nature structural & molecular biology

Human mitochondrial mTERF wraps around DNA through a left-handed superhelical tandem repeat

Nereida Jiménez-Menéndez[1,2], Pablo Fernández-Millán[1], Anna Rubio-Cosials[1], Carme Arnan[1,2], Julio Montoya[3], Howard T Jacobs[4], Pau Bernadó[2], Miquel Coll[1,2], Isabel Usón[1,5] & Maria Solà[1]

The regulation of mitochondrial DNA (mtDNA) processes is slowly being characterized at a structural level. We present here crystal structures of human mitochondrial regulator mTERF, a transcription termination factor also implicated in replication pausing, in complex with double-stranded DNA oligonucleotides containing the *tRNA*$^{Leu}_{UUR}$ gene sequence. mTERF comprises nine left-handed helical tandem repeats that form a left-handed superhelix, the Zurdo domain.

Human mitochondrial regulator mTERF binds to several strategic mtDNA sites, where it controls processes such as transcription termination or pausing of the replication machinery[1,2]. Among them, the sequence from the *tRNA*$^{Leu}_{UUR}$ gene shows the highest affinity of binding to mTERF[3] (**Supplementary Fig. 1**). We set out to characterize the structural basis of this interaction. During recombinant mTERF production, spontaneous proteolysis of the full-length protein (Arg56–Ala399) yielded a shorter variant (Arg99–Ala399; mTERF ΔN) yet DNA-binding form, based on EMSA. The three-dimensional structure of mTERF ΔN was determined from SeMet-derivatized crystals and subsequently used to solve the full-length mTERF structure at 3.1-Å resolution (see **Supplementary Methods** and **Supplementary Table 1**). Both constructs crystallized only in the presence of cognate dsDNA oligonucleotides, consistent with an intrinsic flexibility of the protein in the absence of bound DNA. mTERF consists of nine structural repeats, TERF-I–TERF-IX, which correspond to tandem sequence motifs identified throughout the MTERF family[4–6]. The repeats are flanked by an N-terminal extended segment (Arg55–Glu73) and a C-terminal α-helix C (Lys385–Ala399; **Fig. 1a**). Within each repeat, roughly 35 residues fold into three α-helices (H1, H2 and H3) that are arranged in a left-handed triangular superhelix around a central hydrophobic core (**Fig. 1b**,**c**). The nine repeats successively rotate 7–20°, building up a solenoid-like structure, which further twists to the right, presenting a convex and a concave face (**Fig. 1a**). We call this structure the Zurdo ('left-handed' in Spanish) domain. In each repeat, H1 and H2, the longest and intermediate helices, respectively, intersect at 60°. The ladder of H1-H2 pairs forms the convex surface of the solenoid, and H3 helices build up the concave wall. Previously characterized helical-repeat solenoids, such as Armadillo, HEAT and pumilio domains[7–10], showed right-handed helical connectivity (**Fig. 1b**). Left-handed superhelices have previously been reported only for small helical bundles present in diverse proteins[11,12]. The combination of left-handed helical connectivity with tandem propagation is unique to mTERF.

A structure-guided sequence alignment of the mTERF repeats and their mapping onto several orthologs generated a logo of 54 positions (**Supplementary Fig. 1**). Repeats II–V and VIII (but, notably, not VII) show four residues within H1 (at logo positions 14–17) with the pattern BOBB (B for hydrophobic, O for hydrophilic residues). Position 16 contacts position 14 of the downstream repeat, position 15 is exposed to solvent and position 17 is buried in the hydrophobic core

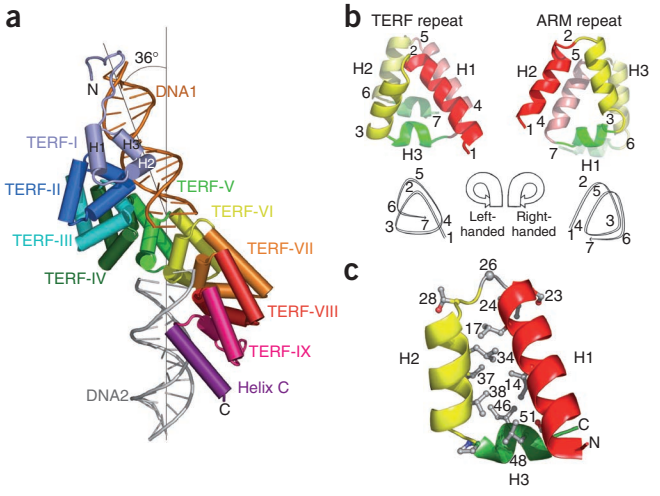

Figure 1 The mTERF Zurdo domain. (**a**) Repeats I–IX consist of three helices, H1, H2 and H3, shown here for TERF-I. In the mTERF crystal, one protein molecule binds two symmetry-related oligonucleotides, DNA1 (orange) and DNA2 (gray), related by ~36°. N and C, N- and C-terminal ends. (**b**) Explanation of left- and right-handedness. Left, starting from point 1 and following the numbers, the TERF superhelix curves to the left (bottom). Right, the Armadillo (ARM) repeat shows right-handed tracing. (**c**) Ball-and-stick representation on TERF-VII of residues mutated and numbered according to the mTERF repeats alignment and logo (see **Supplementary Fig. 1**).

[1]Institut de Biologia Molecular de Barcelona, Parc Cientific de Barcelona, Barcelona, Spain. [2]Institute for Research in Biomedicine, Parc Científic de Barcelona, Barcelona, Spain. [3]Departamento de Bioquímica y Biología Molecular y Celular, Universidad de Zaragoza - CIBER de Enfermedades Raras, Zaragoza, Spain. [4]Institute of Medical Technology and Tampere University Hospital, University of Tampere, Tampere, Finland. [5]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain. Correspondence should be addressed to M.S. (maria.sola@ibmb.csic.es).

Received 24 April; accepted 25 May; published online 13 June 2010; doi:10.1038/nsmb.1859

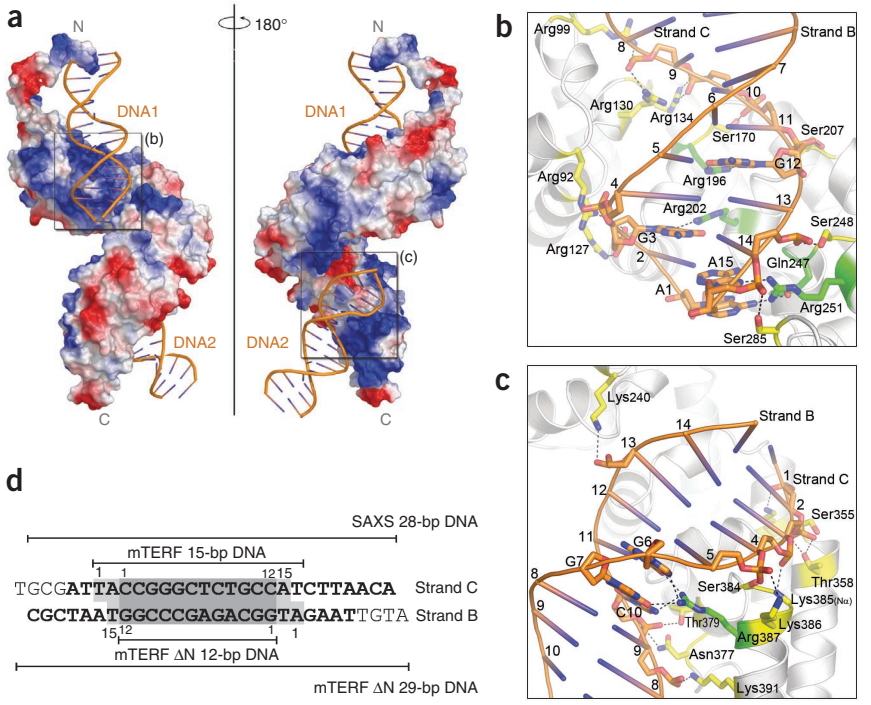

Figure 2 Protein-DNA interactions between mTERF and a 15-bp DNA. (**a**) Electrostatic potential mapped onto the mTERF Connolly surface is shown in two 180°-related views. DNA1 and DNA2 are symmetrically related oligonucleotides that bind to two different areas of a single protein molecule (see text). (**b**) Side chains of N-terminal residues contacting DNA1 are shown as sticks. DNA strands are named according to the PDB coordinates file. DNA bases are indicated in one-letter code. (**c**) Contacts between C-terminal residues and DNA2 bases shown as in **b**. (**d**) Sequence of 12-bp (dark gray), 15-bp (light gray) and 28- and 29-bp oligonucleotides used for crystallization and SAXS experiments. Strands named as in **b**.

also comprising positions 34, 37 and 38 from H2 and 46 and 48 from H3. This hydrophobic core spans the entire protein. In all tandem repeats except the first, logo position 26 (usually glycine), between H1 and H2, shows main chain conformation angles corresponding to a left-handed helix. In addition, the main chain kinks by about 90° at position 44 (usually a proline at the beginning of H3). Glycine and proline tend to disrupt regular secondary-structure elements and here confer the triangular shape of the TERF repeat. Finally, positions 23, 28 and 51 are involved in N- or C-terminal capping of helices H1, H2 and H3, respectively. All these conserved residue define the overall structure of the mTERF repeat (**Fig. 1c**).

Protein variants mTERF and mTERF ΔN crystallized in complexes with a 15-mer and a 12-mer dsDNA, respectively, which included the sequence of the $tRNA^{Leu}_{UUR}$ gene[13] (**Fig. 2**). In both crystals, which belong to different space groups, two crystallographically related, identical oligonucleotides (DNA1 and DNA2) are bound by two distinct positively charged regions of a single protein moiety (**Fig. 2a** and **Supplementary Fig. 2**). In the full-length complex, repeats I–VI contact one half of DNA1, whereas repeats VIII and IX plus helix C approach the other half of DNA2. DNA1 and DNA2 do not contact each other and are oriented at an angle of 36° (**Fig. 1a**), in agreement with previous assays that estimated a bending of the termination site of 35° upon protein binding[14]. The only repeat not engaged in DNA binding, TERF-VII, lacks the polar residue at logo positions 50/51 that is involved in DNA contacts in the other repeats (**Supplementary Fig. 1b**).

This feature suggests for repeat VII a hinge role between the N- and C-terminal DNA-binding subdomains.

The full-length protein is engaged in 19 nonspecific interactions with both dsDNA backbone phosphates (**Supplementary Fig. 1b**). Eight of the residues involved further perform C-capping of α-helices at the protein-DNA interface, thus diminishing the repulsion between peptide-bond carbonyls and DNA phosphates. N-terminal repeats I–VI contribute, through residues at logo positions 39, 45/49 and 50/51, to backbone recognition. Two conserved guanines (G12 from strand C and G3 from B) of the binding sequence[2] are positioned for contacts with Arg169 and Arg202, respectively (**Fig. 2b** and **Supplementary Fig. 2a**,**b**). An adjacent conserved guanine (G4 from strand B) is weakly contacted by Glu165. In the C-terminal subdomain, residues from repeats VIII and IX (logo positions 9 and 13 of H1 and 45 and 51 of H3) contact the phosphate backbone of DNA2 and surround Arg387 from helix C, which interacts through the major groove with three bases thereof (**Fig. 2c** and **Supplementary Fig. 2a**,**c**). Additional contacts contribute to stabilizing the termini of DNA1 and DNA2 within the concave protein surface. Overall, the mTERF residues involved in DNA contacts are conserved throughout the vertebrate MTERF1 subfamily members[4], suggesting that similar complexes could be formed by these proteins.

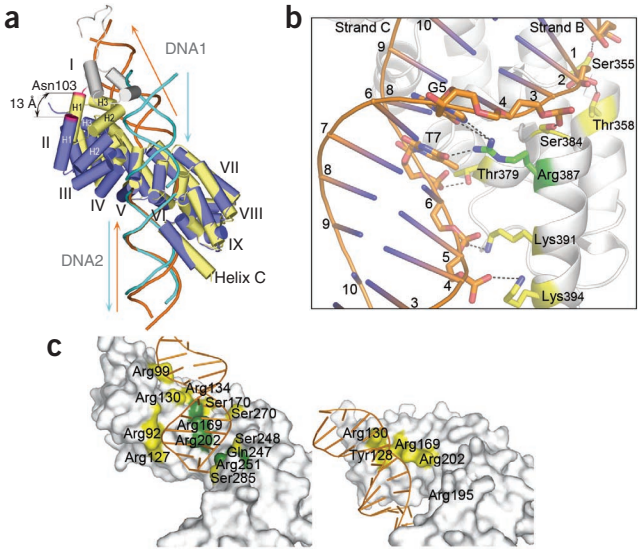

Figure 3 Excision of TERF-I results in the reorganization of the protein structure. (**a**) Structural superimposition of C-terminal repeats of mTERF (yellow) and mTERF ΔN (blue) reveals the relative shift in mTERF and mTERF ΔN N-terminal repeats and DNA1 oligonucleotides (dsDNA of mTERF in orange; dsDNA of mTERF ΔN in cyan). TERF-I, in gray, is excised in mTERF ΔN. TERF-II helices H1–H3 are labeled, and the maximal displacement at Asn103 (magenta) is shown. (**b**) Contacts between mTERF ΔN C-terminal repeats and DNA2 bases, shown as in **Figure 2b**. (**c**) Connolly surface of N-terminal repeats that contact DNA1 in mTERF (left) and mTERF ΔN (right), both in the same orientation. Green, residues contacting the DNA bases; yellow, residues contacting the backbone phosphates.

We obtained full-length mTERF in complex with longer dsDNA oligonucleotides (up to 29 base pairs (bp), **Fig. 2d**), which showed higher stability but did not crystallize. To assess whether a continuous dsDNA bridging DNA1 and DNA2 could be bound by mTERF, we analyzed a complex of the protein with a DNA of 28 bp by small-angle X-ray scattering (SAXS) in solution (**Fig. 2d** and **Supplementary Methods**). Based on the crystallographic structure, we constructed 30 models by translating a canonical 28-bp B-DNA in a stepwise base pair manner from DNA1 to DNA2 sites, respecting the crystalline protein-DNA contacts and introducing the required bending. The model that best fit the experimental data showed the concave surface of mTERF wrapping around the major groove of a continuous DNA that protruded on one side of the protein solenoid (**Supplementary Fig. 3**). The quality of the fit validates the model for the particle in solution. The DNA sequence is coherent with the sequence of DNA1 and suggests that the C-terminal domain slightly rearranges to accommodate the DNA. This intrinsic flexibility of the domain is supported by the consistent, slight variation of the angle between repeats when comparing the mTERF and mTERF ΔN solenoids, whose pitches differ by 13 Å (**Fig. 3a**). The N-terminal excision abolishes the interactions with bases of DNA1, which is displaced to form an unbent pseudocontinuous DNA fiber by base stacking with DNA2 (**Fig. 3**). If C-terminal subdomain repeats VII–IX plus helix C of the two structures are superimposed, the respective DNA2-phosphate positions match perfectly (**Fig. 3**). However, the mTERF ΔN structure unambiguously shows DNA2 bound in opposite orientations, with Arg387 contacting different bases in the major groove (**Fig. 3b** and **Supplementary Fig. 2d**). Furthermore, mTERF ΔN in complex with a 15-bp DNA yielded a structure in space group $C222_1$ where two protein moieties symmetrically contacted a single, two-fold–disordered DNA molecule. Overall, this underpins the nonspecific binding capacity of the mTERF ΔN variant to a DNA major groove.

In summary, nine left-handed helical mTERF repeats of three helices each give rise to a twisted solenoid that binds a continuous bent dsDNA. This suggests a similar binding mode, causing DNA bending, to the mtDNA termination site *in vivo*. A number of sequence-specific proteins have been shown to cause an alteration in DNA structure upon binding[15]. The only transcriptional termination factor reported to exert this effect is DNA polymerase I–specific transcription termination factor I, which binds DNA as a monomer and induces a bending of 40° (ref. 16), but no protein–DNA complex structure exists for comparison. The structure of the Zurdo domain is likely relevant to other TERF family members, which notwithstanding must have member-specific features to accomplish their particular functions.

Accession codes. Protein Data Bank: Coordinates and structure factors for mTERF and the mTERF ΔN variant have been deposited under accession codes 3N6S and 3N7Q, respectively.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We acknowledge J. Colom for help during protein purification, G.M. Sheldrick for in-house data collection and S. Brockhauser for help with minikappa. This study was supported by the Spanish Ministerio de Ciencia e Innovación (grants BFU2006-09593, BFU2009-07134, BIO2009-10576, BFU2008-02372, FIS-PI070045), Generalitat de Catalunya (SGR2009-1366 and SGR2009-1309) and the European Community (LSHG-2006-031220). A.R.-C. and P.F.-M. hold fellowships from CSIC and the Ministerio de Ciencia e Innovación. P.B. holds a Ramon y Cajal contract. We acknowledge the Academy of Finland, Tampere University Hospital Medical Research Fund and Sigrid Juselius Foundation. We are grateful to European Molecular Biology Laboratory—Grenoble Outstation, the European Synchrotron Radiation Facility and the Automated Crystallography Platform (Barcelona Science Park) for their support.

AUTHOR CONTRIBUTIONS

N.J.-M. and C.A. contributed to cloning; N.J.-M. to protein production and crystallization; N.J.-M. and P.B. to SAXS studies; N.J.-M., I.U. and M.S. to X-ray structure determination; and N.J.-M., P.F.-M. and A.R.-C. to figure preparation. M.S. and M.C. provided materials and infrastructure. N.J.-M., H.T.J. and the rest of authors participated in manuscript writing and discussion. M.S. designed and supervised the project.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Martin, M., Cho, J., Cesare, A.J., Griffith, J.D. & Attardi, G. *Cell* **123**, 1227–1240 (2005).
2. Hyvarinen, A.K. *et al. Nucleic Acids Res.* **35**, 6458–6474 (2007).
3. Fernandez-Silva, P., Martinez-Azorin, F., Micol, V. & Attardi, G. *EMBO J.* **16**, 1066–1079 (1997).
4. Linder, T. *et al. Curr. Genet.* **48**, 265–269 (2005).
5. Roberti, M., Bruni, F., Polosa, P.L., Gadaleta, M.N. & Cantatore, P. *Nucleic Acids Res.* **34**, 2109–2116 (2006).
6. Roberti, M. *et al. Biochim. Biophys. Acta* **1787**, 303–311 (2009).
7. Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. *Cell* **94**, 193–204 (1998).
8. Huber, A.H. & Weis, W.I. *Cell* **105**, 391–402 (2001).
9. Wang, X., McLachlan, J., Zamore, P.D. & Hall, T.M. *Cell* **110**, 501–512 (2002).
10. Sibanda, B.L., Chirgadze, D.Y. & Blundell, T.L. *Nature* **463**, 118–121 (2010).
11. Bailey, S. *et al. J. Biol. Chem.* **278**, 15304–15312 (2003).
12. Kazmirski, S.L., Podobnik, M., Weitze, T.F., O'Donnell, M. & Kuriyan, J. *Proc. Natl. Acad. Sci. USA* **101**, 16750–16755 (2004).
13. Kruse, B., Narasimhan, N. & Attardi, G. *Cell* **58**, 391–397 (1989).
14. Shang, J. & Clayton, D.A. *J. Biol. Chem.* **269**, 29112–29120 (1994).
15. Wu, H.M. & Crothers, D.M. *Nature* **308**, 509–513 (1984).
16. Smid, A., Finsterer, M. & Grummt, I. *J. Mol. Biol.* **227**, 635–647 (1992).