Solution structure of the DNA-binding domain of human telomeric protein, hTRF1

Tadateru Nishikawa[1], Aritaka Nagadoi[1], Shoko Yoshimura[2], Saburo Aimoto[2] and Yoshifumi Nishimura[1]*

Background: Mammalian telomeres consist of long tandem arrays of the double-stranded TTAGGG sequence motif packaged by a telomere repeat binding factor, TRF1. The DNA-binding domain of TRF1 shows sequence homology to each of three tandem repeats of the DNA-binding domain of the transcriptional activator c-Myb. The isolated c-Myb-like domain of human TRF1 (hTRF1) binds specifically to telomeric DNA as a monomer, in a similar manner to that of homeodomains. So far, the only three-dimensional structure of a telomeric protein to be determined is that of a yeast telomeric protein, Rap1p. The DNA-binding domain of Rap1p contains two subdomains that are structurally closely related to c-Myb repeats. We set out to determine the solution structure of the DNA-binding domain of hTRF1 in order to establish its mode of DNA binding.

Results: The solution structure of the DNA-binding domain of hTRF1 has been determined and shown to comprise three helices. The architecture of the three helices is very similar to that of each Rap1p subdomain and also to that of each c-Myb repeat. The second and third helix form a helix-turn-helix (HTH) variant. The length of the third helix of hTRF1 is similar to that of the second subdomain of Rap1p.

Conclusions: The hTRF1 DNA-binding domain is likely to bind to DNA in a similar manner to that of the second subdomain of Rap1p. On the basis of the Rap1p–DNA complex, a model of the hTRF1 DNA-binding domain in complex with human telomeric DNA was constructed. In addition to DNA recognition by the HTH variant, a flexible N-terminal arm of hTRF1 is likely to interact with DNA.

Introduction

Telomeres are the protein–DNA complexes that protect the ends of eukaryotic linear chromosomes from degradation and fusion [1]. Mammalian telomeres are composed of long tandem arrays of the double-stranded telomeric repeat, TTAGGG [1–5], packaged by the telomeric protein, telomeric repeat binding factor (TRF1) [6–11]. The human TRF1 (hTRF1) consists of 439 amino acids comprising three presumed functional domains: an N-terminal acidic domain, a central TRF-specific dimerization domain, and a C-terminal DNA-binding domain [7,10,11]. The DNA-binding domain of hTRF1 consists of a 53 amino acid sequence that shows sequence homology to each of three tandem repeats of the c-Myb DNA-binding domain [12,13]. The c-Myb protein is a transcriptional activator that binds to the consensus sequence TAACNG [14–18]. Each of the three repeats in c-Myb has a very similar tertiary structure containing three helices [19–21]; the second and third helix in each repeat form a helix-turn-helix (HTH) variant motif [22–24].

The length of telomeres seems to be related to cell division number; after cell development in somatic cells the length of telomeres will be shortened during the course of cell division, whereas in oncogenic cells the length of telomeres remains constant [25–29]. Recently, it was found that the length of telomeres in oncogenic cells is maintained by an enzyme — the telomerase that adds telomeric DNA sequences [30,31]. The access of the telomerase to DNA seems to be controlled by TRF1 [32]. To understand the mechanism of chromosome stability and cell senescence it will be essential to know the tertiary structure of telomeres.

So far, only the structure of the yeast Rap1p DNA-binding domain in complex with telomeric DNA has been determined [33]. Rap1p is a constituent of the telomeres of the budding yeast *Saccharomyces cerevisiae*, which consist of 300–450 base pairs of an irregularly repeated sequence motif, $C_{2–3}A(CA)_{1–6}$, in contrast to the vertebrate regularly repeated sequence [34–37]. The DNA-binding domain of Rap1p contains two subdomains that bind DNA in a tandem orientation. The structure shows that despite lacking any significant sequence similarity, the Rap1p subdomains are structurally closely related to the c-Myb DNA-binding repeats [33]. In contrast, the hTRF1 DNA-binding domain

contains only a single c-Myb repeat homologous domain. Recently it has been shown that the isolated c-Myb-like domain of hTRF1 binds specifically, and with significant affinity, to telomeric DNA as a monomer [38]. It has been shown that the isolated c-Myb-like domain recognizes a binding site centered on the sequence GGGTTA, and that its DNA-binding mode is similar to that of each of the Rap1p subdomains [38].

We report here the solution structure of the c-Myb-like domain of hTRF1 determined by nuclear magnetic resonance (NMR) spectroscopy. The structure consists of three helices. The architecture of the three helices is very similar to that of each Rap1p subdomain and also to that of each c-Myb repeat. The length of the third helix of the hTRF1 DNA-binding domain is most similar to that of the second subdomain of Rap1p. The hTRF1 DNA-binding domain is quite likely to bind to DNA in a similar manner to the second subdomain of Rap1p.

## Results and discussion
Structure determination
The DNA-binding domain of hTRF1, amino acids 378–430 (see Figure 1), was chemically synthesized and subjected to the usual two-dimensional proton NMR experiments, double quantum filtered correlation spectroscopy (DQF-COSY) [39], nuclear Overhauser effect spectroscopy (NOESY) [40] and total correlation spectroscopy (TOCSY) [41], as reported previously [19,21]. Sequential and short-range nuclear Overhauser effect (NOE) connectivities are shown in Figure 2. These NOE patterns clearly indicate

that the DNA-binding domain of hTRF1 contains three helical regions as found in each of three repeats of the c-Myb DNA-binding domain [21].

To determine the three-dimensional structure of the hTRF1 DNA-binding domain, a set of 546 distance constraints were obtained from NOESY spectra (eight intra-residue, 202 sequential, 182 short-range and 154 long-range). In addition, 27 ϕ angle constraints were obtained from a DQF-COSY spectrum. With these constraints we have carried out distance geometry calculations using four-dimensional simulated annealing [42], as reported previously [19–21].

The 18 calculated structures are shown superimposed in stereo in Figure 3a. The refined average structure, including amino acids forming the hydrophobic core, is shown in Figure 3b. None of the calculated structures showed violations of greater than 0.3 Å, for the distance constraints, or 5°, for the dihedral restraints. Figure 4a shows the root mean square deviation (rmsd) of the backbone and heavy atoms of each residue. The backbone conformations are well defined in our calculations except for the N and C termini which do not have enough NOEs, as shown in Figure 4b. The overall rmsds between the 18 individual structures and the mean coordinates are 0.64 ± 0.09 Å for the backbone atoms and 1.27 ± 0.12 Å for all heavy atoms, excluding the N and C termini. These and other relevant statistics are summarized in Table 1. The Ramachandran plot of the averaged minimized structure (see Figure 5) shows that of all nonglycine/proline residues (48 of the 51

Figure 1
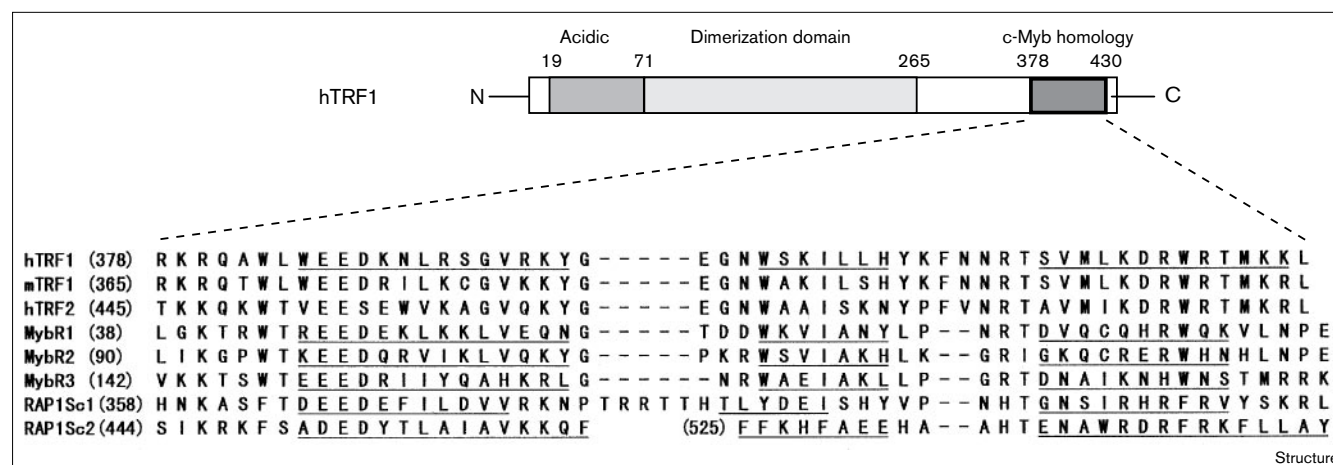
Amino acid sequence of the DNA-binding region of hTRF1. hTRF1 contains three domains: an N-terminal 6 kDa acidic domain; a 30 kDa TRF-specific dimerization domain; and a C-terminal 6 kDa DNA-binding domain. For comparison, the amino acid sequences of murine TRF1, human TRF2, the first (R1), second (R2) and third (R3) repeats of c-Myb, and Rap1p subdomains 1 (Sc1) and 2 (Sc2) were aligned. The three underlined regions of each sequence indicate the helical regions of each domain. Conserved residues are shown in bold. Numbers in brackets at the beginning of the sequences indicate amino acid positions (for Rap1p subdomain 2, it should be noted that only part of the amino acid sequence is shown).

Figure 2

Summary of the sequential and short-range NOEs observed for the DNA-binding domain of hTRF1. The thick, medium and thin bars correspond to strong, medium and weak NOEs, respectively. Helical regions are shaded. The corresponding amino acid sequence of hTRF1 is given at the top of the figure together with sequence numbering.
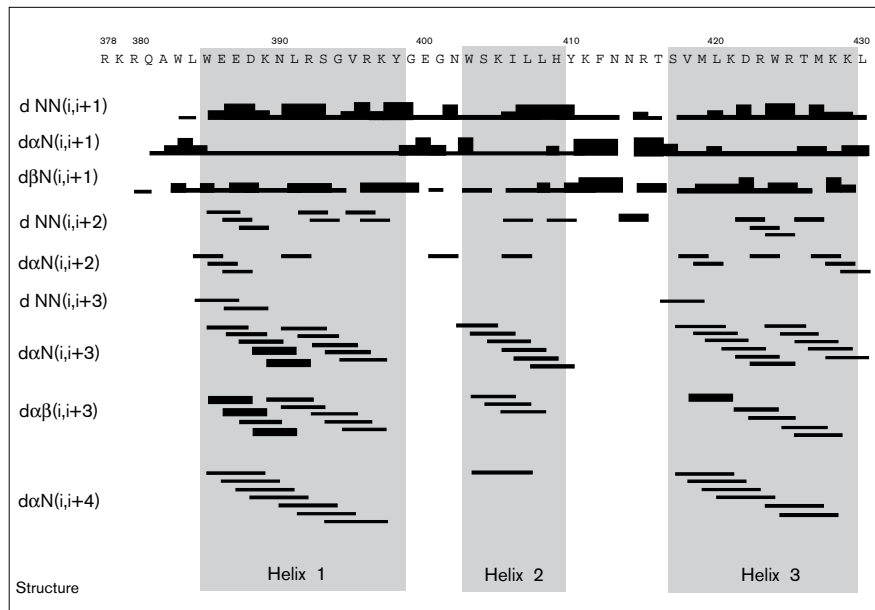

Figure 3

Stereoviews of the hTRF1 DNA-binding domain structures. Backbone atoms and sidechain atoms are shown in yellow and red, respectively; the N and C termini are marked. **(a)** The best-fit superposition of the 18 structures; three tryptophans and one phenylalanine residue form the hydrophobic core. **(b)** The refined average structure again showing the amino acids that form the hydrophobic core.
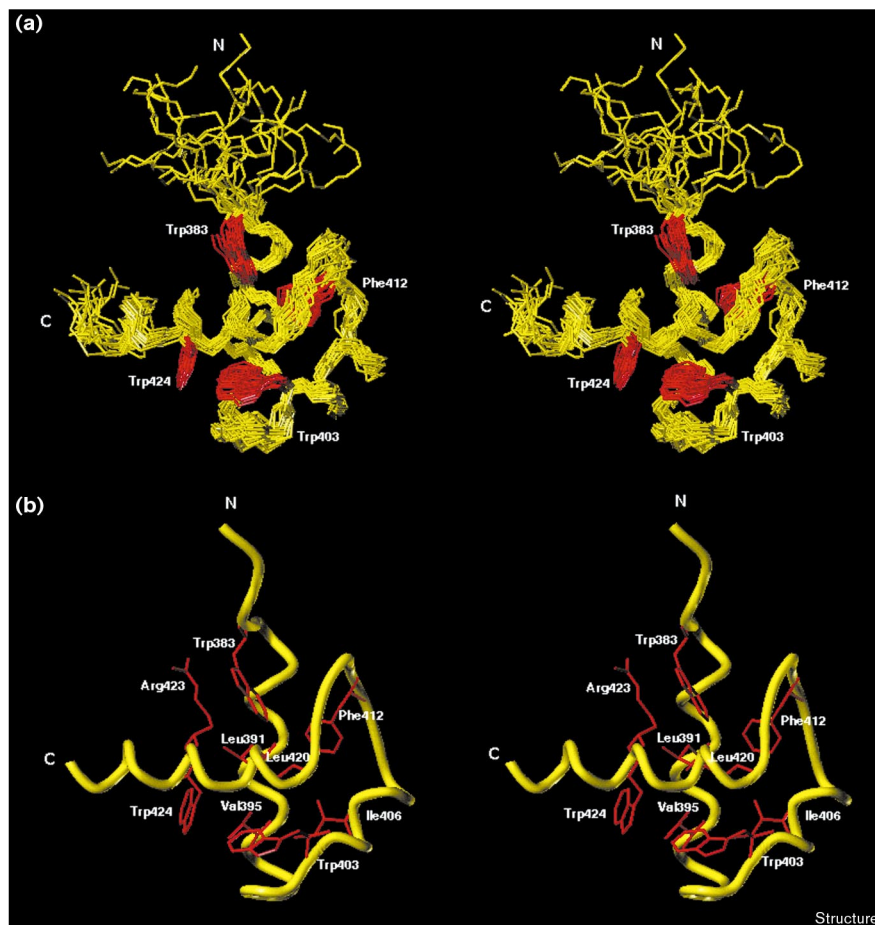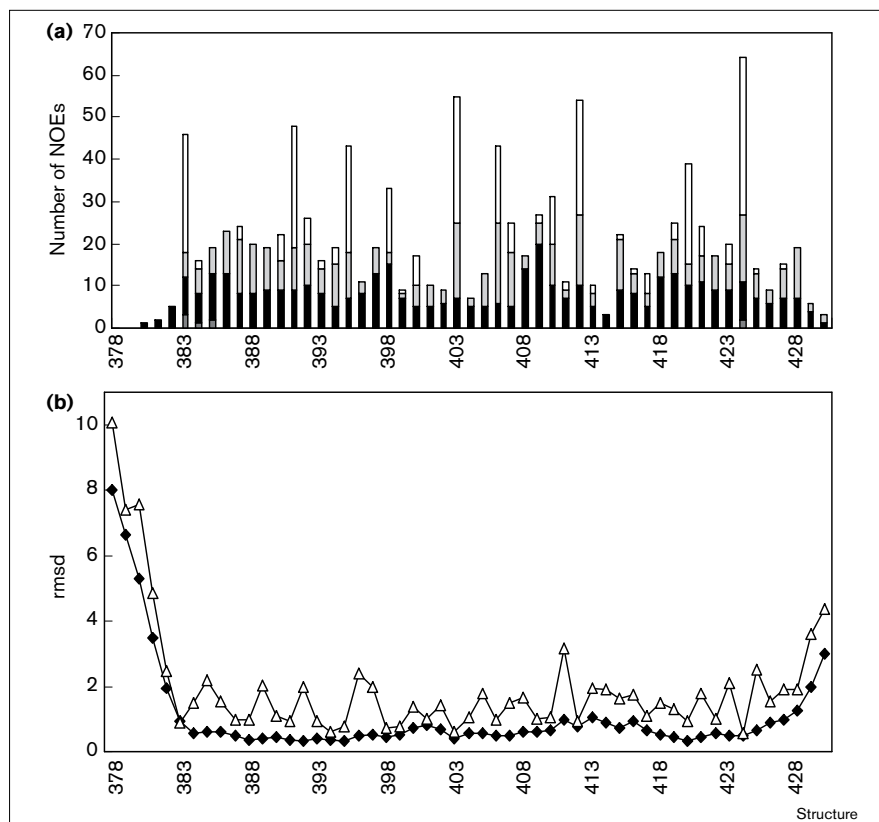
Figure 4

Summary of structural data versus residue number. **(a)** The distribution of distance constraints is shown for each residue by bars representing the number of intraresidue (dark gray bar), sequential (solid bar), medium-range (gray bar), and long-range (open bar) NOEs, cumulatively stacked. **(b)** The root mean square deviations (rmsds) between coordinates for the individual and the mean structures are represented at each residue for the backbone atoms (as a filled diamond) and for the heavy atoms (as an open triangle).

residues) 40 residues are in the most favored regions (A, B and L) and eight residues are in the additional allowed regions (a, b, l and p). The plot was generated using the program PROCHECK [43].

The structure consists of three helical regions, helix 1 (Trp 385–Tyr 398), helix 2 (Trp 403–His 409) and helix 3 (Ser 417–Lys 429), together with two turns, turn 1 (Gly 399–Asn 402) and turn 2 (Tyr 410–Thr 416). All three helices and both turns are well defined. The three helices are maintained by the hydrophobic core formed by residues Trp383, Leu391, Val395, Trp403, Ile406, Phe412, Leu420, Arg423 and Trp424 as shown in Figure 3b.

Helices 2 and 3 form an HTH variant motif containing a turn three amino acids longer than the corresponding turn in the prototypic HTH proteins [22–24]. This relatively long turn (turn 2) is stabilized by Phe412, located at the center of the turn, which protrudes into the hydrophobic core formed by Leu391 from helix 1, Ile406 from helix 2 and Leu420 from helix 3. The stabilized long turn seems to have a role in exposing two presumably functional asparagine residues, Asn413 and Asn414, into the solvent. In addition, the turn causes Lys411 to also be exposed to solvent but in the opposite direction to that of Asn414.

Comparison with each of the c-Myb repeats
Figure 6 shows a superposition of the structure of the hTRF1 DNA-binding domain obtained here, with the structure of the third repeat of c-Myb [21]; several side-chains in the hydrophobic core are illustrated. The back-bone architecture of the three helices and the conformations of three conserved tryptophan residues are similar in the hTRF1 DNA-binding domain and the c-Myb repeat. The rmsds between the backbone atoms in the refined average structure of the hTRF1 DNA-binding domain (amino acids 383–398, 403–409 and 417–426) and the corresponding atoms of each of the first, second and third repeats of c-Myb are 1.30 Å, 1.22 Å and 1.32 Å, respectively. On the other hand, turn 2 of hTRF1 is two amino acids longer than the corresponding turn in each of the c-Myb repeats.

Helix 3 of hTRF1 is slightly longer than the corresponding third helix in each of the c-Myb repeats. In a specific DNA complex of the minimal DNA-binding domain of c-Myb, consisting of two tandem repeats, two DNA-binding helices from the two repeats are closely packed in the major groove of DNA recognizing a specific base sequence in a cooperative manner [20]. The length of helix 3 of hTRF1, which is longer than that of c-Myb, suggests that it is unlikely to interact with DNA in a similar manner to c-Myb.

Table 1

Structural statistics and root mean square deviations for 18 NMR structures of the hTRF1 DNA-binding domain.

Structural restraints*

Distance restraints	
total	546
intraresidue	8
sequential	202
short-range |i−j|≤5	182
long-range |i−j|>5	154
Dihedral angle restraints (backbone φ)	27

Statistics for structure calculations

|	
Rmsds from idealized geometry	
bonds (Å) <SA> / <SA>r[†]	0.005 ± 0.000 / 0.005
angles (°) <SA> / <SA>r[†]	0.640 ± 0.019 / 0.594
impropers (°) <SA> / <SA>r[†]	0.332 ± 0.042 / 0.293
Rmsds of input restraint violations	
distances (Å) <SA> / <SA>r[†]	0.005 ± 0.003 / 0.002
dihedral angles (°) <SA> / <SA>r[†]	0.163 ± 0.286 / 0.000
Atomic rmsds (Å)	
residues 383–424	
backbone	0.64 ± 0.09
all heavy atoms	1.27 ± 0.12

*None of the structures have distance violations ≥ 0.3 Å or dihedral angle violations ≥ 5°. [†]<SA> represents the 18 NMR-derived structures; <SA>r represents the restrained energy minimized average structure from the 18 structures. For <SA>, the root mean square deviation (rmsd) is the average rmsd and the standard deviation for the 18 structures.

Comparison with the Rap1p DNA-binding domain

Figure 7 shows a superposition of the structures of the hTRF1 DNA-binding domain and Rap1p subdomain 2 [33]; several sidechains of the hydrophobic core residues are also shown. The backbone architectures of the three helices are very similar between the hTRF1 DNA-binding domain and Rap1p subdomains. The rmsds between the backbone atoms in the refined average structure of the hTRF1 DNA-binding domain (amino acids 383–398, 403–409 and 417–426) and the corresponding atoms of Rap1p subdomains 1 and 2 are 2.42 Å and 1.56 Å, respectively. In contrast, the structures of turns 1 and 2 of hTRF1 are different from the corresponding regions of each of the Rap1p subdomains. Turn 1 of hTRF1 consists of four amino acids, whereas between the first and second helix the Rap1p subdomain 1 has a nine amino acid loop and subdomain 2 has an extensively longer and partly disordered loop with 59 amino acids (not shown in Figure 7). Turn 2 of hTRF1 is two amino acids longer than the corresponding turn of each Rap1p subdomain.

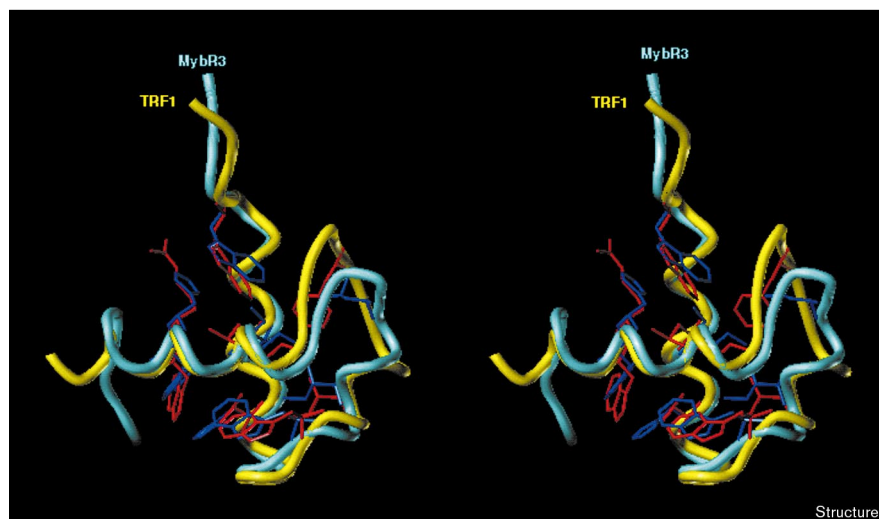## Model structure of the hTRF1 DNA-binding domain in complex with telomeric DNA

As stated previously, the DNA-recognition modes of Rap1p and c-Myb are different. In Rap1p, the two DNA-recognition helices from subdomains 1 and 2 bind spaced apart from each other in the major groove of DNA, each recognizing a separate ACACC sequence [33]. In contrast,

Figure 5
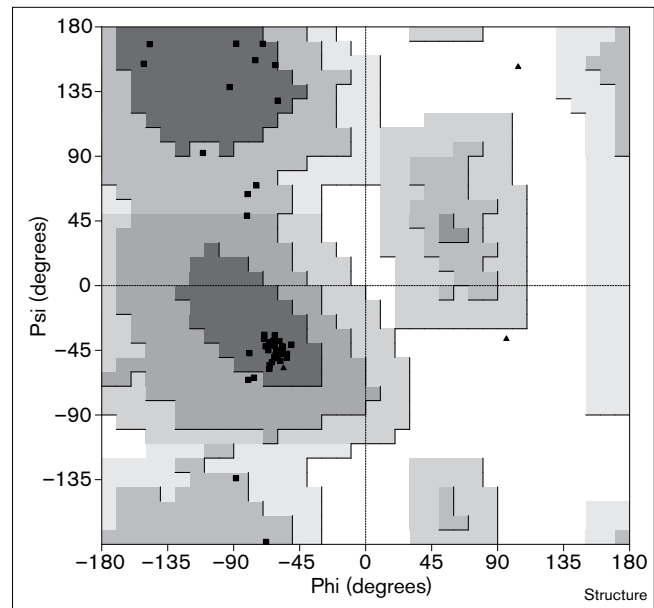
Ramachandran plot of φ and ψ dihedral angles for the averaged structure of the hTRF1 DNA-binding domain. The plot was generated using the program PROCHECK [43]. The dark gray, medium gray, light gray and white areas indicate most favored, additional allowed, generously allowed, and disallowed regions, respectively. No residues were observed in the latter two regions, except for glycine residues. Glycine residues are shown as triangles and all other residues as squares.

the two DNA-recognition helices from the two repeats of c-Myb are closely packed in the major groove of DNA recognizing a single TAACNG sequence cooperatively [20]. The structure of the hTRF1 DNA-binding domain has shown that the presumed recognition helix of hTRF1 is a little longer than that of each c-Myb repeat, rather similar to that of Rap1p subdomain 2. In addition, the amino acids used for direct interaction between the two recognition helices in c-Myb could not be identified in hTRF1. It would seem impossible to pack two hTRF1 DNA-binding domains closely in the major groove of DNA as found in the c-Myb–DNA complex. Instead, the hTRF1 DNA-binding domain could be arranged on the DNA in a similar manner to each of the Rap1p subdomains. Recently, by using gel retardation, primer extension and DNase I footprinting analyses, it has been shown that the isolated c-Myb-like motif of hTRF1 recognizes a binding site centered on the sequence GGGTTA and that its binding mode is very similar to that of engrailed homeodomain [38]. It has been also suggested that in addition to the recognition of the HTH motif, a flexible N-terminal arm of hTRF1 is likely to interact with DNA in the minor groove [38]. Taking these results into consideration, we have built a model of the hTRF1 DNA-binding domain in a complex with
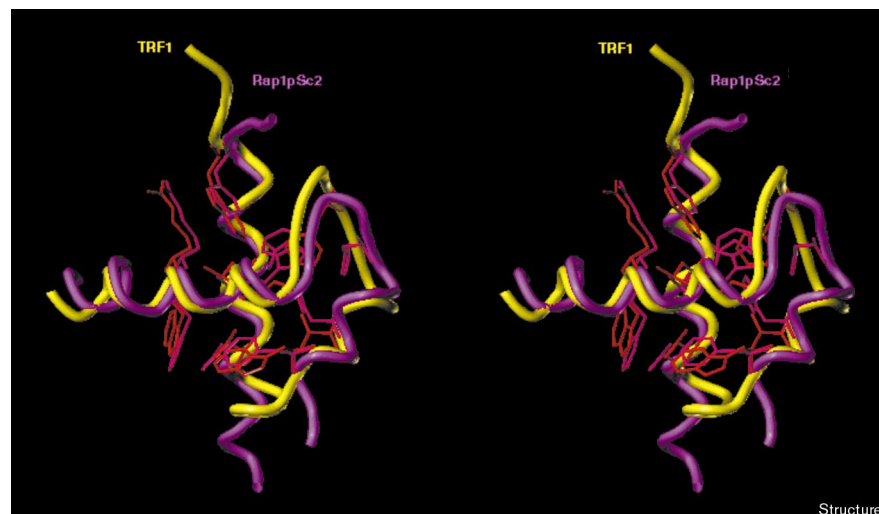
Figure 6

Stereoview best-fit superposition of the hTRF1 DNA-binding domain and the third c-Myb repeat (c-MybR3). Amino acid sidechains forming the hydrophobic core are indicated in stick representation. The backbone and sidechain atoms of the hTRF1 DNA-binding domain are shown in yellow and red, respectively, those for c-Myb R3 are shown in cyan and blue.

telomeric DNA (shown in Figure 8). The model is constructed from the crystalline structure of the Rap1p DNA-binding domain in a complex with DNA [33]. The moiety of subdomain 2 in the complex was replaced by the average structure of the hTRF1 DNA-binding domain by superimposing its three helices onto the corresponding helices of Rap1p subdomain 2; the yeast telomeric DNA sequence was changed into the human telomeric sequence by superimposing the common GGT sequence in both [2]. No attempt was made to optimize the interaction between the protein and DNA. On the basis of the Rap1p structure, in the flexible N-terminal region of hTRF1, a lysine residue at position 379 or an arginine residue at position 380 is likely to interact with

an adenine in the minor groove, in a similar manner to that seen for Lys360 and Lys446 in Rap1p [33,38]. The presence of an N-terminal arm in hTRF1 suggests a similar binding mode to homeodomains [44] and Rap1p subdomains [33]. Asn414 from turn 2 is likely to interact with the DNA phosphate backbone. In the major groove of DNA, Val418, Asp422, Arg425 and Lys421 from the third helix are likely to interact with thymine, adenine/cytosine, guanine and guanine bases, respectively. This model is consistent with that of Konig *et al.* [38]. As each Rap1p subdomain recognizes the ACACC sequence, the hTRF1 DNA-binding domain alone could recognize the TAACC sequence [38]; in the model the repeated TTAGGG sequence is treated as a repeated TAACCC sequence.

Figure 7

Stereoview best-fit superposition of the hTRF1 DNA-binding domain and Rap1p subdomain 2 (Rap1pSc2). Amino acid sidechains forming the hydrophobic core are indicated in stick representation. The backbone and sidechain atoms of the hTRF1 DNA-binding domain are shown in yellow and red, respectively, those for subdomain 2 are shown in purple and magenta.

Figure 8
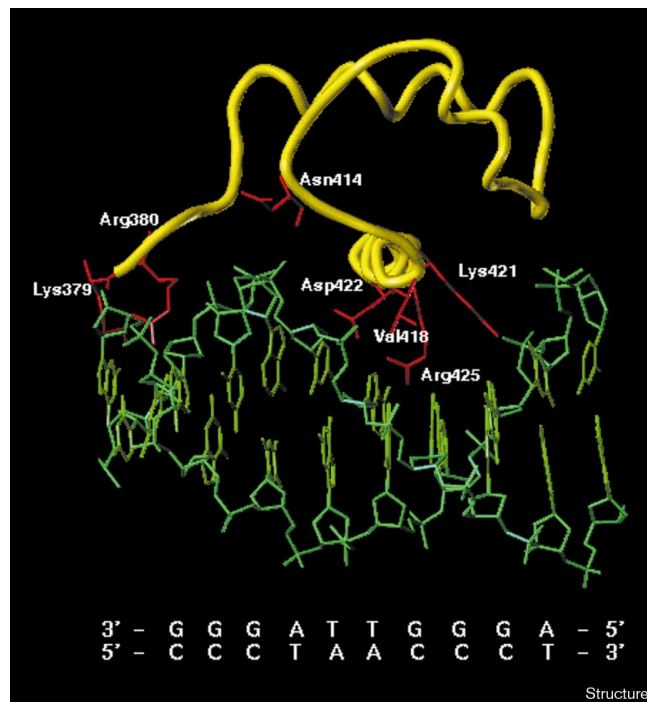
A model structure of the hTRF1 DNA-binding domain in complex with DNA. The backbone and sidechain atoms of the hTRF1 DNA-binding domain are shown in yellow and red, respectively; the backbone and base atoms of telomeric DNA are shown in green. Residues involved in key interactions with the DNA are labeled. The sequence of the telomeric DNA is shown below the model.

The precise details of the interaction between hTRF1 and DNA require further structural analysis, however.

Biological implications

Telomeres, the protein–DNA complexes that protect the ends of eukaryotic linear chromosomes, consist of tandem repeats of guanine-rich sequence motifs packaged by specific DNA-binding proteins. In humans, the double-stranded TTAGGG repeats are bound by human telomere repeat binding factor (hTRF1). hTRF1 contains three domains: an N-terminal acidic domain; a central TRF-specific dimerization domain; and a C-terminal DNA-binding domain that shows sequence homology to each of three tandem repeats of the transcriptional activator c-Myb. Recently, it was shown that the isolated c-Myb-like domain of hTRF1 can specifically recognize one GGGTTA sequence, in a similar fashion to that of homeodomains [38]. On the other hand, at least two repeats are essential for the sequence-specific DNA-binding of c-Myb. Similarly, the specific DNA-binding domain of the yeast telomeric protein Rap1p contains two subdomains that are structurally closely related to c-Myb repeats.

Here, we report the solution structure of the hTRF1 DNA-binding domain. The structure comprises three helices. The architecture of the three helices is very similar to that of each Rap1p subdomain and also to that of each c-Myb repeat. The second and third helix form a helix-turn-helix (HTH) variant containing a longer turn than the corresponding turn in prototypic HTH proteins, c-Myb repeats and Rap1p subdomains. The length of the third helix of hTRF1 is rather similar to that of the second subdomain of Rap1p. Although the hTRF1 DNA-binding domain shows strong sequence similarity to each of the c-Myb repeats, and no apparent sequence homology to the Rap1p DNA-binding domain, the DNA-binding mode of hTRF1 is more closely related to that of the second subdomain of Rap1p [33] than that of c-Myb repeats [20]. However, two subdomains of Rap1p are essential for sequence-specific DNA binding, whereas in hTRF1 a single c-Myb-like domain is sufficient [38]. On the basis of the present structure, a model of the hTRF1 DNA-binding domain in complex with human telomeric DNA has been constructed. The specific binding by the single HTH variant of hTRF1 could be aided by the flexible N-terminal arm, similar to homeodomains [38], and by the slightly longer turn in the HTH variant which is likely to interact with phosphate backbone.

The dimerization domain of hTRF1 is located within the central region of the protein. The dimerization domain does not seem to be essential for the sequence-specific binding of hTRF1 [38]. *In vivo*, hTRF1 seems to exist as a dimer [10]; a dimer of hTRF1 contains two specific DNA-binding domains. Both DNA-binding domains complexed with GGGTTA sequence motifs separately are organized into a particular structural unit of telomeres. Such specific tertiary structures of telomeres may participate in the regulation of telomerase activity. Telomerase adds TTAGGG repeated sequence motifs to telomeres and the activity of this enzyme has been implicated in oncogenic cells. By increasing our understanding of telomeric structure and chromosome stability, it may be possible in the future to learn more about the mechanisms of carcinogenesis.

Materials and methods

Sample preparation

The peptide fragment of the hTRF1 DNA-binding domain was chemically synthesized and was purified as described previously [19,21]. The lyophilized sample was dissolved in 100 mM potassium phosphate buffer (KPB) pH 6.8 and 1 mM NaN_3 and the sample concentration was 1.6 mM.

NMR spectroscopy

NMR spectra were recorded at 600 MHz or 500 MHz on a Bruker DMX-600 or AMX2-500 spectrometer. The temperature during data acquisition was set to 300K. Quadrature detection was made by the TPPI method. For water signal suppression, weak presaturation was used. DQF-COSY spectra [39], NOESY spectra [40] with mixing times of 50, 150 and 200 ms, and TOCSY spectra [41] with mixing times of 80 and 100 ms were recorded and relaxation-compensated DIPSI-2 mixing schemes with

z filtration [45] were used in the pulse sequence of the TOCSY experiments. In 90% H_2O/10% D_2O and in D_2O, 600 t_1 increments (zero-filled to 1024 data points) each of 2048 real data points were recorded with a spectral width of 8064 Hz. In addition a DQF-COSY spectrum with 1024 t_1 increments (zero-filled to 2048 data points) each of 4096 real data points (zero-filled to 8192 data points) was recorded in 90% H_2O/10% D_2O. Sequence-specific assignments could be completed from Arg380 to Leu430. The N-terminal two residues, Arg378 and Lys379, could not be identified in NOESY spectra, however.

Structure calculations

Interproton distance constraints were derived from the cross-peak intensities of the NOESY spectra with mixing times of 50, 150 and 200 ms, using similar assumptions of previous calculations [19–21]. From the NOE intensities, the distances between backbone protons were classified into four ranges, 1.9–3.0, 1.9–4.0, 1.9–5.0 and 1.9–6.0 Å, corresponding to strong, medium, weak and very weak NOEs, respectively. Similarly, the interproton distances involving sidechain protons were classified into four ranges, 1.9–4.0, 1.9–4.5, 1.9–5.5 and 1.9–6.5 Å, corresponding to strong, medium, weak and very weak NOEs.

In addition, the ϕ angles were restrained by estimating the $^3J_{HN\alpha}$ coupling constants from F2 high resolution DQF-COSY spectrum with corrections against deviated values due to broadened signals. The restrained angle ranges were as follows: $-90° < \phi < -40°$ for $^3J_{HN\alpha} <$ 6.5 Hz and $-160° < \phi < -80°$ for $^3J_{HN\alpha} > 8.5$ Hz.

Structures were constructed from random-coil conformations using the four-dimensional simulated-annealing program EMBOSS [42] with the experimentally derived distance and angle constraints, as reported previously [19–21]. All calculations were performed using the workstation Indigo 2 Impact (Silicon Graphics, Inc.).

Accession numbers

The coordinates of 18 calculated structures have been deposited with the Brookhaven Protein Data Bank. The PDB ID code is 1BA5.

Acknowledgements

References

1. Zakian, V.A. (1995). Telomeres: beginning to understand the end. *Science* **270**, 1601-1607.
2. Konig, P. & Rhodes, D. (1997). Recognition of telomeric DNA. *Trends Biochem. Sci.* **22**, 43-47.
3. Moyzis, R.K., *et al.*, & Wu, J.-R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc. Natl Acad. Sci. USA* **85**, 6622-6626.
4. de Lange, T., *et al.*, & Varmus, H.E. (1990). Structure and variability of human chromosome ends. *Mol. Cell. Biol.* **10**, 518-527.
5. Hanish, J.P., Yanowitz, J.L. & de Lange, T. (1994). Stringent sequence requirements for the formation of human telomeres. *Proc. Natl Acad. Sci. USA.* **91**, 8861-8865.
6. Zohng, Z., Shieu, L., Kaplan, S. & de Lange, T. (1992). A mammalian factor that binds telomeric TTAGGG repeats *in vitro. Mol. Cell. Biol.* **12**, 4834-4843.
7. Chong, L., *et al*., & de Lange, T. (1995). A human telomeric protein. *Science* **270**, 1663-1666.
8. Bilaud, T., *et al*., & 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.
9. Broccoli, D., *et al*., & de Lange, T. (1997). Comparison of the human and mouse gene encoding the telomeric protein, TRF1: chromosomal localization, expression and conserved protein domains. *Human Mol. Genet.* **6**, 69-76.
10. Bianchi, A., Smith, S., Chong, L., Elias, P. & de Lange, T. (1997). TRF1 is a dimer and bends telomeric DNA. *EMBO J.* **16**, 1785-1794.
11. Smith, S. & de Lange, T. (1997). TRF1, a mammalian telomeric protein. *Trends Genet.* **13**, 21-26.
12. Gonda, T.J., Gough, N.M., Dunn, A.R. & de Biaquire, J. (1985). Nucleotide sequence of cDNA clones of the murine myb protooncogene. *EMBO J.* **4**, 2003-2008.
13. Klempnauer, K.-H. & Sippel, A.E. (1987). The highly conserved amino-terminal region of the protein encoded by v-myb oncogene function as a DNA-binding domain. *EMBO J.* **6**, 2719-2725.
14. Biedenkapp, H., Borgmeyer, U. Sippel. A. & Klempnauer, K.-H. (1989). Viral myb oncogene encodes a sequence-specific DNA binding activity. *Nature* **335**, 835-837.
15. Weston, K. & Bishop, J.M. (1989). Transcriptional activation by the v-myb oncogene and its cellular progenitor, c-myb. *Cell* **58**, 85-93.
16. Ness, S.A., Marknell, A. & Graf, T. (1989). The v-myb oncogene product binds to and activates promyelocyte specific mim-1 gene. *Cell* **59**, 1115-1125.
17. Nakagoshi, H., Nagase, T., Kanei-Ishii, C., Ueno, Y. & Ishii, S. (1990). Binding of the c-myb protooncogene product to the simian virus 40 enhancer stimulates transcription. *J. Biol. Chem.* **265**, 3479-3483.
18. Tanikawa, J., *et al*., & Sarai, A. (1993). Recognition of specific DNA sequences by the c-myb protooncogene product: role of three repeat units in the DNA-binding domain. *Proc. Natl Acad. Sci. USA* **90**, 9320-9324.
19. Ogata, K., *et al*., & Nishimura, Y. (1992). Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl Acad. Sci. USA* **89**, 6428-6432.
20. Ogata, K., *et al*., & Nishimura, Y. (1994). Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell* **79**, 639-648.
21. Ogata, K., *et al*., & Nishimura, Y. (1995). Comparison of the free and DNA-complexed forms of the DNA-binding domain from c-Myb. *Nat. Struct. Biol.* **2**, 309-320.
22. Brennan, R.G. & Matthews, B.W. (1989). The helix-turn-helix DNA binding motif. *J. Biol. Chem.* **264**, 1903-1906.
23. Harrison, S.C. & Aggarwal, A.K. (1990). DNA recognition by proteins with the helix-turn-helix motif. *Annu. Rev. Biochem.* **59**, 933-969.
24. Pabo, C.O. & Sauer, R.T. (1992). Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61**, 1053-1095.
25. Counter, C.M., *et al*., & Bacchetti, S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* **11**, 1921-1929.
26. Counter, C.M., Hirte, H.W., Bacchetti, S. & Harley, C.B. (1994). Telomerase activity in human ovarian carcinoma. *Proc. Natl Acad. Sci. USA* **91**, 2900-2904.
27. Broccori, D., Godley, L.A., Donenhower, L.A., Varmus, H.A. & de Lange, T. (1996). Telomerase activation in mouse mammary tumors: lack of detectable telomere shortening and evidence for telomerase RNA with cell proliferation. *Mol. Cell. Biol.* **16**, 3763-3772.
28. Autexier, C. & Greide, C.W. (1996). Telomerase and cancer. *Trends Biochem. Sci.* **21**, 387-391.
29. Zakian, V. (1997). Life and cancer without telomerase. *Cell* **91**,1-3.
30. Blasco, M.A., *et al*., & Greider, G.W. (1997). Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell* **91**, 25-34.
30. Morin, G.B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* **59**, 521-529.
31. Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundbladt, V. & Cech, T.R. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276**, 561-567.
32. van Steensel, B. & de Lang, T. (1997). Control of telomere length by the human telomeric protein TRF1. *Nature* **385**, 740-743.
33. Konig, P., Girald, R., Chapman, L. & Rhodes, D. (1996). The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. *Cell* **85**, 125-136.
34. Conrad, M.N., Wright, J.H., Wolf, A.J. & Zakian, V.A. (1990). RAP1 protein interacts with yeast telomeres *in vivo*: overproduction alters telomere structure and decreases chromosome stability. *Cell* **63**, 739-750.
35. Shore, D. (1994). RAP1: a protean regulator in yeast. *Trends Genet.* **10**, 408-412.
36. Krauskopf, A. & Blackburn, A. (1996). Control of telomere growth by interactions of RAP1 with the most distal telomeric repeats. *Nature* **383**, 354-357.

37. Marcand, S. & Gilson, E. & Shore, D. (1997). A protein-counting mechanism for telomere length regulation in yeast. *Science* **275**, 986-990.
38. Konig, P., Fairall, L. & Rhodes, R. (1998). Sequence specific DNA recognition by the Myb-like domain of the human telomere binding protein TRF1: a model for the protein/DNA complex. *Nucleic Acids Res.* 1731-1740.
39. Rance, M., *et al.*, & Wuthrich, K. (1983). Improved spectral resolution in COSY ¹H NMR spectra of proteins via double quantum filtering. *Biochim. Biophys. Res. Commun.* **171**, 479-485.
40. Macura, S. & Ernst, R.R. (1980). Elucidation of cross relaxation in liquids by two-dimensional NMR spectroscopy. *Mol. Phys.* **41**, 95-117.
41. Bax, A. & Davis, D.C. (1985). MLEV-17 based two dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.* **65**, 355-360.
42. Nakai, T., Kidera, A. & Nakamura, H. (1993). Intrinsic nature of the three dimensional structure of proteins as determined by distance geometry with good sampling properties. *J. Biomol. NMR* **3**, 19-40.
43. Lagkowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R. & Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structure solved by NMR. *J. Biomol. NMR* **8**, 477-486.
44. Kissinger, C.R., Liu, B., Martin-Blanco, E., Kornberg, T.B. & Pabo, C.O. (1990). Crystal structure of an engrailed homeodomain–DNA complex at 2.8 Å resolution. *Cell* **63**, 579-590.
45. Griesinger, C., Otting, G., Wuthrich, K. & Ernst, R.R. (1988). Clean TOCSY for ¹H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110**, 7870-7872.