WILEY PROTEINS

RESEARCH ARTICLE

Human telomerase protein: Understanding how the catalytic activity is suppressed under single substitutions of some conserved residues. A computational study

Fernando E. Herrera[1] | Silvano J. Sferco[1,2] (ID)

[1]Physics Department, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral. Ciudad Universitaria, Santa Fe, Argentina

[2]Instituto de Física del Litoral (IFIS Litoral, UNL-CONICET), Santa Fe, Argentina

Correspondence
S.J. Sferco, Instituto de Física del Litoral (IFIS Litoral, UNL-CONICET), Güemes 3450, 3000 Santa Fe, Argentina.
Email: silvano.sferco@santafe-conicet.gov.ar

Abstract

The reverse transcriptase domain in telomerase proteins contains the essential conserved residues to catalyze the addition of a single nucleotide to the ends of DNA strands of most eukaryotic cells. In human telomerase protein, mutations in the conserved residues K902, R631, K626, D712, D868, and D869 are known to suppress catalytic activity. To understand these results, a computational model was constructed to simulate a ternary complex consisting of a model of the protein reverse transcriptase domain, a DNA/RNA double helix, an incoming dNTP, and two Mg^{2+} ions. Three independent Molecular Dynamics Simulations were performed for the wild type and the mutated K902N, R631Q, D712A, D868A, and D869A complexes. Binding Free Energies and alanine-scanning studies were also performed. Using the two-metal-ion mechanism for the nucleotide addition, deviations from the wild type which stop the activity of the human protein, were identified in each mutated complex. The K902N and R631Q mutations might stop the catalytic activity preventing the exit of the pyrophosphate from the catalytic pocket. Additionally, evidence that the same mechanism probably applies to the K626A, R631A, and K902A mutations is presented. For D712A mutation, the pentavalent intermediate state might not form; therefore, the catalytic reaction might not even begin. For the D868A mutation, the O3′-hydroxyl might lose coordination with the Mg ion and the reaction might not either start. Finally, from the limited sampling carried out in this work, we did not obtain any evidence to identify the mechanism by which the D869A mutation cancels the activity of telomerase.

KEYWORDS

D712, D868, hTERT, K626, K902, K902N, R631, R631Q, telomerase

1 | INTRODUCTION

Telomerase is a ribonucleoprotein enzyme responsible for maintaining the length and integrity of telomeres in most eukaryotic cells. It catalyzes the addition of a new deoxynucleotide to the 3′ end of a DNA strand, using an RNA template. It is a reverse transcriptase enzyme, minimally formed by a protein subunit called Telomerase Reverse Transcriptase (TERT) and by a RNA subunit (TER) containing the template.[1] Telomerase activity in humans was observed in germline cells, but it was practically absent in normal somatic cells,[2] suggesting a possible role in aging.[3,4] On the other hand, telomerase activity was observed in many human cancer cells.[2] These findings raised the

interest in the telomerase enzyme, being considered a target for drugs against cancer.[5,6]

In TERT proteins, four general domains may be recognized: the N-terminal (TEN), the telomerase RNA-binding (TRBD), the reverse transcriptase (RT), and the C-terminal (CTE) domains (outlined in Supporting Information Figure S1). Although all domains are necessary for the catalytic activity of the enzyme, the RT domain is responsible for the reverse transcriptase activity itself. The RT domain includes several conserved motifs (1, 2, A, B′, C, D and E), similar to those already known for nontelomerase RT proteins.[1,7,8] Each motif is characterized by one or two evolutionarily conserved residues. In the case of human TERT (hTERT) protein, the conserved residues are: K626 (motif 1),

R631 (motif 2), D712 (motif A), Q833 and G834 (motif B′), D868 and D869 (motif C), K902 (motif D), and G932 (motif E),[7,8] also shown in Supporting Information Figure S1.

The experimental three-dimensional (3D) structure at atomic level for the RT of human TERT protein remains unknown. The only 3D structure experimentally known at this level for a TERT including the RT domain corresponds to the red flour beetle *Tribolium castaneum*. It was obtained using X-ray diffraction in two configurations: the protein alone (PDB ID codes 3DU5 and 3DU6, with resolutions of 3.25 Å and 2.71 Å, respectively)[9] and forming a complex with a DNA/RNA hybrid helix (PDB ID code 3KYL, with a resolution of 2.7 Å).[10] For the case of human telomerase, the Cryo-Electron Microscopy (EM) structure of the substrate-bound holoenzyme was very recently published, achieving a reconstruction with a resolution of 7.7 Å, when the 3D structure of the telomerase protein of *T. castaneum* was used to fit into its Cryo-EM density.[11]

Additional structural insights may be obtained from other nontelomerase RT proteins, in particular, for the known RT of HIV-1 in complex with double-stranded nucleic acids. Structural comparison of the RNA/DNA bonded TERT from *T. castaneum* and HIV-1 RT proteins showed similarity in the overall domain organization and nucleic acids binding.[9,10] Similarly to the HIV-1 case,[12] the 3D structure of the RT domain in telomerases is described in terms of a human right hand model, with subdomains called fingers, palm and thumb. Motifs 1 and 2 are in the fingers, motifs A, B′, C, D, and E are in the palm and the CTE domain is forming the thumb subdomain.[9] In both, HIV-1 and TERT, the triad of conserved aspartates (one from motif A, and the two consecutive ones from motif C), define what is called the *active site* of the protein, whereas *the nucleotide-binding pocket* is located at the interface of the fingers and palm subdomains.[9,10,12,13] Furthermore, this active site is quite similar to the already identified active site in DNA-dependent DNA polymerases, which is coordinated by two metal ions. Therefore, we consider that the two-metal-ion catalyzed mechanism proposed for the nucleotide addition reaction in polymerases,[14,15] is also valid for the TERT proteins as well.[1,16] The two-metal-ion mechanism is schematically presented in Figure 1 for hTERT. Two Mg^{2+} ions are supposed to be coordinated by the three conserved aspartates, by the phosphates of an incoming dNTP and by the 3′-hydroxyl of the DNA primer. The phosphoryl transfer reaction is assumed to include an intermediate pentavalent transition state on the P_α atom, followed by the formation of a new phosphodiester bond (with a new deoxynucleotide attached to the primer) and the release of inorganic pyrophosphate (PP_i).[14,15] In this work, we used two-metal-ion mechanism in an attempt to gain insight into the molecular details involved when some conserved residues of RT domain are singly mutated, causing the drop of telomerase activity.

Below, we briefly review the results published on the mutations on the RT conserved residues of hTERT that cancel its telomerase activity (less than 1% of wild type). The catalytic activity of hTERT drops when any of the three Asp conserved residues mutate to Ala.[17] Additionally, for the *in vitro* reconstituted hTERT, its activity was also monitored in the presence of point mutations. Results show that reconstituted telomerase activity in human telomerase is negligible when mutations in any of the three conserved Asp,[18] or mutations to Ala in any of the D712, D868, K626, R631 residues[19] are present.

Studying some genetic human diseases, other point mutations that also abolish the activity of human telomerase were identified. The K902N mutation was related to the autosomal dominant dyskeratosis congenita disease and to the haploinsufficiency of telomerase.[20] Moreover, point mutations of K902 residue to Ala, Asn, or Gln, provided no evidence of telomerase activity.[20] Likewise, the R631Q point mutation, which also abolishes telomerase activity, was related to several diseases: autosomal dominant dyskeratosis congenita,[21] myelodysplastic syndrome,[22] acute myeloid leukemia,[22] and idiopathic pulmonary fibrosis disease.[23] In summary, concerning the catalytic activity of hTERT, the point mutations of residues K902, R631, and K626 appear to be as crucial as the mutations in any of the aspartic acid residues.

From the theoretical point of view, a 3D theoretical model for the complete hTERT protein was published several years ago.[24] The aim of that work was to understand the processive mechanism of human telomerase. Unlike that work, this paper focuses on the RT domain of hTERT and on single amino acid substitution of its conserved residues. Therefore, we decided to start from scratch and to develop a model by ourselves. A comparison will be made in section 3. The processive mechanisms of telomerase will not be addressed in this work.

The original goal of this paper was to understand why mutations K902N and R631Q (involved in several genetic diseases) abolish the activity of human telomerase. Motivated by our findings, we extended the analysis to include the mutations of the conserved triad of aspartic acids as well. With these purposes, a theoretical model for the RT domain of hTERT was obtained applying standard homology methods. It was used to build ternary chimeric complexes that were further refined by Molecular Dynamics (MD) simulations. Three systems were studied thoroughly: (i) wild type (WT) hTERT complex, (ii) K902N point mutated complex and, (iii) R631Q point mutated complex. Two configurations were analyzed: Configuration I (C-I) just before the catalytic reaction took place (Figure 1A), and Configuration II (C-II), immediately after the reaction occurred, with the release of PP_i (Figure 1B). The MD trajectories were further analyzed calculating the Binding Free Energies, using the Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) method. Additionally, the alanine-scanning method was used on the K902, R631, and K626 residues. Finally, we considered complexes that have mutations to Ala in the conserved Asp residues (D712A, D868A, and D869A) studied in C-I. Additionally, these complexes were also used to perform new short MD simulations starting from a distorted configuration in order to check the stability of C-I. Comparing the results from WT and mutated complexes, we are able to gain certain understanding of the atomic mechanisms by which mutations on the K902, R631, K626, D712, or D868 conserved residues abolish telomerase activity.

2 | MATERIALS AND METHODS

2.1 | Homology modeling

In order to model the 331 residues of the RT domain of hTERT (UniProt code: O14746; RT domain: 605-935), the following procedure was followed: (1) Confirmation of conserved motifs and residues:
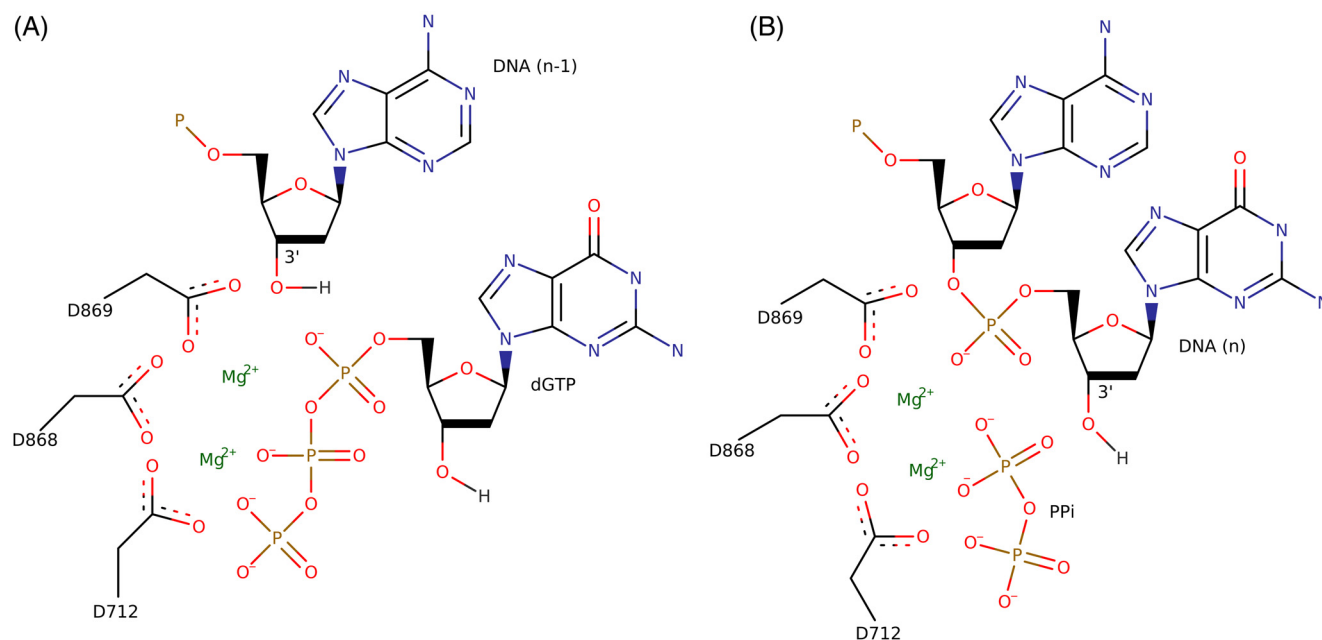
FIGURE 1 Schematic representation of two configurations for the WT complex system studied in this work. A, Configuration I (C-I), just before the catalytic reaction occurs, with RT-hTERT protein (represented by the lateral chains of the conserved aspartates), the DNA primer (represented by the DNA[n-1], the last nucleotide of the primer), the incoming dNTP (represented by the dGTP) and two Mg^{2+} ions, all placed at their catalytic site. B, Configuration II (C-II), just after the catalytic reaction takes place, showing RT-hTERT protein (represented as in A), the DNA primer (represented by two deoxynucleotide, with the last nucleotide, DNA[n], recently included), inorganic pyrophosphate (PPi), and two Mg^{2+} ions. For clarity, the RNA template and water solvent molecules are not shown [Color figure can be viewed at wileyonlinelibrary.com]

using the M-coffee meta server,[25] a multi sequence alignment among RT of telomerases of several species (known at the protein level) and the RT of HIV-1 protein was made and subsequently inspected manually (Supporting Information Figure S2). The results of our multisequence alignment agree with previous studies in the literature.[8,9] In particular, the same conserved residues mentioned in the Introduction were found. (2) Search of the homologous template: It was performed using the HHserver.[26] The RT domain of TERT of *T. castaneum* (PDB codes: 3DU5, 3DU6, and 3KYL, the same protein in all cases)[9,10] was identified as a possible template, with 24% of sequence identity with the RT hTERT target. (3) Protein 3D modeling: due to the low percentage of sequence identity obtained, the fold recognition modeling package I-TASSER was used.[27,28] A search for templates using different threading methods is made by this method, obtaining the RT of *T. castaneum* as template (the same PDB codes mentioned above) in all cases. Using ten different alignments with the same template, I-TASSER obtains five 3D models ranked by the C-score values.[27,28] These models were further refined using the High Resolution Protein Structure Refinement package implemented in the ModRefiner server.[29] Finally, the models were further screened by physicochemical and statistical approaches, as well as for their ability to bind a double helix, like those found in *T. castaneum* (PDB code 3KYL)[10] and in the RT of HIV-1 (PDB code 1RTD).[13] Supporting Information Figure S3 (top panel) shows our final 3D homology model obtained for the RT of hTERT, validated in color code, using the QMEAN quality estimation method.[30] The zoom in Supporting Information Figure S3 shows the lateral chains of the conserved residues, all of them forming the nucleotide-binding pocket. The zone nearest the three lateral chains of the aspartates is the catalytic site of

telomerases. The Insertion Finger Domain (IFD in Supporting Information Figure S1) is poorly modeled since the IFD of the template is notably shorter than that in humans. However, and similarly to what was found in *T. castaneum*, in our model the human IFD folds outside the nucleotide-binding pocket. Moreover, the catalytic site and the nucleotide-binding pocket are acceptably modeled according to QMEAN (Supporting Information Figure S3, top panel). Finally, this homology model of RT hTERT was used to build the complexes with DNA/RNA studied in this work.

2.2 | Complex formation: (RT hTERT protein + DNA/RNA + dGTP +2 Mg^{2+} ions)

The RT domains of the *T. castaneum* (PDB code 3KYL)[10] and HIV-1 (PDB code 1RTD, p66 subunit),[13] show quite a good structural superposition, in particular, for the catalytic pocket region (not shown). Moreover, the nucleotide double helices, even if they are of different kinds (hybrid DNA/RNA in 3KYL, but a double helix dideoxy-DNA/DNA in 1RTD) also appeared remarkably superposed in the catalytic region. A similar agreement was found for the position of one of the Mg^{2+} ions (only one crystallized in 3KYL, but two of them in 1RTD). Furthermore, the catalytic pocket of our RT hTERT protein model also fits very well into the RT structure of *T. castaneum* (from where it was derived) and the HIV-1. Because of all these superposition agreements, a chimeric complex: protein-DNA/RNA-dNTP-2 Mg^{2+} ions was built, assembling different parts from different sources, at the positions dictated by the superposition. Therefore, the hybrid double helix of DNA/RNA from *T. castaneum* (3KYL), and the incoming dNTP together with Mg^{2+} ions, both from HIV-1 RT (1RTD), were attached to the obtained protein RT model. In particular, the incoming dNTP

taken from HIV-1 superpose very well with the last nucleotide of the primer DNA in the hybrid DNA/RNA of *T. castaneum*. Then, in our complex, the last nucleotide in the DNA primer of *T. castaneum* was replaced by the incoming nucleotide taken from the HIV-1 structure. The DNA/RNA from *T. castaneum* is just an engineering continuous strand that forms a double helix when crystallized, with Watson-Crick (WC) binding pairs except at the n-1 position, where a Hoogsteen base pair is found. In our model, the hybrid DNA/RNA double helix of *T. castaneum* was cut and only the nine base pairs located in, and directed away from the catalytic site were conserved. The Hoogsteen pair was changed to be a WC pair. The nucleotide sequences of DNA/RNA of *T. castaneum* were further adapted to the human case. Likewise, the incoming dNTP was changed from the dTTP (HIV-1 case in 1RTD), into a dGTP (corresponding to the human case). Therefore, using the procedure described above, a chimeric ternary complex for RT hTERT was built. Figure 1 schematically presents only the relevant parts for the catalytic process of this complex. Figure 1A depicts the C-I that corresponds to the dGTP placed at the catalytic site, just before the phosphodiester bond reaction occurs. On the other hand, Figure 1B schematically presents the C-II corresponding to the configuration just after the catalytic reaction was completed, when the dGTP was broken in a dGMP (incorporated to the DNA strand), and an inorganic PP_i, to be released from the catalytic pocket.

2.3 | Molecular dynamics simulations

All MD simulations were performed using the AMBER package version 12.[31-34] Parameters from the Amber ff12SB force field were used[31] for all protein residues and standard nucleotides; the parameter for the magnesium ions were taken from Ref 35, while parameters for dGTP were taken from Ref 36. Periodic boundary conditions were taken into account in all simulations. A weak-coupling algorithm[37] was used to couple the simulation boxes with an isotropic pressure of 1 atm and a reference temperature of 300 K. Relaxation times were chosen to be 5 ps and 2 ps for pressure and temperature coupling, respectively. All bonds involving hydrogen were constrained using the SHAKE algorithm.[38] The time step for all simulations was set to 2 fs. The positions of C_α for the first and last ten residues of the protein model were kept fixed during the MD simulation. Additionally, the positions of the RNA and DNA nucleotides, except for the last two DNA nucleotides, were also restricted. All the remaining protein atoms, those of the two nonfixed DNA nucleotides, the incoming dGTP atoms, the Mg^{2+} ions, as well as all water atoms, were free to move in the MD simulations. A first round of implicit water equilibration and MD simulation was carried out during 200 ns at a temperature of 300 K. Then, the complex was immersed into a truncated octahedron of TIP3P water molecules[39] and additionally, the bonds of the water molecules were constrained using the SETTLE algorithm.[40] The complete systems were neutralized with counter-ions. A direct cutoff for nonbonded interactions of 10 Å and the particle mesh Ewald for long-range electrostatics were applied.[41] A minimal distance of 10.1 Å between the surface complex and the edge of the simulation box was employed. For the MD simulations in explicit water, a standard protocol beginning with two initial energy minimizations was used: the first one only for the solvent molecules, constraining the

position of all non-solvent atoms, followed by a second minimization for the entire system. In order to equilibrate the WT and mutated complexes at 300 K, each system was slowly heated from 100 K over a period of 100 ps under NVT conditions, followed by MD simulations at 300 K up to 10 ns. Subsequently, a production run of 200 ns was performed for each complex under NPT conditions. Supporting Information Figure S3 (lower panel) shows the WT complex in C-I after 200 ns of MD simulations, with the catalytic binding pocket zoomed.

Six complexes were considered for the MD simulations of RT hTERT: the WT complex, and the K902N-, R631Q-, D712A-, D868A-, and D869A-mutated complexes. The mutations were implemented changing the corresponding lateral chains in the complex model in explicit water, before equilibration. Then, the starting MD configuration was the same for the WT and the mutated complexes. The WT, the K902N and R631Q complexes were exhaustively studied in both configurations, C-I and C-II during 200 ns, whereas the D712A, D868A, and D869A complexes were studied in C-I by 200 ns. Three sets of independent trajectories (with different random Maxwell-Boltzmann velocities) were considered for all systems, both in C-I and C-II. To check the stability of C-I, all complexes were further studied starting form a distorted configuration (related to C-I, see the steered MD configuration below), during additional 20 ns. As an example, our simulated WT system in C-I, contained 47 639 atoms in total (5353 protein atoms from 331 residues, 583 nucleotide atoms, 2 Mg atoms, 13 900 water molecules, and 1 Cl^- ion to neutralize the system).

2.4 | Binding free energies and alanine scanning

Binding free energies were calculated using post-processed trajectories according to the Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) approximation.[42,43] They were calculated for the WT, K902N, and R631Q systems, in C-I and C-II. The change in binding free energies ($\Delta G_{binding} = G_{complex} - [G_{receptor} + G_{ligand}]$) in solution (by using implicit water) was calculated. Particularly for the calculation in C-I, the Ligand was chosen to be dGTP, whereas for the calculation in C-II, PP_i was chosen instead. The remaining part of the complex was taken as the Receptor (protein + DNA/RNA + 2 Mg^{2+} ions) for both C-I and C-II. Finally, after removing water molecules and counter-ions, and using the last 40 ns of the trajectories for averaging (considering 200 frames, each frame at 200 ps), the binding free energies were calculated using the scripts available in the AmberTools12 package.[42] Binding Free Energy was calculated using the Onufriev, Bashford, and Case generalized Born model[43] to take the desolvation free energy into account. Furthermore, alanine-scanning calculations were performed on the WT complex, replacing one at the time, the K902, R631, and K626 residues to alanine, in C-I and C-II as well, using the last 40 ns of the WT original trajectories.

2.5 | Steered MD configuration

The distorted, C-I-related configuration (called the steered MD or just sMD configuration) was considered because of the following: (i) Due to its construction, C-I of the WT complex corresponds to the configuration where dNTP is located just before the phosphoryl transfer reaction occurs. (ii) After 200 ns of MD simulations for

WT, K902N, and R631Q, dNTP remains in the same configuration (see section 3). The idea was to test the stability of C-I for the mutated complexes. This new configuration was obtained following a steered MD simulation procedure.[44] After 200 ns on WT, K902N and R631Q, a steered MD was implemented by additional 10 ns in each system, in such a way that the $O_{3'}$ - P_α distance was forced to be 6 Å (almost twice the value of WT). At this distance, when running a new, normal MD, the WT complex does not recover the correct configuration for the catalytic reaction to occur (ie, it does not recover C-I). Then, several MD simulations were performed, starting with a diminished $O_{3'}$ - P_α distance, until a threshold value of 4.55 Å was found. All starting WT configurations with distances lower than or equal to this value, were able to recover C-I. Therefore, the WT configuration with the $O_{3'}$ - P_α distance equal to 4.55 Å is called "steered MD (sMD) configuration" here. The same procedure was applied to obtain sMD configurations (the same distance of 4.55 Å as in WT) corresponding to the K902N and R631Q complexes. They were used as the starting configurations, for new normal MD simulations on WT, K902N, and R631Q complexes, during 20 ns. Finally, the WT sMD configuration was also used as the starting configuration to perform 20 ns of MD simulations on D712A, D868A, and D869A complexes (obtained from single residue substitutions on the WT sMD configuration).

3 | RESULTS

3.1 | Molecular model for the WT RT hTERT complex

Having obtained a theoretical 3D model for the RT domain of the hTERT complex in addition to the 3D model for the complete hTERT previously found in the literature,[24] it is worth comparing both models. We do not have the coordinates of the previous model then we only compare how they were obtained. We consider that both final models (ours, for the RT domain of the WT hTERT complex, and that from Ref. 24 for the full protein WT hTERT complex), are those obtained at the end of the respective equilibration procedures. Some differences arise when comparing the steps used to build the complexes: (i) Obtaining a homology model for the RT domain of the hTERT protein: even if the particular methodologies used were different, they may be considered as equivalent. Moreover, both works have used the same 3KYL *T. castaneum* 3D structure as the template for this domain. Then, we may expect both homology models of the RT domain of the hTERT protein to be very similar. (ii) Building the chimeric WT hTERT complex: Even if both models obtain DNA/RNA from the 3KYL structure of *T. castaneum*, there are differences in the complex cofactors used by each model. In our model, in addition to the hybrid DNA/RNA helix, we include a dGTP and 2 Mg^{2+} ions (both taken from an HIV-1 structure), while these cofactors were not included in the model of Ref. 24. (iii) Equilibrating the complex: We used a careful procedure to equilibrate the complex in explicit water including several minimizations and 10 ns of MD simulations under NVT conditions (see section 2). In Reference 24, a minimization and 0.2 ns of MD

simulation under NVT conditions were used (no mention was made to any solvent). Nevertheless, we would not expect significant differences between models for RT WT hTERT.

Finally, in Ref. 24, the model for the full hTERT complex (only with DNA/RNA) was used as a starting point to perform coarse-grained elastic network models, in order to study telomerase processivity. Instead, our model of RT domain of WT hTERT complex was used to perform MD simulations at atomic level, of some mutations on conserved residues that abolish telomerase activity.

3.2 | Independent MD simulations for WT, K902N, and R631Q in C-I and C-II

After three independent runs of 200 ns of MD simulations, for the WT, K902N, and R631Q complexes in C-I, we obtained essentially the same qualitative results (raw data concerning some structural parameters are given in Supporting Information Figures S4-S7). For C-II, the three independent runs gave the same qualitative results for WT, but some differences are observed for the K902N and R631Q systems (Supporting Information Figure S8). These differences will be discussed in the corresponding subsection below. Moreover, Table 1 collects the average and SD values for the distances presented in Figures 3–5, corresponding to the minimized structures after the three 200 ns of MD simulations. Figures 2–5, presented in the following subsections, correspond to the first MD simulation run of each complex, chosen as representative of all the independent runs.

TABLE 1 C-I (top) and C-II (bottom) average and SD values (in Å) corresponding to the distances presented in Figures 3–5, from the three independent runs of 200 ns of MD simulations, followed by a final minimization

Configuration I			
	WT	**K902N**	**R631Q**
O3' - P_α	3.22 ± 0.01	3.30 ± 0.09	3.26 ± 0.03
K902[a] - P_α	3.9 ± 0.5	6.0 ± 0.3	3.5 ± 0.1
K902[a] - P_γ	3.4 ± 0.1	6.7 ± 0.4	3.5 ± 0.2
R631[b] - P_α	3.74 ± 0.06	3.7 ± 0.5	-
R631[b] - P_β	3.8 ± 0.1	4.0 ± 0.4	5 ± 1
K626 - P_γ	3.39 ± 0.05	3.45 ± 0.06	3.8 ± 0.2
K626 - P_α	-	-	4 ± 1
Configuration II			
	WT	**K902N**	**R631Q**
Old P_α - P_β	6.9 ± 0.3	5 ± 1	4.6 ± 0.6
K902[a] - P_β	3.5 ± 0.3	7 ± 2	-
K902[a] - P_γ	3.9 ± 0.8	6 ± 2	4.8 ± 0.6
R631[b] - P_β	4.1 ± 0.4	5 ± 1	-
R631[b] - P_β	4.1 ± 0.9	-	-
R631[b] - old P_α	-	6 ± 1	-
K626 - P_β	4.2 ± 0.5	4.4 ± 0.6	-
K626 - P_γ	3.7 ± 0.4	3.4 ± 0.1	-

[a] For the K902N mutation this residue is Asn.
[b] For the R631Q mutation this residue is Gln.

## 3.3 | Wild type complex: Watson-Crick pairing stability

We considered the stability of the Watson-Crick (WC) interactions (between the incoming nucleotide and the corresponding base from the RNA template) as prerequisite for considering the incoming dNTP as located in the active site of telomerase, just before the catalytic reaction takes place. In order to check that the WC H-bond distances (for the incoming dGTP and the last nucleotide of the DNA primer with their corresponding bases from the RNA template) were calculated as a function of the simulation time. They are presented in Supporting Information Figure S9 for WT in C-I (first MD run). In the same figure, the running average distances (averages taken inside a moving window of 2 ns, as the simulation progresses) were also reported. Even if some dispersion was observed during the simulation, the running average distances approach the expected H-bond distances for the WC pairs. Therefore, the incoming dGTP could be considered to be positioned in the active site.

## 3.4 | Wild type complex: Octahedral coordination of Mg^{2+} ions

In our results, both Mg^{2+} ions are found octahedrally coordinated by several key conserved residues and by the incoming dNTP, similarly to what was found in *T. castaneum*,[9,10] HIV-1,[13] and several experimental polymerases.[14,15] The coordination distances as a function of the simulation time, obtained for both Mg^{2+} ions in the C-I of WT (first MD run), are presented in Supporting Information Figure S9. In spite of some distance dispersion, in general running average distances are quite similar, indicating that an almost perfect octahedral coordination for both Mg^{2+} ions in C-I is achieved. Similar results were found for C-II (not shown).

In Figure 2, representative pictures of the structural coordination of both Mg^{2+} ions (called A and B), for the human WT complex in the C-I and C-II configurations, are presented. They were obtained by minimization after 200 ns of MD simulations of the first run on each configuration. In the C-I of WT (Figure 2A), the Mg^{2+} A ion is found coordinated by the O$_{3'}$ of the DNA primer and by one oxygen atom from the P$_\alpha$ group (besides the three conserved Asp and a water molecule). When all these coordinations are satisfied, the incoming dGTP seems to be located at the catalytic site, in such a way that the O$_{3'}$- P$_\alpha$ distance is appropriate for the initiation of the nucleotide addition reaction. According to the two-metal-ions mechanism, the proximity of the Mg^{2+} A ion with the O$_{3'}$, lowers the affinity of O$_{3'}$ for hydrogen, facilitating the O$_{3'}$ attack on P$_\alpha$, allowing the formation of the pentavalent transition state.[14,15] In C-I (Figure 2A,B) both Mg^{2+} ions stabilize the structure and the electric charge of the transition state. On the other hand, for WT in C-II, the Mg^{2+} B ion (Figure 2D) accompanies the way out of PP$_i$, while the Mg^{2+} A ion remains inside the catalytic pocket and is still coordinated by the three conserved Asp (Figure 2C). In the functional enzyme, the next step after the nucleotide incorporation is the single nucleotide translocation step. As this translocation involves other domains than RT of hTERT,[1,24] it is not captured by our model. However, in the Discussion, we will return to Figure 2 and the one translocation step when discussing the results of D869A.
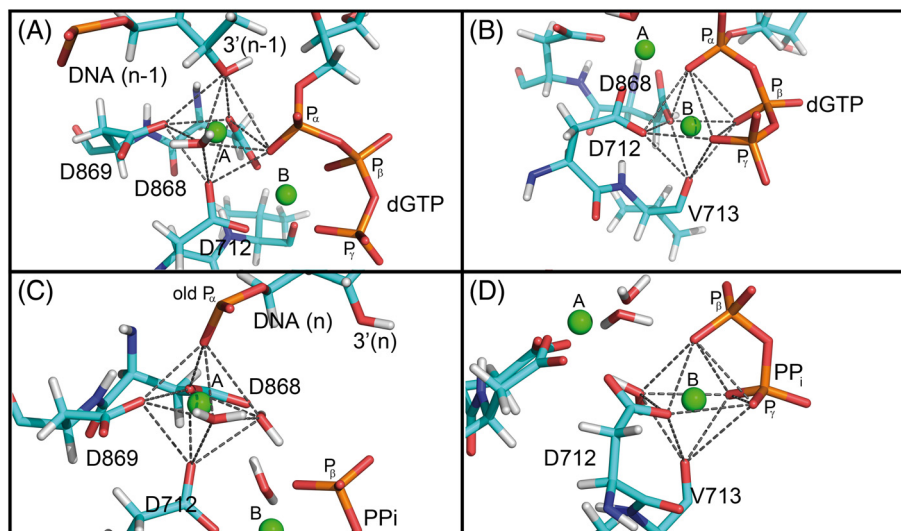
FIGURE 2 Coordinations for both Mg^{2+} ions (A and B spheres; in green in the online figure), for the WT complex after the first run of 200 ns of MD simulations. A, Mg^{2+} A ion coordination in C-I. It is octahedrally coordinated by the O$_{3'}$ of the last nucleotide of the DNA primer; by the lateral chains of the three conserved aspartic acid residues (D712, D868, and D869); by an O atom belonging to the P$_\alpha$ group of the incoming dGTP; and finally, by the O atom of a water molecule. B, Mg^{2+} B ion coordination in C-I. It is octahedrally coordinated by three O atoms, each belonging to the P$_\alpha$, P$_\beta$ and P$_\gamma$ groups, respectively, of the incoming dGTP; by the lateral chains of two of the conserved aspartic acids (D712 and D868); and by the carbonyl O belonging to the backbone of V713 residue. C, Mg^{2+} A ion coordination in C-II. It is coordinated by the old P$_\alpha$ group (now integrated to the DNA), the lateral chains of the three conserved aspartic acids and two water molecules. D, Mg^{2+} B ion coordination in C-II. It is coordinated by three O atoms of P$_\beta$ and P$_\gamma$ groups of PP$_i$; the lateral chain of the conserved D712 residue; the carbonyl O atom of V713 residue; and a water molecule. The stick models correspond to a final minimization after 200 ns of explicit water MD simulation and were obtained using the PyMOL software[47] [Color figure can be viewed at wileyonlinelibrary.com]
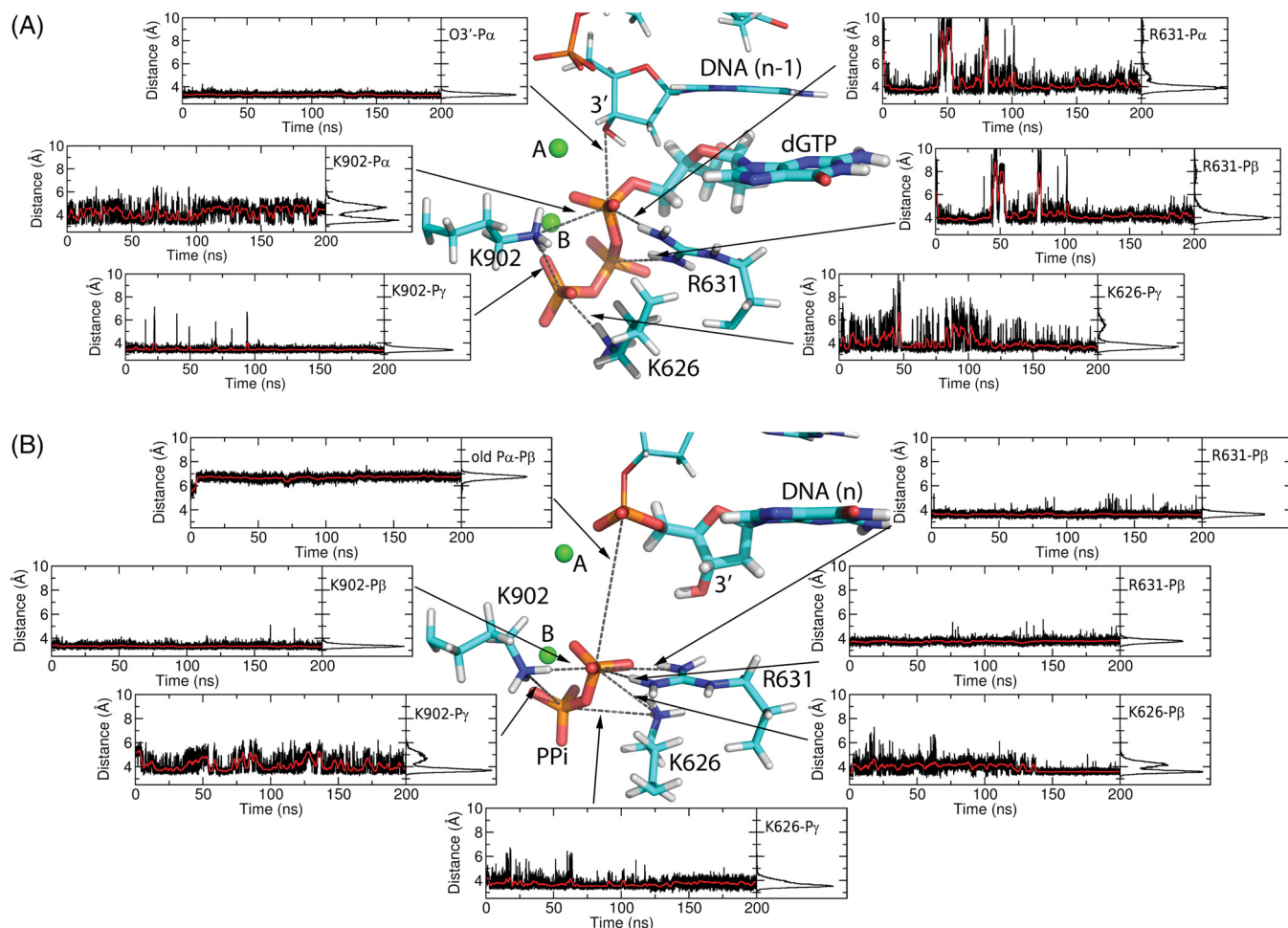
FIGURE 3 A, Contact distances in the WT hTERT complex as a function of the MD simulation time, for the C-I. The distances between the N atoms of the lateral chains of each residue (K902, R631, and K626) and the P atoms of the phosphate groups from the incoming dGTP are shown. The distance between $O_{3'}$ and P_α is also included. B, Same distances as in A, but for the C-II. In this case, distances represent the contacts with the P atoms of PP_i. The distance between the old P_α and P_β is also included in this case. Representative stick models correspond to a final minimization after 200 ns of explicit water MD simulation, obtained using the PyMOL software[47] [Color figure can be viewed at wileyonlinelibrary.com]

3.5 | Wild type complex: Contacts of K902, R631, and K626 residues

Selected contacts formed by the conserved residues K902, R631, and K626, in the WT complex are presented in Figure 3 for both configurations, as a function of the simulation time. Before the reaction (C-I, Figure 3A), it is found that all three conserved residues have contacts with the phosphate groups of the incoming dGTP, suggesting they could also play some role in the stabilization of the dNTP at the reaction site. In particular, the P_α atom makes contact with the N atom of the lateral chain of the K902 residue. Along the simulation time, this N atom explores two contact distances with P_α, whereas the running average distance between the P_α and the $O_{3'}$ atom remains constant in average. Moreover, P_α also makes contact with both N of the lateral chain of R631 residue (for the sake of clarity, only one of them was included in Figure 3A, being the other N contact completely similar). Residue R631 also makes contact with P_β, whereas K626 and K902 residues make contact with the P_γ atom. The distance between $O_{3'}$ and P_α along the simulation time is remarkably stable around the mean value of 3.2 Å. Basically, the same average distance value is obtained

for the independent WT MD runs, as it is shown in Supporting Information Figure S4 and Table 1, confirming that this configuration should correspond to the catalytic one.

Concerning C-II, Figure 3B shows that PP_i tends to leave from the catalytic pocket (indicated here by the increasing distance of the broken bond between the old P_α and P_β), accompanied by the lateral chains of the K902, R631, and K626 residues. The raw data for this WT distance for the three independent MD runs are presented in Supporting Information Figure S8 and the average value from the minimized structures is collected in Table 1. They show that the same qualitative result is obtained: average old $P_\alpha-P_\beta$ distances greater than 6 Å. Our result for K902 residue is in agreement with the proposition that it accompanies the exit of PP_i and stabilizes the charge generated by this leaving group.[16] Our results show that not only K902, but also R631 and K626 residues as well, contribute to this job.

Finally, we would like to mention that in both configurations of WT, there are other residues making contact with the incoming dGTP. This is the case the N899 residue that makes contact with the P_γ group in both C-I and C-II. Since, it is not a conserved residue and it
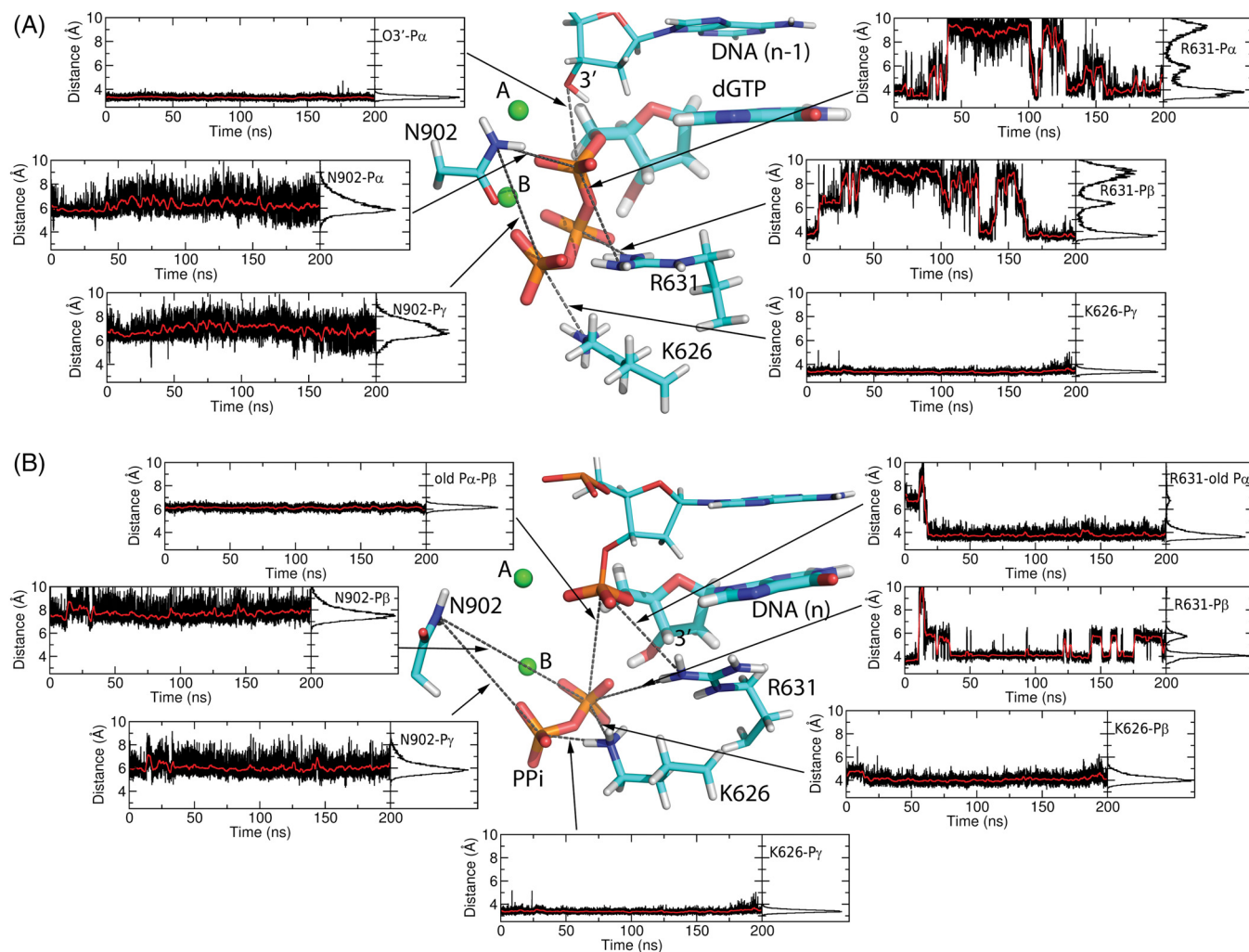
FIGURE 4 A, Contact distances in the K902N hTERT-mutated complex as a function of the MD simulation time, for the C-I. Same distances as in Figure 3, except that the N atom of the lateral chain of N902 replaces the corresponding N atom of K902 residue. B, Same distances as in A, but for C-II. Representative stick models correspond to a final minimization after 200 ns of explicit water MD simulation, obtained using the PyMOL software[47] [Color figure can be viewed at wileyonlinelibrary.com]

was not reported in the literature as essential for telomerase activity, we did not include it in Figure 3.

3.6 | K902N-mutated complex: Contacts of N902, R631, and K626 residues

Concerning the K902N mutated complex in both configurations, some selected distances involving N902, R631, and K626 residues are presented in Figure 4, as a function of the simulation time, for the first MD run. Comparing Figures 4A and 3A for C-I, it is evident that the N902-mutated residue is incapable of maintaining similar contacts with the P_α and P_γ groups (average distances greater than 6 Å), as those formed by the wild type K902 residue. In Figure 4A, R631 residue appears as trying to compensate the absence of K902 interactions, still maintaining contacts with P_α and P_β, but showing greater dispersion than in the WT complex. Finally, K626 residue also maintains contact with P_γ (showing a significantly lower dispersion than in the WT complex). In spite of the absence of the K902 contacts and

the greater fluctuations of the R631 residue, the distance between $O_{3'}$ and P_α, along the simulation time for the three MD runs shows a similar behavior to that in the WT complex. All of them have little dispersion around almost the same average value of WT (Supporting Information Figure S4 and Table 1). Since the distance $O_{3'}$-P_α remains on average the same, it is possible to infer that once the incoming dGTP is placed in the reaction site, the K902N mutation does not significantly alter the WT reaction distance, keeping the appropriate value for the catalytic reaction to take place. This result was unexpected to a certain point, considering that this mutation experimentally abolishes the catalytic activity of telomerase. Additional tests using MD from a sMD configuration will be presented under Discussion.

Figure 4B shows the MD simulation results for the K902N mutated complex, in C-II for the first MD run. Comparing Figures 4B and 3B, it is evident that N902 residue is also incapable of forming similar contacts with PP_i as those formed by K902 in WT. K626 residue still keeps good contacts with P_γ and P_β. Concerning R631
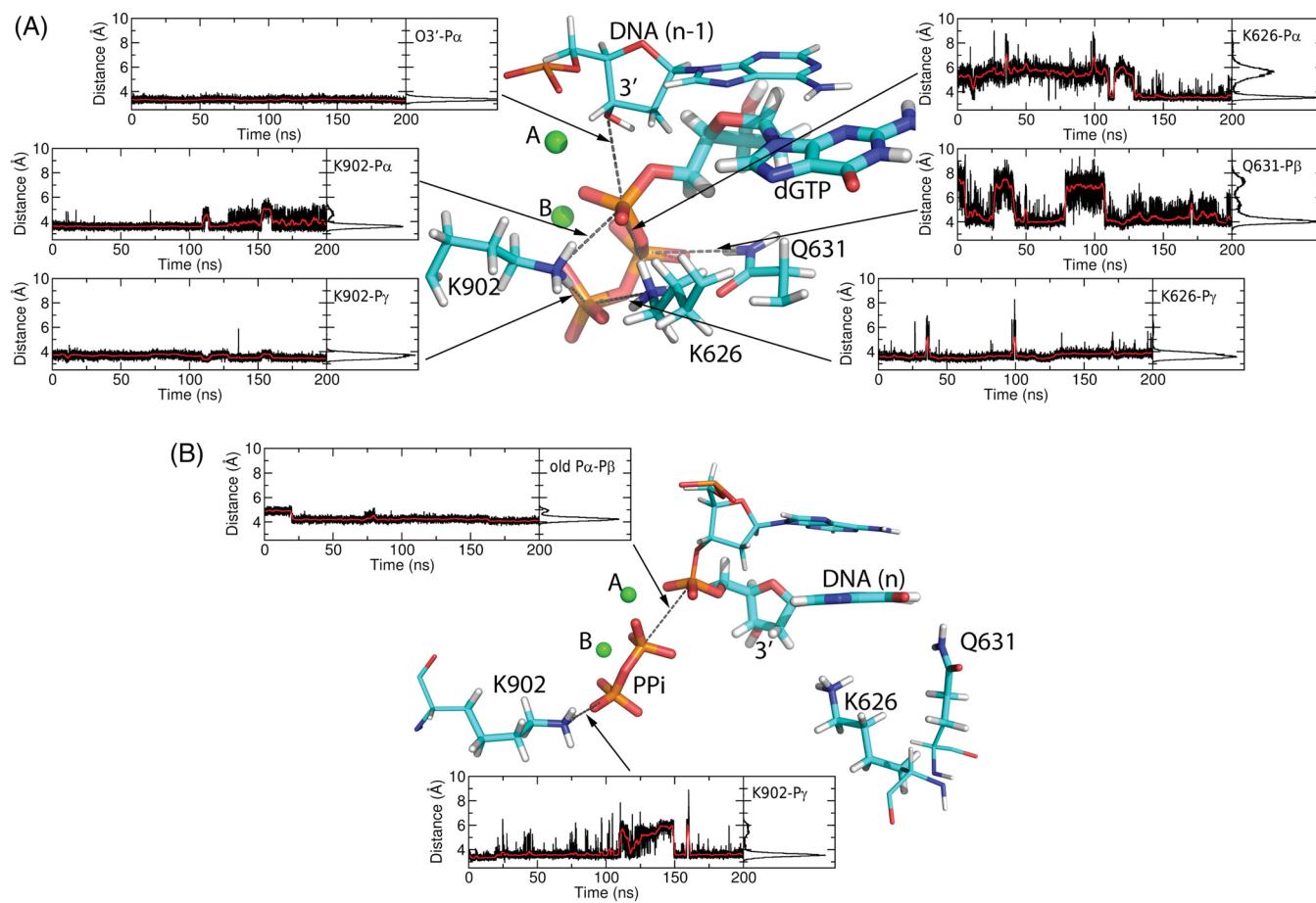
FIGURE 5 A, Contact distances in the R631Q hTERT-mutated complex as a function of the MD simulation time, for C-I. Same distances as in Figure 3, except that the N atom of the lateral chain of Q631 replaces the corresponding N atoms of R631 residue. B, Same distances as in A, but for C-II. Representative stick models correspond to a final minimization after 200 ns of explicit water MD simulation, obtained using the PyMOL software[47] [Color figure can be viewed at wileyonlinelibrary.com]

residue, it makes contact with P_β of PP_i (as in the WT, but with more dispersion), but unlike the WT case, it still continues making contact with P_α (incorporated to the DNA in this configuration). The old P_α and the P_β average distance is, in C-II of K902N complex of Figure 4B, smaller than that of the WT for the same simulation time. The additional MD runs for this complex in C-II show that this distance could be even shorter (Supporting Information Figure S8) with an average value (Table 1) lower than that of the WT, indicating the impossibility for PP_i to leave from the catalytic pocket. This is partly due to the absence of contacts with N902 residue, and partly due to the contact that R631 residue still maintains with the old P_α after the reaction. These results are a first indication that the mutation K902N might alter the normal way out of PP_i once the first nucleotide addition reaction is accomplished, thus stopping the activity of human telomerase.

3.7 | R631Q-mutated complex: Contacts of K902, Q631, and K626 residues

We consider now the R631Q complex in both configurations. Selected distances concerning K902, Q631, and K626 residues for this complex are presented in Figure 5, for both configurations as a function of the simulation time, obtained for the first run of MD

simulations. For C-I before the reaction (Figure 5A), K902 residue has similar contacts with the P_α and P_γ groups, like those formed in WT (being the distance with the P_α, atom less dispersive than in WT). The mutated Q631 residue makes contact (with some important fluctuations) with P_β, whereas in WT a bit stronger contacts are formed with the P_α and P_β groups by R631 residue. Additionally, K626 residue shows similar contacts with P_γ as in WT but, in this case, also forms contact with P_α (contact not present in WT). Once again, it is as if the conserved nonmutated residues try to compensate for those contacts that are missing due to mutation. Interestingly, as in the case of the K902N complex, once the incoming dGTP is placed at the reaction site, the R631Q mutation does not seem to alter the distance between $O_{3'}$ and P_α either. This distance remains the same as in the WT complex, with very low dispersion, along the simulation time. This result is confirmed by the multiple runs of MD simulations (Supporting Information Figure S4 and Table 1).

However, when Figures 5B and 3B for C-II are compared, more striking differences arise. In the R631Q complex, Q631 and K626 residues have completely lost their contacts with the P groups of PP_i. K902 is the only residue that still makes contact with it, but only with P_γ (whereas, in the WT complex it has contacts with both P). Finally, the distance between the old P_α and P_β is averaged around 4.2 Å
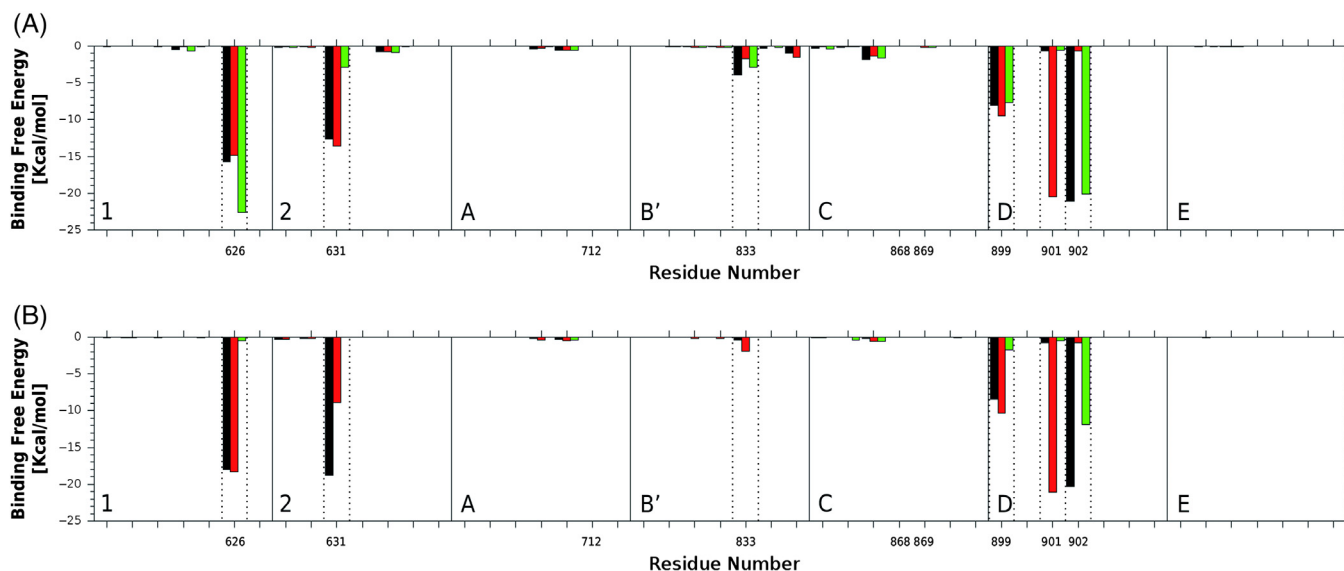
FIGURE 6 Binding free energy as a function of the amino acid residue in the sequence of WT and mutated hTERT complexes: (A) in C-I and (B) and C-II. Numbers correspond to the conserved residues in hTERT. Only the residues inside the RT motifs are shown. Vertical lines help identifying the residues for which the binding free energy is given. Each colored vertical bar corresponds to the following complexes: WT (black), K902N (red), and R631Q (green). The difference in free energy between a bound receptor + ligand complex and the unbound components was calculated with the methodology explained under Section Material and Methods.[33,34] In both configurations, the part of the complex formed by protein + DNA/RNA + 2 Mg^{2+} ions was considered to be the receptor, while the dGTP (in C-I) or PPi (in C-II) were chosen as ligand

along 200 ns, from the MD run shown in Figure 5B. The results from the two additional MD runs for this distance (Supporting Information Figure S8), show average values lower than 5.5 Å, with Q631 residue making contact with the old P$_\alpha$ (not shown). The minimized average values given in Table 1 confirm this tendency. All these results show that the R631Q mutation might not favor the exit of PP$_i$ from the catalytic pocket, thus stopping the catalytic activity of telomerase after the first single nucleotide incorporation.

4 | DISCUSSION

4.1 | Molecular dynamics analysis

From the MD simulations results (Figures 2–5 and S4-S10), we could make the following statements: (i) The C-I of the WT complex captures the catalytic position of the incoming dGTP, just before the phosphodiester reaction. This conclusion is drawn considering three stability criteria: the stability of the WC pairs (Supporting Information Figure S9); the stability of the octahedral coordination of both Mg^{2+} ions (in particular, the Mg^{2+} A ion coordinated by O$_{3'}$ and one O from the P$_\alpha$ group of dGTP, Figures 2A and Supporting Information S10); and the stability of the O$_{3'}$ - P$_\alpha$ distance at an average distance of 3.2 Å (suitable for the phosphoryl transfer reaction to take place, Figures 3A and Supporting Information S4, Table 1). (ii) K902N or R631Q mutations do not seem to affect the catalytic position of dGTP before

the reaction, even if the remaining native contacts are altered. This conclusion is reached following the same three stability criteria used for the case of WT, but now applied to the K902N and R631Q complexes in C-I (Figures 4A and 5A, respectively, and Supporting Information Figure S4). (iii) For WT in C-I, the K902, R631, and K626 residues make contacts with the three phosphate groups of dGTP (K902 with P$_\alpha$ and P$_\gamma$; R631 with P$_\alpha$ and P$_\beta$; K626 with P$_\gamma$, Figure 3A). (iv) For WT after the catalytic reaction, the K902, R631, and K626 residues make contacts with PP$_i$, accompanying its exit from the catalytic pocket. The Mg^{2+} B ion is coordinated by the phosphate groups of PP$_i$ and goes out with them (Figure 3B). (v) For the K902N complex in C-II, PP$_i$ is unable to exit the catalytic pocket as in the WT protein, due to the lost contacts with the mutated K902 residue, and because the R631 residue remains in contact with the old P$_\alpha$, even after the reaction is complete (Figure 4B). (vi) For R631Q complex in C-II, PP$_i$ remains inside the catalytic pocket, now due to the lost contacts with the mutated R631 residue as well as with the K626 residue. The only WT contact that remains (that of the K902) is not sufficient to accompany the way out of PP$_i$ that becomes stabilized inside the catalytic pocket (Figure 5B). (vii) For the K902N and R631Q complexes, in both C-I and C-II, the non-mutated WT conserved residues try to compensate for the absence of the mutated residues (Figures 4 and 5). (viii) For the K902N and R631Q complexes, once the dGTP is placed at the catalytic site in C-I, the incorporation of the first nucleotide to the DNA strand seems to be possible. Then, as PP$_i$ is no longer able to exit

TABLE 2 Change in the binding free energies, from an alanine-scanning of K626, R631, K902, and N899 residues in the human telomerase WT complex, in C-I and C-II. In both configurations, the ligand and the receptor were the same as in Figure 6

Δ(ΔG binding free energy) (kcal/Mol)	K626A	R631A	K902A	N899A
Configuration I (C-I)	−37 ± 5	−36 ± 6	−63 ± 8	−21 ± 3
Configuration II (C-II)	−36 ± 5	−33 ± 3	−49 ± 8	−22 ± 3

the catalytic pocket, telomerase activity stops. This result is in complete agreement with the experimental evidence that both K902N and R631Q mutations abolish telomerase activity.[20–23] In conclusion, K902N or R631Q mutations drop telomerase activity to zero, because the mutated complexes might interrupt the process by stopping the PP_i exit after the first nucleotide addition.

4.2 | Binding free energy analysis

Due to the similar contacts with the incoming dGTP (in C-I), and PPi (in C-II), we decided to perform an energetic analysis for the WT, K902N, and R631Q complexes. The last 40 ns of the corresponding trajectories of the first MD run were processed in order to calculate binding free energies by using the MM-GBSA method, as it was explained in section 2. Figure 6 shows the binding free energy for the WT (black), K902N (red), and R631Q (green) systems, in C-I (Figure 6A) and C-II (Figure 6B), as a function of the residues inside the RT motifs for hTERT. For the WT complex in C-I (black, Figure 6A), the conserved residues making contacts with the dGTP and contributing to the binding free energy are, in order of importance, K902, K626, and R631. For WT in C-II (black, Figure 6B), the same conserved residues are found, with the contribution of K902, R631, and K626 in order of importance in this case. As it was already mentioned under section 3 for the WT complex, the non-conserved N899 residue makes contacts with P_γ of dGTP (in C-I) and PP_i (in C-II). Figure 6 shows that it makes a less important contribution to the binding free energy than the conserved residues, in both configurations. An even smaller contribution was found for the conserved Q833 residue only in C-I of the WT system. For the K902N complex in C-I (red, Figure 6A), practically the same contribution of K626 and R631 residues to the binding free energy as in WT was found. However, the mutated N902 residue does not contribute to the binding free energy. Interestingly, the nonconserved neighbor R901 residue appears to compensate for the absence of K902 contribution. The same situation holds for the K902N complex in C-II (red, Figure 6B). It should be noted that the nonconserved R901 residue does not contribute to the binding free energy in WT. Just for completeness, we mention that the R901W mutation in human TERT lowers telomerase activity to less than 25% of WT, but it does not completely eliminate it.[45] Finally, for the R631Q complex in C-I (green, Figure 6A), Q631 residue contributes very little, whereas K626 residue reinforces its contribution, compensating for the absence of R631 contribution. For R631Q in C-II, the contributions of K626 and Q631 are practically lost, and only a fraction of the WT contribution of K902 residue remains. Summarizing, we may conclude that: (i) all these results are consistent with those of the previous subsection; and (ii) the contribution of K626 conserved residue to the binding free energy seems to be as important as the other conserved residues.

4.3 | Alanine-scanning analysis

Going a step forward, we also analyzed the change in the WT binding free energy when selected residues in the WT were mutated to alanine (see section 2).[42,43] The alanine-scanning analysis was applied to the conserved K626, R631, and K902 residues and we included the

nonconserved N899 residue to check its relative importance when mutated to Ala.

Table 2 shows alanine-scanning results, presenting the change in the WT binding free energy for each residue analyzed. The residues that show the highest values (K626, R631, and K902) are precisely those that greatly affect telomerase activity. The most important contribution to the WT binding free energy is that of K902 residue in both configurations. R631 and K626 residues show comparable contributions, being that of K626 somewhat greater. These results are in complete agreement with the experimental results: mutations K626A,[19] R631A,[19] and K902A[20] cancel the catalytic activity of human telomerase. Finally, the N899 contribution shown in Table 2 is lower than the other contributions, but not negligible, with a similar value in both configurations. Concerning this nonconserved N899 residue, to our knowledge there is no experimental evidence studying its single amino acid substitution. Nevertheless, from the binding free energy and alanine-scanning results, we have found some evidence that mutations in this residue could affect, to some extent, the activity of human telomerase.

4.4 | MD started from the sMD configurations: WT, K902N, and R631Q complexes

From the results shown above, it is evident that considering the C-I of whatever system (WT or mutated complexes), the $O_{3'}-P_\alpha$ distance remains, in average, the same along the different runs of 200 ns of MD simulations. In order to analyze the stability of C-I itself, we next consider a distorted (C-I related) configuration, called the sMD configuration, for the WT, K902N, and R631Q complexes. In the sMD configuration the $O_{3'}-P_\alpha$ distance was forced to be 4.55 Å (the maximal value of the distorted distance where WT recovers C-I, see section 2). The sMD was used as the starting configuration for other MD simulations. Since they converge quickly, we performed only one MD run of 20 ns each for each complex, just to check if the systems recover C-I. Results are presented in Table 3, where the representative $O_{3'}-P_\alpha$ and the Mg A–Mg B distances are given. The comparison was made among their values before and after 20 ns of MD simulations and those obtained after 200 ns of normal first run of MD for each complex. In all cases, a final minimization was performed and the comparison was made using the minimized distances.

After only 20 ns of MD simulation started at the sMD configuration, the three complexes studied showed $O_{3'}-P_\alpha$ and Mg A-Mg B distances in agreement with those obtained after 200 ns of normal MD (for each complex in C-I). The conclusion is that the three complexes: WT, K902N, and R631Q are able to recover essentially the catalytic positions, with the dGTP placed as in C-I. In other words, the K902N and R631Q mutations might not prevent the system from recovering the catalytic C-I (at least when started from the sMD configuration). These results are, again, quite unexpected because some disturbance at the entrance of dGTP could be expected. However, from the results shown in Table 3, both mutations do not appear to disturb the dGTP entry into the catalytic site. Table 3 shows results which are totally compatible with the previous conclusion that the effect of any of the K902N and R631Q mutations in human telomerase activity might be

TABLE 3 $O_{3'} - P_\alpha$ and Mg A − Mg B characteristic distances in WT, K902N, and R631Q complexes after 20 ns of normal MD simulations (followed by a minimization), started from the corresponding sMD Configuration (sMDC) of each complex

	Distance $O_{3'} - P_\alpha$ (Å)			Distance Mg A − Mg B (Å)		
	Normal MD started at the sMDC		Normal MD	Normal MD started at the sMDC		Normal MD
	0 ns	20 ns	200 ns	0 ns	20 ns	200 ns
		(min)	(min)		(min)	(min)
WT	4.55	3.27	3.22	4.01	3.53	3.60
K902N	4.55	3.18	3.25	4.41	3.55	3.64
R631Q	4.55	3.24	3.29	4.22	3.61	3.56

These values are compared to those obtained after the first run of 200 ns of MD simulations (followed by a minimization) for each complex in the catalytic C-I (Figures 3A, 4A, and 5A, respectively).

to prevent PP$_i$ from leaving the catalytic pocket (Figures 4B and 5B), thus stopping the catalytic activity of hTERT.

Motivated by these results a natural question arises: what is(are) the mechanism(s) by which the D712A, D868A, and D869A mutations cancel telomerase catalytic activity? To address this question two new MD simulations were considered: (i) a single run of 20 ns of normal MD for the WT, where each conserved aspartate, one at a time, was mutated to alanine, starting from the WT sMD configuration. (ii) Three independent runs of 200 ns of normal MD for the mutated D712A, D868A, D869A complexes (obtained similarly to WT, K902N, and R631Q complexes) in C-I. These results are presented in the following subsections.

4.5 | MD started from the WT sMDC: D712A, D868A, and D869A mutations

Starting from the WT sMD configuration, where each conserved aspartate, one at a time was mutated to alanine, 20 ns of MD simulations were implemented. Results are presented in Table 4, where the same parameters shown in Table 3 are analyzed. The corresponding values for the WT after 200 ns in C-I (first run) are included for comparison. Again, a final minimization was performed for each system and the comparison was made using the minimized distances. From these results, it is possible to observe that after only 20 ns of MD simulation, the D712A and D868A mutated systems do not recover the WT distances. Both of them increase the distance between Mg ions (4.87 Å compared to the WT value of 3.60 Å), whereas the D712A-mutated complex notably increases the $O_{3'} - P_\alpha$ distance (6.28 Å) with respect to that of the WT system (3.22 Å). The only mutated system that seems to recover the C-I of the WT is the D869A complex ($O_{3'} - P_\alpha$ distance: 3.26 Å compared to 3.22 Å; Mg A-Mg B distance:

3.59 Å compared to 3.60 Å). These results are the first evidence that the mechanism to cancel telomerase activity might probably be different for each conserved aspartate mutant complex.

4.6 | MD started from the C-I configuration: D712A, D868A, and D869A complexes

The D712A, D868A, and D869A complexes were built and equilibrated, in the same way as the K902N and R631Q complexes in C-I (See section 2). Then, fresh MD simulations were performed on each of these complexes, for three independent runs of 200 ns each. The idea was to analyze if MD simulations started from C-I with Asp substitution from the beginning are stable along the simulation time. In particular, we are interested in the coordination of both Mg^{2+} ions and the $O_{3'}$-P_α distances. The results for some structural parameters, for the three independent MD simulation runs on the D712A, D868A, and D869A complexes in C-I, are presented in Figure 7 in a box plot representation. For comparative purposes, we included the results for the same structural parameters, also for the three independent runs of WT, K902N, and R631Q complexes in C-I. Figure 7 compares the following structural parameters: the $O_{3'}$-P_α distance (Figure 7A); the $O_{3'}$-Mg A distance (Figure 7B); the Mg A-Mg B distance (Figure 7C), and the $O_{3'}$-P_α-$O_{\alpha\beta}$ angle (Figure 7D). The raw results, as a function of the simulation time for the three independent runs of each system, are given in Supporting Information Figures S4-S7. For comparison, the experimental values of the same parameters corresponding to human DNA polymerase η are also included in Figure 7. They correspond to a crystallized structure where the nucleotidyl-transfer reaction is just taking place (PDB ID code 4ECV).[46] In essence, at this catalytic step both the human DNA polymerase η and hTERT, perform the addition of one nucleotide to the DNA strand by using the same

TABLE 4 $O_{3'} - P_\alpha$ and Mg A − Mg B characteristic distances in WT with the D712A, D868A, and D869A mutations (one at a time), after 20 ns of normal MD simulations (followed by a minimization), started from the WT sMD configuration (sMDC)

	Distance $O_{3'} - P_\alpha$ (Å)			Distance Mg A − Mg B (Å)		
	Normal MD started at the WT sMDC		WT normal MD in C-I	Normal MD started at the WT sMDC		WT normal MD in C-I
	0 ns	20 ns	200 ns	0 ns	20 ns	200 ns
		(min)	(min)		(min)	(min)
D712A	4.55	6.28	3.22	4.01	4.87	3.60
D868A	4.55	3.71	3.22	4.01	4.87	3.60
D869A	4.55	3.26	3.22	4.01	3.59	3.60

The distances are compared to those obtained after the first run of 200 ns of MD simulations (followed by a minimization) for the WT complex in the catalytic C-I (Figure 3A).
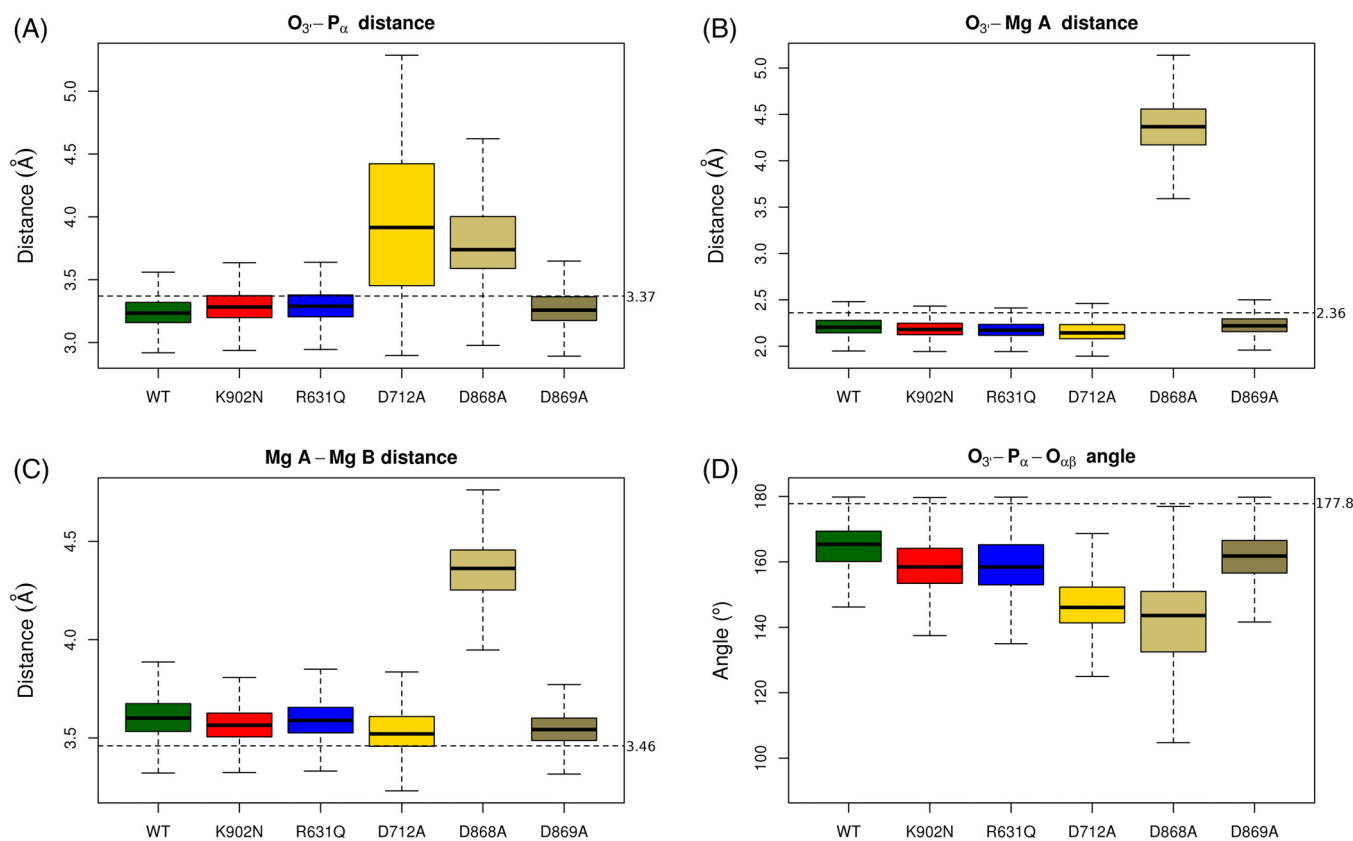
FIGURE 7 Characteristic structural parameters result from MD simulations for the WT, K902N, R631Q, D712A, D868A, and D869A complexes, all in C-I, presented as box plots. Each complex was studied performing three independent MD simulations of 200 ns. Each box plot includes 60 000 data points. A, $O_{3'}$ - P_{α} distance. B, $O_{3'}$ - Mg A distance. C, Mg A - Mg B distance. D, $O_{3'}$ - P_{α} - $O_{\alpha\beta}$ angle. The $O_{\alpha\beta}$ is the bridge O atom between the P_{α} and P_{β} of the incoming dGTP (Figure 2A). The dashed lines correspond to the experimental values of human DNA polymerase η structure, caught just making a phosphodiester bond (PDB ID code 4ECV).[46] As usual, each box encloses values between the first and third quartiles (enclosing 30 000 data points in this figure). The second quartile (the median) is indicated inside each box. The "whiskers" below/above the box are calculated as 1.5*IQR (or including the full range of values if they are found to be less than 1.5*IQR), for each system. IQR means InterQuartile range, and corresponds to the height of each box. The box plots were obtained post processing the raw data (presented in Figures S4-S7), with the R statistical package[48,49] [Color figure can be viewed at wileyonlinelibrary.com]

two-metal-ion mechanism. Then, it would be expected that the nucleotide-binding pocket of C-I in WT hTERT is, to some extent, comparable to the structure of the human DNA polymerase reported in 4ECV. The authors of this study have shown evidence (See Figure 2d in Reference 46), that the $O_{3'}$ -P_{α}-$O_{\alpha\beta}$ angle would approach planarity. Therefore, this parameter was incorporated in the comparison of Figure 7.

When considering the $O_{3'}$-P_{α} distance (Figure 7A), it is evident that D712A complex shows the greatest dispersion of values, with a median greater than the WT value. All three independent production runs for this system (Supporting Information Figure S4) started from the same equilibrated structure (and random Boltzmann velocities) shows the system exploring different distances with important variations from the median of the respective run. This result indicates that the D712A complex might allow different values of $O_{3'}$-P_{α} distance from those found in WT, compromising then the catalytic activity. Concerning the $O_{3'}$-P_{α} distance in the D868A complex, a similar discussion can be held, although this complex shows a lesser dispersion of results than the D712A complex. Figures 7A and S4 also show a median greater than that of WT for the D868A complex. Conversely, the box plot values for the D869A complex appear to be similar to

those of the WT. Finally, our results for the WT $O_{3'}$-P_{α} distance are in acceptable agreement with the experimental value of this distance in the structure of human DNA polymerase η (3.37 Å).

Figure 7B shows the box plot results for the $O_{3'}$-Mg A distance, where all the complexes show they are compatible with those of WT, except for the D868A system. The raw data for the $O_{3'}$-Mg A distance are given in Supporting Information Figure S5, for all complexes. The three MD runs for the D868A complex cluster far from the WT value. With a median of 4.4 Å for the $O_{3'}$-Mg A distance, the Mg A ion is too far from $O_{3'}$, as to contribute to the two-metal-ions mechanism proposed for the catalytic reaction. This is a clear indication that this mutation might not be functional. Supporting Information Figure S5 shows that the D712A complex also explores other values in addition to those of WT for this distance. Again, the $O_{3'}$-Mg A distances for the D869A complex are similar to that of the WT. Our WT results for the $O_{3'}$-Mg A distance are in reasonable agreement with that in the human DNA polymerase η structure.

Regarding the Mg A-Mg B distance (Figures 7C and S6), it is the box plot of the D868A mutation which is clearly different from all others. This is a further indication that severe distortions from the WT are present in the D868A complex. The three independent runs for

this distance, corresponding to the D712A complex, also show strong variations along the simulation time, but clustering at similar values as the WT (as indicated in the respective box plot). Once again, the D869A runs show results similar to the WT. The experimental human DNA polymerase η value for the distance between the Mg^{2+} ions is still in agreement with our results.

Finally, considering the $O_{3'}$-P_{α}-$O_{\alpha\beta}$ angle, Figures 7D and Supporting Information S7 show that this structural parameter is the one which has the most important variations among the complexes studied. The box plots for the D712A and D868A mutations are those that deviated most from the WT values (and also from the planarity). The box plot of this angle for the D868A complex is the largest one having the greatest dispersion. The K902N-, R631Q-, and D869A-mutated systems have similar angle box plots placed at values near that of the WT.

From all the results shown in Figure 7, it is possible to conclude that they represent another way to show that K902N and R631Q mutations do not alter the characteristic parameters of C-I. Completely different results were obtained when the D712A and D868A mutated complexes were considered. From Figure 7A, it is possible to conclude that the D712A mutation does not stabilize dGTP at the catalytic position, with $O_{3'}$-P_{α} distance spread along a wide set of values and with a median larger than the WT. Accordingly, it is not possible for the catalytic reaction to occur. Concerning the D868A mutation, it also shows a spreading of the $O_{3'}$-P_{α} distance, with a median larger than the WT value, but with less dispersion than the D712A complex. However, and unlike the D712A case, the D868A complex shows the box plots of the $O_{3'}$-Mg A and the Mg A-Mg B distances, as very different from the WT (Figure 7B,C). This is a clear indication that the D868A mutation should generate a distorted structure, where the catalytic reaction would be unable to occur. The D868A complex has, in addition, the greatest deviation with respect to WT for the $O_{3'}$-P_{α}-$O_{\alpha\beta}$ angle. On the other hand, taking into account the three independent MD executions of 200 ns performed and all the structural parameters analyzed, the D869A complex is the only one that provides similar box plots to those of the WT.

## 4.7 | Octahedral coordination of Mg^{2+} ions in mutated Asp complexes

In order to contribute to a better understanding of the previous results, the coordination of both Mg^{2+} A and B ions, in C-I is presented in Figure 8 for each mutated Asp complex. For easier comparison, the WT result was also included in Figure 8, and all the coordination residues (or groups) are summarized in Supporting Information Table SI. For the WT complex in C-I, the residues coordinating both Mg^{2+} ions are given in the caption of Figure 2. It is important for the present discussion to consider here that in WT: (i) the D712 and D868 residues coordinate both Mg^{2+} ions, whereas the D869 residue coordinates only the Mg^{2+} A ion; and (ii) both octahedra share a common vertex at the O atom from the P_{α} group.

Concerning the D712A-mutated complex, two representative structures (largely visited during the MD simulations) are presented (Figure 8A,B). In both structures, D868 and D869 residues are coordinating now both Mg^{2+} ions. For the Mg^{2+} B ion, the double

coordination of D869 is the only change with respect to the WT. However, the interesting differences come from the Mg^{2+} A octahedron. In Figure 8A, the Mg^{2+} A ion is found coordinated twice by the P_{α} group: with the same O atom that coordinates in the WT and with $O_{5'}$ (bridging P_{α} and sugar), coordination that is not present in the WT. In Figure 8A, the other coordination atoms of Mg^{2+} A ion are the same as in WT. However, the D712A complex explores other configurations besides that presented in Figure 8A. In the Figure 8B, $O_{5'}$ does no longer coordinate Mg^{2+} A. Instead, a new coordination with another water molecule is formed (Figure 8B), or this coordination is replaced by a new one with O carbonyl belonging to L866 residue (not shown, present in the third MD run). In all cases, coordination with $O_{3'}$ is maintained (as illustrated in Figure 7B), but the $O_{3'}$-P_{α} distance is not constrained to keep the WT values anymore. It spreads over a large interval with a median distance larger than that of WT. Comparing Figure 8A,B, it is as if the Mg^{2+} A octahedron pivoted around the common vertex without losing the coordination with the $O_{3'}$ atom. In the WT, this pivoting configuration does not occur due to the presence of the D869 original contact. Summarizing, when D712A is present, a conformational change should occur in such a way that D869 residue leaves its WT contact, in order to coordinate both Mg^{2+} octahedra (replacing the lacking D712 residue). To replace the original contact of D869 in WT, the Mg^{2+} A ion seems to explore contacts with $O_{5'}$, or with additional water, or even with O carbonyl from L866 residue. According to the two-metal-ion mechanism, the lowering of pKa of hydroxyl is still active, but due to the possible and different $O_{3'}$-P_{α} distances, the pentavalent transition state might not be formed, thus stopping the catalytic activity of hTERT.

We now consider the complex including D868A mutation. Due to the dispersion of values for $O_{3'}$-P_{α} distance obtained in Figure 7A, two representative structures are also analyzed in Figure 8C,D. The two octahedra in Figure 8C,D are not sharing a common vertex anymore. However, they continue to be linked by D712 residue (as in WT), and they now are also linked by two different O atoms belonging to the P_{α} group (the same O atom is linked in WT). Additionally, in Figure 8D, both octahedra are linked together by two O atoms belonging to P_{γ} (there is no coordination of this group to the Mg^{2+} A ion in WT). In Figure 8C, D869 residue is found coordinating the Mg^{2+} A ion twice, while only one contact with the same ion is found in Figure 8D. As a result of all these structural changes, the Mg^{2+} A ion in both configurations loses its coordination with $O_{3'}$ (illustrated in Figure 7B), and relaxes the values of the $O_{3'}$-P_{α} distance with respect to the WT (illustrated in Figure 7A). These two characteristic features of the D868A complex, according to the two-metal-ion mechanism, might be more than sufficient to prevent the catalytic reaction from starting. The Mg^{2+} A ion would not be able to reduce pKa of 3' hydroxyl, as well as the larger distance $O_{3'}$-P_{α} might not be the appropriate one to start the reaction. Adding to these constraints, $O_{3'}$ is found in Figure 8C,D forming a H bond with the $O_{5'}$ bridge atom (H bond not present in the WT). Finally, the long distance between the Mg^{2+} ions (illustrated in Figure 7C) is a consequence of the fact that both octahedra do no longer share a common vertex.

Finally, when the D869A mutation is present, the only change compared with WT is the replacement by a water molecule of the
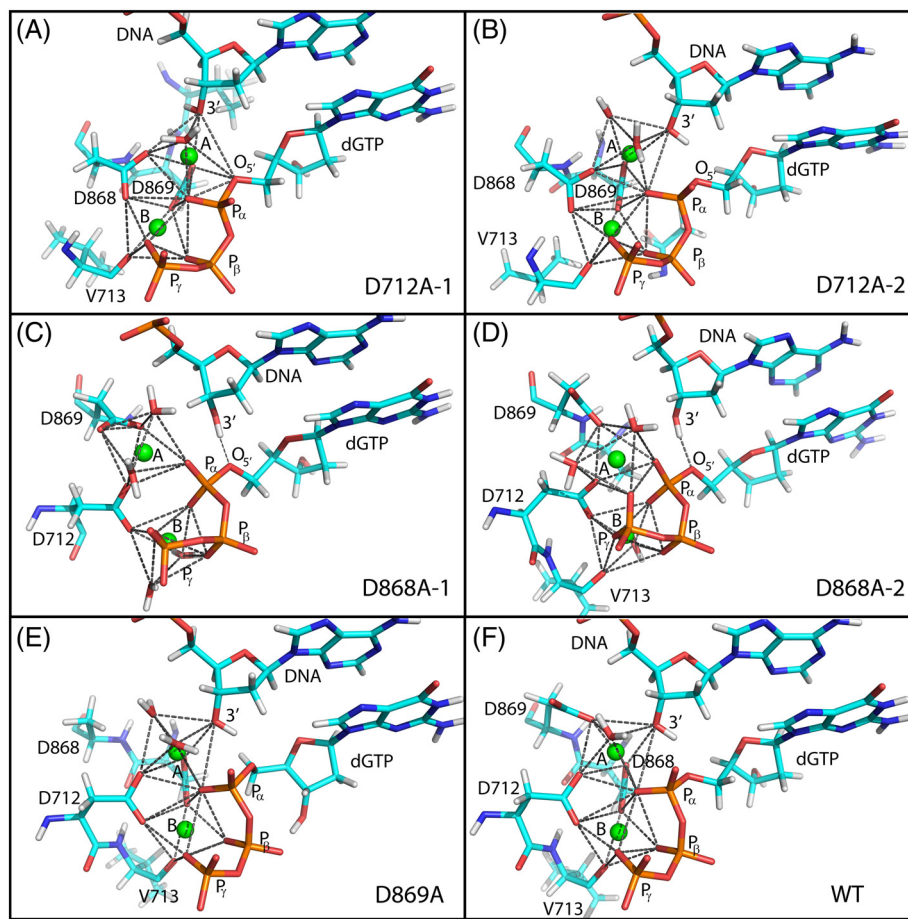
FIGURE 8 Octahedral coordinations for both Mg^{2+} A and B ions (green spheres in the online figure), in C-I, for representative configurations of D712A (A) and (B); D868A (C) and (D); D869A (E); and WT (F) complexes. Figures were generated after a minimization of energy, performed at the end of 200 ns of MD simulations for each system (except for Figure 8A, where a minimization after 50 ns was used). Figures C and D were generated at the end of two independent 200 ns runs of MD simulations for the D868A complex. All the other figures correspond to the end of the first 200 ns MD simulation runs, also depicted in Figures 2–5. To contribute to the interpretation of this figure, the atomic coordinations for both Mg^{2+} ions and for all complexes presented are collected in the Supporting Information Table SI. Figures were obtained using the PyMOL software[47] [Color figure can be viewed at wileyonlinelibrary.com]

coordination lost due to this mutation (only affecting Mg^{2+} A octahedron). All other coordinated atoms remain as in the WT and the complex resembles the WT (illustrated in Figure 7). Therefore, from the limited sampling of the conformational space we made (with our three independent MD runs), we have no evidence by which the D869A mutation could stop the catalytic activity of hTERT. However, it is experimentally known that this mutation abolishes the catalytic activity in human telomerase.[17–19] Moreover, in the WT complex, D869 conserved residue makes contact only with the Mg^{2+} A ion (Figure 2A) and precisely this ion is the one that remains inside the catalytic pocket after the chemical reaction (Figure 2C). Comparing Figure 2A,C, some insight may be obtained about how elongation may proceed. A concerted movement of some residues, as well as the DNA/RNA hybrid, should move the Mg A octahedron (in Figure 2C), in such a way that the coordination with the O of the old P_α is lost while coordinating the new $O_{3'}$ of the primer (as in Figure 2A), therefore the reaction can continue. Thus, it could be speculated that a possible mechanism for this mutation to stop the activity might be the participation of D869 residue in the single nucleotide translocation mechanism. When mutated to Ala the one nucleotide translocation step might not be completed (Mg A does not reach to coordinate the

new $O_{3'}$) and the reaction might stop. A full hTERT model and larger samplings of the conformational space are necessary to clarify this issue.

5 | CONCLUSIONS

This work addresses the computational study of single amino acid substitutions on several conserved residues that abolish the catalytic activity of human telomerase protein. We focused on the conserved residues inside the RT domain of hTERT. In particular, we studied K902N, R631Q, D712A, D868A, and D869A mutations, and used the two-metal-ion mechanism to identify atomic deviations from the WT whereby they are able to cancel telomerase activity. The conclusions of this study are: (i) we are able to identify the possible mechanisms involved in these mutations to abolish human telomerase activity, except for the D869A mutation. (ii) Mutations K902N and R631Q might abolish the catalytic activity because they prevent the exit of PPi from the catalytic pocket. We also show evidence that probably the same holds for K902A, R631A, and K626A mutants. (iii) The D712A mutation might not allow the addition reaction of a single

nucleotide to even begin, since the intermediate state of pentavalent transition could not be formed. In other words, the complex including this mutation might not able to accommodate the incoming dNTP at the correct position corresponding to the catalytic site. (iv) The D868A mutation might also inhibit the initiation of the single nucleotide reaction, because the $O_{3'}$ atom does no longer coordinate the Mg^{2+} A ion. (v) From this study, we have no evidence on the mechanism by which the D869A mutation suppresses the catalytic activity, as it is experimentally found. Larger samplings of the conformational space of the complex containing this mutation are necessary to understand why D869A mutation abolishes the catalytic activity of human telomerase.

ORCID

Silvano J. Sferco (iD) https://orcid.org/0000-0002-3736-0945

REFERENCES

1. Wu RA, Upton HE, Vogan JM, Collins K. Telomerase mechanism of telomere synthesis. *Annu Rev Biochem*. 2017;86:439-460.
2. Kim NW, Piatyszek MA, Prowse KR, et al. Specific association of human telomerase activity with immortal cells and cancer. *Science*. 1994;266:2011-2015.
3. Geserick C, Blasco MA. Novel roles for telomerase in aging. *Mech Ageing Dev*. 2006;127:579-583.
4. Aubert G, Lansdorp PM. Telomeres and aging. *Physiol Rev*. 2008;88: 557-579.
5. Shay JW, Bacchetti S. A survey of telomerase activity in human cancer. *Eur J Cancer*. 1997;33:787-791.
6. Harley CB. Telomerase and cancer therapeutics. *Nat Rev Cancer*. 2008;8:167-179.
7. Lingner J, Hughes TR, Shevchenko A, Mann M, Lundblad V, Cech TR. Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science*. 1997;276:561-567.
8. Nakamura TM, Morin GB, Chapman KB, et al. Telomerase catalytic subunit homologs from fission yeast and human. *Science*. 1997;277: 955-959.
9. Gillis AJ, Schuller AP, Skordalakes E. Structure of the Tribolium castaneum telomerase catalytic subunit TERT. *Nature*. 2008;455:633-637.
10. 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*. 2010;17:513-518.
11. Nguyen THD, Tam J, Wu RA, et al. Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature*. 2018;557: 190-195.
12. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 a resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science*. 1992;256:1783-1790.

13. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*. 1998;282:1669-1675.
14. Steitz TA. A mechanism for all polymerases. *Nature*. 1998;391: 231-232.
15. Steitz TA. DNA polymerases: structural diversity and common mechanisms. *J Biol Chem*. 1999;274:17395-17398.
16. Sekaran VG, Soares J, Jarstfer MB. Structures of telomerase subunits provide functional insights. *Biochim Biophys Acta*. 1804;2010: 1190-1201.
17. Harrington L, Zhou W, McPhail T, et al. Human telomerase contains evolutionarily conserved catalytic and structural subunits. *Genes Dev*. 1997;11:3109-3115.
18. Beattie TL, Zhou W, Robinson MO, Harrington L. Reconstitution of human telomerase activity in vitro. *Curr Biol*. 1998;8:177-180.
19. Weinrich SL, Pruzan R, Ma L, et al. Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nat Genet*. 1997;17:498-502.
20. Armanios M, Chen J-L, Chang Y-PC, et al. Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc Natl Acad Sci USA*. 2005;102: 15960-15964.
21. Basel-Vanagaite L, Dokal I, Tamary H, et al. Expanding the clinical phenotype of autosomal dominant dyskeratosis congenita caused by TERT mutations. *Haematologica*. 2008;93:943-944.
22. Kirwan M, Vulliamy T, Marrone A, et al. Defining the pathogenic role of telomerase mutations in myelodysplastic syndrome and acute myeloid leukemia. *Hum Mutat*. 2009;30:1567-1573.
23. Diaz de Leon A, Cronkhite JT, Katzenstein A-LA, Godwin JD, Raghu G, Glazer CS, Rosenblatt RL, Girod CE, Garrity ER, Xing C, Garcia CK. Telomere lengths, pulmonary fibrosis and telomerase (TERT) mutations. *PLoS One* 2010;5:e10680, Telomere lengths, pulmonary fibrosis and telomerase (TERT) mutations.
24. Steczkiewicz K, Zimmermann MT, Kurcinski M, 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*. 2011; 108:9443-9448.
25. Wallace IM, O'Sullivan O, Higgins DG, Notredame C. M-coffee: combining multiple sequence alignment methods with T-coffee. *Nucleic Acids Res*. 2006;34:1692-1699.
26. Soding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res*. 2005;33:W244-W248.
27. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. 2008;9:40.
28. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*. 2010;5: 725-738.
29. Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J*. 2011;101:2525-2534.
30. Benkert P, Künzli M, Schwede T. QMEAN server for protein model quality estimation. *Nucleic Acids Res*. 2009;37:W510-W514.
31. Case DA, Darden TA, Cheatham III TE, et al. AMBER 12, 2012, University of California, San Francisco.
32. Pearlman DA, Case DA, Caldwell JW, et al. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput Phys Commun*. 1995;91:1-41.
33. Case DA, Cheatham TE, Darden T, et al. The amber biomolecular simulation programs. *J Comput Chem*. 2005;26:1668-1688.
34. Salomon-Ferrer R, Case DA, Walker RC. An overview of the amber biomolecular simulation package. *Wiley Interdiscip Rev Comput Mol Sci*. 2013;3:198-210.
35. Allnér O, Nilsson L, Villa A. Magnesium ion–water coordination and exchange in biomolecular simulations. *J Chem Theory Comput*. 2012;8: 1493-1502.
36. Meagher KL, Redman LT, Carlson HA. Development of polyphosphate parameters for use with the AMBER force field. *J Comput Chem*. 2003;24:1016-1025.

37. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 1984;81:3684-3690.

38. Ryckaert J-P, Ciccotti G, Berendsen HJ. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*. 1977;23:327-341.

39. Price DJ, Brooks CL. A modified TIP3P water potential for simulation with Ewald summation. *J Chem Phys*. 2004;121:10096-10103.

40. Miyamoto S, Kollman PA. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem*. 1992; 13:952-962.

41. Darden T, York D, Pedersen L. Particle mesh Ewald: an N·log (N) method for Ewald sums in large systems. *J Chem Phys*. 1993;98: 10089-10092.

42. Case DA, Berryman JT, Betz RM, et al. *AMBERTOOLS 12*. 2012 San Francisco, CA: University of California.

43. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins Struct Funct Genet*. 2004;55:383-394.

44. Grubmüller H, Heymann B, Tavan P. Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*. 1996;271:997-999.

45. Marrone A, Walne A, Tamary H, et al. Telomerase reverse-transcriptase homozygous mutations in autosomal recessive dyskeratosis congenita and Hoyeraal-Hreidarsson syndrome. *Blood*. 2007;110:4198-4205.

46. Nakamura T, Zhao Y, Yamagata Y, Hua Y, Yang W. Watching DNA polymerase η make a phosphodiester bond. *Nature*. 2012;487:196-201.

47. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger L. PyMOL.

48. Chambers JM, Cleveland WS. Kleiner B, et al. *Graphical Methods For Data Analysis*. Wadsworth, 1983.

49. Murrell P. R graphics. (Chapman & Hall/CRC, 2006).

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.