The effect of Pot1 binding on the repair of thymine analogs in a telomeric DNA sequence

Jacob A. Theruvathu, Agus Darwanto, Chia Wei Hsu and Lawrence C. Sowers*

Department of Pharmacology and Toxicology, The University of Texas Medical Branch, 3.330 Basic Science Building, 301 University Blvd, Galveston, TX 77555-0617, USA

Received March 21, 2014; Revised June 05, 2014; Accepted June 23, 2014

ABSTRACT

Telomeric DNA can form duplex regions or single-stranded loops that bind multiple proteins, preventing it from being processed as a DNA repair intermediate. The bases within these regions are susceptible to damage; however, mechanisms for the repair of telomere damage are as yet poorly understood. We have examined the effect of three thymine (T) analogs including uracil (U), 5-fluorouracil (5FU) and 5-hydroxymethyluracil (5hmU) on DNA–protein interactions and DNA repair within the GGTTAC telomeric sequence. The replacement of T with U or 5FU interferes with Pot1 (Pot1pN protein of Schizosaccharomyces pombe) binding. Surprisingly, 5hmU substitution only modestly diminishes Pot1 binding suggesting that hydrophobicity of the T-methyl group likely plays a minor role in protein binding. In the GGTTAC sequence, all three analogs can be cleaved by DNA glycosylases; however, glycosylase activity is blocked if Pot1 binds. An abasic site at the G or T positions is cleaved by the endonuclease APE1 when in a duplex but not when single-stranded. Abasic site formation thermally destabilizes the duplex that could push a damaged DNA segment into a single-stranded loop. The inability to enzymatically cleave abasic sites in single-stranded telomere regions would block completion of the base excision repair cycle potentially causing telomere attrition.

INTRODUCTION

The ends of all linear eukaryotic chromosomes have repeating sequences, telomeres, which can simultaneously form duplex, single-stranded, and unusual deoxyribonucleic acid (DNA) structures. An array of proteins binds to these structures, providing protection and facilitating telomere maintenance (1–4). Accelerated telomere erosion can occur from telomere damage or deficiencies in telomere-binding pro-}

Teines, leading to premature cellular aging and senescence (5–7). The important role for telomere–protein complex formation is underscored by the high frequency of telomere end-fusions and inappropriate recombination that occur in the absence of telomere protection proteins (7,8). Currently, little is known about the impact of telomere damage on telomere–protein interactions or how such damage might be repaired.

Telomere repeat sequences in most organisms are five to nine bases in length, with the G- and T-rich strand forming a single-stranded 3’ overhang (9). In the fission yeast Schizosaccharomyces pombe, the telomere is 350–500 bases in length, terminating with a 3’-G and T-rich overhang of 30 to 50 bases (10,11). The single-stranded G- and T-rich strand can be bound by multiple Pot1 (protection of telomere 1) proteins. An N-terminal domain of S. pombe Pot1 comprising the first 185 amino acids, referred to as Pot1pN, has been overexpressed and examined in several structural and biophysical studies (12,13). Pot1pN and its parent Pot1 bind with high affinity to the repeat sequence GGT-
TAC. Pot1pN has one oligonucleotide/oligosaccharide-binding fold that docks with the single-stranded repeat sequence (12–14). The crystal structure of Pot1pN bound to a GGTTA oligonucleotide reveals that within the complex, the G and T bases form two unusual base pairs in which the T-methyl group is positioned in a hydrophobic cleft between the G and T bases. The folded and base-paired oligonucleotide sequence interacts with the protein through both nonspecific base-stacking and sequence-specific hydrogen-bonding interactions. Discrimination between DNA and ribonucleic acid binding is afforded by the absence of 2’-hydroxyl groups in the sugar moieties and the presence of the thymine methyl groups (15). The replacement of thymine by uracil could generate an energetically disfavored cavity, potentially explaining the substantial decrease in Pot1pN binding to oligodeoxynucleotides containing uracil in place of thymine (12,16). The critical role of the thymine methyl group in the Pot1-telomere complex prompted us to consider the potential consequences of additional thymine substitutions.

*To whom correspondence should be addressed. Tel: +1 409-772-9678; Fax: +1 409-772-9648; Email: lasowers@utmb.edu

Present Address:
Agus Darwanto, MolecularMD, 840 Memorial Drive, Cambridge, MA 02139, USA.

© The Author(s) 2014. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
The DNA of all organisms undergoes continuous cycles of damage and repair (17,18). Chemical damage to DNA bases can result in miscoding and mutation during DNA replication as well as interfere with sequence-specific DNA–protein interactions in both telomeric and non-telomeric sequences. Among the many types of DNA damage, modifications of the thymine methyl group can occur frequently. In replicating cells, thymidine residues can be replaced by 2′-deoxyuridine, especially during folate deficiency (19–22). The chemotherapy agent 5-fluorouracil (5FU) can also be incorporated into replicating DNA and facilitate uracil incorporation by interfering with thymidine monophosphate (dTMP) metabolism (23–25). The methyl group of thymine can be oxidized, forming 5-hydroxymethyluracil (5hmU) either directly in DNA or in the nucleotide pool (26,27). It is known that the modified pyrimidines U, 5FU and 5hmU interfere with sequence-specific DNA–protein interactions when in duplex DNA (28–30). However, the impact of 5FU and 5hmU on telomere-binding proteins has not been previously investigated. While base damage is expected to occur across the chromosomes, telomeric DNA might be particularly susceptible to endogenous DNA damage (31). It is known that guanine residues in telomeres are more likely to become oxidized (32), and further, that endogenous reactions including oxidation can diminish the affinity of telomere binding proteins that localize to duplex regions (33). In contrast, Ghost et al. showed that the presence of 8-oxoG can enhance Pot1 binding in telomere D-loop regions (34).

Most endogenous DNA damage is repaired by the base excision repair (BER) pathway in which damaged bases are recognized by specific glycosylases generating an abasic site. The resulting abasic site is cleaved by an apurinic endonuclease, followed by polβ-mediated repair synthesis and ligation (35–37). While the BER pathway has been studied extensively within duplex, B-form DNA, substantially less is known about how the components of this pathway function in single-stranded regions or in unusual DNA structures. The modified pyrimidines including U, 5FU and 5hmU can be removed from DNA by members of the uracil DNA glycosylase family (19–25,38–41). Glycosylase removal of a base generates an abasic site (AP site, apurinic/apyrimidinic site), as does spontaneous hydrolytic loss of normal and damaged bases (42,43). In multiple organisms, uracil misincorporation during DNA synthesis is one of the most common types of endogenous damage (44–46). Uracil misincorporation into telomeres in human cells correlates with telomere shortening and dysfunction (47). The intracellular nucleotide pool can be protected from fraudulent nucleotides by a series of dNTPases including dUTPase (46) and 8-oxodGTPase (MTH1) (48) that cleave modified 2′-deoxynucleoside monophosphates (dNMPs). Depletion of dUTPase can result in embryonic lethality, cell cycle arrest and apoptosis (45,46) demonstrating the importance of excluding uracil from DNA.

In this study, we synthesized a series of model DNA substrates containing an array of damage products within the GGTTAC motif including U, 5FU, 5hmU and a tetrahydrofuran (THF) abasic site. We investigated the consequence of each of these substitutions on the binding of the N-terminal domain of S. pombe Pot1 (Pot1pN), the potential for competition between Pot1 and DNA glycosylases, the possibility that Pot1 binding could interfere with glycosylase removal of the modified pyrimidine and the capacity of an abasic endonuclease to cleave an abasic site within a telomeric repeat. Collectively, our results indicate that DNA damage in telomeric regions can pose a significant biological challenge and the implications of these findings for telomere maintenance are discussed. Our results also suggest potential mechanisms for chemotherapy-induced toxicity which have not been previously considered.

**MATERIALS AND METHODS**

**Oligonucleotide synthesis and characterization**

Oligonucleotides were synthesized by standard phosphoramidite methods as previously described (49,50). The sequences of the oligonucleotides used in this study are shown in Figure 1. The telomere sequence is comprised of three GGTTAAC repeats in which the first two repeats have C in place of N. The third repeat is the canonical binding sequence GGTTAC flanked by As on the 3′ and 5′ side. This sequence was designed to have a single, high affinity Pot1 binding site at the 3′-end and was of sufficient length to conduct both Pot1 binding and DNA repair studies. In the study described here, the 2′-deoxyribonucleoside analogs of U, 5FU or 5hmU replaced T at positions T3 or T4 as described in the text. Oligonucleotides were also prepared in which the THF model abasic site replaced G at G1 or G2 and T at T3 or T4. An oligonucleotide sequence containing only normal bases and complementary to the primary sequence examined here was also prepared and was used to create duplex structures. All oligonucleotides prepared for this study were analyzed by mass spectrometric methods (51). Melting temperatures of oligonucleotides duplexes (4 μM) were determined from measurement of ultraviolet optical absorbance at 260 nm as a function of temperature in a buffer containing 10-mM NaCl, 150-mM KCl, 20-mM tris and 1-mM ethylenediaminetetraacetic acid (EDTA) (52).

**Preparation of Pot1pN**

Pot1pN, comprising amino acids 1–185 based on the cDNA sequence NM_001019882, was cloned and isolated as previously described (13). The purified protein was characterized by mass spectrometry following trypsin digestion (Supplementary Figure S1).

**Pot1 binding reaction and electrophoresis mobility shift assay**

Oligonucleotides containing canonical bases or base analogs (Figure 1) were 5′,32P-end labeled with T4 polynucleotide kinase and γ−32PATP and column purified using standard procedures. Labeled oligonucleotides (25 nM) were then incubated in a buffer containing 25-mM HEPES/NaOH, pH 7.5, 50-mM NaCl, 40-mM KCl, 7% glycerol, 1-mM EDTA and 0.1-mM dithiothreitol (DTT) at room temperature for 30 min with Pot1pN at various concentrations up to 640 nM (total volume 30 μl) as described further. Following the addition of loading buffer (7.5 μl, 60% 0.25x TBE and 40% glycerol) binding reactions were loaded onto 10% non-denaturing polyacrylamide (37.5:1,
acrylamide/bis-acrylamide) gels that had been equilibrated for 90 min at 200 V and 4°C. Samples were electrophoresed at 250 V for 2 h at 4°C. Visualization and quantification of the gels was carried out using a Storm 860 imager and ImageQuant 5.0 Software (GE Healthcare).

**Binding model and data analysis**

Pot1pN is a monomer protein that binds to a single hexanucleotide binding motif (13). In the electrophoretic mobility shift assay (EMSA) studies reported here, oligonucleotides containing only one functional hexanucleotide binding motif were used to mimic a non-cooperative binding model. The Pot1pN binding analysis was carried out using the non-cooperative, single binding model described earlier (53). In these studies, the percentage of oligonucleotide bound to Pot1pN is measured as a function of the total Pot1pN concentration. The dissociation constant, $K_d$, corresponds to the concentration of Pot1pN when 50% of the oligonucleotide is bound. In the studies reported here, the oligonucleotide sequence has one high affinity GGTTAC binding site and two low affinity GGTAC binding sites. As the T to C substitution increases the $K_d$ by a factor of more than $10^5$ (12) we are assuming only one binding site and therefore reporting our results as $K_d$ rather than $K_d$ apparent. The $K_d$ values reported here are the averages of four independent measurements ± SD.

**Binding competition between Pot1pN and glycosylases**

Uracil-DNA glycosylase (UNG) was obtained from New England Biolabs, Beverly, MA, USA. The human hSMUG1 was cloned in our laboratory as previously described (41). DNA substrates (50 nM) were incubated with 200-nM glycosylase and 800-nM Pot1pN in the binding buffer described above for a total volume of 10 μl at room temperature for 15 min. Reactions were stopped by adding 5 μl of 0.1-M NaOH and heated at 95°C for 30 min to cleave the phosphodiester DNA backbone. The reaction conditions used here were designed to generate cleavage of 50% of the oligonucleotides upon treatment of a uracil-containing oligonucleotide with UNG. Following addition of loading buffer (98% formamide, 0.01-M EDTA, 1-mg/ml xylene cyanol and 1-mg/ml bromophenol blue) containing 50 pmol of an unlabeled competitor oligonucleotide, oligonucleotides were separated by electrophoresis on 20% denaturing polyacrylamide gels (8-M urea) and visualized as described above.

**AP endonuclease assay**

Duplexes for AP endonuclease (APE1) cleavage assays were prepared by adding a 2-fold excess of an unlabeled complementary strand to the labeled substrate of interest followed by heating (95°C for 5 min) and slow cooling to room temperature. Duplex or single-stranded substrates (50 nM) were incubated with human APE1 (200 nM, Trevigen, Gaithersburg, MD, USA) in buffer (10-mM HEPES-KOH (pH 7.4), 100-mM KCl, 10-mM MgCl$_2$) in a total volume of 10 μl at
37°C for 60 min. The reactions were stopped by adding an equal volume of loading buffer, and oligonucleotides were electrophoresed on 20% denaturing polyacrylamide gels (8M urea). Gels were visualized as described above.

RESULTS

Pot1 and Pot1pN proteins are known to bind to the GGT-TAC site with high affinity. In the studies reported here, a 20-mer oligonucleotide was constructed containing one Pot1pN GGT-TAC binding site on the 3’-end. Modifications were made within the G1G2T3T4AC sequence by substituting thymine analogs for thymine at positions T3 and T4 and a chemically stable (THF) abasic site at positions G1, G2, T3 and T4.

Pyrimidine analogs can interfere with Pot1pN binding

Previously, it has been shown that the minimal-length for the telomeric single-stranded DNA required to bind Pot1pN is 6 nucleotides (GGTTAC) by gel filtration chromatography, EMSA and filter binding assays (13). In this study, the dissociation constant (Kd) for the interaction of Pot1pN with each of the single-stranded oligonucleotides was measured by EMSA. The binding assays were performed at Pot1pN protein concentrations ranging from 0 to 640 nM (Figures 2 and 5). The Kd was determined for each oligonucleotide–Pot1pN complex using non-linear regression to fit the equation for a simple, non-cooperative monomeric binding mode (Table 1). Previously, the Kd for the binding of Pot1pN to the telomere oligonucleotide was reported to be 83 nM (13) for the GGT-TAC six-mer. The Kd for the 20-mer oligonucleotide containing normal DNA bases obtained in our study was 66 ± 5 nM, which is comparable to data reported earlier.

When T3 was replaced by 5hmU, the observed dissociation constant (Kd) Table 1) was essentially unchanged. In contrast, the replacement of T3 by U and 5FU diminished binding by a factor of about 2.4 as seen by the corresponding increase in Kd values. Substitution at the T4 position showed more significant changes in the Kd values (Table 1). Replacing T4 with 5hmU slightly decreased binding (1.9-fold), whereas replacement with 5FU decreased the binding significantly (13-fold). The substitution at T4 with U had the greatest effect, increasing the Kd by 52-fold.

The size of the uracil 5-substituent appears to be more important than hydrophobicity in modulating the binding affinity of Pot1pN

The decreased Pot1pN binding affinity observed with uracil and 5FU substitution was not unexpected. In contrast, the replacement of the hydrophobic T-methyl group with the more hydrophilic 5-hydroxymethyl group of 5hmU was anticipated to interfere with G-T base pair formation within the Pot1pN–telomere complex and disrupt Pot1pN binding. We therefore considered that substituent size might be a critical factor. While the size of a 5-substituent can be easily calculated for spherically symmetric substituents including a hydrogen atom and a methyl group, a different approach must be used for non-symmetric substituents including a 5-hydroxymethyl group. Previously, we investigated the relative sizes of a series of uracil-5-substituents in a glycosylase repair specificity study and evaluated substituent size by calculating the solvent accessible surface area (SASA) for the entire pyrimidine. The relative pyrimidine size is calculated as the SASA for a given pyrimidine divided by the SASA for thymine (41). Within the series examined here, the differences in pyrimidine size can be attributed to differences in 5-substituent size. Relative pyrimidine size for the analogs examined here is plotted versus the Pot1pN dissociation constant (Kd). (ln Kd) is observed with thymine. Binding affinity decreases (Kd increases) as the size of the modified pyrimidine increases (5hmU) or decreases (U, 5FU) relative to thymine. The effect is greater with substitution at T4. The relative affinity of Pot1pN for pyrimidines in this series is T>5hmU>5FU>U.

Pot1pN binding can interfere with glycosylase cleavage of pyrimidine analogs

UNG cleaves U and 5FU from single-stranded oligonucleotides, whereas hSMUG1 cleaves pyrimidine oxidation damage products including 5hmU. The capacity of these glycosylases to cleave in the telomeric sequence, in the absence of Pot1pN, is shown in Figure 4. In order to examine the ability of a glycosylase to interact and repair modified bases in the presence of Pot1pN binding, substituted oligonucleotides were incubated with a high concentration of Pot1pN and probed with UNG and hSMUG1, as shown in Figure 4. UNG efficiently cleaved U from the oligonucleotide even in the presence of Pot1pN. UNG could also cleave 5FU at the T4 position but had more difficulty at T3 when Pot1pN was present. In contrast, hSMUG1 was unable to cleave 5hmU at either position in the presence of Pot1pN.

The formation of an abasic site inhibits Pot1pN binding

DNA bases can be lost from the DNA molecule by spontaneous hydrolytic depurination or depyrimidination or by glycosylase cleavage. To investigate the effect of a nuclease loss, we utilized oligonucleotides containing stable abasic sites at the T positions (T3 and T4) as well as at the adjacent G positions (G1 and G2). The abasic site at the G1 position decreased the binding affinity by 2.7-fold. However, the abasic site at other positions had a more dramatic effect on the binding affinity (Figure 5 and Table 1). The observed decreased binding affinity corresponds to an increase in the Kd values of 75-fold, 200-fold and >10 000-fold for oligonucleotides containing an abasic site at positions G2, T1 and T4, respectively.

APE1 can cleave an abasic site within the Pot1pN binding motif when placed within a duplex region, but not when single-stranded

We examined the ability of APE1 to cleave abasic sites within the telomere sequence (Figure 6). Cleavage was observed for the four abasic site-containing oligonucleotides examined here when present in a duplex. However, APE1
Figure 2. Pyrimidine substitution at position T₄ decreases Pot1pN binding affinity. (a) EMSA of single-stranded oligonucleotides containing thymine or a thymine analog incubated with increasing concentrations of Pot1pN. (b) Non-linear regression analysis of average percent binding versus Pot1pN concentration. The error bars represent standard deviation calculated from four independent measurements.

Table 1. Pot1pN binding affinity is significantly dependent upon the structure and position of the modification within the binding motif

<table>
<thead>
<tr>
<th>G₁</th>
<th>G₂</th>
<th>T₃</th>
<th>T₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>-</td>
<td>161 ± 10.8</td>
<td>3410 ± 209</td>
</tr>
<tr>
<td>5hmU</td>
<td>-</td>
<td>67.0 ± 6.0</td>
<td>129 ± 8.80</td>
</tr>
<tr>
<td>5FU</td>
<td>-</td>
<td>144 ± 11.6</td>
<td>844 ± 89.4</td>
</tr>
<tr>
<td>Abasic site</td>
<td>176 ± 20.5</td>
<td>4990 ± 278</td>
<td>13 400 ± 995</td>
</tr>
</tbody>
</table>

Dissociation constant (Kᵯ) values (nM) of Pot1pN binding to single-stranded oligonucleotides containing thymine analogs or abasic sites at positions G₁, G₂, T₃ or T₄ are shown. A control containing no modifications has a Kᵯ of 66 ± 5 nM. The values reported here are averages of four measurements ± SD.

Abasic site formation within the Pot1pN binding site destabilizes duplex formation

The effect of each substitution on the melting temperature of the corresponding duplex formed between the model substrate oligonucleotide and its complement containing all normal DNA bases was measured (Table 2). The replacement of T by U, 5FU or 5hmU had no significant effect on melting temperatures. In contrast, the replacement of G or T with an abasic site reduced duplex melting temperatures by 5–10°C (Table 2).

DISCUSSION

The repetitive DNA sequence motifs characteristic of telomeres can be found in multiple configurations including duplex, single-stranded and loop regions (1–4). Multiple proteins bind to each of the DNA configurations, protecting them from being recognized as DNA repair intermediates or from participating in illegitimate chromosomal end fusions (5–7). However, DNA bases are persistently damaged in all organisms and the magnitude of the damage could be greater in telomeric regions. It is as yet unknown how telomere protection proteins might interfere with the legitimate repair of telomere sequences, and, therefore, incomplete DNA repair within telomeres might represent a previously unrecognized mechanism of telomere attrition.

In this manuscript, we have focused on repair within the GGTTAC sequence of the S. pombe telomere. Multi-
Figure 4. Glycosylase repair diminishes as Pot1 binding affinity increases. Competition between Pot1pN binding and glycosylase activity on single-stranded oligonucleotides with pyrimidine substitutions at positions T3 and T4 is shown. (a) Uracil substitution and UNG cleavage. (b) 5FU substitution and UNG cleavage. (c) 5hmU substitution and hSMUG1 cleavage. Lane 1: negative control without the addition of Pot1pN or a glycosylase; Lane 2: positive control with the addition of a glycosylase; Lane 3: simultaneous incubation with both Pot1pN and a glycosylase.

Figure 5. An abasic site at positions G1, G2, T3 or T4 results in decreased Pot1pN binding affinity. (a) EMSA of single-stranded oligonucleotides containing thymine or an abasic site incubated with increasing concentrations of Pot1pN. (b) Non-linear regression analysis of average percent binding versus Pot1pN concentration. The error bars represent standard deviation calculated from four independent measurements.
Table 2. Pyrimidine substitution does not observably change duplex melting temperatures, but abasic site substitution substantially decreases thermal stability.

<table>
<thead>
<tr>
<th></th>
<th>G1</th>
<th>G2</th>
<th>T3</th>
<th>T4</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>-</td>
<td>-</td>
<td>70.3 ± 0.5</td>
<td>70.1 ± 0.1</td>
</tr>
<tr>
<td>5hmU</td>
<td>-</td>
<td>-</td>
<td>68.8 ± 0.4</td>
<td>69.6 ± 0.6</td>
</tr>
<tr>
<td>5FU</td>
<td>-</td>
<td>-</td>
<td>70.1 ± 0.1</td>
<td>70.0 ± 0.1</td>
</tr>
<tr>
<td>Abasic site</td>
<td>58.7 ± 0.5</td>
<td>60.8 ± 0.5</td>
<td>63.5 ± 0.6</td>
<td>65.8 ± 0.5</td>
</tr>
</tbody>
</table>

Melting temperatures (Tm) values (°C) of the duplex (4 μM) containing thymine analogs or abasic sites at positions G1, G2, T3 or T4 are shown. A control containing no modifications has a Tm of 69.9 ± 0.5°C. The values reported here are averages of four measurements ± SD.

Figure 6. APE1 is able to cleave an abasic site in duplex DNA, but no cleavage is observed when the abasic site is located in a single-stranded region. Cleavage of an oligonucleotide containing an abasic site at positions G1, G2, T3 or T4 by APE1 is shown. (a) Abasic site in a single-stranded telomeric sequence. (b) Abasic site in a duplex telomeric sequence.

ple GGTTAC sequences can be found in a 3′-overhanging single-stranded configuration bound to Pot1 (protection of telomere protein 1) proteins (2,5). Multiple studies conducted on this sequence with the N-terminal region of Pot1, Pot1pN, have established the sequence specificity within this repeat and have revealed that the sequence folds into a compact structure with unusual G-T base pairs (10,12–13). The 2′-hydroxyl group in ribonucleotides disrupts this structure, as does the replacement of the thymine methyl group with the 5-hydrogen of uracil, and both substitutions disrupt Pot1 binding (15).

The capacity of uracil substitution to disrupt Pot1 binding led us to consider the likelihood that other pyrimidine substitutions in which the 5-position is modified might also interfere with Pot1–DNA interactions. We therefore constructed a series of oligonucleotides containing U, 5FU and 5hmU (Figure 1). Levels of U substitution into DNA can increase substantially under folate stress, 5FU chemotherapy treatment can increase levels of both U and FU in DNA and the thymine oxidation damage product 5hmU can become incorporated into DNA either from the nucleotide pool or from in situ thymine oxidation (19–27). The parent oligonucleotide was constructed from three GGTTAC repeats with As flanking the 3′-repeat and in which the first two sites were mutated to GGTCAC, generating a single, high affinity Pot1 binding site. The overall length of the sequence allowed both DNA binding and DNA repair studies to be conducted on the same sequences.

In accord with previous studies, we confirmed the replacement of T by U diminished Pot1pN binding, and that the effect was much greater at the second T position (Table 1 and Figure 2). Similarly, the replacement of T by 5FU diminished Pot1pN binding although the increase of the Kd's was smaller. In previous studies, the T methyl group was described to be in a hydrophobic pocket formed by the folding of the GGTTAC sequence with formation of two unusual G-T base pairs (12,15). We therefore expected that the replacement of T with 5hmU would be highly disruptive due to the replacement of a hydrophobic methyl group with a hydrophilic 5-hydroxymethyl group. Surprisingly, we observed that the replacement of T with 5hmU had only a modest effect on Pot1pN binding (Table 1 and Figure 2).

The minor effect of 5hmU replacement on Pot1pN binding suggested that hydrophobicity was not the primary factor for the observed binding differences within the series examined here. The replacement of the 5-methyl group of T with the 5-hydrogen atom of U would create an energetically disfavored hole between the paired G and T bases. We therefore considered that the size of the 5-substituent might be the more important factor and that substituents that more completely fill the hole without disrupting the G-T base pairs would be less disruptive to Pot1 binding. In a previous study, we estimated the sizes of T, U, 5FU and
5hmU by calculating the SASA of each (41). Within this homologous series, differences in pyrimidine size can be attributed to differences in the 5-substituent size. The relative size of each pyrimidine is obtained by dividing the SASA of that pyrimidine by the SASA of thymine. As shown in Figure 3, a linear relationship is observed between the ln $K_d$ and the relative pyrimidine size for U, 5FU and T, with a greater effect at the T$_4$ position over the T$_3$ position. The decreased binding with 5hmU could indicate that the 5-hydroxymethyl group slightly exceeds the optimum size of the hole or might suggest a secondary hydrophobic effect. An examination of the crystal structure of the GGTTA–Pot1pN complex presented by Lei et al. (12) reveals that the pockets formed around the T-methyl groups are lined with multiple potential hydrogen bond acceptors for the 5-hydroxymethyl group of 5hmU which could explain the modest effect of 5hmU substitution on Pot1pN binding.

Having established that some pyrimidines substantially disrupt Pot1pN binding, whereas others have little effect, we next asked if Pot1pN binding would compete with glycosylase repair. Both U and 5FU are recognized and cleaved from single-stranded oligonucleotides by UNG, whereas the oxidized pyrimidine, 5hmU is removed by the single-strand specific uracil glycosylase hSMUG1 in the absence of Pot1pN, as shown in Figure 4 (lanes 2). The addition of Pot1pN to the reaction has no observable effect upon U cleavage (Figure 4a) which is expected as Pot1pN binds only weakly to U-containing sequences. The 5FU at position 4 is cleaved by UNG in the presence of Pot1pN. However, Pot1pN both binds to the sequence containing 5FU at position 3 and inhibits UNG cleavage. Sequences containing 5hmU at either position are bound by Pot1pN, and Pot1pN binding inhibits SMUG1 cleavage. These results establish that Pot1 binding could conceal some types of DNA damage from glycosylase cleavage, whereas other could prevent Pot1 binding and expose the site to glycosylase initiation of the BER cycle.

The cleavage of a base from DNA by a glycosylase generates an abasic site. In duplex DNA, the phosphodiester DNA backbone at an abasic site is cleaved by an apurinic/apyrimidinic (AP) endonuclease, creating a substrate for polβ repair synthesis. In this study, we examined Pot1pN binding to an abasic site as well as APE1 cleavage within the telomere repeat sequence using oligonucleotides containing a THF abasic site analog. Unlike the abasic site generated by a glycosylase, the reduced THF analog is stable to chemical hydrolysis but can be cleaved by APE1. The placement of the THF abasic site at positions G$_2$, T$_3$ or T$_4$ substantially decreases Pot1pN binding (Figure 5). Surprisingly, an abasic site at G$_1$ has only a modest effect upon Pot1pN binding. Abasic sites at the G positions could result from spontaneous depurination or guanine damage followed by glycosylase repair.

The THF abasic site, when placed at all four positions, G$_1$ through T$_4$, is efficiently cleaved by APE1 when in duplex DNA [Figure 6, lower panel (b)]. However, no APE1 cleavage is observed in single-stranded oligonucleotides. The absence of APE1 cleavage in single-stranded regions has been previously reported (54). Because DNA damage frequently thermally destabilizes duplex DNA configurations, we measured the melting temperatures of oligonucleotide duplexes containing each of the modified pyrimidines and THF abasic sites examined here (Table 2). Oligonucleotide duplex melting temperatures are usually reproducible to within ~0.5°C. Differences in experimental melting temperatures resulting from the pyrimidine substitutions examined here are at or near experimental error, and, therefore, in the system examined here, these substitutions have no significant effect upon oligonucleotide thermal stability. In contrast, the replacement of T or G with a THF abasic site reduces observed melting temperatures of duplexes formed by the telomere repeat sequence by 5–10°C, in accord with previously reported values in other DNA sequences (55). Because the telomere sequence is repeated and the G-T-rich strand is in excess over its complement, single-stranded regions could appear dynamically in loops along the entire length of the telomere.

Collectively, the results of this study suggest that DNA damage in telomeric sequences, arising from either the fraudulent incorporation of thymine analogs or chemical damage to DNA bases within the telomeric repeat, could create previously unrecognized challenges for telomere maintenance. The well-established BER cycle operating within duplex DNA regions is illustrated in the left side of Figure 7 for the case of uracil misincorporation. The GGTUA$^C$ sequence could be found in a duplex (black) or in a single-stranded region (blue). As the replacement of T with U has no significant effect upon thermal stability, the presence of U would not alter the duplex-single-stranded equilibrium. When single-stranded, the GGTUA$^C$ sequence can fold and bind to Pot1 proteins. However, in the case of uracil substitution, the misincorporated uracil would interfere with Pot1 binding, causing the uracil residue to be exposed. A uracil residue within this sequence is readily removed by UNG, generating an abasic site. The decreased duplex stability generated by an abasic site within a repetitive sequence could promote DNA sliding, making it more likely that the abasic-site-containing region was in a single-stranded region. The localization of DNA damage within other repetitive sequences, in particular abasic sites, has been recently demonstrated (56,57).

An abasic site located within a duplex region can be cleaved by APE1, creating a substrate for polβ-mediated repair synthesis. In contrast, when the abasic site is located within a single-stranded region, APE1 activity is usually undetectable (54). The results of our study are in accord with these previous findings. It is perhaps advantageous that APE1 does not cleave abasic sites when not in a duplex as cleavage of a single strand could cause telomere shortening. In the absence of APE1 cleavage, however, an abasic site generated by a glycosylase can undergo spontaneous hydrolytic cleavage with a half-life of ~3 days (58). Therefore, DNA repair can be initiated within the telomeric sequence but stalls at the APE1 cleavage step, potentially resulting in spontaneous hydrolysis and loss of the distal telomere.

Our studies reveal that attempted repair of base damage or base analogs within single-stranded telomeric sequences poses a threat to telomeres, suggesting that some alternative repair pathways might exist. Others have shown that abasic sites located within some unusual DNA structures are cleaved by APE1 (59–61). Alternatively, Wiederhold et al. (62) have described an APE1-independent
BER pathway that relies on NEIL (endonuclease VIII like) family glycosylases to cleave abasic sites in unusual DNA structures. In support of this alternative pathway, Zhou et al. (63) demonstrated that NEIL1 and NEIL3 can cleave damaged bases including 8-oxoguanine and thymine glycol within quadruplex-forming telomere repeat sequences. The analogs examined here have not as yet been examined in these systems. While the results described above follow reasonably from the in vitro studies with model systems described here, extrapolation into biological systems is more complicated. In the study described here, we placed a single, high affinity GGATTC sequence at the 3′-end of the oligonucleotide. Telomeres in vivo are comprised of an array of multiple repeats. The presence of multiple repeats has been shown to generate cooperative Pot1 binding that could diminish or amplify the effect of a single pyrimidine substitution (13). Additionally, Pot1 proteins often form complexes with other proteins that could either diminish or amplify the effect of DNA modification or damage (64–66). Recently, Nandakumar et al. examined the binding of hPOT1 and the hPOT1-hTPP1-N heterodimer to the sequence GGT-TAGGGTTAG containing tandem human Pot1 binding sites (15). They report that a single 2′-deoxyuridine substitution, generating GGTUAGGGTTAG, did not change the hPOT1 $K_d$. Interestingly, the observed $K_d$ relative to the control sequence increased 7-fold (diminished binding) for the heterodimer complex upon uracil substitution. The results reported in our study suggest that further examination of the effect of DNA damage within telomere repeats in more complicated systems involving multiple repeats and multiple interacting proteins is warranted.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGMENT**

We thank Dr Kangling Zhang and Hui Tang for assistance with mass spectrometric characterization of POT1 protein.

**FUNDING**

National Institutes of Health, National Institute of General Medical Sciences [GM50351]. Funding for open access charge: National Institutes of Health, National Institute of General Medical Sciences [GM50351].

**Conflict of interest statement.** None declared.

**REFERENCES**


Downloaded from https://academic.oup.com/nar/article-abstract/42/14/9063/1276853 by guest on 24 April 2019