

# SCIENTIFIC REPORTS



OPEN

## Structural, functional, and stability change predictions in human telomerase upon specific point mutations

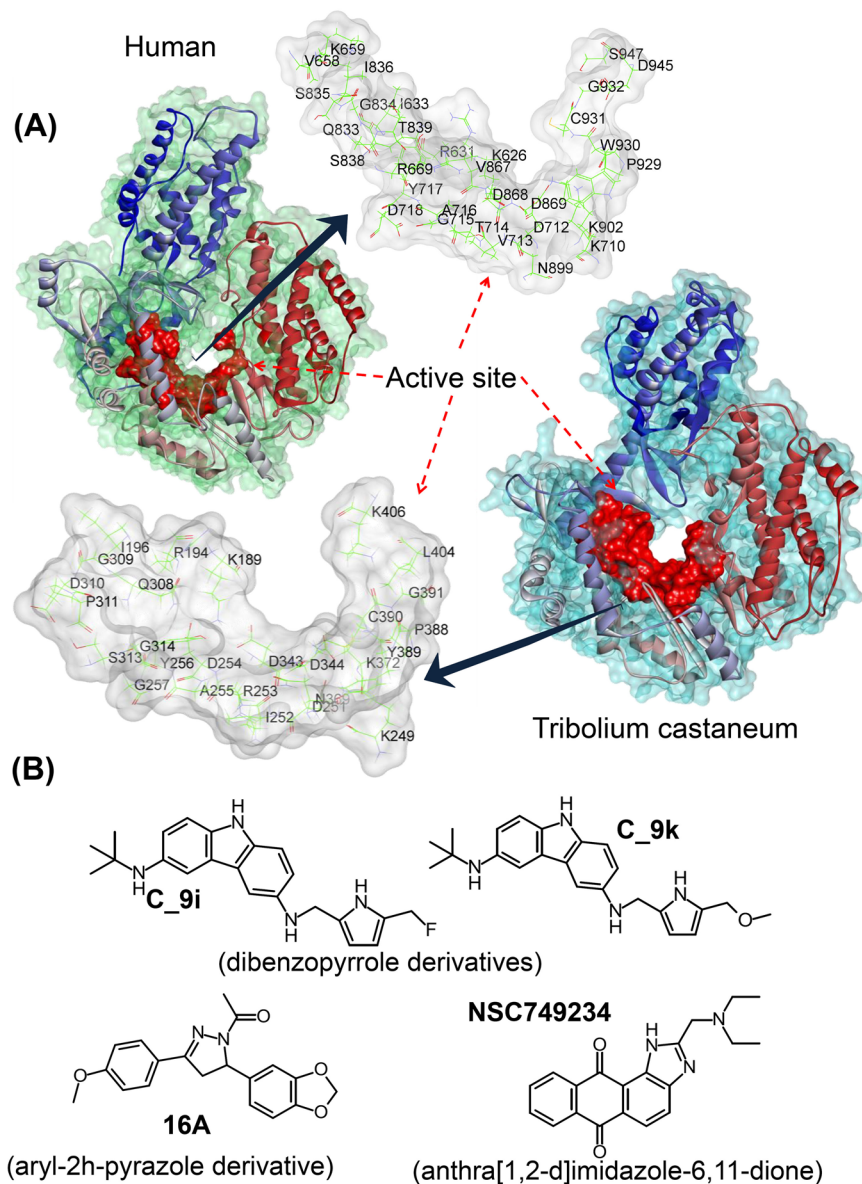
U. Kalathiya<sup>1,2</sup>, M. Padariya<sup>1</sup> & M. Baginski<sup>1</sup>

Overexpression of telomerase is one of the hallmarks of human cancer. Telomerase is important for maintaining the integrity of the ends of chromosomes, which are called telomeres. A growing number of human disease syndromes are associated with organ failure caused by mutations in telomerase (hTERT or hTR). Mutations in telomerase lead to telomere shortening by decreasing the stability of the telomerase complex, reducing its accumulation, or directly affecting its enzymatic activity. In this work, potential human telomerase mutations were identified by a systematic computational approach. Moreover, molecular docking methods were used to predict the effects of these mutations on the affinity of certain ligands (C\_9i, C\_9k, 16A, and NSC749234). The C\_9k inhibitor had the best binding affinity for wild-type (WT) telomerase. Moreover, C\_9i and C\_9k had improved interactions with human telomerase in most of the mutant models. The R631 and Y717 residues of WT telomerase formed interactions with all studied ligands and these interactions were also commonly found in most of the mutant models. Residues forming stable interactions with ligands in molecular dynamics (MD) were traced, and the MD simulations showed that the C\_9k ligand formed different conformations with WT telomerase than the C\_9i ligand.

Telomerase, a ribonucleoprotein (RNP), acts as a reverse transcriptase that synthesises telomeric DNA repeats at the ends of chromosomes<sup>1</sup>. Telomerase is active in the early stages of life, maintaining the telomere length, and it becomes inactive in most somatic cells during adulthood<sup>2,3</sup>. The ability of a telomere to provide genomic stability decreases over time, owing to both the natural loss of telomeric structure with each cell division (the end replication problem) and the loss of telomerase activity. In addition, this process leads to ageing<sup>4,5</sup>. However, in cancer cells, telomerase becomes reactivated, and it continuously maintains the short length of telomeres in rapidly dividing cells, leading to their immortality<sup>6,7</sup>. Most cancers have adapted mechanisms to protect lengths of telomeres, and this protection is achieved by telomerase activation in approximately 90% of human cancers. Therefore, the role of telomerase in cancer and ageing makes it an important target for cancer therapies and age-associated disorders<sup>8–11</sup>. Additionally, telomere-mediated disorders, such as dyskeratosis congenita, aplastic anaemia, and idiopathic pulmonary fibrosis, have telomerase mutations<sup>12–14</sup>.

Telomerases from evolutionarily distant organisms share a conserved structural organisation<sup>15</sup>, which contains a template-encoding RNA (telomerase RNA or hTR, TER) and a primary protein component (telomerase reverse transcriptase, TERT) with various functional domains (telomerase essential N-terminal domain, TEN; telomerase RNA-binding domain, TRBD; reverse transcriptase domain, RT; and C-terminal extension, CTE)<sup>16–19</sup>. Telomerase-specific domains at the N- and C-terminals of TERTs are not present in any of the viral reverse transcriptases<sup>20</sup>. The N-terminal domain is required for enzymatic function in the assembly of protein with its integral RNA component and in homodimerisation of the protein<sup>20–23</sup>. The C-terminal domain is required for telomerase-specific activity other than catalytic function and in the telomeric nucleotide addition process<sup>24,25</sup>. The fact that mutations of some key residues in telomerase, which are known to affect the conventional reverse transcriptase catalytic activity, also negatively influence telomerase activity strongly argues that the TERT domain is

<sup>1</sup>Department of Pharmaceutical Technology and Biochemistry, Faculty of Chemistry, Gdansk University of Technology, Gdansk, Poland. <sup>2</sup>Present address: International Centre for Cancer Vaccine Science; University of Gdansk, Gdansk, Poland. Correspondence and requests for materials should be addressed to U.K. (email: [umesh.kalathiya@pg.edu.pl](mailto:umesh.kalathiya@pg.edu.pl) or [umesh.kalathiya@ug.edu.pl](mailto:umesh.kalathiya@ug.edu.pl))



**Figure 1.** (A) The structure of the catalytic subunit (TERT) and active-site residues in *Tribolium castaneum* (PDB: 3DU6)<sup>11</sup> and the human<sup>19</sup> telomerase model. (B) The structure of ligand molecules, namely, C<sub>9i</sub><sup>33</sup>, C<sub>9k</sub><sup>33</sup>, 16A<sup>34</sup>, and NSC749234<sup>36</sup>, which were selected to study with human telomerase.

the catalytic domain of the enzyme complex<sup>25–28</sup>. Inherited mutations in both human TERC (the template region of an integral RNA component) and human TERT part of protein lead to rare bone marrow failure syndromes, autosomal dominant dyskeratosis congenita, and acquired aplastic anaemia<sup>28–31</sup>. Many experiments have shown the effects of specific amino acid substitutions or deletions on the enzymatic activity of telomerase and have provided useful knowledge about key residues. In addition, an analysis of such single-residue substitutions and their functional consequences at the molecular level using a complete structural model of telomerase<sup>19</sup> will be a useful complement to such structural data.

The key active-site residues in the TERT catalytic subunit of *Tribolium castaneum* telomerase have been published by Gillis *et al.* in<sup>11</sup> (PDB: 3DU6; Fig. 1A). Steczkiewicz *et al.*<sup>19</sup> provided the first structural model of human telomerase, using computational methods derived from the X-ray structure of the full-length *T. castaneum* telomerase (PDB: 3DU6; apo form)<sup>11</sup> and the enzyme in complex with a RNA:DNA hairpin (PDB: 3KYL)<sup>32</sup>. In the present study, we used the TERT catalytic subunit of the human telomerase model<sup>19</sup> and defined/predicted the active-site residues based on *T. castaneum* telomerase structure<sup>11,33–35</sup> (Fig. 1A). Moreover, residues of telomerase involved in biological functions were also covered in active site<sup>19</sup>. By superimposing the human telomerase structure over *T. castaneum* structure, we identified the active-site residues at the same structural location in both structures. Active site residues of *T. castaneum*/human telomerase were as follows: K189/K626, R194/R631, I196/I633, K249/K710, D251/D712, I252/V713, R253/T714, D254/G715, A255/A716, Y256/Y717, G257/D718, Q308/Q833, G309/G834, D310/S835, P311/I836, S313/S838, G314/T839, D343/D868, D344/D869, N369/N899, K372/

K902, P388/P929, Y389/W930, C390/C931, G391/G932, L404/D945, and K406/S947<sup>11,19</sup>. Moreover, residues V658, K659, R669, V867, R972, and K973 of human model were also selected in the current study because previous studies have shown the importance of these residues in different biological functions of human telomerase (Fig. 1A)<sup>19</sup>.

To the best of our knowledge, there have been limited theoretical studies on telomerase that analyse the effect of the mutations at the molecular level. Here, we studied different potential mutations of telomerase enzyme and their effects when binding to various ligands (known as potential inhibitors) by means of molecular docking and molecular dynamics (MD) simulations. The present study provided valuable insights into the nature of potential structural changes as a result of mutations, especially at the functionally important regions or residues of the active site.

Four recently designed or identified telomerase inhibitors, namely, C<sub>9</sub>i<sup>33</sup>, C<sub>9</sub>k<sup>33</sup>, 16A<sup>34</sup>, and NSC749234<sup>36</sup>, were selected to study with the mutated human telomerase model (Fig. 1B). The C<sub>9</sub>i and C<sub>9</sub>k compounds are derivatives of dibenzopyrrole, and *in silico* analysis of this new chemical scaffold has shown potential telomerase-binding properties<sup>33</sup>. Compound 16A<sup>34</sup> is from a series of novel aryl-2h-pyrazole derivatives containing an oxygen-bearing heterocyclic group. Compound 16A has potent inhibition activity for telomerase and good activity against human melanoma cell B16-F10<sup>34</sup>. The NSC749234 compound<sup>36</sup> is a derivative of anthra[1,2-d]imidazole-6,11-dione and has been evaluated for telomerase inhibition, hTERT expression and suppression of cancer cell growth *in vitro*<sup>36</sup>.

To study potential mutations of residues in active-site of human telomerase, two different approaches were followed. For the first approach, the Blast tool implemented in the UniProtKB/Swiss-Prot<sup>37</sup> database was used to search natural mutants in proteins (having similar sequences) other than telomerase. This approach takes advantage of the fact that evolution has been performing its own massively parallel mutagenesis and selection experiments over time. Thus, these natural sequence variations found in other proteins can be more likely to occur also in telomerase active site. In the second approach, the mutant (MT) series was compared against wild-type (WT) using the Molecular Operating Environment (MOE; Chemical Computing Group Inc.), and for a particular residue all possible point mutations were searched to determine mutational consequences on telomerase<sup>38</sup>. Mutations that have shown similar behaviour with all four ligands as well as those obtained the highest and lowest energies compared to WT telomerase were studied by molecular docking to reveal potential differences in ligand-binding properties. Furthermore, all-atomic-level studies were performed by MD simulations to better understand the protein dynamics and the effects of the mutations on protein structure.

## Materials and Methods

**Structural model of human telomerase.** Steczkiewicz *et al.*<sup>19</sup> employed a theoretical modelling approach to generate the entire 3D structure of the human TERT, TEN, and TRBD bound to a DNA substrate and its RNA template. They used an advanced meta-profile comparison method, Meta-BASIC<sup>39</sup>, to map the human telomerase protein sequence onto sequences of the determined structures from *Tribolium* and *Tetrahymena*. The mappings obtained were confirmed by using a variety of fold recognition methods. Steczkiewicz *et al.*<sup>19</sup> built three-dimensional models separately for TEN and the other components of the human telomerase complex including a hybrid RNA: DNA double helix formed between the RNA template and the single-stranded telomeric DNA substrate and assembled them by using protein-protein docking, guided by relevant experimental data (Fig. S1A). In this study, we used the TERT catalytic subunit of the human telomerase model provided by Steczkiewicz *et al.*<sup>19</sup>.

**Selection of natural variants.** Short sequences (containing active-site residues) of human telomerase were submitted to the Blast tool implemented in UniProtKB/Swiss-Prot<sup>37</sup> to identify variations in similar sequences (natural mutants). Residues that were predicted as natural mutants in protein other than telomerase and mutations occurring with unusual frequency in different organisms were selected for the present study. From the screening results, four residues, namely, R631, Y717, D868, and D869, were considered because they have shown frequent mutations or variants in other organisms (Fig. S2). In total, seven natural mutant models, namely, R631Q, R631P, Y717G, D868N, D868K, D869R, and D869S, were generated. Model structures of human telomerase with novel natural mutations were built with the help of the ‘Built Mutant’ protocol incorporated in the modelling environment of the Discovery Studio Client v18.1 programme [Dassault Systemes, BIOVIA Corp., San Diego, CA, USA]. The ‘Built Mutant’ protocol uses the Modeler programme (version: 9v8)<sup>40</sup> to mutate residues to specified types and optimises the conformation of both the mutated residues and neighbouring residues. Side chains of the residues in the models were further refined to correct their conformation and orientation using the ‘Chi Rotor’ modelling programme<sup>41</sup>.

The generated natural mutant models and ligand structures were energy minimised with default parameters of the ‘smart minimiser’ algorithm applying CHARMM forcefield<sup>42</sup> in the Discovery Studio Client v18.1 programme. Following energy minimisation, flexible molecular docking was performed using CDOCKER, a CHARMM-based docking tool<sup>43</sup> from Discovery Studio Client v18.1, to analyse ligand affinity towards different mutated human telomerase models. In CDOCKER, 100 random ligand conformations were generated in the active site of the telomerase structure. The temperature was set to 700 K for 2000 steps and cooled to 300 K for 5000 steps. The grid extension was set to 8 Å, and 10 ligand binding poses were ranked according to their CDOCKER energies.

**MOE point mutations.** The ‘Sequence Design’ protocol of the protein design module incorporated in MOE, which calculates frequency and probability of amino acids at residue mutation sites (mutation expression), was used in the present study, and a ligand affinity score was obtained when ligand binds to WT and mutated telomerase (point mutation). The goal of this method is to modulate physical protein properties such as stability or affinity. Sequence Design is meant to handle situations where the “combinatorial explosion” due to the number

of residues and mutations reaches a point where producing all mutants is not feasible due to sterical hindrances. In addition, for this method the wild type mutation is always allowed to compete since in some cases, if no better mutant is possible, the best mutant will be simply the wild type<sup>44</sup>. This approach was followed for four different telomerase-ligand complexes, in which all 33 functional residues were mutated by 20 amino acids [A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, and V]. In each point mutation, the rotamer explorer RMSD limit was set to 0.25 Å, energy window to 10 kcal/mol and residues farther than 4.5 Å were kept fixed. LowMode MD, the search method in MOE, was used with default parameters for a thorough conformational search because it is a recent and particularly relevant addition to the MOE toolbox for a wide set of structures from small molecules to peptides, macrocycles, and protein loops. The CHARMM27 force field was used, and the maximum number of search iterations was 10000. Mutation analysis in MOE showed that mutation of 16 residues from 33 functional residues affected ligand affinity. Furthermore, considering these results, mutations of these 16 residues were further studied by the CDOCKER docking programme in Discovery Studio Client v18.1.

**Docking of WT human telomerase.** Protein-ligand binding affinity was computed using the CDOCKER docking programme for WT human telomerase when in complex with four different ligands, using the same parameters used for the docking of natural mutants. Furthermore, rigid and induced fit (flexible) dockings were performed using the MOE software package<sup>44</sup>. In MOE, all four complexes were subjected to energy minimisation using the CHARMM27 forcefield, and docking was performed with the ‘Rigid Receptor’ and ‘Induced Fit’ docking protocols. One thousand ligand conformations were generated, and the docking was performed using the ‘Triangle Matcher’ placement method, which is the most efficient method for well-defined binding sites. All 1000 conformations per ligand were scored using the ‘London dG’ scoring function, submitted to a refinement step based on molecular mechanics and rescored with the ‘GBVI/WSA dG’ scoring function<sup>45,46</sup>. GBVI/WSA dG, a forcefield-based scoring function, determines the binding free energy (kcal/mol) of the ligand from a given pose<sup>45,46</sup>.

**Molecular dynamics.** MD simulations were performed for eleven different systems as follows: (i) WT human telomerase in apo-form; (ii) and (iii) WT human telomerase with ligand C\_9i and C\_9k (CDOCKER binding mode); (iv) and (v) WT human telomerase with ligand C\_9i and C\_9k (MOE rigid binding mode); (vi) and (vii) WT human telomerase with ligand C\_9i and C\_9k (MOE induced fit binding mode); (viii) and (ix) mutated Y717H human telomerase with C\_9i and C\_9k (CDOCKER binding mode); and (x) and (xi) mutated Y717R human telomerase with C\_9i and C\_9k (CDOCKER binding mode). All simulations were performed with the GROMACS 4.6.5 simulation suite<sup>47</sup> using the Gromos96 43a1 forcefield parameter set<sup>48</sup>. The PRODRG server was used for generating topologies of ligands that can be used in GROMACS<sup>49</sup>. The SPC water condition was selected, and the box was defined as a dodecahedron. Periodic boundary conditions were applied in all directions, and Na<sup>+</sup>Cl<sup>-</sup> counter ions (yielding a 150 mM NaCl solution) were added to neutralise the system. The system was energy minimised (50,000 steps of steepest descents) to relax any steric conflicts. Energy minimisation was then followed by NPT equilibration for 1000 ps (1 ns). The particle mesh Ewald method<sup>50</sup> was employed to account for the long-range electrostatic interactions, and the LINCS algorithm<sup>51</sup> was used to restrain bond lengths. To maintain constant temperature and pressure (300 K and 1 bar), the V-rescale thermostat<sup>52</sup> and Parrinello-Rahman barostat<sup>53</sup> were used, respectively. The production run was performed for 100 ns with leapfrog integrator<sup>54</sup>, and the coordinates were saved every 10 ps.

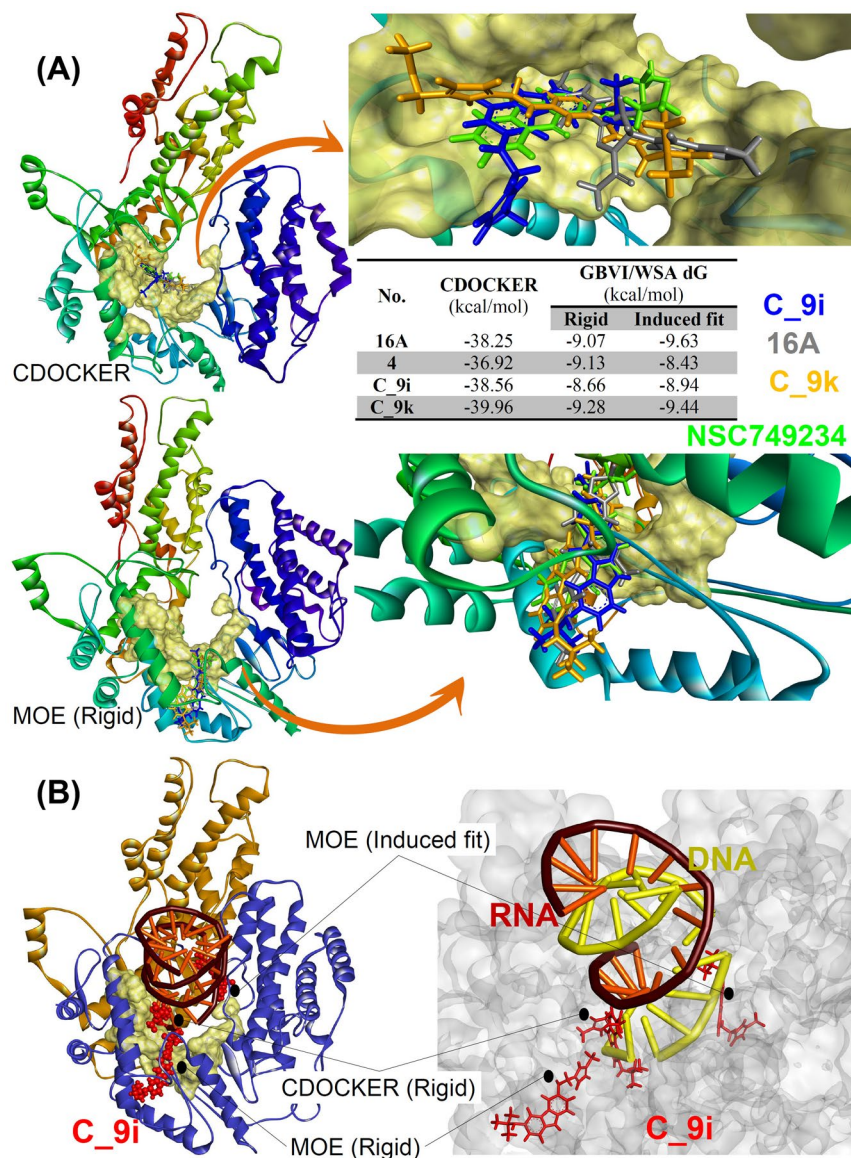
Ligand binding free energy was calculated using Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) method<sup>55</sup>. Binding energy of each snapshot was calculated for each complex using g\_mmpbsa tool of Gromacs<sup>56</sup>. The entropy contribution was not included in the binding energy. Principal component analysis (PCA) was carried out to obtain a mass-weighted covariance matrix of the protein atom displacement that is indicative of dominant and collective modes of the protein from the overall dynamics of the MD trajectory. The covariance matrix was diagonalized to extract a set of eigenvectors and eigenvalues that reflect concerted motion of the molecule<sup>57</sup>. To yield the eigenvalues and eigenvectors by calculating and diagonalizing the covariance matrix, Gromacs in-built tool g\_covar was used. To analyze and plot the eigenvectors, the g\_anaig tool was used<sup>58</sup>.

## Results and Discussion

**WT human telomerase-ligand binding.** Before analysing the effect of mutations on ligand binding at the structural level, we examined the binding affinity of four different inhibitors (C\_9i, C\_9k, 16A, and NSC749234) for the native (WT) protein. For the WT *human* telomerase model, both flexible and rigid docking were performed to enhance the sampling space of protein-ligand interactions. The C\_9k inhibitor had the best docking score in flexible (CDOCKER) and rigid (MOE) docking, and it had a good binding score in flexible docking of MOE relative to other compounds (Fig. 2).

The docking analysis of WT telomerase showed that Arg631 and Tyr717 residues formed potential interactions with all four ligands and that the Asp868 residue also formed interactions with all ligands, except compound 16A, in docking studies of CDOCKER (Fig. S3). In rigid docking of MOE, Arg865 was dominant and formed interactions with all ligands, except ligand C\_9k. Moreover, Arg669 formed interactions with two different ligands, namely, NSC749234 and C\_9i (Fig. S4). In MOE flexible or induced fit docking, the Val997, Ile1004, and Asn571 residues interacted most efficiently with the 16A and C\_9i ligands, respectively (Fig. S5). Superposition of all four compounds from rigid docking showed that all studied ligands occupied the same binding cleft in the human telomerase model structure (Fig. 2).

The similar binding mode of the ligand in different docking programs is not always the best binding affinity conformation. Here also, top ranked binding complexes from different docking programs have shown different binding mode of the compound C\_9i with human telomerase. However, in top ranked pose from all docking



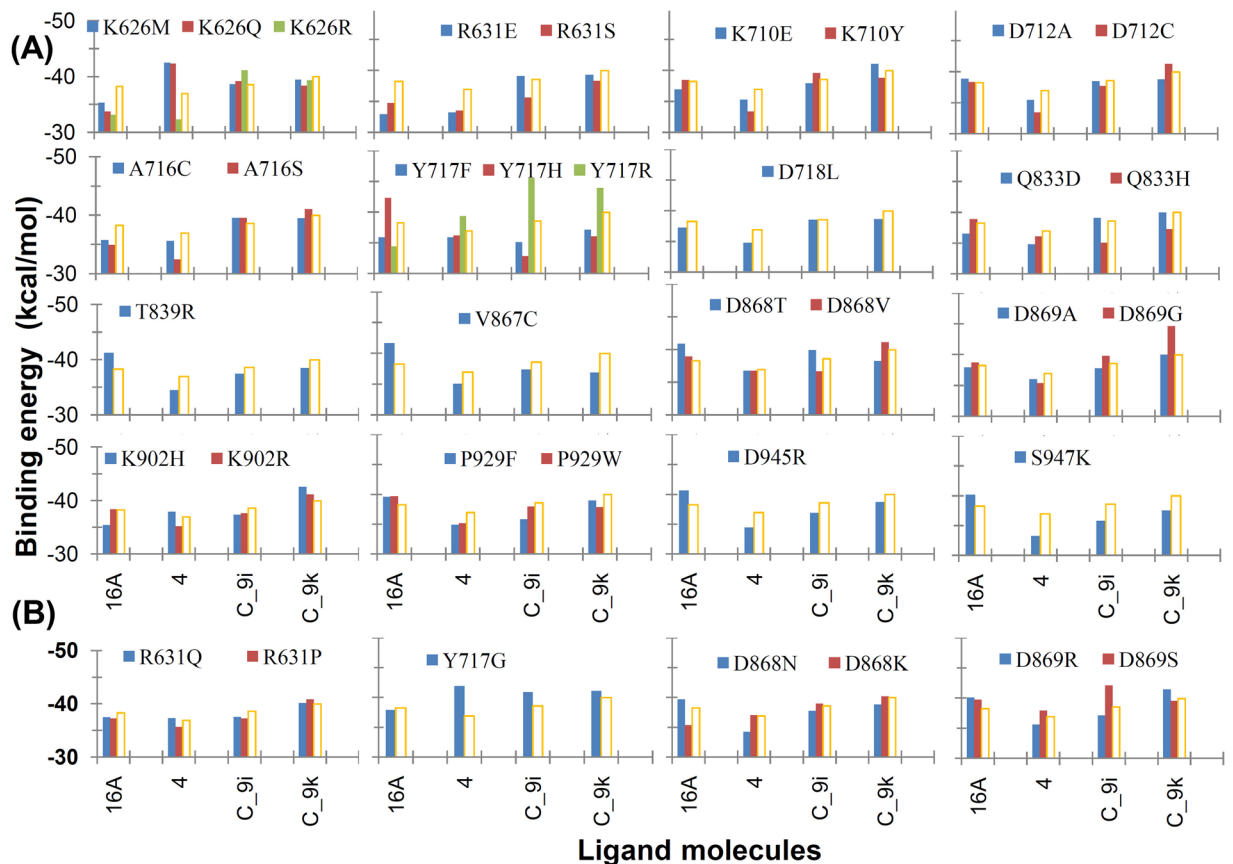
**Figure 2.** (A) Interaction energies and binding of the C<sub>9i</sub>, C<sub>9k</sub>, 16A, and NSC749234 inhibitors in the active site of WT human telomerase according to CDOCKER and MOE rigid (GBVI/WSA dG) docking. The active site is shown as a surface model, and the inhibitor is shown as a stick model. (B) Binding modes of compound C<sub>9i</sub> according to CDOCKER, MOE induced fit and rigid docking. In all, C<sub>9i</sub> binds close to DNA binding region.

programs, ligand C<sub>9i</sub> occupies the regions near to the DNA binding site (Fig. 2B). All compounds were also able to form same binding mode in different docking program though with not highest affinity. One of the example is shown for compound C<sub>9i</sub> from MOE Induced fit and Rigid docking (Fig. S1B).

**Mutated human telomerase interactions.** By applying the MOE point mutations approach, the ligand affinity score was computed for the different possible mutations. From the 33 active-site residues, mutations of 16 residues (Lys626, Arg631, Lys710, Asp712, Ala716, Tyr717, Asp718, Gln833, Thr839, V867, Asp868, Asp869, Lys902, Pro929, Asp945, and Ser947) resulted in changes in the ligand affinity (Fig. S6, Tables S1 and S2). On contrary, other studied mutations of the 14 other residues (Ile633, V658, K659, R669, Val713, Thr714, Ser835, Ile836, Ser838, Asn899, Trp930, Cys931, R972, and K973) had no effect on the ligand affinity score, showing that these residues were far from the ligand-binding region. Moreover, for three residues, namely Gly715, Gly834, and Gly932, mutations were not possible in the MOE point mutations approach.

Mutations of the K626, K710, T839, D868, and P929 residues in the presence of ligand 16A consistently obtained higher affinity scores than for the WT, whereas any variants of R631 had lowered affinity compared to the WT. Mutations of the A716, Y717, D718, and V867 residues did not show much effect on ligand affinity. Most of the mutations of D712, Q833, K902, D945, and S947 obtained almost similar ligand affinity scores with



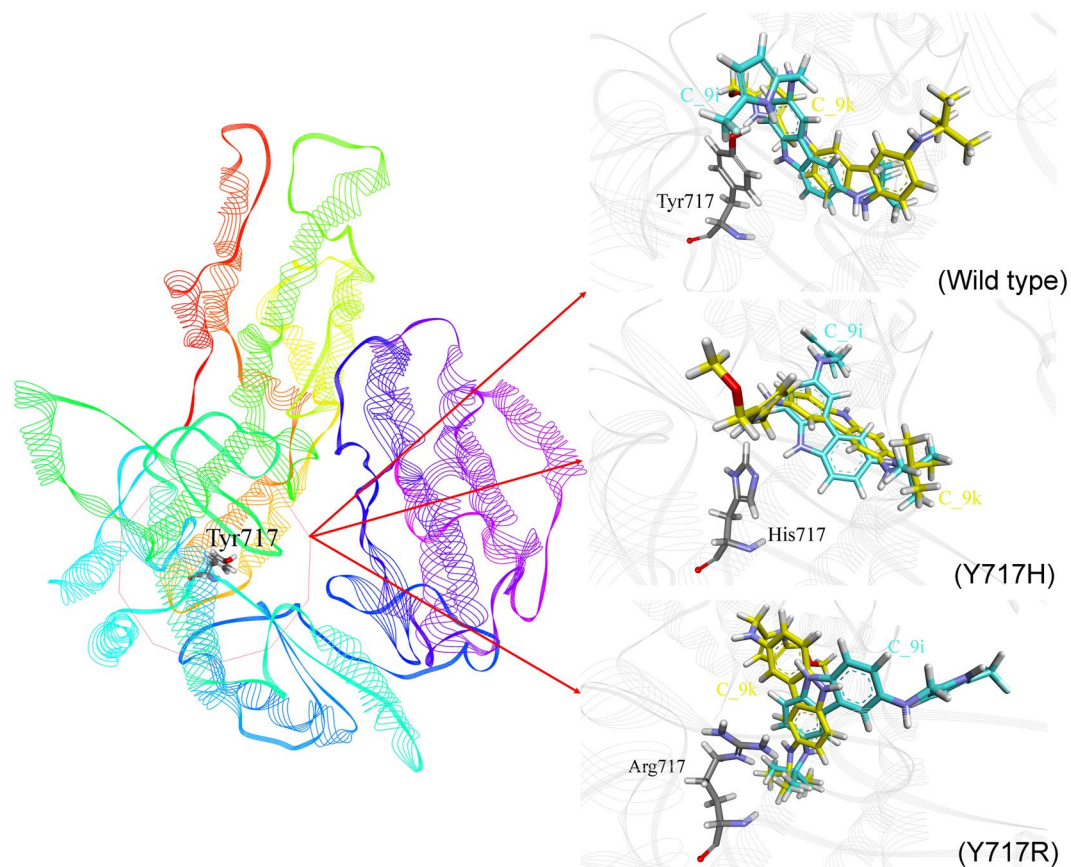


**Figure 3.** CDOCKER interaction (binding) energy for mutations obtained from (A) MOE point mutations approach and (B) naturally occurring mutations approach. The orange bar represents the interaction energy for WT telomerase.

a few exceptions having higher affinity than for WT (D712Q, Q833R, K902R, D945R, S947W, S947Y, S947R, and S947K). Among the point mutations of D869, the D869A and D869G mutations had considerably decreased affinity for ligand 16A, while the other mutations obtained almost similar scores (Tables S1 and S2). For the NSC749234 ligand, mutations of the R631, T839, and D868 residues always obtained higher affinity scores than for the WT, whereas any mutation of Y717, Q833, and D869 had lower affinity than for the WT. The A716, D718, and V867 variants did not show considerable effects in ligand affinity. Most of the mutants of D712, P929, D945, S947, K626, K710, and K902 had similar ligand affinity scores with some exceptions having higher/lower affinity (D712C, P929R, P929K, D945R, S947W, S947F, S947Y, S947R, S947K, K626Q, K626N, K710F, K710A, K710C, K710N, K710S, K710E, K710D, K710G, K902C, K902S, K902N, K902D, and K902H) than WT (Tables S1 and S2).

Predicted MOE point mutations of K626 and R631 showed reduced ligand affinity compared to WT in the presence of the C<sub>9i</sub> ligand, whereas D712, A716, Y717, D718, Q833, T839, V867, D868, and S947 mutants obtained ligand affinity scores almost the same as the wild type mutation. Most of the mutations of K710, K902, P929, and D945 obtained almost the same ligand affinity scores with few exceptions showing considerably higher affinity than WT (K710F, K710R, K902E, K902R, K902Q, K902E, K902H, P929W, P929F, and D945R). Among the D869 mutations, D869A and D869G had considerably lower affinity to the C<sub>9i</sub> ligand, while other D869 mutations obtained almost the same scores (Tables S1 and S2). In the case of the C<sub>9k</sub> ligand, mutation of T839 was only possible by T839R and obtained a higher affinity score than for WT, whereas mutations of K626, V867, and D868 did not show considerable effects on ligand affinity. Most mutations of R631, K710, D712, A716, Y717, D718, D869, K902, P929, and D945 residues obtained similar ligand affinity scores with a few exceptions having lower/higher affinity (R631S, D712A, A716S, A716N, Y717F, Y717N, Y717D, D869G, S947M, S947R, S947Q, S947E, K710F, K710Y, K710M, K710R, K710Q, D718L, K902R, K902Q, P929N, D945W, S947W, S947F, S947Y, and S947K) than WT (Tables S1 and S2).

Seven mutant models, namely, R631Q, R631P, Y717G, D868N, D868K, D869R, and D869S, obtained from the naturally occurring mutations approach and 29 mutations (of 16 residues) that obtained considerably higher or lower ligand affinity scores in the MOE mutation approach were validated by the CDOCKER docking programme in Discovery Studio Client v18.1. In CDOCKER docking of four studied ligands, protein-ligand interaction energy and interacting residues were identified for all variants (Figs 3 and 4, Tables S3 and S4). The docking score or interaction energy of the WT telomerase-16A complex was  $-38.25$  kcal/mol, whereas the docking score ranged from  $-32.91$  to  $-42.27$  kcal/mol for the mutant model. Similarly, the docking score was  $-36.92$  kcal/



**Figure 4.** Binding conformation of the C<sub>9i</sub> and C<sub>9k</sub> ligands with respect to the WT and mutated human telomerase models (Y717H and Y717R) according to CDocker docking.

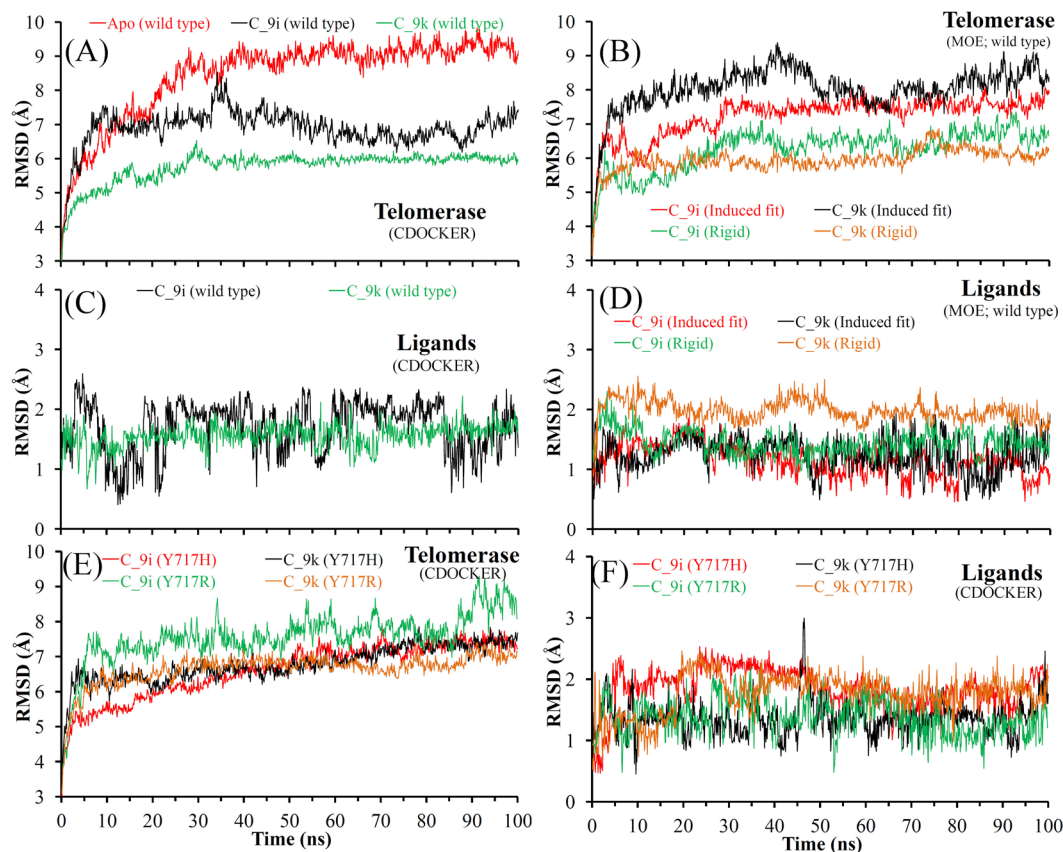
mol for the WT telomerase-NSC749234 complex, and it ranged from  $-32.37$  to  $-42.50$  kcal/mol for the mutant models. For the dibenzopyrrole derivatives (C<sub>9i</sub> and C<sub>9k</sub>), the docking score to the WT telomerase was  $-38.56$  and  $-39.96$  kcal/mol, respectively. For the mutant models, the docking score of C<sub>9i</sub> ranged from  $-32.83$  to  $-45.53$  kcal/mol, and the docking score of C<sub>9k</sub> ranged from  $-36.05$  to  $-44.63$  kcal/mol (Tables S3 and S4).

Comparing binding of the dibenzopyrrole derivatives (C<sub>9i</sub> and C<sub>9k</sub>) with the other two studied inhibitors (16A and NSC749234), the dibenzopyrrole derivatives had improved interactions with telomerase in most of the studied mutant models. Inhibitor 16A showed similar binding in the majority of studied natural mutants. Most of the mutations reduced binding of the NSC749234 inhibitor to telomerase or had similar binding as to the wild type one (Fig. 3). The Arg631 and Tyr717 residues of WT telomerase formed interactions with all four studied ligands and were also found frequently in most of the mutant models. The Asp868 residue in WT model formed interactions with all ligands, except 16A. Compound 16A formed interactions with Cys931 and Gly932, and compound C<sub>9i</sub> formed an interaction with Gln833 in the WT telomerase (Table S3).

Work by Drosopoulos and Prasad, showed that substitutions at Val867 lead to significant changes in overall enzyme activity and telomere repeat extension rate, but have little effect on polymerase processivity. And also indicated that Val867 plays a key role in nucleotide incorporation and suggested that this residue may provide important interactions with incoming substrate nucleotides<sup>59</sup>. In our analysis, mutations of V867 does not affect the ligand affinity of compound C<sub>9i</sub> and C<sub>9k</sub>.

Previous studies suggest that three conserved uncharged residues Y717, Q833, and V867 form a hydrophobic pocket adjacent to the catalytic aspartates and take part in nucleotide binding<sup>19,32</sup>. Residue mutation at V867 has been shown to alter human telomerase substrate specificity, and at Q833, mutations cause hypersensitivity to substrate analogs<sup>19,59,60</sup>. In our docking analysis, WT telomerase residue Tyr717 formed potential interactions with all four ligands. Moreover, mutations of all three residues Tyr717, Gln833, and Val867 resulted some changes in the ligand affinity.

Docking analysis of the WT and mutant models indicated the effect of mutation on the ligand binding efficiency. The overall analysis suggested that two novel mutations, namely, Y717H and Y717R, showed considerable effects on binding affinity. The Y717H mutant reduced the interaction energy, and the Y717R mutant improved the binding towards dibenzopyrrole derivatives (C<sub>9i</sub> and C<sub>9k</sub>) compared to other variants. Moreover, both C<sub>9i</sub> and C<sub>9k</sub> derivatives obtained conformations or binding modes similar to WT telomerase, whereas they changed their conformation in the Y717H and Y717R mutant models, which resulted in lower and higher binding affinity, respectively (Fig. 4). Considering the impact of these mutations on the binding affinity, mutated



**Figure 5.** RMSDs of (A) WT telomerase in apo-form and in complex with C<sub>9i</sub>/C<sub>9k</sub> from MD of CDOCKER binding mode, (B) WT telomerase in complex with C<sub>9i</sub>/C<sub>9k</sub> from MD of MOE induced fit or rigid binding mode, (C,D) ligand C<sub>9i</sub>, C<sub>9k</sub> from MD of CDOCKER and MOE induced fit or rigid binding mode, respectively, (E,F) mutated human telomerase (Y717H and Y717R) and ligand molecules from the complex (MD of CDOCKER binding mode). RMSDs were computed for all atoms, excluding hydrogens.

Y717H and Y717R human telomerase models with C<sub>9i</sub> and C<sub>9k</sub> were selected to investigate the alteration in structural-dynamics properties by the MD approach.

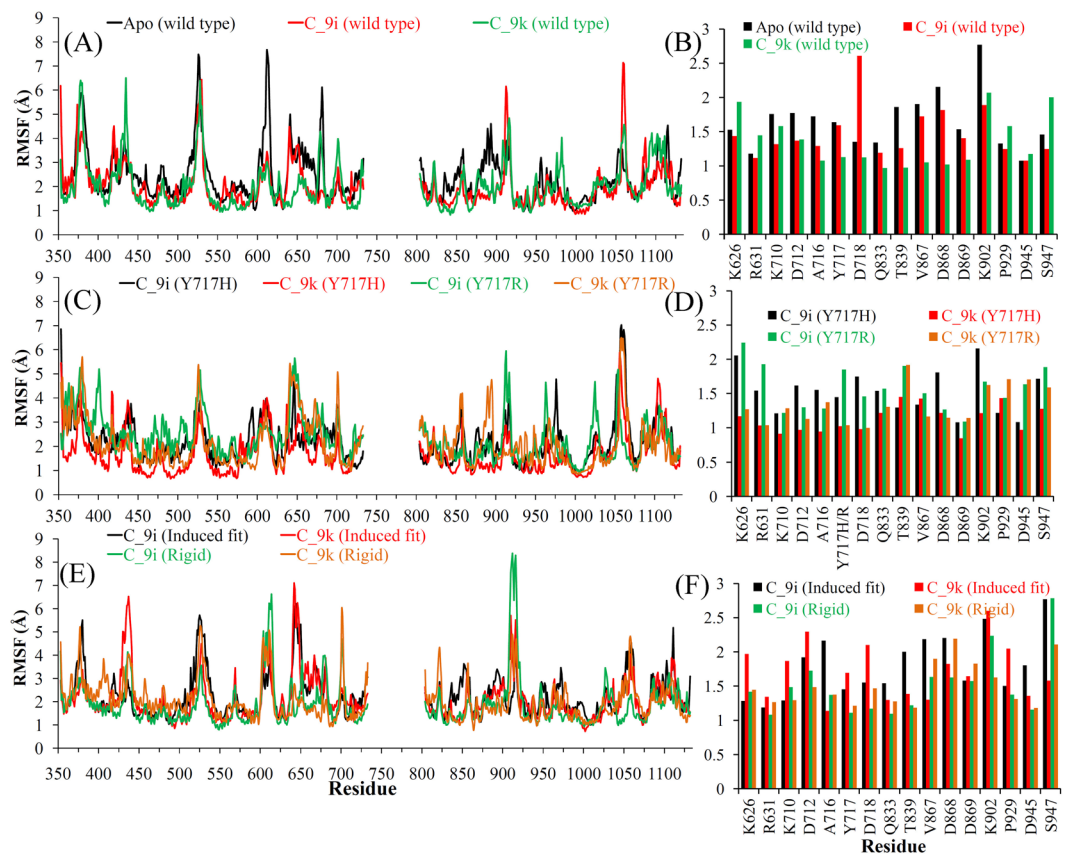
**MD simulations.** MD simulations were performed for WT telomerase in the apo-form and in complex with either C<sub>9i</sub> or C<sub>9k</sub> ligand. The top ranked binding affinity pose of compound C<sub>9i</sub> and C<sub>9k</sub> have obtained different binding mode in CDOCKER, MOE rigid and induced fit docking. Considering this, simulations were performed for each of the binding mode taken as the starting structure for the MD (Fig. S7). In addition, MD was performed on two mutated human telomerase models, namely, Y717H and Y717R, with the respective C<sub>9i</sub> and C<sub>9k</sub> ligands. For mutant models, the best docked conformation of both ligands into the active site of human telomerase obtained from CDOCKER was taken as the starting structure for the MD run.

The two most common measures of structural fluctuations, root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF), were analysed for all simulated systems. As shown in Fig. 5A, WT telomerase had high RMSD values in the apo-form compared to when in complex with the C<sub>9i</sub> or C<sub>9k</sub> ligand. The RMSD analysis of MD simulations starting from the CDOCKER and MOE rigid docking poses suggested more stable telomerase structure (with lower RMSDs) when in complex with the C<sub>9k</sub> ligand than with the C<sub>9i</sub> ligand. The reverse was observed for MOE induced fit docking poses, telomerase structure has obtained lower RMSDs when in complex with the C<sub>9i</sub> ligand than with the C<sub>9k</sub> ligand (Fig. 5). The RMSD of telomerase structure in complex with C<sub>9k</sub> was higher by approximately 1 Å in both mutant models compared to WT (Fig. 5A,E). In the case of systems with the C<sub>9i</sub> ligand, the Y717H mutation had similar RMSDs as for WT telomerase, whereas the Y717R mutation obtained higher RMSDs by 1 to 2 Å (Fig. 5A,E). The RMSDs of both ligands C<sub>9i</sub> and C<sub>9k</sub> showed similar behaviour (~2 Å) in all simulated systems (Fig. 5).

These findings suggested that structural fluctuations of WT human telomerase were found less pronounced in the presence of ligand compared to the apo-form, and similar behaviour was observed in 16 specific residues with only a few exceptions (for which mutations affected the ligand affinity; Figs 3 and 6). Overall the RMSF analysis of mutated models with the C<sub>9i</sub> and C<sub>9k</sub> ligands showed similar fluctuations in all four systems. However, the RMSFs of the 16 specific residues indicated that there were fewer structural fluctuations for the mutated models in complex with the C<sub>9k</sub> ligand compared to the C<sub>9i</sub> ligand (Fig. 6).

Hydrogen bond (H-bond) interactions between human telomerase and ligands (C<sub>9i</sub> and C<sub>9k</sub>) were analysed. The residues forming relatively stable contacts with the ligands were identified by taking into account





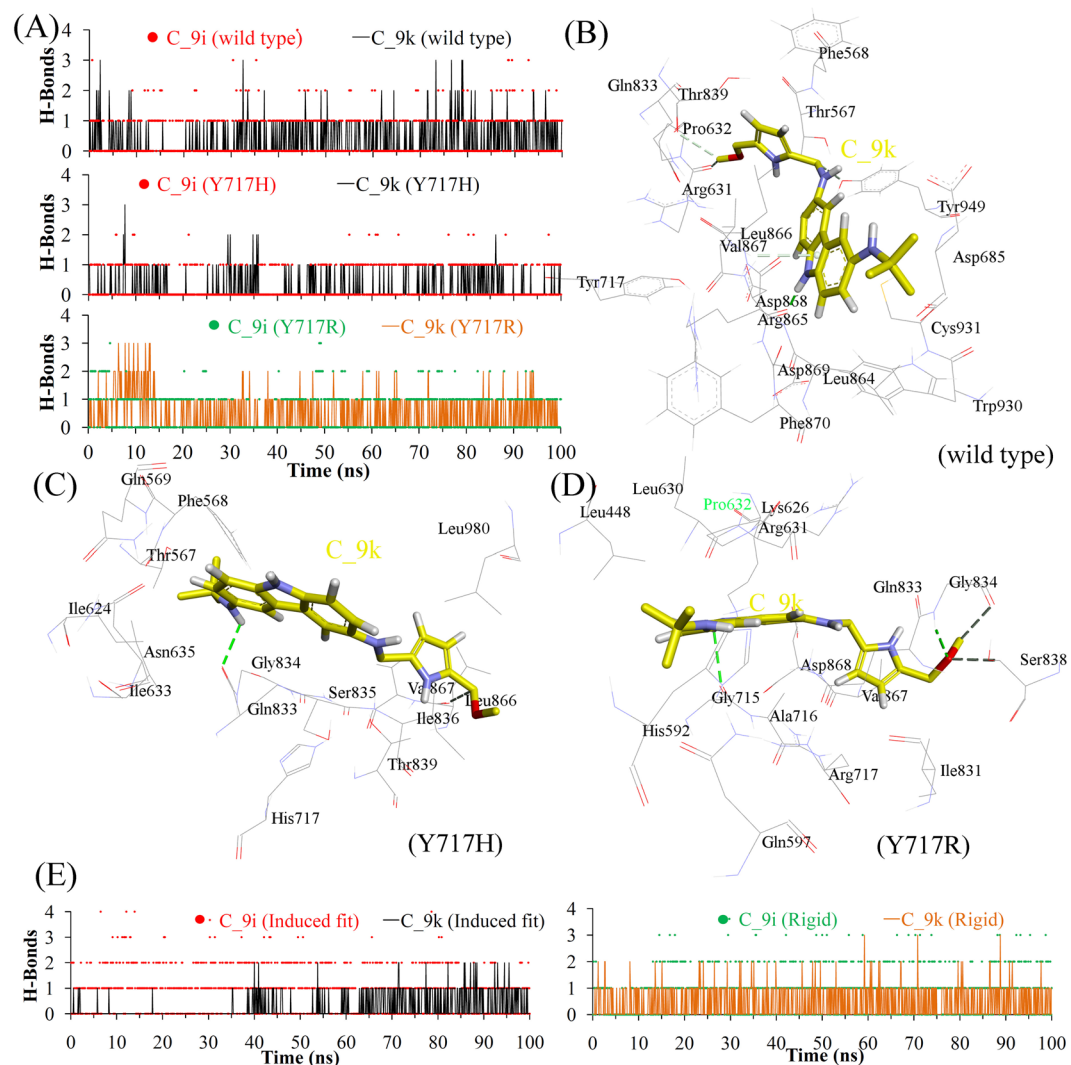
**Figure 6.** Residue-based RMSF of (A,B) WT telomerase in apo-form and in complex with C\_9i/C\_9k from MD of CDOCKER binding mode, (C,D) mutated human telomerase (Y717H and Y717R) in complex with C\_9i/C\_9k from MD of CDOCKER binding mode, (E,F) WT telomerase in complex with C\_9i/C\_9k from MD of MOE induced fit or rigid binding mode. RMSFs were computed for C $\alpha$  atoms.

the occupancy time (>1%). The residues forming H-bond interactions in MD simulations starting from the CDOCKER poses are given in Table S5. Asp868, a single residue, formed conserved interactions with the C\_9i ligand in all the simulated models. The H-bond interactions of the Thr567 and Gly834 ligand with the C\_9i ligand were common to the WT and Y717H mutated models. Moreover, the interaction between the Gln569 residue and the C\_9i ligand in WT telomerase was the same as the Y717R mutated model. For the C\_9k ligand, the Gln833 and Thr839 interacting residues were common to both Y717H and Y717R mutated models. In addition, the Asp868 residue formed an interaction with the C\_9k ligand in the WT and Y717R models (Table S5). Additionally, the residues forming H-bond interactions in MD simulations starting from the MOE rigid and induced fit docking poses are given in Table S6.

The analysis of H-bond statistics suggested that both the C\_9i and C\_9k ligands showed more interactions with the WT human telomerase model compared to the studied mutant models (Y717H and Y717R). The complex of ligand with WT telomerase model was stabilised by two to three H-bond interactions, whereas in the mutant models, the ligand mostly formed one to two H-bonds (Fig. 7). The residues of the human telomerase model interacting with the inhibitors (C\_9i or C\_9k) and the binding mode of inhibitors at the active site of WT and mutant models were traced and are presented in Fig. 7. Additionally, 2D interaction diagrams showing different interactions were further analysed for all simulated models and are shown in Fig. S8.

To trace the dynamics of ligand binding to the human telomerase model, simulated structures of WT and mutant models were analysed from the beginning and end of the MD simulations (Figs 8, S9 and S10). Both the C\_9i and C\_9k ligands showed similar binding modes with WT telomerase in docking, whereas the C\_9k ligand had different conformations in the MD simulation of CDOCKER binding mode (Figs 4 and S9). For the mutant models, conformation analysis of the C\_9i and C\_9k ligands from the MD simulations correlated well with the docking findings. As shown in Fig. 8, the C\_9k ligands had less conformational change in the Y717R model than in the Y717H model, which correlated with the observation that C\_9k had a higher binding affinity in the Y717R model compared to the Y717H model. In contrast, the C\_9i ligand showed different conformations in the Y717R model compared to those in the Y717H model in the MD simulation, and the findings were similar to the docking results (Figs 4 and S10).

To better estimate binding energy than just taking into account scores from docking programs, we performed MM-PBSA calculations taking into account MD trajectories. The analysis was performed three protein systems, namely WT and mutated telomerase (at Y717H or Y717R) in complex with ligand C\_9i. Comparative analysis can be performed using the obtained MM-PBSA analysis since the entropic contributions were not included in



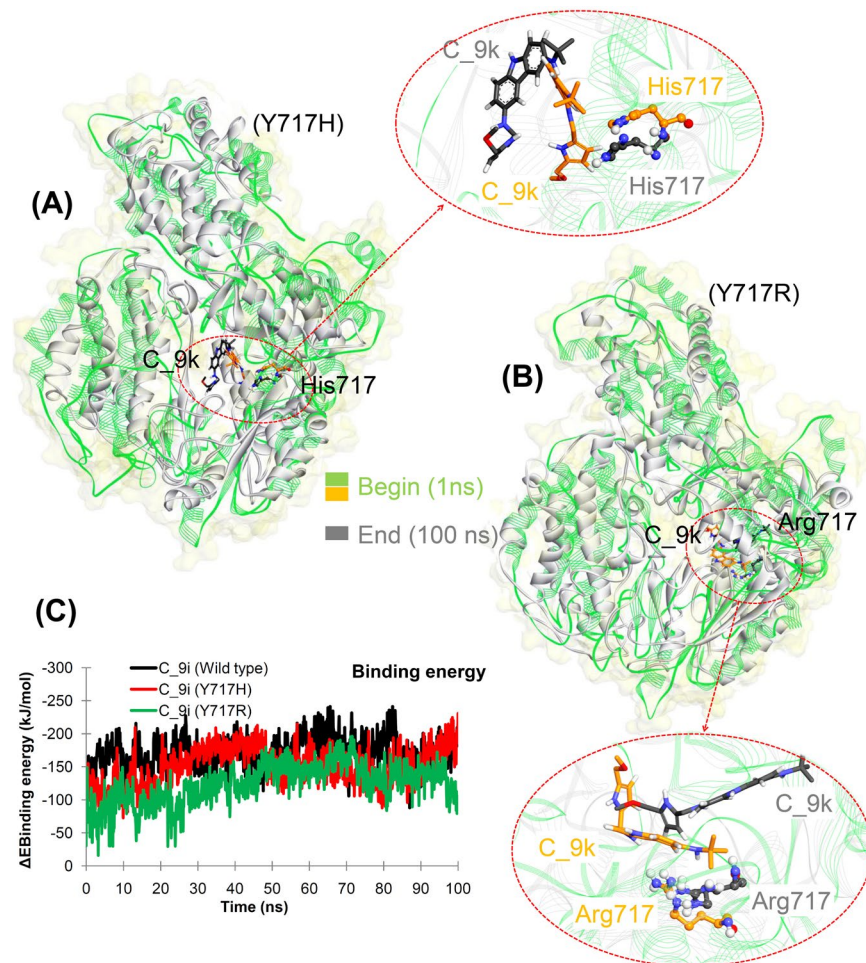
**Figure 7.** H-bond interactions of human telomerase with inhibitors (C<sub>9</sub>i and C<sub>9</sub>k). (A) Number of intermolecular H-bonds between the protein and ligand from MD of CDOCKER binding mode. (B–D) Binding pattern of the C<sub>9</sub>k ligand in the WT as well as Y717H and Y717R mutant models from MD of CDOCKER binding mode. (E) Number of intermolecular H-bonds between the protein and ligand from MD of MOE induced fit or rigid binding mode.

the binding energy. As seen in Fig. 8C, ligand C<sub>9</sub>i obtained the best binding energy when in complex with WT telomerase and results also indicated that C<sub>9</sub>i may behave similarly in Y717H mutated complex. Lower affinity of this ligand was recorded towards the second mutant Y717R. This observation correlates well with the energy contribution from the van der Waals, electrostatic, polar, and non-polar interactions (Fig. S11A). Furthermore, a quantitative assessment of the binding energy at individual residue had been studied as well (Fig. S11B). The residues with high contribution are placed in the binding region.

In order to analyse potential movement of residues in the binding region, the principal component analysis of MD trajectory was performed. This analysis was performed for WT human telomerase and mutant model Y717H with ligand C<sub>9</sub>i. Presentation of significant motions projected along first eigenvector shows that introduced mutation Y717H drive the ligand movement. Due to that ligand motions, the movement of protein fragments in telomerase active site was observed. In WT human telomerase bound with C<sub>9</sub>i, no such protein motions were observed relative to the mutant model (Figs S12 and S13).

## Conclusions

The human telomerase model structure was screened for the prediction of natural variants, and a MOE point mutations protocol was applied to obtain a large-scale structural analysis of the mutations. WT and variants of human telomerase were designed to study the binding mode of the ligands (C<sub>9</sub>i, C<sub>9</sub>k, 16A, and NSC749234) by molecular docking and MD simulations. The C<sub>9</sub>k inhibitor preferentially binds to WT human telomerase compared to the other compounds. Site-directed mutation (point mutations) obtained by MOE studies pointed to mutations of 16 residues that may influence ligand affinity towards human telomerase (K626, R631, K710,



**Figure 8.** Overlay of the binding mode of the C\_9k inhibitor at the beginning and at the end of the MD simulations of (CDOCKER binding mode), (A) Y717H and (B) Y717R mutant models. (C) Calculated MM-PBSA binding energy of compound C\_9i with telomerase (from MD of CDOCKER binding mode).

D712, A716, Y717, D718, Q833, T839, V867, D868, D869, K902, P929, D945, and S947). Compared to the 16A and NSC749234 ligands, the dibenzopyrrole derivatives (C\_9i and C\_9k) had improved their interactions with human telomerase in most of the studied mutant models. The R631 and Y717 residues of WT telomerase formed interactions with all four studied ligands and were also found frequently in most of the mutant models. In the MD studies, the D868 residue formed stable H-bond interactions with the C\_9i ligand in all simulated models, and the T567 and G834 residues interacting with C\_9i were the same in the WT and Y717H mutated models. Moreover, an interaction between Q569 and C\_9i was the same in the Y717R and WT telomerase models. The Q833 and T839 residues interacting with the C\_9k ligand were the same in both mutated models (Y717H and Y717R). D868 formed interactions with the C\_9k ligand in the WT and Y717R models. In molecular docking, both the C\_9i and C\_9k ligands showed similar binding modes with WT telomerase, whereas the C\_9k ligand showed different conformations in the MD simulations. The detailed large-scale structural analysis of mutations/variants in human telomerase significantly enhances the predictive power of existing approaches that can be utilised in the development of new/improved potent inhibitors against telomerase.

## References

1. Shay, J. W. & Wright, W. E. Role of telomeres and telomerase in cancer. *Semin. Cancer Biol.* **21**, 349–353 (2011).
2. Blackburn, E. H. Telomeres: no end in sight. *Cell* **77**, 621–623 (1994).
3. Greider, C. W. & Blackburn, E. H. Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell* **43**, 405–413 (1985).
4. Wright, W. & Hayflick, L. Nuclear control of cellular aging demonstrated by hybridization of anucleate and whole cultured normal human fibroblasts. *Exp. Cell Res.* **96**, 113–121 (1975).
5. Wright, W. E. & Shay, J. W. Cellular senescence as a tumor-protection mechanism: the essential role of counting. *Curr. Opin. Genet. Dev.* **11**, 98–103 (2001).
6. Bodnar, A. G. Extension of life-span by introduction of telomerase into normal human cells. *Science* **279**, 349–352 (1998).
7. Campisi, J., Kim, S., Lim, C. & Rubio, M. Cellular senescence, cancer and aging: the telomere connection. *Exp. Gerontol.* **36**, 1619–1637 (2001).
8. Shay, J. & Bacchetti, S. A survey of telomerase activity in human cancer. *Eur. J. Cancer* **33**, 787–791 (1997).
9. Hahn, W. C. *et al.* Inhibition of telomerase limits the growth of human cancer cells. *Nat. Med.* **5**, 1164–1170 (1999).

10. Zhang, X., Mar, V., Zhou, W., Harrington, L. & Robinson, M. O. Telomere shortening and apoptosis in telomerase-inhibited human tumor cells. *Genes Dev.* **13**, 2388–2399 (1999).
11. Gillis, A. J., Schuller, A. P. & Skordalakes, E. Structure of the *tribolium castaneum* telomerase catalytic subunit TERT. *Nature* **455**, 633–637 (2008).
12. Fogarty, P. F. *et al.* Late presentation of dyskeratosis congenita as apparently acquired aplastic anaemia due to mutations in telomerase RNA. *Lancet* **362**, 1628–1630 (2003).
13. Armanios, M. Y. *et al.* Telomerase mutations in families with idiopathic pulmonary fibrosis. *N. Eng. J. Med.* **356**, 1317–1326 (2007).
14. Vulliamy, T. J. Mutations in dyskeratosis congenita: their impact on telomere length and the diversity of clinical presentation. *Blood* **107**, 2680–2685 (2006).
15. Kelleher, C., Teixeira, M., Förstemann, K. & Lingner, J. Telomerase: biochemical considerations for enzyme and substrate. *Trends Biochem. Sci.* **27**, 572–579 (2002).
16. Autexier, C. & Lue, N. F. The Structure and function of telomerase reverse transcriptase. *Annu. Rev. Biochem.* **75**, 493–517 (2006).
17. Wyatt, H. D., West, S. C. & Beattie, T. L. In TERTpreting telomerase structure and function. *Nucleic Acids Res.* **38**, 5609–5622 (2010).
18. Sekaran, V. G., Soares, J. & Jarstfer, M. B. Structures of telomerase subunits provide functional insights. *Biochim. Biophys. Acta Proteins Proteom.* **1804**, 1190–1201 (2010).
19. Steczkiewicz, K. *et al.* Human telomerase model shows the role of the TEN domain in advancing the double helix for the next polymerization step. *Proc. Natl. Acad. Sci. USA* **108**, 9443–9448 (2011).
20. Xia, J., Peng, Y., Mian, I. S. & Lue, N. F. Identification of functionally important domains in the N-Terminal region of telomerase reverse transcriptase. *Mol. Cell Biol.* **20**, 5196–5207 (2000).
21. Ji, H., Platts, M. H., Dharamsi, L. M. & Friedman, K. L. Regulation of telomere length by an N-terminal region of the yeast telomerase reverse transcriptase. *Mol. Cell Biol.* **25**, 9103–9114 (2005).
22. Bosoy, D., Peng, Y., Mian, I. S. & Lue, N. F. Conserved N-terminal motifs of telomerase reverse transcriptase required for ribonucleoprotein assembly *in Vivo*. *J. Biol. Chem.* **278**, 3882–3890 (2002).
23. Moriarty, T. J., Huard, S., Dupuis, S. & Autexier, C. Functional multimerization of human telomerase requires an RNA interaction domain in the N terminus of the catalytic subunit. *Mol. Cell Biol.* **22**, 1253–1265 (2002).
24. Banik, S. S. *et al.* C-terminal regions of the human telomerase catalytic subunit essential for *In Vivo* enzyme activity. *Mol. Cell Biol.* **22**, 6234–6246 (2002).
25. Lue, N. F., Lin, Y. & Mian, I. S. A conserved telomerase motif within the catalytic domain of telomerase reverse transcriptase is specifically required for repeat addition processivity. *Mol. Cell Biol.* **23**, 8440–8449 (2003).
26. Lingner, J. Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276**, 561–567 (1997).
27. Bosoy, D. & Lue, N. F. Functional analysis of conserved residues in the putative “Finger” domain of telomerase reverse transcriptase. *J. Biol. Chem.* **276**, 46305–46312 (2001).
28. Xin, Z. *et al.* Functional characterization of natural telomerase mutations found in patients with hematologic disorders. *Blood* **109**, 524–532 (2007).
29. Dokal, I. & Vulliamy, T. Dyskeratosis congenita: its link to telomerase and aplastic anaemia. *Blood Rev.* **17**, 217–225 (2003).
30. Yamaguchi, H. *et al.* Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. *N. Eng. J. Med.* **352**, 1413–1424 (2005).
31. Vulliamy, T. J. *et al.* Mutations in the reverse transcriptase component of telomerase (TERT) in patients with bone marrow failure. *Blood Cells Mol. Dis.* **34**, 257–263 (2005).
32. Mitchell, M., Gillis, A., Futahashi, M., Fujiwara, H. & Skordalakes, E. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat. Struct. Mol. Biol.* **17**, 513–518 (2010).
33. Kalathiya, U., Padariya, M. & Baginski, M. Molecular modeling and evaluation of novel dibenzopyrrole derivatives as telomerase inhibitors and potential drug for cancer therapy. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **11**, 1196–1207 (2014).
34. Luo, Y. *et al.* Synthesis, biological evaluation, 3D-QSAR studies of novel aryl-2H-pyrazole derivatives as telomerase inhibitors. *Bioorg. Med. Chem. Lett.* **23**, 1091–1095 (2013).
35. Zhang, Y. *et al.* Design, synthesis and biological evaluation of heterocyclic azoles derivatives containing pyrazine moiety as potential telomerase inhibitors. *Bioorg. Med. Chem.* **20**, 6356–6365 (2012).
36. Chen, C. *et al.* Structure-based design, synthesis and evaluation of novel anthra[1,2-d]imidazole-6,11-dione derivatives as telomerase inhibitors and potential for cancer polypharmacology. *Eur. J. Med. Chem.* **60**, 29–41 (2013).
37. Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M. & Bairoch, A. UniProtKB/Swiss-Prot. *Plant Bioinform.* **406**, 89–112 (2007).
38. Long, W. F. & Labute, P. Calibrative approaches to protein solubility modeling of a mutant series using physicochemical descriptors. *J. Comput. Aided Mol. Des.* **24**, 907–916 (2010).
39. Ginalski, K., von Grothhuss, M., Grishin, N. V. & Rychlewski, L. Detecting distant homology with Meta-BASIC. *Nucleic Acids Res.* **32**, W576–W581 (2004).
40. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinform.* **54**, 5.6.1–5.6.37 (2016).
41. Spassov, V. Z., Yan, L. & Flook, P. K. The dominant role of side-chain backbone interactions in structural realization of amino acid code. ChiRotor: a side-chain prediction algorithm based on side-chain backbone interactions. *Protein Sci.* **16**, 494–506 (2007).
42. Brooks, B. R. *et al.* CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–614 (2009).
43. Wu, G., Robertson, D. H., Brooks, C. L. & Vieth, M. Detailed analysis of grid-based molecular docking: a case study of CDOCKER-a CHARMM-based MD docking algorithm. *J. Comput. Chem.* **24**, 1549–1562 (2003).
44. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Disc.* **3**, 935–949 (2004).
45. Wojciechowski, M. & Lesyng, B. Generalized born model: analysis, refinement, and applications to proteins. *J. Phys. Chem. B* **108**, 18368–18376 (2004).
46. Labute, P. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using london dispersion instead of atomic surface area. *J. Comput. Chem.* **29**, 1693–1698 (2008).
47. Hess, B., Kutzner, C., Spoel, D. V. & Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
48. Schuler, L. D., Daura, X. & Gunsteren, W. F. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comput. Chem.* **22**, 1205–1218 (2001).
49. Schüttelkopf, A. W. & Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr D Biol Crystallogr.* **60**, 1355–1363 (2004).
50. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
51. Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
52. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
53. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
54. Gunsteren, W. F. & Berendsen, H. J. A Leap-frog algorithm for stochastic dynamics. *Mol Simul.* **1**, 173–185 (1988).

55. Kollman, P. A. *et al.* Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33**, 889–897 (2000).
56. Kumari, R., Kumar, R. & Lynn, A. G. *mmpbsa* - a GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **54**, 1951–1962 (2014).
57. Amadei, A., Linssen, A. B. & Berendsen, H. J. Essential dynamics of proteins. *Proteins* **17**, 412–425 (1993).
58. Aalten, D., Findlay, J., Amadei, A. & Berendsen, H. Essential dynamics of the cellular retinol-binding protein evidence for ligand-induced conformational changes. *Protein Eng. Des. Sel.* **8**, 1129–1135 (1995).
59. Drosopoulos, W. C. & Prasad, V. R. The active site residue Valine 867 in human telomerase reverse transcriptase influences nucleotide incorporation and fidelity. *Nucleic Acids Res.* **35**, 1155–1168 (2007).
60. Smith, R. A., Anderson, D. J. & Preston, B. D. Hypersusceptibility to substrate analogs conferred by mutations in human immunodeficiency virus type 1 reverse transcriptase. *J. Virol.* **80**, 7169–7178 (2006).

## Acknowledgements

The hardware and software resources were provided by the PL-Grid Infrastructure (Poland), and this work was supported by grant no. 2014/13/B/NZ7/02207 from the National Science Centre (Krakow, Poland).

## Author Contributions

U.K. designed project, performed research, and prepared figures. M.P. partially analysed data and performed research. M.B. partially designed research and revised data. All authors reviewed the manuscript and approved for publication.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-45206-y>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019