

## ORIGINAL ARTICLE

# Molecular dynamics simulation of telomeric single-stranded DNA and POT1

Masaaki Kaburagi<sup>1,2</sup>, Hironao Yamada<sup>3</sup>, Takeshi Miyakawa<sup>3</sup>, Ryota Morikawa<sup>3</sup>, Masako Takasu<sup>3</sup>, Takamitsu A Kato<sup>4</sup> and Mitsuru Uesaka<sup>2</sup>

Telomeres, which consist of single-, double- and four-stranded DNA, shorten after each round of cell division. The repeated telomeric DNA sequence 5'-TTAGGG-3' does not encode genetic information and is not replicated completely. We performed molecular dynamics (MD) simulations of telomeric single-stranded DNA (ssDNA) and protection of telomere 1 (POT1) for 100 ns. We calculated the distance between C<sub>α</sub> (POT1) and O5' (telomeric ssDNA) to verify the binding system for 100 ns MD. We then calculated the distance between the bases of the telomeric DNA ends and the root-mean-square deviation and gyration radius in the single and binding states. Moreover, we compared the root-mean-square fluctuations (RMSFs) between the single and binding states and calculated the number of hydrogen bonds between POT1 and telomeric DNA. There are many hydrogen bonds between Gln94 and the first guanine of the closest 5'-TTAGGG-3' sequence in telomeric single-stranded DNA, and the RMSF between the single and binding states has a large difference between Gln94 and guanine. Overall, we found that Gln94 and guanine are important components of the binding system and are related to the stability of this system.

*Polymer Journal* (2016) 48, 189–195; doi:10.1038/pj.2015.82; published online 14 October 2015

## INTRODUCTION

DNA consists of deoxyribose sugars, phosphoric acids and the nitrogenous bases adenine (A), thymine (T), guanine (G) and cytosine (C). The length of telomeric DNA, which is located at the terminal ends of linear eukaryotic chromosomes,<sup>1</sup> is a factor in determining the lifetime of a normal cell.<sup>2</sup> Telomeric DNA is shortened after each round of cell division because the repeated telomeric DNA sequence 5'-TTAGGG-3', which does not encode genetic information, is not replicated completely. Telomeres are stabilized by specialized T-loop and D-loop structures.<sup>3</sup> T-loops and D-loops comprise single-stranded DNA (ssDNA), double-stranded DNA (dsDNA) and specific proteins that maintain the telomeric structure. Telomeric repeat factor (TRF) 1 and TRF2 are telomeric dsDNA-binding proteins,<sup>4,5</sup> whereas protection of telomere 1 (POT1) is a telomeric ssDNA-binding protein.<sup>3,6</sup>

Of the DNA bases, guanine most readily reacts with OH radicals, which are oxidizing agents. OH radicals are often formed when water is exposed to ionizing or ultraviolet radiation or during metabolic activity within the cell.<sup>7</sup> Mutations can arise within a DNA sequence as the result of guanine oxidation to 7,8-dihydro-8-oxoguanine (8-oxoG). Thus, gaining mechanistic insight into the genomic maintenance of telomeric DNA is important for studying oncogenesis<sup>8,9</sup> and the biological effects of ionizing radiation.<sup>10,11</sup>

POT1 protects single-stranded telomeric DNA and controls telomere length. When bound to telomeric ssDNA, POT1 prevents replication protein A from binding.<sup>12</sup> The replication protein A

protein consists of three subunits and is involved in the DNA damage checkpoint pathway. POT1 is also a key protein that prevents the activation of Ataxia Telangiectasia and Rad3-related kinase, which is involved in DNA damage repair.<sup>13</sup>

Here, we performed molecular dynamics (MD) simulations of telomeric dsDNA and TRF1 to further explore the telomeric protein-binding system.<sup>14</sup> We found structural differences in the telomeric dsDNA depending on the binding of TRF1. Furthermore, we investigated the relationship between telomeric ssDNA and POT1 by performing MD simulations, which revealed a novel role for POT1 in maintaining telomeric ssDNA.

POT1 and telomeric ssDNA-binding systems are important for studying telomere maintenance and cell longevity, and several reports have explored this relationship. Ramos *et al.*<sup>15</sup> simulated and analyzed POT1 and a telomeric ssDNA fragment in the bound state. Their initial structure was obtained from an X-ray diffraction experiment (PDBID: 1QZH<sup>16</sup>). Chatterjee *et al.*<sup>17</sup> performed MD simulations of an ssDNA fragment without protein for 400 ps and calculated the root-mean-square deviation (RMSD) of the ssDNA in the single state. Their sequence of ssDNA was a fragment of p53-coding DNA (130–140 codon sequence). The RMSD values of this ssDNA under the constant particle number, the constant volume, constant temperature (NVT) and the constant particle number, the constant pressure, the constant temperature (NPT) conditions were also calculated for 4 and 5 ns. These RMSD values increased during the simulations. Jaiswal

<sup>1</sup>Research Planning Department, Fukushima Research Infrastructural Creation Center, Japan Atomic Energy Agency, Fukushima, Japan; <sup>2</sup>Department of Nuclear Engineering and Management, School of Engineering, University of Tokyo, Tokyo, Japan; <sup>3</sup>Department of Molecular Life Sciences, School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, Tokyo, Japan and <sup>4</sup>Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO, USA  
Correspondence: M Kaburagi, Iwaki Office, Japan Atomic Energy Agency, 8F Taira Central Building, 7-1 Aza-Omachi, Taira, Iwaki, Fukushima 970-8026, Japan.  
E-mail: kaburagi.masaaki@jaea.go.jp

Received 4 May 2015; revised 21 July 2015; accepted 22 July 2015; published online 14 October 2015

*et al.*<sup>18</sup> performed MD simulations of the binding systems of POT1 and peptides for 3 ns and calculated the RMSD, root-mean-square fluctuation (RMSF) and number of hydrogen bonds (H-bonds) in the binding systems.

Luscombe *et al.* reported a three-dimensional analysis of the protein–DNA interaction at an atomic level.<sup>19</sup> Specifically, they studied the structures of protein–DNA binding systems based on a crystal analysis. Lei *et al.*<sup>20</sup> studied the H-bonds of the crystal structures between POT1 and telomeric ssDNA. In this work, we studied the H-bonds between POT1 and telomeric ssDNA using MD simulation and compared the results of this simulation with the X-ray crystal structure.

## MATERIALS AND METHODS

### POT1 protein and telomeric ssDNA

The telomeric ssDNA fragment sequence in 3KJP<sup>21</sup> was defined as 5′-GGTTAGGGTTAG-3′, which is a long telomeric ssDNA fragment. The number of bases in the telomeric ssDNA is indicated from the 5′ terminus, and the initial guanine base in 3KJP is G2. The POT1 and telomeric ssDNA structures were individually obtained from the X-ray crystal structure (PDBID: 3KJP). The single state is defined as the state of being unbound, and the binding state is defined as the state of being bound. POT1 and telomeric ssDNA in the single and binding states were used to analyze the structural dynamics of POT1 and telomeric ssDNA.

### MD simulation

The GROMACS version 4.5.5 package<sup>22</sup> was used to perform all-atom MD simulations. Conformation figures were drawn using VMD software,<sup>23</sup> and AMBER99SB-ILDN force field<sup>24,25</sup> was used to assess atomic interactions. The TIP3P water model,<sup>26–28</sup> which is a three-point electrostatic interaction model, was used as a solvent. First, to neutralize the total charge of each simulated system, 4 Cl<sup>−</sup> ions were inserted into the POT1 single system, 10 Na<sup>+</sup> ions were inserted into the telomeric ssDNA single system and 6 Na<sup>+</sup> ions were inserted into the POT1 and telomeric ssDNA-binding system. Next, Na<sup>+</sup> and Cl<sup>−</sup> ions were adjusted in each system to reach a NaCl concentration of 0.16 mol l<sup>−1</sup>. Furthermore, 75 Na<sup>+</sup> ions, 69 Cl<sup>−</sup> ions and 22 127 water molecules were inserted into the binding system, and 65 Na<sup>+</sup> ions, 69 Cl<sup>−</sup> ions and 20 805 water molecules were inserted into the POT1 single system. Finally, 28 Na<sup>+</sup> ions, 18 Cl<sup>−</sup> ions and 5960 water molecules were inserted into the telomeric ssDNA single system.

Energy minimization was performed before MD simulations using the steepest descent method followed by the conjugated gradient method to remove any large force contacts. Periodic boundary conditions using the Particle Mesh Ewald method<sup>29,30</sup> were applied to calculate long-range Coulomb interactions. For non-bonded interactions, the cutoff was 10 Å. Using the linear constraint

solver algorithm, these constraints were applied to all bonds, including hydrogen bonds between heavy atoms.<sup>31</sup> The leapfrog method was selected to integrate atom dynamics with a step length of 2.0 fs. The temperature was controlled at 300 K during 100 ps NVT-MD, with protein and DNA molecules restrained. The velocity scaling method<sup>32</sup> was applied to control the temperature with a 0.1-ps time-step. In addition, pressure was applied at 1 bar during 100 ps NPT-MD, with protein and DNA molecules restrained. The isotropic Parrinello-Rahman<sup>33</sup> method was applied to control the pressure in 2-ps time-step increments. The compressibility of water is  $4.5 \times 10^{-5} \text{ bar}^{-1}$ . NVT-MD and NPT-MD simulations were performed to maintain the solvent in thermal equilibrium. Here, NVT-MD and NPT-MD are defined as pre-MD. After pre-MD, NPT-MD was performed without protein or DNA restraints for 100 ns, which is defined as the production-MD. In our study, pre-MD and production-MD for POT1 and telomeric ssDNA-binding state, POT1 single state and ssDNA single state were performed. These simulations were independently performed three times with different initial configurations and velocity distributions of water molecules.

## RESULTS

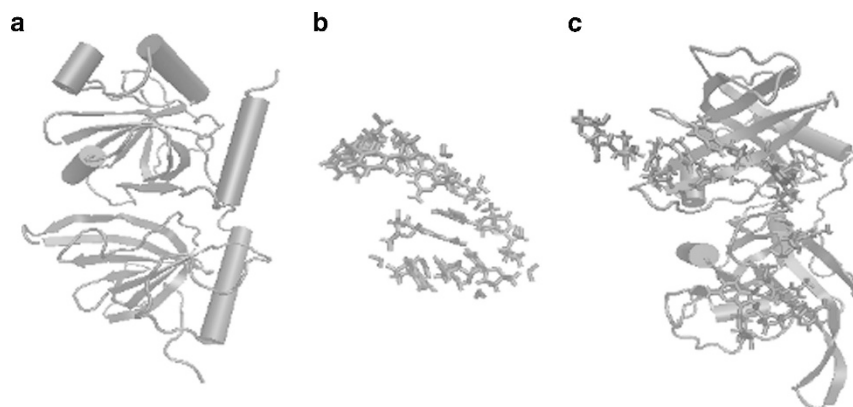
### ssDNA and POT1 conformation in single and binding states

To examine POT1 and telomeric ssDNA structures in the single and binding states in water, 100 ns production-MD was used to illustrate the conformation of molecules under both conditions (Figure 1). The structure of telomeric ssDNAs differed between the single and binding states. Based on an analysis of the crystal structure, telomeric ssDNA in the single state changed from the binding state structure into a C-shape (Figure 1b). However, the structure of telomeric ssDNA in the binding state did not change into a C-shape during the 100 ns production-MD.

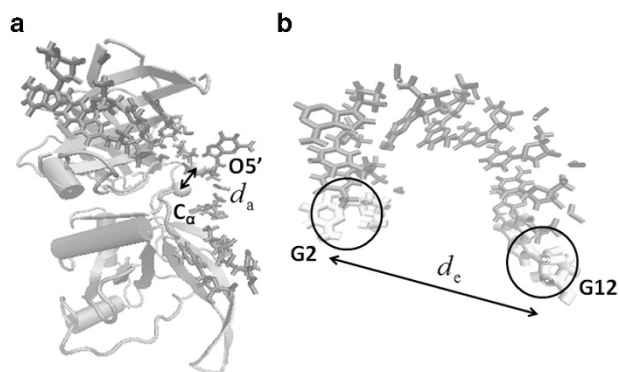
### Assessing POT1 and telomeric ssDNA-binding stability

We evaluated the distance between POT1 and telomeric ssDNA in the single state to determine the stability of the binding system. The access distance ( $d_a$ ) was defined as the distance between the closest C<sub>α</sub> (POT1) and O5′ (telomeric ssDNA) pair (Figure 2a). To investigate the binding state of telomeric ssDNA and POT1 after 100 ns production-MD,  $d_a$  values were calculated in each sample using a 20-ps time-step. The ensemble average of  $d_a$  was calculated for each sample in water under the three different initial conditions.

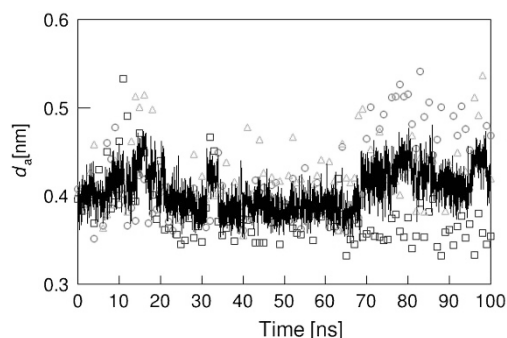
As shown in Figure 3, the  $d_a$  value averaged over the three initial conditions continued to be between 0.35 and 0.48 nm. The  $d_a$  value averaged over time was 0.40 nm, and the deviation was less than 0.10 nm. These small  $d_a$  values indicated a constant binding state



**Figure 1** Structural analysis following 100 ns production-molecular dynamics. (a) Protection of telomere 1 (POT1) in the single state, (b) telomeric single-stranded DNA (ssDNA) in the single state and (c) POT1 (blue) and telomeric ssDNA (red) in the binding state. A full color version of this figure is available at the *Polymer Journal* journal online.



**Figure 2** (a) The access distance ( $d_a$ ) between  $C_\alpha$  (lower yellow sphere) in protection of telomere 1 and  $O5'$  (upper yellow sphere) in telomeric single-stranded DNA (ssDNA). (b) End-to-end distance ( $d_c$ ) between the telomeric ssDNA ends. A full color version of this figure is available at the *Polymer Journal* journal online.



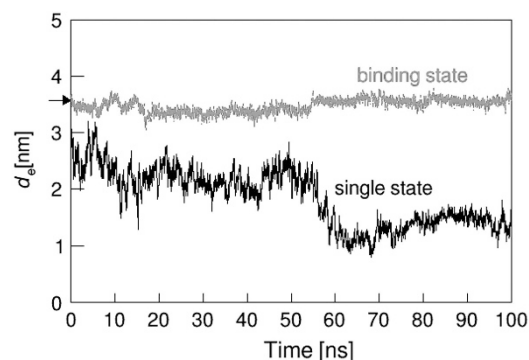
**Figure 3** The ensemble average of access distance ( $d_a$ ) between protection of telomere 1 and the telomeric single-stranded DNA (line). The three different values of  $d_a$  for the three different water initial conditions (circle, square, triangle) plotted over 1.0-ns time steps. A full color version of this figure is available at the *Polymer Journal* journal online.

between POT1 and telomeric ssDNA. The  $d_a$  values of each sample for the three different initial conditions are also plotted by 1-ns intervals.

#### Calculating the distance between single-state telomeric ssDNA ends

The telomeric ssDNA structure in the single state changed from the structure in the binding state into a C-shape (Figure 1b). This C-shape change was further analyzed by measuring the distance ( $d_c$ ), which is the distance between the centers of mass (COM) of G2 and G12 (Figure 2b). The  $d_c$  values in the single and binding states were calculated using a 20-ps time-step. The ensemble average of  $d_c$  was also calculated. When the telomeric ssDNA structure in the single state changed from the structure in the binding state into a C-shape, the  $d_c$  value decreased.

The  $d_c$  values in the single and binding states are shown in Figure 4. The  $d_c$  values in the binding state were between 3.0 and 3.8 nm, with an average value of 3.5 nm for the 100-ns production-MD. The  $d_c$  value in the binding state was 3.6 nm for the initial crystal structure, but the crystal structure in the single state has not yet been identified to the best of our knowledge. The  $d_c$  values in the binding state for the 100-ns production-MD remained near the  $d_c$  value for the crystal structures. Conversely, the  $d_c$  values in the single state decreased from between 50 and 60 ns to between 0.80 and 3.2 nm. The  $d_c$  values in the single state were smaller than all  $d_c$  values in the binding state for the 100-ns production-MD and the initial crystal structure. The  $d_c$



**Figure 4** The distance between telomeric single-stranded DNA G2 and G12 centers of mass ( $d_c$ ) in the single (black) and binding (red) states and  $d_c$  for the initial crystal structure (arrow). A full color version of this figure is available at the *Polymer Journal* journal online.

temporal fluctuations in the single state were larger than those in the binding state.

#### Root-mean-square deviation

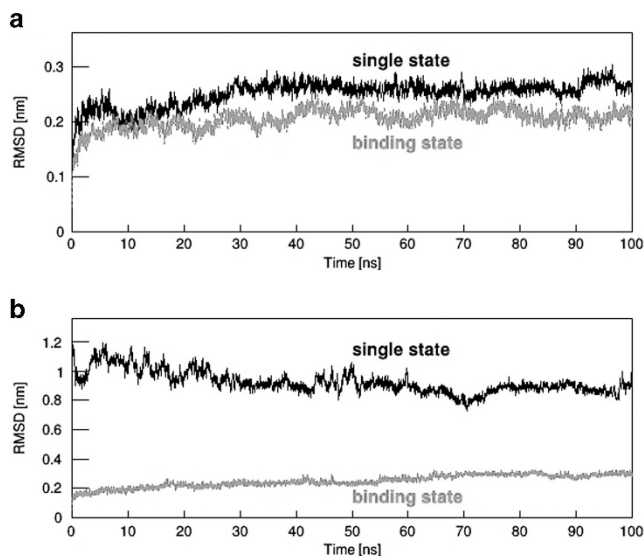
The previous section showed that the telomeric ssDNA structures differ between the single and binding states, whereas the POT1 structures do not differ between states (Figure 1). We quantitatively evaluated the structural differences in these molecules between the single and binding states. Thus, we analyzed the results of the MD of each molecule as a time series using the RMSD of the POT1 backbone and all telomeric ssDNA atoms in the single and binding states over a 20-ps time-step. The backbone is the main chain of a protein or  $-NC_\alpha CNC_\alpha C-$  in this calculation. As shown in equation (1), the RMSD of certain atoms in a molecule with respect to the atoms in the reference structure was calculated. The initial binding crystal structure was used as the reference structure. The COM of the reference and MD-snapshot structures were then superimposed. Next, the MD-snapshot structures were rotated to minimize the use of least-square fitting by RMSD.

$$\text{RMSD}(t) = \left[ \frac{1}{M} \sum_{i=1}^N m_i (\mathbf{r}_i(t) - \mathbf{r}_i(0))^2 \right]^{\frac{1}{2}} \quad (1)$$

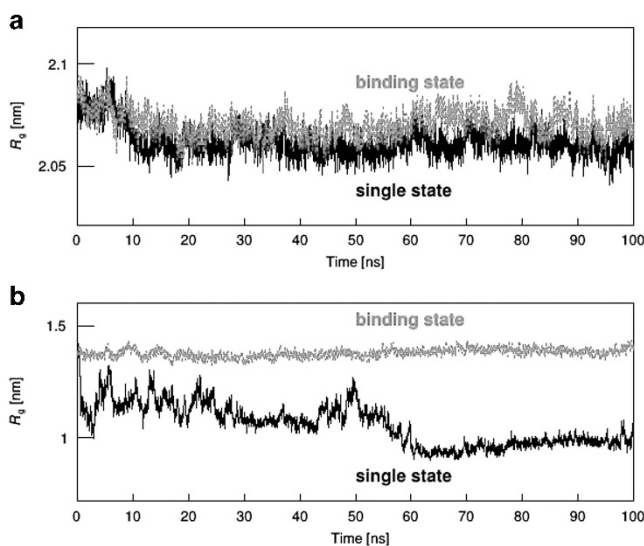
Here, the POT1 and telomeric ssDNA reference structures served as the initial crystal structure. The coordinate  $\mathbf{r}_i(0)$  represents the  $i$ -th atom of the initial crystal structure (PDBID: 3KJP). The coordinate  $\mathbf{r}_i(t)$  represents the position of the  $i$ -th atom at time  $t$ .  $N$  is the number of either backbone atoms (POT1) or all atoms (the telomeric ssDNA). The mass of the  $i$ -th atom is represented by  $m_i$ , and the mass  $M$  is the sum of the masses of either the backbone atoms (POT1) or all atoms (the telomeric ssDNA).

The RMSD values of POT1 are shown in Figure 5a. The average RMSD values calculated after 40 ns in the single and binding states were 0.25 and 0.20 nm, and the deviations were 0.03 nm in each case. The difference between the RMSD values in the single and binding states was less than 0.10 nm. These results show that the single and binding structures of POT1 are similar.

In addition, we quantitatively evaluated the difference between the single and binding structures of POT1. We computed the average single and average binding structures of POT1 after 40 ns in the 100-ns production-MD. The samples were individually analyzed under the three different initial conditions, and we calculated the RMSD values of nine pairs between the average structures in the single and binding structures after least-square fitting. The maximum value of the RMSD between two average single and average binding structures of



**Figure 5** The root-mean-square deviation (RMSD) in the single (black) and binding (red) states for the (a) protection of telomere 1 backbone and (b) telomeric single-stranded DNA. A full color version of this figure is available at the *Polymer Journal* journal online.



**Figure 6** The  $R_g$  values for the single (black) and binding (red) states for the (a) protection of telomere 1 backbone and (b) telomeric single-stranded DNA. A full color version of this figure is available at the *Polymer Journal* journal online.

POT1 was 0.27 nm, which was sufficiently small for the size of POT1. Overall, the results indicated that the single and binding structures of POT1 were the same.

The RMSD values of telomeric ssDNA are shown in Figure 5b. The average RMSD values calculated after 65 ns in the single and binding states were 0.92 and 0.24 nm, respectively, and the deviations were 0.11 and 0.06 nm, respectively. The average RMSD values and deviations were larger in the single state than in the binding state. The telomeric ssDNA RMSD values in the binding state gradually increased, but the increase was small. The results show that the telomeric ssDNA structures in the single and binding states were different.

### Radius of gyration

To study molecular size changes over time, the gyration radius ( $R_g$ ) of the POT1 backbone or of all telomeric ssDNA atoms in each single and binding state was calculated using a 20-ps time-step. The  $R_g$  value was calculated as the mean squared distance of the atoms from the COM of the molecule, as shown in equation (2).

$$R_g(t) = \left[ \frac{1}{M} \sum_{i=1}^N m_i (\mathbf{r}_i(t) - \mathbf{r}_{\text{COM}}(t))^2 \right]^{\frac{1}{2}} \quad (2)$$

Here, the coordinate  $\mathbf{r}_{\text{COM}}(t)$  is the molecule's COM position at time  $t$ . The average  $R_g$  values were calculated for each sample in water under different initial conditions. The  $R_g$  values of POT1 in the single and binding states are shown in Figure 6a. The difference between the binding and single state values was less than 0.03 nm. The  $R_g$  values of telomeric ssDNA in the single and binding states are shown in Figure 6b. The single state values changed between 0.90 and 1.4 nm; however, the binding state values were stable between 1.3 and 1.4 nm during the 100-ns production-MD simulation. The deviations in the single and binding states were 0.15 and 0.06 nm, respectively. In the single state, these values dramatically changed between 50 and 60 ns, which was accompanied by a large RMSD change.

### Root-mean-square fluctuation

To determine which parts of one molecule within the binding system are influenced by the other, we compared the fluctuation of either POT1 residues or telomeric ssDNA bases in the single state with those in the binding state. In our study, RMSF is the COM positions of POT1 residues or telomeric ssDNA bases, as shown in equation (3).

$$\text{RMSF} = \left[ \frac{1}{T_{\text{total}}} \sum_{t=t_0}^{t_f} |\mathbf{r}(t) - \bar{\mathbf{r}}|^2 \Delta t \right]^{\frac{1}{2}} \quad (3)$$

Here, the coordinate  $\mathbf{r}(t)$  is the COM position of the POT1 residue or the telomeric ssDNA base.  $T_{\text{total}}$  is the total time from the initial time  $t_0$  to the final time  $t_f$ ;  $T_{\text{total}} = t_f - t_0$ . During the 100-ns production-MD, the initial time  $t_0$  of the RMSF calculation is 40 ns in the POT1 and 65 ns in the telomeric ssDNA because the RMSD and  $R_g$  values did not largely increase or decrease after those times. The average COM position of a residue or a base in each sample is  $\bar{\mathbf{r}}$ , as shown in equation (4).

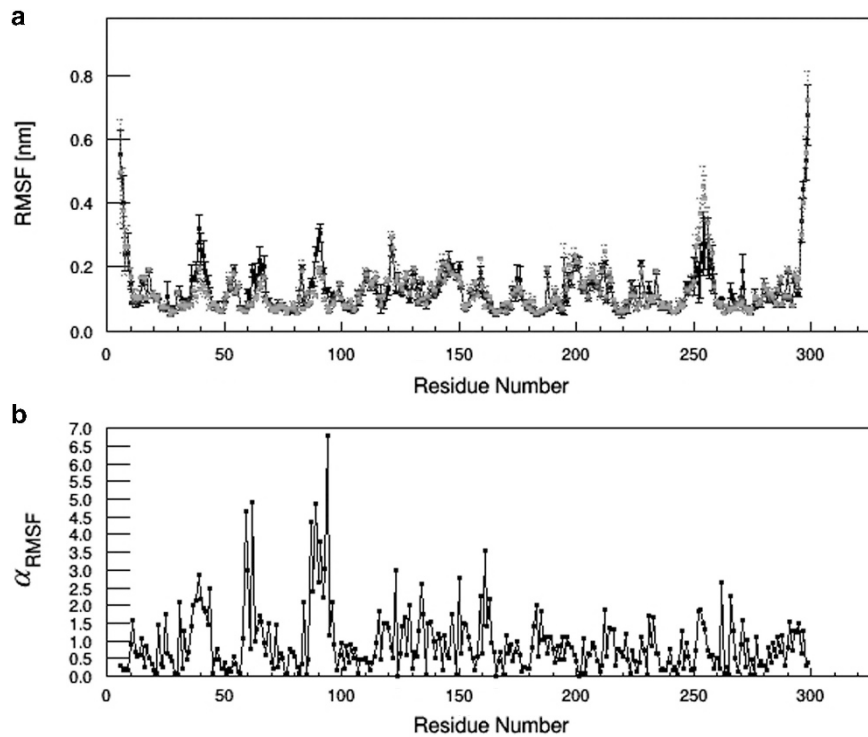
$$\bar{\mathbf{r}} = \frac{1}{T_{\text{total}}} \sum_{t=t_0}^{t_f} \mathbf{r}(t) \Delta t \quad (4)$$

The RMSF values of POT1 residues and telomeric ssDNA bases were calculated for each sample in the single and binding states. The average values and RMSF variance ( $S_r^2$ ) were calculated. The variances of each sample in water were calculated for the different initial conditions.

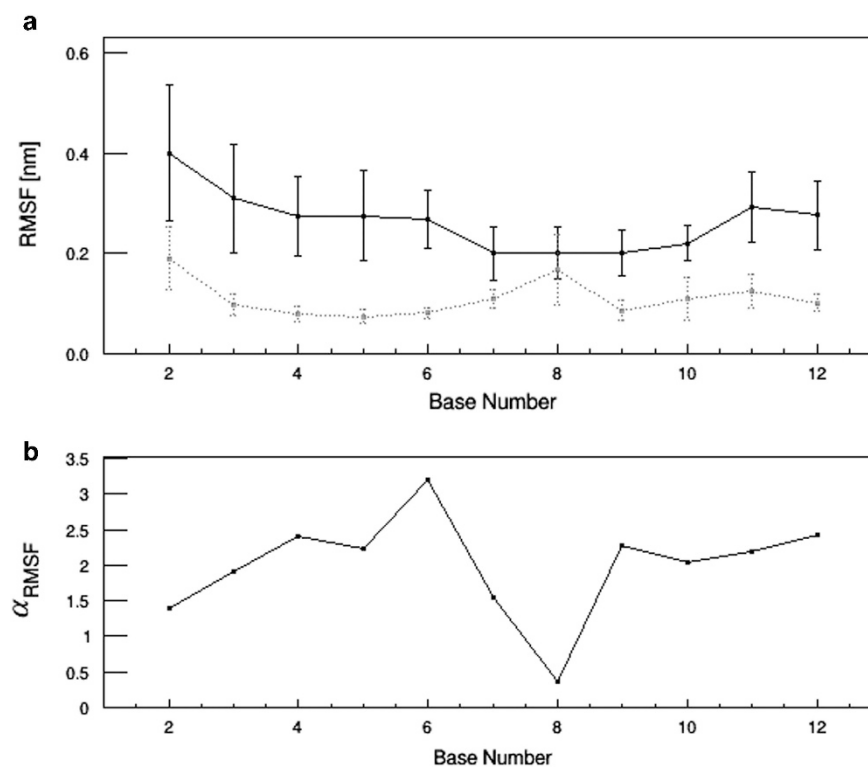
We evaluated the difference in RMSF values between POT1 and telomeric ssDNA in both the single and binding states. The binding value is defined as  $\alpha_{\text{RMSF}}$  and was used to compare the values of RMSF in the single and binding states, as shown in equation (5).

$$\alpha_{\text{RMSF}} = \sqrt{\frac{(\text{RMSF}(\text{binding state}) - \text{RMSF}(\text{single state}))^2}{s_r^2(\text{binding state}) + s_r^2(\text{single state})}} \quad (5)$$

The RMSF and  $s_r$  values of POT1 residues in the single and binding states are shown in Figure 7a, and the  $\alpha_{\text{RMSF}}$  values are shown in Figure 7b. The maximum  $s_r$  values were 0.10 nm at Glu254 in the single state and 0.16 nm at Ala6 in the binding state. In POT1, only the  $\alpha_{\text{RMSF}}$  value of Gln94 was more than 6.0.



**Figure 7** (a) The root-mean-square fluctuation (RMSF) of the protection of telomere 1 (POT1) residues in the single (solid black line) and binding (red dotted line) states and the variance (error bar) after 40 ns. (b)  $\alpha_{RMSF}$  of POT1 residues. A full color version of this figure is available at the *Polymer Journal* journal online.



**Figure 8** (a) The root-mean-square fluctuation (RMSF) of telomeric single-stranded DNA (ssDNA) bases in single (solid black line) and binding (red dotted line) states and the variance (error bar) after 65 ns. (b)  $\alpha_{RMSF}$  of telomeric ssDNA bases. A full color version of this figure is available at the *Polymer Journal* journal online.

The average RMSF and  $s_r$  values of telomeric ssDNA in the single and binding states are shown in Figure 8a. The dynamics of atoms in this short dsDNA may differ from those in longer dsDNAs, especially because water molecules influence both ends of telomeric ssDNA. The  $\alpha_{\text{RMSF}}$  values are shown in Figure 8b. The RMSF values of telomeric ssDNA in the single state are larger than those in the binding state. The maximum  $s_r$  values were 0.14 nm at G2 in the single state and 0.07 nm at G8 in the binding state. In telomeric ssDNA, only the  $\alpha_{\text{RMSF}}$  value of G6 was more than 3.0.

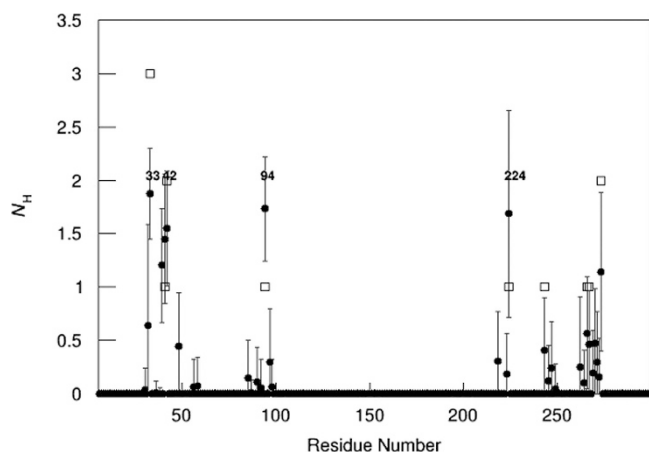
Upon evaluation of  $\alpha_{\text{RMSF}}$ , the difference in the RMSF of Gln94 and G6 was found to be the largest between the single and binding states of POT1 and telomeric ssDNA.

### Hydrogen bonds between POT1 and telomeric ssDNA

Of the non-covalent interactions, such as van der Waals' forces and hydrophobic bonds, hydrogen bonding (H-bond) constitutes the strong interaction between molecules. To determine the parts of the binding system that strongly interacted with other molecules, the H-bonds between POT1 and telomeric ssDNA were investigated. To this end, the three-dimensional geometrical-based hydrogen (H)-bond criteria<sup>34</sup> were used. These criteria are based on the angle and distance between an acceptor (A) and donor (D) pair. In our study, the criteria were a D-H-A angle of the H-bond of less than 30° and a distance between a D-A pair of less than 0.32 nm. To study the interactions between a POT1 residue and a telomeric ssDNA base, the number of H-bonds between telomeric ssDNA and POT1 was calculated using a 20-ps time-step for 100 ns production-MD. The number of H-bonds in the initial crystal structure (PDBID: 3KJP) was also calculated. The average number of H-bonds ( $N_{\text{H}}$ ) over time was calculated for POT1 residues and telomeric ssDNA bases for three different initial water conditions.

The  $N_{\text{H}}$  values of POT1 for the 100-ns production-MD and the values in the initial crystal structures are shown in Figure 9. The  $N_{\text{H}}$  values were larger than 1.5 in Lys33, Asp42, Gln94 and Asp224 in the 100-ns production-MD and less than 1.1 in Lys33, Asp42, and Asp224 in the initial crystal structure. There was a large difference between  $N_{\text{H}}$  for the 100-ns production-MD and  $N_{\text{H}}$  of the initial crystal structure.

The  $N_{\text{H}}$  values of telomeric ssDNA for 100-ns production-MD and initial crystal structure values are shown in Figure 10. The  $N_{\text{H}}$  values were greater than or equal to 2.0 in T4, G6 and G12 for the 100-ns production-MD and in T4, G6, T9 and G12 for the initial crystal



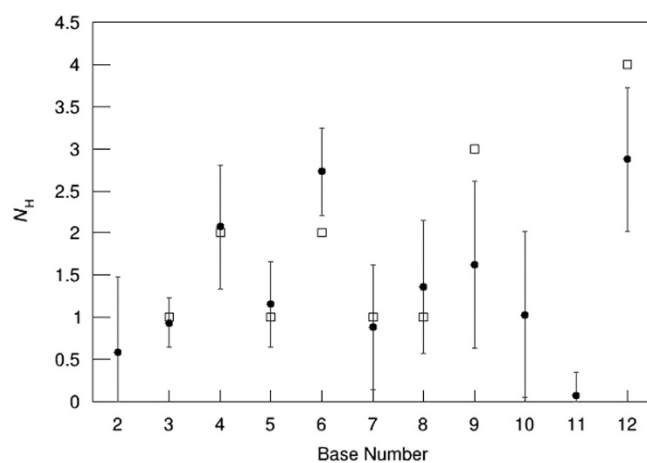
**Figure 9** The  $N_{\text{H}}$  of protection of telomere 1 for the 100-ns production-molecular dynamics (closed circle) and  $N_{\text{H}}$  of the initial crystal structure (open square). A full color version of this figure is available at the *Polymer Journal* journal online.

structure. The number of H-bonds was higher in G6 than in T9 for the 100-ns production-MD but lower in G6 than in T9 for the crystal structure.

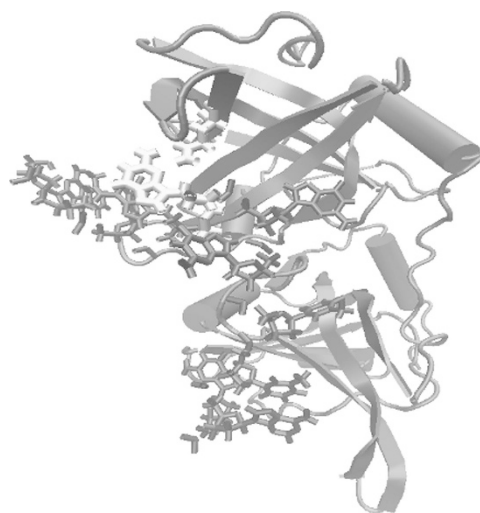
The  $N_{\text{H}}$  value between Gln94 and G6 was calculated to be 1.74 with a deviation of 0.49. In addition, the  $N_{\text{H}}$  value between Gln94 and all telomeric ssDNA bases except for G6 was zero, and the  $N_{\text{H}}$  value between G6 and all POT1 residues except for Gln94 was 0.99. Thus, the H-bonds of Gln94 only pair with G6 in telomeric ssDNA, but G6 forms H-bonds with other POT1 residues. The geometric positions of Gln94 and G6 (Figure 11) indicate that Gln94 and G6 are close to each other in this binding system.

### DISCUSSION

Our initial structure (PDBID: 3KJP<sup>21</sup>) includes more residues of POT1 and a longer sequence of telomeric ssDNA than the initial structure studied by Ramos *et al.* (PDBID: 1QZH<sup>16</sup>). Our MD simulations are for the single and binding systems of POT1 and telomeric ssDNA and



**Figure 10** The  $N_{\text{H}}$  of the telomeric single-stranded DNA for the 100-ns production-molecular dynamics (closed circle) and  $N_{\text{H}}$  of the initial crystal structure (open square). A full color version of this figure is available at the *Polymer Journal* journal online.



**Figure 11** Snapshots of protection of telomere 1 (blue) and telomeric single-stranded DNA (red) after 100-ns production-molecular dynamics, where G6 and Gln 94 are in yellow. A full color version of this figure is available at the *Polymer Journal* journal online.

were conducted for a longer time. Furthermore, we calculated other physical quantities, not only for the binding system but also for the single systems. Our simulations under NPT conditions, which were longer than those conducted by Chatterjee *et al.*,<sup>17</sup> showed that the RMSD decreased because the structure of the telomeric ssDNA changed. Our analysis of  $d_c$ , RMSD and  $R_g$  also revealed a large difference between the single structure of telomeric ssDNA and the structure of telomeric ssDNA attached to POT1.

Jaiswal *et al.*<sup>18</sup> studied the binding systems of POT1 and peptides, but our study focused on the binding systems of POT1 and telomeric ssDNA and compared the single and binding systems of POT1 and telomeric ssDNA.

The time series of telomeric ssDNA  $d_c$  and  $R_g$  values in the single states also indicates how the telomeric ssDNA structure in the single state does not change upon removal of POT1. We found that the structure of telomeric ssDNA in the single state changed into a C-shape from the structure in the binding states and that the telomeric ssDNA shape is sustained by POT1.

We calculated and compared the number of H-bonds in the binding system of the initial crystal structure and during the 100-ns production-MD between POT1 and telomeric ssDNA. Our calculations show that the number of H-bonds between POT1 and telomeric ssDNA is very different between the 100-ns production-MD and the crystal structure.

Overall, our study suggest that Gln94 and G6 are important parts of the binding systems of POT1 and telomeric ssDNA.

## CONCLUSIONS

In this study, we investigated the features of the POT1 and telomeric ssDNA single and binding systems. Our study shows that the telomeric ssDNA structure is sustained by POT1; however, the converse is not true: telomeric ssDNA binding does not alter the POT1 structure. In addition, our study shows not only that Gln94 frequently forms H-bonds between POT1 and telomeric ssDNA exclusively with G6 but also that G6 forms frequent H-bonds with other residues. Overall, we found that G6 and Gln94 are important components of the POT1 and telomeric ssDNA-binding system.

In the future, MD simulations of longer telomeric ssDNA sequences should be performed. Of the DNA bases, guanine is the most chemically sensitive to OH radicals, which often oxidize guanine to form 8-oxoG. Therefore, MD simulations using telomeric ssDNA in which G6 is replaced by other bases, such as 8-oxoG, should also be performed.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

- 1 Blackburn, E. H. Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. *FEBS Lett.* **579**, 859–862 (2005).
- 2 Blackburn, E. H. Telomere states and cell fates. *Nature* **408**, 53–56 (2000).
- 3 Loayza, D. & Lange, T. D. POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **423**, 1013–1018 (2003).
- 4 Bianchi, A., Smith, S., Chong, L., Elias, P. & Lange, T. D. TRF1 is a dimer and bends telomeric DNA. *EMBO J.* **16**, 1785–1794 (1996).
- 5 Steensel, B. V., Smogorzewska, A. & Lange, T. D. TRF2 protects human telomeres from end-to-end fusions. *Cell* **92**, 401–413 (1998).
- 6 Baumann, P. & Cech, T. R. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **292**, 1171–1175 (2001).
- 7 Begusova, M., Michalik, M. S. D. S. V. & Charlier, M. RADACK, a stochastic simulation of hydroxyl radical attack to DNA. *J Biomol Struct Dyn.* **19**, 141–158 (2001).
- 8 Chang, S. Cancer chromosomes going to POT1. *Nat. Genet.* **45**, 473–475 (2003).

- 9 Ramsay, A. J., Quesada, V., Foronda, M., Conde, L., Martínez-Trillos, A., Villamor, N., Rodríguez, D., Kwarciak, A., Garabaya, C., Gallardo, M., López-Guerra, M., López-Guillermo, A., Puente, X. S., Blasco, M. A., Campo, E. & López-Otin, C. POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. *Nat. Genet.* **45**, 526–530 (2003).
- 10 Li, P., Hou, M., Lou, F., Björkholm, M. & Xu, D. Telomere dysfunction induced by chemotherapeutic agents and radiation in normal human cells. *Int. J. Biochem. Cell Biol.* **44**, 1531–1540 (2012).
- 11 Tang, T., Zhou, F. X., Lei, H., Yu, H. J., Xie, C. H., Zhou, Y. F. & Liu, S. Q. Increased expression of telomere-related proteins correlates with resistance to radiation in human laryngeal cancer cell lines. *Oncol. Rep.* **21**, 1505–1509 (2009).
- 12 Baumann, P. & Price, C. Pot1 and telomere maintenance. *FEBS Lett.* **584**, 3779–3784 (2010).
- 13 Denchi, E. L. & Lange, T. D. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* **448**, 1068–1071 (2007).
- 14 Kaburagi, M., Fukuda, M., Yamada, H., Miyakawa, T., Morikawa, R., Takasu, M., Kato, T. A. & Uesaka, M. Molecular dynamics simulation of telomere and proteins. *JPS. Conf. Proc.* **1**, 016018 (2014).
- 15 Ramos, R. M., Fernandes, L. F. & Moreira, I. S. Extending the applicability of the O-ring theory to protein–DNA complexes. *Comp. Biol. Chem.* **44**, 31–39 (2013).
- 16 Lei, M., Podell, E. R., Baumann, P. & Cech, T. R. DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature* **426**, 198–203 (2003).
- 17 Chatterjee, S., Gersten, B., Thakur, S. & Burin, A. Molecular dynamics simulations of a single stranded (ss) DNA. *Mol. Sim.* **33**, 573–576 (2007).
- 18 Jaiswal, A. & Lakshmi, P. T. V. Molecular inhibition of telomerase recruitment using designer peptides: an *in silico* approach. *J. Biol. Struct. Dyn.* **9**, 37–41 (2014).
- 19 Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. Amino acid–base interactions: a three dimensional analysis of protein–DNA interactions at an atomic level. *Nucl. Acids Res.* **29**, 2860–2874 (2001).
- 20 Lei, M., Podell, E. R. & Cech, T. R. Structure of human POT1 bound to telomeric single stranded DNA provides a model for chromosome end-protection. *Nat. Struct. Mol. Biol.* **11**, 1223–1229 (2004).
- 21 Nandakumar, J., Podell, E. R. & Cech, T. R. How telomeric protein POT1 avoids RNA to achieve specificity for single-stranded DNA. *Proc. Natl Acad. Sci. USA* **107**, 651–656 (2010).
- 22 Hess, B., Kutzner, C., Spoel, D. V. D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
- 23 Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics Model* **14**, 33–38 (1996).
- 24 Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. Improved side-chain torsion potentials for the Amber. ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
- 25 Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987).
- 26 Izaguirre, J. A., Reich, S. & Skeel, R. D. Longer time steps for molecular dynamics. *J. Chem. Phys.* **110**, 9853–9864 (1999).
- 27 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
- 28 Steinbach, P. J. & Brooks, B. R. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem.* **15**, 667–683 (1994).
- 29 Darden, T., Perera, L., Li, L. & Pedersen, L. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* **7**, 55–60 (1999).
- 30 Price, D. J. & Brooks, C. L. III A modified TIP3P water potential for simulation with Ewald summation. *J. Chem. Phys.* **121**, 10096–10103 (2004).
- 31 Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
- 32 Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
- 33 Parrinello, M. & Rahman, A. Crystal structure and pair potential: a molecular dynamics study. *Phys. Rev. Lett.* **45**, 1196 (1980).
- 34 Mills, J. E. J. & Dean, P. M. Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. *J. Comput. Aided. Mol. Des.* **10**, 607–622 (1996).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>