The N-terminal domains of TRF1 and TRF2 regulate their ability to condense telomeric DNA

Anaïs Poulet[1,2], Sabrina Pisano[1,2,3], Cendrine Faivre-Moskalenko[4], Bei Pei[3], Yannick Tauran[7], Zofia Haftek-Terreau[1], Frédéric Brunet[6], Yann-Vaï Le Bihan[5], Marie-Hélène Ledu[5], Fabien Montel[1], Nicolas Hugo[1], Simon Amiard[1], Françoise Argoul[1], Annie Chaboud[7], Eric Gilson[2,3,8] and Marie-Josèphe Giraud-Panis[1,2,3,*]

[1]Université de Lyon, Laboratoire Joliot-Curie, CNRS USR3010, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, F-69364 Lyon, [2]Université de Lyon, Laboratoire de Biologie Moléculaire de la cellule, CNRS UMR5239, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, F-69364 Lyon, [3]University of Nice, Laboratory of Biology and Pathology of Genomes, UMR 6267 CNRS U998 INSERM, 28 avenue Valombrose, Faculté de Médecine, F-06107 Nice, [4]Laboratoire de Physique, CNRS UMR5672, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, F-69364 Lyon, [5]CEA/DSV/IBiTec-S/SB2SM, Laboratoire de Biologie Structurale et Radiobiologie, bat 144, CEA Saclay, F-91191 Gif-sur-Yvette, [6]Institut de Génomique Fonctionnelle de Lyon, CNRS UMR5242 Ecole Normale Supérieure de Lyon, 46, allée d'Italie, F-69364 Lyon, [7]IFR 128, BioSciences Gerland – Lyon Sud, Plateau de Production et Analyse des Protéines, IBCP, CNRS UMR 5086, 21, avenue Tony Garnier, F-69367 Lyon and [8]Department of Medical Genetics, CHU of Nice, Nice, France

ABSTRACT

TRF1 and TRF2 are key proteins in human telomeres, which, despite their similarities, have different behaviors upon DNA binding. Previous work has shown that unlike TRF1, TRF2 condenses telomeric, thus creating consequential negative torsion on the adjacent DNA, a property that is thought to lead to the stimulation of single-strand invasion and was proposed to favor telomeric DNA looping. In this report, we show that these activities, originating from the central TRFH domain of TRF2, are also displayed by the TRFH domain of TRF1 but are repressed in the full-length protein by the presence of an acidic domain at the N-terminus. Strikingly, a similar repression is observed on TRF2 through the binding of a TERRA-like RNA molecule to the N-terminus of TRF2. Phylogenetic and biochemical studies suggest that the N-terminal domains of TRF proteins originate from a gradual extension of the coding sequences of a duplicated ancestral gene with a consequential progressive alteration of the biochemical properties of these proteins. Overall, these data suggest that the N-termini of TRF1 and TRF2 have evolved to finely regulate their ability to condense DNA.

INTRODUCTION

Telomeres are specialized structures protecting the natural termini of linear chromosomes from degradation and illicit repair (1). They are assembled through associations between telomeric DNA, a TTAGGG repeat containing sequence that ends with a single stranded 3′ overhang and telomere-specific proteins. The transcription of telomeric DNA produces a UUAGGG repeat containing RNA, TERRA, which is anticipated to play fundamental roles in telomere biology (2).

Several protein complexes associate with telomeric DNA. Among these, the shelterin complex in mammals involves six proteins (TRF1, TRF2, RAP1, TIN2, TPP1 and POT1) (3,4). Binding of this complex to DNA is mediated by the two double-stranded DNA (dsDNA)-binding proteins TRF1 (5,6) and TRF2 (7,8) and the G-tail binding protein POT1 (9). The other three

proteins do not interact directly with DNA (10–14). Other complexes containing fewer members of the shelterin complex have also been described (15). Telomeres have the capacity to fold into t-loops, lasso-like structures that have been observed on purified telomeres of diverse origins (16–20). *In vitro*, formation of t-loops depends on TRF2 and has been proposed to involve the *cis*-oriented invasion of the telomeric dsDNA by the G tail (17,21). One interesting concept that emerged from recent biochemical studies on TRF2 is that its biological functions seem closely driven by its intrinsic properties. Its capacity to bind telomeric DNA ends (22,23) has been implicated in the inhibition of the non-homologous end-joining pathway (24). Its N-terminal domain binds the center of Holliday junctions, protecting them from resolvase cleavage (25), a property possibly explaining TRF2's ability to protect the t-loops from resolution. TRF2 has also been shown to stimulate the invasion of a telomeric single-strand inside a homologous duplex in free DNA (26) and in the context of the nucleosomal fiber (27), a process thought to participate in the formation of t-loops and chromatin loops. This stimulation was proposed to result from a change in topology generated by the condensation of the DNA. This latter property was shown to primarily involve the C-terminal Telobox/Myb-like and the homodimerization TRFH domains of the protein (26). Despite the close resemblance between these domains in TRF1 and TRF2, TRF1 exhibits a very different behavior towards DNA. In this study, we investigated this paradox and show that TRF1 and TRF2 are not as different as hitherto assumed. Through studies of different mutants and chimeras of these proteins, we uncover the important regulatory role played by the N-termini in the biochemical properties of these proteins.

MATERIALS AND METHODS

Plasmids and oligonucleotides

pTelo2 and PLTelo are both pUC19-based plasmid containing 650 bp of human telomeric repeats.

Proteins

Cloning of protein genes are described in Supplementary Data. TRF proteins and mutants were: TRF2 (3-500), TRF2$^{\Delta B}$ (47-500), TRF2$^{\Delta D}$ (3-44, 244-500),TRF2$^{\Delta B \Delta D}$ (244-500), TRF2$^{\Delta L}$ (3-248, 445-500), TRF2$^{\Delta D}$TRF1^D (TRF2 3-44 TRF1 65-264 TRF2 244-500), TRF1 (2-439), TRF2$^{hA \Delta B}$ (hTRF1 2-67 TRF2 47-500), TRF2$^{mA \Delta B}$ (mTRF12-63 TRF2 47-500), TRF2$^{cA \Delta B}$ (cTRF1 2-27 TRF2 47-500) and TRF1$^{\Delta A}$ (65-248). All proteins were fused to a 6-histidines tag and produced as published, a gel filtration chromatography step was added when necessary (26). A Coomassie blue-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) gel of the wild type and mutant proteins is shown in Supplementary Figure S1A. Un-tagged TRF2 protein was also used in a control topology experiment (see below). In that case the tagged protein was cleaved using Tabacco Etch virus Protease I, the Histidine tag

removed by Ni-column chromatography followed by two chromatographic steps (heparin and gel filtration).

Strand Invasion assay, electrophoretic mobility shift assay and topology assays

Strand invasion assay, electrophoretic mobility shift assay (EMSA) and topology assays were performed as published (26). A control topology using un-tagged TRF2 protein was performed to control the absence of effect of the Histidine Tag (Supplementary Figure S1B).

Surface plasmon resonance (SPR)-based invasion assay

The experiments were performed on a Biacore T100 (GE Healthcare) using C1 sensorchips (GE Healthcare). Binding of the T15G probe and the R100 control (a non-telomeric random 100 bases oligonucleotide, Supplementary Data) was performed via a streptavidin–biotin interaction in conditions recommended. A signal of 150 RU was achieved for both the captured telomeric (T15G) and control (R100) oligonucleotides. pTelo2 (50 nM) and proteins (200 nM) were pre-incubated in HBS-ET buffer (10 mM HEPES pH 7.4, 300 mM NaCl, 3 mM EDTA, 0.05% Tween 20) during 15 min at room temperature. Samples were injected at 10 µl/min and the washing step performed using HBS-ET buffer. Surfaces were regenerated by sequential injections of water, 1 M NaCl, and 0.1% SDS during 30 s. Each sensorgram was corrected by subtraction of the signal obtained from the control flow cell (R100).

Atomic force microscopy imaging

DNA was incubated with proteins for 20 min at 25°C in 5 mM HEPES pH 7.4, 150 mM KCl. The protein/DNA molar ratios used: 10/10 for TRF1, TRF2, TRF2$^{hA \Delta B}$ and TRF2$^{\Delta L}$; 3/6 for TRF1$^{\Delta A}$ and TRF2$^{\Delta B}$; 6/9 for TRF2$^{mA \Delta B}$; 10/9 for TRF2$^{cA \Delta B}$. Samples were crosslinked with glutaraldehyde (0.1% final) for 30 min on ice and applied on freshly cleaved mica surfaces treated with 10 mM MgCl$_2$. After 2 min, mica was washed with deionized water and dried. Imaging was performed on a Nanoscope IIIa equipped with E-scanner (Digital Instruments Inc., Santa Barbara, CA, USA), in air under Tapping Mode using silicon tips. Images were recorded at 1.5–2.0 Hz over scan areas 1 µm wide (512 × 512 pixels). Raw scanning force microscopy (SFM) images were flattened using the manufacturer's software and converted into TIF files. Contour lengths (CLs) were measured by the read-through length method using SigmaScan Pro software (SPSS Inc., Chicago, IL, USA). Volumes calculated as half-ellipsoids as published (26). Between 150 and 300 objects were scored for each condition. Detailed information on the construction of the 2D probability density maps are given in Supplementary Data.

Phylogenetic studies

Pblast. The Telobox sequence from the human TRF1 and TRF2 was blasted against the NCBI protein database. The resulting alignment was used to generate the

PhyML tree. We downloaded the protein families defined in the Ensembl database version 56 (as of September 2009) (www.ensembl.org/) ENSFM00250000004074 and ENSFM00250000007334 and also added sequences from the NCBI database (www.ncbi.nlm.nih.gov/). All alignments were manually verified using SeaView to exclude redundant and improperly annotated sequences (28). We downloaded cDNAs from NCBI and also checked ESTs data. The best alignment carried out using ClustalW and Muscle (29,30) was considered and manually refined. All nucleotidic sequences retrieved were translated before our phylogenetic analyses.

RESULTS

The acidic domain of TRF1 inhibits its ability to condense DNA

The TRFH domain of TRF2 (here called D domain, Figure 1A) plays a critical role in the ability of TRF2 to condense DNA and to stimulate telomeric invasion (26). In view of the structural homology between these D domains in TRF1 and TRF2, it was therefore surprising to observe that TRF1 inefficiently condensed DNA. Two hypotheses could explain this difference: (i) the TRF1 and TRF2 D domains could be functionally different or (ii) the D domain of TRF1 could also be capable of DNA condensation but this property is inhibited in the full-length protein. To investigate this question, we have constructed several mutants of TRF2 (Figure 1A, Supplementary Figure S1A) and analyzed their ability to condense DNA using a topology assay and atomic force microscopy (AFM) (26). The former assay is based on the analysis of the topology of DNA by gel electrophoresis after incubation with the protein in the presence of wheat germ topoisomerase I. This enzyme can remove the supercoils, located in the unbound part of a closed DNA molecule, generated by the binding on this DNA of a topologically active protein. Modification of DNA topology caused by this type of proteins can be visualized through the appearance of topoisomers on a gel (Figure 1B). This experiment allows the measurement of the average number of supercoils created (Figure 1B, right) and these turns can be characterized as being positive or negative by using chloroquine in the experiment. Indeed, this drug increases the rate of migration of positively supercoiled DNA and conversely decreases this rate for negatively supercoiled DNA compared to controls (Supplementary Figure S1C). AFM allows the direct visualization of the DNA–protein complexes and the measure of both the contour length (CL) of the DNA and the volume of the complexes (Figure 1C). From these numbers, color-coded 2D-probability density maps were drawn showing the probability ($p(x,y)$) of a given complex to have a volume x and a y DNA CL. As shown in Figure 1C, most of the TRF1 complexes exhibit a small volume and a long DNA CL, implying a lack of DNA condensation. This is confirmed by the topology assay showing that TRF1 inefficiently creates supercoils in DNA (Figure 1B). In contrast, TRF2 causes a significant decrease in DNA CL (Figure 1C), thus showing DNA condensation. TRF2 creates positive supercoils (Figure 1B and Supplementary Figure S1) that can also be observed by AFM imaging of the DNA resulting from the topology assay (Supplementary Figure S2). Comparison of the different mutants schematically shown in Figure 1A reveals several key points: (i) deleting the acidic domain of TRF1 (TRF1$^{\Delta A}$) greatly increases TRF1's ability to condense DNA, thus transforming it into a TRF2$^{\Delta B}$-like protein; (ii) adding the A domain of TRF1 on TRF2$^{\Delta B}$ (TRF2$^{hA\Delta B}$) does the reverse, transforming TRF2 in a TRF1-like protein; (iii) the linker or hinge domain that separates the D domain and the Myb-like domains has little role in this function (TRF2$^{\Delta L}$); (iv) the B domain of TRF2 seems to stimulate TRF2 ability to modify DNA topology (TRF2$^{\Delta B}$); (v) the TRFH domain of TRF2 is required (TRF2$^{\Delta D}$, TRF2$^{\Delta B\Delta D}$); and (vi) the TRFH domains of TRF1 and TRF2 seem interchangeable, since the chimeric protein TRF2$^{\Delta D}$TRF1^{D} is very efficient in modifying DNA topology.

In summary, both TRF1 and TRF2 have the ability to condense DNA, but the N-terminal acidic domain of TRF1 prevents this condensation.

TRF2 dramatically increases the rate of telomeric strand invasion

As a consequence of this change of DNA topology, TRF2 was shown to increase the invasion of telomeric double strand by an homologous single-stranded probe (26). In order to verify that the acidic domain could also be responsible for the lower efficiency of TRF1 in telomeric invasion, we analyzed the invasion activity using the pTelo2 telomeric plasmid and a single-stranded probe containing 15 TTAGGG telomeric repeats (T15G) with two different assays: an invasion assay based on gel electrophoresis thus measuring the association at steady state; an assay using SPR technology to monitor the same association in real-time. In the latter experiments, we measured the changes in refractive index (in arbitrary units RU) due to the binding of the pTelo2 plasmid on the T15G single-stranded probe immobilized on the chip (31). An SPR sensorgram presents two phases (Supplementary Figure S3A): injection, when the analyte is continuously injected and washing to follow dissociation. First, the samples were injected until the signal reached a plateau. With pTelo2 alone the response was very weak (RU max around 10) and reached a plateau in about 11 h. However, in the presence of TRF2, the signal reached the plateau in about 15 min for a response 60 times higher. This suggests that TRF2 dramatically increases the association rate. Then, we performed comparative analysis between TRF1 and TRF2 on different substrates. Samples were injected for only 3 min and the measure performed 1 min after the injection stop (measure point in Supplementary Figure S3B). Results are presented as histograms of the corresponding measure points. The response obtained in the presence of TRF2 was 44 times higher than the one obtained with pTelo2 alone and therefore corresponding to a stimulation of invasion of 44-fold. This increase could not be due to
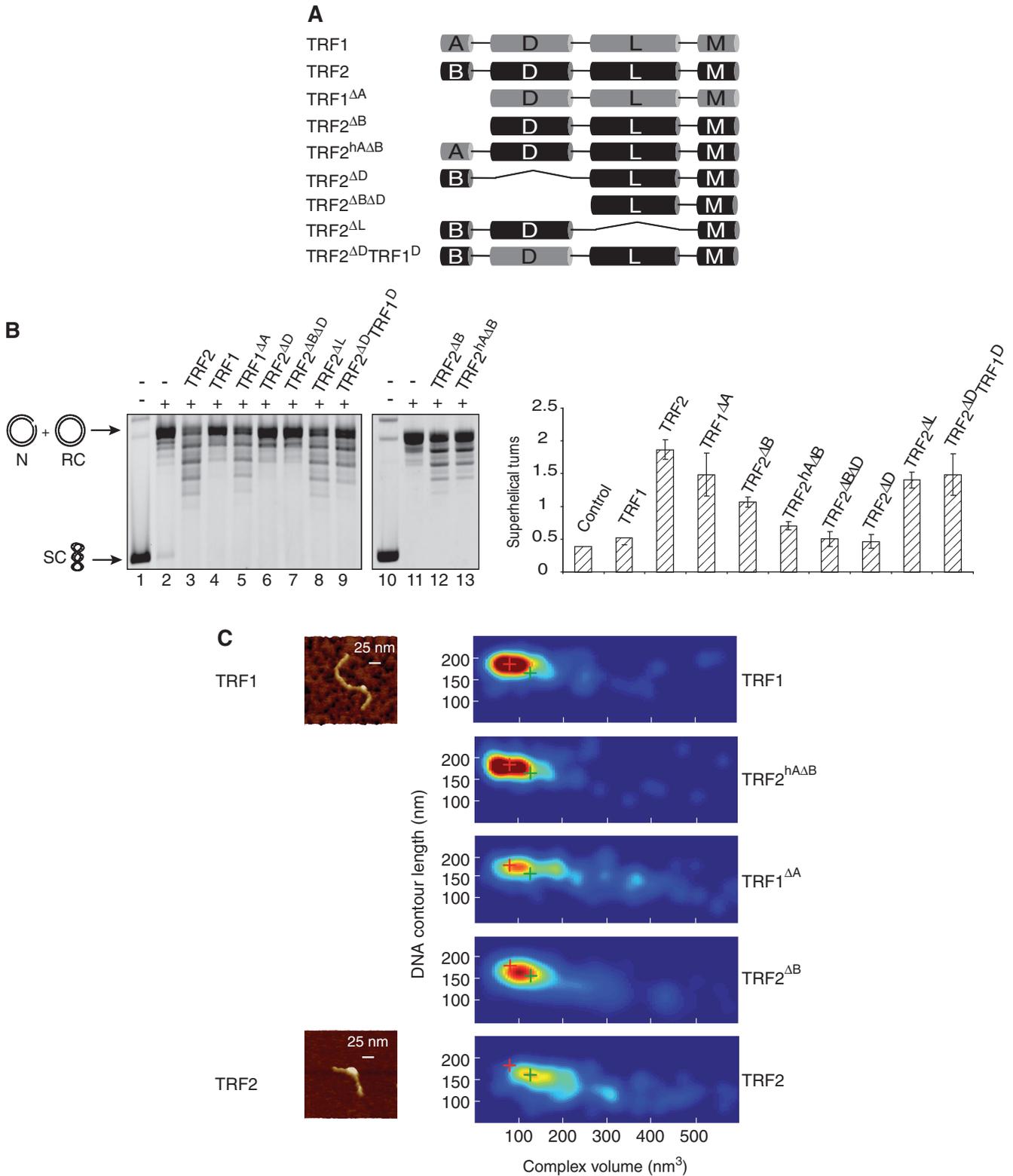
Figure 1. The acidic domain of TRF1 prevents DNA condensation and also inhibits the modification of DNA topology. (**A**) Schematic representation of the TRF1 and TRF2 wild type and mutant proteins. (**B**) Topology assay performed with the telomeric plasmid pLTelo and 1 μM of each protein using wheat germ topoisomerase I (WG Topo I). Samples were analyzed by TBE agarose gel electrophoresis. SC stands for supercoiled plasmid, RC for relaxed circular, N for nicked. The plot on the right panel shows the average number of supercoils created by the proteins calculated from gels such as the one shown in (B). Error bars show standard deviation. (**C**) 2D-probability density maps of volumes and DNA contour length (CL) measured by AFM for the complexes formed with TRF1, TRF2$^{hA\Delta B}$, TRF1$^{\Delta A}$, TRF2$^{\Delta B}$ and TRF2. The red and green crosses indicate the position in the graph of the majority of TRF1 and TRF2 complexes, respectively.

the binding of TRF2 on T15G since no change in SPR response could be seen when using pUC19 (Figure 2B). The dissociation phase was slow, suggesting that the major effect of the protein was on the association rate. TRF1 had a much lower effect than TRF2 (<10%, Supplementary Figure S3C and G). Gel-based assays performed with the same proteins (Supplementary Figure S3D, E and G) give the same trend. As expected, the effect measured for TRF2 was higher in SPR since, in this experiment, even unstable complexes would be scored. As seen in the invasion assay (26), the TRF2 effect was more potent on a closed supercoiled plasmid than on a linear molecule (Supplementary Figure S3F), showing the importance of topological constraints. Overall, data obtained with the two methods are in good agreement which suggests that both experiments monitor the same phenomenon. Furthermore, we can say that the main effect of TRF2 on telomeric invasion is to greatly increase the association rate of the double-stranded target with the invading single strand.

The acidic domain of TRF1 inhibits its ability to promote strand invasion

Data obtained for a large panel of mutants clearly show a remarkable correlation between the ability to condense and to promote strand invasion (Figure 2). Proteins inefficient in condensing DNA (TRF1, TRF2$^{hA\Delta B}$) were also poor in stimulating invasion in both gel-based and SPR assays. Conversely, efficient proteins in DNA condensation (TRF2, TRF1$^{\Delta A}$, TRF2$^{\Delta B}$ and TRF2$^{\Delta D}$TRF1^{D}) were found to be active in strand invasion. In summary, we clearly establish that TRF proteins have the inherent capacity of condensing DNA, modifying DNA topology and stimulating invasion but this intrinsic property is nearly lost in TRF1 through the presence of an acidic domain at its N-terminus while in TRF2 it is stimulated by the presence of a basic domain. This prompted us to investigate how these domains evolved.

Evolution of the N-terminal domains of TRF1 and TRF2 in the vertebrate's lineage

As telomeric DNA is highly conserved, we began our analysis by blasting (Pblast) the human Telobox DNA-binding domain against Metazoans. As expected (5,7), we found that TRF Telobox sequences form a distinct monophyletic group (Supplementary Figure S4). Moreover, it is possible to distinguish a signature differentiating TRF1 from TRF2 (Supplementary Figure S5). Studies of the available genomes of Prochordates, the Urochordates Ciona intestinalis and Ciona savignyi, and the Cephalochordate Amphioxus, Branchiostoma floridae, reveal a single copy of the telobox with sequences highly similar to that of the two human TRFs. Of note, the genome of the lamprey (*Petromyzon marinus*), one of the most basal organisms that split off from the Gnathostomes about 540 MYA ago, shows the two forms of telobox. Altogether, these results are in favor of a duplication of a unique ancestral gene during one of the two rounds of genome duplications (32) that occurred early in the Chordate lineage.
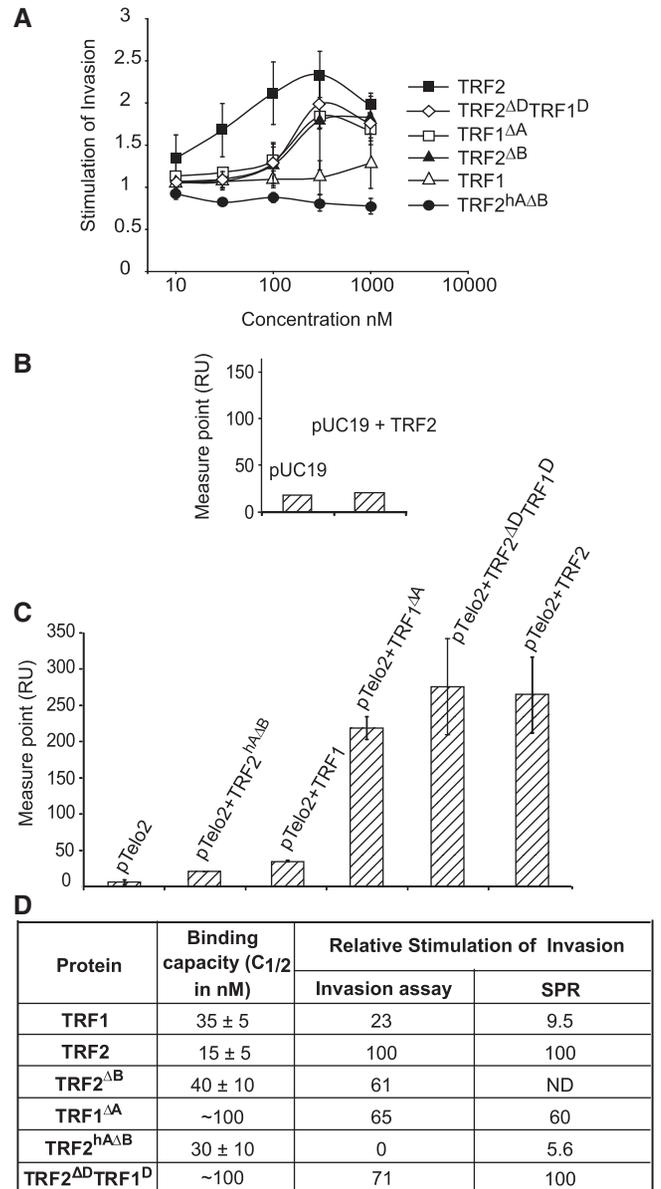

Figure 2. TRF1 can stimulate invasion but is prevented to do so by its N-terminal acidic domain. (**A**) Invasion assays performed with supercoiled pTelo2 and increasing concentrations (10, 30, 100, 300 nM and 1 μM) of TRF1, TRF2, TRF2$^{\Delta D}$TRF1^{D}, TRF1$^{\Delta A}$, TRF2$^{\Delta B}$ and TRF2$^{hA\Delta B}$. (**B**) Histogram of the SPR values obtained after injection of a control plasmid without telomeric repeats (pUC19, 50 nM) with or without pre-incubation with TRF2 (200 nM). (**C**) Histogram of the SPR values obtained after injection of supercoiled pTelo2 (50 nM) with or without pre-incubation with TRF1, TRF2, TRF2$^{\Delta D}$TRF1^{D}, TRF1$^{\Delta A}$ and TRF2$^{hA\Delta B}$ (200 nM). Error bars show standard deviation. (**D**) Summary table showing, for each protein, the concentration necessary for binding half the quantity of a double-stranded DNA in a standard EMSA (binding capacity), the relative stimulation of invasion calculated at maximum effect and the relative increase in SPR response both using TRF2 as reference.

Protein	Binding capacity ($C_{1/2}$ in nM)	Relative Stimulation of Invasion	
		Invasion assay	SPR
TRF1	35 ± 5	23	9.5
TRF2	15 ± 5	100	100
TRF2$^{\Delta B}$	40 ± 10	61	ND
TRF1$^{\Delta A}$	~100	65	60
TRF2$^{hA\Delta B}$	30 ± 10	0	5.6
TRF2$^{\Delta D}$TRF1^{D}	~100	71	100

Outside Eutherians, only a few N-terminal sequences are available (Figure 3A, accession numbers in Supplementary Figure S6). However, some key species allow the plotting of a plausible scenario for the evolution of these domains. The TRF2 B domain is very short in
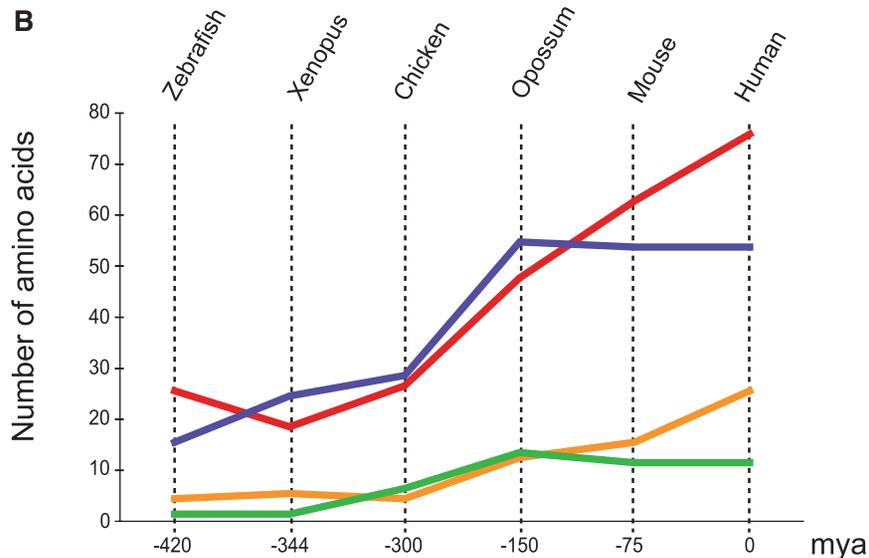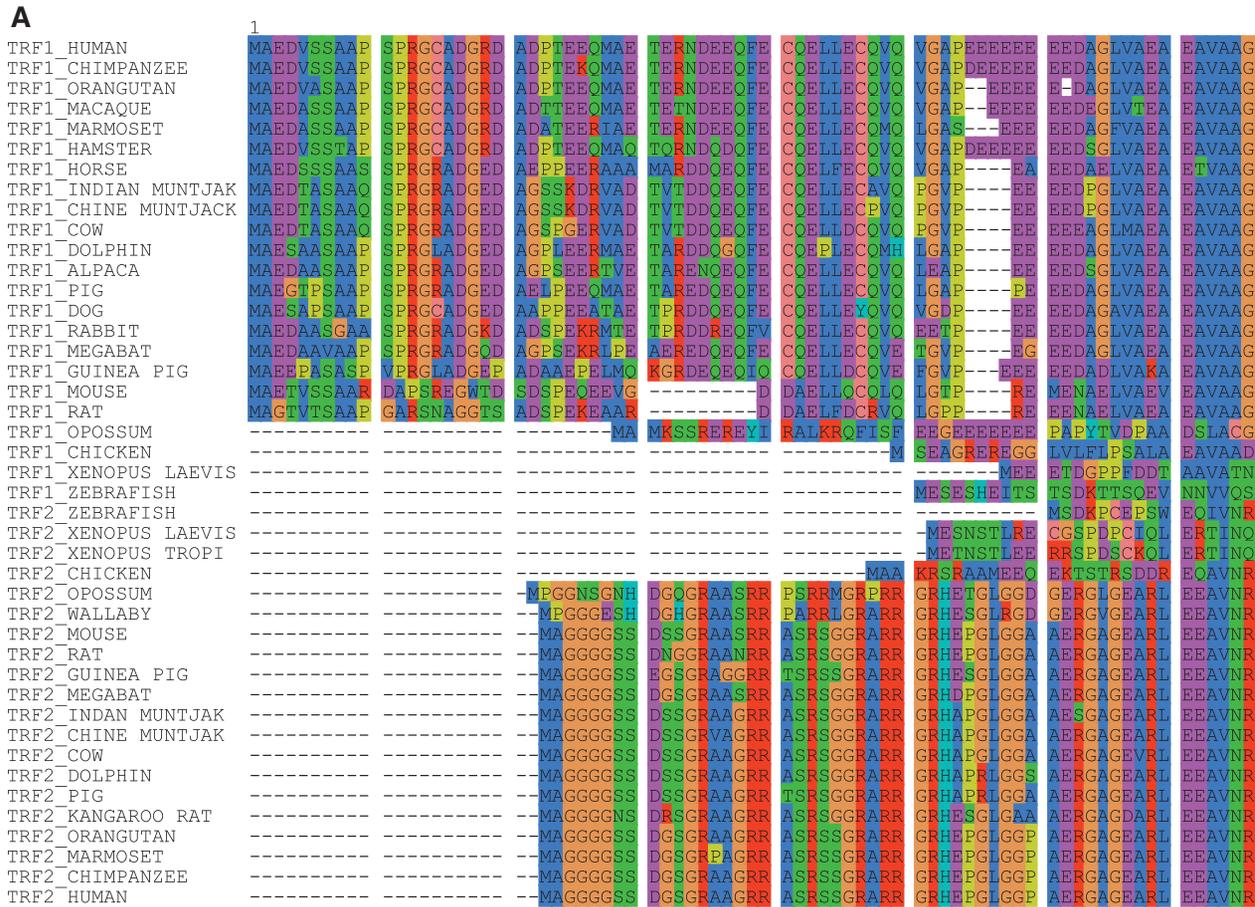
Figure 3. Evolutionary analyses of TRF genes. (**A**) Alignment of the N-terminal sequences of TRF proteins in different species. The N-terminal domains were defined by alignment of the protein sequences through their TRFH domain. (**B**) Curves showing the variations of the length in amino acids of the N-terminal domain of TRF1 (red) or TRF2 (blue) and of the number of acidic and basic residues in these domains (TRF1 in orange and TRF2 in green) as a function of the time since speciation on the lineage leading to the Eutherians.

zebrafish (16 residues), and increases in Xenopus, and even further in chicken. The gray short-tailed opossum has a B domain very similar to that of the placental mammals. For TRF1, the Xenopus sequences are

shorter than the zebrafish domain (19 residues), but show a high proportion of acidic residues. The chicken has a longer domain (27 residues), with several acidic residues. The opossum has also a long domain (48 residues)

but the Eutherians have the longest acidic domains (from 63 to 76 residues) which are also the richest in acidic residues. Thus, the evolution leading to the Eutherians seems to correlate with a gradual increase in both length and acidic/basic residues composition of the N-terminus of TRF proteins (Figure 3B).

Acquisition of these domains does not seem to result from exon addition but rather an extension of the first TRF exon by a continuous gain of the 5′ non-coding sequence, since: (i) we observe a gradual increase in the length and (ii) base composition of the 5′UTR is strikingly very similar to that of the corresponding A/B domains. In the vicinity of the human terf1 and terf2 genes, we calculated a striking GC content of 0.639 (5′UTR) and 0.677 (A domain) in TRF1, and 0.708 (5′UTR) and 0.793 (B domain) in TRF2. This suggests that the GC content of the 5′UTR is correlated to that of the A/B domain. Moreover, GC-rich nucleotidic sequences, when incorporated into a coding sequence, tend to yield larger

numbers of charged residues (particularly D, E, R, S and G, which are hallmarks of the N-terminal domains of TRF1 and TRF2).

Altogether, pending on sequences available so far, the A/B domains of TRF proteins seem to originate from a gradual gain of amino acids by successive and iterative incorporation of the 5′UTR in the coding sequence all along the lineage leading to the Eutherians.

Importance of the length of the N-terminal acidic domain of TRF1

To test the consequences of the evolution of the N-terminal domains on the biochemical properties of TRF1 and TRF2, we analyzed the capacity of the acidic domain from mouse and chicken to inhibit strand invasion and DNA condensation (Figure 4). We fused these N-termini to TRF2$^{\Delta B}$ (TRF2$^{mA\Delta B}$ and TRF2$^{cA\Delta B}$ for mouse and chicken sequences, respectively) and characterized these chimeric proteins. Mouse (63 residues) and
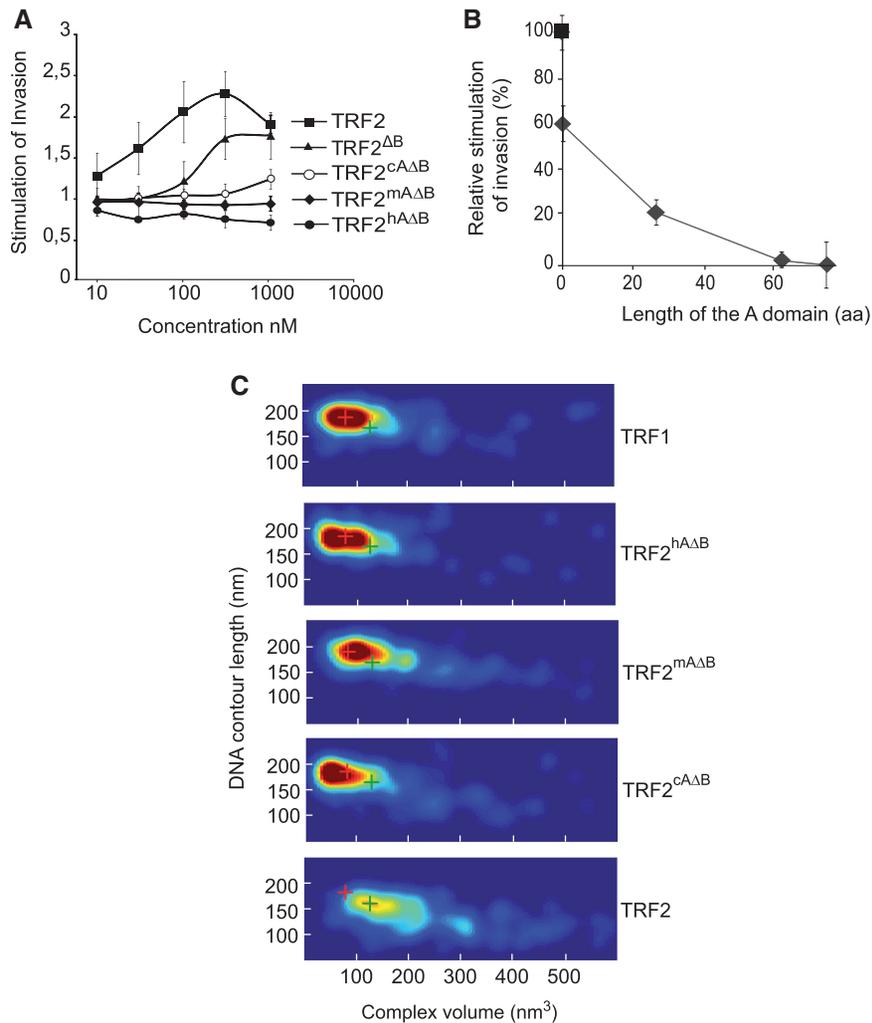

Figure 4. Importance of the length of the N-terminal acidic domain of TRF1. (**A**) Invasion assays performed with supercoiled pTelo2 and increasing concentrations (10, 30, 100, 300 nM and 1 μM) of TRF2, TRF2$^{\Delta B}$, TRF2$^{hA\Delta B}$, TRF2$^{mA\Delta B}$ and TRF2$^{cA\Delta B}$. (**B**) Variations of the relative stimulation of invasion at maximum as a function of the length of the acidic N-terminal domain. (**C**) 2D-probability density maps of volumes and DNA CLs measured by AFM for the complexes formed with TRF1, TRF2, TRF2$^{cA\Delta B}$, TRF2$^{mA\Delta B}$ and TRF2$^{hA\Delta B}$. The red and green crosses indicate the position in the graph of the majority of TRF1 and TRF2 complexes, respectively.

chicken domains (27 residues) are shorter than the human one (76 residues), but have similar isoelectric points (3.65, 3.73 and 4.4 for human, mouse and chicken, respectively).

Invasion assays (Figure 4A and B) and AFM (Figure 4C) experiments show a similar trend and suggest a correlation between the length of the acidic domain and the ability to inhibit DNA condensation and to stimulate invasion. The longer the A domain, the less strand invasion and DNA condensation was observed. Our results indicate that evolution of TRF1 proteins might have been associated with a progressive inhibition of their capacity to condense DNA and to stimulate telomeric invasion.

Telomeric RNA negatively regulates TRF2-mediated DNA condensation

The binding of molecules to the N-termini of TRF1 and TRF2 is expected to have an effect on their behavior regarding strand invasion and DNA condensation. For TRF2, one possible candidate is TERRA, the telomeric RNA that was recently proposed to interact with the N-terminal basic domain of TRF2 (33). To study this, we performed topoisomerase I assays and AFM experiments in the presence of a G-RNA containing three (UUAGGG) repeats and a C-RNA bearing the $(CCCUAA)_3$ sequence. As seen in Figure 5A and B, while the C-RNA had little effect on TRF2-mediated modification of topology, the G-RNA greatly reduced the ability of TRF2 to condense DNA. This was not due to a reduction in the binding of duplex DNA since no important changes in the amount of DNA bound by EMSA were observed in the presence of either RNA (Figure 5C). In accordance with published results (33), direct binding of TRF2 on these RNAs (Figure 5D) showed that TRF2 strongly prefers the G-RNA to the C-RNA, which could explain their differential effect. Of note, a faint band of C-RNA-TRF2 complex is visible at the highest concentration of protein, suggesting that recognition of the C-RNA by TRF2 is possible. AFM experiments show that adding G-RNA to TRF2 prior to DNA binding decreases its capacity to form condensed complexes. TRF2 complexes strikingly resemble those obtained for TRF1 (Figure 5E). The C-RNA has a much weaker but not negligible effect probably due to the residual binding mentioned above.

Collectively, these experiments show that the binding of G-RNA on TRF2 severely hinders its capacity to condense DNA and to modify DNA topology.

DISCUSSION

Our phylogenetic studies on the terf1 and terf2 genes suggest that they originated from the duplication of an ancestral gene and that their N-termini gradually increased in size through iterative incorporation of sequences from their 5′UTR. These additions had profound consequences on the behavior of these proteins. They created binding sites for several telomere-interacting molecules (3) and conferred to TRF2 the ability to bind Holliday junctions and to protect them from resolvase cleavage (25). Here, we reveal that these domains also make important contributions to the ability of TRF1 and TRF2 to condense telomeric DNA and to stimulate telomeric invasion. Indeed, we show that TRF1 inefficiency in modifying topology and stimulating invasion can be attributed to the acidic nature of its N-terminus. Conversely, the basic nature of the N-terminus of TRF2 causes an increase in these capacities for TRF2. Analyzing the acidic domains of TRF1 proteins from chicken, mouse and human, we have found that the length rather than the pI seems important, indicating that the functional divergence between TRF1 and TRF2 may have gradually increased all along the Chordate lineage.

TRF2-mediated DNA condensation leads to the untwisting/unwrithing of the surrounding constrained DNA which is thought to stimulate telomeric invasion, a reaction involved in the folding of telomeric DNA into t-loop (26). Our data strongly reinforce this view since we observe a striking correlation between the capacity of TRF2 and its mutants to condense DNA, to modify DNA topology and to stimulate invasion. Furthermore, we have uncovered that stimulation of telomeric invasion is mediated by a striking increase in the kinetics of the invasion reaction (steady state is reached in 15 min in the presence of TRF2, 11 h in its absence).

Overall, these data suggest a role for TRF2 in the regulation of DNA topology on telomeres. In accordance, a recent study (34) shows that TRF2 acts together with topoisomerase 2 to protect telomeric DNA from replicative DNA damage. One major future challenge will be to elucidate the mechanism and the precise role of TRF2 in this process.

One interesting idea raised by our work is that modification(s) of the N-terminal domains of TRF1 and TRF2 might result in a TRF1 protein exhibiting TRF2-like behavior or *vice versa*. To our knowledge, modifications of TRF1 N-terminal domain have not been reported so far. However, one cannot exclude the possibility of a regulation through the binding of a protein partner. A tankyrase-binding motif exists in most of the TRF1 N-termini from mammals but tankyrases stand as poor candidates since mouse and rat TRF1s lack this motif and their ultimate role is the inhibition of TRF1 DNA binding (3). In contrast, the N-terminal basic domain of TRF2 is the target of several binding partners and, at least, one post-translational modification. Arginines 17 and 18 in this domain were shown to be methylated (35). Proteins such as the Werner protein as well as ORC1 were shown to recognize this domain (36,37). Similarly, the telomeric RNA, TERRA, was recently proposed to interact with the N-terminus of TRF2 (33). TERRA is a very attractive candidate for regulating the functions of the N-terminus of TRF2 since it could both mask the positive charges of this domain and repulse DNA. To investigate this hypothesis we performed AFM experiments and topology assays in the presence of small RNA molecules containing telomeric repeats. The effect of the G-rich small RNA was striking with a total abrogation of the capacity of TRF2 to condense DNA (Figure 5). Although these RNA molecules are far shorter than the natural TERRA, which can reach 9 kb, and lack the sub-telomeric sequences present in TERRA, it is tempting
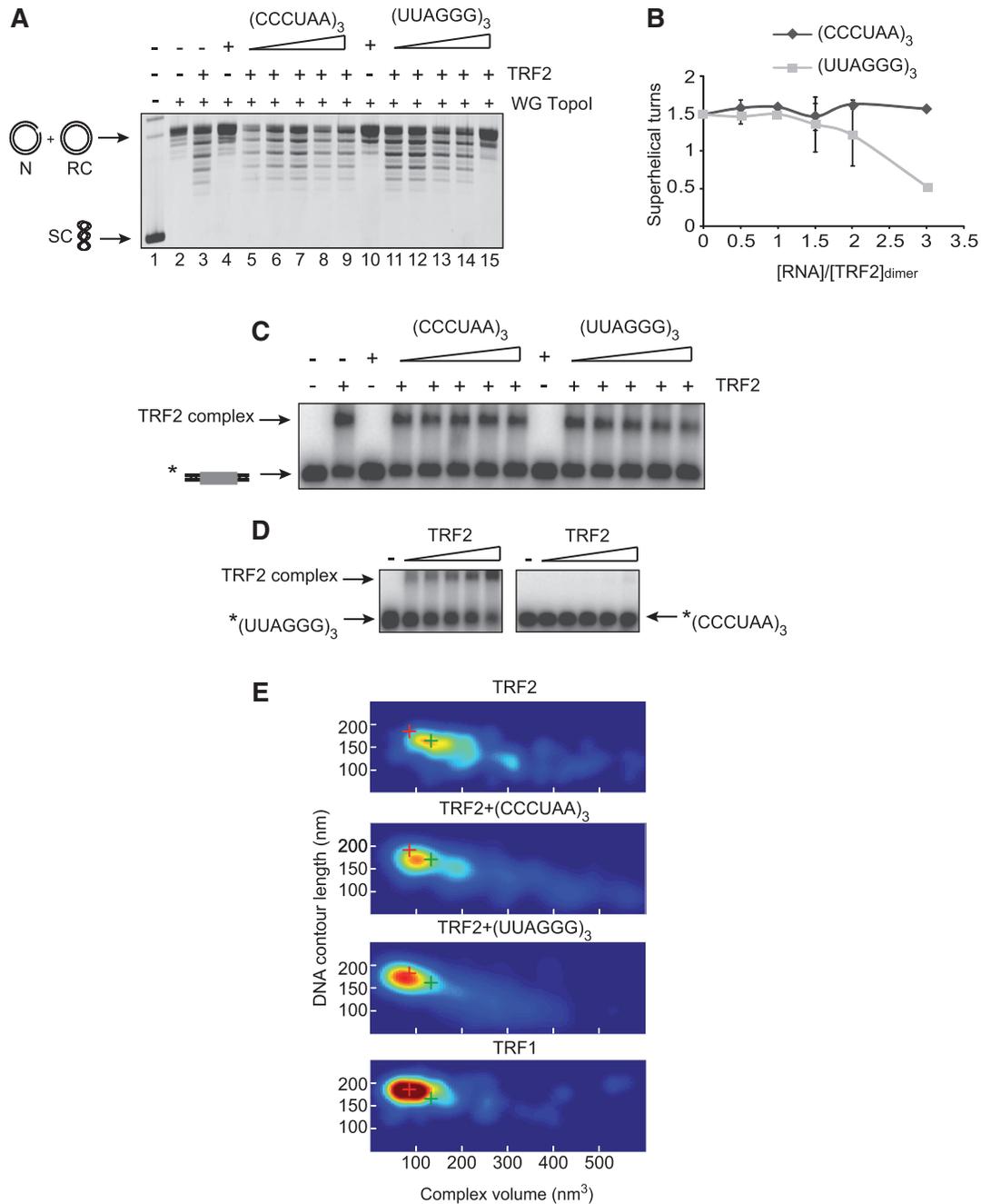
Figure 5. Telomeric RNA negatively regulates TRF2-mediated DNA condensation. (**A**) Topology assay performed using the telomeric plasmid pLTelo and 1 μM of TRF2 in the presence of wheat germ topoisomerase I (WG Topo I) and increasing amounts of either G- or C-RNA (0.5, 1, 1.5, 2 and 3 μM). SC stands for supercoiled plasmid, RC for relaxed circular, N for nicked. (**B**) Plot of data in (A) showing the variations of the average number of supercoils as a function of the ratio between the concentration of dimers of TRF2 and the concentration of G-RNA (light gray curve) or C-RNA (dark gray curve). Error bars correspond to standard deviation. (**C**) EMSA using 5 nM of a double-stranded telomeric probe and 20 nM of TRF2 in the presence of increasing amounts of G- and C- RNA (10, 20, 30, 40 and 60 nM). (**D**) EMSA using 5 nM of labeled G- and C-RNA and increasing amounts of TRF2 (10, 30, 50, 100 and 300 nM). (**E**) 2D-probability density maps of volumes and DNA CLs measured by AFM for complexes formed with TRF1 and TRF2 in the presence of a ratio of 3 G- and C-RNA per TRF2 dimer. The red and green crosses indicate the position in the graph of the majority of TRF1 and TRF2 complexes, respectively.

to speculate that TERRA may impact TRF2 behavior concerning DNA topology and even t-loop formation. One could imagine that TERRA could guide TRF2 complexes from a DNA-condensation/t-loop proficient type of complex to a more TRF1/shelterin proficient type. It follows that TERRA could impede t-loop

formation, an event that could be deleterious to the cell but may also be a necessary step during replication. In accordance, TERRA has been proposed to regulate some aspects of telomere replication (33). TERRA could even be involved in the removal of the t-loop. Indeed, the N-terminal domain of TRF2 has been implicated in the

protection of the t-loop (38) and Holliday junctions (25) against resolution events. One could speculate that TERRA binding could alleviate this protection and cause t-loop HR. In accordance with this hypothesis, it has been observed that depletion in Nonsense-mediated decay (NMD) proteins that increase the number of TERRA foci in human cells causes sudden telomere loss and a marked increase in telomeric fragments (39). This would make TERRA a crucial factor in regulating telomeric folding and state.

This work also raises the question of why different N-termini have evolved for TRF1 and TRF2 during Chordate evolution. We show here that these domains regulate the ability of TRF to condense telomeric DNA. Strikingly, the N-terminal part of TRF2 might facilitate heterochromatin formation by binding both ORC1 and TERRA (33), suggesting that the N-terminal parts of the TRF proteins are crucial to regulate telomeric chromatin condensation both through intrinsic properties and interactions with chromatin components. This is expected to be crucial for proper organization and dynamics of telomeric chromatin during the cell cycle, but it could also lead to long-range chromatin interactions between telomeres and non-telomeric chromosomal loci. Recent studies from our group and that of Zhou Songyang (40,41) have revealed the presence of TRF2 outside telomeres, specifically, on centromeric/pericentromeric satellite DNA and intersterstitial telomeric sequences. Often located in the proximity of genes or within introns, these TRF2 binding sites may be involved in the regulation of the corresponding genes and thus participate in the cell transcriptional program. In view of the present results, it is tempting to speculate that the acquisition of specialized N-termini by TRF proteins during the evolution of chordates may have thus contributed to the establishment of new transcriptional regulatory programs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Information, Supplementary Figures 1–6.

ACKNOWLEDGEMENTS

FUNDING

Conflict of interest statement. None declared.

REFERENCES

1. Blackburn,E.H. (2005) Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. *FEBS Lett.*, **579**, 859–862.
2. Luke,B. and Lingner,J. (2009) TERRA: telomeric repeat-containing RNA. *EMBO J.*, **28**, 2503–2510.
3. Palm,W. and de Lange,T. (2008) How shelterin protects mammalian telomeres. *Annu. Rev. Genet.*, **42**, 301–334.
4. Xin,H., Liu,D. and Songyang,Z. (2008) The telosome/shelterin complex and its functions. *Genome Biol.*, **9**, 232.
5. Bilaud,T., Koering,C.E., Binet-Brasselet,E., Ancelin,K., Pollice,A., Gasser,S.M. and Gilson,E. (1996) The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. *Nucleic Acids Res.*, **24**, 1294–1303.
6. Chong,L., van Steensel,B., Broccoli,D., Erdjument-Bromage,H., Hanish,J., Tempst,P. and de Lange,T. (1995) A human telomeric protein. *Science*, **270**, 1663–1667.
7. Bilaud,T., Brun,C., Ancelin,K., Koering,C.E., Laroche,T. and Gilson,E. (1997) Telomeric localization of TRF2, a novel human telobox protein. *Nat. Genet.*, **17**, 236–239.
8. Broccoli,D., Smogorzewska,A., Chong,L. and de Lange,T. (1997) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat. Genet.*, **17**, 231–235.
9. Baumann,P. and Cech,T.R. (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science*, **292**, 1171–1175.
10. Houghtaling,B.R., Cuttonaro,L., Chang,W. and Smith,S. (2004) A dynamic molecular link between the telomere length regulator TRF1 and the chromosome end protector TRF2. *Curr. Biol.*, **14**, 1621–1631.
11. Kim,S.H., Kaminker,P. and Campisi,J. (1999) TIN2, a new regulator of telomere length in human cells. *Nat. Genet.*, **23**, 405–412.
12. Li,B., Oestreich,S. and de Lange,T. (2000) Identification of human Rap1: implications for telomere evolution. *Cell*, **101**, 471–483.
13. Liu,D., O'Connor,M.S., Qin,J. and Songyang,Z. (2004) Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *J. Biol. Chem.*, **279**, 51338–51342.
14. Ye,J.Z., Hockemeyer,D., Krutchinsky,A.N., Loayza,D., Hooper,S.M., Chait,B.T. and de Lange,T. (2004) POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev.*, **18**, 1649–1654.
15. O'Connor,M.S., Safari,A., Xin,H., Liu,D. and Songyang,Z. (2006) A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly. *Proc. Natl Acad. Sci. USA*, **103**, 11874–11879.
16. Cesare,A.J., Quinney,N., Willcox,S., Subramanian,D. and Griffith,J.D. (2003) Telomere looping in P. sativum (common garden pea). *Plant J.*, **36**, 271–279.
17. Griffith,J.D., Comeau,L., Rosenfield,S., Stansel,R.M., Bianchi,A., Moss,H. and de Lange,T. (1999) Mammalian telomeres end in a large duplex loop. *Cell*, **97**, 503–514.
18. Munoz-Jordan,J.L., Cross,G.A., de Lange,T. and Griffith,J.D. (2001) t-loops at trypanosome telomeres. *EMBO J.*, **20**, 579–588.
19. Nikitina,T. and Woodcock,C.L. (2004) Closed chromatin loops at the ends of chromosomes. *J. Cell Biol.*, **166**, 161–165.
20. Tomaska,L., Makhov,A.M., Griffith,J.D. and Nosek,J. (2002) t-Loops in yeast mitochondria. *Mitochondrion*, **1**, 455–459.
21. Stansel,R.M., de Lange,T. and Griffith,J.D. (2001) T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. *EMBO J.*, **20**, 5532–5540.
22. Yanez,G.H., Khan,S.J., Locovei,A.M., Pedroso,I.M. and Fletcher,T.M. (2005) DNA structure-dependent recruitment of telomeric proteins to single-stranded/double-stranded DNA junctions. *Biochem. Biophys. Res. Commun.*, **328**, 49–56.
23. Yoshimura,S.H., Maruyama,H., Ishikawa,F., Ohki,R. and Takeyasu,K. (2004) Molecular mechanisms of DNA end-loop formation by TRF2. *Genes Cells*, **9**, 205–218.

24. Bae,N.S. and Baumann,P. (2007) A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends. *Mol. Cell*, **26**, 323–334.
25. Poulet,A., Buisson,R., Faivre-Moskalenko,C., Koelblen,M., Amiard,S., Montel,F., Cuesta-Lopez,S., Bornet,O., Guerlesquin,F., Godet,T. *et al.* (2009) TRF2 promotes, remodels and protects telomeric Holliday junctions. *EMBO J.*, **28**, 641–651.
26. Amiard,S., Doudeau,M., Pinte,S., Poulet,A., Lenain,C., Faivre-Moskalenko,C., Angelov,D., Hug,N., Vindigni,A., Bouvet,P. *et al.* (2007) A topological mechanism for TRF2-enhanced strand invasion. *Nat. Struct. Mol. Biol.*, **14**, 147–154.
27. Baker,A.M., Fu,Q., Hayward,W., Victoria,S., Pedroso,I.M., Lindsay,S.M. and Fletcher,T.M. (2011) The telomere binding protein TRF2 induces chromatin compaction. *PLoS One*, **6**, e19124.
28. Galtier,N., Gouy,M. and Gautier,C. (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci.*, **12**, 543–548.
29. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
30. Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R. *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
31. Piliarik,M., Vaisocherova,H. and Homola,J. (2009) Surface plasmon resonance biosensing. *Methods Mol. Biol.*, **503**, 65–88.
32. Dehal,P. and Boore,J.L. (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol*, **3**, e314.
33. Deng,Z., Norseen,J., Wiedmer,A., Riethman,H. and Lieberman,P.M. (2009) TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol. Cell*, **35**, 403–413.
34. Ye,J., Lenain,C., Bauwens,S., Rizzo,A., Saint-Léger,A., Poulet,A., Benarroch,D., Magdinier,F., Morere,J., Amiard,S. *et al.* (2010) TRF2 and Apollo cooperate with Topoisomerase 2alpha to protect human telomere from replicative damage. *Cell*, **142**, 230–242.
35. Mitchell,T.R., Glenfield,K., Jeyanthan,K. and Zhu,X.D. (2009) Arginine methylation regulates telomere length and stability. *Mol. Cell. Biol.*, **29**, 4918–4934.
36. Deng,Z., Dheekollu,J., Broccoli,D., Dutta,A. and Lieberman,P.M. (2007) The origin recognition complex localizes to telomere repeats and prevents telomere-circle formation. *Curr. Biol*, **17**, 1989–1995.
37. Li,B., Jog,S.P., Reddy,S. and Comai,L. (2008) WRN controls formation of extrachromosomal telomeric circles and is required for TRF2DeltaB-mediated telomere shortening. *Mol. Cell. Biol.*, **28**, 1892–1904.
38. Wang,R.C., Smogorzewska,A. and de Lange,T. (2004) Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell*, **119**, 355–368.
39. Azzalin,C.M., Reichenbach,P., Khoriauli,L., Giulotto,E. and Lingner,J. (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science*, **318**, 798–801.
40. Simonet,T., Zaragosi,L.E., Philippe,C., Lebrigand,K., Schouteden,C., Augereau,A., Bauwens,S., Ye,J., Santagostino,M., Giulotto,E. *et al.* (2011) The human TTAGGG repeat factors 1 and 2 bind to a subset of interstitial telomeric sequences and satellite repeats. *Cell Res.*, **21**, 1028–1038.
41. Yang,D., Xiong,Y., Kim,H., He,Q., Li,Y., Chen,R. and Songyang,Z. (2011) Human telomeric proteins occupy selective interstitial sites. *Cell Res.*, **21**, 1013–1027.