



Human Pif1 helicase is a G-quadruplex DNA binding protein with G-quadruplex DNA unwinding activity

Cyril Mark Sanders

► To cite this version:

Cyril Mark Sanders. Human Pif1 helicase is a G-quadruplex DNA binding protein with G-quadruplex DNA unwinding activity. *Biochemical Journal*, 2010, 430 (1), pp.119-128. 10.1042/BJ20100612 . hal-00506523

HAL Id: hal-00506523

<https://hal.science/hal-00506523>

Submitted on 28 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Human Pif1 helicase is a G-quadruplex DNA binding protein with G-quadruplex DNA unwinding activity

Cyril M. Sanders

Address for correspondence: Cyril Sanders, PhD, Institute for Cancer Studies, University of Sheffield, Beech Hill Rd., Sheffield, S10 2RX, UK.

Tel: +44 114 2712482; Fax: +0114 2713892; E-mail: c.m.sanders@sheffield.ac.uk

Running head: hPif1 G4 DNA interactions

Abbreviations used: BLM, Bloom's syndrome helicase; DMS, dimethylsulphate; dsDNA, double-stranded DNA; DTT, dithiothreitol; FANCD1, Fanconi anaemia group D protein; GST, glutathione-S-transferase; G4, G-quadruplex; hPif1, human Pif1 protein; hPifHD, human Pif1 protein helicase domain; NP-40, nonyl phenoxypolyethoxylethanol; RPA, replication protein A; ssDNA, single-stranded DNA; TB, Tris-borate; TBE, Tris-borate EDTA; TE, Tris-EDTA; WRN, Werner syndrome helicase.

SYNOPSIS

Pif1 proteins are helicases that in yeast are implicated in the maintenance of genome stability. One activity of *Saccharomyces cerevisiae* Pif1 is to stabilize DNA sequences that could otherwise form deleterious G-quadruplex (G4) structures by acting as a G4 resolvase. Evidence is presented here that human Pif1 (hPif1, nuclear form) is a G4 DNA binding and resolvase protein and that these activities are properties of the conserved helicase domain (amino acids 206-620 of 641, hPifHD). hPif1 preferentially bound synthetic G4-DNA relative to ssDNA, dsDNA and a partially single-stranded duplex DNA helicase substrate. G4 DNA unwinding but not binding required an extended (>10 nucleotide) 5' ssDNA tail and in competition assays G4 DNA was an ineffective suppressor of helicase activity compared to ssDNA. These data suggest a distinction between the determinants of G4-DNA binding and the ssDNA interactions required for helicase action and that hPif1 may act on G4 substrates by binding alone or as a resolvase. Human Pif1 could therefore have a role in processing G4 structures that arise in the single-stranded nucleic acid intermediates formed during DNA replication and gene expression.

Key words: Pif1 helicase, G-quadruplex DNA, DNA repair, protein-DNA interactions, genome stability, DNA replication

INTRODUCTION

DNA can adopt structures other than the canonical B-form duplex stabilized by Watson-Crick base pairing. Structures such as hairpins, including the complex *i*-motifs [1,2], triplex DNA [3] and non-B duplex DNA (*e*-motif; [4]) frequently contain triplet repeat sequences prone to replication errors. The expansion of these triplets promotes the formation of non-B form DNA during replication and repair resulting in genetic instability. In turn, this is thought to be the underlying cause of several hereditary diseases including Huntington's disease [5]. G quadruplex (G4) DNA is an unusual DNA secondary structural form arising from the unique ability of guanosine to hydrogen bond with itself in solution [6]. This can result in highly stable planar rings of four guanines stabilized through Hoogsteen bonds. Although G4 structures were first observed with guanosine, its mononucleotides and synthetic poly(G) DNA sequences [7], they were later found in natural sequences from chromosomal telomeres, rDNA gene clusters and immunoglobulin gene switch regions [8-10]. More recently, they have been shown to be particularly enriched in the upstream promoter regions of genes from higher and lower eukaryotes and may have a role in gene regulation [11,12]. The minimum requirement for formation of these unimolecular anti-parallel quadruplexes is four separate stretches of at least three or four contiguous guanines [7].

Formation of G4 DNA *in vivo* has been observed by electron microscopy in human G-rich sequences transcribed in bacteria [13] and immunologically with antibodies raised to telomeric G4 DNA that react with *Stylocheilichia lemnae* macronuclei [14]. Genetic studies in *C. elegans* have also demonstrated that G4 DNA forming sequences are intrinsically mutagenic and are dependent on the FANCD1 homologue *dog-1* to prevent genomic instability at G4 DNA sites [15]. In man, Fanconi anaemia is a heritable cancer predisposition disorder of which FANCD1 is one of a group of proteins associated with the condition (reviewed in [16]). FANCD1 is a helicase superfamily 2 protein that unwinds DNA in the 5'→3' direction. The *in vitro* substrates of FANCD1 include forked molecules and D-loops suggesting functions in DNA replication re-start, recombination and repair [17,18]. Recently it has been shown that FANCD1 can also unwind G4 DNA *in vitro* [19]. Potentially, FANCD1 may have a role in stabilizing the 370 000 or so G4 tracts predicted in the human genome [20,21], and G4 instability could be related to the genomic instability seen in human tumours.

FANCD1 however is not the only human helicase known to resolve G4 structures *in vitro*. The Bloom's and Werner's syndrome RecQ family helicases, BLM and WRN, unwind DNA in the 3'→5' direction and function in several aspects of DNA metabolism including transcription, replication and repair (reviewed in [22]). Both helicases also unwind G4 DNA *in vitro* and the preferred substrate for BLM appears to be G4 DNA [23,24]. These data suggest that failure to unwind G4 DNA may also contribute to the genetic instability observed in Bloom's and Werner's cells. Moreover, studies of the *S. cerevisiae* Sgs1 helicase indicate that the G4 DNA processing activity of the RecQ family helicases is conserved [11,25]. It has also been established recently that the yeast Pif1 protein, a 5'→3' helicase [26], processes G-quadruplex forming sequences *in vivo* and *in vitro* [27]. This raises the interesting question of whether the human enzyme functions similarly, given that there is not always an equivalence of function even between conserved proteins of closely related species. Here, the G4 DNA binding and unwinding properties of a recombinant nuclear isoform of human Pif1 (hPif1) have been analysed. Surprisingly, hPif1 bound G4 sequences effectively compared to ssDNA and partially single-stranded duplex DNA substrates. This activity was intrinsic to the core helicase domain and was not dependent on a lengthy (≥55 bases) ssDNA tail otherwise required for the enzyme to dissociate synthetic G4 DNA test substrates. These data imply that the enzyme may have direct and indirect roles in processing G4 structures that arise in the single-stranded nucleic acid intermediates formed during DNA replication and gene expression.

EXPERIMENTAL

Expression and Purification of hPif1 proteins

The purified full-length nuclear form of hPif1 and the core helicase domain (amino acids 206 to 620 of 641) were generated in *E. coli* BL21(DE3) as glutathione S-transferase (GST) fusion proteins as previously described [28]. Briefly, the proteins were purified free of the GST moiety following GST chromatography, thrombin cleavage and ion exchange and size exclusion chromatography.

DNA substrates and DNA labeling

All oligonucleotides were purchased from Sigma Aldrich. The substrate PST55 consists of a 20 base pair duplex with a 55 base 5' T tail, as described before [28]. The G-quadruplex (G4) DNA substrates were formed from oligonucleotides with either a run of four G residues (GGGG), or two runs of four G residue separated by four T residues (GGGGTTTGGGG). To avoid confusing nomenclature, the substrates with these repeats are named V4 (one four G repeat) or OX (two four G repeats) in reference to the DNA sequences where the G-quadruplex forming elements were first described (the V4 DNA sequence and the *Oxytrichia nova* telomere (OX); see reference [19]). The oligonucleotide substrates contained these motifs (V4/OX) with 5' T tail extensions of variable length as indicated. The V4 substrate also had a single 3' C residue required for G-quadruplex stability (see Figure 1). G4 DNA was formed using the general guidelines described by Sen and Gilbert [7]. Oligonucleotides (1 mM in TE buffer) were heated to 95°C for 2 minutes, cooled on ice, centrifuged, and the solution adjusted to 250 μ M oligo in 0.1 M KCl, 1.9 M NaCl, 1 x TE buffer. After overlaying with mineral oil the solution was incubated for 16-20 hours at 60°C. G4 DNA was stable when stored at -30°C. Substrates were end-labeled with 32 P using polynucleotide kinase and [γ - 32 P]ATP (6000 Ci/mmol) and resolved on 8% (19:1) poly-acrylamide gels formed with 1 x TBE running buffer and 5 mM KCl (1 x TBE/5 mM KCl running buffer). After identification by autoradiography, bands were excised and DNA soak-eluted in 0.1 x TE, 15 mM KCl, 50 mM NaCl before quantification. G-quartet formation was confirmed by methylation protection [29] and analysis of products by urea-PAGE (10% gels).

Helicase reactions

Helicase reactions were performed in 20 mM HEPES-NaOH pH 7.5, 20 mM NaCl, 5 mM MgCl₂, 2 mM ATP, 1 mM DTT, 0.1 mg/ml BSA and 0.1 % v/v NP-40 as described previously [28]. Reactions were incubated at 20°C for 30 minutes and terminated by the addition of 0.2 volumes of 120 mM EDTA, 0.6 % w/v SDS, 60 % v/v glycerol and 0.1 % w/v bromophenol blue. Products were resolved on 8 % (19:1) poly-acrylamide gels and visualized and quantified following exposure of dried gels to a phosphorimaging plate (Fujifilm). Graphed data are for a minimum of three repeats with standard error of the mean (SEM) indicated.

DNA binding reactions and gel-shift assay

The reaction buffer for hPif1 DNA binding was 20 mM HEPES-NaOH pH 7.5, 75 mM NaCl, 5 % v/v glycerol, 1 mM DTT, 0.1 mg/ml BSA and 0.1 % v/v NP-40. Reactions were incubated for 20 minutes at 20°C before resolving complexes on 5 % poly-acrylamide gels (29:1) using 0.25 x TBE running buffer. To analyze complexes formed in the presence of ATP, the

binding reactions also contained 5 mM ATP/MgCl₂. The gels were cast with 0.25 x Tris-borate buffer containing no EDTA (TB) and 0.1 mM ATP/MgCl₂ and the running buffer was 0.25 x TB, 0.1 mM ATP/MgCl₂. However, comparable results were obtained even when the complexes were run on standard TBE gels. Visualization, quantification and graphical representation of the data were as described above.

ATPase assays

ATPase activity was measured as previously described [28] using the charcoal-binding assay of Iggo and Lane [30] to determine phosphate released. The reaction buffer was 20 mM HEPES-NaOH pH 7.5, 75 mM NaCl, 5 mM MgCl₂, 5 mM ATP, 0.0125 mM [γ -³²P]ATP (6000 Ci/mmol), 1 mM DTT, 0.1 mg/ml BSA and 0.1 % v/v NP-40. Reactions (20 μ l) were incubated at 20°C for 10 minutes.

RESULTS

Formation and analysis of G-quadruplex (G4) DNA

The sequences employed for G4 DNA formation (GGGGC, V4 and GGGGTTTTGGGG, OX; G-quadruplex forming guanines underlined) were derived from known G4 forming DNA sequences but the 5' flanking ssDNA sequences were modified to be runs of T residues. The structures of the V4/OX G4 substrates with 55 T residue 5' extensions (T55-OX G4 and T55-V4 G4) are shown in Figure 1A. Figure 1B shows the ^{32}P -labeled products of G4 annealing reactions named according to the nomenclature of Sen and Gilbert [7]. The distribution of products was the same as those observed previously for the native OX and V4 sequences [19]. Products corresponding to G4 DNA and G'2 DNA, as indicated, were easily identified with the tetramolecular parallel G4 products most prominent, as expected for the hybridization conditions. G'2 DNAs are bimolecular anti-parallel quadruplexes [7]. In the case of the V4 substrate, the high molecular weight products have both been designated as G4 DNA [19] and are likely to be tetramolecular parallel and anti-parallel quadruplexes. The V4 product of lowest mobility and greatest abundance was analysed in detail in Figure 1C, where the formation of G-quartets was confirmed by methylation protection. For each substrate, the brackets donate the G-runs that are accessible to DMS modification in the ssDNA oligonucleotides but not the G4 DNA. Subsequent analysis focused on G4 structures formed with the OX motif (two runs of four guanines) given their superior stability.

hPif unwinds G4 DNA

hPif1 requires an extended 5' ssDNA tail of 55 bases or more for optimal initiation of unwinding of simple partially single-stranded duplex test substrates (ss/dsDNA substrates). Our reference substrate for hPif1 unwinding is therefore a 20 bp duplex with a 55 base 5' T tail, PST55 [28]. Accordingly, G4 DNA substrates formed using oligonucleotides with a 55 base 5' T tail were tested first in the unwinding assay. Strand displacement of ^{32}P -labelled substrates was measured by separating reaction products on native poly-acrylamide gels. As shown in Figure 2A, T55-V4 G4 (lanes 1-10) and T55-OX G4 (lanes 11-20) DNA were both unwound by full-length hPif1 and the core helicase domain hPifHD. With ss/dsDNA substrates like PST55, the unwinding activity of full-length hPif1 we have observed rises steeply from little or no unwinding at 1 nM to a peak at 2 nM and decreases thereafter. The reason why this activity profile is observed is unclear, although a high degree of co-operativity between monomers could be required to initiate unwinding [28], while non-specific protein-DNA binding could suppress unwinding at high protein concentrations. The unwinding activity for G4 substrates observed with hPif1 or hPifHD is therefore relatively robust in comparison to the substrate PST55 where a short duplex length (20 base pairs) is displaced under the same biochemical conditions, as shown in the graph in Figure 2A. It should be stressed however that an absolute comparison of relative unwinding of G4 compared to duplex DNA substrates is problematic since they will each be influenced differently by factors such as DNA context, substrate length and competing re-annealing reactions. The data also show that the unwinding activity of the full-length enzyme appears to be superior to the hPifHD core helicase domain. Curiously, T55-V4 G4 DNA was more efficiently unwound than T55-OX G4 DNA by hPifHD, while the converse was observed with hPif1, the full-length enzyme. The V4 annealing product designated G4* in Figure 1B was also resolved by the enzyme (not shown). As shown in Figure 2B, the efficiency of G4 DNA unwinding by hPif1 was also highly sensitive to the enzyme-substrate ratio. For example, at 4 nM hPif1 and substrate (T55-OX G4) concentrations of 0.05, 0.1 and 0.2 nM (lanes 3, 10 and 17 respectively), 28.3, 8 and 0 pmols of substrate were unwound per ml of reaction under the standard assay conditions. These data may also indicate that unwinding of synthetic G4 DNA substrates

requires cooperation between monomers that are limiting at high substrate concentrations. This could be compounded if the specific activity of the enzyme preparations were low such that the stoichiometry of protein:DNA is different from expected. However, measurements of DNA binding activity indicate that both hPif1 and PifHD have high DNA binding activity (see below). Furthermore, as expected for helicase action, no unwinding of the substrate T55-OX G4 was detected without ATP in the reactions and re-formation of G4 DNA from denatured substrate was minimal under the reaction conditions used (see Supplementary Figure 1).

Unwinding of G4 DNA by hPif1 requires substrates with an extended 5' ssDNA tail

As mentioned above, the unwinding of ss/dsDNA test substrates by hPifHD/hPif1 requires an extended 5' ssDNA tail [28]. Accordingly, the requirements of 5' ssDNA tail length for unwinding of G4 DNA substrates were tested using the OX G-quartet motif (GGGGTTTGGGG) with 55, 25 and 10 base poly T 5' extensions, substrates T55-OX G4, T25-OX G4 and T10-OX G4 respectively (Figure 3). T25-OX G4 (lanes 6-10) was unwound less efficiently than T55-OX G4 (lanes 1-5) while unwinding of T10-OX G4 was barely detectable (lanes 11-15), as summarized in the graphed data below the autoradiogram. A substrate with 5 base poly T 5' extensions also failed to unwind (T5-OX G4, data not shown). Therefore, as with ss/dsDNA molecules the initiation of unwinding of G4 DNA substrates requires an extended 5' ssDNA tail.

The specific requirement for an extended ssDNA tail for G4 DNA unwinding by hPif1 was tested further using a competition assay (Figure 4). Helicase reactions with the substrate T55-OX G4 were assembled with increasing concentrations of unlabelled competitor DNAs, T10-OX G4, T55-OX G4, T55 ssDNA and the reference helicase substrate PST55, and unwinding determined under the standard assay conditions. To construct the graphical data, the fraction unwound (%) relative to the reaction with no competitor (set at 100 %) was plotted against the molar excess of competitor. As shown in Figure 4A, T10-OX G4 (lanes 3-7) was a relatively ineffective inhibitor of unwinding compared to all other substrates with 55 bases of ssDNA. The concentration of competitors T55 and PST55 required to give 50% inhibition of unwinding were 6-8 fold less than that of T10-OX G4 (graph on the right). T55-OX G4 was a very effective competitor and greater than 50% inhibition was obtained with a molar equivalent of unlabelled substrate. This is consistent with the results described above (Figure 2B), where high ratios of substrate to enzyme were inhibitory for unwinding. Sensitivity of the unwinding reaction of substrate PST55 to the addition of competitor DNAs was also tested in parallel under the same conditions (Figure 4B). These reactions displayed the same profile of sensitivity to the various competitors as substrate T55-OX G4 (T55-OX G4>PST55>T55>T10-OX G4), except that T10-OX G4 DNA resulted in a modest but progressive stimulation of unwinding (~ 0.5 fold), over the concentration range tested (see graph on the right). Also, a slight stimulation of unwinding was observed with T55 ssDNA and T55-OX G4 at low concentrations of competitor (2.5 fold molar excess), followed by inhibition of unwinding. Such stimulatory effects have also been observed for other helicases in DNA competition assays (see references 19 and 23). Although the reasons for these effects are unclear, the presence of additional DNA could stabilize the enzyme non-specifically or the "active" form of the enzyme could be a multimeric complex whose assembly is promoted by DNA. These data described above are therefore consistent with the requirement for an extended ssDNA interaction in substrates unwound by hPif1.

hPif1 binds G4 DNA

Our previous analysis of hPifHD/hPif1 had shown a correlation between helicase substrate binding and unwinding, as well as high affinity binding to ssDNA molecules greater than 30-35 residues [28]. DNA binding assays with hPif1 were performed in the absence of nucleotide cofactors and complexes resolved on poly-acrylamide gels before visualization by phosphorimaging. hPif1 bound T55-OX G4 DNA with apparent high affinity in the gel-shift assay, readily shifting the G4 DNA to high molecular weight complexes that failed to enter the gel. The mobility-shift pattern at low protein concentrations was considered difficult to interpret, as complexes were difficult to resolve from free probe. However, retardation of the probe is evident as an up-shift and broadening of the probe band (Figure 5, lanes 17-20). Notwithstanding this, binding to T55-OX G4 is difficult to interpret as it has both G4 and extended ssDNA character. Binding to the substrate T10-OX G4 was also tested. hPif1 failed to bind to the 22 base ssDNA precursor of T10-OX G4 as expected (T10-OX ss, lanes 1-4), or the same DNA annealed to its complementary strand (T10-OX ds, lanes 5-8). However, hPif1 bound T10-OX G4 DNA with apparent high affinity, and higher than the reference helicase substrate PST55 (lanes 9-12 compared to lanes 13-16 and graphed data shown below the autoradiogram). Similar results were obtained with the substrate T5-OX G4 with only a 5 base T tail (data not shown). These observations were surprising since in all previous cases we have observed a correlation between the ability of hPif1/hPifHD to bind a substrate in the gel-shift assay and unwind it in the helicase assay ([28] and unpublished data). Furthermore, a distinct "ladder-like" binding pattern was observed with substrate T10-OX G4 as the protein concentration increased (lanes 10-12). This was not observed with other partially single-stranded duplex DNA substrates (lanes 14-16) or ssDNA molecules (see below). It is possible to discern approximately eight discrete bands in the ladder.

G4 DNA binding activity is intrinsic to hPifHD

Binding of the core helicase domain hPifHD to G4 DNA was also tested, since the initiation of unwinding observed with the helicase domain (Figure 2A) could be simply due to the extended ssDNA tail in the test substrates. As shown in Figure 6A, hPif1 and hPifHD both bound the T10-OX G4 substrate with similar affinity in the gel-shift assay (lanes 2-4 compared to 5-7 and the graphs below for hPif1 binding (B) and hPifHD binding (C)). The ladder-like pattern of binding to G4 DNA was also observed with hPifHD, and not with the ssDNA substrate T55 shown for comparison where binding extents were similar (lanes 12-14). The binding affinity of hPif1 for the T55 oligonucleotide also appeared moderately lower than that for T10-OX G4 DNA (lanes 9-11 compared to 2-4 and graphical data below on the left). However, hPifHD appeared to bind T55 and T10-OX G4 to similar extents (graph on the right). Overall, the gel-shift analysis with the various DNA substrates (ssDNA, G4 and partially single-stranded; Figures 5 and 6) indicate a high level of DNA binding activity for the hPif1/hPifHD proteins. Supplementary Figure 2, showing saturation DNA binding experiments with T10-OX G4 DNA, supports the notion that a major proportion of the protein in the enzyme preparations is active for DNA binding.

G4 DNA binding with ATP and stimulation of ATPase activity

Since it is conceivable that DNA binding by hPif1 is substantially altered by nucleotide cofactors, binding to hPif1 DNA substrates was compared with or without ATP using the gel-shift assay. Substrate unwinding complicates an assessment of DNA binding under the exact same helicase conditions (low salt and presence of ATP). However, higher salt concentrations inhibit unwinding [28], making it possible to test binding directly in the presence of ATP. As described previously for hPifHD [28], full-length hPif1 binding to the reference substrate PST55 was reduced

in the presence of ATP (Figure 7A, lanes 2-4 compared to 6-8). At sub-saturating protein concentrations binding is reduced 6-8 fold (graphed data on the right, Figure 7A). With the G4 DNA substrate T10-OX G4, DNA binding was reduced only minimally in the presence of ATP, by about approximately 0.5 fold, at sub-saturating protein concentrations (Figure 7B, lanes 2-4 and 6-8 and graphed data). Furthermore, the ladder-like hPif1 binding pattern observed with T10-OX G4 in the absence of cofactors was also observed in the presence of ATP (lane 8). Therefore in contrast to the ss/dsDNA substrate PST55, substantial differences in T10-OX G4 DNA binding are not observed with and without ATP.

We have shown previously that the ATPase activity of hPif1 is dependent on ssDNA binding and that ssDNA binding is DNA length dependent [28]. In that study, effective DNA binding and stimulation of ATPase activity was observed with poly T oligonucleotides 30 bases or greater. Accordingly, the ability of G4 DNA to stimulate hPif1 ATPase activity was tested in comparison to the substrates used in other aspects of this study. The protein concentration was fixed at 50 nM, the concentration of DNAs at 25, 50 or 200 nM and phosphate release determined as a measure of ATP hydrolysis. As shown in Figure 7C, little ATPase activity was measured in the absence of DNA. The stimulation of ATPase activity by T10-OX G4 DNA was low compared to the substrates containing 55 T ssDNA and comparable to the ssDNA precursor of T10-OX G4 (T10-OX ssDNA) at low ratios of DNA to protein (25 nM DNA:50 nM protein, graph on the left). As the relative concentration of DNA was increased in the reaction, to equal and then exceed the protein concentration, the stimulation of ATPase activity increased but did not surpass that of the T55 ssDNA containing DNA molecules. Like the helicase substrate PST55, T55 ssDNA also showed reduced DNA binding affinity in the presence of ATP (data not shown). Therefore, although hPif1 has substantially higher affinity for T10-OX G4 DNA compared to PST55 (and T55 ssDNA) in the presence of ATP (Figure 7A&B), binding of G4 DNA alone is a poor stimulator of nucleotide hydrolysis.

DISCUSSION

hPif1 is a 5' → 3' helicase and member of a highly conserved protein family [26,32]. Genetic studies in yeast and biochemical analysis of the human enzyme indicate that Pif1 has roles in telomere stability, Okazaki fragment processing and the maintenance of genomic stability [28,32-35]. hPif1 is thus one of a cohort of proteins, including the RecQ helicases WRN and BLM, that recognize and process specific DNA structures that could arise during transcription or DNA replication, recombination and repair [17,18,22]. As well as these replication intermediates composed of partially single-stranded duplex DNA, other non-B form DNA structures can form. Some of these, such as G-quadruplex structures, could prevent the completion of replication for example, by impeding polymerase action [36]. New evidence is presented here that human Pif1 binds and unwinds G4 DNA structures *in vitro*. This activity therefore appears to be conserved between yeast and human Pif1 proteins, suggesting that the enzyme may be a G4 resolvase in mammalian cells.

hPif1 requires an extended 5' ssDNA tail (≥55 base) in order to initiate efficient unwinding of simple partially single-stranded duplex (ss/dsDNA) substrates [28]. Using synthetic G-quadruplex structures with 55 base T tails (T55-V4 and T-55-OX G4), robust unwinding of G4 DNA was obtained in comparison with the ss/dsDNA substrate PST55, under the same biochemical conditions. However, as noted above, such direct quantitative comparisons between different substrates are problematic since unwinding will be influenced by DNA context (the G4 substrate is quadrivalent for T55 ssDNA for example), duplex/G4 length and competing re-annealing reactions that may not be equivalent. The G4 resolvase activity is intrinsic to the conserved hPif1 helicase domain, hPifHD (Figure 2). The N-terminal PINT domain, recently shown to have DNA binding properties [37], does not appear to influence substrate choice, but does appear to increase the activity of the full-length enzyme relative to the core helicase module, hPifHD, on all substrates tested. As with the unwinding of ss/dsDNA substrates, initiation of G4 DNA unwinding requires an extended ssDNA tail and negligible unwinding of T10 (T10-OX G4) or T5 (T5-OX G4; not shown) tailed G4 substrates was observed. Along with the observation that hPif1/hPifHD only binds ssDNA molecules greater than 30 bases with significant affinity [28], these observations support the notion that hPif1 interacts extensively with ssDNA during unwinding (Figures 2&3). This view is upheld by helicase competition assays (Figure 4), demonstrating preferential inhibition of G4 DNA unwinding by competitors with extended ssDNA tails (55 T residues).

Gel-shift assays demonstrate that hPif1 binds G4 DNA with much greater affinity relative to ssDNA and dsDNA molecules of similar length, but also with increased affinity compared to the ss/dsDNA substrate PST55 (Figures 5,6&7). However, T10-OX G4 DNA (not unwound) was a relatively poor inhibitor of helicase activity in the competition unwinding assay (Figure 4). These observations suggest that the hPif1 helicase core, hPifHD, may contain distinct DNA binding determinants for structure-specific DNA recognition and binding of the ssDNA required for initiation of unwinding. A structure-specific DNA binding domain (DBD) may be critical in targeting hPif1 to substrates containing G4 DNA for example. On the other hand, initiation of DNA unwinding itself may require only the interactions made with a relatively long stretch (~ 55 residues) of ssDNA, explaining the inability of T10-OX G4 to inhibit unwinding effectively. Since the helicase action dissociates the recruiting DNA structure, it is readily conceivable that the structure-specific protein-DNA interaction is uncoupled or dispensable once unwinding is initiated or the enzyme bound to sufficient ssDNA. Indeed, continued interaction with the same DNA molecule through the structure-specific DBD could be an impediment to unwinding. Many helicases translocate on ssDNA during unwinding and this is coupled to ATP hydrolysis. The notion of two distinct DNA binding modules in the helicase core, as described above, is further

substantiated by the observation that although substrate T10-OX G4 binds hPif1 with higher affinity relative to ssDNA containing substrates, it stimulates ATPase activity poorly compared to the latter (Figure 7).

The observations that hPif1 can unwind G4 DNA as well as partially single-stranded duplex substrates like PST55 and more complex structures such as those resembling stalled DNA replication forks [28] is not unlike the findings for other structure-specific helicases such as BLM and WRN. hPif1 therefore appears to use different features of DNA structure to affect substrate recognition. However, there appears to be a striking difference in DNA-complex formation observed by gel-shift for G4 DNA compared to the ss/dsDNA substrates PST55 and also ssDNA. With G4 substrates, a distinct “ladder-like” pattern of DNA binding formed, with or without ATP (Figure 7). This mode of DNA binding appears cooperative and is more reminiscent of the binding of hexameric helicases to DNA [38], although there is no evidence that a hexamer specifically forms with hPif1. The significance of this mode of DNA binding and the structural features underlying it remain to be determined.

The data presented above raise two interrelated questions concerning the *in vivo* substrates and functions of hPif1. They indicate that hPif1 can bind discrete G4 DNA structures. Thus it is conceivable that hPif1 could bind G4 DNA in numerous contexts, whenever single stranded nucleic acid intermediates arise that contain the appropriate guanine composition for G4 DNA formation [15]. This may happen during DNA recombination, repair and replication where local G4 DNA formation out-competes binding of ssDNA binding protein (Replication protein A, RPA), potentially stalling these processes. Equally, the hPif1 G4 substrate may form during transcription in DNA, or within RNA as the yeast enzyme is also active on RNA containing substrates [39]. As has been argued previously [27], since human telomeres contain consensus sequences for G4 DNA formation, the association of hPif1 with telomeres and telomerase [32] could indicate that hPif1 is a telomere G4 DNA binding protein. In yeast, the involvement of Pif1 in telomere stability has been well established [40]. However, the data presented here also suggest that without sufficient flanking 5' ssDNA (>10 nucleotide), G4 DNA may be bound but unwinding will not initiate and G4 DNA will not be resolved. It remains to be determined whether a G4 DNA binding event alone can create the local conditions (sufficient ssDNA) needed for unwinding or whether direct resolvase activity is restricted to DNA substrates with more defined structural features. It is noteworthy that in our previous analysis of hPif1 substrates specificity employing ss/dsDNA molecules there was always a positive correlation between DNA binding affinity and unwinding *in vitro*, contrasting with the observations for G4 substrates. Furthermore, synthetic structures resembling stalled DNA replication forks were effectively bound and unwound by the enzyme. Significantly, the addition of a second dsDNA arm to gapped (10-20 nucleotides) duplex molecules to create such substrates effectively reduced the requirement for extended ssDNA tail to initiate unwinding [28].

As the *in vivo* substrates of Pif1 remain uncertain so does the requirement for the enzyme in mammalian cells. Although highly conserved, the enzyme is essential in fission yeast [41] but not in Baker's yeast and knock out mice have no readily apparent phenotype [42]. The activities of hPif1 may be restricted to maintaining the stability of specific G4 DNA-forming repeat sequences in mammalian cells, as in yeast cells [26]. This activity would be hard to detect and of questionable significance given that Pif1 null mice are normal, a contrasting condition to WRN and BLM cells where DNA instability can be readily detected [43]. It is perhaps more tempting to speculate that hPif1 activity is normally redundant, with sufficient G4 resolvase activity provided by the many other helicases known to have this activity (*e.g.* WRN, BLM, FANCD1 and G4R1; [19,23,24,44]). It is noteworthy that these enzymes, like hPif1, also appear to act upon and resolve multiple contrasting DNA structures. A requirement for Pif1 in mammalian cells may therefore be apparent

only under certain conditions, possibly when one of these other G4 resolvase systems is compromised.

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20100612

ACKNOWLEDGEMENTS

I am grateful to Mark Meuth and Jon Sayers for critically reading the manuscript and Tresa George for help with protein purification.

FUNDING

This work was supported by a Yorkshire Cancer Research programme grant [S305PA] awarded in part to CMS.

REFERENCES

1. Mitas, M., Yu, A., Dill, J., Kamp, T., Chambers, E., and Haworth, I. (1995) Hairpin properties of single-stranded DNA containing a GC-rich triplet repeat: (CTG)₁₅. *Nucleic Acids Res.* **23**, 1050-1059
2. Castati, P., Chen, X., Moyzis, R., Bradbury, E., and Gupta, G. (1996) Structure-function correlations of the insulin-linked polymorphic region. *J. Mol. Biol.* **264**, 534-545
3. Gacy, A., Goellner, G., Spiro, C., Chen, X., Gupta, G., Bradbury, E., Dyer, R., Mikesell, M., Yao, J., Johnson, A., Richter, A., Melançon, B., and McMurray, C. (1998) GAA instability in Friedreich's ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. *Mol. Cell.* **1**, 583-593
4. Zheng, M., Huang, X., Smith, G., Yang, X., and Gao, X. (1996) Genetically unstable CXG repeats are structurally dynamic and have a high propensity for folding. An NMR and UV spectroscopic study. *J. Mol. Biol.* **264**, 323-336
5. McMurray, C. (1999) DNA secondary structure: A common and causative factor for expansion in human disease. *Proc. Natl. Acad. Sci. USA.* **96**, 1823-1825
6. Sanger, W. (1984) "Principals of nucleic acid structure" Springer-Verlag, New York, Berlin, Heidelberg, and Tokyo
7. Sen, D. and Gilbert, W. (1992) Guanine quartet structures. *Methods in Enzymology* **211**, 191-199
8. Sen, D. and Gilbert, W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364-366
9. Williamson, J., Raghuraman, M., and Cech, T. (1989) Monovalent cation induced structure of telomeric DNA: the G-quartet model. *Cell* **59**, 871-880
10. Hanakahi, L., Sun, H., and Maizels, N. (1999) High affinity interactions of nucleolin with G-G-paired rDNA. *J. Biol. Chem.* **274**, 15908-15912
11. Hershman, S., Chen, Q., Lee, J., Kozak, M., Yue, P., Wang, L-S., and Johnson, F. (2008) Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **36**, 144-156
12. Du, Z., Zhao, Y., and Li, N. (2008) Genome wide analysis reveals regulatory role of G4 DNA in Gene transcription. *Genome Res.* **18**, 233-241
13. Duquette, M., Handa, P., Vincent, J., Taylor, A., and Maizels, N. (2004) Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes. Dev.* **18**, 1618-1629
14. Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H., Plückthun, A. (2001) In vitro generated antibodies specific to telomeric guanine-quadruplex DNA react with *Stylonychia lemnae* macronuclei. *Proc Natl. Acad. Sci. USA* **88**, 8572-8577
15. Kruisselbrink, E., Guryev, V., Brouwer, K., Pontier, D., Cuppen, E., and Tijsterman, M. (2008) Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCD1-defective *C. elegans*. *Current Biol.* **18**, 900-905
16. Mathew, C. (2006) Fanconi anaemia genes and susceptibility to cancer. *Oncogene* **25**, 5875-5884
17. Cantor, S., Bell, D., Ganesan, S., Kass, E., Drapkin, R., Grossman, S., Wahrer, D., Sgori, D., Lane, W., Haber, D., and Livingston, D. (2001) BACH1, a novel helicase-like protein, interacts directly with BRCA 1 and contributes to its DNA repair function. *Cell* **105**, 149-160

18. Gupta, R., Sharma, S., Sommers, J., Jin, Z., Cantor, S., and Brosch, R. (2005) Analysis of the DNA substrate specificity of the human BACH1 helicase associated with breast cancer. *J. Biol. Chem.* **280**, 25450-25460
19. London, T., Barber, L., Mosedale, G., Kelly, G., Balasubramanian, S., Hickson, I., Boulton, S., and Hoim, H. (2008) FANCI is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J. Biol. Chem.* **283**, 36132-36139
20. Huppert, J. and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **33**, 2908-2916
21. Todd, A., Johnston, M., and Neidle, S. (2005) Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **33**, 2901-2907
22. Chu, W. and Hickson, I. (2009) RecQ helicases: multifunctional genome caretakers. *Nature Reviews Cancer* **9**, 644-654
23. Sun, H., Karow, J., Hickson, I., and Maizels, N. (1998) The Bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem.* **273**, 27587-27592
24. Fry, M. and Loeb, L. (1999) Human Werner syndrome DNA helicase unwinds tetrahedral structures of the fragile X syndrome repeat sequence d(CGG)_n. *J. Biol. Chem.* **274**, 12797-12802
25. Huber, M., Lee, D., and Maizels, N. (2002) G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition. *Nucleic Acids Res.* **30**, 3954-3961
26. Lahaye, A., Leterme, S., and Foury, F. (1993) Pif1 DNA helicase from *Saccharomyces cerevisiae*. Biochemical characterization of the enzyme. *J. Biol. Chem.* **268**, 26155-26161
27. Ribeyre, C., Lopes, J., Boulé, J.-B., Piazza, A., Zakiam, A., Mergny, J.-L., and Nicolas, A. (2009) The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences *in vivo*. *PLOS Genet.* **5**(5), e1000457
28. George, T., Wen, Q., Griffiths, R., Ganesh, A., Meuth, M., Sanders, C. (2009) Human Pif-1 helicase unwinds synthetic DNA structures resembling stalled DNA replication forks. *Nucleic Acids Res.* **37**, 6419-6502
29. Maxam, A. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc Natl. Acad. Sci. USA* **74**, 560-564
30. Iggo, R. and Lane, D. (1989) Nuclear protein p68 is an RNA-dependent ATPase. *EMBO. J.* **8**, 1827-1831
31. Sen, D. and Gilbert, W. (1992) Novel DNA superstructures formed by telomere-like oligomers. *Biochemistry* **32**, 65-70
32. Zhang, D.-H., Zhou, B., Huang, Y., Xu, L.-X., and Zhou, J.-Q. (2006) Pif1 helicase, a potential *Escherichia coli* RecD homologue, inhibits telomerase activity. *Nucleic Acids Res.* **34**, 1393-1404
33. Foury, F. (1990) These d'agregation, Universite Catholique de Louvain, Belgium, Louvain, Belgium.
34. Zhou, J.-Q., Monson, E., Teng, S.-C., Schulz, V., and Zakian, V. (2000) Pif1p helicase, a catalytic inhibitor of telomerase in yeast. *Science* **289**, 771-774
35. Budd, M., Reis, C., Smith, S., Myung, K., and Campbell, J. (2006) Evidence suggesting that Pif1 helicase functions in DNA replication with the Dna2 helicase/nuclease and DNA polymerase δ . *Mol. Cell. Biol.* **26**, 2490-2500
36. Woodford, K., Howell, R., and Usdin, K. (1994) A novel K(+)-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes. *J. Biol. Chem.* **269**, 27029-27035
37. Gu, Y., Masuda, Y., and Kamiya, K. (2008) Biochemical analysis of human Pif helicase and functions of its N-terminal domain. *Nucleic Acids Res.* **36**, 6295-6308

38. Sanders, C. and Stenlund, A. (2001) Mechanism and requirements for bovine papillomavirus, type 1, E1 initiator complex assembly promoted by the E2 transcription factor bound to distal sites. *J. Biol. Chem.* **276**, 23689-23699
39. Boule, J. and Zakian, V. (2007) The yeast Pif1p DNA helicase preferentially unwinds RNA-DNA substrates. *Nucleic Acids Res.* **35**, 5809-18
40. Vega, L., Phillips, J., Thornton, B., Benanti, J., Onigbanjo, M., Toczyski, D., and Zakian, V. (2007) Sensitivity of yeast strains with long G-tails to levels of telomere-bound telomerase. *PLoS Genet.* **3**(6), e105
41. Tanaka, H., Ryu, G.-H., Seo, Y.-S., Tanaka, K., Okayama, H., MacNeill, S., and Yuasa, Y. (2002) The fission yeast *pfh1+* gene encodes an essential 5' to 3' DNA helicase required for the completion of S-phase. *Nucleic Acids Res.* **30**, 4728-4739
42. Snow, E., Mateyak, M., Paderova, J., Wakeham, A., Iorio, C., Zakian, V., Squire, J., and Harrington, L. (2007) Murine Pif1 interacts with telomerase and is dispensable for telomere function *in vivo*. *Mol. Cell. Biol.* **27**, 1017-1026
43. Wu, L. and Hickson, I. (2006) DNA helicases required for homologous recombination and repair of damaged replication forks. *Annu. Rev. Genet.* **40**, 279-306
44. Creacy, S., Routh, E., Iwamoto, F., Nagamine, Y., Akman, S., and Vaughn, J. (2008) G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *J. Biol. Chem.* **283**, 34626-34634

Figure Legends

Figure 1. Formation of G-quadruplex (G4) structures. (A) T55-OX G4 and T55-V4 G4 DNA structures. The ssDNA sequences of each molecule are indicated; one G residue in each planar quartet is also labelled. (B) PAGE analysis and purification of T55-tailed G4 products. The tetramolecular parallel G4 molecule with the greatest stability was the most abundant product of the annealing reactions in each case. The bimolecular anti-parallel quadruplex structure (G'2) of T55-OX is also indicated. For the V4 sequence, both high molecular weight products have been designated as G4 DNA (G4 and G4*) as the conditions employed would not favour formation of bimolecular structures. The ladder of very low mobility products seen above the T55-OX G4 products are likely to be G4 "superstructures" that have been observed previously [31]. The desired tetramolecular products were excised from the gel and recovered by soak-elution. In the case of V4, the lowest mobility species was analysed further. All samples shown were run on the same gel and the same adjustments made to the whole image from which the image group in part B was constructed. (C) Confirmation of G-quartet formation. Purified DNA was subject to DMS protection analysis (N7 guanine). The brackets denote the G-runs that are accessible in the single-stranded oligonucleotide but protected in the G4 DNA.

Figure 2. Unwinding of G4 DNA by hPif1 and hPifHD. (A) Lanes 1-10, unwinding of ³²P-labelled T55-V4 G4 DNA by hPif1 and hPifHD (0.1 nM substrate, 2, 4 and 6 nM hPif1 and 2.5, 10 and 40 nM hPifHD). Lanes 8-10 include a boil control to dissociate the G4 substrate and markers for V4-G4* and V4 ssDNA. Lanes 11-20, as 1-10 but analysis of substrate T55-OX G4. Lanes 21-28, unwinding of the reference partially single-stranded duplex substrate PST55 (20 base pair duplex, 5' T 55 extension ³²P end-labeled) by the two enzyme species. (B) Dependence of T55-OX G4 unwinding on substrate concentration. Substrate concentrations were fixed at 0.05, 0.1 and 0.2 nM and hPif1 titrated (2, 4, 6, 8 and 10 nM) under standard assay conditions. Apparent unwinding was suppressed at higher substrate concentrations.

Figure 3. 5' ssDNA tail length and G4 DNA unwinding. G4 DNA substrates based on the OX G-quartet motif were generated with 55, 25 or 10 base poly T 5' extensions, T55-OX G4, T25-OX G4 and T10-OX G4 respectively, and analyzed in the hPif1 unwinding assay (0.1 nM probe, 2, 4, 6 nM hPif1). No appreciable unwinding of T10-OX G4 (lanes 11-15) was observed and the unwinding efficiency of T25-OX G4 (lanes 6-10) was lower than T55-OX G4 (lanes 1-5). Graphed data is shown below the autoradiogram.

Figure 4. Competition assays substantiate the need for ssDNA in the initiation of DNA unwinding. (A) Competition assay with ³²P-labeled substrate T55-OX G4. Helicase assays (0.1 nM substrate, 6 nM hPif1) were assembled with increasing concentrations of purified unlabelled competitor DNAs (1 to 25 fold molar excess), and unwinding determined. Helicase action was highly sensitive to competitors with T55 ssDNA composition (T55, PST55 and T55-OX G4) compared to the G4 DNA T10-OX G4, as summarized in the graph. (B) Reactions performed in parallel with substrate PST55. Reactions showed similar patterns of sensitivity to the various competitors, except there was a modest stimulation of unwinding (~0.5 fold) of substrate PST55 by T10-OX G4 DNA over the concentration range tested.

Figure 5. hPif1 binds G4 DNA. Binding reactions without nucleotide cofactor (0.1 nM substrate, 0.1, 0.25 and 1 nM hPif1) were assembled and reaction products resolved on native

poly-acrylamide gels (gel-shift assay, 0.25 x TBE). hPif1 did not bind the ssDNA precursor of T10-OX G4, T10-OX ss (lanes 1-4), or the same oligonucleotide annealed to its complementary strand (5' C₄A₄C₄A₁₀, substrate T10-OX ds, lanes 5-8). T10-OX G4 DNA (lanes 9-12) was bound with apparent high affinity, greater than the helicase substrate PST55 (lanes 13-16 and graphed data below) and generated a ladder-like DNA binding pattern. T55-OX G4 was also bound with high affinity, shifting all the probe at the highest protein concentration, as with substrate T10-OX G4 (lanes 17-20). All reactions were performed in parallel and all samples shown were run on the same gel and the same adjustments made to the whole image from which the image group was constructed.

Figure 6. G4 DNA binding activity is intrinsic to hPifHD. (A) DNA binding of hPifHD and hPif1 to G4 DNA (T10-OX G4) and ssDNA (T55) was compared by gel-shift assay (0.1 nM DNA, 0.1, 0.25 and 1 nM protein). A ladder-like pattern of DNA binding was observed for both hPif1 and hPifHD and binding activities were similar (lanes 2-4 compared to 5-7 and graphed data below). The ladder-like pattern of DNA binding was not observed with ssDNA (lanes 8-14). (B) Graphed data for hPif1 DNA binding. (C) Graphed data for hPifHD DNA binding.

Figure 7. DNA binding with and without ATP and stimulation of hPif1 ATPase activity by DNA binding substrates. DNA binding reactions (0.1 nM ³²P-labeled probe, 0.1, 0.25 and 1 nM hPif1) were performed with or without 5 mM ATP/MgCl₂ in the binding buffer. Complexes were resolved on 0.25 x Tris borate buffered gels with or without 0.1 mM ATP/MgCl₂ in the gel and running buffer. However, comparable results were also obtained when reactions were resolved on standard 0.25 x TBE gels (data not shown). (A) As with hPifHD [28], DNA binding of hPif1 to substrate PST55 was reduced in the presence of ATP, as summarized in the graphical data on the right. Unwinding is almost completely inhibited at 75 mM NaCl used in these assays. Furthermore, the 20 base oligonucleotide of the substrate was radio-labeled as it is too small to be bound by the enzyme. Comp. denotes the DNA-protein complexes formed. (B) Binding to substrate T10-OX G4 was analysed in parallel to PST55. Only minimal differences in G4 DNA binding activity were observed with and without ATP (C) ATPase reactions (50 nM hPif1, 1 pmol per reaction) were assembled with or without DNAs at 25, 50 or 200nM. Phosphate release (total for 20 µl reaction indicated, 10 minute incubation) was determined using the charcoal-binding assay. T10-OX G4 DNA (and its ssDNA precursor) was ineffective at stimulating ATPase activity compared to the substrates with 55 base poly T sequences.

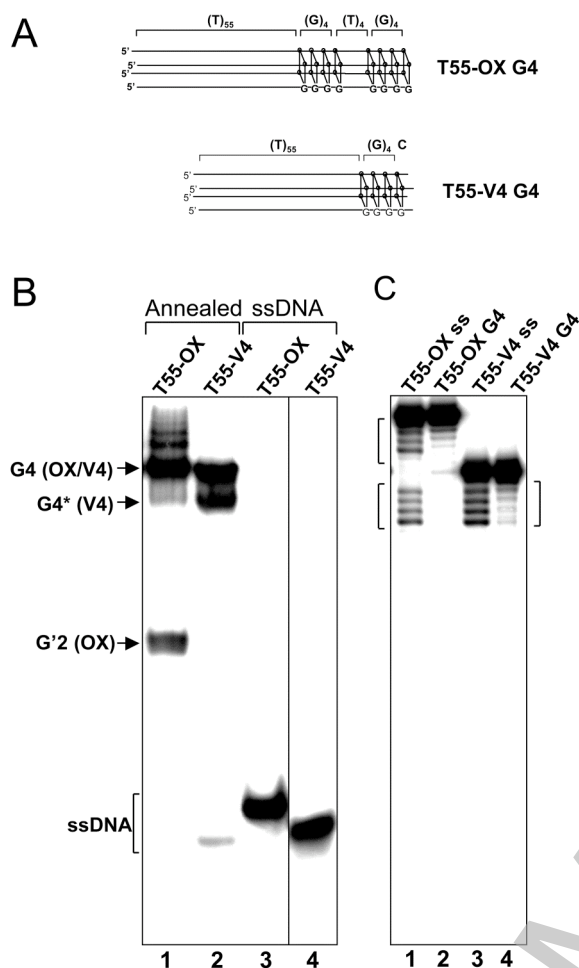
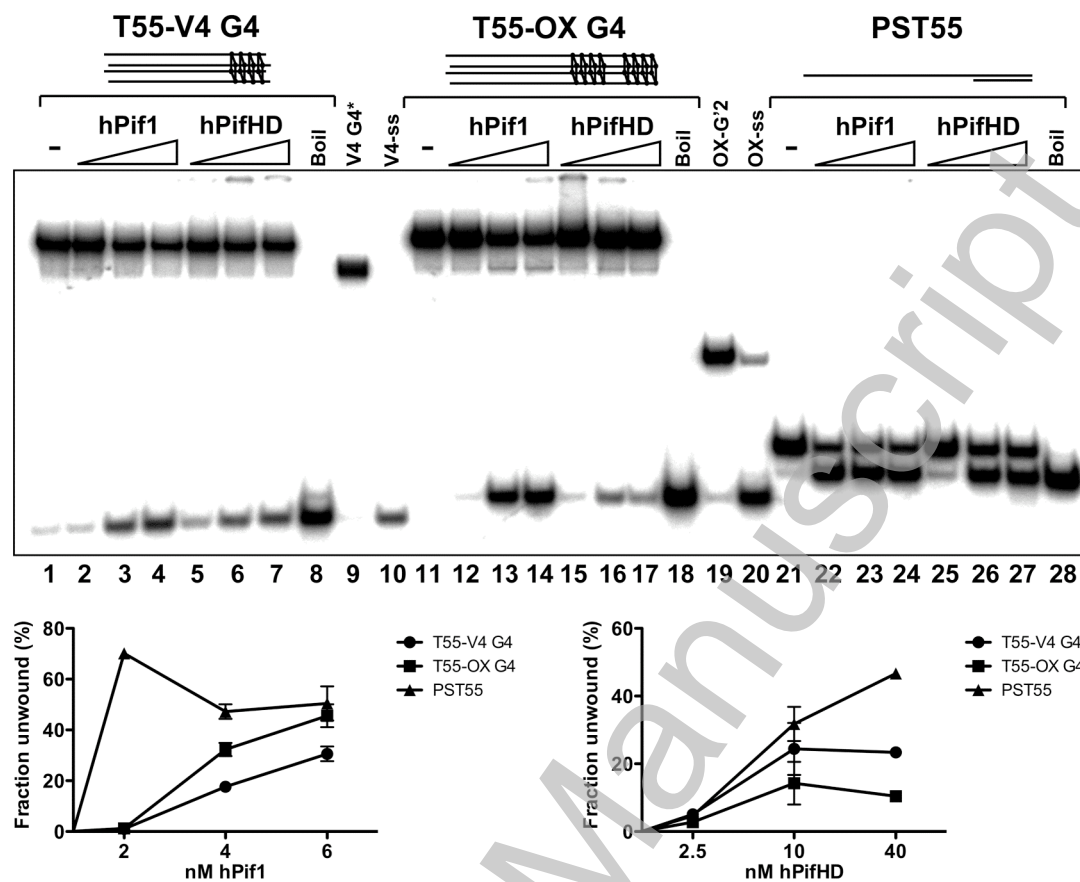


Figure 1.

A



B

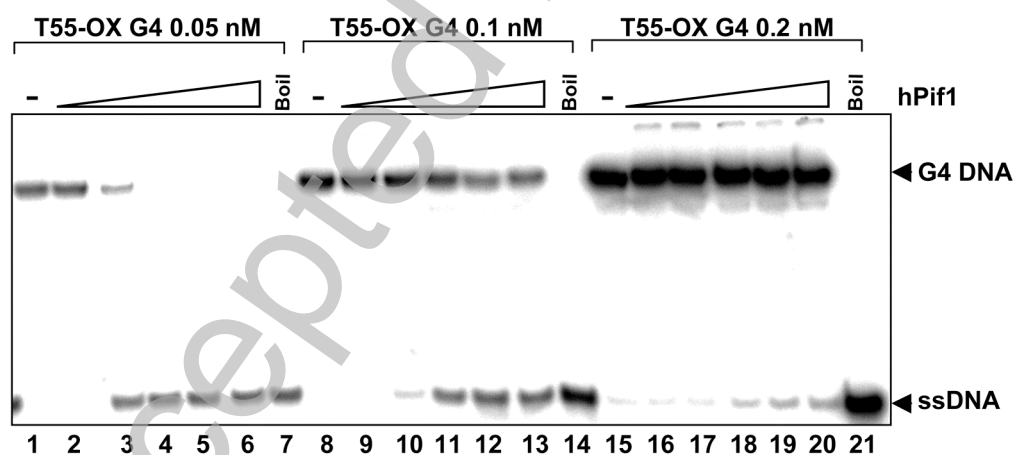


Figure 2.

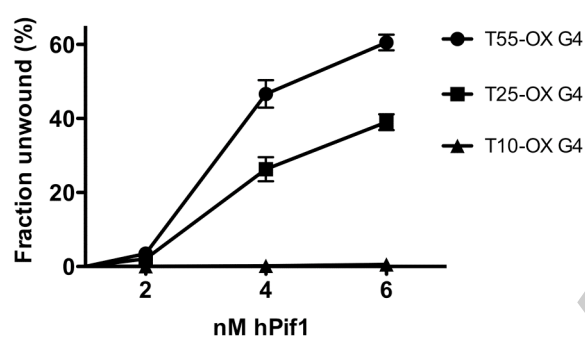
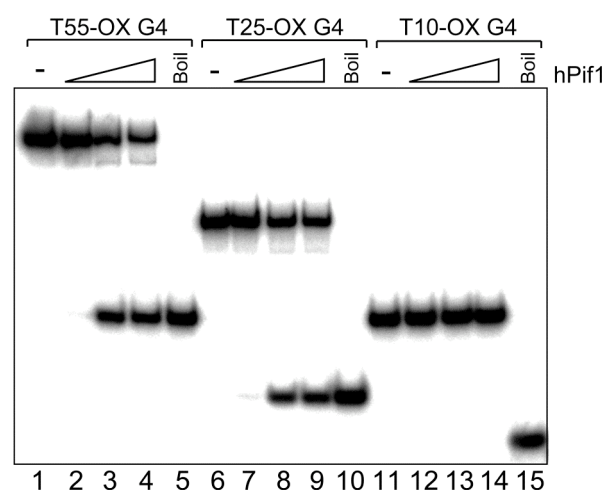


Figure 3.

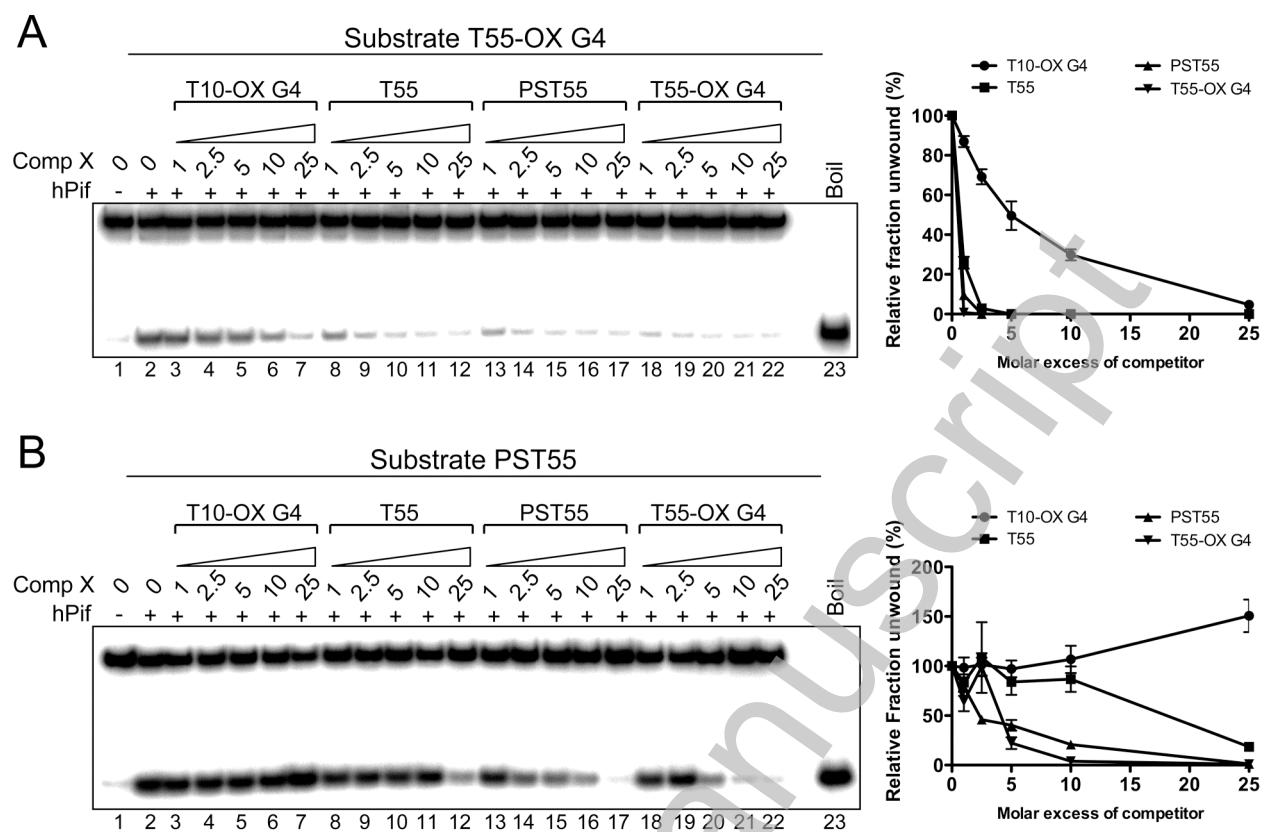


Figure 4.

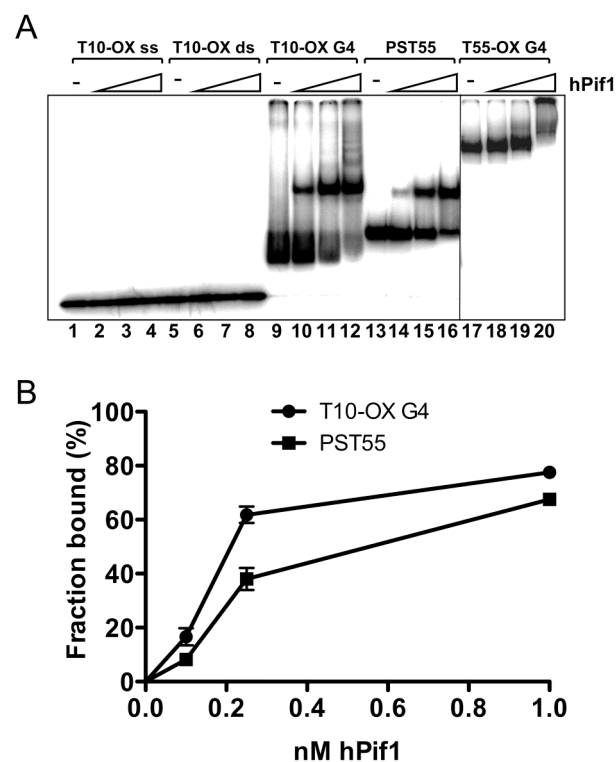


Figure 5.

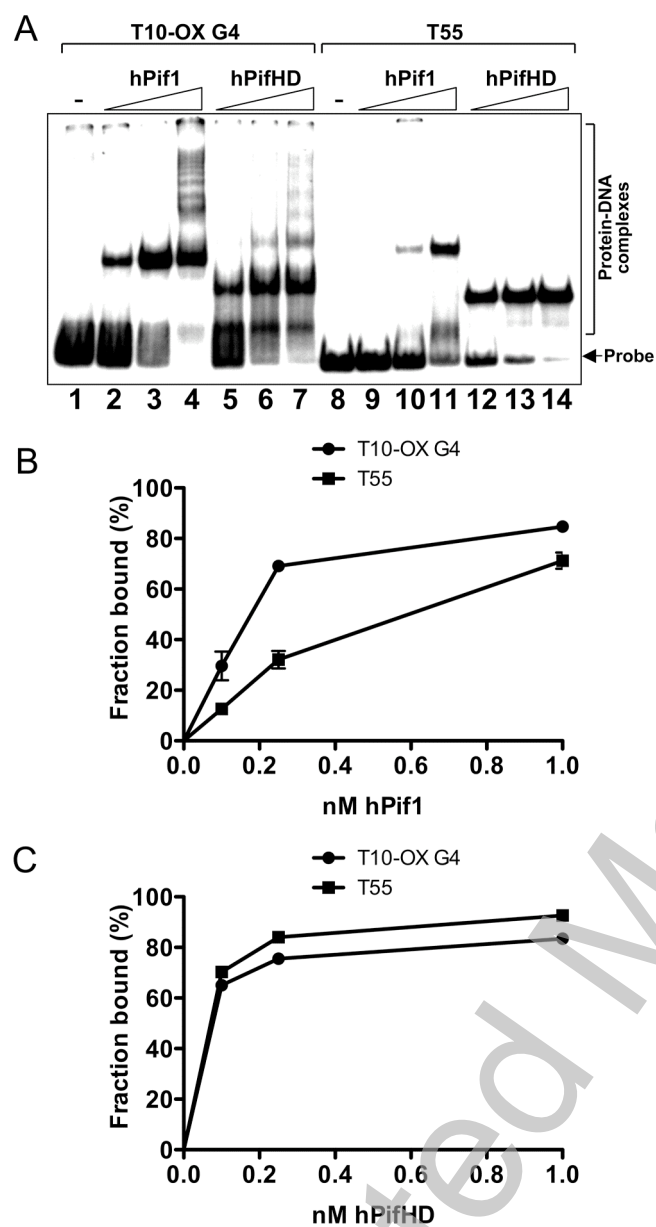


Figure 6.

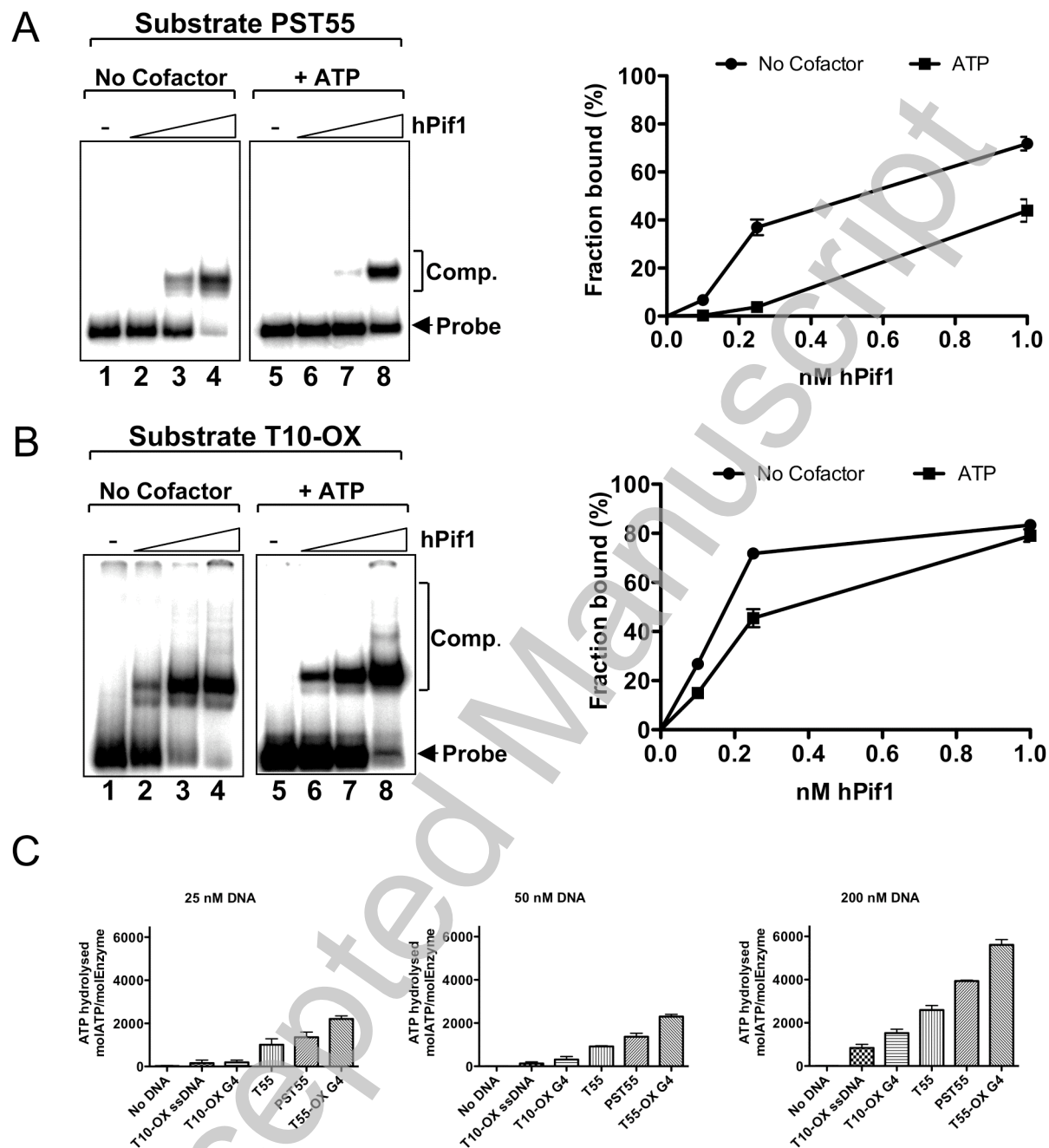


Figure 7.