixing equal volumes of the protein-DNA complex and reservoir buffer (0.2 M Imidazole-Malate pH 8.5, 20% PEG 4K). The reservoir solution with 25% glycerol was used as a cryoprotectant.

The X-Ray diffraction data set was collected on beam line 19U1 at Shanghai Synchrotron Radiation Facility (SSRF) at wavelength 0.978 Å. The data set was indexed, integrated and scaled with HKL2000[1]. The initial crystallographic phases were calculated using PHENIX.autoSolve by the combination of single-wavelength anomalous dispersion of zinc signals and molecular replacement which was carried out by Phaser employing the DNA molecule of TRF1-telomeirc DNA complex as the search model (PDB code: 1W0T)[2, 3]. An initial model was automatically build by Buccaneer, and then was further build and refined using Coot and Phenix.refine[2, 4, 5].

Fluorescence-based DNA-binding assays

This protein stock was diluted in 1/2 series in Buffer A to the lowest desired

concentration. The FAM-labeled Double-stranded DNA probe stock was diluted to 100 nM in Buffer A. 100 μL of this diluted DNA probe was mixed with 100 μL of 1/2 series protein stock and incubated for 30 min. The fluorescence polarization was measured using a SpectraMax M5 plate reader (Molecular Devices) at 20℃. Curves were fit individually using the equation [mP] = [maximum mP] × [C]/(K_D + [C]) +[baseline mP], where mP is millipolarization, and [C] is protein concentration. K_D values and the fitting figures were derived (from two experimental replicates) by fitting the experimental data to the equation employing a fitting script written in python.

SEC-MALS

Analytical SEC was performed using an Akta Pure L system (GE healthcare) with a superdex75 10/300 column (GE healthcare). The system was coupled on-line to an 8-angel MALS detector (DAWN HELEOS II, Wyatt Technology) and a differential refractometer (Optilab T-rEX, Wyatt Technology). Molar masses were determined using ASTRA 7.0.1 software.

Immunofluorescence

Full length TZAP amplified by PCR from a human brain cDNA library and cloned into a modified pRK5 vector with a FLAG tag. Mutants were generated using MutanBEST kit (Takara) and confirmed by DNA sequencing. For plasmid transfections in U2OS cells, plasmid was transfected with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions.

For Immunofluorescence assay, cells cultured on coverslips were washed with cold PBS, fixed with 4% polyformaldehyde, permeabilized with PBS + 0.1% Triton X-100, and blocked with 5% BSA. Then primary antibodies were diluted and incubated for 1 h at RT. The following antibodies were used: Rabbit anti TRF1 (11899-1-AP, Proteintech), Mouse anti FLAG (F1804, Sigma). After three washes, samples were incubated with TRITC-conjugated goat anti-rabbit-IgG and FITC-conjugated goat anti-mouse-IgG (both from Jackson Immunoresearch) as secondary antibodies for 30 min at RT. After three additional washes, samples were treated with DAPI reagent and the images were acquired with an Olympus DP71X microscope.

References

1 Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* 1997; **276**:307-326.

2 Adams PD, Afonine PV, Bunkoczi G *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography* 2010; **66**:213-221.

3 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic

software. *Journal of applied crystallography* 2007; **40**:658-674.
4 Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta crystallographica Section D, Biological crystallography* 2010; **66**:486-501.
5 Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D* 2006; **62**:1002-1011.