J Mol Model (2012) 18:3035–3049
DOI 10.1007/s00894-011-1313-z

ORIGINAL PAPER

Theoretical prediction of the binding free energy for mutants of replication protein A

Claudio Carra · Janapriya Saha · Francis A. Cucinotta

Received: 1 November 2011 / Accepted: 16 November 2011 / Published online: 10 December 2011

Abstract The replication protein A (RPA) is a heterotrimeric (70, 32, and 14 kDa subunits), single stranded DNA (ssDNA) binding protein required for pivotal functions in the cell metabolism, such as chromosomal replication, prevention of hairpin formation, DNA repair and recombination, and signaling after DNA damage. Studies based on deletions and mutations have identified the high affinity ssDNA binding domains in the 70 kDa subunit of RPA, regions A and B. Individually, the domain A and B have a low affinity for ssDNA, while tandems composed of AA, AB, BB, and BA sequences bind the ssDNA with moderate to high affinity. Single and double point mutations on polar residues in the binding domains leads to a reduction in affinity of RPA for ssDNA, in particular when two hydrophilic residues are involved. In view of these results, we performed a study based on molecular dynamics simulation aimed to reproduce the experimental change in binding free energy, ΔΔG, of RPA70 mutants to further elucidate the nature of the protein-ssDNA interaction. The MM-PB(GB)SA methods implemented in Amber10 and the code FoldX were used to estimate the binding free energy. The theoretical and experimental ΔΔG values correlate better when the results are obtained by MM-PBSA calculated on individual trajectories for each mutant. In these conditions, the correlation coefficient between experimental and theoretical ΔΔG reaches a value of 0.95 despite the overestimation of the energy change by one order of magnitude. The decomposition of the MM-GBSA energy per residue allows us to correlate the change of the affinity with the residue polarity and energy contribution to the binding. The method revealed reliable predictions of the change in the affinity in function of mutations, and can be used to identify new mutants with distinct binding properties.

Keywords Amber · Binding · MM-PBSA · Molecular dynamics · Mutation · Replication protein · RPA

C. Carra (✉) · J. Saha
Universities Space Research Association,
Columbia, MD, USA
e-mail: claudio.carra-1@nasa.gov

F. A. Cucinotta
NASA-JSC Space Radiation Health Project,
2101 NASA Parkway,
Houston, TX 77058, USA

Introduction

The eukaryotic replication protein A (RPA), a heterotrimeric ssDNA-binding protein, plays a major role in cell activities, such as chromosomal replication, repair processes and recombination pathways. More recently, RPA was shown to have functions in signaling following DNA damage [1–9]. In the DNA repair process, RPA contributes to the organization of the sequential assembly of DNA processing proteins along the ssDNA [10, 11]. RPA binds to ssDNA with high affinity and has a specific role in protein-protein interaction [3, 8], in particular, in the early enzymatic stages of nucleotide excision repair (NER) and DNA replication [12, 13]. In the latter, RPA promotes the priming of the polymerase α/primase to the DNA and stabilizes the ssDNA formed from the strand separation catalyzed by helicase [14]. In the former, RPA participates at the first step of the repair process, before the incision and possibly plays a crucial role in the recognition of DNA damage [15–17]. The interaction of RPA with duplex DNA is not fully elucidated, but the affinity is lower, about three orders of magnitude with respect to ssDNA, with some variations due to experimental conditions [8, 18]. The preferential binding behavior of RPA is toward a damaged double stranded DNA, dsDNA, rather than an undamaged chain, and the affinity is a function of the structure of damaged DNA; in particular, the binding is more favorable where disruptive damages lead to a ssDNA formation [19–21]. Individual substructures of RPA have been studied, however the entire geometry of the full length of the protein is still missing even with some interesting attempts to shed light on the overall structure [22, 23].

The human homologue subunits of RPA is composed by three subunits named according to their molecular weight, RPA70, RPA32 and RPA14 [3, 8, 24, 25]. The X-ray structure of the core DNA binding domain, RPA70AB, has been solved in complex with dC_8, revealing OB (oligonucleotide/oligo-saccharide-binding) domains folded similarly to the corresponding OB domains identified in other known ssDNA-binding proteins (SSBs) [2, 26, 27]. RPA is composed of five ssDNA binding units, denoted as A, B, C, D and F, in order of decreasing affinity. The first three are located in the subunit RPA70 linked together by a flexible peptide chain, F is connected to the domain A by a ~60 amino acid flexible linker [28], and the remaining binding domain, D, is found in RPA32 subunit. The binding modes and affinities depend on the length of the substrate, with 8–10, 12–23 and 28–30 nucleotides (nt) respectively [29–31]. The role of the N terminus of RPA70 is more elusive. Despite some affinity to ssDNA, its contribution to the overall binding is hypothesized to be mostly regulatory [28, 32, 33]. Due to the presence of several binding modes, RPA can assume different binding conformations, as suggested by scanning transmission electron microscopy and gel filtration experiments [34].

The X-ray structure of the binding site reveals 18 polar and non-polar residues, located in L12 and L45 region of the protein, which interact with the phosphate backbone. The high number of polar contacts is rather puzzling in view of the very low binding selectivity that RPA has for the ssDNA base sequence [35]. The non-specific binding characteristic can be justified as a dynamic remodeling of the binding sites occurring upon the formation of the complex [36]. The molecular mechanism describing the RPA binding to DNA is still under debate. Although the residues interacting with the 8-mer oligomer have been identified by X-ray studies, the contribution of individual amino acids to the binding is still unclear. The mechanism proposed for the binding predicts a first step initiated by RPA70AB, which binds the first 8–10 nt of a ssDNA with polarity from 5' to 3' order. The individual affinity of each subunit for the ssDNA is rather small, with greatest values measured for RPA70A where the dissociation equilibrium constant, K_D, is estimated to be ~2 μM. In the following step, RPA70B, separated from RPA70A by a short linkage, binds the ssDNA, with an overall affinity enhanced by several orders of magnitude [35]. The binding progressively involves all of the remaining subunits, reaching a high affinity binding mode that forms a stable complex with a ~30 nt oligomer [29, 30, 35, 37–39].

The linker between the two major binding domains, RPA70A and RPA70B, is expected to be flexible, as independently confirmed by NMR studies on RPA70AB in solution [35]. RPA70A has an affinity for a variety of proteins, including the papilloma virus E1 helicase [40], the SV40T antigen [40, 41], XPA [42], and the human Rad51 recombinase [43]. Mutations in Rad51 show that the interaction between the N-terminus of Rad51 with RPA70A plays a crucial role in the RPA displacement from ssDNA leading to the formation of Rad51-ssDNA nucleoprotein filament [43]. Several DNA damage signaling and processing proteins revealed an interaction with RPA; Rad17 [44, 45], Rad9 [46], ATRIP [9, 47, 48], 53BPI [49], BRCA2 [50], Mre11-Rad50-Nbs1 (MRN complex) [51], and nucleolin [52, 53].

RPA binding shows a low cooperativity and its strength is mostly a function of the length of the substrate, with minor variation for changes in the nucleotide sequence and experimental conditions, even though a 50-fold preference is observed for polypyrimidine tracks [54–57]. Wold et al. [39, 58, 59] provided an extensive study of correlation between the change of the association equilibrium constant, K_A, of a series of mutations and truncations of the entire RPA protein. Due to the similarity of the system studied with the solved crystal structure of RPA70AB interacting with a octamer ssDNA, we focused our investigation on the point mutations that Wold et al. performed on the RPA70 [39] to evaluate the major residue contributions to the binding. Experimentally, mutations on polar residues gave the highest loss in affinity, but also non-polar residues alter the K_A considerably. As a consequence, the binding is not a function of only few residues but of a combination of multiple interactions [39].

The goal of this paper is to find a correlation between experimental changes of RPA affinity in function of mutations with the theoretical prediction of the calculated values for the corresponding systems. Moreover, the study aims to identify the residues contributing most to the binding and their energy contributions to the overall values. Our results are in good agreement with experimental data, in particular the correlation coefficient between the experimental and the theoretical values relative to the change in binding free energies in function of point mutations. The agreement confirms that the hydrophilic residues have a major contribution in the binding and contribute because of a large van der Walls (vdW) energy component. The optimal agreement is reached when the MM-PBSA technique is used on single trajectories. Alanine scan methods gave a lower correlation with the experimental data.

Materials and methods

The geometry of the of RPA70AB interacting with dC_8, RPA-C8, was constructed by considering part of the crystal structure determined by Bochkarev et al. [60] (pdb id 1jmc). Before starting the simulation, the X-ray structures were optimized using the WHAT IF series of programs [61]. The molecular dynamics simulations and the data analysis were carried out with the AMBER 10 [62] package using the parmbsc0 force field [63]. The solute was modeled in a periodic box with a 8 Å buffer of water molecules explicitly described by the TIP3P [64] model. The particle mesh Ewald (PME) method [65] was used to treat the long range electrostatic interactions.

The equilibration of RPA-C8 and the corresponding mutants was performed as follows. An initial optimization of 20,000 cycles, the first 10,000 by steepest descent then followed by the conjugate gradient method. The complex is constrained to relax the solvent. Then a further optimization of 30,000 cycles with no constraints on the whole system was carried out to lead to a final relaxed geometry. The first equilibration was carried out with a weak restraint on the complex for 100 ps at constant volume, constantly increasing the temperature from 0 to 300 K. The equilibration continued for 200 ps at a constant pressure of 1 atm, by keeping the temperature constant with the Langevin temperature equilibration scheme [66] using a collision frequency of 1.0 ps^{-1}. Under these conditions, the restraints were gradually removed. The production run was carried out without restraints for 24 ns. During the MD calculation, hydrogen stretching motions were removed using SHAKE bond constraints [67], allowing a longer time step of 2 fs without introducing any instability. Free binding energies of the complex were calculated with the molecular mechanics Poisson Boltzmann and generalized Born surface area method, MM-BP(GB)SA. The binding free energy, ΔG, is calculated by taking the average energy difference between the complex, $\overline{G}_{comp}$, and the reactants (receptor and ligand),$\overline{G}_{rec} + \overline{G}_{lig}$.

$$\Delta G = \overline{G}_{comp} - \left(\overline{G}_{rec} + \overline{G}_{lig}\right) \tag{1}$$

where the average free energy $\overline{G}$, for the complex, RPA70AB-dC_n, receptor, RPA70AB, and ligand, dC_n, is composed by:

$$\overline{G} = E_{MM} + G_{PB/GB} + G_{SA(np)} - T\Delta S \tag{2}$$

where E_{MM} is the molecular mechanics interaction energy in "gas phase" within the system, $G_{PB/GB}$ is the component of the electrostatic energy calculated with the Poisson-Boltzmann (PB) [68], or generalized Born (GB) [69] method. One of the advantages of using the MM-PB(GB)SA method is that the "nonphysical" annihilation [70, 71] or decoupling [72, 73] of the species alone in solution or bounded to a substrate is not required anymore. Moreover, it is not necessary to model the partially unbound states as demanded using the umbrella sampling. It has already been shown that the MM-PB(GB)SA approach was able to reproduce well qualitatively the binding free energy of such systems [74–76]. Amongst the MM-GBSA methods available, we considered the approach of Onufriev et al., IGB2 [77]. The G_{SA} represents the non-polar contribution to the solvation free energy which is determined with solvent-accessible-surface-area-dependent terms (SA) [78] approach. The term $T\Delta S$ is the conformational entropy change of the solute. The grid size used to solve the Poisson-Boltzmann equation was 0.5 Å, and the values of interior dielectric constant and exterior dielectric constant were set to 1 and 80, respectively. The gas phase and the solvation free energies were calculated over 400 snapshots taken at 20 ps interval from the last 8 ns of the MD trajectories. To get closer to the experimental conditions, the concentration of the salt in the bulk of the solution was set to 50 mM. The solute binding entropy was calculated by normal mode, NM, analysis where the standard state is assumed to be at 1 M [79].

The relative binding free energy, ΔΔG, was calculated as a difference between the binding free energy of the wild type system, RPA-C8, and the mutated structure. Mutations were generated by FoldX [80, 81] and each geometry was "repaired" according to FoldX protocol. A series of calculations were performed to estimate the binding free energy using the MM-PBSA and MM-GBSA methods, ΔG^{PB}, and ΔG^{GB}, respectively. The binding entropy, $-T\Delta S$, was calculated by normal mode analysis and added to the MM-PB (GB)SA energies to obtain the standard binding free energy, $(PB)G_{tot}=\Delta G^{PB} - T\Delta S$, and $(GB)G_{tot}=\Delta G^{GB} - T\Delta S$, respectively.

A different approach, still based on a MM-PB(GB)SA calculations, extracts the snapshots from the RPA−C8 trajectory and replaces the selected residues with alanine. The applet ALSCAN implemented in Amber 10 will be used for this purpose. The last approach involves the use of FoldX [81, 82] for the estimation of the binding free energy for RPA-C8 and mutants.

Results

The model of the complexes of RPA70AB interacting with dC_8, RPA-C8, was constructed on the crystal structure of the active complex determined by Bochkarev et al. [60] (pdb id: 1jmc). We considered several single and double mutations on the binding loop, Fig. 1a, according to the work of Wold et al. [39]. As a comparison, we analyzed the change in binding free energy for all the alanine mutations with all residues in contact with the ssDNA, Fig. 1b.

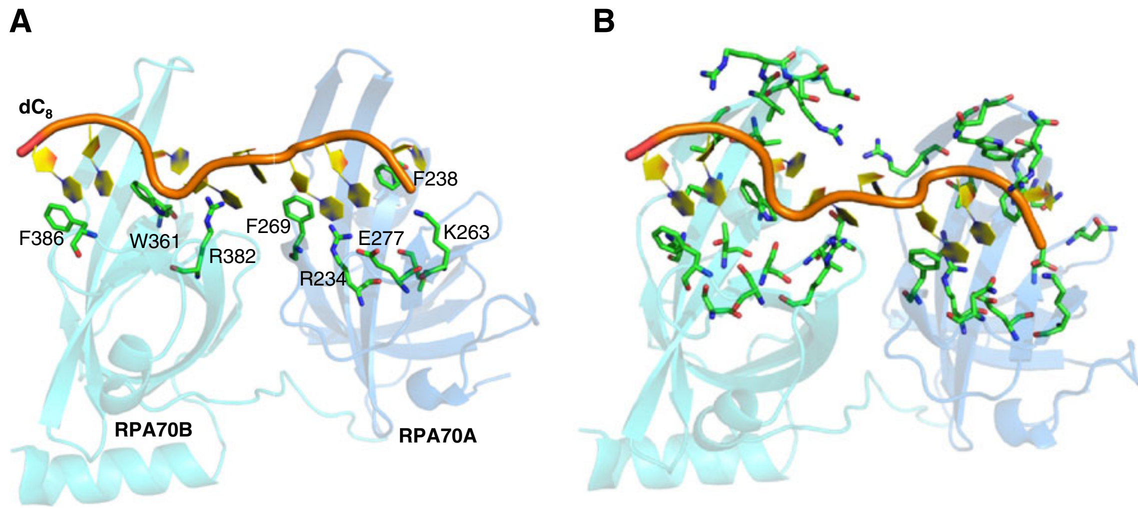

Fig. 1 X-ray structure for the structure used in the modeling, representing RPA70AB binding to the oligomer dC_8. (**a**) The position of the point mutations considered, (**b**) the contact points of the protein to the ligand, derived with the FoldX code

The data reported in Fig. 2 represent the correlation between the relative binding free energy, ΔΔG, experimentally measured and theoretically determined by alanine scan. The experimental ΔΔG were calculated from the values of the association equilibrium constants, K_A [39], according to the following expression:

$$\Delta\Delta G_{\exp} = \mathrm{RT}\, ln\left(K_{A,wt}/K_{A,Ala}\right) \tag{3}$$

where $K_{A,wt}/K_{A,Ala}$ is the ratio between the association rate constant of the wild RPA, $K_{A,wt}$, and the association equilibrium constant of the alanine mutated RPA, $K_{A,Ala}$. The theoretical prediction of the ΔΔG was based on the alanine scan performed by the ALSCAN applet in Amber 10. The resulting values are overestimated with respect to experimental results. The slope of the liner regression between the theoretical and experimental prediction of ΔΔG, is 9.4 and 7.8 for ΔG^{PB} and ΔG^{GB} respectively. The correlation coefficient, R, is within the range of the accuracy of the method [83], 0.71 for ΔG^{PB} and 0.77 for ΔG^{GB} respectively. As a comparison, the corresponding values of ΔΔG were calculated with FoldX. The estimation of the change in binding free energy is much closer to the experimental values, with a slope of the linear interpolation of 0.89; however, the R value is significantly lower, 0.44.

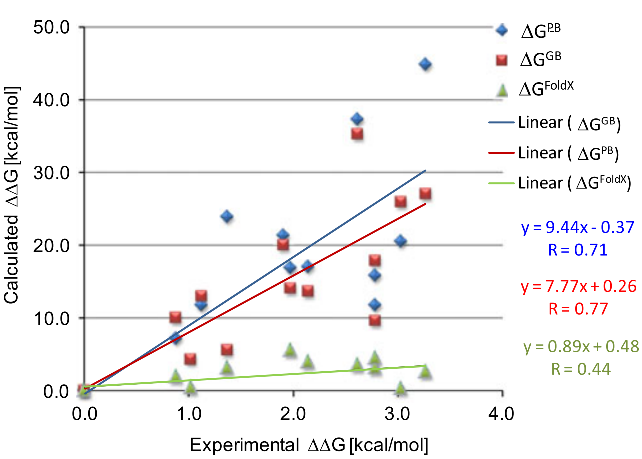

Fig. 2 The ΔΔG analysis based on the alanine scan method for the experimental single and double mutations. The data are calculated by MM-PBSA, in blue, and mm-GBSA in red. The experimental geometry is used for the same analysis performed with the FoldX code, in green. The linear interpolations for each set of data are shown, with the corresponding values for the slope and R^2

To increase the accuracy of the predictions, the individual modeling of each mutation was examined. This approach improves considerably the agreement with experiments. Even though the overestimation is the same order of magnitude of the previous analysis, the slope of the linear interpolation is 11.9 and 9.46 for ΔG^{PB} and ΔG^{GB} respectively and the R value is higher, 0.88 for ΔG^{PB} and 0.78 for ΔG^{GB}, Fig. 3a. The data calculated by MM-PBSA show a better correlation with experimental values as compared to the calculations with MM-GBSA.

The values of the binding entropy, $-T\Delta S$, are calculated for each system by averaging the results on 20 snapshots chosen in an equidistant interval from the last 8 ns of the simulation. The term is then subtracted to the $\Delta G^{PB/GB}$ values. The resulting estimation, however, seems to diverge for the ΔΔG calculated without the $T\Delta S$ contribution, Fig. 3b, showing lower R values for both MM-PBSA and MM-GBSA.

The values that are more spread over the liner interpolation curve have a curious common characteristic: they all refer to the R234 residue. By removing the single and double mutations which include R234, the R values increase appreciably to 0.95 by MM-PBSA and 0.82 by MM-GBSA,

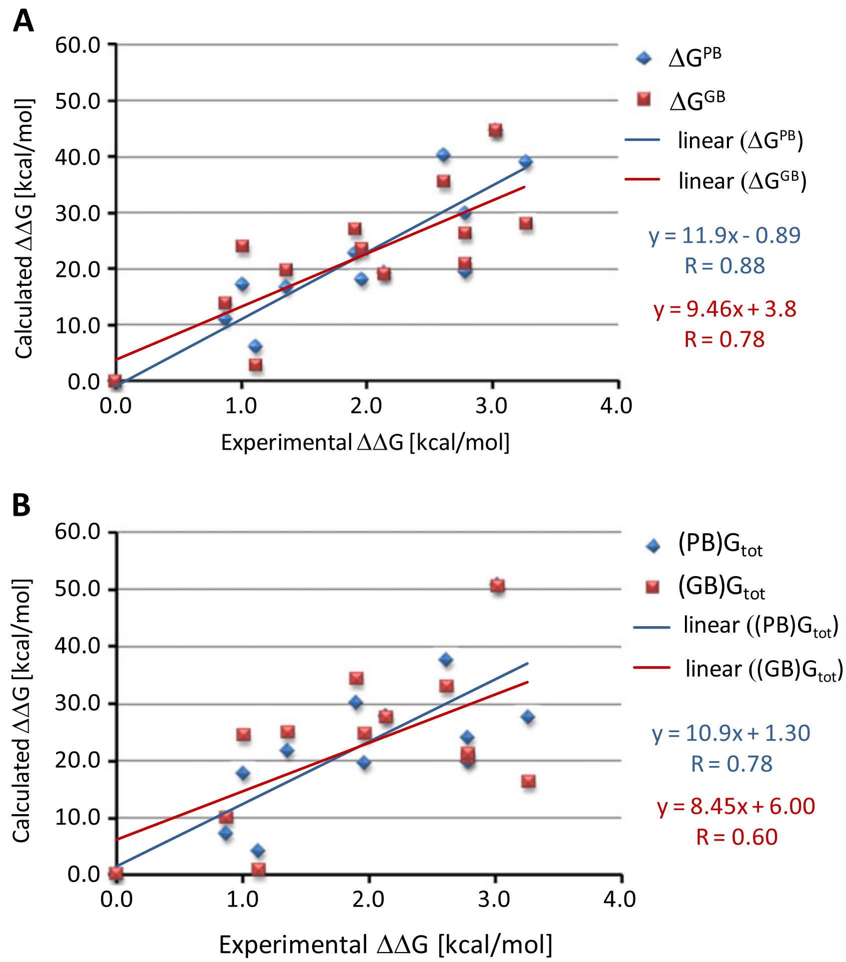

Fig. 3 The ΔΔG analysis based on individual trajectories method for the experimental single and double mutations. The data are calculated by MM-PBSA, in blue, and MM-GBSA in red. (**a**) only the $\Delta G^{PB/GB}$ is used for the ΔΔG estimation, (**b**) the $T\Delta S$ correction is introduced

Fig. 4a. The R234 residue is the only amino acid among the mutations considered by Wold et al., which has two hydrogen bonds toward a base. The R382, on the contrary, reveals a polar interaction with an oxygen atom of the backbone of the ssDNA. Except for R234, all the hydrogen bond interactions in the binding loop of RPA70AB for the mutations considered are not specifically oriented toward the bases, Fig. 4b. E277 shows one hydrogen bond with the adjacent base, but the energy contribution is negligible (see further in the text).

The MM-GBSA energy, ΔG^{GB}, calculated by IGB2 protocol [77], can be decomposed in terms of residue contributions, Fig. 5. The ΔG^{GB} term includes the Coulombic interaction summed with the solvation free energy, $G_{Coul+GB}$, the non-polar contribution, G_{np}, and the van der Waals, G_{vdW} energy. The strongest values are found for the Arg residue, particularly in positions 382, 335, 210, and 216 with contributions of −9.9 kcal mol^{-1}, −8.6 kcal mol^{-1}, −7.2 kcal mol^{-1}, and −6.5 kcal mol^{-1} respectively. The energy contributions for the mutations considered, labeled with a star (*), are quite different and are a function of the polarity of the residue. Arg382 has the highest binding character, dominated by a $G_{Coul+GB}$ term, where the G_{vdW} value is almost negligible. On the contrary, hydrophobic residues such as Phe 238, 269, and 386 have minor binding nature, with ΔG^{GB} values between −3 kcal mol^{-1} and −4 kcal mol^{-1}, but the energy contribution is characterized by a dominating G_{vdW} term. Lys 263 and Glu277, despite their strong polarity, have a minor role in the binding with ΔG^{GB} values close to zero.

To compare the MM-PB(GB)SA results with a different theoretical protocol, the same residues where mutated to alanine with the FoldX program and for each of them the binding free energy, ΔG^{FoldX}, was calculated. In Fig. 6a, for each residue, the components ΔG^{FoldX} are calculated according to the following equation:

$$\Delta G^{FoldX} = \Delta G_{solvH} + \Delta G_{solvP} + \Delta G_{ele} + \Delta G_{vdW} + \Delta G_{hbond} + T\Delta S \quad (4)$$

where ΔG_{solvH} and ΔG_{solvP} are the differences in solvation energy for non-polar and polar groups respectively when these change from the unfolded to the folded state. The ΔG_{ele} is the electrostatic contribution of charged groups including the helix dipole; the ΔG_{vdW} is the sum of the van

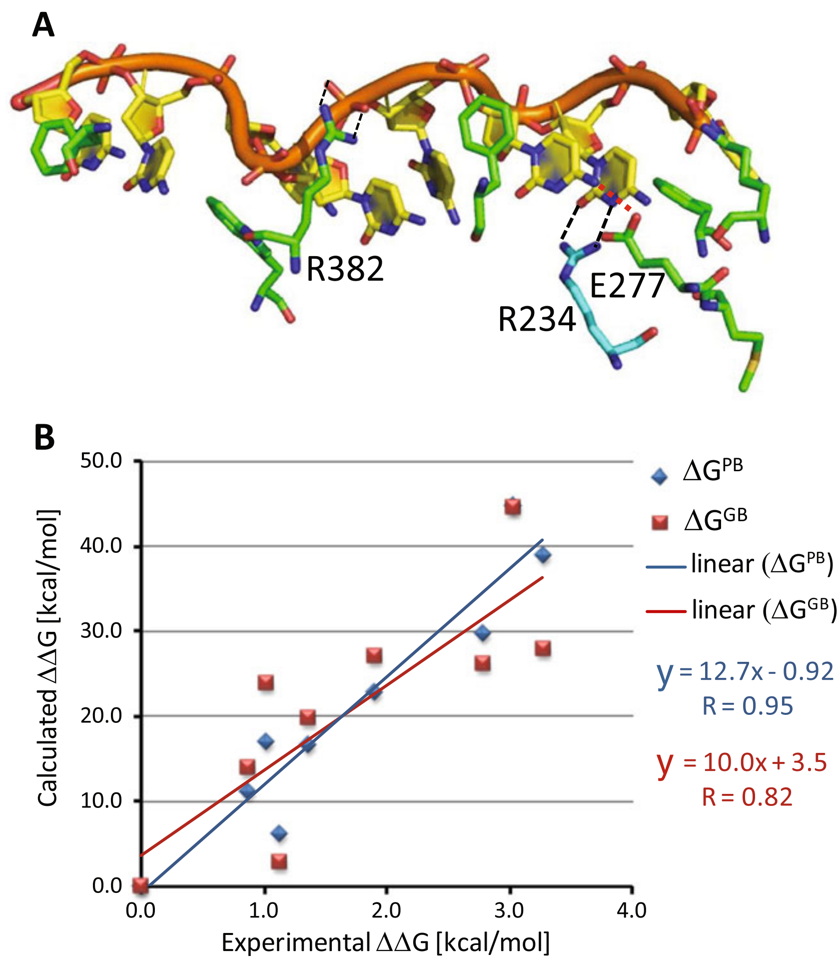

Fig. 4 (**a**) residues considered for the mutations. The labeled ones have a strong polar contact with dC_8: R382 with the backbone; R234 shows two hydrogen bonds with the cytosine, black dashed lines; E277 one hydrogen bond, red dotted line. (**b**) The experimental results versus the calculated ΔΔG binding where the R234 data are removed

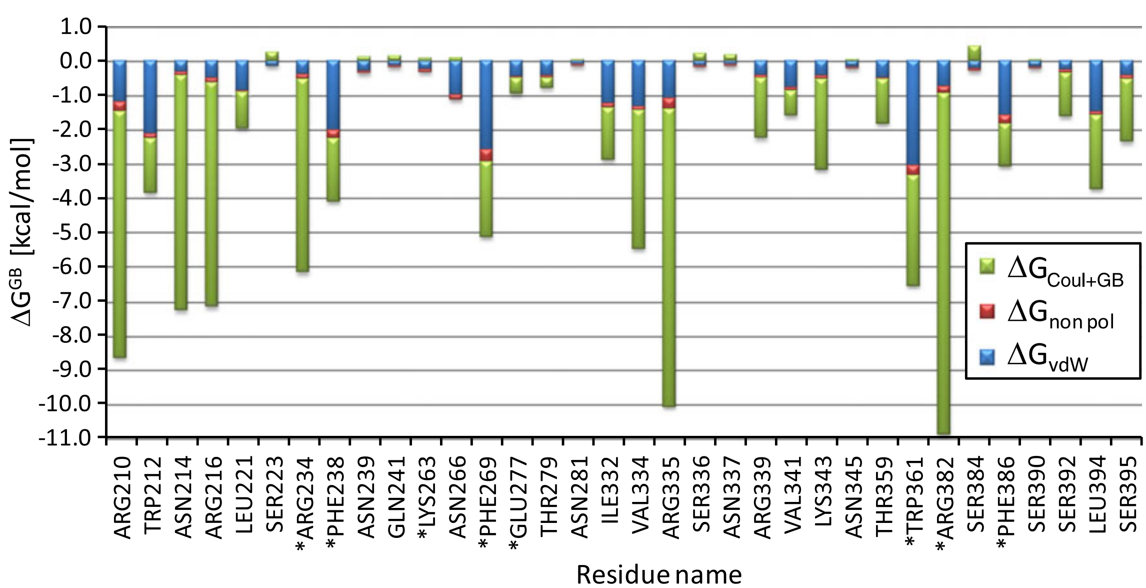

Fig. 5 Decomposition of ΔG^{GB} per residue in contact with dC_8 calculated by MM-GBSA for RPA70AB. The three main energy components are illustrated, van der Waals energy (vdW), sum of Coulombic interactions + polar solvation free energy, and the non-polar (non pol) part of the solvation free energy. The residues labeled with a star (*) are the studied mutations

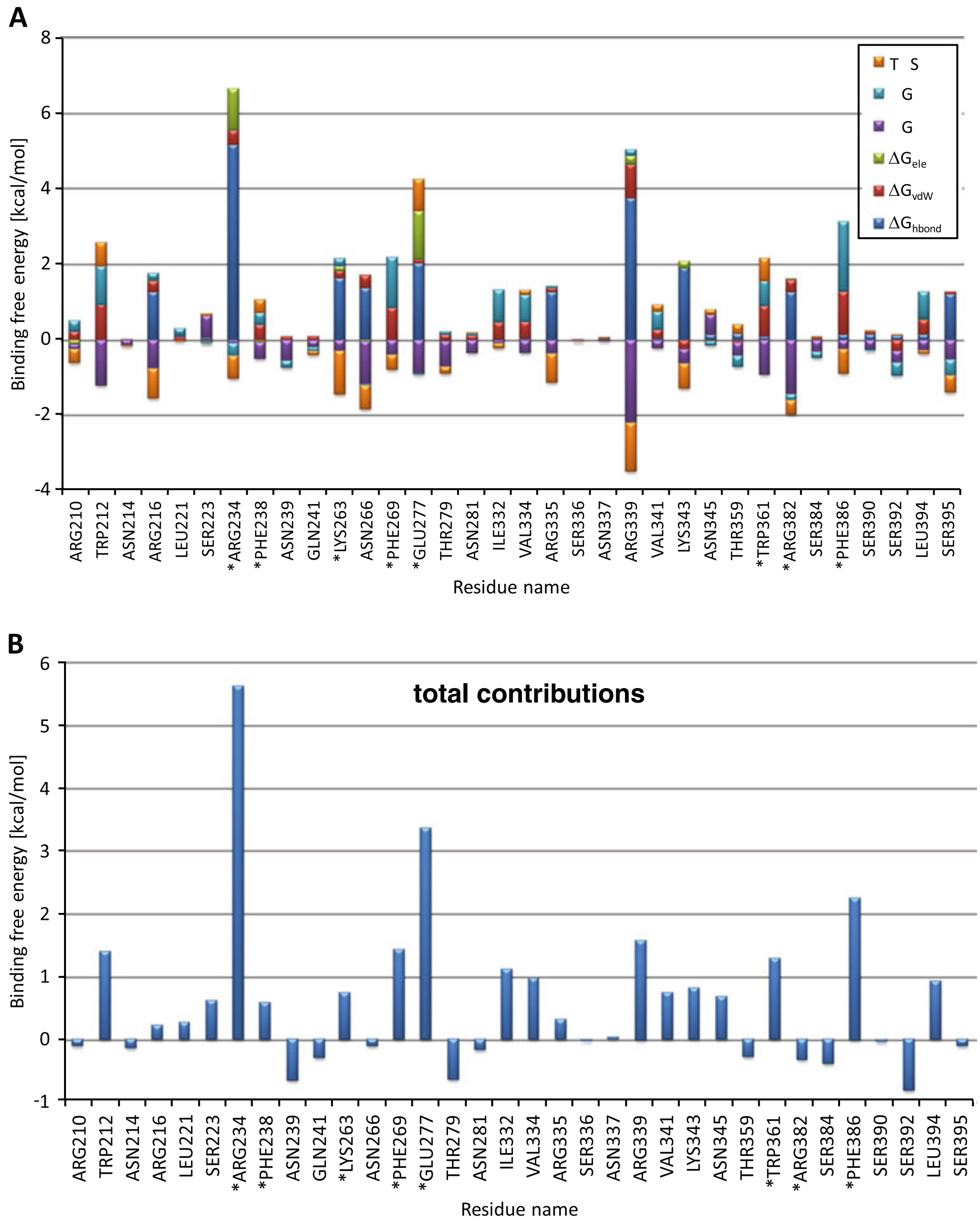

Fig. 6 (**a**) Components of the binding free energy by residues in contact with dC_8, calculated by FoldX. The residues labeled with a star (*) are the studied mutations. (**b**) Sum of energy contributions per residues

der Waals contributions of all atoms with respect to the same interactions with the solvent. The ΔG_{hbond} is the free energy difference between the formations of an intra-molecular hydrogen bond compared to inter-molecular hydrogen bond formation (with solvent). $T\Delta S$ is the sum of the entropy cost of fixing the backbone in the folded state and the entropic cost of fixing a side chain in a particular conformation.

The total difference in binding free energy for the same residues mutated to alanine was also calculated, Fig. 6b. According to the simulation, the highest loss in binding is due to the mutation of Arg in position 234 with a corresponding increase in binding free energy of 5.6 kcal mol^{-1}. A strong loss in $\Delta G^{\mathrm{FoldX}}$is also encountered with the mutation of Glu277, where the $\Delta G^{\mathrm{FoldX}}$ is 3.4 kcal mol^{-1} higher than RPA-C8. Additionally, the position 386 seems critical to preserve the binding affinity; in fact its mutation to Ala increases $\Delta G^{\mathrm{FoldX}}$ by 2.2 kcal mol^{-1}. The alanine scan reveals that there are single point mutations for which the affinity is increased with respect to RPA70AB. In particular, the mutation Ser392 and Asn239 should increase the $\Delta G^{\mathrm{FoldX}}$ by 0.8 kcal mol^{-1} and 0.6 kcal mol^{-1} respectively. In contrast to experimental findings, the alanine replacement in position 382 should lead to a higher affinity with a drop of $\Delta G^{\mathrm{FoldX}}$ value by −0.3 kcal mol^{-1}.

Another aspect examined in this study is the standard binding free energy, $(PB)G_{tot}$ and $(GB)G_{tot}$, and their energy composition, Tables 1, 2, 3, and 4. The two major contributions of the MM-PB(GB)SA with a bonding character are the gas phase Coulombic energy, H_{elec}, and van der Waals energy, H_{vdW}, whose sum is labeled as H_{gas}. In general, the anti-bonding contributions come from the polar solvation component, G_{PB}. Often, the non-polar part, G_{np}, provides a minor contribution. The sum of G_{PB} and G_{np} is defined as polar and non-polar solvation term, G_{solv}.

The mean values of the energy decomposition, including the binding entropy are also shown. The binding character is overestimated, which is a common feature of the single trajectory MM-PB(GB)SA approach. The data relative to free energy contributions calculated for double point mutations are reported in Table 1. The MM-GBSA method overestimates the binding free energy and correlates poorly with the experimental data. Thus, in the following session dedicated to the energy analysis, only the MM-PBSA results will be discussed.

The loss of binding character is particularly considerable for mutations involving polar residues, in particular for K263A-R382A where the ΔG^{PB} value is −82.0 kcal mol^{-1} compared to the RPA-C8, where the ΔG^{PB} is estimated to be −126.7 kcal mol^{-1}. A very similar value is obtained by the equally polar R234A-R382A mutation, −86.6 kcal mol^{-1}, followed by a mutation that involves positively and negatively charged residues, both in direct contact with the ssDNA base, E277A-R382A, with a ΔG^{PB} value of −87.8 kcal mol^{-1}. The loss in binding free energy is substantially lower for hydrophobic mutations, as observed for F238A-F269A and W361A-F386A with a ΔG^{PB} value of −115.7 kcal mol^{-1} and −124.5 kcal mol^{-1} respectively. Intermediate results are reached by K263A-E277A, K263A-R382A, and R234A-E277A, where the ΔG^{PB} values are −97.0 kcal mol^{-1}, −108.6 kcal mol^{-1}, and −107.3 kcal mol^{-1} correspondingly. The components to the binding free energy reveal how the Coulombic energy, H_{elec}, has the major contribution, in particular for non-polar mutations, as encountered for F238A-F269A and W361A-F386A with values over −800 kcal mol^{-1}. The polar solvation free energy has the major antibonding character, which compensates the strong binding character of H_{elec}. The solute binding entropy contributions have similar values for all the mutants, in proximity of −40 kcal mol^{-1} with the exception of E277A-R238A, where $T\Delta S$ reach −55.9 kcal mol^{-1}. Single mutations alter ΔG^{PB} estimations with a lower magnitude with respect to the double mutations, as shown in Table 2.

The position 382 seems extremely critical, and it shows the highest loss in binding with a ΔG^{PB} value of −104.0 kcal mol^{-1}. A similar change is observed for the R234A, where the ΔG^{PB} value is −107.4 kcal mol^{-1}. Moderate loss in binding is encountered for the remaining E277A and K263A where the $\Delta G_{gas+sol}$ is −110.2 kcal mol^{-1} and −109.7 kcal mol^{-1}. The strongest component of $T\Delta S$ is estimated for the mutations R234A, R382A and E277A characterized by the loss of residues having two polar contacts. The lowest value of binding entropy is calculated for K263A where one polar contact is lost. Amongst the components of the free energy, the Coulombic term has a strong negative value for E277A, −936.1 kcal mol^{-1}, in comparison to −832.3 kcal mol^{-1} of RPA70AB, and with R382 which has a value of −454.1 kcal mol^{-1}.

Another approach, alanine scan, was utilized to evaluate the change of binding free energy by alanine replacement on the snapshots derived from the RPA-C8 trajectory. In this approach, all the residues of all mutations preserve the same orientation of the generating structure, thus attributing the change of free energy to only local point mutation. The change of entropy contribution is not considered because it is assumed to be negligible by a single or double residue replacement. Despite the preservation of the protein conformation along the considered mutants, the components of the binding free energy differed widely, Table 3.

The Coulombic and the van der Waals energy results are lower for all mutations with respect to the corresponding cases individually modeled; however, the polar solvation free energy results show a systematic increase. The results calculated by alanine scan, Tables 3–4, show a consistent increase in the ΔG^{PB} and ΔG^{GB} values.

Single residue replacement, Table 4, yields similar results to the double mutation in terms of increase of the binding free energy, however, consistent with the previous case, Table 2, the affinity results are higher with respect to the double mutations. The non-polar free energy contribution is almost unaltered for all cases.

Discussion

The molecular mechanism which regulates the interaction of RPA with ssNDA is poorly understood despite the successful mapping by X-ray and NMR of the core residues of RPA interacting with a fragment of dC_8 [36, 60]. To underline the role of single residues in binding, extensive studies on single and double point mutations of RPA were conducted by Wold et al. [31, 39, 58].

Our study aims to reproduce theoretically the change in binding free energy, $\Delta\Delta G$, and correlate the calculated values to the published experimental results. In the investigation, the MM-PB(GB)SA method was used to estimate the binding free energy, in comparison with the data calculated by FoldX. More recently, the MM-PBSA method has been developed [69, 84] and applied to HIV reverse transcriptase [85], avidin [86], neuraminidase [87], cathespin D [88], growth factor receptor binding protein 2 [89], metalloprotease [90, 91] histone deacetylase [92], DNA glycosylase [76], and protein-protein interface [93].

The mutations examined [39] are shown in the crystal structure of RPA-C8, Fig. 1a. To analyze the role of the binding sites, L12 and L45 located in both RPA70 A and B subunits, mutations on polar and non-polar residues were

Table 1 MM-PB(GB)SA energy components of the double mutants calculated for 800 equidistant snapshots extracted from the last 8 ns of the each trajectory individually modeled. $T\Delta S$ is calculated by normal mode analysis

	E277A-R382A		F238A-F269A		K263A-E277A		K263A-R382A		R234A-E277A		R234A-K263A		R234A-R382A		W361A-F386A	
Contrib	Mean	sem^h	Mean	sem	Mean	sem	Mean	sem	Mean	sem^h	Mean	sem^h	Mean	sem^h	Mean	sem^h
$H_{elec}{}^a$	−573.7	7.9	−817.9	7.9	−692.6	5.3	−266.5	7.4	−753.5	8.1	−459.5	6.4	−340.1	8.0	−858.9	8.6
$H_{vdW}{}^b$	−122.3	1.8	−118.0	1.7	−107.6	1.7	−109.8	1.7	−119.8	1.8	−112.4	1.6	−116.4	1.7	−121.0	1.8
$H_{gas}{}^c$	−695.9	8.4	−935.9	8.0	−800.2	6.0	−376.3	7.7	−873.4	8.9	−572.0	6.8	−456.5	8.3	−979.9	8.8
$G_{np}{}^d$	−17.8	0.1	−18.3	0.1	−16.5	0.1	−15.7	0.1	−17.4	0.1	−16.9	0.1	−17.4	0.1	−18.1	0.1
$G_{PB}{}^e$	626.0	7.4	838.6	7.2	719.6	4.9	310.0	7.0	782.2	7.2	481.5	5.7	387.3	7.2	873.5	7.7
$G_{solv}{}^f$	608.2	7.4	820.2	7.2	703.2	4.9	294.3	6.9	764.8	7.1	464.7	5.7	369.9	7.2	855.4	7.6
$\Delta G^{PB\ g}$	−87.8	3.6	−115.7	3.3	−97.0	3.6	−82.0	3.4	−108.6	3.8	−107.3	3.2	−86.6	3.4	−124.5	3.7
G_{np}	−17.8	0.1	−18.3	0.1	−16.5	0.1	−15.7	0.1	−17.4	0.1	−16.9	0.1	−17.4	0.1	−18.1	0.1
G_{GB}	617.4	7.4	843.9	7.3	717.9	4.8	312.3	6.9	789.9	7.3	485.6	5.8	385.2	7.3	877.6	7.8
$G_{solv}{}'$	599.6	7.4	825.5	7.2	701.4	4.8	296.6	6.9	772.6	7.2	468.7	5.7	367.8	7.3	859.5	7.8
$\Delta G^{GB\ g}$	−96.3	3.4	−110.4	3.1	−98.7	3.3	−79.7	3.4	−100.8	3.6	−103.2	3.1	−88.7	3.2	−120.5	3.5
$T\Delta S^i$	−55.9	3.6	−39.7	3.5	−37.7	4.0	−49.3	3.6	−44.8	3.6	−43.8	3.8	−40.8	3.7	−41.6	3.7
$(PB)G_{tot}{}^l$	−31.9	5.1	−76.0	4.8	−59.3	4.0	−32.7	4.9	−63.8	5.2	−63.5	4.9	−45.8	5.0	−82.9	5.2
$(GB)G_{tot}$	−40.4	4.9	−70.7	4.7	−61.0	5.2	−30.4	4.9	−56.0	5.1	−59.4	4.9	−47.9	4.9	−78.9	5.1

The standard state is assumed to be at 1 M. The units are in kcal mol^{-1} . Ionic strength is set at 50 mM. $^{(a)}$ H_{elec}: coulombic energy; $^{(b)}$ H_{vdW}: van der Waals energy; $^{(c)}$ $H_{gas}=H_{elec}+H_{vdW}$; $^{(d)}$ G_{np} non-polar solvation free energy; $^{(e)}$ G_{PB}: polar solvation free energy calculated by solving the Poisson-Boltzmann equation PB, in italic by solving the generalized Born equation, GB; $^{(f)}$ $G_{solv}=G_{np}+G_{PB/GB}$; $^{(g)}$ $\Delta G^{PB/GB}=H_{gas}+G_{solv}$; $^{(h)}$ Standard error of mean values; The MM-GBSA value are shown in italic; $^{(i)}$ $T\Delta S$: total solute entropy contribution; $^{(l)}$ $(PB/GB)G_{tot}=\Delta G^{PB/GB}-T\Delta S$

Table 2 MM-PB(GB)SA energy components of the single mutant calculated for 800 equidistant snapshots extracted from the last 8 ns of the each trajectory individually modeled. $T\Delta S$ is calculated by normal mode analysis

	RPA-C8		E277A		K263A		R234A		R382A	
Contrib	Mean	sem[h]	Mean	sem	Mean	sem	Mean	sem	Mean	sem
H_{elec}[a]	−832.3	8.2	−936.1	7.7	−588.6	6.9	−640.6	6.6	−454	17.4
H_{vdW}[b]	−129.3	1.6	−122.0	1.7	−109.6	1.8	−118.0	1.6	−117.8	1.7
H_{gas}[c]	−961.6	8.4	−1058.1	7.8	−698.1	7.5	−758.6	7.0	−571.9	8.0
G_{np}[d]	−18.1	0.1	−18.3	0.1	−16.3	0.1	−17.5	0.1	−16.6	0.1
G_{PB}[e]	853.0	7.7	966.2	7.2	604.7	6.1	668.7	5.9	484.6	6.9
G_{solv}[f]	834.9	7.7	947.9	7.1	588.4	6.1	651.2	5.9	468.0	6.8
G^{GB}[g]	−126.7	3.4	−110.2	3.6	−109.7	3.5	−107.4	3.2	−104.0	3.3
G_{np}	−18.1	0.1	−18.3	0.1	−16.3	0.1	−17.5	0.1	−16.6	0.1
G_{GB}	855.5	7.6	972.0	7.0	614.0	6.3	670.9	5.9	491.3	6.9
G_{solv}'	837.4	7.6	953.7	7.0	597.7	6.3	653.4	5.9	474.7	6.8
G^{PB}[g]	−124.2	3.3	−104.4	3.5	−100.4	3.3	−105.2	3.1	−97.2	3.3
$T\Delta S$[i]	−43.5	3.7	−48.6	3.7	−44.1	3.6	−51.9	4.0	−50.7	3.6
(PB)G_{tot}[l]	−83.2	5.1	−61.6	5.2	−65.6	5.0	−55.5	5.2	−53.2	4.9
(GB)G_{tot}	−80.7	4.9	−55.8	5.1	−56.3	4.9	−53.3	5.1	−46.5	4.9

The standard state is assumed to be at 1 M. The units are in kcal mol^{-1} . Ionic strength is set at 50 mM. [(a)] H_{elec}: coulombic energy; [(b)] H_{vdW}: van der Waals energy; [(c)] H_{gas}=H_{elec}+H_{vdW}; [(d)] G_{np} non-polar solvation free energy; [(e)] G_{PB}: polar solvation free energy calculated by solving the Poisson-Boltzmann equation PB, in italic by solving the generalized Born equation, GB; [(f)] G_{solv}=G_{np}+$G_{PB/GB}$; [(g)] $\Delta G^{PB/GB}$ =H_{gas}+G_{solv}; [(h)] Standard error of mean values; The MM-GBSA value are shown in italic; [(i)] $T\Delta S$: total solute entropy contribution; [(l)] (PB/GB)G_{tot}=$\Delta G^{PB/GB}$ – $T\Delta S$

performed and tested. It has been established that point mutations on either the subunit A or B cause a reduction in the affinity. Our investigation examined some of the mutations studied by Wold et al. involving single and double point mutations at 234, 263, 277 and 382 for the polar residues and 361, 386, 238 and 269 for the non-polar interaction. Moreover, we analyzed the energy contribution of all the residues interacting with the ssDNA, Fig. 1a, to identify if additional residues, not yet experimentally considered, might play a critical role in the binding. The theoretical model, however, has some differences compared to the experimental system. In our approach, the interacting oligomer dT_{30} is replaced with the dC_8 present in the X-ray structure. This approximation is justified by experimental studies performed on different oligomers such as dT_n [54] and dC_n [94] revealing a mild selectivity of RPA toward the nature of the ssDNA. In addition, when shorter oligomers like dC_8 and dT_{10} were used, the corresponding K_A value for both systems were estimated to be close to 7×10^7 M^{-1} [35], consistent with the low selectivity of RPA binding. Further experiments revealed that the RPA70AB affinity is almost unaffected by the increase in length of the ssDNA once it reaches ten bases [94], thus validating the assumption that dC_8 can be a suitable ligand model. One limitation, however, may arise by the length of the RPA used in our model, which comprises the first two subunits RPA 70AB, the only structure available by X-ray crystallography. Experimental data cover mutations encompassing the entire RPA protein; however, we selected the results relative to mutations in the RPA70AB region obtained utilizing full length RPA70ABCD (RPA1). The binding affinity of the remaining domains is significantly lower compared to the AB fragment and their role in the binding is assumed to be minor. Hence, we assume that RPA70AB interacting with a dC_8 oligomer is a suitable model for studying the ssDNA binding interactions of RPA1. Moreover, the correlation between the theoretical and experimental $\Delta\Delta G$ is excellent, validating the applicability of the model.

The first approach used to estimate a correlation between the calculated end experimental $\Delta\Delta G$ was based on the alanine scan method, Fig. 2. In this approach, mutants have geometry very close to the wild type structure. The results show a moderate correlation with a correlation coefficient, R, value of 0.71 and 0.76, when the energies are calculated by MM-PBSA and MM-GBSA respectively. The known overestimation of the binding free energy calculated by MM-PB(GB)SA is reflected also in the $\Delta\Delta G$ values. The corresponding slope of the linear regression between the calculated and the experimental $\Delta\Delta G$ is about one order of magnitude higher. This behavior is not unknown in the literature [95], but the correlation parameter is most relevant for the prediction of the $\Delta\Delta G$ change. The use of FoldX program yields results of a different nature. Although the prediction of $\Delta\Delta G$ is much closer to the experimental data, as shown by the slope of the linear regression close to unity, the R value, 0.43, reveals a low correlation between the

Table 3 MM-PB(GB)SA energy components of the double mutants individually modeled and calculated for 800 equidistant snapshots extracted from the last 8 ns of the same trajectory of RPA70AB

Contrib	E277A-R382A		F238A-F269A		K263A-E277A		K263A-R382A		R234A-E277A		R234A-K263A		R234A-R382A		W361A-F386A	
	Mean	sem[h]	Mean	sem	Mean	sem	Mean	sem	Mean	sem[h]	Mean	sem[h]	Mean	sem[h]	Mean	sem[h]
H_{elec} [a]	−711.1	8.3	−861.2	8.0	−827.0	7.8	−409.7	8.1	−817.1	8.1	−515.8	8.0	−399.9	8.3	−860.1	8.0
H_{vdW} [b]	−121.9	1.7	−111.0	1.7	−121.8	1.7	−121.3	1.7	−122.7	1.7	−122.1	1.7	−122.3	1.7	−110.2	1.7
H_{gas} [c]	−833.0	8.7	−972.2	8.4	−948.8	8.3	−531.1	8.5	−939.8	8.5	−637.9	8.4	−522.2	8.7	−970.3	8.4
G_{np} [d]	−18.5	0.1	−18.4	0.1	−18.3	0.1	−18.0	0.1	−18.6	0.1	−18.1	0.1	−18.3	0.1	−18.0	0.1
G_{PB} [e]	762.9	7.8	864.4	7.5	845.5	7.3	436.3	7.5	842.2	7.6	538.5	7.5	444.3	7.8	866.6	7.5
G_{solv} [f]	744.4	7.7	846.1	7.5	827.2	7.3	418.3	7.5	823.6	7.6	520.4	7.4	426.1	7.7	848.6	7.4
ΔG^{PB} [g]	−88.6	3.6	−126.1	3.6	−121.6	3.6	−112.8	3.6	−116.2	3.6	−117.5	3.6	−96.1	3.6	−121.6	3.6
G_{np}	*−18.5*	*0.1*	*−18.4*	*0.1*	*−18.3*	*0.1*	*−18.0*	*0.1*	*−18.6*	*0.1*	*−18.1*	*0.1*	*−18.3*	*0.1*	*−18.0*	*0.1*
G_{GB}	*746.2*	*7.7*	*868.0*	*7.5*	*844.5*	*7.3*	*442.6*	*7.6*	*841.9*	*7.6*	*541.6*	*7.5*	*443.2*	*7.8*	*868.9*	*7.5*
G_{solv}	*727.7*	*7.7*	*849.7*	*7.4*	*826.2*	*7.3*	*424.6*	*7.5*	*823.3*	*7.5*	*523.5*	*7.4*	*424.9*	*7.7*	*850.9*	*7.4*
ΔG^{GB}	*−105.3*	*3.4*	*−122.5*	*3.4*	*−122.6*	*3.4*	*−106.5*	*3.4*	*−116.5*	*3.4*	*−114.4*	*3.4*	*−97.2*	*3.4*	*−119.3*	*3.4*

The standard state is assumed to be at 1 M. The units are in kcal mol^{-1} . Ionic strength is set at 50 mM. [(a)] H_{elec}: coulombic energy; [(b)] H_{vdW}: van der Waals energy; [(c)] $H_{gas}=H_{elec}+H_{vdW}$; [(d)] G_{np} non-polar solvation free energy; [(e)] G_{PB}: polar solvation free energy calculated by solving the Poisson-Boltzmann equation PB, in italic by solving the generalized Born equation, GB; [(f)] $G_{solv}=G_{np}+G_{PB/GB}$; [(g)] $\Delta G^{PB/GB}=H_{gas}+G_{solv}$; [(h)] Standard error of mean values; The MM-GBSA value are shown in italic

experimental and theoretical estimation. The alanine scan, which uses the same conformational geometry gathered from the wild type complex, is an inappropriate approach to study the change of RPA affinity in function of mutations. This result suggests that a conformational change has to take place upon mutation during the binding process.

To circumvent this problem, we individually modeled each mutation and calculated the binding free energy with a single trajectory method, Fig. 3. In the first case, Fig. 3a, we included the solute binding entropy term, $T\Delta S$, in the change of the free energy estimate, whereas in the second case, only the $\Delta G^{PB/GB}$ was considered, Fig. 3b. The slope of the linear interpolation for both cases is quite similar, with values comprised from 8.45 to 11.9. Regardless of the use of $T\Delta S$, there is a systematic overestimation of the $\Delta\Delta G$ when the MM-PB(GB)SA method is used. On the other hand, the correlation coefficient R is in optimal accord with experimental data. When the change in ΔG^{PB} and ΔG^{GB} are correlated with the experimental $\Delta\Delta G$, i.e., the $T\Delta S$ term is omitted, the resulting R value raises to 0.88 and 0.78 respectively, Fig. 3b. This is a significant improvement with respect to the data gathered from the alanine scan, Fig. 2. This finding shows that despite the overestimation of the $\Delta\Delta G$ by MM-PB(GB)SA method, the correlation with experiments is very satisfactory and some prediction on change of the affinity of RPA70AB using new mutants would be achievable with reasonable accuracy. The agreement, however, is better when the MM-PBSA method is used compared to the MM-GBSA approach. The introduction of the entropy does not improve results; the R values are slightly lowered to 0.78 and 0.60 by PB and GB respectively. We used 20 snapshots to estimate the $T\Delta S$ term, thus the decrease in agreement might be attributed to the limited number of snapshots used.

The data obtained suggests that the MM-PB(GB)SA can be used with good accuracy to predict and reproduce the changes in affinity as a function of mutation; however, it is necessary to model each mutant independently, avoiding methods based on the alanine scan which neglect conformational changes that may affect the binding.

In the geometry of the interacting complex, the amino acids mutated in this study are shown in Fig. 4a. Here, it is noticeable that R234 is the only amino acid which has two hydrogen bond interactions with the cytosine moiety of the dC_8 ligand (dotted black lines). The residue E277 shows a polar contact with the adjacent base; however, the position of the hydrogen (not shown in the picture) is not aligned with the O−N axes. The strength of the corresponding hydrogen bond that will result is highly diminished because of its high directionality character. This is consistent with the very low value of ΔG^{GB} calculated for E277, Fig. 5, and with the corresponding moderate loss in affinity which the corresponding mutant shows. The remaining hydrogen bond

Table 4 MM-PB(GB)SA energy components of the single mutant individually modeled and calculated for 800 equidistant snapshots extracted from the last 8 ns of the same trajectory of RPA70AB

	E277A		K263A		R234A		R382A	
Contrib	Mean	sem	Mean	sem	Mean	sem	Mean	sem
H_{elec} [a]	−996.6	8.1	−695.3	7.9	−685.4	8.1	−579.4	8.2
H_{vdW} [b]	−123.4	1.7	−122.8	1.7	−123.8	1.7	−123.0	1.7
H_{gas} [c]	−1120.1	8.5	−818.1	8.3	−809.2	8.5	−702.4	8.6
G_{np} [d]	−18.8	0.1	−18.3	0.1	−18.6	0.1	−18.4	0.1
G_{PB} [e]	1029.3	7.6	702.7	7.4	710.8	7.6	608.4	7.7
G_{solv} [f]	1010.5	7.5	684.3	7.3	692.3	7.6	589.9	7.7
G^{PB} [g]	−109.5	3.6	−133.8	3.6	−116.9	3.6	−112.5	3.6
G_{np}	−18.8	0.1	−18.3	0.1	−18.6	0.1	−18.4	0.1
G_{GB}	1012.4	7.5	708.8	7.4	709.4	7.6	610.2	7.7
G_{solv}’	993.6	7.5	690.5	7.3	690.8	7.5	591.7	7.6
G^{GB}	−126.5	3.4	−127.6	3.4	−118.4	3.4	−110.7	3.4

The standard state is assumed to be at 1 M. The units are in kcal mol^{-1} . Ionic strength is set at 50 mM. [(a)] H_{elec}: coulombic energy; [(b)] H_{vdW}: van der Waals energy; [(c)] H_{gas}=H_{elec}+H_{vdW}; [(d)] G_{np} non-polar solvation free energy; [(e)] G_{PB}: polar solvation free energy calculated by solving the Poisson-Boltzmann equation PB, in italic by solving the generalized Born equation, GB; [(f)] G_{solv}=G_{np}+$G_{PB/GB}$; [(g)] $\Delta G^{PB/GB}$ =H_{gas}+G_{solv}; [(h)] Standard error of mean values; The MM-GBSA value are shown in italic

interaction is from R382, but it is directed toward the backbone of the ssDNA, hence not particularly specific to the nature of the ligand. The only residue that might underline a difference between the model system, where the ligand is dC_8, and the experimental system, where the ligand is dT_{30}, is the R234 which is the only amino acid amongst the mutations considered that has a strong and specific interaction with a base. Curiously, if we remove all data relative to the single and double mutations containing R234, the R value increases considerably showing an optimal correlation. The new R values are 0.95 by MM-PBSA and 0.82 by MM-GBSA, Fig. 4b. This finding is indeed very interesting since it suggests that if the corresponding data were measured using dC_8 as a ligand, we could expect a very good correlation between the ΔΔG value calculated by MM-PBSA and the experiments.

The ΔG^{GB} energies per residue for amino acids in close contact with the dC_8 are depicted in Fig. 5. The ΔG^{GB} values comprised by the sum of Coulombic interaction and polar solvation free energy $\Delta G_{Coul+GB}$ are strongly anticorrelated with the van der Waals energy, ΔG_{vdW}, and the non-polar contribution, ΔG_{np}. The values of each term are consistent with the hydrophobicity and hydrophilicity of the residue. The polar amino acids, such as Arg 382, 234, 335, 210, 216, 214, and 339 have a strong polar contribution in the ΔG^{GB} value, contrary to all Phe residues which have a higher ΔG_{vdW} contribution with respect to the $\Delta G_{Coul+GB}$ component. There is good agreement with the mutation of residues with high ΔG^{GB} values and corresponding reduction of K_A. For single mutations, the highest values of ΔG^{GB} are found for R234 and R382, −6.1 kcal mol^{-1} and −10.9 kcal mol^{-1} respectively, and experimentally for the same mutation, a strong loss in affinity with ratios of relative bindings of 0.027 and 0.04. Small changes in affinity are encountered by mutations at positions 277 and 263, consistent with the corresponding ΔG^{GB} values which are estimated to be lower than 1 kcal mol^{-1}. Double mutations involving non polar residues influence the affinity moderately. The F238A-F269A and W261-F386A mutations have a relative binding of 0.23 and 0.15 respectively, where the ΔG^{GB} are −4.1 kcal mol^{-1} and −5.1 kcal mol^{-1} for the first couple and −6.5 kcal mol^{-1} and −3.1 kcal mol^{-1} for the second couple. In both cases, the ΔG_{vdW} terms contribute by more than half to the ΔG^{GB} term. The qualitative accuracy in the correlation between the ΔG^{GB} and the K_A values suggest that additional mutations might lead to a strong loss of binding. The R335, R216, and N214 have a significant value of ΔG^{GB} contrary to S233, N214, E241 N281, N337, S336 and S384, all of which show a negligible energetic contribution to the binding.

The results indicate that hydrophobic and hydrophilic properties of a residue are insufficient to estimate its binding capability. Even with a strong polar ligand such as dC_8, non-polar amino acids might have a significant binding character due to the ΔG_{vdW} component.

A different agreement is reached by the calculations of the change of binding free energy as a function of the alanine scan on the contact sites by FoldX, Fig. 6. The highest loss in affinity is correctly reproduced with an increase of the binding free energy to 5.6 kcal mol^{-1}; however, the free energy calculated for the other mutations seem to correlate poorly with experimental data, as already observed, Fig. 2. The absolute values calculated by MM-PB (GB)SA for each mutation on individual trajectory or with the alanine scan method are reported to underline how the change in ΔG^{GB} is substantial even for a single point mutation, Tables 1, 2, 3, and 4.

Conclusions

This study demonstrated the correlation between theoretical and experimentally observed ΔΔG values for single and double point mutations of RPA70 binding an oligomer ssDNA. The MM-PBSA results are the most accurate and yield a good value of R^2, especially if the mutations are individually modeled and each trajectory is used to estimate the binding free energy. The ΔΔG values, however, are one order of magnitude overestimated. Similar analysis is performed using the code FoldX with good qualitative agreement, but the results are less accurate. The introduction of binding entropy did not improve the results. The alanine scan technique, based on the single trajectory of the complex of RPA70AB, returned a less accurate prediction suggesting that the conformational change resulting from a mutation might play a significant role in the binding process. We further deduce that omission of the mutation involving the only amino acid that forms a direct hydrogen bond with the base of ssDNA leads to a considerable increase in the value of R, thus highlighting how the MM-PBSA method can be used successfully to predict the correlation between the experimental and theoretical ΔΔG.

Acknowledgments We gratefully acknowledge the support for this work from the NASA (National Aeronautics and Space Administration) Space Radiation Risk Assessment Project.

References

1. Binz SK, Sheehan AM, Wold MS (2004) Replication protein A phosphorylation and the cellular response to DNA damage. DNA Repair (Amst) 3:1015–1024
2. Bochkarev A, Bochkareva E (2004) From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. Curr Opin Struct Biol 14:36–42
3. Iftode C, Daniely Y, Borowiec JA (1999) Replication protein A (RPA): the eukaryotic SSB. Crit Rev Biochem Mol Biol 34:141–180
4. Johnson A, O'Donnell M (2005) Cellular DNA replicases: components and dynamics at the replication fork. Annu Rev Biochem 74:283–315
5. Machida YJ, Hamlin JL, Dutta A (2005) Right Place, Right Time, and Only Once: Replication Initiation in Metazoans. Cell 123:13–24
6. Shechter D, Costanzo V, Gautier J (2004) ATR and ATM regulate the timing of DNA replication origin firing. Nat Cell Biol 6:648–655
7. Stauffer ME, Chazin WJ (2004) Structural mechanisms of DNA replication, repair, and recombination. J Biol Chem 279:30915–30918
8. Wold MS (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu Rev Biochem 66:61–92
9. Zou L, Elledge SJ (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science 300:1542–1548
10. Kowalczykowski SC (2000) Some assembly required. Nat Struct Biol 7:1087–1089
11. Yuzhakov A, Kelman Z, Hurwitz J, O'Donnell M (1999) Multiple competition reactions for RPA order the assembly of the DNA polymerase delta holoenzyme. EMBO J 18:6189–6199
12. Shivji KK, Kenny MK, Wood RD (1992) Proliferating cell nuclear antigen is required for DNA excision repair. Cell 69:367–374
13. Nichols AF, Sancar A (1992) Purification of PCNA as a nucleotide excision repair protein. Nucleic Acids Res 20:2441–2446
14. Braun KA, Lao Y, He Z, Ingles CJ, Wold MS (1997) Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. Biochemistry 36:8443–8454
15. Sancar A (1996) DNA excision repair. Annu Rev Biochem 65:43–81
16. Sancar A (1995) Excision repair in mammalian cells. J Biol Chem 270:15915–15918
17. He Z, Henricksen LA, Wold MS, Ingles CJ (1995) RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. Nature 374:566–569
18. Lao Y, Lee CG, Wold MS (1999) Replication protein A interactions with DNA. 2. Characterization of double-stranded DNA-binding/helix-destabilization activities and the role of the zinc-finger domain in DNA interactions. Biochemistry 38:3974–3984
19. Burns JL, Guzder SN, Sung P, Prakash S, Prakash L (1996) An affinity of human replication protein A for ultraviolet-damaged DNA. J Biol Chem 271:11607–11610
20. Patrick SM, Turchi JJ (1998) Human replication protein A preferentially binds cisplatin-damaged duplex DNA in vitro. Biochemistry 37:8808–8815
21. Patrick SM, Turchi JJ (1999) Replication protein A (RPA) binding to duplex cisplatin-damaged DNA is mediated through the generation of single-stranded DNA. J Biol Chem 274: 14972–14978
22. Pretto DI, Tsutakawa S, Brosey CA, Castillo A, Chagot ME, Smith JA, Tainer JA, Chazin WJ (2010) Structural dynamics and single-stranded DNA binding activity of the three N-terminal domains of the large subunit of replication protein A from small angle X-ray scattering. Biochemistry 49:2880–2889
23. Cai L, Roginskaya M, Qu Y, Yang Z, Xu Y, Zou Y (2007) Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler. Biochemistry 46:8226–8233
24. Fanning E, Klimovich V, Nager AR (2006) A dynamic model for replication protein A (RPA) function in DNA processing pathways. Nucleic Acids Res 34:4126–4137
25. Mer G, Bochkarev A, Chazin WJ, Edwards AM (2000) Three-dimensional structure and function of replication protein A. Cold Spring Harb Symp Quant Biol 65:193–200
26. Murzin AG (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. EMBO J 12:861–867
27. Theobald DL, Mitton-Fry RM, Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. Annu Rev Biophys Biomol Struct 32:115–133
28. Daughdrill GW, Ackerman J, Isern NG, Botuyan MV, Arrowsmith C, Wold MS, Lowry DF (2001) The weak interdomain coupling observed in the 70 kDa subunit of human replication protein

A is unaffected by ssDNA binding. Nucleic Acids Res 29: 3270–3276

29. Bochkareva E, Belegu V, Korolev S, Bochkarev A (2001) Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding. EMBO J 20:612–618
30. Bochkareva E, Korolev S, Lees-Miller SP, Bochkarev A (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. EMBO J 21:1855–1863
31. Bastin-Shanower SA, Brill SJ (2001) Functional analysis of the four DNA binding domains of replication protein A. The role of RPA2 in ssDNA binding. J Biol Chem 276:36446–36453
32. Binz SK, Lao Y, Lowry DF, Wold MS (2003) The phosphorylation domain of the 32-kDa subunit of replication protein A (RPA) modulates RPA-DNA interactions. Evidence for an intersubunit interaction. J Biol Chem 278:35584–35591
33. Bochkareva E, Kaustov L, Ayed A, Yi GS, Lu Y, Pineda-Lucena A, Liao JC, Okorokov AL, Milner J, Arrowsmith CH, Bochkarev A (2005) Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. Proc Natl Acad Sci U S A 102:15412–15417
34. Blackwell LJ, Borowiec JA, Mastrangelo IA (1996) Single-stranded-DNA binding alters human replication protein A structure and facilitates interaction with DNA-dependent protein kinase. Mol Cell Biol 16:4798–4807
35. Arunkumar AI, Stauffer ME, Bochkareva E, Bochkarev A, Chazin WJ (2003) Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains. J Biol Chem 278:41077–41082
36. Bhattacharya S, Botuyan MV, Hsu F, Shan X, Arunkumar AI, Arrowsmith CH, Edwards AM, Chazin WJ (2002) Characterization of binding-induced changes in dynamics suggests a model for sequence-nonspecific binding of ssDNA by replication protein A. Protein Sci 11:2316–2325
37. Iftode C, Borowiec JA (2000) 5' –>3' molecular polarity of human replication protein A (hRPA) binding to pseudo-origin DNA substrates. Biochemistry 39:11970–11981
38. de Laat WL, Appeldoorn E, Sugasawa K, Weterings E, Jaspers NG, Hoeijmakers JH (1998) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. Genes Dev 12:2598–2609
39. Wyka IM, Dhar K, Binz SK, Wold MS (2003) Replication protein A interactions with DNA: differential binding of the core domains and analysis of the DNA interaction surface. Biochemistry 42:12909–12918
40. Loo YM, Melendy T (2004) Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA. J Virol 78:1605–1615
41. Park CJ, Lee JH, Choi BS (2005) Solution structure of the DNA-binding domain of RPA from Saccharomyces cerevisiae and its interaction with single-stranded DNA and SV40 T antigen. Nucleic Acids Res 33:4172–4181
42. Daughdrill GW, Buchko GW, Botuyan MV, Arrowsmith C, Wold MS, Kennedy MA, Lowry DF (2003) Chemical shift changes provide evidence for overlapping single-stranded DNA- and XPA-binding sites on the 70 kDa subunit of human replication protein A. Nucleic Acids Res 31:4176–4183
43. Stauffer ME, Chazin WJ (2004) Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA. J Biol Chem 279:25638–25645
44. Ellison V, Stillman B (2003) Biochemical characterization of DNA damage checkpoint complexes: clamp loader and clamp complexes with specificity for 5' recessed DNA. PLoS Biol 1:E33
45. Zou L, Liu D, Elledge SJ (2003) Replication protein A-mediated recruitment and activation of Rad17 complexes. Proc Natl Acad Sci USA 100:13827–13832
46. Wu X, Shell SM, Zou Y (2005) Interaction and colocalization of Rad9/Rad1/Hus1 checkpoint complex with replication protein A in human cells. Oncogene 24:4728–4735
47. Ball HL, Myers JS, Cortez D (2005) ATRIP binding to replication protein A-single-stranded DNA promotes ATR-ATRIP localization but is dispensable for Chk1 phosphorylation. Mol Biol Cell 16:2372–2381
48. Namiki Y, Zou L (2006) ATRIP associates with replication protein A-coated ssDNA through multiple interactions. Proc Natl Acad Sci USA 103:580–585
49. Yoo E, Kim BU, Lee SY, Cho CH, Chung JH, Lee CH (2005) 53BP1 is associated with replication protein A and is required for RPA2 hyperphosphorylation following DNA damage. Oncogene 24:5423–5430
50. Wong JM, Ionescu D, Ingles CJ (2003) Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2. Oncogene 22:28–33
51. Robison JG, Elliott J, Dixon K, Oakley GG (2004) Replication protein A and the Mre11.Rad50.Nbs1 complex co-localize and interact at sites of stalled replication forks. J Biol Chem 279: 34802–34810
52. Daniely Y, Borowiec JA (2000) Formation of a complex between nucleolin and replication protein A after cell stress prevents initiation of DNA replication. J Cell Biol 149:799–810
53. Kim K, Dimitrova DD, Carta KM, Saxena A, Daras M, Borowiec JA (2005) Novel checkpoint response to genotoxic stress mediated by nucleolin-replication protein a complex formation. Mol Cell Biol 25:2463–2474
54. Kim C, Paulus BF, Wold MS (1994) Interactions of human replication protein A with oligonucleotides. Biochemistry 33:14197–14206
55. Liu Y, Yang Z, Utzat CD, Liu Y, Geacintov NE, Basu AK, Zou Y (2005) Interactions of human replication protein A with single-stranded DNA adducts. Biochem J 385:519–526
56. Kim C, Wold MS (1995) Recombinant human replication protein A binds to polynucleotides with low cooperativity. Biochemistry 34:2058–2064
57. Kim C, Snyder RO, Wold MS (1992) Binding properties of replication protein A from human and yeast cells. Mol Cell Biol 12:3050–3059
58. Lao Y, Gomes XV, Ren Y, Taylor JS, Wold MS (2000) Replication protein A interactions with DNA. III. Molecular basis of recognition of damaged DNA. Biochemistry 39:850–859
59. Walther AP, Gomes XV, Lao Y, Lee CG, Wold MS (1999) Replication protein A interactions with DNA. 1. Functions of the DNA-binding and zinc-finger domains of the 70-kDa subunit. Biochemistry 38:3963–3973
60. Bochkarev A, Pfuetzner RA, Edwards AM, Frappier L (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. Nature 385:176–181
61. Vriend G (1990) WHAT IF: a molecular modeling and drug design program. J Mol Graph 8(52–56):29
62. Case DA, Darden TA, Cheatham III TE, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Pearlman DA, Crowley M, Walker RC, Zhang W, Wang B, Hayik S, Roitberg A, Seabra G, Wong KF, Paesani F, Wu X, Brozell S, Tsui V, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Beroza P, Mathews DH, Schafmeister C, Ross WS, Kollman PA (2008) Amber 10. University of California, San Francisco
63. Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE 3rd, Laughton CA, Orozco M (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. Biophys J 92:3817–3829
64. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, M.L. K (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935

65. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N•log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
66. Pastor RW, Brooks BR, Szabo A (1988) Mol Phys 65:1409–1419
67. Miyamoto S, Kollman PA (1992) J Comput Chem 13:952–962
68. Honig B, Nicholls A (1995) Classical electrostatics in biology and chemistry. Science 268:1144–1149
69. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE 3rd (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897
70. Jorgensen WL, Buckner JK, Boudon S, Tirado-Rives J (1988) Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water. J Chem Phys 89:3742–3746
71. Jorgensen WL (1989) Free energy calculations: a breakthrough for modeling organic chemistry in solution. Acc Chem Res 22:184–189
72. Hermans J, Wang L (1997) Inclusion of Loss of Translational and Rotational Freedom in Theoretical Estimates of Free Energies of Binding. Application to a Complex of Benzene and Mutant T4 Lysozyme. J Am Chem Soc 119:2707–2714
73. Roux B, Nina M, Pomès R, Smith JC (1996) Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study. Biophys J 71:670–681
74. Reyes CM, Kollman PA (2000) Structure and thermodynamics of RNA-protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. J Mol Biol 297:1145–1158
75. Gohlke H, Kiel C, Case DA (2003) Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. J Mol Biol 330:891–913
76. Olufsen M, Smalas AO, Brandsdal BO (2008) Electrostatic interactions play an essential role in DNA repair and cold-adaptation of uracil DNA glycosylase. J Mol Model 14:201–213
77. Onufriev A, Bashford D, Case DA (2000) Modification of the generalized Born modelsuitable for macromolecules. J Phys Chem B 104:3712–3720
78. Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free energies usingmacroscopic solvent models. J Phys Chem 98:1978–1988
79. Gohlke H, Case DA (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. J Comput Chem 25:238–250
80. Guerois R, Nielsen JE, Serrano L (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol 320:369–387
81. Schymkowitz JW, Rousseau F, Martins IC, Ferkinghoff-Borg J, Stricher F, Serrano L (2005) Prediction of water and metal binding sites and their affinities by using the Fold-X force field. Proc Natl Acad Sci USA 102:10147–10152
82. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L (2005) The FoldX web server: an online force field. Nucleic Acids Res 33:W382–W388
83. Huo S, Massova I, Kollman PA (2002) Computational alanine scanning of the 1:1 human growth hormone-receptor complex. J Comput Chem 23:15–27
84. Massova I, Kollman PA (2000) Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Perspect Drug Discov Des 18:113–135
85. Zhou Z, Madrid M, Evanseck JD, Madura JD (2005) Effect of a bound non-nucleoside RT inhibitor on the dynamics of wild-type and mutant HIV-1 reverse transcriptase. J Am Chem Soc 127:17253–17260
86. Kuhn B, Kollman PA (2000) Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. J Med Chem 43:3786–3791
87. Masukawa KM, Kollman PA, Kuntz ID (2003) Investigation of neuraminidase-substrate recognition using molecular dynamics and free energy calculations. J Med Chem 46:5628–5637
88. Huo S, Wang J, Cieplak P, Kollman PA, Kuntz ID (2002) Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: insight into structure-based ligand design. J Med Chem 45:1412–1419
89. Wang W, Lim WA, Jakalian A, Wang J, Wang J, Luo R, Bayly CI, Kollman PA (2001) An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculations, and sequence analysis. J Am Chem Soc 123:3986–3994
90. Donini OA, Kollman PA (2000) Calculation and prediction of binding free energies for the matrix metalloproteinases. J Med Chem 43:4180–4188
91. Hou TJ, Guo SL, Xu XJ (2002) Predictions of Binding of a Diverse Set of Ligands to Gelatinase-A by a Combination of Molecular Dynamics and Continuum Solvent Models. J Phys Chem B 106:5527–5535
92. Yan C, Xiu Z, Li X, Li S, Hao C, Teng H (2008) Comparative molecular dynamics simulations of histone deacetylase-like protein: binding modes and free energy analysis to hydroxamic acid inhibitors. Proteins 73:134–149
93. Wong S, Amaro RE, McCammon JA (2009) MM-PBSA Captures Key Role of Intercalating Water Molecules at a Protein-Protein Interface. J Chem Theor Comput 5:422–429
94. Pfuetzner RA, Bochkarev A, Frappier L, Edwards AM (1997) Replication protein A. Characterization and crystallization of the DNA binding domain. J Biol Chem 272:430–434
95. Hou T, Wang J, Li Y, Wang W Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51:69–82