DNA–protein interactions at the telomeric repeats of *Schizosaccharomyces pombe*

Margaret Duffy and Alistair Chambers*

Department of Genetics, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH, UK

ABSTRACT

Gel retardation assays using a probe containing the repeat region of a *Schizosaccharomyces pombe* chromosomal telomere identified four specific DNA–protein complexes in *S.pombe* total protein extracts (I, I′, IIa and IIb). The proteins responsible for these complexes bound to the telomeric repeat region irrespective of whether or not the repeats were in close proximity to the end of a DNA molecule, and none of them bound strongly to single-stranded DNA. The protein responsible for complex I (TeRF I) was separated from the activity responsible for complexes IIa and IIb (TeRF II) using heparin–Sepharose chromatography. Both factors were efficiently cross-competed by an oligonucleotide containing the 18 bp sequence 5′-GGTT-ACAGGTTACAGGTT-3′, which corresponds to two complete telomeric repeat units. Mutation of the T residues at positions 4 and 11 in the oligonucleotide dramatically reduced binding by TeRF II, but had no affect on binding by TeRF I. The protein responsible for complex I′ did not bind strongly to either the wild-type or mutant oligonucleotide.

INTRODUCTION

Telomeres are the physical ends of eukaryotic chromosomes and they have a number of important roles within the cell. These roles include preventing chromosome fusion, protecting the chromosomes from exonucleolytic attack and facilitating the complete replication of chromosomes (1,2). In most eukaryotes, with the notable exception of *Drosophila*, telomeres are made up of multiple copies of a short (5–8 bp) repeat unit (3,4). This repeat is usually rich in T and G residues in the DNA strand which runs 5′→3′ towards the chromosome end. In ciliated protozoa this strand overlaps its partner strand at the very end of the chromosome to form a short 3′ extension (5,6). In the budding yeast, *Saccharomyces cerevisiae*, telomeres gain long single-stranded extensions during S phase of the cell cycle (7). The functions of telomeres are thought to be achieved via interactions between specific proteins and the telomeric repeat sequences, and by the formation of specific DNA structures at the chromosome ends. Both the single-stranded 3′ extension and the double-stranded repeat region are the targets for specific DNA binding proteins in a range of different organisms (8). In the ciliated protozoan *Oxytricha nova*, a dimeric protein containing α- and β-subunits

interacts with the single-stranded 3′ extension (9–11). In *Euplotes*, a monomeric protein with homology to the α-subunit of the *Oxytricha* protein, appears to have a similar role (12). The budding yeast, *S.cerevisiae*, has been used extensively as a model system to investigate telomere function. In yeast, telomeres consist of an average of 300 bp of a variable TG rich sequence, abbreviated as TG_{1-3} (3). Several different telomere binding proteins have been described. The best characterised of these is the multi-functional protein Rap1p (13). This protein binds to both the double-stranded repeat region and the single-stranded extension of yeast telomeres, and these interactions appear to be important for telomere function (8,14–19). Reduction in the amount of functional Rap1p in the yeast cell, by growing a *rap1^{ts}* strain at a semi-permissive temperature, resulted in a gradual reduction in telomere length. This could be reversed by returning the cells to a permissive temperature (19). Similarly, overexpression of Rap1p caused an increase in telomere length and an increase in the rate of chromosome loss and mitotic recombination (8). Immuno-fluorescence experiments have shown that the majority of Rap1p in yeast cells is associated with the telomeres and that this association requires the products of the *SIR3* and *SIR4* genes (20,21). *In vitro* experiments suggest that Rap1p may bind once per 18 bp of telomeric DNA, resulting in up to 25–30 binding sites per telomere (22). The presence of Rap1p at telomeres is particularly intriguing because this protein is also a transcription factor, involved both in activating and silencing transcription at a range of loci (23,24,13). It interacts with the UAS of many yeast housekeeping genes, including genes encoding ribosomal proteins, components of the translational machinery and glycolytic enzymes (25–32). Rap1p also interacts with the silencers at *HML* and *HMR* and plays a role in repressing the inactive mating type genes (30,33–36). Recently a *RAP1* gene was cloned from the closely related yeast *Kluyveromyces lactis* (37). The protein product of this gene binds a similar DNA sequence to budding yeast Rap1p and contains a conserved region which in the budding yeast gene encodes a domain of Rap1p involved in telomere function (37). It is not yet known if the *K.lactis* protein interacts with *K.lactis* telomeres. A second budding yeast gene, *TBF1*, encodes an essential TTAGGG repeat binding factor, with a molecular weight of 63 kDa (38). This protein binds to two TTAGGG sequences proximal to the TG_{1-3} repeat sequences, but its role in telomere function is unclear. Recently, three further genes were isolated from a yeast gene library based on their ability to make protein products which specifically interact with TG_{1-3} DNA *in vitro* (39). However, the role, if any, that these proteins play at telomeres *in vivo* is unclear.

---

* To whom correspondence should be addressed

Telomere binding factors have also been identified in higher eukaryotes. *Xenopus laevis* contains a well characterised activity called *Xenopus* telomere end factor (XTEF) which interacts with two repeats of the sequence TTAGGG in the single-stranded telomeric extension (40). Single strand TTAGGG binding factors have also been purified from mouse liver extracts (41). Activities have been identified in human cells which interact with both the double-stranded and single-stranded forms of the TTAGGG human telomeric repeat (42,43). The single-stranded binding factors also interact with RNA and are components of hnRNPs (42). The human double-stranded telomeric repeat binding factor (TRF) requires six copies of TTAGGG to form an efficient substrate and was initially estimated to be ~50 kDa in size (43). The gene encoding TRF has been isolated from a human cDNA library (44). It encodes a protein of 439 amino acids, with a predicted molecular weight of 50.3 kDa. The protein has a region at the N-terminus rich in aspartic and glutamic acid residues and a region close to the C-terminus containing an myb-like DNA binding motif (44).

Fission yeast, *Schizosaccharomyces pombe*, is distantly related to both budding yeasts and higher eukaryotes. A comparison between telomere organisation and function in fission yeast and other eukaryotic organisms may therefore provide important insights into the evolution and roles of these key structures. Fission yeast has a genome of about the same size as budding yeast, but a haploid chromosome number of only three. The chromosomal telomeres have been implicated in the specific movements chromosomes make as haploid nuclei fuse during karyogamy, and the movements the fused nucleus makes within the cell prior to meiosis (45). These movements appear to be mediated via attachment of the telomeres to the spindle pole body (45). Four *S.pombe* chromosomal telomeres have been cloned and sequenced (46). They are ~300 bp in length and are made up of a repeat unit of consensus $C_{1-8}G_{0-1}T_{0-2}GTA_{1-3}$ (46,47). Although this consensus accurately describes the sequences at fission yeast telomeres, it suggests that the repeat unit is very variable, and it relates to the sequence of the A and C rich DNA strand, rather than the T and G rich strand. To facilitate comparison with the repeats in other organisms, we have used the simpler consensus 5′-TTA-CAG$_{1-8}$-3′, which describes the majority of telomeric repeats in *S.pombe*. Although cloned *S.pombe* telomeres have been available for a number of years, no direct telomere binding proteins have been characterised. The only protein so far shown to be localised to *S.pombe* telomeres is the chromodomain protein Swi6p. This protein associates with both centromeres and telomeres in *S.pombe* and is required for proper centromere function (48). In this paper we describe the identification and initial characterisation of fission yeast protein factors which interact with the double-stranded repeat region of the telomeric DNA.

MATERIALS AND METHODS

Yeast strains and media

The haploid *S.pombe* wild-type strains *975h+* and *975h−* were used throughout. They were routinely grown using YEPD medium on plates and in liquid culture (49).

Plasmid construction

The starting plasmid for the telomere sub-clones was pNSU28 which contains 0.9 kb of *S.pombe* telomeric DNA in pUC19 (46). The 0.9 kb telomeric fragment was isolated from pNSU28 by digestion with *Eco*RI. The isolated *Eco*RI fragment was then cut with *Rsa*I and a 425 bp fragment containing the telomere repeat region plus telomere associated sequence was isolated. This fragment was ligated into the *Sma*I site of plasmid pSP56 to generate plasmid pAJ25 (50). pAJ25 was cut with *Eco*RI and *Bam*HI to release the cloned 425 bp fragment. This was then digested with *Hae*III to release a fragment containing 140 bp of telomeric repeat sequence plus telomere associated sequence. This was end-filled using Klenow polymerase and dNTPs to convert the *Bam*HI end to a blunt end. It was then cloned into the *Hin*cII site of pSP56 such that the extreme telomeric sequences were at the 5′ end of the polylinker and the telomere associated sequences were at the 3′ end. This generated plasmid pAJ34. The telomeric sub-fragment used in most retardation assays was isolated by digestion of pAJ34 with *Bam*HI and *Pst*I.

Protein extracts and gel retardation assays

Schizosaccharomyces pombe cultures were grown to mid-log phase in 50 ml YEPD medium. Cells were harvested and washed twice with 1 ml 25 mM NaPO$_4$ pH 7.5, then resuspended in 300 μl ice-cold 25 mM NaPO$_4$ pH 7.5, containing 1 mM PMSF. The cells were broken with glass beads by vortexing for 2 min then centrifuged briefly in a microfuge. The supernatant was collected and recentrifuged for 10 min at 4°C. The supernatant was again collected and the protein concentration determined using the Bradford assay (51). Typically 50 ml of cells yielded ~1 mg protein, of which 5 μg was used in each gel retardation assay. Retardation assays were performed at room temperature in a total reaction volume of 20 μl. The labelled DNA fragment for use in the assay was generated by end-labelling ~120 ng of isolated DNA fragment using [γ-^{32}P]ATP (>185 TBq/mmol; Amersham International plc) and T4 polynucleotide kinase (Life Technologies, Inc.). Protein extract (5 μg) was incubated with 2 μg poly(dI:dC) and 2 ng of the labelled DNA fragment in a binding buffer containing 5% glycerol, 1 mM EDTA, 10 mM β-mercaptoethanol, 25 mM Tris–HCl pH 7.5, 25 mM NaCl and 20 mM KCl. After incubation for 30 min at room temperature, DNA–protein complexes were separated by electrophoresis at 180V for ~90 min using 16 cm-long 5% polyacrylamide gels containing 0.5× TBE.

Competitor DNA fragments and oligonucleotides

Unlabelled competitor DNAs consisted of either gel isolated DNA fragments or annealed oligonucleotide pairs. The standard telomere competitor fragment was obtained by digestion of pAJ34 with *Bam*HI and *Pst*I and isolation of the 140 bp telomere fragment. A competitor fragment with telomeric repeats away from the ends was obtained by digestion of pAJ34 with *Sph*I and *Pvu*II and isolation of a 810 bp fragment. A control fragment for these experiments was isolated by digestion of plasmid pSP65 with the same enzymes and isolation of a 670 bp fragment.

The telomeric repeat oligonucleotides consisted of the sequences:

AC1 5′-GATCTCAGCT**GGTTACAGGTTACAGGTT**G-3′
AC2 5′-GATCC**AACCTGTAACCTGTAACC**AGCTGA-3′
AC3 5′-**GGTTACAGGGGGG**TT-3′
AC4 5′-AA**CCCCCCCTGTAACC**-3′
AC5 5′-**TTACAGGTTACAGG**-3′
AC6 5′-**CCTGTAACCTGTAA**-3′
AC7 5′-GATCTCAGCT**GGTGACAGGTGACAGGTT**G-3′
AC8 5′-GATCC**AACCTGTCACCTGTCACC**AGCTGA-3′

In each case the telomeric sequences are shown in bold type.

These were annealed in pairs to generate double-stranded oligo-nucleotides for use as competitor DNAs.

A control oligonucleotide used in some experiments consisted of the sequence:

AS1 5′-GATCCTAAATATAAAAA-3′

Heparin–Sepharose fractionation

Heparin–Sepharose (Pharmacia) was treated according to the manufacturer's instructions. A column was prepared at 4°C with a 2.5 ml bed volume of heparin–Sepharose in Z buffer (10 mM Tris pH 7.5, 1 mM EDTA, 10% glycerol, 100 mM KCl, 50 mM β-mercaptoethanol, 1 mM PMSF). Total protein extract (10–12 ml; 50 mg protein) was applied to the column using a BioRad EconoSystem. Proteins were eluted using increasing concentrations of KCl ranging from 100 mM to 1 M in Z buffer. Fractions of 0.5 ml were collected automatically and dialysed overnight at 4°C against a large volume of Z buffer. Five μl of each dialysed fraction was used in gel retardation assays.

DNase I footprinting

pAJ34 was digested with *Bgl*II and *Sma*I and the telomere DNA fragment was isolated. This was radioactively labelled at the *Bgl*II end by end filling using Klenow polymerase in the presence of [α-^{32}P]dCTP (Amersham). The end-labelled fragment was incubated in standard binding buffer with 10 μl of a fraction eluted from heparin–Sepharose at 400 mM and 3 μg poly(dI:dC) in a total volume of 50 μl. Binding was allowed to proceed for 30 min on ice before addition of diluted DNase I at room temperature. One μl of a 1:100 dilution of Promega DNase I was allowed to digest the DNA for 30 s and the reaction terminated by the addition of 100 μl stop buffer (50 mM Tris pH 8, 2% SDS, 10 mM EDTA pH 8, 0.4 mg/ml proteinase K, 100 μg/ml glycogen). The reaction was then incubated at 37°C for 30 min and 70°C for a further 2 min. It was then extracted once with phenol–chloroform and dried in a vacuum concentrator. The dried samples were resuspended in sequencing gel loading buffer and subjected to electrophoresis on a 8% denaturing polyacrylamide gel. Approximately 1200 c.p.m. was loaded per lane. The marker was the product of the Maxam and Gilbert A+G reaction on the same DNA.

RESULTS

Telomere binding factors in a fission yeast total protein extract

In order to identify factors which interact with the double-stranded regions of fission yeast telomeres, we isolated a 140 bp DNA fragment from a cloned *S.pombe* chromosomal telomere (46). This fragment consisted of 102 bp of telomeric repeat sequence and 38 bp of telomere associated sequence. The 102 bp of repeat sequence contained a total of 13 repeats, the majority of which conformed to the consensus 5′-TTACAG$_{1–8}$-3′ (Fig. 1A). The isolated telomeric DNA was radioactively labelled and tested in gel retardation assays with a fission yeast total protein extract (Fig. 1B). Three DNA–protein complexes were detected; a strong complex (I) and two fainter, lower mobility complexes (IIa and IIb) (lanes 2 and 5). The complexes were cross-competed by the addition of an excess of the unlabelled telomere fragment (lanes 6 and 7) but not by equivalent amounts of an unlabelled, non-specific DNA fragment,

of approximately the same size (lanes 3 and 4). This suggested that the complexes were the result of specific DNA–protein interactions. The telomeric DNA fragment used in these initial retardation assays contained telomeric repeat sequence, a short region of telomere associated sequence and a small amount of plasmid polylinker sequence. Shorter fragments which contained only the telomeric repeat sequence generated the same pattern of complexes as the original probe fragment (data not shown), indicating that complex formation did not require the telomere associated sequence or the polylinker.

In the radioactively labelled probe fragment, and in the unlabelled competitor DNA fragment, the telomeric repeats were in close proximity to one end of the duplex DNA molecule. It was possible that proximity to a DNA end might be a requirement for efficient binding by telomere binding proteins. To test this, we isolated a competitor DNA fragment containing the telomeric repeat sequences flanked by 232 and 474 bp of plasmid DNA plus telomere associated sequence, and a control fragment containing only the plasmid DNA. These DNA fragments were used as unlabelled competitor DNAs in retardation reactions containing the radioactively labelled telomeric fragment (Fig. 1C). The competitor fragment containing the telomeric sequence efficiently competed complex formation (lanes 5 and 6) whereas the competitor DNA containing the plasmid sequences caused only a slight reduction in the intensity of the complexes (lanes 2 and 3). These results suggest that telomeric sequences form a good target for telomeric binding proteins even when situated >200 bp away from a DNA end.

The retardation gel shown in Figure 1C was subjected to electrophoresis for a longer time than that shown in Figure 1B. As a consequence, an additional complex was revealed by the greater resolving power of the assay. The complex, which migrated slightly faster than complex I, was termed complex I′. Complex I′ was cross-competed efficiently only by the DNA fragment containing the telomeric repeats (lanes 5 and 6), suggesting that it was also formed by a specific DNA–protein interaction.

At least two different proteins interact with the telomeric repeat region

The pattern of complexes detected in the gel retardation assays could be explained in several ways: first, multiple binding by a single protein could account for the multiple complexes detected; secondly, different forms of a single protein might generate the complexes; thirdly, the lower mobility complexes may be ternary complexes resulting from protein–protein interactions between proteins present in the extract and a single DNA binding protein; finally, the different complexes might result from DNA binding by two or more proteins with overlapping specificities. Experiments in which increasing amounts of total protein extract were added to retardation assays containing the radioactively labelled telomere fragment demon-strated that as the ratio of protein:DNA increased, the amounts of all three complexes increased in parallel. Complexes IIa and IIb were not preferentially formed at high protein concentrations (data not shown). This suggested that multiple binding by the complex I protein was not responsible for the formation of complexes IIa and IIb. To distinguish between the other possibilities the total protein extract was fractionated by passing it through heparin–Sepharose and eluting bound proteins using a gradient of KCl. Fractions were tested in gel retardation assays using the telomere fragment as a probe. The proteins responsible for complexes I, IIa and IIb eluted between 350 and 500 mM KCl (Fig. 2). Complexes IIa and IIb
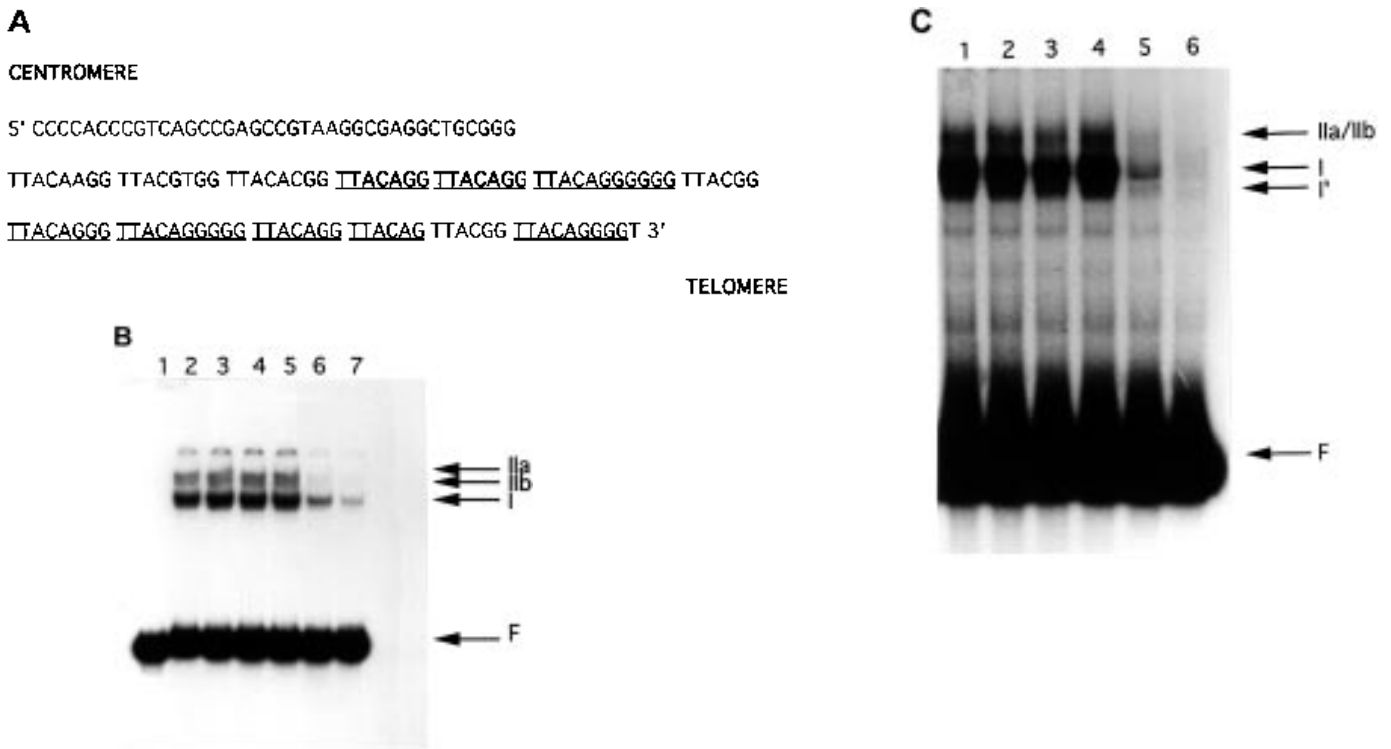
A

CENTROMERE

5' CCCCACCCGTCAGCCGAGCCGTAAGGCGAGGCTGCGGG

TTACAAGG TTACGTGG TTACACGG <u>TTACAGG</u> <u>TTACAGG</u> <u>TTACAGGGGGG</u> TTACGG

<u>TTACAGGG</u> <u>TTACAGGGGGG</u> <u>TTACAGG</u> <u>TTACAG</u> TTACGG <u>TTACAGGGGT</u> 3'

TELOMERE

B

C

Figure 1. (**A**) Sequence of the 140 bp of telomeric DNA present in the DNA fragment used in gel retardation assays (46,47). The fragment contains 13 telomeric repeats, plus 38 bp of telomere associated sequence. The sequence is shown running in the 5′→3′ direction, from the centromere towards the telomere. Eight of the 13 repeats conform to the simple consensus TTACAG$_{(1–8)}$ and these are underlined in the figure. The sequence corresponding to the oligonucleotide AC1/2 is shown in bold type. (**B**) Gel retardation assays using a fission yeast total protein extract and the radioactively labelled 140 bp telomere fragment. Lane 1, fragment alone; lanes 2–7, fragment plus total protein extract; lanes 3 and 4 contain 7- and 14-fold molar excess of an unlabelled non-specific competitor DNA fragment; lanes 6 and 7 contain 7- and 14-fold molar excess of unlabelled telomere fragment. I, IIa and IIb are the DNA–protein complexes generated, F indicates the position of the unbound labelled fragment. (**C**) Gel retardation assays using a fission yeast total protein extract and the radioactively labelled 140 bp telomere fragment. Cross competition by a competitor DNA fragment with centrally located repeat units. Lanes 1, 2 and 3, fragment, total protein extract and a 0-, 4- and 16-fold molar excess of control fragment, lanes 4, 5 and 6, fragment, total protein extract and a 0-, 4- and 16-fold molar excess of a competitor fragment containing the telomeric repeats in a central location. I, IIa and IIb are specific DNA–protein complexes. I′ is the extra complex revealed by extended electrophoresis of the reactions.

were detected in fractions ranging from ~350 to 400 mM, although in these fractions a small amount of a faster migrating complex was also detected (lanes 2–5). This faster migrating complex was probably the result of the presence of small amounts of the complex I protein. These fractions generated more of complex IIb than complex IIa, suggesting that the proteins responsible for these two complexes did not copurify during the fractionation procedure. Later fractions, eluting between ~430 and 500 mM, gave rise to a strong complex I, but produced only extremely faint complexes IIa and IIb (lanes 6–12). These fractions also generated a new high mobility complex (NC), which was probably formed by a degradation product of the protein responsible for complex I. These results confirmed that complexes IIa and IIb were not formed as a result of multiple binding by the complex I protein because some of the fractions (lanes 8 and 9) contained relatively large amounts of the complex I protein but did not produce strong complexes IIa and IIb. It is also unlikely that the three complexes were generated by differently modified forms of a single protein, unless that modification led to significantly different properties on heparin–Sepharose fractionation. The most likely explanation for these results is that complex I resulted from binding by one factor [telomere repeat factor I (TeRF I)] and that complexes IIa and IIb resulted from binding by different forms of a second factor (telomere repeat factor II (TeRF II)]. The protein responsible for

Complex I′ was not clearly present in any of the fractions collected, although the fractions in lanes 2–5 do contain a complex below complex I, which might have been complex I′. Complex I′ probably resulted from binding by a third protein which also recognised sequences present in the telomere fragment (see later).

DNA–protein interactions within the telomeric DNA

In order to localise the positions within the telomeric sequence at which DNA–protein interactions occurred, DNase I footprinting was performed. A partially purified protein fraction which generated complexes I, IIa and IIb in gel retardation assays (see Fig. 2, lane 5) and the telomeric DNA fragment labelled at the 3′ end of the AC rich strand were used (Fig. 3). The positions on the gel of the individual repeats can be determined from the positions of the AA*G repeated pattern in lane 2 (Maxam and Gilbert A+G reaction), corresponding to the conserved TTAC of each repeat on the opposite strand. We reasoned that if a regular defined footprint was obtained it would be evidence to suggest that only one protein interacted directly with the telomeric DNA. The footprint produced contained various regions of both protection and hypersensitivity (compare lanes 1 and 3 control ladders with lane 4 containing the protein fraction). The most clearly protected bands are indicated by arrows on the figure. These show no
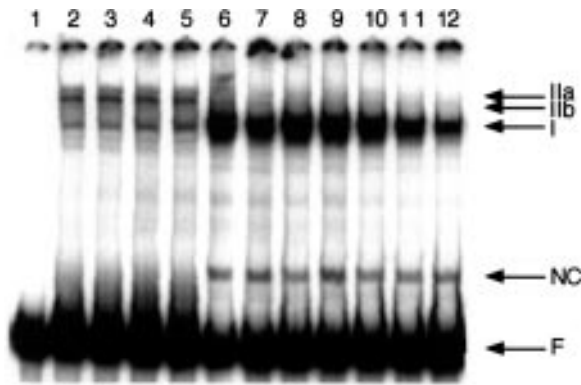
Figure 2. Gel retardation assays using radioactively labelled 140 bp telomere fragment and heparin–Sepharose fractions. Lane 1, fragment alone; lanes 2–5, fragment plus 5 μl of sequential fractions eluted between 350 and 400 mM KCl. Lanes 6–12, fragment plus 5 μl of sequential fractions eluting between 430 and 500 mM KCl. I, IIa, IIb and NC (new complex) indicate the positions of DNA–protein complexes. F indicates the position of the unbound labelled fragment.

regular pattern in relation to the telomeric repeats. Five clear hypersensitivities are also indicated. None of these are within the TTAC region of the repeat, they all occur within the 3′ end of the repeat unit. The pattern of footprint obtained might have been the result of a single DNA binding protein interacting with different repeats in different ways, but is perhaps more likely to have resulted from the presence of different proteins with overlapping DNA binding specificities.

Two adjacent copies of the telomeric repeat sequence form a binding site for both TeRF I and TeRF II

Because the gel retardation assays and footprinting experiments suggested multiple interactions between proteins in the extract and the telomeric sequences, we synthesised a series of double-stranded oligonucleotides and used these to examine the DNA sequence requirements for production of the different complexes. Initially we tested two oligonucleotides designated AC1/2 and AC3/4. AC1/2 contained the sequence 5′-GG TTACAGGTTA-CAGG TT-3′. This corresponds to two complete repeats with two base pairs of flanking sequence at each end. AC3/4 contained the sequence 5′-GG TTACAGGGGGG TT-3′. This is a particular variant of a single repeat in which six G/C pairs are present, again with two base pairs of flanking sequence at each end. These oligonucleotides were added as unlabelled competitor DNAs to gel retardation reactions containing the 140 bp telomere fragment and an *S.pombe* total protein extract (Fig. 4). The double repeat oligonucleotide (AC1/2, lanes 2 and 3) efficiently cross-competed complexes I, IIa and IIb, suggesting that this oligonucleotide contained the recognition sequences for both TeRF I and TeRF II. Interestingly it did not cross-compete formation of complex I′, suggesting that this complex must be formed by a third factor with different sequence requirements to TeRF I and TeRF II.

In contrast with AC1/2, identical amounts of the single repeat oligonucleotide (AC3/4, lanes 4 and 5) failed to compete any of the complexes, indicating that the sequence present in this oligonucleotide was not sufficient for binding by any of the factors. In order to check that the length of the oligonucleotide, rather than the presence or absence of particular sequences, was not a factor in

Figure 3. DNase I footprinting using end-labelled telomere fragment and a partially purified protein fraction. Lane 1, DNase I treatment of naked DNA (half counts loaded compared with lane 3); lane 2, Maxam and Gilbert A+G reaction on probe DNA; lane 3, DNase I treatment of naked DNA; lane 4, DNase I footprinting in presence of protein fraction. Arrows indicate protected bands; * indicates positions of hypersensitivities.

determining competition, we also tested a longer oligonucleotide containing one complete repeat unit plus extra flanking sequences. This longer oligonucleotide also failed to cross compete any of the complexes in an efficient manner (data not shown). To investigate further the sequence requirements for complex formation we synthesised another oligonucleotide designated AC5/6. This contained two complete repeats but lacked the four base pairs of flanking sequence present in AC1/2. Addition of AC5/6 to the retardation reactions failed to cross-compete any of the complexes (lanes 6 and 7), suggesting that the four base pairs which this oligonucleotide lacks include one or more base pairs which are required for binding by both TeRF I and TeRF II.

The telomeric repeat contains two highly conserved T residues at the 5′ end and a more variable run of G residues at the 3′ end. In order to test whether the conserved T residues were important for binding by either factor, we synthesised a further oligonucleotide designated AC7/8. This was identical to AC1/2 but contained mutations within the highly conserved T residues of both repeats. To maintain the mainly TG rich nature of the DNA strand running 5′→3′ we mutated one T residue in each repeat to a G residue. The double repeat in AC7/8 was therefore converted to 5′-GG T**G**ACAGG T**G**ACAGG TT-3′. When AC7/8 was added to retardation assays containing the 140 bp telomere fragment and
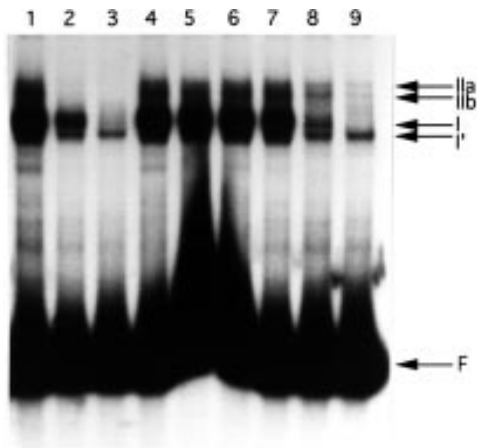
Figure 4. Gel retardation assays using a fission yeast total protein extract and the radioactively labelled 140 bp telomere fragment. Competition by double-stranded oligonucleotides containing telomeric repeat sequences. Lanes 1–9 contain the fragment plus total protein extract; lanes 2 and 3 contain 60- and 150-fold molar excess of unlabelled AC1/AC2 oligonucleotide repeats; lanes 4 and 5 contain 60- and 150-fold molar excess of unlabelled AC3/4 oligonucleotide repeats; lanes 6 and 7 contain 60- and 150-fold molar excess of unlabelled AC5/6 oligonucleotide repeats; lanes 8 and 9 contain 60- and 150-fold molar excess of unlabelled AC7/8 oligonucleotide repeats. F indicates the position of the unbound labelled fragment. I′, I, IIa and IIb indicate the positions of DNA–protein complexes.

Figure 5. Gel retardation assays using a fission yeast total protein extract and the radioactively labelled 140 bp telomere fragment. Competition by single-stranded oligonucleotides containing telomeric repeat sequences. Lane 1 contains the fragment alone; lanes 2–10 contain the fragment plus total protein extract; lanes 3 and 4, 6 and 7, and 9 and 10 contain respectively 150- and 900-fold molar excess of unlabelled AC1, AC2 and AS1 oligonucleotides. F indicates the position of the unbound labelled fragment. I′, I, IIa and IIb indicate the positions of DNA–protein complexes.

a *S.pombe* total protein extract, it was found to cross-compete the formation of complex I as efficiently as AC1/2, however, competition of complexes IIa and IIb was much reduced. Like the original double repeat oligonucleotide it also failed to compete complex I′. These results suggested that the two bp mutated in this oligonucleotide are important for binding by TeRF II but not for binding by TeRF I, the oligonucleotide therefore dramatically reduced the level of complex I, but had less effect on the levels of complexes IIa and IIb.

The telomere binding factors do not interact strongly with single-stranded DNA

Telomeres in several different organisms have been shown to end in a single-stranded 3′ extension, usually only a few repeats in length (5,6). In budding yeast, telomeres gain long single-stranded extensions during the S phase of the cell cycle (7). Single-stranded extensions are the targets for several telomere binding proteins, including the budding yeast protein Rap1p, which promotes the formation of G tetrad structures (15). Although it is not known if such extensions are ever found at fission yeast telomeres, we have tested whether any of the telomere binding factors we have identified will bind to single-stranded DNA, using a range of single-stranded oligonucleotides as competitors. Because binding by budding yeast Rap1p to single-stranded DNA is less efficient than to double-stranded DNA, we used the single-stranded oligonucleotides at higher concentrations than the double-stranded oligonucleotides (Fig. 5). Three single-stranded oligonucleotides were tested; AC1 corresponds to the TG rich strand of the AC1/2 double repeat oligonucleotide, AC2 corresponds to the AC rich strand, and AS1 is a control oligonucleotide. None of the oligonucleotides cross competed the formation of any of the complexes (Fig. 5, lanes 3 and 4, 6 and 7, 9 and 10), even when used at high concentrations. This
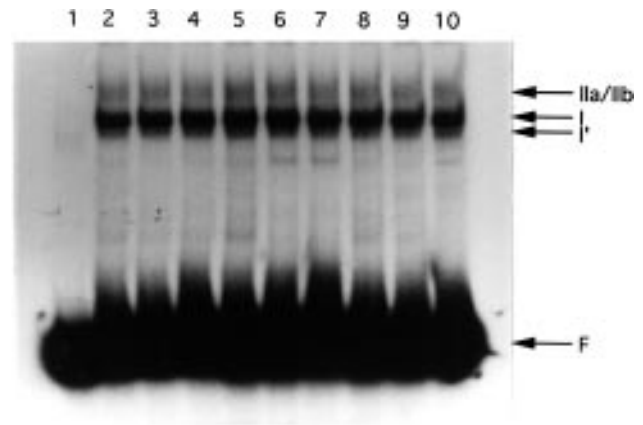
indicated that neither of the individual strands of the double repeat is a strong substrate for either TeRF I or TeRF II.

DISCUSSION

Schizosaccharomyces pombe telomeres, like those in other eukaryotes, are made up of many copies of a short repeat unit (46,47). We have now shown that proteins are present in *S.pombe* which can interact specifically with these telomeric repeats. Four specific DNA–protein complexes were detected by gel retardation assays, using a telomere fragment and a total protein extract. These complexes were produced as a result of binding by at least three different factors (see below). All four complexes were produced irrespective of the proximity of the repeats to the end of the DNA fragment, indicating that the factors we have identified are not end-specific. Fractionation of the total protein extract using heparin–Sepharose separated the factor responsible for complex I away from the factor responsible for complexes IIa and IIb. The partially purified factors were termed TeRF I and TeRF II respectively. Complexes IIa and IIb were probably produced by differently modified forms of TeRF II because IIa and IIb were generated by the same column fractions, and their DNA binding specificities appeared identical in cross-competition experiments. The existence of differently modified forms could be a property of certain telomere binding proteins because the human telomere repeat binding factor TRF also produced multiple complexes in gel retardation assays when isolated from HeLa cells (44). The factor responsible for complex I′ was not unambiguously identified in any of the fractions tested.

We have probed the DNA recognition specificities of the different factors using a series of unlabelled oligonucleotides as competitor DNAs in gel retardation assays. These experiments demonstrated that the factors responsible for the different complexes had similar, but distinct, recognition sequences. An oligonucleotide (AC3/4) containing one complete telomeric repeat failed to compete any of the complexes, suggesting that a single repeat is not a strong binding

site for TeRF I, TeRF II or the complex I′ factor. An oligonucleotide (AC1/2) containing two complete telomere repeat units cross-competed formation of complexes I, IIa and IIb, but not complex I′. AC1/2 therefore contained binding sites for both TeRF I and TeRF II but not for the protein which formed complex I′. The presence of binding sites for TeRF I and TeRF II in an oligonucleotide containing just two complete repeat units, suggests that these factors may not be analogous to the human telomere binding factor TRF, which requires six repeat units for strong binding (43). The oligonucleotide AC1/2 contained additional telomeric sequences, flanking the two complete repeats. These sequences comprised two G/C bp at the 5′ end and two T/A bp at the 3′ end. When these extra bases were absent (AC5/6) the ability of the oligonucleotide to compete complex formation was lost. When an oligonucleotide identical to AC1/2, but containing mutations of the second T in each repeat (AC7/8), was used as a competitor DNA, complex I was still efficiently competed, but competition of complexes IIa and IIb was much reduced. The mutations in this oligonucleotide have therefore dramatically reduced the strength of binding of TeRF II, but had little effect on TeRF I binding. The T/A bp which was mutated in each repeat within this oligonucleotide is highly conserved and is present in all telomeric repeat units in *S.pombe*. The lack of effect of these mutations on TeRF I binding may indicate that the blocks of G/C bp within the oligonucleotide, and the spacing between these blocks, are the critical factors for strong binding by this protein. The difference in DNA recognition sequence requirements for formation of the complexes confirmed that complexes IIa and IIb were produced by a different factor to complex I.

Individual single-stranded oligonucleotides from AC1/2 failed to cross-compete formation of any of the complexes, even when present at relatively high levels. This suggested that none of the factors interacts efficiently with single-stranded DNA. Our experiments were designed to identify strong single-stranded binding activities. It remains possible that one or more of our factors possesses a weak single-stranded binding activity that requires the presence of multiple copies of a recognition sequence for complex formation *in vitro*, analogous to the interaction between budding yeast Rap1p and single-stranded DNA (15).

Budding yeast Rap1p is currently the best characterised telomere binding factor in any yeast species. We have now identified two new telomere binding factors in fission yeast. Is either of these factors a good candidate to be a homologue of Rap1p? Rap1p is an abundant protein which interacts with a consensus sequence generated in telomeric DNA by juxtaposition of particular variants of the TG$_{(1–3)}$ repeat unit. Both TeRF I and TeRF II have recognition sequences formed by two copies of the telomeric repeat unit. However, TeRF I appears to be an abundant DNA binding activity, whilst TeRF II is less abundant. Because Rap1p in budding yeast also binds to the promoters of many glycolytic and ribosomal protein genes (25–32) we have performed preliminary experiments to test whether TeRF I or TeRF II binds to a ribosomal protein gene promoter from fission yeast. Cross competition experiments suggested that neither factor interacted with the single ribosomal protein gene promoter (from the *K5* ribosomal protein gene) that we tested (data not shown).

Telomere binding factors play a role in telomere function and also appear to be important in controlling the length of telomeres. In budding yeast, changes in the level of Rap1p have been shown to result in changes to the lengths of telomeres (8,19). In addition, it has been proposed recently that factors which bind to the

telomeric repeats of *Kluyveromyces lactis* play a role in negatively modulating the activity of telomerase (52). As changes in telomerase activity and in the lengths of telomeres have been implicated in carcinogenesis and cellular ageing, characterisation of telomere binding proteins and an understanding of their roles, may be of key importance in understanding these processes (53–55). If the interactions which we have identified *in vitro* prove to be important *in vivo*, *S.pombe* may be a useful model organism in this regard.

ACKNOWLEDGEMENTS

REFERENCES

1 Blackburn, E.H. (1991) *Nature*, **350**, 569–573.
2 Blackburn, E.H. (1994) *Cell*, **77**, 621–623.
3 Zakian, V.A. (1989) *Ann. Rev. Genet.*, **23**, 579–604.
4 Levis, R.W., Ganesan, K.H., Tolar, L.A. and Sheen, F.M. (1993) *Cell*, **75**, 1083–1093.
5 Pluta, A.F., Kaine, B.P. and Spear, B.B. (1982) *Nucleic Acids Res.*, **10**, 8145–8154.
6 Klobutcher, L.A., Swanton, M.T., Donini, P. and Prescott, D.M. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 3015–3019.
7 Wellinger, R.J., Wolf, A.J. and Zakian, V.A. (1993) *Cell*, **72**, 51–60.
8 Conrad, M.N., Wright, J.H., Wolf, A.J. and Zakian, V.A. (1990) *Cell*, **63**, 739–750.
9 Gottschling, D.E. and Zakian, V.A. (1986) *Cell*, **47**, 195–205.
10 Price, C.M. and Cech, T.R. (1987) *Genes Dev.*, **1**, 783–793.
11 Gray, J.T., Celander, D.W., Price, C.M. and Cech, T.R. (1991) *Cell*, **67**, 807–814.
12 Price, C.M., Skopp, R., Krueger, J. and Williams, D. (1992) *Biochemistry*, **31**, 10835–10843.
13 Shore, D. and Nasmyth., K. (1987) *Cell*, **51**, 721–732.
14 Gilson, E., Muller, T., Sogo, J., Laroche, T. and Gasser, S.M. (1994) *Nucleic Acids Res.*, **22**, 5310–5320.
15 Giraldo, R. and Rhodes, D. (1994) *EMBO J.*, **13**, 2411–2420.
16 Kyrion, G., Boakye, K.A. and Lustig, A.J. (1992) *Mol. Cell. Biol.*, **12**, 5159–5173.
17 Kyrion, G., Liu, K. and Lustig, A.J. (1993) *Genes Dev.*, **7**, 1146–1159.
18 Longtine, M.S., Wilson, N.M., Petracek, M.E. and Berman., J. (1989) *Curr. Genet.*, **16**, 225–239.
19 Lustig, A.J., Kurtz, S. and Shore., D. (1990) *Science*, **250**, 549–553.
20 Klein, F., Laroche, T., Cardenas, M.E., Hofmann, J.F.X., Schweizer, D. and Gasser, S.M. (1992) *J. Cell Biol.*, **117**, 935–948.
21 Palladino, F., Laroche, T., Gilson, E., Axelrod, A., Pillus, L. and Gasser, S.M. (1993) *Cell*, **75**, 543–555.
22 Gilson, E., Roberge, M., Giraldo, R., Rhodes, D. and Gasser, S.M. (1993) *J. Mol. Biol.*, **231**, 293–310.
23 Buchman, A.R., Lue, N.F. and Kornberg, R.D. (1988) *Cell*, **8**, 5086–5009.
24 Buchman, A.R., Kimmerly, W.J., Rine, J. and Kornberg, R.D. (1988) *Mol. Cell. Biol.*, **8**, 210–225.
25 Huet, J. and Sentenac, A. (1987) *Proc. Natl Acad. Sci. USA* **84**, 3648–3652.
26 Capieaux, E., Vignais, M.L., Sentenac, A. and Goffeau, A. (1989) *J. Biol. Chem.*, **264**, 7437–7446.
27 Vignais, M.L., Woudt, L.P., Wassenaar, G.M., Mager, W.M., Sentenac, A. and Planta, R.J. (1987) *EMBO J.*, **6**, 1451–1457.
28 Brindle, P.K., Holland, J.P., Willett, C.E., Innis, M.A. and Holland., M.J. (1990) *Mol. Cell. Biol.*, **10**, 4872–4885.
29 Chambers, A., Tsang, J.H., Stanway, C., Kingsman, A.J. and Kingsman, S.M. (1989) *Mol. Cell. Biol.*, **9**, 5516–5524.
30 Giesman, D., Best, L. and Tatchell, K. (1991) *Mol. Cell. Biol.*, **11**, 1069–1079.
31 Hurd, H.K. and Roberts, J.W. (1989) *Mol. Cell. Biol.*, **9**, 5359–5372.
32 Scott, E.W. and Baker., H.V. (1993) *Mol. Cell. Biol.*, **13**, 543–550.

33 Sussel, L. and Shore., D. (1991.) *Proc. Natl Acad. Sci. USA* **88**, 7749–7753.
34 Hardy, C.F.J., Balderes, D. and Shore, D. (1992) *Mol. Cell. Biol.*, **12**, 1209–1217.
35 Kurtz, S. and Shore, D. (1991) *Genes Dev.*, **5**, 616–628.
36 Liu, C., Mao, X. and Lustig, A.J. (1994) *Genetics*, **138**, 1025–1040.
37 Larson, G.P., Castanotto, D., Rossi, J.J. and Malafa, M.P. (1994) *Gene*, **150**, 35–41.
38 Brigati, C., Kurtz, S., Balderes, D., Vidali, G. and Shore, D. (1993) *Mol. Cell. Biol.*, **13**, 1306–1314.
39 Lin, J. and Zakian, V.A. (1994) *Nucleic Acids Res.*, **22**, 4906–4913.
40 Cardenas, M.E., Bianchi, A. and de Lange, T. (1993) *Genes Dev.*, **7**, 883–894.
41 McKay, S.J. and Cooke, H. (1992) *Nucleic Acids Res.*, **20**, 1387–1391.
42 Ishikawa, F., Matunis, M.J., Dreyfuss, G. and Cech, T.R. (1993) *Mol. Cell. Biol.*, **13**, 4301–4310.
43 Zhong, Z., Shiue, L., Kaplan, S. and de Lange, T. (1992) *Mol. Cell. Biol.*, **12**, 4834–4843.
44 Chong, L., Steensel, B.V., Broccoli, D., Erdjument-Bromage, H., Hanish, J., Tempst, P. and De Lange, T. (1995) *Science*, **270**, 1663–1671.
45 Chikashige, Y., Ding, D.Q., Funabiki, H., Haraguchi, T., Mashiko, S., Yanagida, M. and Hiraoka, Y. (1994) *Science*, **264**, 270–273.
46 Sugawara, N. (1989) *PhD thesis, Harvard University*,
47 Sugawara, N. and Szostak, J.W. (1986) *Yeast*, **2 (suppl.)**, 373.
48 Ekwall, K., Javerzat, J.P., Lorentz, A., Schmidt, H., Cranston, G. and Allshire, R. (1995) *Science*, **269,** 1429–1431.
49 Hawthorne, D.C. and Mortimer, R.K. (1960) *Genetics*, **45**, 1085–1110.
50 Ogden, J.E., Stanway, C., Kim, S., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) *Mol. Cell. Biol.*, **6**, 4335–4343.
51 Bradford, M.M. (1976) *Anal. Biochem.*, **72**, 248–254.
52 McEachern, M.J. and Blackburn, E.H. (1995) *Nature*, **376**, 403–408.
53 Counter, C.M., Avilion, A.A., LeFeuvre, C.E., Stewart, N.G., Greider, C.W., Harley, C.B. and Bacchetti, S. (1992) *EMBO J.*, **11**, 1921–1929.
54 Harley, C.B., Futcher, A.B. and Greider, C.W. (1990) *Nature*, **345**, 458–460.
55 Hastie, N.D., Dempster, M., Dunlop, M.G., Thompson, A.M., Green, D.K. and Allshire, R.C. (1990) *Nature*, **346**, 866–868.