

Title:

Rif1 promotes association of G-quadruplex (G4) by its specific G4 binding and oligomerization activities

Authors:

***Hisao Masai¹, Rino Fukatsu¹, Naoko Kakusho¹, Yutaka Kanoh¹, Kenji Moriyama¹, Yue Ma², Keisuke Iida³, Kazuo Nagasawa²**

Affiliations:

¹Department of Genome Medicine, Tokyo Metropolitan Institute of Medical Science, Kamikitazawa, Setagaya-ku, Tokyo 156-8506, Japan

²Department of Biotechnology and Life Science, Faculty of Technology, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan

³Molecular Chirality Research Center, Synthetic Organic Chemistry, Department of Chemistry, Graduate School of Science, Chiba University, Chiba, 263-8522, Japan.

Running title: Binding of Rif1 to G-quadruplex

Key words: G-quadruplex, Rif1, DNA replication timing, nucleic acid structure, chromatin architecture

***Correspondence to: Dr. Hisao Masai, Telephone: +81-3-5316-3220, Fax: +81-3-5316-3145, E-mail: masai-hs@igakuken.or.jp**

Rif1 is a conserved protein regulating replication timing and binds preferentially to the vicinity of late-firing/ dormant origins in fission yeast. The Rif1 binding sites on the fission yeast genome have an intrinsic potential to generate G-quadruplex (G4) structures to which purified Rif1 preferentially binds. We previously proposed that Rif1 generates chromatin architecture that may determine replication timing by facilitating the chromatin loop formation. Here, we conducted detailed biochemical analyses on Rif1 and its G4 binding. Rif1 prefers sequences containing long stretches of guanines and binds preferentially to the multimeric G4 of parallel or hybrid/mix topology. Rif1 forms oligomers and binds simultaneously to multiple G4. We present a model on how Rif1 may facilitate the formation of chromatin architecture through its G4 binding and oligomerization properties.

DNA replication proceeds in accordance with temporal and spatial program that is intrinsic to each cell type ¹. In higher eukaryotes, replication timing may be determined by megabase chromosome domains that dictate the temporal units for DNA replication ^{2,3}. Rif1, a conserved nuclear factor originally identified as a telomere binding protein in yeasts ^{4,5}, was recently shown to play a major role in defining the replication timing domains in both yeast ^{6,7} and mammalian cells ⁸⁻¹⁰.

In fission yeast *rif1Δ* cells, late/ dormant origins are fired in the presence of HU (hydroxy urea) or fired early even during normal S phase and initiation at some early-firing origins was reduced ⁶. In mammalian cells, replication timing domain structures were dramatically altered in Rif1-depleted or knockout cells ⁸⁻¹⁰. It was shown that chromatin loop length became longer in Rif1-depleted cells, suggesting that Rif1 may be involved in generation of chromatin loop structures ⁸. Immunostaining indicated that Rif1 is localized at nuclear periphery in Triton- and DNaseI-insoluble structures, showing that it associates with nuclear matrix structures, probably tethering chromatin fibers at nuclear periphery. It has been known that mid-to-late replication foci are localized at nuclear periphery ^{1,11}. Thus, we proposed that Rif1 may generate chromatin compartments that define mid-to-late replicating chromosome domains. Analyses of Rif1 binding sites and chromatin interactions by the 4C-seq assays in mouse ES cells also indicated that Rif1 confined the chromatin interactions within each replication timing domain ¹².

It was also reported that Rif1 carries binding sites for phosphatase (PP1; Protein Phosphatase 1) ¹³⁻¹⁸ and that the recruited phosphatase inhibits initiation by antagonizing the

phosphorylation events mediated by Cdc7 kinase. Therefore, Rif1's abilities to organize functional chromatin domains and to recruit a phosphatase contribute to temporal and spatial regulation of DNA replication.

Rif1 has also been implicated in repair of double-stranded DNA breaks¹⁹⁻²². It suppresses homologous recombination-dependent repair by inhibiting an end-resection reaction and stimulates non-homologous end-joining repair²³⁻²⁷. Recent reports indicate additional roles of Rif1 at the replication forks and S phase regulation during early embryogenesis²⁸⁻³⁰. It would be an interesting possibility that Rif1 may also regulate DSB repair through its ability to organize chromatin architecture at nuclear periphery.

Analyses of Rif1 binding sequences (Rif1BS) with ChIP-seq led to the identification of Rif1CS (Rif1 binding consensus sequence) containing 5-6 runs of guanine residues³¹. We then showed that Rif1BS can adopt G-quadruplex(G4)-like structures *in vitro* and the purified Rif1 protein binds to G4 structures. Strong correlation between the ability of Rif1BSs to form G4 *in vitro* and binding of Rif1 to these sequences *in vivo* indicates that Rif1 specifically recognizes G4 structures that are indeed formed in cells.

In order to clarify how Rif1 interacts with G4 DNA and contributes to the formation of specific nuclear architecture, we have conducted detailed analyses of target sequences of fission yeast Rif1 protein, and also biochemically characterized this protein. We found that Rif1 preferentially binds to multimeric G4 structures with parallel or hybrid/mix-type topology containing 5-6 runs of guanine and show that Rif1 protein forms oligomers and promotes association of multiple DNAs containing G4 structures. On the basis of these data, we will present a model on how Rif1 may interact with G4 DNA and how it may contribute to the establishment of replication timing domains.

Results

Purification of the full-length fission yeast Rif1 protein

Fission yeast Rif1 (hereafter, referred to as "Rif1"; Rif1 from other species will be specified) protein is 1,400 amino acid long, composed of the N-terminal HEAT (Huntingtin, Elongation factor 3, A subunit of protein phosphatase 2A, and TOR) - and ARMADILLO-type repeats¹³ and a C-terminal unknown domain. We expressed the full-length Rif1 in human embryonic kidney cells 293T³² in an N-terminally His₆ and C-terminally FLAG₃-tagged form. We first showed that the presence of the tags at the N- and C-termini of the protein does not affect its function by showing i) expression of the tagged protein in $\Delta rif1$ *hsk1-89*¹⁸ abrogated the bypass of the Hsk1 function, and ii) the telomere length is not significantly affected by the tags. To assess the function of N-terminally His₆- and C-terminally FLAG₃-tagged Rif1 protein, we expressed the His₆-Rif1-Flag₃ protein in *rif1* Δ cells. We took advantage of the fact that Hsk1 (the homologue of Cdc7 kinase) function is bypassed by *rif1* deletion. Expression of the functional Rif1 in *hsk1-89 rif1* Δ cells (viable at 30°C) inhibits the growth of the strain at 30°C (non-permissive for *hsk1-89*). This reflects the ability of Rif1 to inhibit the firing of late origins. We first examined the expression level of non-tagged Rif1 and His₆-Rif1-Flag₃ cloned on pREP81 in the presence of low levels of thiamine with western using the antibody that can detect the endogenous Rif1 protein. At 2 μ M or 5 μ M thiamine, both non-tagged and tagged proteins were expressed at, respectively, ~5 fold or ~3 fold more than the endogenous Rif1 protein (**Fig. 1A**). Cell cycle profiles of *rif1* Δ cells carrying pREP81 plasmid expressing non-tagged or tagged Rif1 exhibited the identical pattern (**Fig. 1B**). We then examined the effect of both proteins on the growth of *hsk1-89 rif1* Δ cells at 30°C. *hsk1-89* harboring a vector can not grow at 30°C, whereas *hsk1-89 rif1* Δ harboring the vector can grow at this temperature. *hsk1-89 rif1* Δ harboring pREP81-Rif1 grew poorly at 30°C due to expression of the wild-type Rif1 protein. Similarly, *hsk1-89 rif1* Δ harboring pREP81-His₆-Rif1-Flag₃ grew poorly (**Fig. 1C**). The extent of the growth was similar between non-tagged and tagged Rif1, indicating that His₆-Rif1-Flag₃ proteins retain the ability to inhibit the origin firing. The results were similar at both 5 μ M and 15 μ M thiamine.

We also examined the telomere functions of His₆-Rif1-Flag₃. The length of DNA fragments containing telomeres in the wild-type cells (YM71) was ~1,000 bp under our experimental condition, whereas that in *rif1* Δ cells was ~1,200 bp and an additional fragments were detected at around 2.5 kb. The telomere length returned to the wild-type level in *rif1* Δ cells harboring pREP81-His₆-Rif1-Flag₃, indicating that the double-tagged Rif1 retains the telomere regulation function as well (**Fig. 1D**).

The protein was purified by consecutive anti-FLAG and nickel affinity columns. The protein is highly prone to degradation, giving rise to a 70 kDa truncated polypeptide, which is derived

from the C-terminal segment (indicated by its reactivity to the C-terminal FLAG tag). In order to obtain a full-length form of Rif1 that is devoid of the 70 kDa degradation polypeptide, we conducted monoQ column fractionation or glycerol gradient fractionation, which permitted us to partially remove the degradation product (**Supplementary Fig. S1A**, lane 1 and data not shown). Using these preparations, we examined the affinity of the full-length Rif1 to a G4 substrate, T₆G₂₄ that forms a propeller-type parallel-stranded G4 containing three G-tetrad layers and three single-guanine propeller loops³³ (**Supplementary Fig. S1B**; see a drawing in **Fig. 2**). The monoQ fraction contains ~1:2 ratio of the full-length and the 70kDa polypeptide, and the apparent K_d of this preparation for this G4 DNA was ~0.3 nM (**Supplementary Fig. S1C**), when only the full-length polypeptide was considered. Even if the coexisting 70 kDa polypeptide (present at 4-fold molar excess than the full-length and on the assumption it binds to the substrate with affinity same as the full-length) were considered, the K_d would be less than 1.5 nM.

We examined the DNA binding property of the 70 kDa polypeptide. We were able to obtain the fractions containing only the 70 kDa polypeptide free from the full-length protein by glycerol gradient centrifugation or monoQ ion exchange chromatography (**Supplementary Fig. S1A**, lanes 2 and 3). We conducted gel shift assays with these preparations along with the full-length fraction. The 70 kDa polypeptide exhibited much reduced affinity compared to the full-length; at 10 nM only a half of the substrate was bound (with estimated K_d >10 nM; **Supplementary Fig. S1B and C**). Two different preparations of the 70 kDa polypeptide showed almost identical affinity to the G4 substrate. Thus, we can safely conclude that the full-length Rif1 protein binds to G4 DNA with extremely high affinity, while the 70 kDa polypeptide binds to the same DNA with at least one order of magnitude lower affinity. Furthermore, we can be reassured that our preparation of Rif1, although not completely devoid of the C-terminally derived polypeptide, represents mostly the binding of the full-length form at a lower concentration, and thus will be used in the following experiments.

Interaction of Rif1 with various G4 DNAs

We examined the binding of Rif1 protein to single-stranded DNAs that are known to form G4 structures. Among those examined, Rif1 bound to Pu24T, CEB1 and 93del more efficiently than others (**Fig. 2**, lanes 6-8, 10-12 and 18-20). These single-stranded DNAs contain 4-6 runs of guanines, suggesting a possibility that consecutive guanines may facilitate the generation of the structures favored by Rif1. Indeed, the G₂₄ oligonucleotide known to generate a parallel-type G4 structure³³ is bound by Rif1 with high affinity (**Fig. 2**, lanes 53-56). Notably, Rif1 binds to telomere-derived sequence only very inefficiently (**Fig. 2**, lanes 21-40). The K_ds for different G4 oligonucleotides were estimated to be 0.5-0.9 nM, values similar to that for T₆G₂₄ (**Supplementary Fig. S2**)

We then further examined the sequence preference of Rif1 by generating series of the derivatives of these good and poor binders. Htelo1 is not bound by Rif1 very efficiently (**Fig. 2**, lanes 25-28 and **Fig. 3**, lane 12). When the TTA spacer between the 5'-proximal two 3Gs was deleted (resulting in the generation of 6G near the 5'-end), binding efficiency was increased (Htelo1_no_spacer; **Fig. 3**, lane 16), while reducing the spacer to "A" did not affect the binding (Htelo1_A; **Fig. 3**, lane 14). Similarly, deletion of the TTA between the 3'-proximal two 3Gs (resulting in the generation of 6G at the 3'-end) led to even more increase of the Rif1 binding (Htelo1_no_spacer_2; **Fig. 3**, lane 18). On the other hand, insertions at the middle of 5'-proximal 6Gs did not significantly affect the Rif1 binding in CEB1 DNA, a relatively good binder (CEB1_A, CEB1_TA and CEB1_TTA; **Fig. 3**, lanes 1-8). Elongation of the 5'-proximal G runs in T95-2T did not increase, or even reduced the Rif1 binding (T95-2T-G, T95_2T_GG and T95_2T_GGG; **Fig. 3, lanes 23-30**), showing that simply the presence of a long G-tract is not sufficient for efficient Rif1 binding (see also **Supplementary Fig. S3**).

Htelo4 is also a poor binder of Rif1 (**Fig. 2**, lanes 37-40; **Fig. 4**, lanes 1-4), but deletion of CTA between the 3'-proximal two 3Gs significantly increased the binding (Htelo4(GGGGGG); **Fig. 4**, lanes 9-12). On the other hand, insertion of 3 nucleotides into the central loop did not affect the Rif1 binding both in the Htelo4 and Htelo4(GGGGGG) (Htelo4_3nt_spacer and Htelo4(GGGGGG)_3nt_spacer; **Fig. 4**, lanes 5-8 and 13-16; see also **Supplementary Fig. S4**). Both Htelo4_3nt_spacer and Htelo4(GGGGGG)_3nt_spacer show the presence of slow-migrating forms of DNA on PAGE which may represent oligomeric assembly of monomeric G4 or intermolecular G4 assembly. Rif1 preferentially binds to these slow-migrating forms. Indeed, those G4 oligonucleotides that are bound with Rif1 exhibit the slow migrating forms on polyacrylamide gel and those forms are preferentially bound by Rif1. In contrast, those that are poor binders generate very little slow migrating forms and predominantly generate a single band. Comprehensive analyses on polyacrylamide gel (12%

PAGE with 10% PEG200) revealed that all the oligonucleotides bound by Rif1 showed slow-migrating forms (**Supplementary Fig. S5**). Thus, it appears that the ability to generate slow-migrating forms (representing oligomers of G4) may be a major determining factor for affinity of Rif1 to G4. The results also suggest that the presence of a long G-tract at the 3'-terminus may facilitate the formation of "multimeric" forms of G4 that are bound by Rif1 with higher affinity. The results also indicate that the loop length between the long 3' G-tracts and the adjacent G-tracts does not significantly affect Rif1 binding.

Conserved (A)GTGGG sequence in Rif1BS can mediate Rif1 binding

We previously identified Rif1CS (Rif1 consensus sequence; CNWWGTGGGGG [W=A or T]) through analyses of 35 Rif1BS sequences³¹. We also noted the presence of additional G-tracts in the vicinity of Rif1CS. Visual inspection of these sequences led us to realize that these additional G-tracts are frequently preceded by (A)GT, making (A)GTG_n(n=3~6) a more plausible conserved sequence motif. We then synthesized (AGTGGG)_n(n=1~4), (AGTGGG)_n(n=1~4), and (AGTGGGG)_n(n=1~4), and used them in gel shift assays to examine whether these conserved motifs are specifically recognized by Rif1. The oligonucleotides were heat-treated/renatured in 50 mM KCl and 40% PEG200 (a reagent that facilitates formation of G4 with molecular crowding effect³⁴) before used in the assays. Among the oligonucleotides synthesized, the single "AGTGGG" was not bound with Rif1, but the single "AGTGGGG" or "AGTGGGGG" was weakly bound (**Supplementary Fig. 6A**, lanes 2, 8 and 14). The oligonucleotides carrying two or more repeats of AGTG_n were bound with Rif1. The most efficient binding was observed with three repeats (**Supplementary Fig. S6A**, lanes 6, 12 and 22 and **Fig. S5B**, lane 6).

We then synthesized (ACAGGGG)_n or (ACAGGGGG)_n (n=1 or 3) to evaluate the significance of the conserved GT present before the G-tracts. Unexpectedly, Rif1 bound to ACAG_n oligonucleotides as efficiently as or slightly more efficiently than to the AGTG_n oligonucleotides (**Supplementary Fig. S6B**, compare lanes 2&4, 6&8, 10&12, and 14&16). The 8mer ACAGGGGG generated significant amounts of smeared slow-migrating forms after heat treatment, to which Rif1 bound (**Supplementary Fig. S6B**, lanes 11 and 12). The results indicate that these small G-rich oligonucleotides can form G4 structures which are preferred targets of Rif1 binding. We also examined whether 6G, 8G and 10G tract oligonucleotide can be bound with Rif1, and showed that 10G can be bound with Rif1, and 8G to a small extent (**Supplementary Fig. S7**). Thus, the conserved AGT sequence before G-tracts does not appear to be crucial for Rif1 recognition, and it is less likely that the GTG_n sequence motif is specifically recognized by Rif1. However, we cannot rule out the possibility that the conserved GT sequence somehow facilitates the formation of a specific DNA structure on a duplex DNA that is preferentially bound by Rif1 in the cells.

CD analyses of various G4-forming oligonucleotides suggest the preference of Rif1 for G4 with specific topology and increased stability

Various topologies are known to exist for G4 DNA. Those include parallel, anti-parallel, hybrid and mix³⁵. These different topologies can be identified by measuring CD (Circular Dichroism) spectra³⁶⁻³⁹. In case of a typical anti-parallel structure, positive cotton effects are observed at 290 and 240 nm, and negative one at 265 nm. CD spectra of a typical parallel structure have positive and negative effects at 265 and 240 nm, respectively. In a typical hybrid structure, characteristic CD spectra are observed at 290 and 240 nm as positive and negative effects. We therefore measured the CD spectra of all the oligonucleotides used for Rif1 binding analyses (**Supplementary Fig. S8** and **Supplementary Table. S1**). Htelo4, derived from telomeres, adopts anti-parallel forms and are not bound by Rif1 (**Fig. 4**, lanes 1-4). Generation of 6G stretch converted Htelo4 and its non-binding derivative (Htelo4_3nt_spacer; anti-parallel form) into good binders (Htelo4(GGGGGG) and Htelo4(GGGGGG)_3nt_spacer), which adopt hybrid forms (**Fig. 4**, lanes 9-16). However, Htelo1~2, which are judged to be hybrid forms, are poor binders (**Fig. 2**, lanes 25-32). Therefore, there may be other factors that converted the non-binders to the good binders. It could be the multimer formation by the latter DNAs (**Supplementary Fig. S5**). Other sequences that are bound by Rif1 are mostly parallel-type, but the derivatives of T95_2T (T95_2T_G, T95_2T_GG, and T95_2T_GGG) are parallel-type and are not efficiently bound by Rif1 (**Fig. 3**, lanes 25-30; T95_2T_GGG was weakly bound by Rif1, **Supplementary Fig. S3**, lane 30). Single-stranded DNA with sequences derived from Rif1BS are generally efficiently bound by Rif1⁴⁰. Rif1-8, one of them, also adopts a parallel-type structure (**Supplementary Fig. S8**).

We next conducted melting assay to evaluate the stability of G4 by CD (**Supplementary Fig. S9**). In this assay, CD of T₆G₂₄, CEB1, CEB1_TA, CEB25, T95_2T and Htelo4 were measured at different temperatures (with 50 mM KCl). Structures of T₆G₂₄, CEB1, and CEB1_TA (good binders) were very stable, maintaining the parallel-type structure even at 95°C. On the other hand, structures of CEB25, T95_2T (low efficiency binders) and Htelo4 (non-binder) are lost at this temperature. The structure of Htelo4 was most unstable, being disrupted at temperatures above 55°C. In order to more precisely determine the melting temperatures of the stable G4, we conducted the melting assays at 10 mM KCl. Parallel-type structures were maintained for T₆G₂₄, CEB1, and CEB1_TA even under this condition. Structures of CEB25 and T95_2T were disrupted at higher temperatures (>75°C for CEB25 and >85°C for T95_2T). Our results suggest that there is general correlation between the heat stability and affinity to RifI.

Formation of G4 is affected by the presence of monovalent cations⁴¹, and it is known that the kind of salt affects the type of topology that a given sequence adopts. Indeed, the topology of some oligonucleotides changed from hybrid-type (in KCl) to anti-parallel (in NaCl). Therefore, we denatured/reannealed CEB1_TA, CEB1_TTA, Htelo1_no_spacer_2 and Htelo4(GGGGGG)_3nt_spacer in the presence of KCl or NaCl or in the absence of salt, and compared their affinity to RifI. These DNAs were efficiently bound with RifI when they were heat denatured in the presence of KCl and PEG200 (**Fig. 3**, lanes 5-8, 17, 18; **Fig. 4**, lanes 13-16). In the presence of KCl, RifI bound efficiently to generate a high-molecular-weight complex on all the DNAs tested, in spite of the absence of PEG200 in the running gel (**Fig. 5A**). In contrast, in the absence of salt, all but T₆G₂₄ exhibited very little or much reduced mobility-shift upon heat denaturation and showed only inefficient binding with RifI. T₆G₂₄, that can adopt a parallel form even in the absence of salt (**Supplementary Fig. S8**), was mobility-shifted upon heat denaturation and showed binding (**Fig. 5B**, lanes 9 and 10). In the presence of NaCl, CEB1_TA and CEB1_TTA showed very little mobility shift after heat treatment, consistent with non-G4 structures indicated by the CD profiles (**Supplementary Fig. S8**), and exhibited much reduced binding with RifI (**Fig. 5B**, lanes 1-4). On the other hand, Htelo1_no_spacer_2 and Htelo4(GGGGGG)_3nt_spacer was significantly mobility-shifted upon heat denaturation in NaCl, suggesting that some structures, most likely anti-parallel G4 structure as speculated from the CD results (**Supplementary Fig. S8**), are formed. However, there was only a low level of binding of RifI to these structures (**Fig. 5C**, lanes 5-8; see also **Supplementary Fig. S10**). These results supported our above speculation and indicated that the topology of G4 affects the efficiency of RifI binding probably through the formation of oligomeric G4s.

RifI may promote association of G4-containing DNAs

We previously proposed that RifI may regulate chromatin architecture by facilitating chromatin loop formation^{8,10}. RifI may hold together chromatin fibers by linking multiple G4 sequences. To explore this possibility, we examined if two G4 DNAs can be simultaneously bound by RifI protein. We have generated a biotin-labeled T₆G₂₄ oligonucleotide, which forms a parallel-type G4 structure³⁰ (**Supplementary Fig. S8**) and is efficiently bound by RifI (**Fig. 2**, lanes 53-56; **Fig. 3**, lanes 31 and 32). We mixed this oligonucleotide with ³²P-labeled T₆G₂₄ or T₆[GA]₁₂ (incapable of forming a G4 structure and not bound by RifI, see lane 34 in **Fig. 3**) in the presence or absence of RifI protein, pulled down the biotin-T₆G₂₄ DNA with streptavidin-beads and examined whether ³²P-labeled DNA is associated with the biotin-T₆G₂₄ DNA. We found that, in the presence of RifI protein, 1.55 % of the input ³²P-labeled T₆G₂₄ DNA was co-pulled down (**Fig. 6A**, lane 8), while ³²P-labeled T₆[GA]₁₂ (not capable of forming G4) was not pulled down under the same condition (**Fig. 6A**, lanes 16-18). A small amount (0.2%) of ³²P-labeled T₆G₂₄ DNA was pulled down even in the absence of RifI protein (**Fig. 6A**, lane 5), suggesting that the G4 DNA can self-associate with each other. This was conducted by washing the pulled down materials with 1M NaCl. When the pulled down materials were washed by binding buffer (50 mM KCl) alone, close to 30% of the input ³²P-labeled G4 DNA was pulled down by biotin-labeled G4 DNA. These results indicate that RifI can promote association of G4 molecules. It is not clear whether this is due to spontaneous disruption and regeneration of multimeric G4 between biotin-T₆G₂₄ and ³²P-T₆G₂₄ or to association between pre-formed G4-structured T₆G₂₄ molecules.

The sequences derived from RifIBS can form G4 structure³¹ and they can also be pulled down with biotin-T₆G₂₄ in the presence of RifI (**Fig. 6B**, lanes 4, 8, 16 and 24). Again, weak association of ³²P-labeled DNA was observed even in the absence of RifI (**Fig. 6B**, lanes 3, 7 and 15). This association is dependent on the presence of G4 structure, since it is not observed with mutant forms of

³²P-labeled DNA that do not form a G4 structure (**Fig. 6B**, lanes 12, 20 and 28). Thus, G4 molecules have potential of self-association.

Rif1 forms oligomers

The ability of Rif1 protein to facilitate the association of G4 structures could result from its potential to form oligomers. The potential oligomerization domains have been identified in the C-terminal segment of mammalian and yeast Rif1 proteins^{22,42,43}. We have analyzed the fission yeast Rif1 protein in size-exclusion chromatography and glycerol gradient centrifugation (**Fig. 7** and **Supplementary Fig. S11**). The full-length Rif1 migrated at three different positions (S, M and F) in glycerol gradient, with estimated subunit compositions of 4, 6, 8 deduced by one method (the Siegel and Monti method)⁴⁴ or 8 and over 10 by another method (the Erickson method)⁴⁵, although it was difficult to make precise estimation due to its extremely large size. The 70 kDa degradation polypeptide, derived from the C-terminal segment, migrates at four positions (S, M1, M2 and F) in glycerol gradient, and its subunit composition was estimated to be 2, 4, 8 and 12. The results indicate that the Rif1 protein exists as various oligomeric forms with a highly elongated shape and suggest a possibility that the oligomeric Rif1 holds together chromatin fibers by binding to multiple G4 sequences through its subunits, promoting their association.

Rif1 may be recruited to telomeres through different mechanisms

It has been known that Rif1 is recruited to telomere in a manner dependent on Taz1, another telomere binding protein⁵. We have examined whether Taz1 stimulates interaction of Rif1 with telomere sequence. For that, we first expressed the full-length Taz1 protein in mammalian cells and purified it (**Supplementary Fig. S12A**). We then used it to examine if it can facilitate the interaction of Rif1 with a typical telomere sequence, GGGTTA. Htelo3, containing repeats of GGGTTA and adopting an anti-parallel G4 (**Supplementary Fig. S8**), was not bound by Rif1 very efficiently, as was shown above (**Fig. 2**, lanes 33-46). We added increasing amount of Taz1 in the absence or presence of Rif1. Taz1 also did not show significant binding to Htelo3 on its own. The presence of both proteins only showed additive complex formation (**Supplementary Fig. S12B**).

We then generated a 21 bp duplex DNA containing repeats of GGTTAC, the fission yeast telomere consensus sequence, and used this DNA as a substrate. Although Rif1 did not bind to this DNA on its own, Taz1 did with K_d of ~30 nM, consistent with the *in vivo* interaction of Taz1 with the telomere. Addition of increasing amount of Rif1 in the presence of Taz1 only slightly increased the amount of the shifted band (**Supplementary Fig. S12C and D**). These results indicate that Taz1 indeed directly binds to the double-stranded telomeric repeat sequences, consistent with the previous report with *in vitro* translated Taz1 protein⁴⁶, but the efficient recruitment of Rif1 to telomere may require some additional factors or telomere chromatin structure.

Discussion

Rif1 is a conserved nuclear protein that appears to play a major role in determining replication timing in both yeast and mammalian cells⁶⁻¹⁰. In the *rif1Δ* mutant of fission yeast, 29.6% of late/ dormant origins (189 out of 638) are fired early in S phase⁶. In yeast, Rif1 does not affect pre-RC formation but inhibits the loading of Cdc45, the step regulated by Cdc7 kinase. Rif1 binding is enriched around the late/ dormant origins on the fission yeast genome, suggesting that Rif1 binding somehow suppresses the pre-RC activation. In mammalian cells, depletion of Rif1 led to the increased level of Mcm phosphorylation mediated by Cdc7 kinase⁸. Rif1 was shown to recruit phosphatase both in yeast and human cells, thereby counteracting the phosphorylation of Mcms catalyzed by Cdc7 kinase¹⁴⁻¹⁸, providing explanation for these earlier observations. It was recently reported that Rif1 facilitates pre-RC formation by protecting ORC1 from degradation¹⁷. We proposed that Rif1 may regulate replication timing by generating the specific chromatin compartments through regulating the chromatin loops^{8,10}. This prediction was supported by the fact that Rif1 binding to chromatin affects the replication initiation over close to 100 kb segment²⁸. Analyses by 4C-seq indeed showed that mouse Rif1 regulates chromatin interactions within and between the replication domains¹².

In yeasts, Rif1 not only binds to telomere but also to selected segments on the chromosome arms^{6,47,48}. Genome-wide analyses of Rif1 binding sites and subsequent analyses of the binding

sequences revealed the presence of a conserved G-rich sequence³¹. Further analyses showed the potential of these sequences to form G4-like structures and specific binding of Rif1 to these non-B DNA structures^{31,40}. In mammalian cells, Rif1 binding is enriched in late-replicating domains, overlapping with LAD (Lamin-associated domains)¹². Mammalian Rif1 was reported to bind to replication fork and Holliday junction DNA²² or to cruciform structures with high specificity⁴⁹. Our analyses indicate that mouse Rif1 also preferentially binds to G4 structures *in vitro*⁵⁰, and a subset of strong Rif1 binding sites on the chromatin are associated with G4-forming sequences (Moriyama, Yoshizawa and Masai, unpublished data). Diverse structures and biological functions of G4 DNA as well as their interacting proteins have been a focus of recent intense researches⁵¹⁻⁵⁴. In this report, we have conducted detailed analyses on the specificity of G4 recognition by fission yeast Rif1 and biochemically characterized this protein. The results suggest a model on how Rif1 may regulate chromatin architecture through its ability to bind simultaneously to multiple G4 and to oligomerize.

Rif1 prefers to bind to oligomeric assembly of G4, the formation of which is potentially facilitated by long runs of guanines

Various single-stranded DNAs known to form G4 structures were examined for their affinity to Rif1. Interestingly, Rif1 did not bind efficiently to telomere-derived sequences (repeats of GGGTTA; **Fig. 2, 3 and 4**). Analyses of derivatives of the non-binders and good binders showed that the presence of long runs of guanine (6G) may be important but the length of the loop may not significantly affect the binding affinity (**Fig. 3 and 4**). However, the presence of 6G is not sufficient for generation of the structure bound by Rif1, since insertion of 3G at one of the GGGT repeats (generating GGGGGGT near the 5'-terminus; T95-2T-GGG) did not improve the binding (**Fig. 3**). In contrast, generation of 6G at the 3'-terminus converted poor binders, Htelo1 and Htelo4, to good binders, Htelo1_no_spacer_2 and Htelo4(GGGGGG), respectively (**Fig. 3 and 4**). Rif1-8, the 70 nt sequence derived from Rif1BS_{1:4255}, formed ladders on PAGE, and these forms were preferentially bound by Rif1⁴⁰ (**Supplementary Fig. S13 and data not shown**). This sequence contains a G5-tract near its 3' end. Thus, the presence of long runs of guanines near the 3'-terminus could be important for generating structures preferred by Rif1 but its location and other sequence context also affects the affinity to Rif1. This conclusion is supported by the presence of 5 or 6 runs of guanine in all the Rif1CS, which appear at least twice within the high-affinity Rif1BS³¹.

We analyzed the numbers of Gs in the G-tracts (n=3 or longer) on Rif1BS on both strands. We aligned the 35 Rif1BS in order of binding efficiency (determined by the intensities of the ChIP peaks) and counted the numbers of Gs in G_{n(n=3 or longer)} in the top 10 and bottom 10 binders. On the G-rich strand, 43.3% and 37.7% were 4G or longer, in the top10 and bottom10 Rif1BS, respectively. On the other hand, on the C-rich strand, 22% and 14.3% were 4G or longer in the top10 and bottom10, respectively (**Supplementary Fig. S14**). These data show that long G-tracts are enriched on the G-rich strand of Rif1BS, and corroborate the above *in vitro* results on Rif1-G4 interactions, showing that long G-tracts would constitute an element required for efficient Rif1 binding.

We also noted that Rif1 selectively binds to the slow-migrating forms generated by heat treatment, but not to the fast-migrating forms of the G4-forming single-stranded DNA, suggesting that Rif1 preferentially binds to the G4 assembly composed of multiple G4-forming sequences or multimerized G4 structures. It has been known that monomeric intramolecular quadruplexes, such as that formed by human telomeric DNA and RNAs, can dimerize by stacking end-to-end. More recently, sequences from the promoter regions of c-kit2 and B-raf or those from an intron of the N-myc gene have been shown by NMR analyses to generate G4 dimers^{55,56}. In these cases, two strands are intertwined, each spanning the entire length of the structures, generating dimeric structures with six or seven consecutive G quartets. It is possible that similar dimeric or oligomeric structures are generated on the Rif1BS-derived sequences which carry multiple long G-tracts.

Those forms that are efficiently bound by Rif1 are generally slow migrating and often appear as smeared bands on PAGE, suggesting that the structures may be oligomers or intermolecular G4 structures. They may not be very stable (partially disrupted during the run on PAGE), or more dynamic than anticipated. Under selective gel electrophoresis conditions, both T₆G₂₄ and Rif1-8, very good binders of Rif1, generate clear ladders of molecules, each of which probably represents a distinct oligomeric form. The ladders are seen even on denaturing polyacrylamide gel after heat denaturation in formamide (**Supplementary Fig. S13 and lane 25 of Supplementary Fig. S16**), suggesting that DNA molecules may be composed of intertwined DNA strands indicative of intrastrand G4 oligomers. Rif1 binding was observed also on simple G-tract sequences, such as 8G and 10G, albeit at a low level,

suggesting the ability of Rif1 to interact with interstrand G4 (**Supplementary Fig. S7**). On the other hands, generation of oligomers through stacking of monomeric forms of intermolecular G4 is also possible, given the self-associating ability of G4 (**Fig. 6**). CD analyses of these single-stranded DNA suggested that topology of G4 may not be a sole determinant for Rif1 recognition (**Supplementary Fig. S8**), although the anti-parallel form may be least preferred by Rif1. This is most clearly shown by the fact the same oligonucleotide adopting different topology under different chemical conditions exhibits differential affinity to Rif1 (higher affinity in “hybrid” conformation than in “mix/ anti-parallel” conformation; **Fig. 5**, **Supplementary Fig. S10** and see also **Supplementary Table S1**).

Rif1BS contains multiple copies of (A)GTG_n to which Rif1 can bind

Although the presence of long G-tracts characterizes Rif1BS, we noted the frequent occurrence of other G-tracts in the vicinity of Rif1CS. Furthermore, these G-tracts (3G or longer) are very frequently preceded by (A)GT. Analyses of the frequency of the dinucleotides preceding G_n(n=3 or longer) within the 1 kb segments surrounding the 35 Rif1BS showed that 42% was GT. This bias was observed only on the G-rich strand on which Rif1CS-G-tract is present, but not on the other C-rich strand (**Supplementary Fig. S15**).

Rif1 binds to oligonucleotides containing repeats of AGTG_n(n=3 or longer) and even to a single copy of AGTGGGG or AGTGGGGG that has been heat-treated. However, these bindings do not depend on the presence of GT, and ACA_n was also bound by Rif1 with similar or better efficiency, showing that sequence GT itself may not be an important determinant for Rif1 recognition *in vitro*. Indeed, Rif1 binds to G10 and to G8, to a small extent (**Supplementary Fig. S7**). The functional significance of this motif needs to be evaluated in the future by mutating the sequences on the genome and examining its effect on Rif1 binding and timing regulation in cells.

Rif1 may promote association of multiple G4 DNAs

Biotinylated T₆G₂₄ can pull down ³²P-labeled T₆G₂₄ DNA in the presence of Rif1 protein (**Fig. 6**). Thus, these experiments indicate the simultaneous binding of Rif1 to multiple DNA molecules. Analyses of Rif1 protein in size exclusion chromatography and glycerol gradient centrifugation indicate formation of oligomers composed of 4 to over 10 protomers (**Fig. 7** and **Supplementary Fig. S11**). The G4 oligonucleotides used in this study, including T₆G₂₄, form G4 structures even without heat treatment (**data not shown**), suggesting that it is not likely that Rif1 facilitates the G4 formation, and our FRET experiments show no evidence for Rif1-mediated stabilization of G4 (Masai, Kanoh and Kakusho, unpublished data). These results support our conclusion that Rif1 promotes association of multiple G4 DNAs rather than facilitating the formation of G4 or stabilizing G4.

We propose that the oligomeric Rif1 may tether multiple chromatin fibers through each subunit binding to different G4, potentially contributing to the formation of chromatin loops in the cells. This process would probably be facilitated by the ability of the N-terminal HEAT repeat segments to cooperatively spread on DNA and to encase DNA ⁵⁷. We speculate that the formation of chromatin loops may be dynamic, dictated by transient and even stochastic interaction of Rif1 and G4 (**Fig. 8**). Unexpectedly, we found that G4 self-associates *in vitro* even in the absence of Rif1. Indeed, biotinylated T₆G₂₄ can pull down not only heat-treated T₆G₂₄ but also Rif1BS-derived single-stranded DNA containing Rif1CS capable of forming G4 structures (**Fig. 6B**). The mutated forms of Rif1BS DNAs are not pulled down, showing that the association is specific to G4-like structures. This interaction is further enhanced by the presence of Rif1 (**Fig. 6**). Thus, an alternative, intriguing possibility is that chromatins make dynamic and stochastic interactions with each other through G4 structures present on the genome, and Rif1 may stabilize or facilitate these interactions (**Fig. 8**). Loss of a specific Rif1 binding site by mutation did not affect the binding of Rif1 to other binding sites including those close to the mutated binding site, as examined by ChIP-seq ³¹, which could suggest that the inter-G4 interactions, if any, may be weak or transient.

In ChIP-chip analyses, we initially identified more than 160 Rif1 binding sites ⁶. In ChIP-seq analyses, we first identified more than 90 Rif1BS, and reduced the numbers to 35 by selecting only the very reliable strong signals ⁵. Since formation and stability of G4 structures are affected by the nucleotide sequence as well as by the microenvironment (salt concentration, transcriptional state *etc.*), there may be many unstable or transient G4 structures on the chromosomes.

We propose that Rif1 may generate a very dynamic and even stochastic chromatin domains. Rif1 bound to a strong site may serve as a hub with which surrounding chromatins interact in a transient and stochastic manner, and this may result in generation of a replication timing domain (**Fig. 8**). This

model fits with recent reports on the effect of Rif1 on chromatin interactions in mammalian cells¹². It was reported that *S. cerevisiae* Rif1 is tethered to nuclear membrane through its palmitoylation⁵⁸. Since the late replicating domains are normally associated with the nuclear periphery, it is of interest whether similar membrane tethering contributes to the Rif1's ability to generate specific chromatin architecture near the nuclear membrane in fission yeast and higher eukaryotes.

All the binding assays in this study have been conducted on G4 structures generated on single-stranded DNAs. However, *in vivo* situations are vastly different. G4 on Rif1BS are generated in a competition with the duplex structure which is more stable. It could be generated during the course of transcription or under the regional topological stress or epigenome modifications. Other proteins, such as Taz1, may affect chromatin interaction of Rif1. Indeed, Rif1 is recruited to telomere through Taz1⁵, probably not through direct DNA binding, and this is consistent with the low affinity of Rif1 to telomere-derived G4 (**Fig. 2, 3, 4** and **Supplementary Fig. S12**). Therefore, there may be multiple pathways for chromatin loading of Rif1 in cells. It would be necessary to analyze the G4 structures generated on a duplex Rif1BS and its interaction with Rif1. It would also be critically important to investigate the *in vivo* structures of Rif1BS DNAs, to demonstrate that G4 is generated at Rif1BS on the chromatin in cells, and to clarify molecular mechanisms of G4 formation in the cells as well as other unknown pathways for chromatin loading of Rif1.

Materials and Methods

Antibodies. The primary antibodies used were M2 (Sigma-Aldrich, F1804) and anti-DDDDK tag MAb (MBL, FLA-1).

Expression and purification of fission yeast Rif1 protein and its derivatives. His₆- and FLAG₃-tagged Rif1 protein or its derivatives were expressed in 293T cells on the ver3-4 vector and purified as previously described^{31,32}. The proteins were further purified with monoQ ion-exchange column or glycerol gradient centrifugation, if necessary.

G4 DNA and Rif1BS DNA. The sequences of the oligonucleotides used in the assays are described in each figure. The oligonucleotides were OPC column-purified. Most of them were further purified by PAGE containing 8M urea. T₆G₂₄ and T₆(GA)₁₂ were always purified from urea-PAGE. The purity of the used oligonucleotides is shown in **Supplementary Fig. S16**. Oligonucleotides were heat-denatured at 96 °C for 3 min and gradually cooled down to room temperature in 50 mM KCl and 40% PEG200^{34,58}. In some experiments, oligonucleotides were heat-denatured and cooled down in 60 mM cacodylate buffer with 60 mM KCl.

Expression and purification of fission yeast Taz1 protein and the substrate DNA. The Taz1 coding frame was amplified by RT-PCR from fission yeast cells by using the primer set (Taz1-N[Bam]: CGGGATCCATGATAAGCGTGCAAAGTACAGAAA and Taz1-C[Bam] CGGGATCCAGATTGATAATTAACAAGCTCTT) and was cloned into ver3-4 vector and N-terminally His₆ and C-terminally FLAG₃-tagged Taz1 was expressed in mammalian cells, and purified by anti-Flag affinity column followed by nickel column. The 21 bp ³²P-end labeled double-stranded DNA containing Taz1 binding site (³²P-TTACAGGTTACAGGTTACAGG/CCTGTAACCTGTAACCTGTAA) was generated by annealing of the two oligonucleotides, purified from PAGE, and used as a substrate for DNA binding assays.

Pull-down assays with biotinylated DNA. 5'-biotinylated T₆G₂₄ oligonucleotide (2.5 pmole) was incubated in gel shift assay buffers (40 mM Hepes-KOH [pH 7.6], 50 mM KCl, 1 mM EDTA, 10 % glycerol and 0.01 % Triton X-100) in the presence or absence of Rif1 protein (140 fmole) with 1-2.5 pmole of ³²P-labeled single-stranded DNA that Rif1 can bind to. Both biotinylated DNA and ³²P-DNA had been heat-treated and denatured in KCl-PEG200 before used in the assays. After incubation for 30 min at room temperature, Dynabeads M-280 Streptavidin was added and the beads were extensively washed by the same buffer or that containing 1 M NaCl. Beads were resuspended in 80 % formamide containing 1 mM EDTA, boiled for 3 min and run on 12 % PAGE containing 8M urea in 0.5x TBE.

Analyses of DNA and protein-DNA complexes on polyacrylamide gels. Labeled DNA fragments or cold DNA fragments were mixed with purified proteins in reaction mixtures (10 μ l or 20 μ l) containing 40 mM Hepes-KOH (pH 7.6), 50 mM KCl, 1 mM EDTA, 10 % glycerol, and 0.01 % Triton X-100. After incubation at room temperature for 30 min, the reaction mixtures were directly applied onto a polyacrylamide gel. DNA and Protein-DNA complexes were analyzed on polyacrylamide gels prepared in 1x TBE, 50 mM KCl and 40 % PEG200 or on those prepared in 1x TBE, 50 mM KCl (or other salt or no salt, where indicated) and 10% glycerol. Other gel electrophoresis conditions were also used, which are indicated in figure legends. For analyses of denatured DNA, DNA were heat-denatured in 95 % formamide containing 5 mM EDTA (and LiCl where indicated) and were analyzed on PAGE containing 8 M urea in 0.5x TBE. The experiments were performed at least two times (in most cases three times or more) independently, and standard deviation and p-value by two-tailed student's t-test were determined and presented, where indicated.

Analytical gel-filtration analyses of Rif1 protein. Fifty μ l of a peak fraction of the Ni²⁺-NTA column (0.22 μ m-filtrated) was applied to pre-equilibrated Superose 6 PC3.2/30 column (GE Healthcare), and run in 20 mM Tris-HCl [pH7.8], 150 mM KCl, 1 mM EDTA, 0.5 mM DTT, 10 % glycerol and protease inhibitors [cOmplete[™] Protease Inhibitor Cocktail; Roche] at 40 μ l/min at 4°C. Eighty μ l-fractions were collected and subjected to SDS-PAGE followed by western blotting with anti-DDDDK antibody (FLA-1, MBL Co.) to detect the C-terminal FLAG₃ tag. Molecular size standards (GE Healthcare) were chromatographed under the same condition, and retention volumes of all the standards were determined by monitoring UV absorbance at 260 nm.

Analytical glycerol-gradient centrifugation analyses of Rif1 protein. In 2.2-ml centrifuge tubes, 15-36% glycerol-gradient was made in 20 mM Tris-HCl (pH7.8), 150 mM KCl, 1 mM EDTA, 0.5 mM DTT and cOmplete[™] Protease Inhibitor Cocktail (Roche). Twenty μ l of a peak fraction of the Ni²⁺-NTA column (0.22 μ m-filtrated) was diluted 5-fold, and layered on top of the gradient. One hundred μ l of molecular size standards was similarly layered on top of another tube. These tubes were centrifuged at 40,000 rpm for 16 hr in Beckman TLS55 rotor at 4°C. 0.1 ml-fractions were removed from top to bottom, and subjected to SDS-PAGE followed by anti-DDDDK blotting. The sedimentation pattern of molecular size standards was analyzed by SDS-PAGE followed by silver-staining.

Analysis of hydrodynamic behaviors of Rif1 and its degradation products. Both *S*- and *R*_s-values of Rif1 and its degradation products were calculated from the data of glycerol-gradient centrifugation and gel-filtration after linear approximation to those of molecular size standards. Then, their native molecular weights were estimated as previously reported^{45,46}.

Fission yeast strains, medium and general techniques. All strains used in this study were previously described⁶. Methods for genetic and biochemical analyses of fission yeast have been described previously^{60,61}. YES media containing 0.5 % yeast extract, 3 % glucose and 0.1 mg/ml each of adenine, uracil, leucine, lysine and histidine were used for cell culture and YES plates were made by adding 2 % agar to YES media. 0.2 mg/ml G418 was added to YES medium for selection of kanMX. For 5-fluoroorotic acid (5-FOA) selection, 0.1 mg/ml 5-FOA was added to the media containing 6.3 g/L synthetic dextrose minimal medium (SD), 2 % glucose and 0.1 mg/ml each of adenine, uracil and leucine.

Circular Dichroism (CD) spectrometry. The oligonucleotides were diluted to 2 μ M in 50 mM Tris-HCl (pH 7.5) without salt, or that with 50 mM KCl or that with 50 mM NaCl. Subsequently, these solutions were annealed by heating at 99 °C for 5 min, then slowly cooled to room temperature, and incubated overnight. Circular Dichroism (CD) spectra were recorded on a J-720 spectropolarimeter (JASCO, Tokyo, JAPAN) using a quartz cell (Agilent, microcell 50 μ L, 10 mm optical path length) with scanning speed of 500 nm/min and a response time of 1 sec over a wavelength range of 230-320 nm. The CD spectra shown are representatives of five averaged scans taken at 25 °C or at various temperatures, as indicated in the figures. See **Supplementary Fig. S8 and S9** for data and **Supplementary Table S1** for summary.

References

1. Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N. & Oda, M. Eukaryotic chromosome DNA replication: where, when, and how? *Annu. Rev. Biochem.* **79**, 89-130 (2010)
2. Pope, B.D. *et al.* Topologically associating domains are stable units of replication-timing regulation. *Nature*. **515**, 402-405 (2014)
3. Rivera-Mulia, J.C. & Gilbert, D.M. Replication timing and transcriptional control: beyond cause and effect-part III. *Curr. Opin. Cell Biol.* **40**, 168-178 (2016)
4. Hardy, C.F.J., Sussel, L. & Shore, D. A RAP1-interacting protein involved in silencing and telomere length regulation. *Genes. Dev.* **6**, 801-814 (1992)
5. Kanoh, J. & Ishikawa, F. spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. *Curr. Biol.* **11**, 1624-1630 (2001)
6. Hayano, M. *et al.* Rif1 is a global regulator of timing of replication origin firing in fission yeast. *Genes. Dev.* **26**, 137-150 (2012)
7. Peace, J.M., Ter-Zakarian, A. & Aparicio, O.M. Rif1 regulates initiation timing of late replication origins throughout the *S. cerevisiae* genome. *PLoS One*. **9**, e98501 (2014)
8. Yamazaki, S. *et al.* Rif1 regulates the replication timing domains on the human genome. *EMBO J.* **31**, 3667-3677 (2012)
9. Cornacchia, D. *et al.* Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J.* **31**, 3678-3690 (2012)
10. Yamazaki, S., Hayano, M. & Masai, H. Replication timing regulation of eukaryotic replicons: Rif1 as a global regulator of replication timing. *Trends Genet.* **29**, 449-460 (2013)
11. Dimitrova, D.S. & Gilbert, D.M. The spatial position and replication timing of chromosomal domains are both established in early G1 phase. *Mol. Cell.* **4**, 983-993 (1999)
12. Foti, R. *et al.* Nuclear Architecture Organized by Rif1 Underpins the Replication-Timing Program. *Mol. Cell.* **61**, 260-273 (2016)
13. Sreesankar, E., Senthilkumar, R., Bharathi, V., Mishra, R.K. & Mishra, K. Functional diversification of yeast telomere associated protein, Rif1, in higher eukaryotes. *BMC Genomics*. **13**, 255 (2012)
14. Hiraga, S. *et al.* Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex. *Genes. Dev.* **28**, 372-383 (2014)
15. Mattarocci, S. *et al.* Rif1 controls DNA replication timing in yeast through the PP1 phosphatase Glc7. *Cell Rep.* **7**, 62-69 (2014)
16. Davé, A., Cooley, C., Garg, M. & Bianchi, A. Protein phosphatase 1 recruitment by Rif1 regulates DNA replication origin firing by counteracting DDK activity. *Cell Rep.* **7**, 53-61 (2014)
17. Hiraga, S.I. *et al.* Human RIF1 and protein phosphatase 1 stimulate DNA replication origin licensing but suppress origin activation. *EMBO Rep.* **18**, 403-419 (2017)
18. Alver, R.C., Chadha, G.S., Gillespie, P.J. & Blow, J.J. Reversal of DDK-Mediated MCM Phosphorylation by Rif1-PP1 Regulates Replication Initiation and Replisome Stability Independently of ATR/Chk1. *Cell Rep.* **18**, 2508-2520 (2017)

19. Silverman, J., Takai, H., Buonomo, S.B., Eisenhaber, F. & de Lange, T. Human Rif1, ortholog of a yeast telomeric protein, is regulated by ATM and 53BP1 and functions in the S-phase checkpoint. *Genes. Dev.* **18**, 2108-2119 (2004)
20. Buonomo, S.B., Wu, Y., Ferguson, D. & de Lange, T. Mammalian Rif1 contributes to replication stress survival and homology-directed repair. *J. Cell. Biol.* **187**, 385-398 (2009)
21. Xu, L. & Blackburn, E.H. Human Rif1 protein binds aberrant telomeres and aligns along anaphase midzone microtubules. *J. Cell. Biol.* **167**, 819-830 (2004)
22. Xu, D. *et al.* Rif1 provides a new DNA-binding interface for the Bloom syndrome complex to maintain normal replication. *EMBO J.* **29**, 3140-3155 (2010)
23. Chapman, J.R. *et al.* RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Mol. Cell.* **49**, 858-871 (2013)
24. Di Virgilio, M. *et al.* Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. *Science*. **339**, 711-715 (2013)
25. Escribano-Díaz, C. *et al.* A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol. Cell.* **49**, 872-883 (2013)
26. Zimmermann, M., Lottersberger, F., Buonomo, S.B., Sfeir, A. & de Lange, T. 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science*. **339**, 700-704 (2013)
27. Feng, L., Fong, K.W., Wang, J., Wang, W. & Chen, J. RIF1 counteracts BRCA1-mediated end resection during DNA repair. *J. Biol. Chem.* **288**, 11135-11143 (2013)
28. Hiraga, S.I. *et al.* Budding yeast Rif1 binds to replication origins and protects DNA at blocked replication forks. *EMBO Rep* **19**, pii: e46222. (2018)
29. Munden, A. *et al.* Rif1 inhibits replication fork progression and controls DNA copy number in *Drosophila*. *Elife* **7**, pii: e39140. (2018)
30. Seller, C.A. & O'Farrell, P.H. Rif1 prolongs the embryonic S phase at the *Drosophila* mid-blastula transition. *PLoS Biol.* **16**, e2005687. (2018)
31. Kanoh, Y. *et al.* Rif1 binds to G quadruplexes and suppresses replication over long distances. *Nat. Struct. Mol. Biol.* **22**, 889-897 (2015)
32. Uno, S., You, Z. & Masai, H. Purification of replication factors using insect and mammalian cell expression systems. *Methods* **57**, 214-221 (2012)
33. Sengar, A., Heddi, B. & Phan, A.T. Formation of G-quadruplexes in poly-G sequences: structure of a propeller-type parallel-stranded G-quadruplex formed by a G₁₅ stretch. *Biochemistry* **53**, 7718-7723 (2014)
34. Miyoshi, D., Nakao, A. & Sugimoto, N. Molecular crowding regulates the structural switch of the DNA G-quadruplex. *Biochemistry* **41**, 15017-15024 (2001)
35. Collie, G. W. & Parkinson, G. N. The application of DNA and RNA G-quadruplexes to therapeutic medicines. *Chem. Soc. Rev.* **40**, 5867-5892 (2011)
36. Rezler, E. M. *et al.* Telomestatin and Diseleno Sapphyrin Bind Selectively to Two Different Forms of the Human Telomeric G-Quadruplex Structure. *J. Am. Chem. Soc.* **127**, 9439-9447 (2005)

37. Paramasivan, S., Rujan, I. & Bolton, P. H. Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods*. **43**, 324-331 (2007)
38. Karsisiotis, A. I. *et al.* Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chem. Int. Ed.* **50**, 10645-10648 (2011)
39. Villar-Guerra, R., Trent, J. O. & Chaires, J. B. G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew. Chem. Int. Ed.* **57**, 7171-7175 (2018)
40. Masai, H. *et al.* Molecular architecture of G-quadruplex structures generated on duplex Rif1-binding sequences. *J Biol Chem.* **293**, 17033-17049 (2018)
41. Ou, T.M. *et al.* G-quadruplexs: targets in anticancer drug design. *ChemMedChem*, **3**, 690-713 (2008)
42. Shi, T. *et al.* Rif1 and Rif2 shape telomere function and architecture through multivalent Rap1 interactions. *Cell*. **153**,1340-1353 (2013)
43. Kobayashi, S. *et al.* Both a unique motif at the C-terminus and an N-terminal HEAT repeat contribute to G-quadruplex binding and origin regulation by the Rif1 protein. *Mol Cell Biol.* **39**. pii: e00364-18 (2019)
44. Siegel, L. M., & Monty, K. J. Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. *Biochim. Biophys. Acta.* **112**, 346-362 (1966)
45. Erickson, H. P. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol. Proced. Online.* **11**, 32-51 (2009)
46. Spink, K.G., Evans, R.J., & Chambers, A. Sequence-specific binding of Taz1p dimers to fission yeast telomeric DNA. *Nucleic Acids Res.***28**, 527-533. (2000)
47. Masai, H. *et al.* Telomere-binding factors in the regulation of DNA replication. *Genes Genet Syst.* **92**, 119-125 (2017)
48. Hafner, L. *et al.* Rif1 Binding and Control of Chromosome-Internal DNA Replication Origins Is Limited by Telomere Sequestration. *Cell Rep.* **23**, 983-992 (2018)
49. Sukackaite, R. *et al.* Structural and biophysical characterization of murine rif1 C terminus reveals high specificity for DNA cruciform structures. *J. Biol. Chem.* **289**, 13903-13911 (2014)
50. Moriyama, K., Yoshizawa-Sugata, N. & Masai, H. Oligomer formation and G-quadruplex binding by purified murine Rif1 protein, a key organizer of higher-order chromatin architecture. *J. Biol. Chem.* **293**, 3607-3624 (2018)
51. Maizels, N. & Gray, L.T. The G4 genome. *PLoS Genet.* **9**, e1003468 (2013)
52. Maizels, N. G4-associated human diseases. *EMBO Rep.* **16**, 910-922 (2015)
53. Rhodes, D. & Lipps, H.J. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **43**, 8627-8637 (2015)
54. Hänsel-Hertsch, R., Di Antonio, M. & Balasubramanian, S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell Biol.* **18**, 279-284 (2017)
55. Wei, D. *et al.* Crystal structure of a promoter sequence in the B-raf gene reveals an intertwined dimer quadruplex. *J. Am. Chem. Soc.* **135**, 19319-19329 (2013)

56. Kuryavyi, V., Phan, A.T. & Patel, D.J. Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. *Nucleic Acids Res.* **38**, 6757-6773 (2010)
57. Mattarocci, S. *et al.* Rif1 maintains telomeres and mediates DNA repair by encasing DNA ends. *Nat. Struct. Mol. Biol.* **24**, 588-595 (2017)
58. Park, S. *et al.* Palmitoylation controls the dynamics of budding-yeast heterochromatin via the telomere-binding protein Rif1. *Proc Natl Acad Sci USA* **108**, 14572–14577 (2011)
59. Zheng, K.W., Chen, Z., Hao, Y.H. & Tan, Z. Molecular crowding creates an essential environment for the formation of stable G-quadruplexes in long double-stranded DNA. *Nucleic Acids Res.* **38**, 327-338 (2010)
60. Alfa, C. Fantes, P. Hyams, J. McLeod, M. & Warbrick, E. Experiments with fission yeast: a laboratory course manual. Cold. Spring. Harbor. Laboratory. Press, Plainview, NY. (1993)
61. Moreno, S. Klar, A. & Nurse, P. Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*. *Methods Enzymol.* 194:795-823 (1991)

Acknowledgments

We thank members of our laboratory for helpful discussion. We thank Mayumi Shindo of our institute for mass spectrometry analyses. This work was supported by JSPS KAKENHI (Grant-in-Aid for Scientific Research (A) [Grant Numbers 23247031 and 26251004] and Grant-in-Aid for Scientific Research on Priority Areas [“non-coding RNA” and “Genome Adaptation”; Grant Numbers 24114520 and 25125724, respectively] to H.M.

Author Contributions

H.M. conducted various DNA binding analyses, and mutant constructions. R.F. expressed and purified the mutant and full-length fission yeast Rif1 proteins, analyzed them biochemically, did gel-shift assays, and helped plasmid constructions. N.K. conducted gel shift assays and pull-down assays. Y.K. conducted genetic analyses of tagged Rif1 protein. K.M. conducted oligomer analyses of Rif1 protein. Y.M. conducted CD analyses and generated G4 binding ligands. K.I. helped the analyses of Rif1BS and K.N. supported the research. H.M. conceived, designed the experiments, and wrote the paper.

Additional information

Supplementary information accompanies this paper at xxxxx.

Competing Interests

The authors declare no competing interest, either financial or non-financial.

Data availability statement.

All data generated and analyzed in this study are included in this published article and its Supplementary Information.

Legends to Figures

Figure 1.

His₆-Rif1-Flag₃ protein is functional.

A and B. The wild-type or *rif1*Δ harboring pREP81 vector, or *rif1*Δ harboring pREP81-Rif1 or pREP81-His₆-Rif1-Flag₃ were grown in SD+2 μM thiamine, and cells were washed with EMM without thiamine and transferred to the EMM medium containing the indicated concentration (2 μM or 5 μM) of thiamine for 16 hrs. **A.** The expression levels of Rif1 and tubulin were examined by western blotting. **B.** Cell cycle was analyzed by FACS. **C.** The 5-fold serial dilution of the cells, indicated in the figure, was spotted on EMM plate containing 5 or 15 μM thiamine. The plates were incubated at 25 °C or 30 °C. The photographs were taken at 4 days after plating. **D.** Function in telomere length regulation. Genomic DNA was extracted from the cells indicated, digested with *Eco*RI and electrophoresed on 1.2 % agarose in TAE buffer. The digested DNA was transferred to nitrocellulose membrane and fixed with UV irradiation, followed by hybridization with ³²P-labeled telomere DNA. Left, EtBr-stained gel; right, Southern blotting of the same gel. M: molecular weight marker (ϕX174 DNA digested by *Hae*III). The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top.

Figure 2.

Binding of Rif1 to various single-stranded DNAs derived from sequences known to form specific G4 configuration.

A. Single-stranded oligonucleotide DNAs (0.2 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated in the presence of increasing amounts of Rif1 protein (0, 0.1, 1 and 10 fmole [0, 0.01, 0.1 and 1 nM, respectively] of the Rif1 full-length polypeptide; the preparation used contains ~10-fold molar excess of degradation products [mainly a 70 kDa polypeptide derived from the C-terminus]) and were analyzed on 12% PAGE (0.5x TBE, 40% PEG200 and 50 mM KCl). List of oligonucleotides used in the assays and their sequences are shown. The schematic drawing of the structure of each G4 and its topology is shown, where known (see **Supplementary Fig. S18** for the summary of the structures and references). The G-tracts longer than 3 in Pu24T, CEB1 and 93del are highlighted in red. The graph shows quantification of the Rif1 binding to each oligonucleotide. The values were calculated by dividing the radioactivity of the shifted bands (complex) by the sum of the free G4-structured DNA and the shifted bands (see also the **Supplementary Fig. S1** and its legend). All the binding assays were conducted in separate experiments two times or more, with similar results, and only the representative data are presented. See **Supplementary Fig. S2** for additional data related to this figure. M: molecular weight marker (ϕX174 DNA digested by *Hae*III). The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top.

Figure 3

Binding of Rif1 to various single-stranded DNAs and their derivatives: effect of spacer sequence and runs of guanine on binding of Rif1.

The single-stranded oligonucleotide DNAs (0.2 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40 % PEG200, were incubated in the absence (-) or presence (+) of Rif1 protein (10 fmole [1nM] of the Rif1 full-length polypeptide; the preparation used contains ~10-fold molar excess of degradation products), and were analyzed on 12% PAGE (1x TBE, 50 mM KCl and 40 % PEG200). List of oligonucleotides used in the assays and their sequences are shown below the panels. The relevant residues for modification are highlighted in red. The graph shows quantification of the Rif1 binding to each oligonucleotide with error bars, conducted as described in the legend to Fig. 2. All the binding assays were conducted in separate experiments two times or more with similar results, and one of the representative data are presented. The results of the same, but independent assays are shown in **Supplementary Fig. S3**. The oligonucleotides showing >50%, >16%, and >6% mobility-shift at 1nM Rif1 were classified as +++, ++, + for Rif1 binding. M: molecular weight marker (ϕX174 DNA digested by *Hae*III). The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top.

Figure 4.

Binding of Rif1 to various single-stranded DNAs and their derivatives: stimulation of Rif1 binding by runs of guanine at the 3'-end of the DNA.

The single-stranded oligonucleotide DNAs (0.5 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated in the presence of increasing amounts of Rif1 protein. With T₆G₂₄ and T₆(GA)₁₂, 10 fmole (1 nM) of Rif1 was added (+). Samples were analyzed on 12% PAGE (1× TBE, 50 mM KCl and 40% PEG200). List of oligonucleotides used in the assays and their sequences are shown below the panels. The relevant residues for modification are highlighted in red. The topology, as determined by CD analyses (**Supplementary Fig. S8**), is indicated on top the lanes for each oligonucleotide. The graph shows quantification of the Rif1 binding to each oligonucleotide, conducted as described in the legends to Fig. 2. M: molecular weight marker (ϕX174 DNA digested by *Hae*III). The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top. The quantification of the binding represents the average of three independent experiments with error bars. The background in the absence of Rif1 protein is subtracted. **The results of the same, but independent assays are shown in Supplementary Fig. S4.**

Figure 5.

Effect of salt on Rif1 binding to selected G4 oligonucleotides whose topologies change in response to salt.

The single-stranded oligonucleotide DNAs (0.25 pmole) as shown, which had been heat-denatured and renatured in 50 mM Tris-HCl (pH 7.5) with 50 mM KCl (**A**), 50 mM NaCl (**B**) or no salt (**C**), were incubated in the presence (+, 10 fmole [1 nM]) of the Rif1 full-length polypeptide; the preparation used contains ~10-fold molar excess of degradation products as well) or absence (-) of Rif1 protein. Samples were analyzed on 8 % PAGE (1x TBE, 10% glycerol) containing 50 mM KCl (**A**), 50 mM NaCl (**B**) or no salt (**C**). The Htelo1_no_spacer_2 and Htelo4(GGGGGG)_3nt_spacer adopt mix or hybrid-type topology, respectively, in KCl, and these forms are bound by Rif1 (**A**). However, they adopt anti-parallel type and are not efficiently bound by Rif1 in NaCl (**B**). The graph shows quantification of the Rif1 binding to each oligonucleotide, conducted as described in the legend to Fig. 2. M: molecular weight marker (ϕX174 DNA digested by *Hae*III). The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top. The quantification of the binding represents the average of three independent experiments with error bars. The background in the absence of Rif1 protein is subtracted. *, p<0.1; **, p<0.05. **The results of the same, but independent assays are shown in Supplementary Fig. S10.**

Figure 6.

Rif1 can facilitate association of G4 DNA molecules.

A. Increasing amounts of biotin-labeled T₆G₂₄ DNA (denatured and renatured in 50 mM KCl and 40% PEG200; lanes 2, 6, 11 and 15, no biotin-labeled T₆G₂₄; lanes 3, 7, 12 and 16, 1 pmole; lanes 4, 8, 13 and 17, 2.5 pmole; lanes 5, 9, 14 and 18, 5 pmole) were mixed with ³²P-labeled DNA templates (1.25 pmole [denatured and renatured in 50 mM KCl and 40% PEG200]; lanes 1-9, T₆G₂₄; lanes 10-18, T₆[GA]₁₂) in the presence (lanes 6-9 and 15-18) or absence (lanes 2-5 and 11-14) of Rif1 protein (100 fmole). Biotin-labeled DNA was pulled down by streptavidin beads, and washed with buffer containing 1M NaCl before resuspended in formamide dye and boiling. Lanes 1 and 10, 5% of the input ³²P DNA. Note that exactly same amount of ³²P-labeled T₆G₂₄ and T₆[GA]₁₂ DNA was used. However, the extent of ³²P end-labeling was four times less efficient with T₆[GA]₁₂ DNA than with T₆G₂₄ for some unknown reason. Lower panels show another set of experiments in which Biotin-labeled T₆G₂₄ (0.625 and 2.5 pmole) was added in the presence or absence of Rif1 protein, as indicated in the figure. The pulled down materials were washed with binding buffer. Fractions of the pulled down materials relative to the input are indicated under each lane. The drawings (right) schematically represent the procedure of the experiments. **B.** Pull-down of biotin-labeled T₆G₂₄ DNA (2.5 pmole) by streptavidin beads was conducted in the presence of various ³²P-labeled DNA templates (1.25 pmole), as indicated. Rif1 protein (100 fmole) was also present, where indicated (+). The pulled down materials were washed with 1M NaCl before resuspended in formamide dye and boiling. I: 5% of the input ³²P DNA. M: molecular weight marker (ϕX174 DNA digested by *Hae*III). The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 nt, from the top. In both **A** and **B**, the pulled down ³²P-labelled DNAs were analyzed on 12% PAGE containing 8M urea (in 0.5x TBE).

Figure 7.

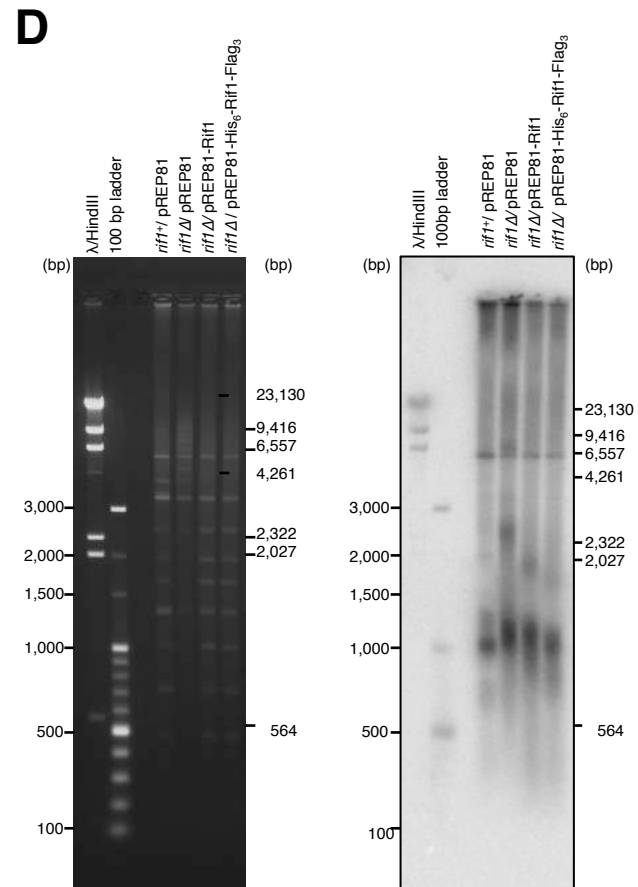
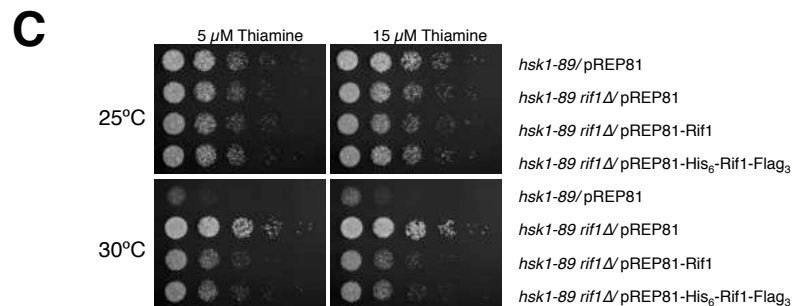
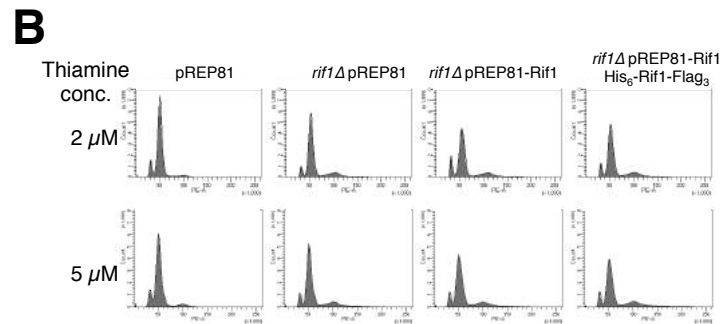
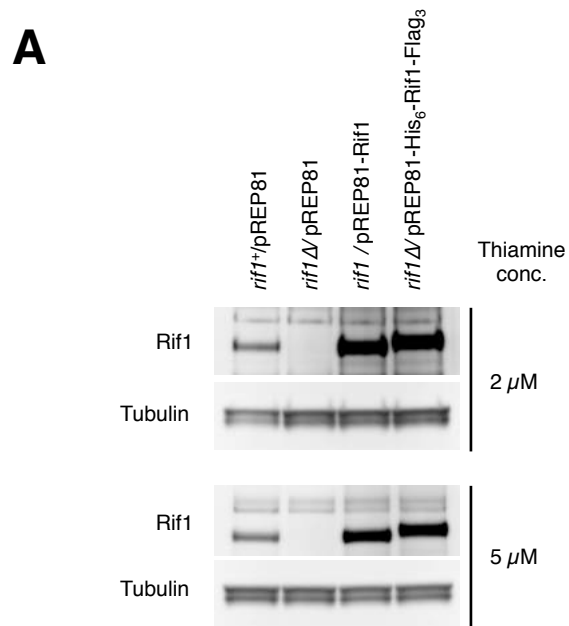
Physicochemical analyses of oligomeric states and molecular shapes of Rif1 polypeptides.

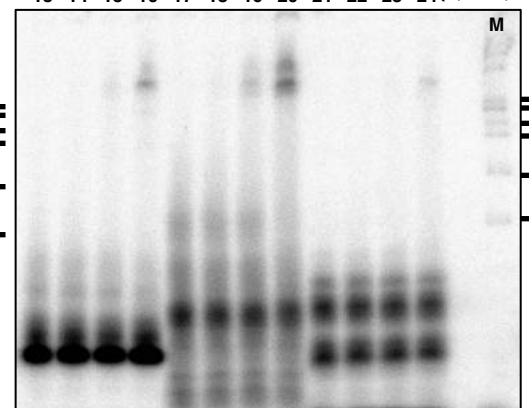
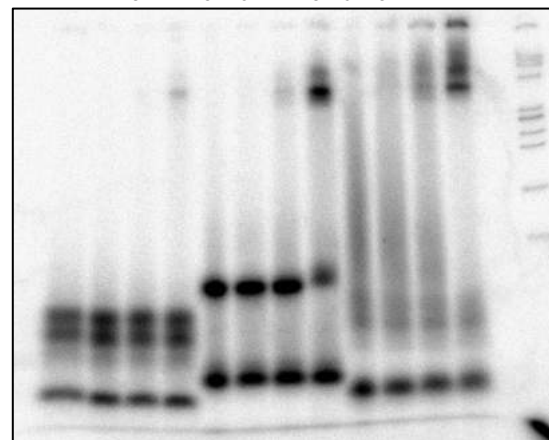
Upper: Size-exclusion (gel-filtration) patterns of the partially purified fraction of His₆-Rif1-FLAG₃ (containing a ~70kDa degradation product derived from the C-terminal segment). Retention volume of the molecular size standard was determined by monitoring A₂₆₀. Lower: Sedimentation pattern of the same fraction through glycerol-gradient centrifugation. The three peaks of the full-length Rif1 are marked as F, M and S in order of sedimentation speed. The four peaks of the ~70 kDa polypeptide are marked similarly as F, M2, M1 and S. Sedimentation pattern of molecular size standards was analyzed by SDS-PAGE followed by silver-staining (data not shown). In both panels, the positions of size markers and those of the full-length and degradation polypeptide are shown by red, green and blue *filled arrowheads*, respectively.

Figure 8.

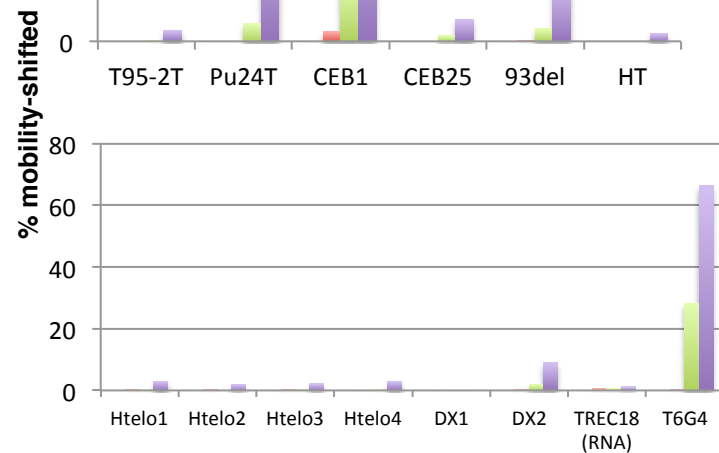
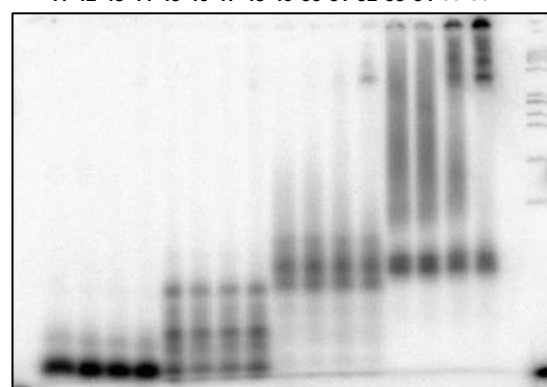
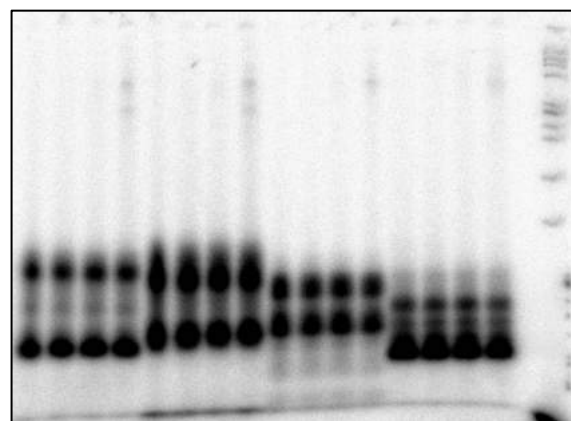
A model on actions of G4 and Rif1 in chromatin organization

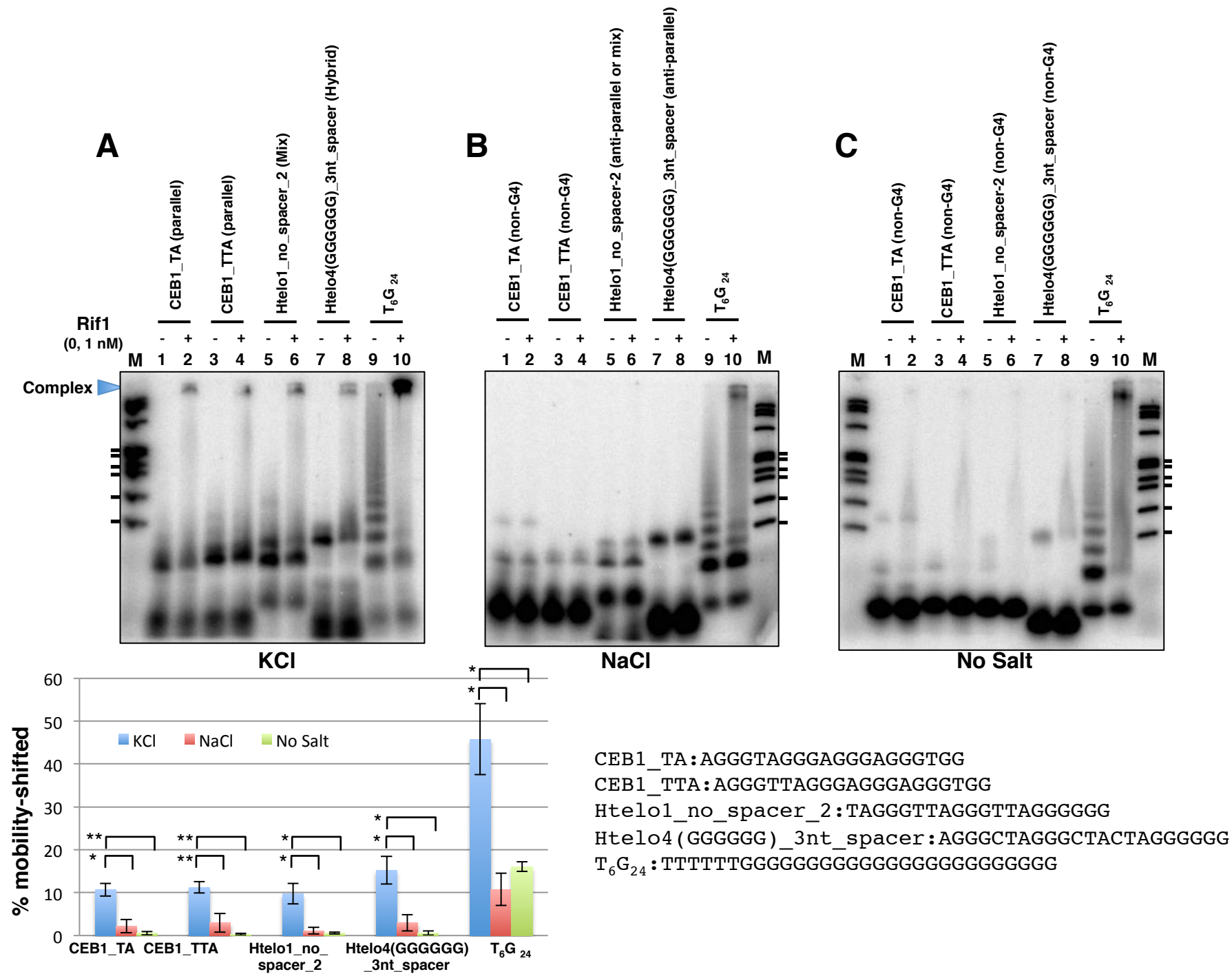
G4 structures generated on the genomes may stochastically interact with each other through their self-association ability. A strong binder of Rif1 (indicated by pink circles), which probably forms more stable G4, would be recognized and bound by Rif1. This may stabilize the transient interactions between multiple G4s through its G4 binding and oligomerization abilities, leading to formation of chromatin compartment related to replication timing. The length of chromatin loop in the drawing is not to the actual size, which should be much larger.

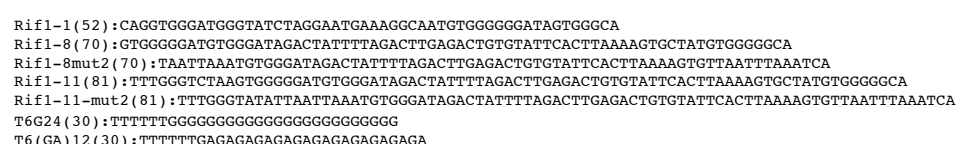
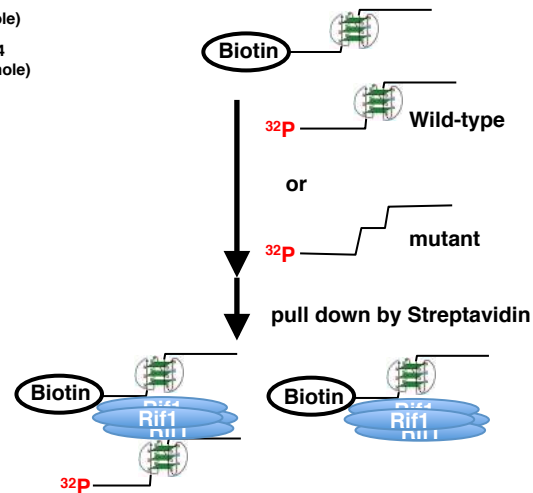


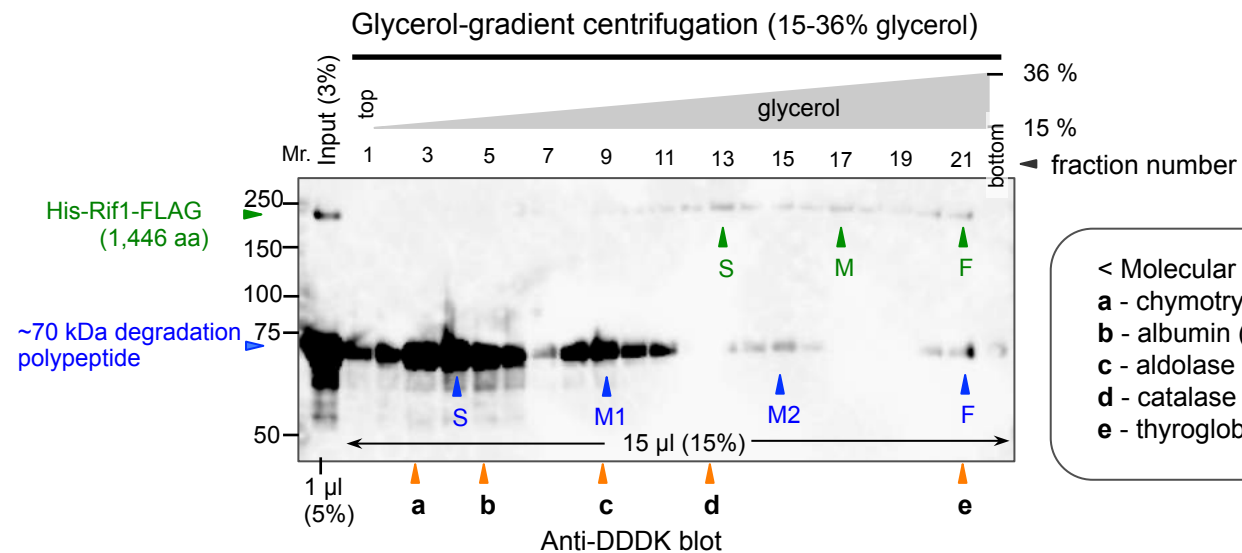
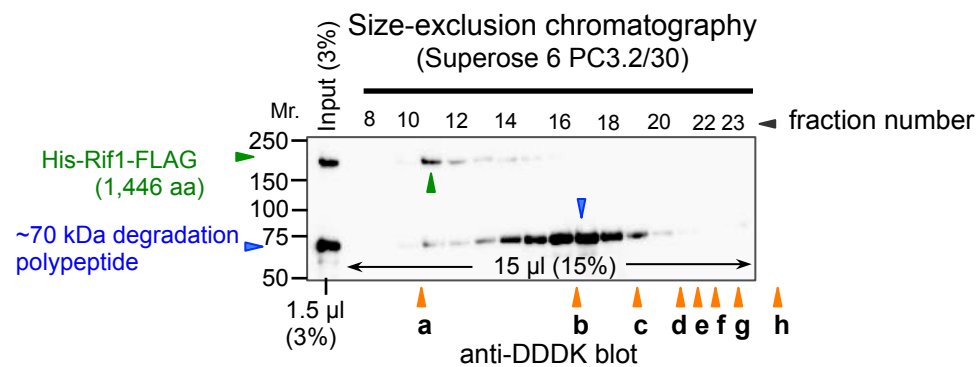


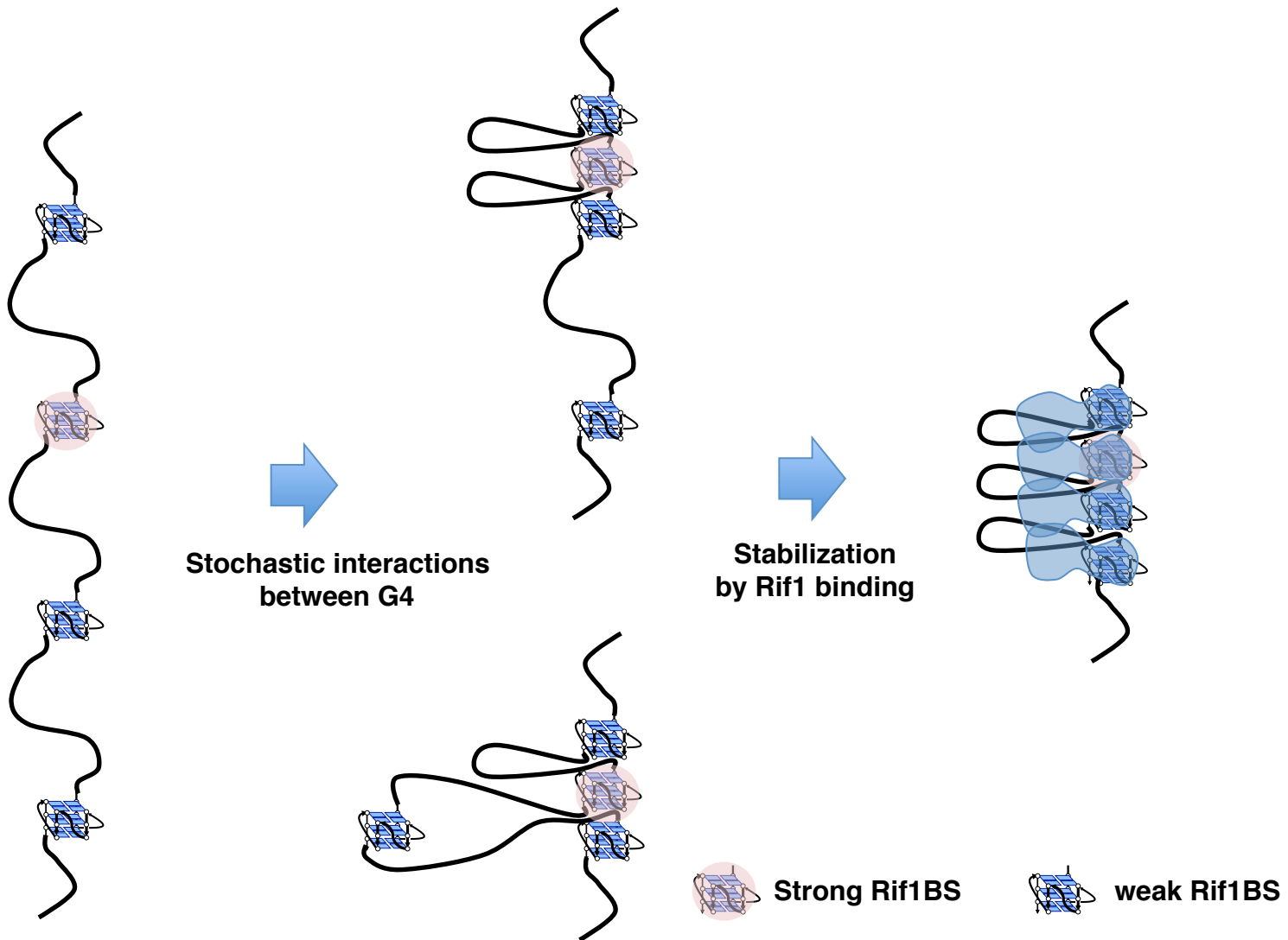
Transcription Factor	shifted
T95-2T	4
Pu24T	6
CEB1	4
CEB25	2
93del	4
HT	3











Supplementary Information for

Title:

Rif1 promotes association of G-quadruplex (G4) by its specific G4 binding and oligomerization activities

Authors:

***Hisao Masai, Rino Fukatsu, Naoko Kakusho, Yutaka Kanoh, Kenji Moriyama, Yue Ma, Keisuke Iida, Kazuo Nagasawa**

Correspondence should be addressed to

Hisao Masai

E-mail: masai-hs@igkakuken.or.jp

Contents

Legends to Supplementary Figures

Supplementary Table S1

Supplementary Figures (Supplementary Figures S1-S16)

Legends to Supplementary Figures

Supplementary Figure S1.

Comparison of DNA binding activities of the full-length and 70 kDa C-terminal polypeptide of Rif1.

A. Protein fractions containing the full-length or 70 kDa polypeptide of Rif1 were analyzed on 5-20% gradient SDS-PAGE and proteins were stained by silver. Lane 1, full-length Rif1 protein purified by consecutive steps of anti-FLAG affinity column, nickel column and monoQ column chromatography (a different fraction from that used in Fig. 1). Lane 2, the 70 kDa polypeptide fraction purified by glycerol gradient centrifugation of FLAG-column purified material. Lane 3, the 70 kDa polypeptide fraction purified by monoQ column chromatography of the material enriched by consecutive steps of anti-FLAG affinity column and nickel column. Closed and open triangles indicate the full-length and the 70 kDa C-terminal polypeptide of Rif1, respectively. In the fractions for lanes 2 and 3, the full-length Rif1 was not detected even with western blotting (not shown). The protein concentrations were estimated by comparison with the BSA marker that was run in the same gel (not shown). Lane 1 and lanes 2/3 were from different gels; therefore, the band intensities cannot be directly compared. Original full-length gels are shown in Supplementary Fig. S17. **B.** The end-labeled T₆G₂₄ DNA (0.25 pmole), which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated with increasing amounts of Rif1 protein, as shown, and were analyzed on 8% PAGE containing 50 mM KCl and 40% PEG200 in 1x TBE. **C.** Band intensities of the gel were measured and relative binding efficiency was estimated as described in the legend to Fig. 2. The concentrations of the protein on the X-axis are based on those of the full-length polypeptide alone. However, the preparation used contains four-fold molar excess of the C-terminal 70 kDa polypeptide. If this polypeptide is capable of binding to G4 and is also taken into account, the concentrations of the polypeptides could be as much as five-times larger. Even that is the case, the K_d would be ~1.5 nM (assuming that the 70 kDa polypeptide binds to the G4 substrate as efficiently as the full-length). However, the former protein binds to G4 with much less affinity than the latter does. Thus, the full-length Rif1 protein mediates high-affinity binding to G4. Molecular weight marker is ϕ X174 DNA digested by *Hae*III.

Supplementary Figure 2 (related to Fig. 2).

Binding of Rif1 to various single-stranded DNAs derived from sequences known to form specific G4 configuration.

Single-stranded oligonucleotide DNAs (0.25 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated in the presence of increasing amounts of Rif1 protein (0, 0.3, 0.75, and 1.5 nM) and were analyzed on 12% PAGE (29:1; 1x TBE, 40% PEG200 and 50 mM KCl). The sequences of the oligonucleotides are shown in Fig. 2. The graph shows quantification of the Rif1 binding to each oligonucleotide. The values were calculated by dividing the radioactivity of the shifted bands (complex) by the sum of the free G4-structured and the shifted bands. The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top. The estimated K_d values are also shown in the table. ND: not determined.

Supplementary Figure S3 (related to Fig. 3).

Binding of Rif1 to various single-stranded DNAs and their derivatives.

The single-stranded oligonucleotide DNAs (0.25 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40 % PEG200, were incubated in the absence (-) or presence (+) of Rif1 protein (10 fmole [1nM] of the Rif1 full-length polypeptide; the preparation used contains ~10-fold molar excess of degradation products), and were analyzed on 12% PAGE (1x TBE, 50 mM KCl and 40 % PEG200). List of oligonucleotides used in the assays and their sequences are shown in Fig. 3.

Supplementary Figure S4 (related to Fig. 4).

Binding of Rif1 to various single-stranded DNAs and their derivatives: stimulation of Rif1 binding by runs of guanine at the 3'-end of the DNA (another one of the triplicate experiments shown in Fig. 4)

The results of an independent set of assay identical to Fig. 4 are shown.

Supplementary Figure S5

Analyses of oligonucleotide used in this study on polyacrylamide gel.

Oligonucleotides shown (0.25 pmole), denatured and reannealed in 50 mM KCl and 40% PEG200, were run on 10% PAGE (29:1; 1xTBE, 50 mM KCl and 10% PEG200), and autoradiographed. The oligonucleotides in red, pink, and black represent good, fair and poor binders of Rif1, respectively. The bands indicated by pink bars show the slow migrating forms. T95_2T_GGG, that shows mobility-shifter forms (indicated by a dotted pink line), is a weak but better binder than T95_2T_G and T95_2T_GG, which show little slow-migrating forms.

Supplementary Figure S6.

Binding of Rif1 to small oligonucleotides derived from Rif1BS consensus sequence: evaluation of AGTG_n sequences and effect of a mutation at the conserved AGT sequence.

The single-stranded oligonucleotide DNAs (0.1 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated in the presence (+, 25 fmole [2.5 nM] of the Rif1 full-length polypeptide; the preparation used contains ~10-fold molar excess of degradation products as well) or absence (-) of Rif1 protein. Samples were analyzed on 14% (A [left]) or 12% (A [right] and B) PAGE (1x TBE, 50 mM KCl and 40% PEG200). The binding assays were conducted in separate experiments two times with similar results, and only the representative data are presented. The graph shows quantification of the Rif1 binding to each oligonucleotide. ds: 12 bp double-stranded DNA (CGCGAATTCGCG). M: molecular weight marker (ϕX174 DNA digested by *Hae*III). The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top.

Supplementary Figure S7.

Binding of Rif1 to small G-tract oligonucleotides.

The single-stranded oligonucleotide DNAs (0.25 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated with increasing amounts of the Rif1 full-length polypeptide, as shown in the figure. Samples

were analyzed on 12% PAGE (1x TBE, 50 mM KCl and 40% PEG200). The ticks represent the sizes of 600, 310, 271/281, 234, 194, 118 and 72 bp, from the top.

Supplementary Figure S8.

CD patterns of the oligonucleotides analyzed in this study.

CD measurement was conducted at 2 μ M DNA in 50 mM Tris-Cl (pH 7.5) without salt (black) or in 50 mM Tris-Cl (pH 7.5) containing 50 mM KCl (red) or 50 mM NaCl (blue). The sequence and relative binding affinity to Rif1 in the presence of KCl are also indicated for each oligonucleotide (-, +, ++, +++). The criteria of binding affinity is defined in the legend to Fig. 3. The topology of DNA estimated from the CD pattern is also presented in red (50 mM KCl) or in blue (50 mM NaCl). CD of Rif1-8, derived from Rif1BS₁₋₄₂₅₅, was measured only in the presence of 100 mM KCl. Summary of this measurement is shown in **Supplementary Table S1**.

Supplementary Figure S9.

Determination of melting temperatures of selected G4-forming sequences used in this study.

CD profiles were measured at various temperatures for a selected set of G4-forming sequences (with 50 mM and 10 mM KCl), as described in the legend to Supplementary Fig. S7. T_m for each sequence was estimated from the melting curve. CEB1 and CEB1_TA are capable of maintaining the parallel form even at 95°C, and thus their T_m are probably higher than 95°C.

Supplementary Figure S10 (related to Fig. 5).

Effect of salt on Rif1 binding to selected G4 oligonucleotides whose topologies change in response to salt (another one of the triplicate experiments shown in Fig. 5).

The results of an independent set of assays identical to Fig. 5 are shown.

Supplementary Figure S11.

Calculation of native molecular weights of the full-length Rif1 and the ~70kDa degraded polypeptide.

The values for both *S* (sedimentation coefficient; *left-most graph and table*) and *R_s* (Stokes radius) were determined. *R_s* was determined according either to Siegel and Monty (*middle graph and middle table*; ref. 44) or to Erickson (*right graph and right table*; ref 45). The native molecular weight (Native Mw.) and the oligomeric state (Multimer), calculated with the equation below the table (left side), are presented in the middle and right tables. *S_{max}/S* values were also calculated and are presented in the same tables. General interpretation of *S_{max}/S*, indicative of molecular shape, is given in a box below the tables.

Supplementary Figure S12.

Effect of Taz1 on interaction of Rif1 with DNA

A. Purification of fission yeast Taz1 protein (663 aa). His₆-Rif1-Flag₃ protein expressed in human embryonic kidney 293T cells was purified by consecutive anti-Flag and nickel columns. Eluates from nickel column were analyzed on SDS-PAGE (5-20% gradient gel) and stained by silver. **B.** ³²P-end labeled Htelo3 DNA (0.25 pmole) was mixed with

the combination of Taz1 and Rif1 proteins, as indicated, and was analyzed on 8% PAGE (29:1; 1x TBE, 40% 50 mM KCl and 1 PEG200). **C.** and **D.** ³²P-end labeled Taz1BS dsDNA (0.25 pmole) was mixed with the combination of Taz1 and Rif1 proteins, as indicated, and was analyzed on 6% PAGE (29:1; 1xTBE and 10% glycerol) (**C**) or 8%PAGE (29:1; 1xTBE, 50 mM KCl, and 40% PEG200) (**D**). Ticks in the ϕ X174/ HaeIII markers (**B**) represent 600, 310, 271/281, 234, 194, 118 and 72 bp.

Supplementary Figure S13.

Ladder formation in denaturing PAGE.

A. ³²P-labeled Rif1-8 (0.25 pmole), denatured and reannealed in 100 mM KCl and 40% PEG200, were run on 6 % PAGE (29:1) with 8M urea in 0.5x TBE, and autoradiographed. **B.** ³²P-labeled T₆G₂₄ (0.25 pmole), denatured and reannealed in 100 mM KCl and 40% PEG200, were run on 8 % PAGE (29:1) with 8M urea in 0.5x TBE, and autoradiographed. In both A and B, DNAs were pretreated with different concentration of formamide in the absence (lanes 1-6) or presence (lanes 7-12) of 1M NaCl. DNAs were heated at 96°C for 1 min before loading onto PAGE (lanes 1-5, 7-11). Ladders are still generated in a denaturing gel after heat denaturation in formamide, suggesting that the structures may not be simply stacked oligomers, but could include intertwined G4 structures made of multiple DNA strands.

Supplementary Figure S14.

The length of G-tracts present in Rif1BS.

The sequences of the 1 kb segments encompassing the 35 Rif1BS were analyzed. All the G-tracts equal or longer than 3 were extracted, and the numbers of 3G~7G were scored for the “top 10” and “bottom 10” Rif1BS. The analyses were conducted for both strands. The G-strand represents the strand on which the G-tracts of Rif1CS are present, and the C-strand the other strand. Fractions of each G-tract are shown.

Supplementary Figure S15.

Frequency of various dinucleotide sequences preceding the G-tracts (n=3 or longer) in Rif1BS.

The sequences of the 1 kb segments encompassing the 35 Rif1BS were analyzed. All the G-tracts equal or longer than 3 were extracted, and the preceding dinucleotides were scored. The analysis was conducted on both strands. The G-strand represents the strand on which the G-tracts of Rif1CS are present, and the C-strand the other strand. Fractions of each dinucleotide are shown.

Supplementary Figure S16.

Analyses of labeled oligonucleotides on denaturing PAGE containing 8M urea.

Oligonucleotides used in this studies were purified by gel filtration, end-labeled by T4 polynucleotide kinase, and denatured in 95% formamide containing 5 mM EDTA and 20 mM LiCl, and were run on 15% PAGE containing 8M urea in 0.5x TBE. T₆G₂₄ appears as a ladder, since it forms secondary structures even in the presence of 8M urea. The bands at the bottom of the gels are free labeled nucleotides that remained after gel filtration purification of labeled oligonucleotides. In lane11, the oligonucleotide (12mer) was too small to be separated on this gel and ran off the gel.

Supplementary Figure S17.

Original full-length gels for data in Supplementary Fig. S1.

Figure S1A lane 1: gel1

Figure S1A lane 2: gel2

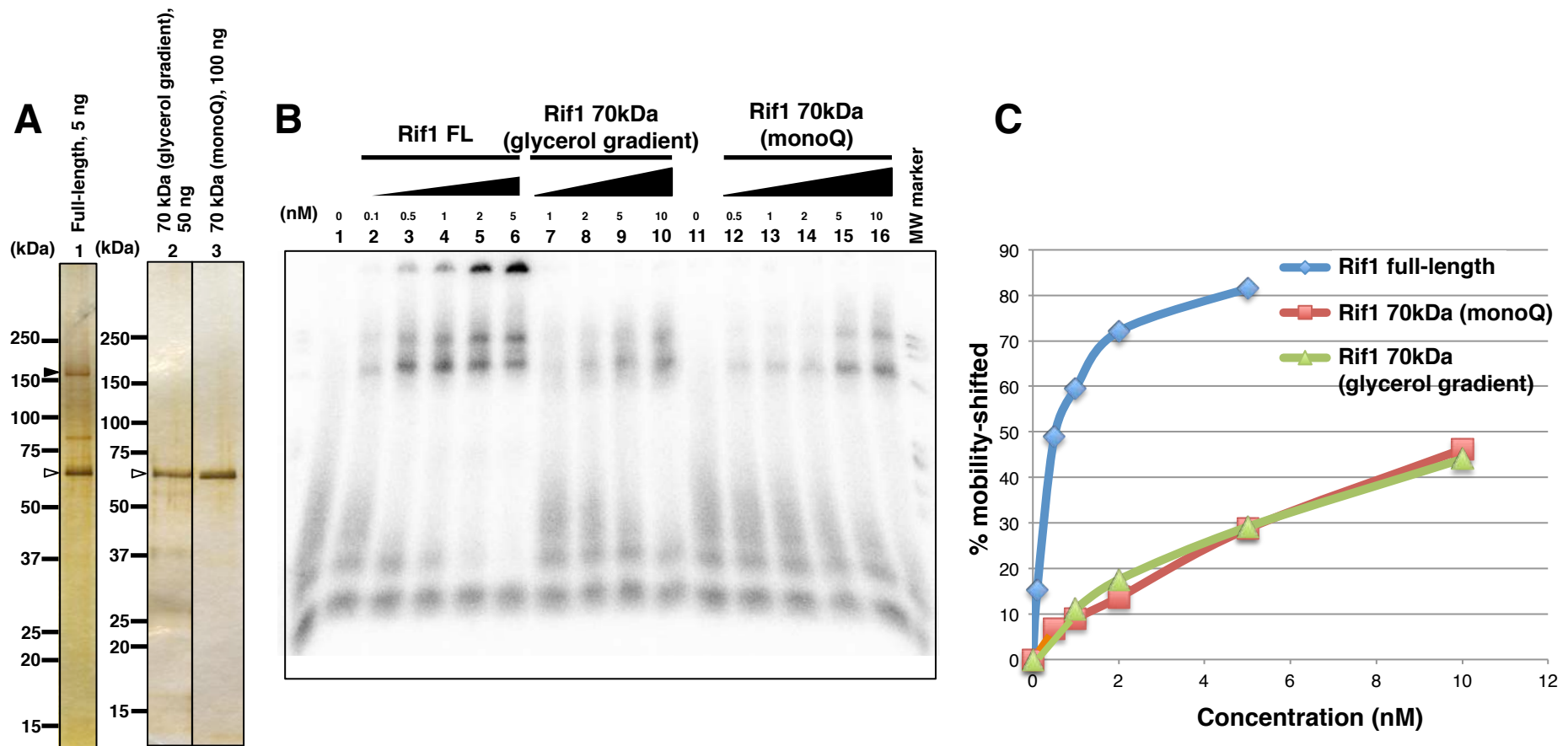
Figure S1A lane 3: gel3

Supplementary Figure S18.

List of the selected oligonucleotides used in this study and their reported structures.

Name	Sequence	Topology (reported)	Topology in KCl (observed)	Topology in NaCl (observed)	Rif1 binding	Rif1 binding in NaCl	Rif1 binding without salt
Pu24T	TGAGGGTGGTGAGGGTGGGGA	Parallel	Parallel	Non-G4?	++		
CEB25	AAGGGTGGGTGTAAGTGTGGGTGGG T	Parallel	Parallel	Parallel?	+		
HT	TTGGGTAGGGTTAGGGTTAGGGA	Hybrid (form1)	Hybrid	Anti-Parallel	-		
Htelo1	TAGGGTTAGGGTTAGGGTTAGGG	Hybrid (form1)	Hybrid	Anti-Parallel	-		
Htelo2	TAGGGTTAGGGTTAGGGTTAGGGTT	Hybrid (form2)	Hybrid	Anti-Parallel	-		
Htelo3	GGGTTAGGGTTAGGGTTAGGGT	Anti-Parallel	Hybrid	Anti-Parallel	-		
DX1 (duplex)	CGCGAATTCGCG		Non-G4	Non-G4	-		
DX2 (hairpin)	ATCTGAGAATCAGAT		Non-G4	Non-G4	-		
TERC18 (RNA)	GGGUUGCGGAGGGUGGGC	Parallel (dimeric)	Parallel	Parallel?	+		
T ₆ G ₂₄	TTTTTTGGGGGGGGGGGGGGGGGG GGGG	a propeller-type parallel	Parallel	Parallel	+++	++	+
Htelo4	AGGGCTAGGGCTAGGGCTAGGG	Chair	Anti-Parallel	Anti-Parallel	-		
Htelo4_3nt spacer	AGGGCTAGGGCTACTAGGGCTAGGG		Anti-Parallel	Anti-Parallel	-		
Htelo4(GGGGG)	AGGGCTAGGGCTAGGGGGG		Hybrid or Anti-Parallel	Mix or Anti-Parallel	++		
Htelo4(GGGGG)_3nt spacer	AGGGCTAGGGCTACTAGGGGGG		Hybrid?	Anti-Parallel	++	+	+
T ₆ (GA) ₁₂	TTTTTTGAGAGAGAGAGAGAGAGAG AGAGA		Non-G4	Non-G4	-		
CEB1	AGGGGGGAGGGAGGGTGG	Parallel (dimeric)	Parallel	Parallel?	++		
CEB1_A	AGGGAGGGAGGGAGGGTGG		Parallel	Parallel?	++		
CEB1_TA	AGGGTAGGGAGGGAGGGTGG		Parallel	Non-G4?	++	-	-
CEB1_TT_A	AGGGTTAGGGAGGGAGGGTGG		Parallel	Non-G4	++	-	-
CEB1_ΔGG	AGGGGGGAGGGAGGGT		Parallel	Parallel?	++		
Htelo1	TAGGGTTAGGGTTAGGGTTAGGG		Hybrid	Anti-Parallel	-		
Htelo1_A	TAGGGAGGGTTAGGGTTAGGG		Hybrid or Mix	Anti-Parallel or Mix	-		
Htelo1_no spacer	TAGGGGGGTTAGGGTTAGGG		Parallel	Anti-Parallel or Mix	++		
Htelo1_no spacer 2	TAGGGTTAGGGTTAGGGGGG		Mix?	Anti-Parallel or Mix	+++	+	+
93del	GGGGTGGGAGGAGGGT	Parallel (dimeric)	Parallel	Parallel	++		
93del_ΔG	GGGTGGGAGGAGGGT		Parallel	Parallel	+		
T95_2T	TTGGGTGGGTGGGTGGGT	Parallel	Parallel	Parallel	+		
T95_2T_G	TTGGGGTGGGTGGGTGGGT		Parallel	Parallel	-		
T95_2T_GG	TTGGGGGTGGGTGGGTGGGT		Parallel	Parallel	-		
T95_2T_GGG	TTGGGGGGTGGGTGGGTGGGT		Parallel	Parallel	-		
Rif1-8	GTGGGGGATGTGGGATAGACTATTTT AGACTTGAGACTGTGTATTCACCTAA AAGTGCTATGTGGGGGCA		Parallel		++		

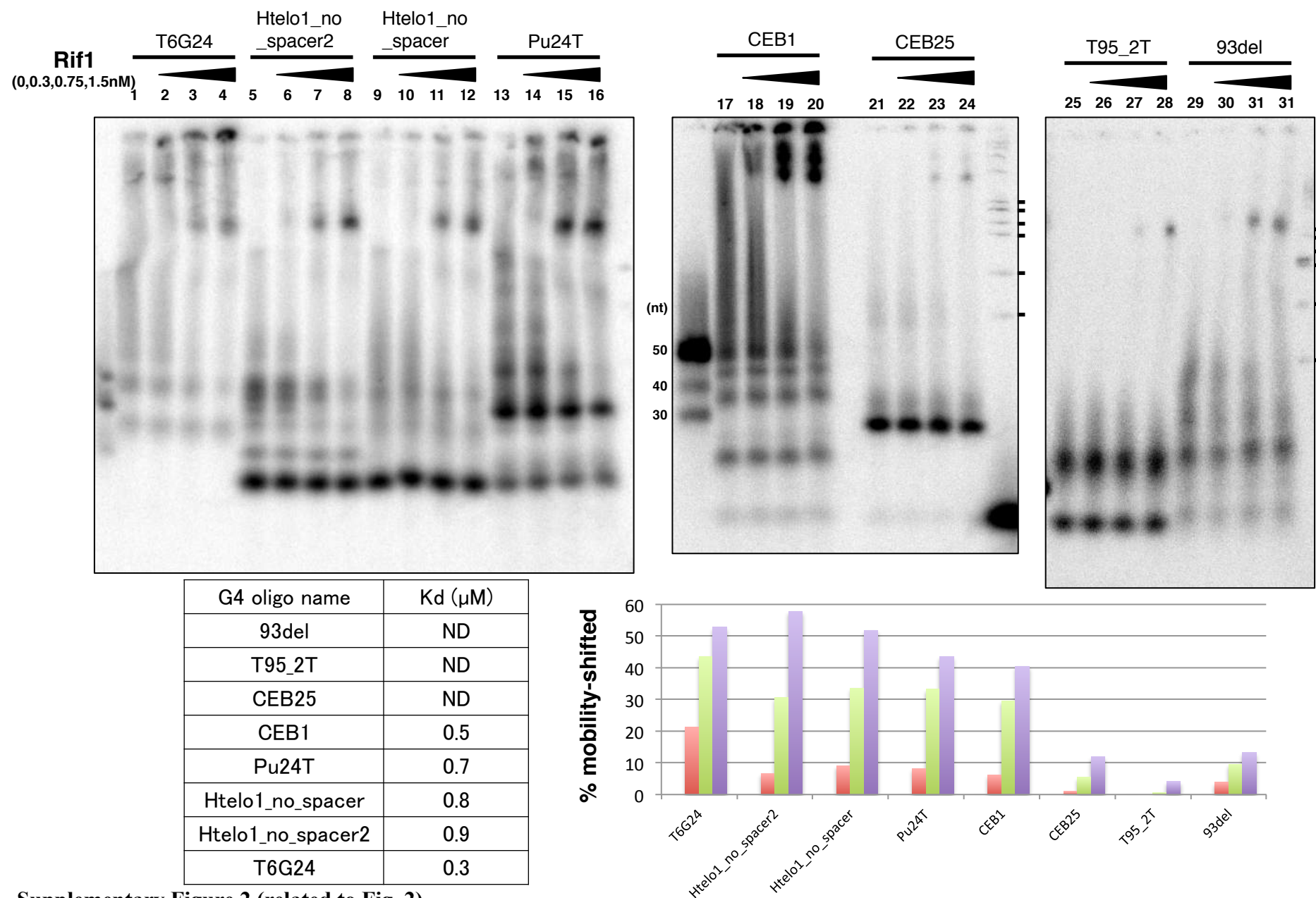
Supplementary Table S1 List of G4-related oligonucleotides used in this study: sequence, topology and Rif1 binding



Supplementary Figure S1.

Comparison of DNA binding activities of the full-length and 70 kDa C-terminal polypeptide of Rif1.

A. Protein fractions containing the full-length or 70 kDa polypeptide of Rif1 were analyzed on 5-20% gradient SDS-PAGE and proteins were stained by silver. Lane 1, full-length Rif1 protein purified by consecutive steps of anti-FLAG affinity column, nickel column and monoQ column chromatography (a different fraction from that used in Fig. 1). Lane 2, the 70 kDa polypeptide fraction purified by glycerol gradient centrifugation of the FLAG-column purified material. Lane 3, the 70 kDa polypeptide fraction purified by monoQ column chromatography of the material enriched by consecutive steps of anti-FLAG affinity column and nickel column. Closed and open triangles indicate the full-length and the 70 kDa C-terminal polypeptide of Rif1, respectively. In the fractions for lanes 2 and 3, the full-length Rif1 was not detected even with western blotting (not shown). The protein concentrations were estimated by comparison with the BSA marker that was run in the same gel (not shown). Lane 1 and lanes 2/3 were from different gels; therefore, the band intensities cannot be directly compared. Original full-length gels are shown in Supplementary Fig. S17. **B.** The end-labeled T_6G_{24} DNA (0.25 pmole), which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated with increasing amounts of Rif1 protein, as shown, and were analyzed on 8% PAGE containing 50 mM KCl and 40% PEG200 in 1x TBE. **C.** Band intensities of the gel were measured and relative binding efficiency was estimated as described in the legend to Fig. 2. The concentrations of the protein on the X-axis are based on those of the full-length polypeptide alone. However, the preparation used contains four-fold molar excess of the C-terminal 70 kDa polypeptide. If this polypeptide is capable of binding to G4 and is also taken into account, the concentrations of the polypeptides could be as much as five-times larger. Even that is the case, the K_d would be ~ 1.5 nM (assuming that the 70 kDa polypeptide binds to the G4 substrate as efficiently as the full-length). However, the former protein binds to G4 with much less affinity than the latter does. Thus, the full-length Rif1 protein would mediate high-affinity binding to G4. Molecular weight marker is ϕ X174 DNA digested by *Hae*III.

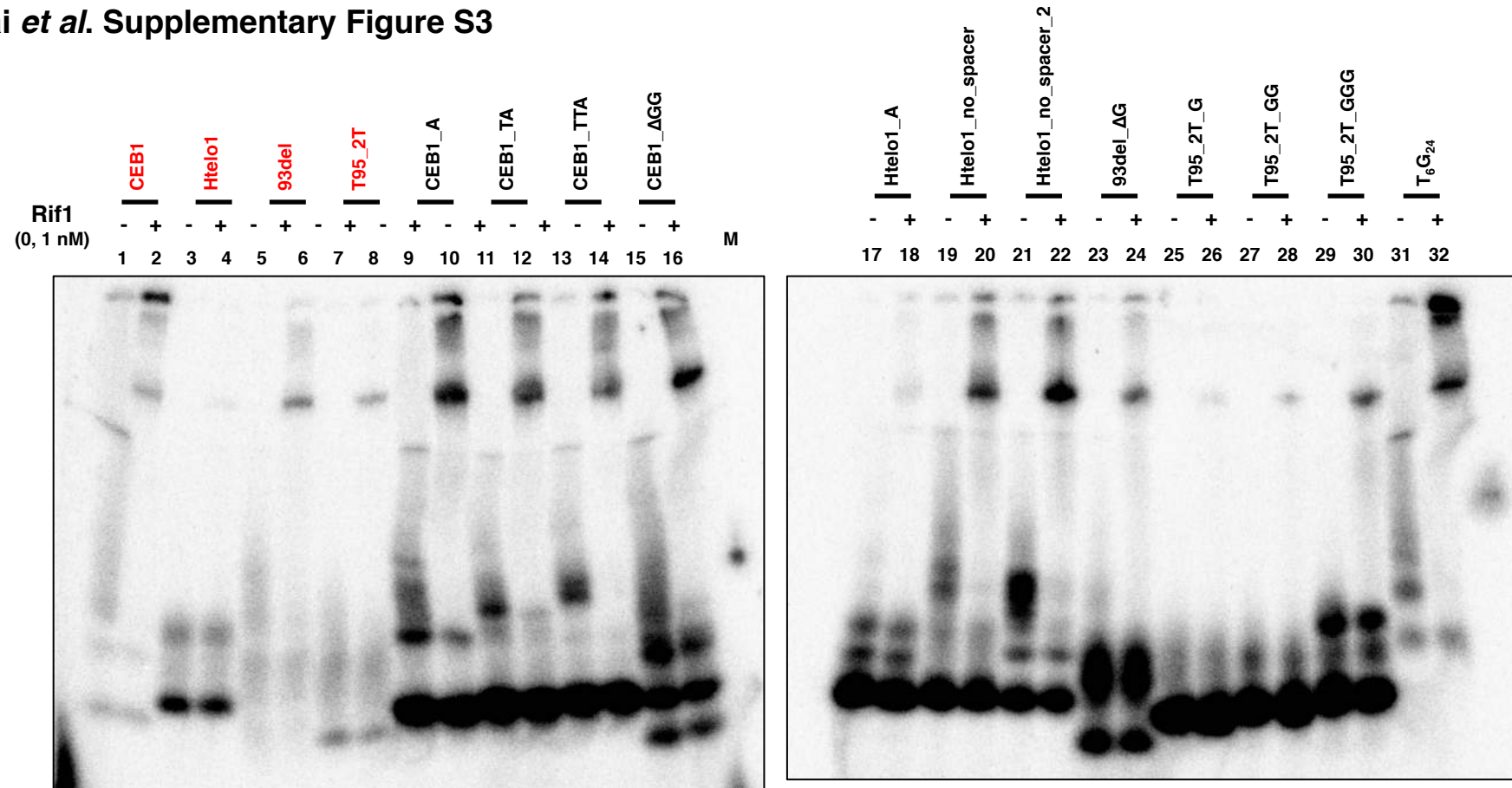


Supplementary Figure 2 (related to Fig. 2).

Binding of Rif1 to various single-stranded DNAs derived from sequences known to form specific G4 configuration.

Single-stranded oligonucleotide DNAs (0.25 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated in the presence of increasing amounts of Rif1 protein (0, 0.3, 0.75, and 1.5 nM) and were analyzed on 12% PAGE (29:1; 1x TBE, 40% PEG200 and 50 mM KCl). The sequences of the oligonucleotides are shown in Fig. 2. The graph shows quantification of the Rif1 binding to each oligonucleotide. The values were calculated by dividing the radioactivity of the shifted bands (complex) by the sum of the free G4-structured and the shifted bands. The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top. The estimated Kd values are also shown in the table. ND: not determined.

Masai *et al.* Supplementary Figure S3

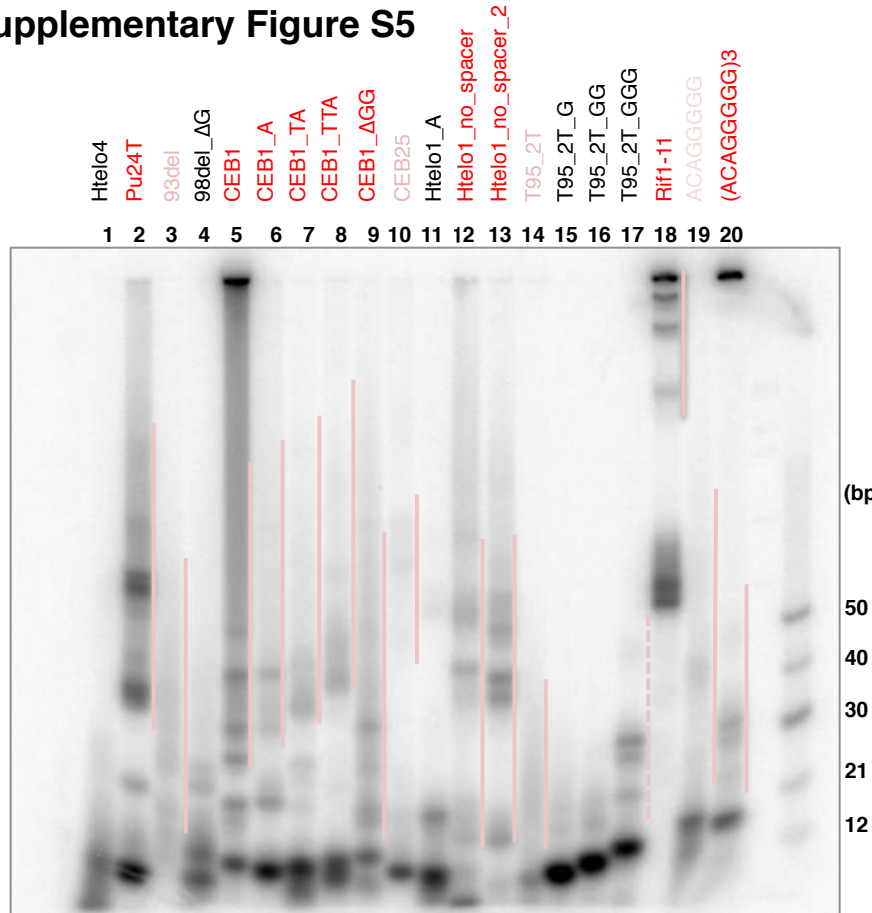


Supplementary Figure S3 (related to Fig. 3).

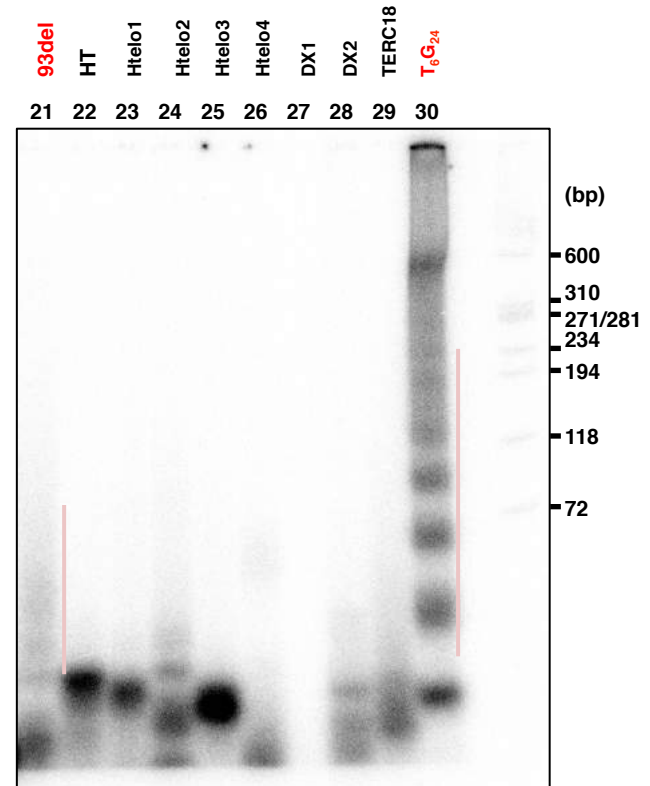
Binding of Rif1 to various single-stranded DNAs and their derivatives: effect of spacer sequence and runs of guanine on binding of Rif1.

The single-stranded oligonucleotide DNAs (0.25 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40 % PEG200, were incubated in the absence (-) or presence (+) of Rif1 protein (10 fmole [1nM] of the Rif1 full-length polypeptide; the preparation used contains ~10-fold molar excess of degradation products), and were analyzed on 12% PAGE (1x TBE, 50 mM KCl and 40 % PEG200). List of oligonucleotides used in the assays and their sequences are shown in Fig. 3.

Masai *et al.* Supplementary Figure S5



Htelo4:AGGGCTAGGGCTAGGGCTAGGG
 Pu24T:TGAGGGTGGTGGGTGGGA
 93del:GGGGTGGGAGGAGGGT
 93del_ΔG:GGGTGGGAGGAGGGT
 CEB1:AGGGGGAGGGAGGGTGG
 CEB1_A:AGGGAAGGGAGGGAGGGTGG
 CEB1_TA:AGGGTAGGGAGGGAGGGTGG
 CEB1_TTA:AGGGTTAGGGAGGGAGGGTGG
 CEB1_ΔGG:AGGGGGAGGGAGGGT
 CEB25:AAGGGTGGGTGTAAGTGTGGGTGGGT
 Htelo1:TAGGGTTAGGGTTAGGGTTAGGG
 Htelo1_A:TAGGGAAGGGTTAGGGTTAGGG
 Htelo1_no_space:TAGGGGGTTAGGGTTAGGG
 Htelo1_no_spacer_2:TAGGGTTAGGGTTAGGGGGG
 T95_2T:TTGGGTGGGTGGGTGGGT
 T95_2T-G:TTGGGGTGGGTGGGTGGGT
 T95_2T-GG:TTGGGGTGGGTGGGTGGGT
 T95_2T-GGG:TTGGGGTGGGTGGGTGGGT
 Rif1-11:TTTGGGTCTA AGTGGGGAT GTGGGATAGA CTATTTTAGA CTTGAGACTG
 TGTATTCAT TAAAAGTGCT ATGTGGGGC A

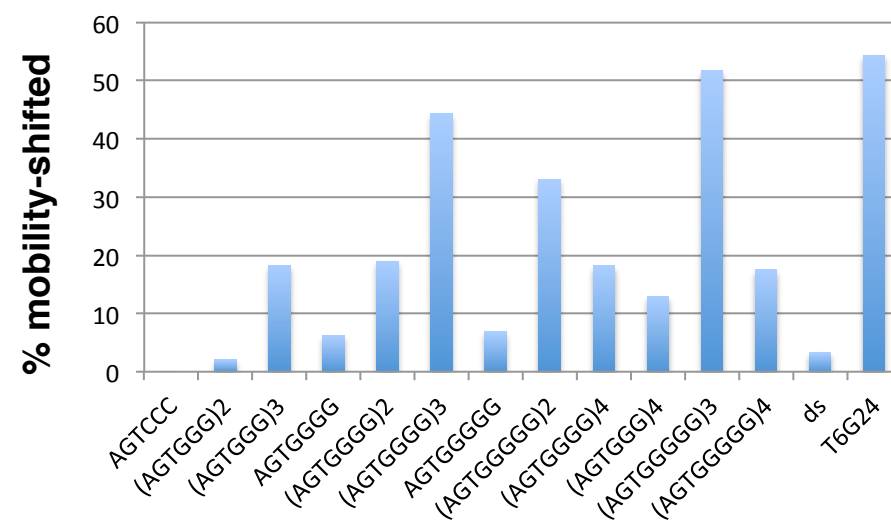
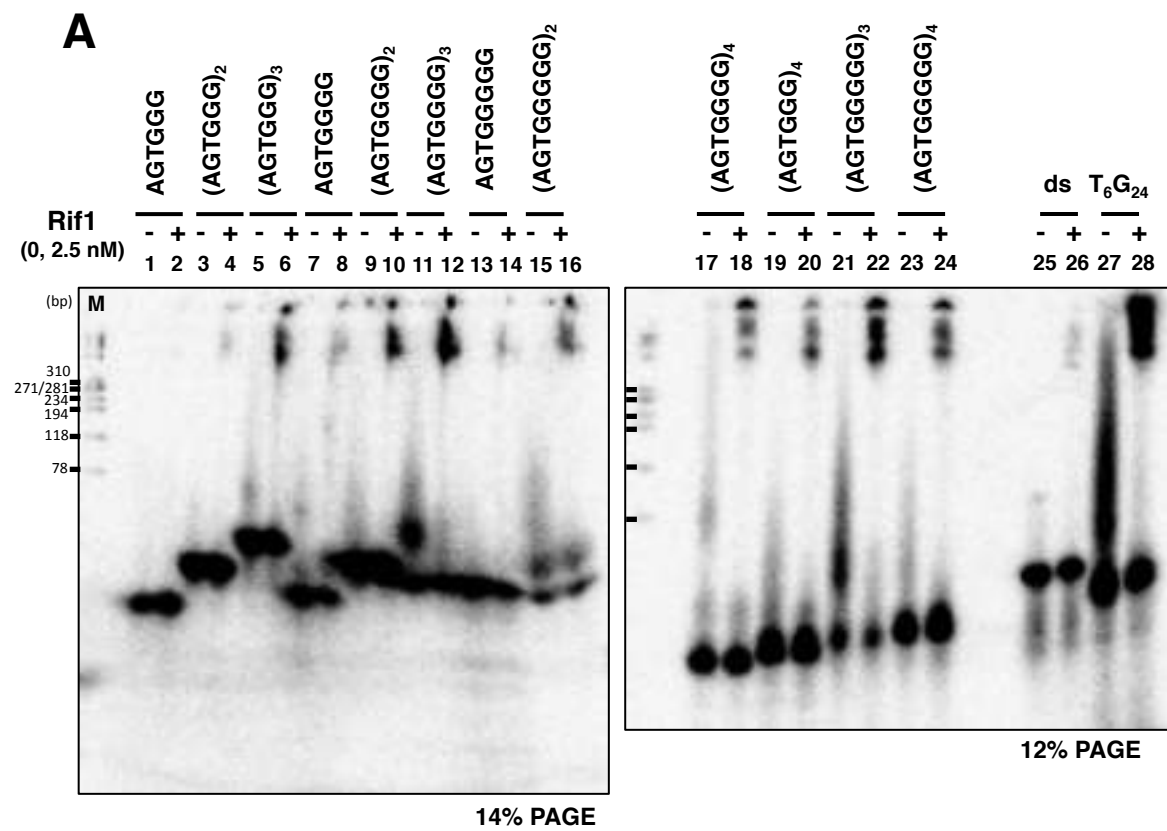


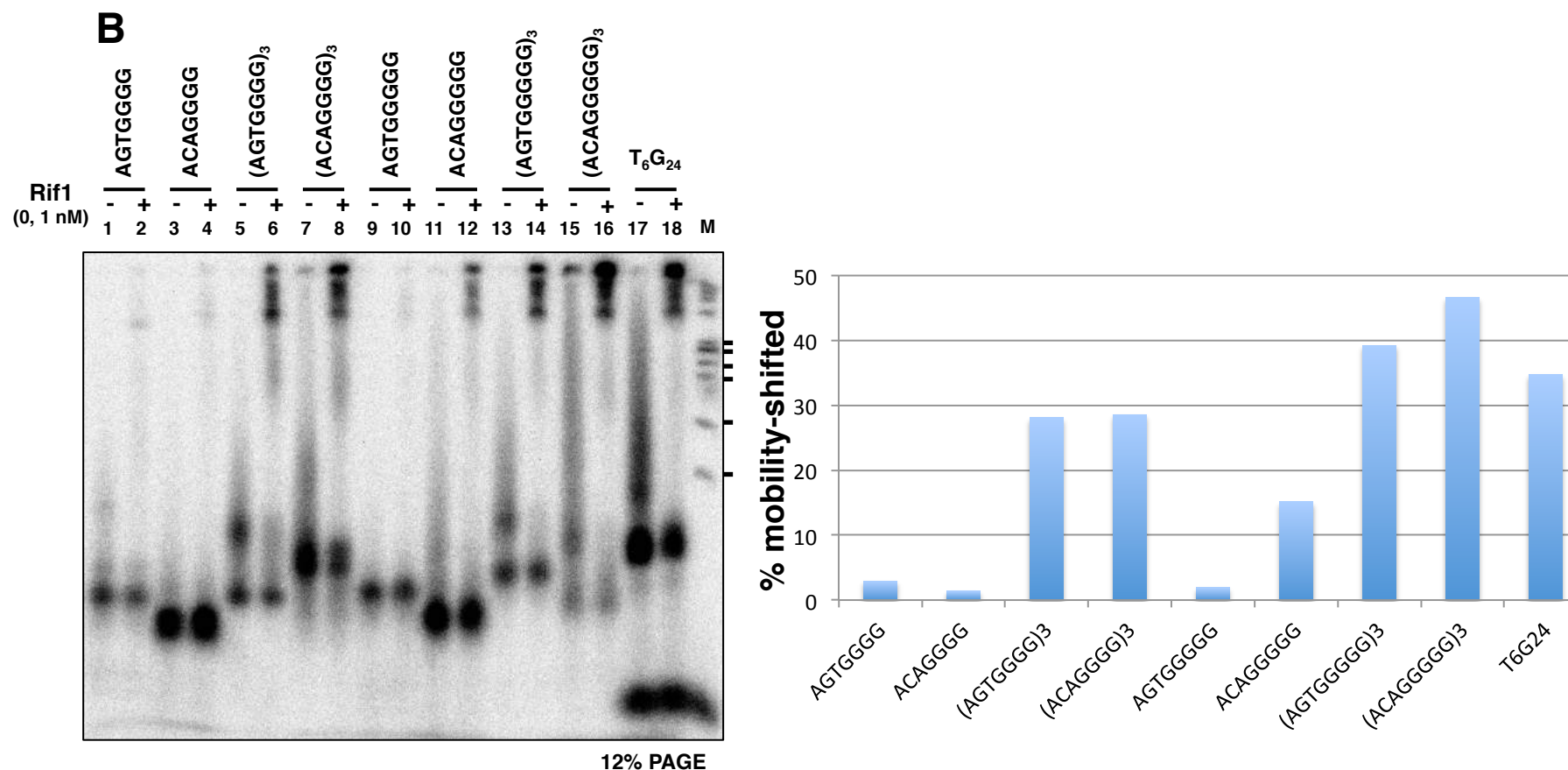
93del:GGGGTGGGAGGAGGGT
 HT:TTGGGTAGGGTTAGGGTTAGGGA
 Htelo1:TAGGGTTAGGGTTAGGGTTAGGG
 Htelo2:TAGGGTTAGGGTTAGGGTTAGGGTT
 Htelo3:GGGTAGGGTTAGGGTTAGGGT
 Htelo4:AGGGCTAGGGCTAGGGCTAGGG
 DX1:CGCGAATTCGCG
 DX2:ATCTGAGAATCAGAT
 TERC18:GGGUUGCGGAGGGUGGGC
 T₆G₂₄:TTTTTTGGGGGGGGGGGGGGGGGGGG

Supplementary Figure S5

Analyses of oligonucleotide used in this study on polyacrylamide gel.

Oligonucleotides shown (0.25 pmole), denatured and reannealed in 50 mM KCl and 40% PEG200, were run on 10% PAGE (29:1; 1xTBE, 50 mM KCl and 10% PEG200), and autoradiographed. The oligonucleotides in red, pink, and black represent good, fair and poor binders of Rif1, respectively. The bands indicated by pink bars show the slow migrating forms.

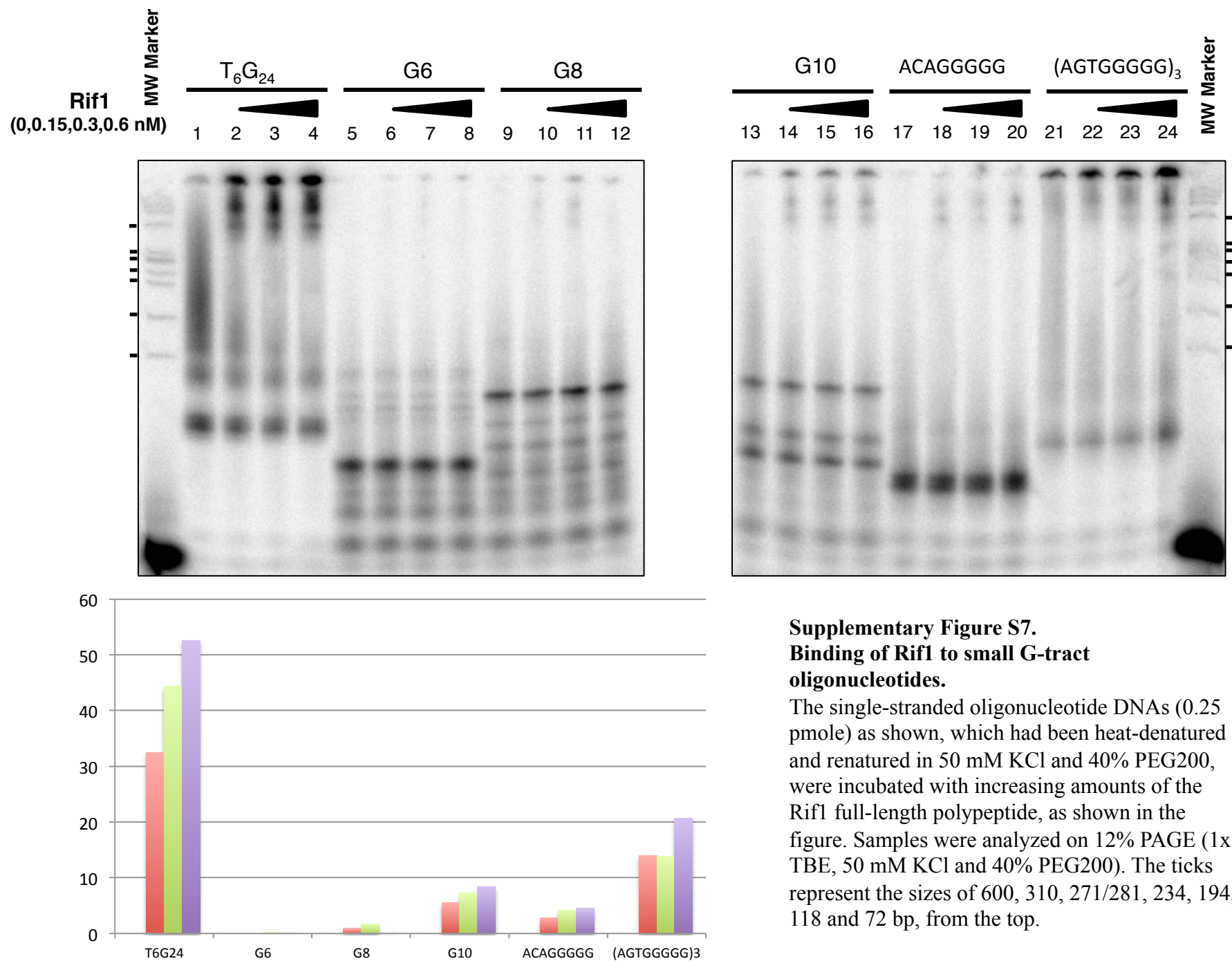




Supplementary Figure S6.

Binding of Rif1 to small oligonucleotides derived from Rif1BS consensus sequence: evaluation of AGT_n sequences and effect of a mutation at the conserved AGT sequence.

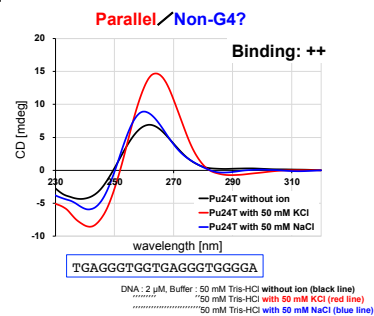
The single-stranded oligonucleotide DNAs (0.1 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated in the presence (+, 25 fmole [2.5 nM] of the Rif1 full-length polypeptide; the preparation used contains ~10-fold molar excess of degradation products as well) or absence (-) of Rif1 protein. Samples were analyzed on 14% (A [left]) or 12% (A [right] and B) PAGE (1x TBE, 50 mM KCl and 40% PEG200). The binding assays were conducted in separate experiments two times with similar results, and only the representative data are presented. The graph shows quantification of the Rif1 binding to each oligonucleotide. ds: 12 bp double-stranded DNA (CGCGAATTCGCG). M: molecular weight marker (ϕX174 DNA digested by *Hae*III). The ticks represent the sizes of 310, 271/281, 234, 194, 118 and 72 bp, from the top.



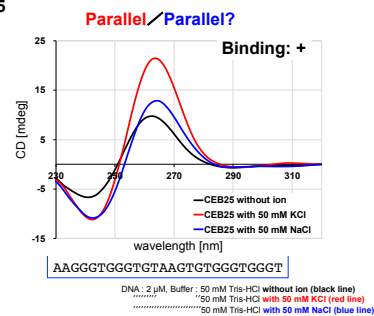
Supplementary Figure S7.
Binding of Rif1 to small G-tract oligonucleotides.

The single-stranded oligonucleotide DNAs (0.25 pmole) as shown, which had been heat-denatured and renatured in 50 mM KCl and 40% PEG200, were incubated with increasing amounts of the Rif1 full-length polypeptide, as shown in the figure. Samples were analyzed on 12% PAGE (1x TBE, 50 mM KCl and 40% PEG200). The ticks represent the sizes of 600, 310, 271/281, 234, 194, 118 and 72 bp, from the top.

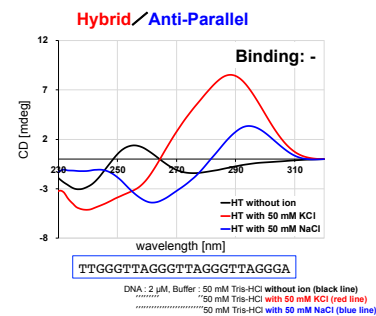
Pu24T



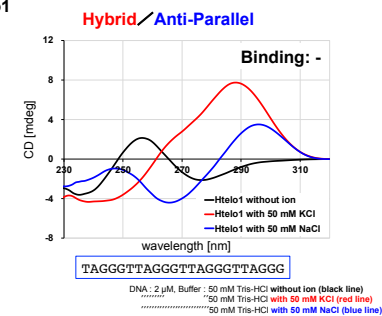
CEB25



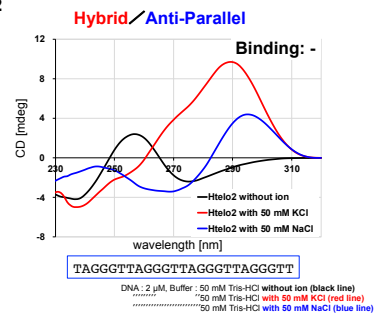
HT



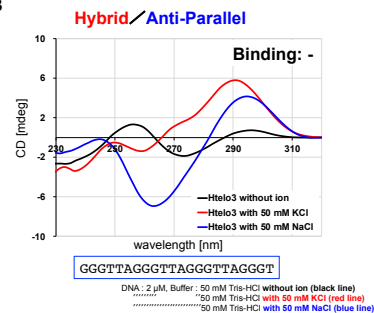
Htelo1



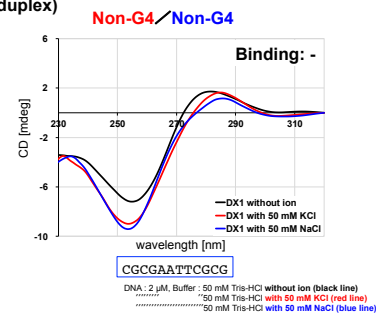
Htelo2



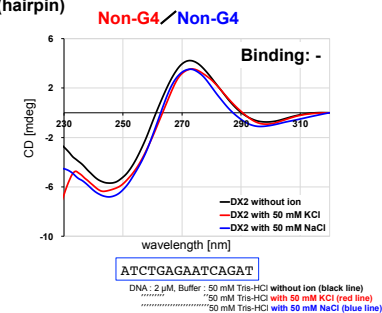
Htelo3



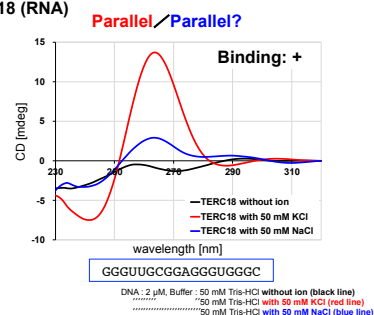
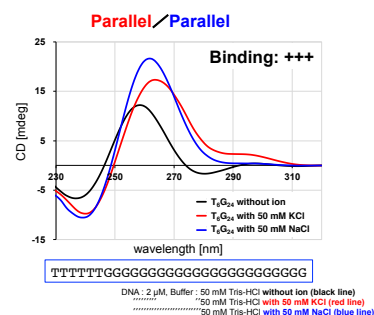
DX1 (duplex)



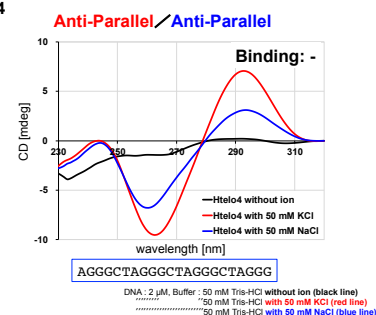
DX2 (hairpin)



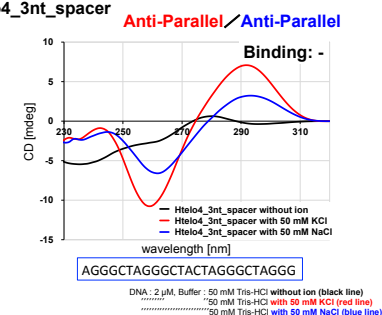
TERC18 (RNA)

 T_6G_{24} 

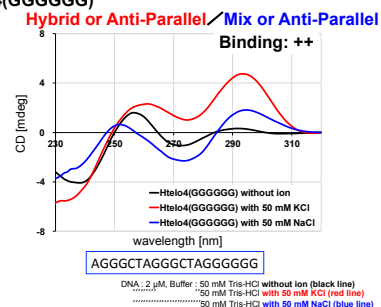
Htelo4



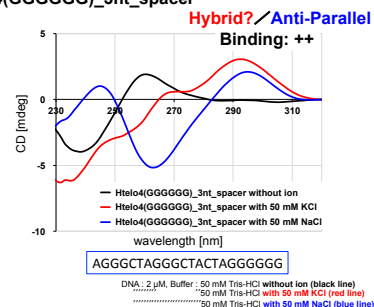
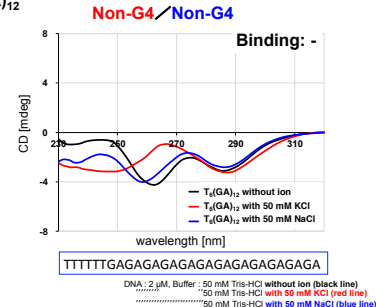
Htelo4_3nt_spacer



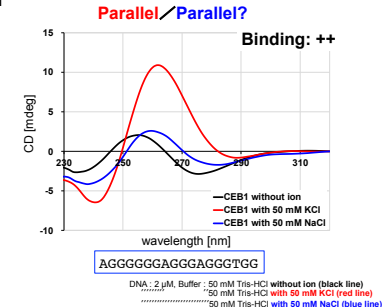
Htelo4(GGGGGG)

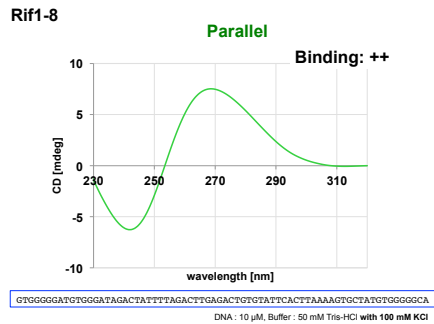
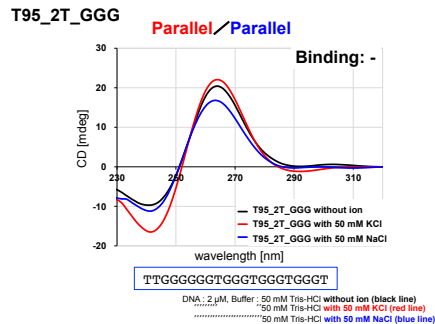
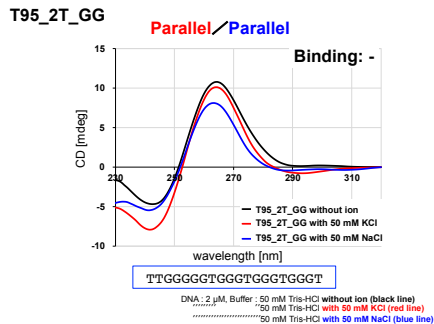
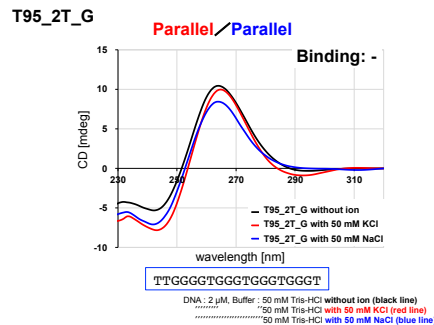
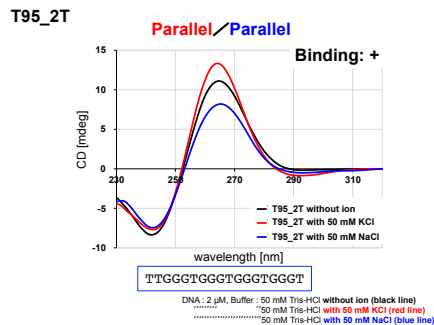
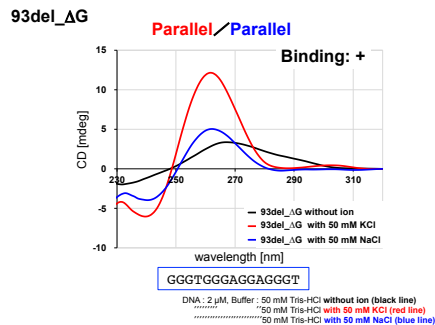
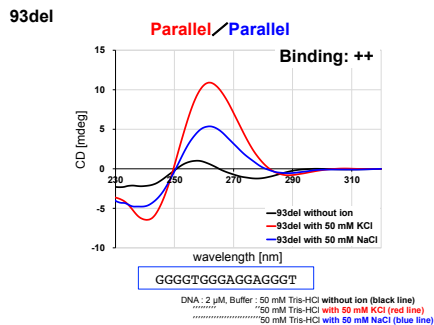
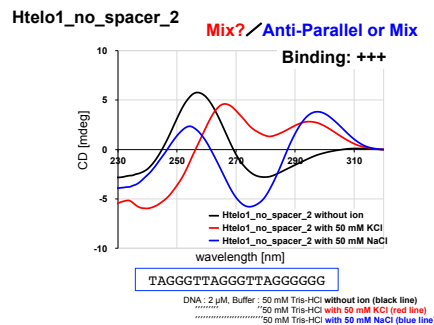
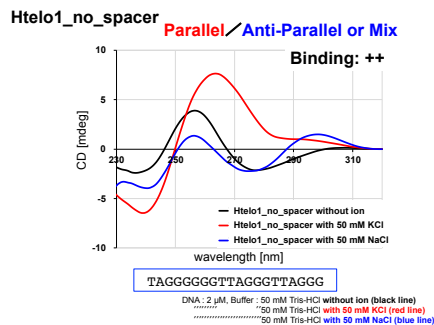
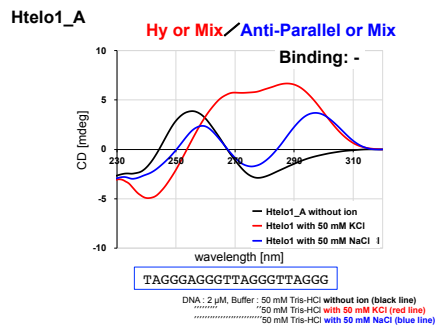
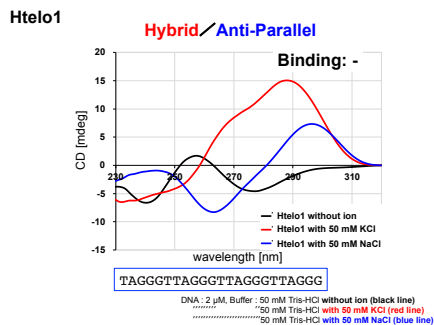
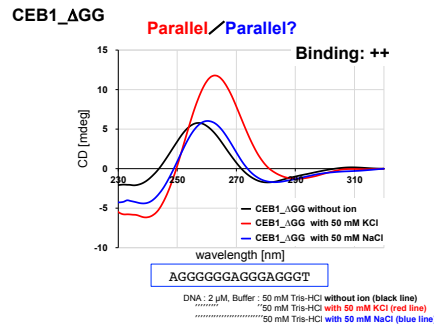
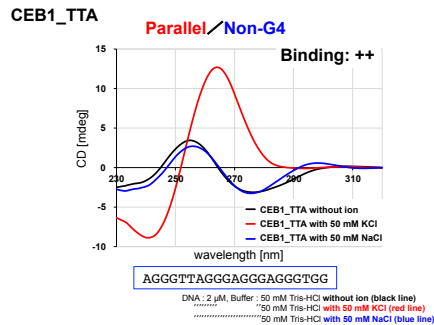
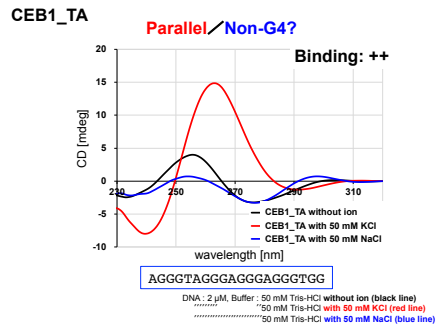
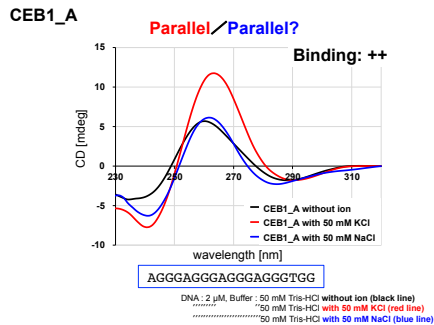


Htelo4(GGGGGG)_3nt_spacer

 $T_6(GA)_{12}$ 

CEB1

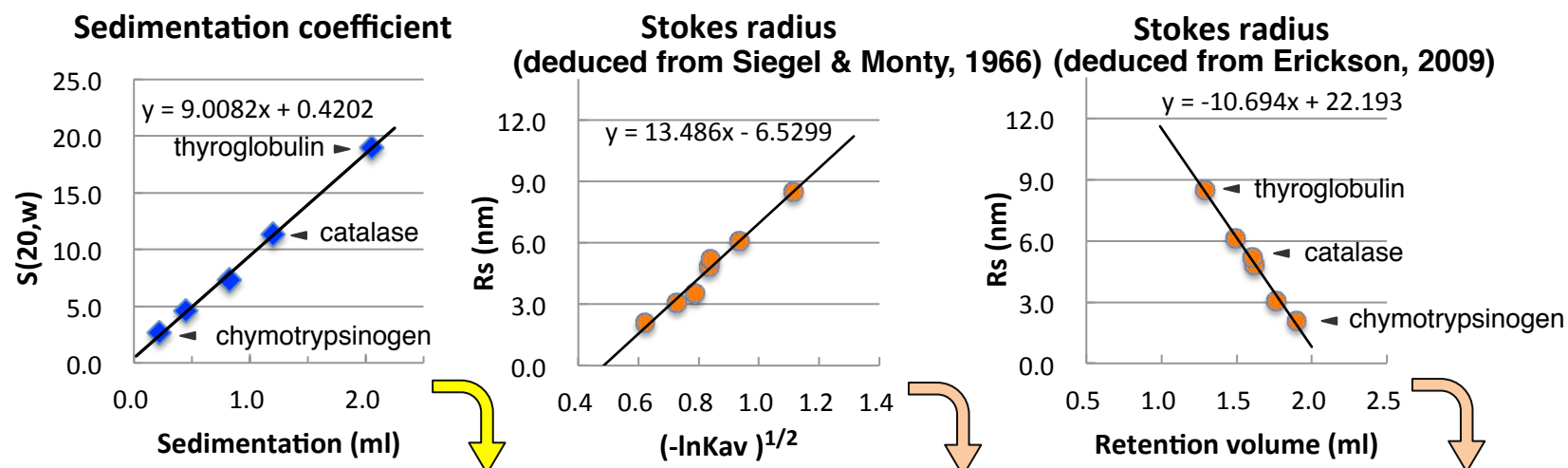




Supplementary Figure S8.

CD patterns of the oligonucleotides analyzed in this study.

CD measurement was conducted at 2 μ M DNA in 50 mM Tris-Cl (pH 7.5) without salt (black) or in 50 mM Tris-Cl (pH 7.5) containing 50 mM KCl (red) or 50 mM NaCl (blue). The sequence and relative binding affinity to RifI in the presence of KCl are also indicated for each oligonucleotide (-, +, ++, +++). The criteria of binding affinity is defined in the legend to Fig. 3. The topology of DNA estimated from the CD pattern is also presented in red (50 mM KCl) or in blue (50 mM NaCl). Summary of this measurement is shown in **Supplementary Table S1**.



	Monomer Mw.	sediment (ml)	S value	retention (ml)	Kav	Rs (nm)	Native Mw.	Multimer	Smax/S	Rs (nm)	Native Mw.	Multimer	Smax/S
Full-length Rif1(S)	161,332	1.250	13.129	0.840	0.0063694	24.483	1,351,555	8.38	3.36	13.368	737,961	4.57	2.25
Full-length Rif1(M)	161,332	1.650	17.404	0.840	0.0063694	24.483	1,791,672	11.11	3.06	13.368	978,268	6.06	2.04
Full-length Rif1(F)	161,332	2.050	21.679	0.840	0.0063694	24.483	2,231,789	13.83	2.84	13.368	1,218,576	7.55	1.90
70 kD polypeptide(S)	70,000	0.350	3.510	1.296	0.2968153	8.512	125,612	1.79	2.58	8.398	123,927	1.77	2.56
70 kD polypeptide(M1)	70,000	0.850	8.854	1.296	0.2968153	8.512	316,885	4.53	1.90	8.398	312,634	4.47	1.88
70 kD polypeptide(M2)	70,000	1.420	14.946	1.296	0.2968153	8.512	534,935	7.64	1.59	8.398	527,759	7.54	1.58
70 kD polypeptide(F)	70,000	2.050	21.679	1.296	0.2968153	8.512	775,939	11.08	1.41	8.398	765,530	10.94	1.39

$$\text{Native Mw.} = SN_0(6\pi\eta R_s)/(1-\nu_2\rho) = 4,205(SR_s)$$

Mw. = Native Molecular weight (Da)

S = Sedimentation coefficient (Svedverg unit = 10^{-13} s)

Rs = Stokes radius (nm)

$N_0 = 6.023 \times 10^{23}$ (Avogadro's number)

$\eta = 0.01$ g/cm·s (viscosity of the solvent, H₂O)

$\nu_2 = 0.73$ cm³/g (partial specific volume of most proteins)

$\rho = 1.0$ g/cm³ (density of the solvent, H₂O)

$$S_{max}/S = R/R_{min}$$

1.2~1.3: globular protein

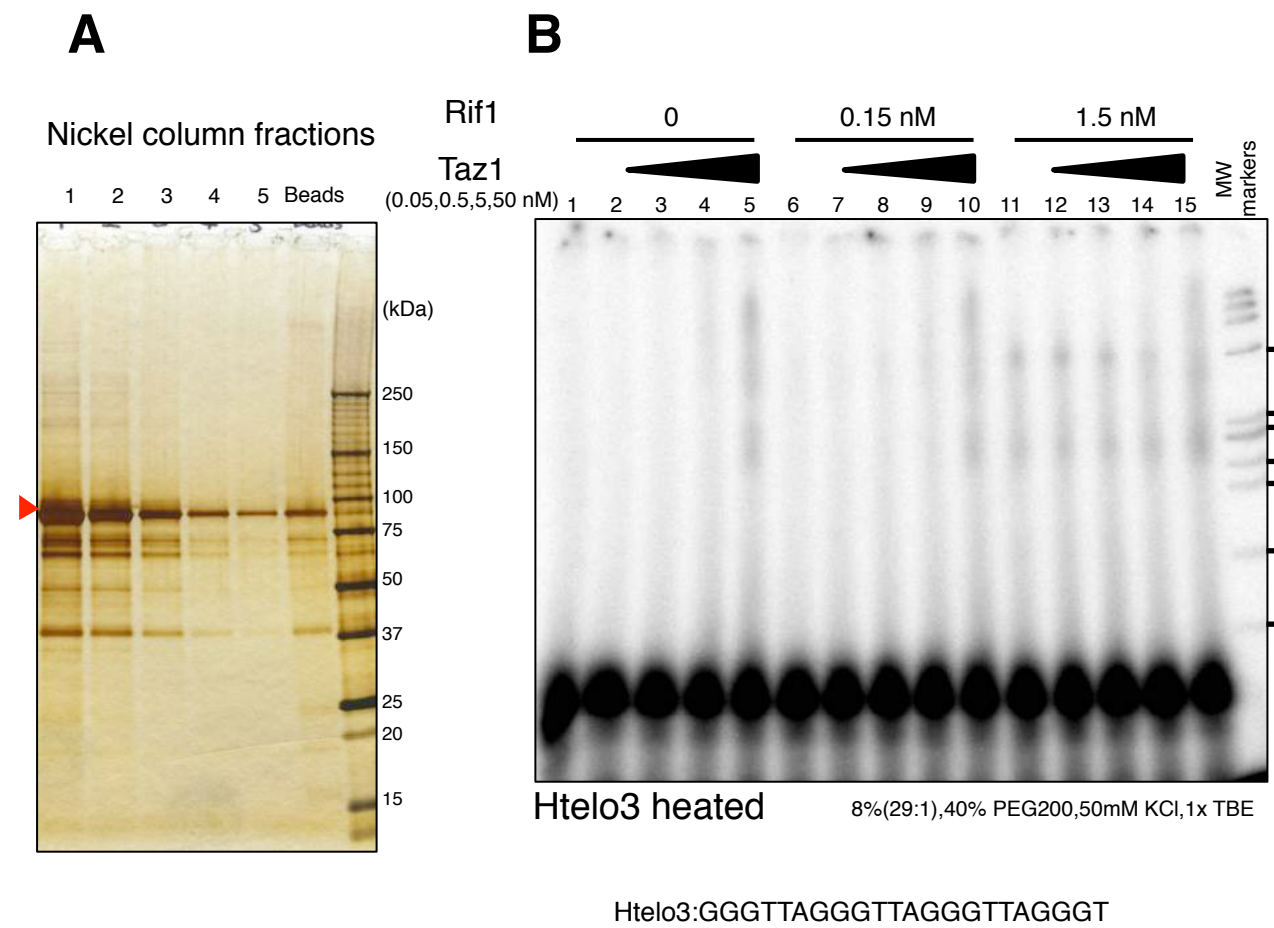
1.5~1.9: moderately elongated

2.0~3.0: highly elongated

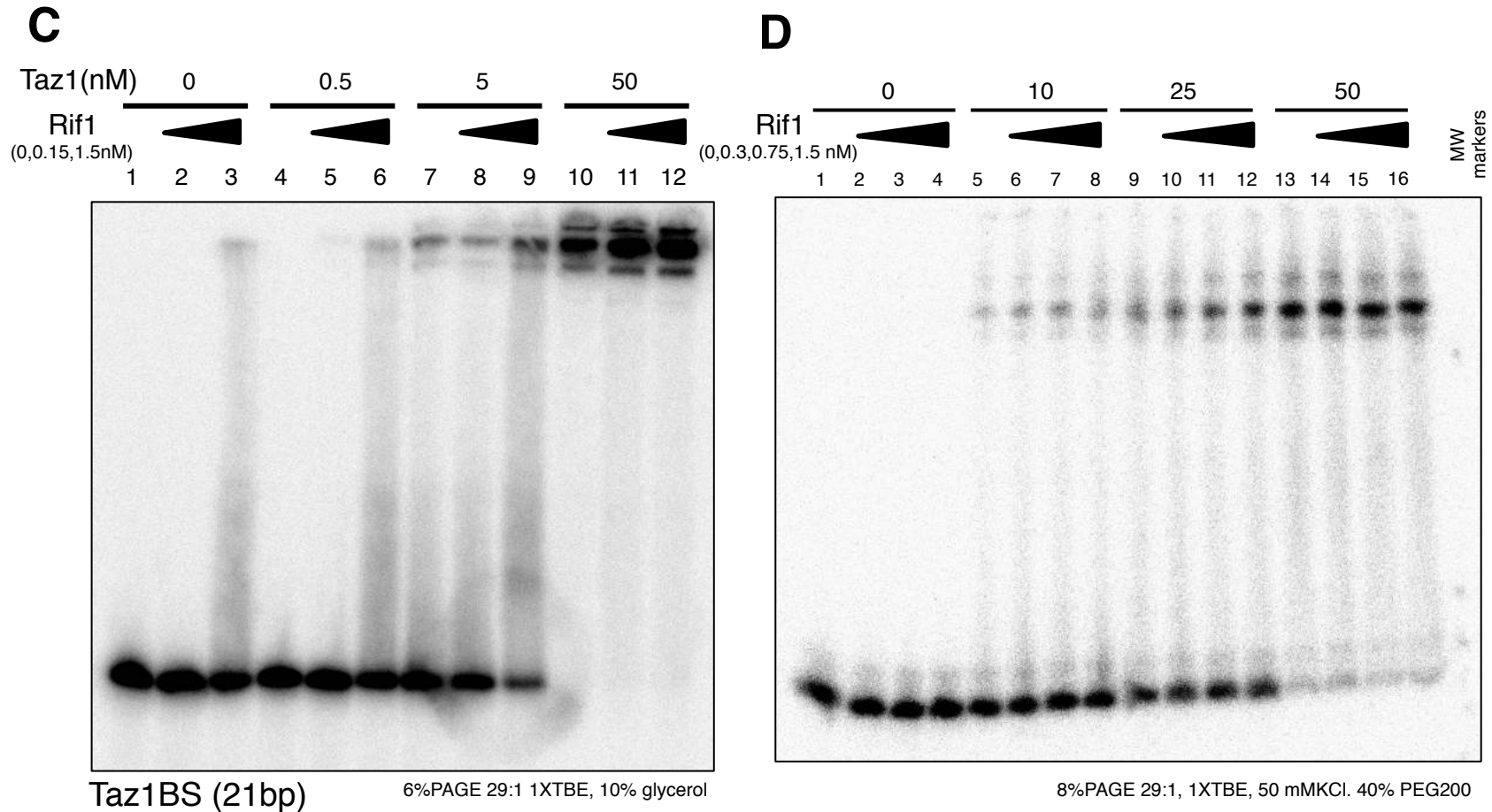
>3.0: very long thread-like molecule

Supplementary Figure S11. Calculation of native molecular weights of the full-length Rif1 and the ~70kDa degraded polypeptide.

The values for both *S* (sedimentation coefficient; left-most graph and table) and *Rs* (Stokes radius) were determined. *Rs* was determined according either to Siegel and Monty (middle graph and middle table; ref. 1) or to Erickson (right graph and right table; ref 2). The native molecular weight (Native Mw.) and the oligomeric state (Multimer), calculated with the equation below the table (left side), are presented in the middle and right tables. *S_{max}/S* values were also calculated and are presented in the same tables. General interpretation of *S_{max}/S*, indicative of molecular shape, is given in a box below the tables.



Supplementary Figure S12

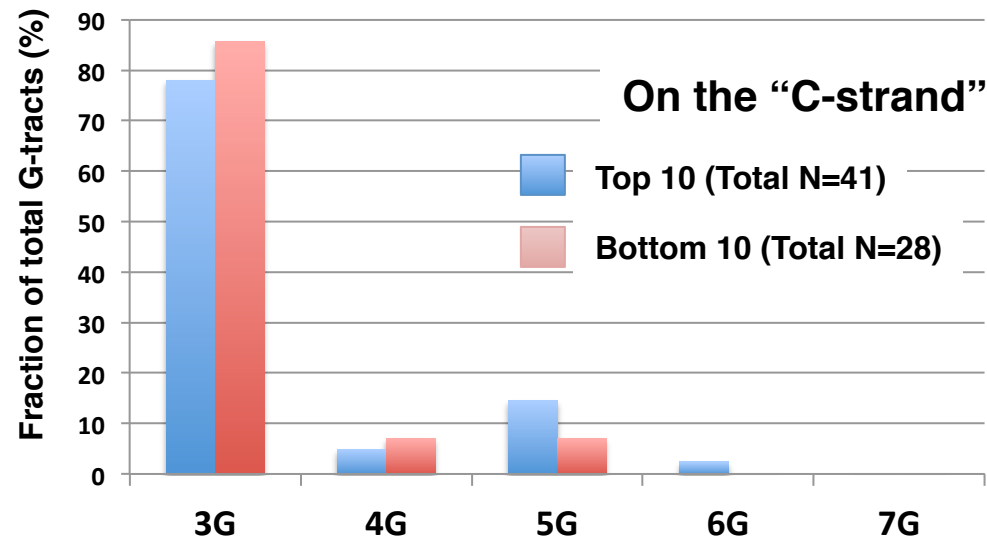
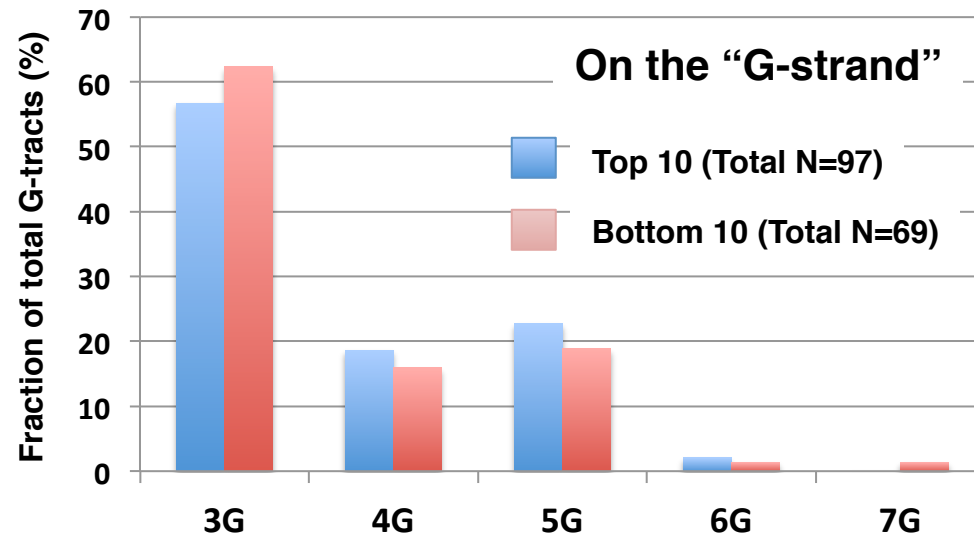


Taz1BS (ds DNA): TTACAGGTTACAGGTTACAGG/ CCTGTAACCTGTAACCTGTAA

Supplementary Figure S12.

Effect of Taz1 on interaction of Rif1 with DNA

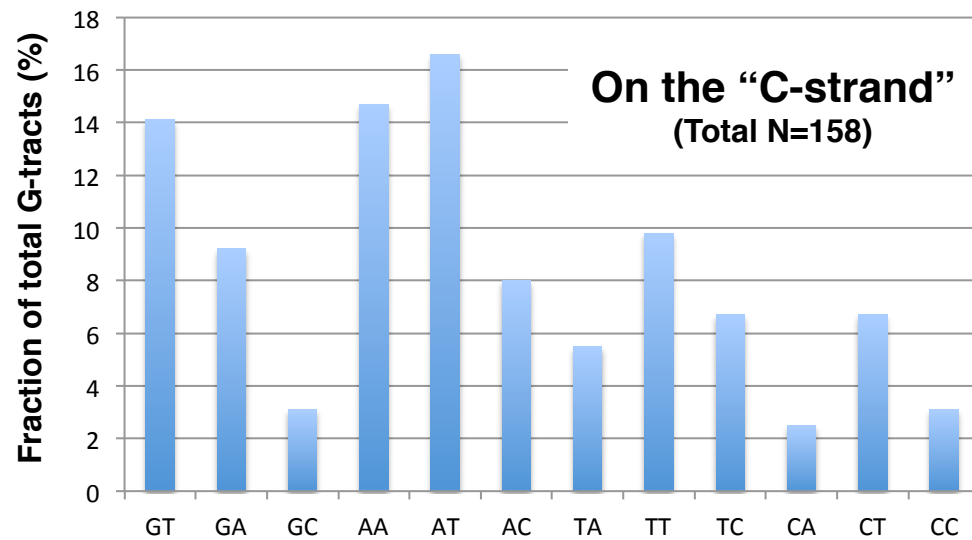
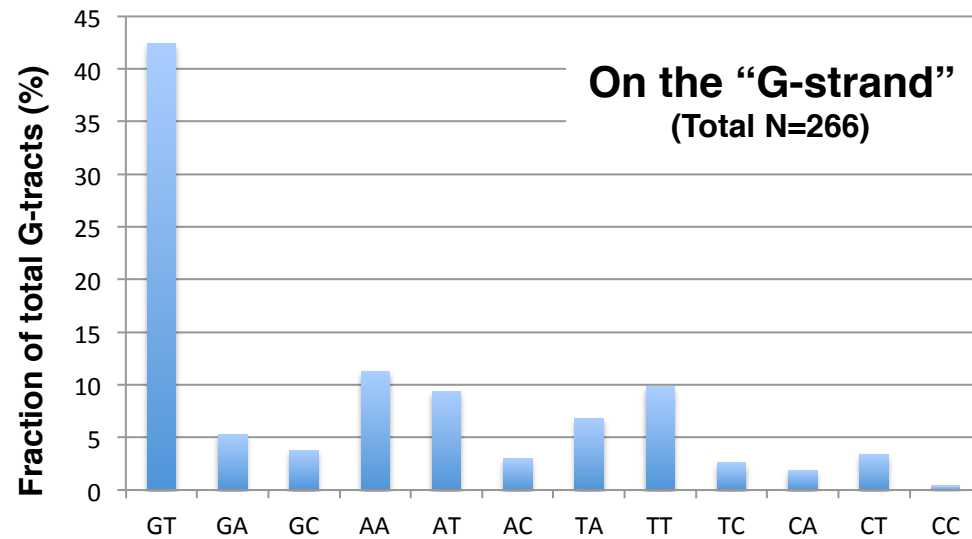
A. Purification of fission yeast Taz1 protein (663 aa). His₆-Rif1-Flag₃ protein expressed in human embryonic kidney 293T cells was purified by consecutive anti-Flag and nickel columns. Eluates from nickel column were analyzed on SDS-PAGE (5-20% gradient gel) and stained by silver. **B.** ³²P-end labeled Htelo3 DNA (0.25 pmole) was mixed with the combination of Taz1 and Rif1 proteins, as indicated, and was analyzed on 8% PAGE (29:1; 1x TBE, 40% 50 mM KCl and 1 PEG200). **C.** and **D.** ³²P-end labeled Taz1BS dsDNA (0.25 pmole) was mixed with the combination of Taz1 and Rif1 proteins, as indicated, and was analyzed on 6% PAGE (29:1; 1xTBE and 10% glycerol) (C) or 8%PAGE (29:1; 1xTBE, 50 mM KCl, and 40% PEG200) (D). Ticks in the fX174/ HaeIII markers (B) represent 600, 310, 271/281, 234, 194, 118 and 72 bp.



Supplementary Figure S14.

The length of G-tracts present in Rif1BS.

The sequences of the 1 kb segments encompassing the 35 Rif1BS were analyzed. All the G-tracts equal or longer than 3 were extracted, and the numbers of 3G~7G were scored for the "top 10" and "bottom 10" Rif1BS. The analyses were conducted for both strands. The G-strand represents the strand on which the G-tracts of Rif1CS are present, and the C-strand the other strand. Fractions of each G-tract are shown.

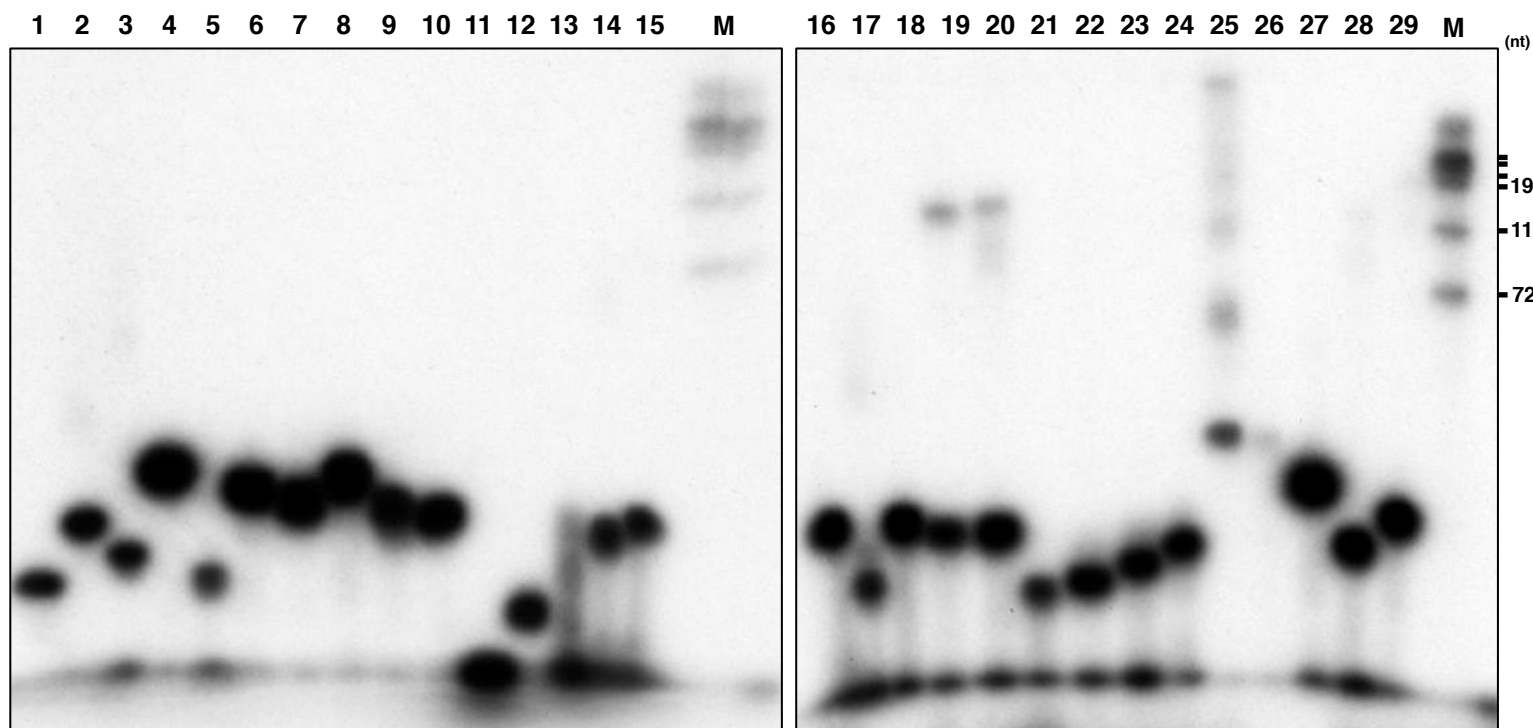


XX in XXG_{n(n=3 or longer)}

Supplementary Figure S15.

Frequency of various dinucleotide sequences preceding the G-tracts (n=3 or longer) in Rif1BS.

The sequences of the 1 kb segments encompassing the 35 Rif1BS were analyzed. All the G-tracts equal or longer than 3 were extracted, and the preceding dinucleotides were scored. The analysis was conducted on both strands. The G-strand represents the strand on which the G-tracts of Rif1CS are present, and the C-strand the other strand. Fractions of each dinucleotide are shown.



Supplementary Figure S16.

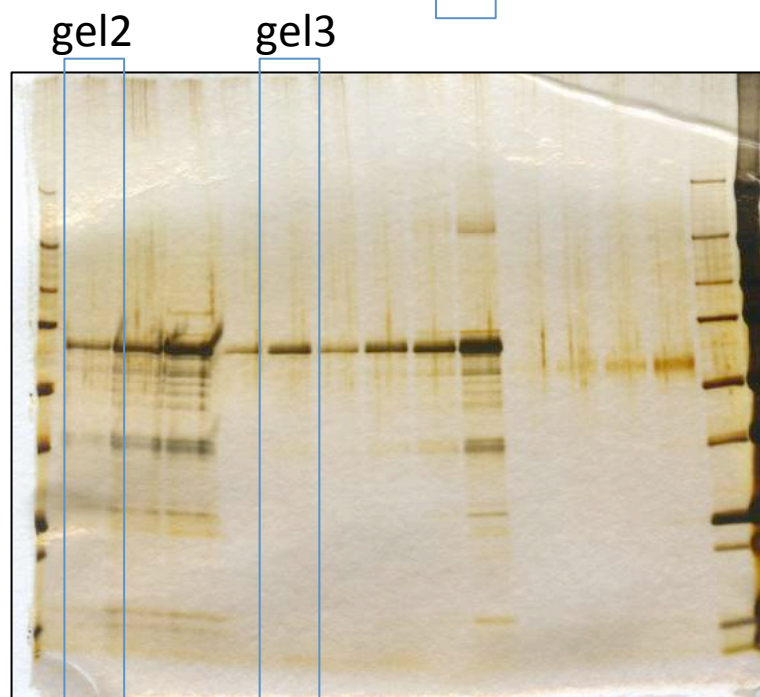
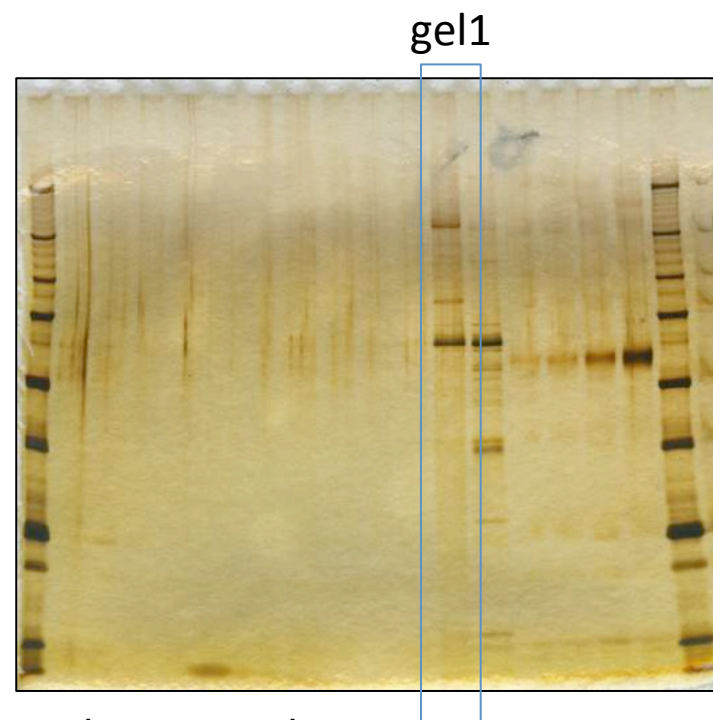
Analyses of labeled oligonucleotides on denaturing PAGE

containing 8M urea.

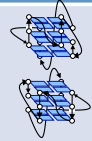



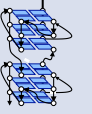





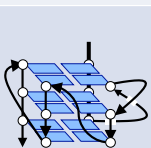
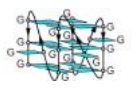
Oligonucleotides used in this studies were end-labeled by T4 polynucleotide kinase, purified by gel filtration, (spun column) denatured in 95% formamide containing 5 mM EDTA and 20 mM LiCl, and were run on 15% PAGE containing 8M urea in 0.5x TBE. T₆G₂₄ appears as a ladder, since it forms secondary structures even in the presence of 8M urea. The bands at the bottom of the gels are free labeled nucleotides that remained after gel filtration purification of labeled oligonucleotides. In lane11, the oligonucleotide (12mer) was too small to be separated on this gel.

No	Name	Sequence	Length (nt)
1	T95_2T	TTGGGTGGGTGGGTGGGT	18
2	Pu24T	TGAGGTTGGTGAGGTTGGGA	21
3	CEB1	AGGGGGAGGGAGGGTGG	18
4	CEB25	AAGGTTGGGTGTAAGTGTGGGTGGG T	26
5	93del	GGGTGGGAGGAGGGT	16
6	HT	TTGGGTTAGGTTAGGTTAGGGA	25
7	Htelo1	TAGGTTAGGTTAGGTTAGGG	23
8	Htelo2	TAGGTTAGGTTAGGTTAGGTT	25
9	Htelo3	GGTTAGGTTAGGTTAGGTT	22
10	Htelo4	AGGCTAGGCTAGGCTAGGG	22
11	DX1	CGCGAATTCGCG	12
12	DX2	ATCTGAGAATCAGAT	15
13	TERC18(RNA)	GGGUUGCGGAGGUGGGC	18
14	CEB1_A	AGGAGAGGAGGAGGGTGG	19
15	CEB1_TA	AGGTTAGGAGGAGGGTGG	20

No	Name	Sequence	Length (nt)
16	CEB1_TTA	AGGGTAGGGAGGAGGGTGG	21
17	CEB1_ΔGG	AGGGGGAGGGAGGGT	16
18	Htelo1_A	TAGGGAGGGTTAGGTTAGGG	21
19	Htelo1_no_space	TAGGGGGTTAGGTTAGGG	20
20	Htelo1_no_space_2	TAGGTTAGGTTAGGGGG	20
21	93del_ΔG	GGGTGGGAGGAGGGT	15
22	T95_2T_G	TTGGGTGGGTGGGTGGGT	19
23	T95_2T_GG	TTGGGTGGGTGGGTGGGT	20
24	T95_2T_GGG	TTGGGTGGGTGGGTGGGT	21
25	T ₆ G ₂₄	TTTTTGGGGGGGGGGGGGGGGGG GGGG	30
26	T ₆ (GA) ₁₂	TTTTTTGAGAGAGAGAGAGAGAGAG AGAGA	30
27	Htelo4_3nt_pacer	AGGCTAGGCTACTAGGCTAGGG	25
28	Htelo4(GGGGGG)	AGGCTAGGCTAGGGGG	19
29	Htelo4(GGGGGG)_3nt_spacer	AGGCTAGGCTACTAGGGGG	22



Supplementary Figure S17.
Original full-length gels for data in
Supplementary Figure S1.
Figure S1A lane 1: gel1
Figure S1A lane 2: gel2
Figure S1A lane 3: gel3

Name	Sequence	Structure	Topology	References
CEB1	AGGGGGGAGGGA GGGTGG		Parallel-type (dimeric form)	Adrian, M., Ang, D. J., Lech, C. J., Heddi, B., Nicolas, A. & Phan, A. T. Structure and conformational dynamics of a stacked dimeric G-quadruplex formed by the human CEB1 minisatellite. <i>J. Am. Chem. Soc.</i> 136, 6297-6305 (2014).
CEB25	AAGGGTGGGTGTA AGTGTGGGTGGGT		Parallel-type	Amrane, S., Adrian, M., Heddi, B., Serero, A., Nicolas, A., Mergny, J. L. & Phan, A. T. Formation of pearl-necklace monomeric G-quadruplexes in the human CEB25 minisatellite. <i>J. Am. Chem. Soc.</i> 134, 5807-5816 (2012).
T95_2T	TTGGGTGGGTGGG TGGGT		Parallel-type	Do, N. Q. & Phan, A. T. Monomer-dimer equilibrium for the 5'-5' stacking of propeller-type parallel-stranded G-quadruplexes: NMR structural study. <i>Chem. Eur J.</i> 18, 14752-14759 (2012).
Pu24T (c-myc)	TGAGGGTGGTGAG GGTGGGGA(AGG)		Parallel-type	Chung, W. J., Heddi, B., Hamon, F., Teulade-Fichou, M. P. & Phan, A. T. Solution structure of a G-quadruplex bound to the bisquinolinium compound Phen-DC(3). <i>Angew. Chem. Int. Ed.</i> 53, 999-1002 (2014). Phan, A. T., Kuryavyi, V., Gaw, H. Y. & Patel, D. J. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. <i>Nat. Chem. Biol.</i> 1, 167-173 (2005). Heddi, B., Cheong, V. V., Martadinata, H. & Phan, A. T. Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution structure of a peptide-quadruplex complex. <i>Proc. Natl. Acad. Sci. U. S. A.</i> 112, 9608-9613 (2015).
93del (aptamer)	GGGGTGGGAGGAG GGT		Parallel-type (dimeric form)	Phan, A. T., Kuryavyi, V., Ma, J. B., Faure, A., Andreola, M. L. & Patel, D. J. An interlocked dimeric parallel-stranded DNA quadruplex: a potent inhibitor of HIV-1 integrase. <i>Proc. Natl. Acad. Sci. U. S. A.</i> 102, 634-639 (2005).
HT (Human telomere)	TTGGGTTAGGGTT AGGGTTAGGGA		Hybrid-type (form 1)	Chung, W. J., Heddi, B., Tera, M., Iida, K., Nagasawa, K. & Phan, A. T. Solution structure of an intramolecular (3 + 1) human telomeric G-quadruplex bound to a telomestatin derivative. <i>J. Am. Chem. Soc.</i> 135, 13495-13501 (2013). Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L. & Patel, D. J. Structure of the human telomere in K ⁺ solution: an intramolecular (3 + 1) G-quadruplex scaffold. <i>J. Am. Chem. Soc.</i> 128, 9963-9970 (2006).
Htelo1 (Human telomere)	TAGGGTTAGGGTT AGGGTTAGGG		Hybrid-type (form 1)	Phan, A. T., Kuryavyi, V., Luu, K. N. & Patel, D. J. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K ⁺ solution. <i>Nucleic Acids Res.</i> 35, 6517-6525 (2007).
Htelo2 (Human telomere)	TAGGGTTAGGGTT AGGGTTAGGGTT		Hybrid-type (form 2)	Phan, A. T., Kuryavyi, V., Luu, K. N. & Patel, D. J. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K ⁺ solution. <i>Nucleic Acids Res.</i> 35, 6517-6525 (2007).
Htelo3 (Human telomere)	GGGTTAGGGTTAG GGTTAGGGT		Anti-parallel-type	Lim, K. W., Amrane, S., Bouaziz, S., Xu, W., Mu, Y., Patel, D. J., Luu, K. N. & Phan, A. T. Structure of the human telomere in K ⁺ solution: a stable basket-type G-quadruplex with only two G-tetrad layers. <i>J. Am. Chem. Soc.</i> 131, 4301-4309 (2009).
Htelo4 (Human telomere)	AGGGCTAGGGCTA GGGCTAGGG		Chair-type	Lim, K. W., Alberti, P., Guedin, A., Lacroix, L., Riou, J. F., Royle, N. J., Mergny, J. L. & Phan, A. T. Sequence variant (CTAGGGG)n in the human telomere favors a G-quadruplex structure containing a G.C.G.C tetrad. <i>Nucleic Acids Res.</i> 37, 6239-6248 (2009).
TERC18 (Human Telomerase RNA (hTERC))	GGGUUGCGGA GGGUGGGC		Parallel-type (proposed) + dimer (in equilibrium)	Martadinata, H. & Phan, A. T. Formation of a Stacked Dimeric G-Quadruplex Containing Bulges by the 5' -Terminal Region of Human Telomerase RNA (hTERC). <i>Biochemistry</i> 53, 1595-1600 (2014).
G ₁₅	GGGGGGGGGG GGGGG		a propeller-type parallel-stranded G-quadruplex	Sengar, A., Heddi, B. and Phan, A. T. Formation of G-quadruplexes in poly-G sequences: structure of a propeller-type parallel-stranded G-quadruplex formed by a G ₁₅ stretch. (2014) <i>Biochemistry</i> 53, 7718-7723.