CHAPTER 5

Investigating the role of XPE in modulating the functioning of XPA

CHAPTER 5A

To study and characterize the protein-protein interactions between the DNA binding proteins, XPA and XPE

CHAPTER 5A

To study and characterize the protein-protein interactions between the DNA binding proteins, XPA and XPE

5A.1 Abstract

The scaffold nature of *Xeroderma pigmentosum* complementation group A (XPA) protein makes it an important member of nucleotide excision repair (NER) that removes bulky DNA lesions with the help of various protein-protein interactions (PPI) and DNA-protein interactions. However, many structural insights of XPA's interaction and the binding patterns with other NER proteins are yet to be understood. Here, we have studied one such crucial PPI of XPA with another NER protein, *Xeroderma pigmentosum* complementation group E (XPE), by using the previously identified binding site of XPA (residues 185-226) in the Assisted Model Building with Energy Refinement (AMBER) force-field mediated dynamic system. We studied the relationship between the $XPA_{185-226}$-XPE complex using three different docked models. The major residues observed in all the models that were responsible for the PPI of this complex were Arg20, Arg47, Asp51, and Leu57 from XPE, and the residues Leu191, Gln192, Val193, Trp194, Glu198, Glu202, Glu205, Arg207, Glu209, Gln216, and Phe219 from $XPE_{185-226}$. During the simulation study, the orientation of XPA was also noticed to be changed by almost 180˚ in Model 1 and 3, which remain unchanged in Model 2, indicating that XPA interacts with XPE with its N-terminal end facing downwards and C-terminal end facing upward. The same was concurrent with the binding of the DNA binding domain (DBD) region of XPA (aa98-239) with XPE. The N-terminal of XPE was stretched for accommodating XPA. Using the per-residue energy decomposition (PRED) analysis for the interface residues of all models, the binding affinity between these proteins was found to be dependent upon R20, R47, and L57 of XPE and the residues L191, V193, W194, E198, E202, E205, R207, and F219 of XPA, respectively. The net binding free energy (BFE) of the $XPA_{185-226}$-XPE protein complex was found to be -48.3718 kcal mol^{-1} for Model 1, -49.09 kcal mol^{-1} for Model 2, and -56.51 kcal mol^{-1} for Model 3, respectively.

5A.2 Introduction

DNA is always under constant threats and attacks from entities of an endogenous or an exogenous nature, which makes the role of DNA repair response a crucial mechanism in all living beings. Nucleotide excision repair (NER) is one such DNA repair pathway that addresses bulky DNA damages, such as cyclobutane pyrimidine dimers (CPD), 6-4 photoproducts (6-4PP), and helix distorting platinum (Pt) crosslinks, which are inflicted upon DNA by various mutagens [48, 352, 353]. This process is mediated in a multi-step

fashion by the coordinated interaction of more than twenty different proteins, which is mainly overseen and systematized by *Xeroderma pigmentosum* complementation group A (XPA) protein, earning itself the title of 'scaffolding protein'. XPA functions as a primary damage recognition protein in both global genome NER (GG-NER) and transcription-coupled NER (TC-NER) [48, 54, 127, 164]. As a result, any alteration in the XPA gene, or the protein function leads to classical *Xeroderma pigmentosum* (XP) disease that is characterized by extreme sun sensitivity, and neurological damages and is often linked with skin cancers [54, 160, 161, 163-165, 169, 170, 361, 362, 364, 366, 371].

The highly conserved XPA, [423-425], consisting of 273 residues [365], with its disordered N- and C-terminals, has been reported to bind and interact with many proteins of NER as well as with the damaged DNA [114-116]. The DNA interacting region of XPA was mapped initially to the globular DNA binding domain (DBD), which spans between 98-219 amino acid residues [142, 367], but since the earlier DBD lacked a significant number of positive residues for the strong bonding with the negatively charged DNA, the DBD of XPA has been now redefined between the 98-239 amino acid residues [58, 126, 176]. Till now only a small number of the XPA's protein-protein interactions (PPI) with other NER proteins have been explored. Some of the well-documented PPIs of XPA are (i) helicase, transcription factor II H (TFIIH) complex [128, 165, 356], (ii) GG-NER damage verifier, *Xeroderma pigmentosum* complementation group C (XPC) protein [123, 354, 370], (iii) excision-repair cross-complementing group 1 endonuclease (*Xeroderma pigmentosum* group F) (ERCC1/XPF) [144-146], and (iv) replication protein A70 (RPA70), which binds to the undamaged DNA [125, 127, 142]. For the remaining, unexplored set of PPIs between XPA and other NER proteins, only a few binding sites of XPA by which it communicates with other NER proteins have been identified, while in some cases, vice-versa, where the binding site of the participating fellow NER proteins through which they interact with XPA is not known.

Among many such PPIs of XPA, one such interaction is with *Xeroderma pigmentosum* complementation group E, also known as the damaged DNA binding protein 2 (DDB2/XPE), which is a subunit of the heterodimer DDB1/DDB2 protein

complex, involved in damage verification and the recognition of DNA lesions in GG-NER [147, 148, 446]. In the earlier *in vitro* and *in vivo* studies by Wakasugi et al. [151], the exact binding site of XPA through which it communicates with XPE was identified. They mapped out the exact binding range of XPA responsible for the interaction by conducting a series of N-terminal residue truncations of XPA in various residue lengths, before identifying its true binding location (aa185-226), which upon truncation in anywhere in-between regions of the protein decreased the binding activity of these two proteins. They also reported a missense mutation in XPA, R207G, which lowered the chances of XPA's recruitment to the NER bubble. This particular mutation further inhibited the binding between XPA and XPE, leading to the incompetency on the part of NER to remove the CPD lesion from the DNA strand in simian vacuolating virus 40 (SV40) transformed human cells and thereby causing a complete NER failure. Since there is no structure for the XPA-XPE complex, and the exact interaction site of XPA on XPE is not known either, thus it becomes important that the key residues mediating the binding of these two proteins be elucidated.

Therefore, in this study, we have attempted to determine the probable binding site of XPE to which $XPA_{185-226}$ binds, and analyze their PPI using the molecular dynamics (MD) approach, which is complementary to the experimental methods. Here, we prepared three models for the protein complex of $XPA_{185-226}$-XPE using the ClusPro web server [124, 345, 447], followed by MD simulation of the systems using the Assisted Model Building with Energy Refinement (AMBER) 14 software package [242] to study their conformational dynamics and stability. Furthermore, using the molecular mechanics Poisson–Boltzmann surface area continuum solvation (MM-PBSA) package [321, 372, 373, 374, 375], we also identified the probable residues responsible for their stable protein formation using the per-residue energy decomposition (PRED) analysis.

5A.3 Materials and methodology

5A.3.1 Molecular docking and the preparation of initial structures

The 3D structure of $XPA_{185-226}$ was modeled by querying the desired protein sequence in Universal Protein Resource (UniProt) database [378] (ID: P23025), followed by which it was submitted to the ITASSER server [448] that works on the protein threading and

homology strategy for the structure prediction. Out of the 5 obtained structures from I-TASSER, we chose the best model based on the C-score, TM score [449], and RMSD score. We then validated the fitness of the modeled structure using the RAMPAGE server [383], ProSA server [382], and Molprobity sever [450]. For the structure of XPE, we retrieved the X-ray crystallized structure of the XPE/DDB1 heterodimer, which was in complex with DNA (PDB ID: 4E5Z) [451] from the Protein Data Bank [377], and removed the DDB1 protein and DNA from the heterodimer complex using the University of California, San Francisco (UCSF) Chimera v.1.12 [314], retaining only XPE (DDB2) for further steps.

Next, we docked the best-represented structure of $XPA_{185-226}$ with XPE using the ClusPro web server. This server, in particular, predicts an ideal docked conformations using three main strategies (i) tests numerous conformations by fast Fourier transform (FFT)-based rigid docking using the PIPER algorithm, wherein the ligand is rotated 70,000 times, and each rotation places the ligand in x, y, and z-axis relative to the receptor on a grid, after which the best score of each rotation is chosen for further testing; (ii) of the 70,000 rotations, 1000 ligands/conformers are chosen based on their lowest energy scores, and (iii) 1000 conformers are clustered with 9 Å Cα RMSD radius to represent the most favorable conformer of the complex from their respective highly populated clusters, which means that the conformer having maximum neighbor within 9 Å is chosen as the cluster center while its neighbors which becomes part of that particular cluster are later removed to retain only one conformer (cluster center) to represent that cluster. This is repeated for all the clusters Lastly, the 10 such cluster center conformers obtained from the populated clusters are energy minimized for 300 steps with a fixed backbone using CHARMM (Chemistry at HARvard Macromolecular Mechanics) forcefield. The 10 model complexes (cluster centers) generated for the $XPA_{185-226}$-XPE complex were ranked based on the number of highly populated clusters, cluster centers, and the lowest energy weighted scores, out of which we chose the top three models for our study.

Using the Amberff99 force field [282], we prepared the initial coordinate and the topology file for the chosen docked model of $XPA_{185-226}$-XPE in the Leap module of the AMBER 14 software package [252]. The whole system was hydrated with water

molecules (model 1=16,879, model 2=18699, model 3=17382) in a cuboid box with a buffer distance of 10 Å using the TIP3P (transferable intermolecular potential with 3 points) water model [296], and were added with charge-balancing counterions (seventeen Cl^- ions each for all models) to neutralize the system.,

5A.3.2 MD simulation of $XPA_{185-226}$-XPE complex

MD study of our three systems/models was done in AMBER 14 suite using the Particle Mesh Ewald algorithm (PME) [388], where the Leap prepared system was subjected to 2 step energy minimization first with the non-bonded interactions cut off of 8 Å. The initial minimization was done using the steepest descent algorithm by imposing restraints over the solute, while the second conjugate gradient minimization method was done without any such restraints. The heating dynamics of the systems were done by gradually increasing the temperature from 0-300 K under constant volume (NVT) conditions, followed by the equilibration at NPT conditions. 40 ns of full long-range MD production runs were done for the equilibrated structure of all the models using the PME algorithm. The Berendsen weak coupling algorithm [294] (0.5 ps of heat bath and 0.2 ps of pressure relaxation) was employed to control the temperature throughout the simulation process. The SHAKE algorithm [292] was used to restrain all the bonds at the time step of 2 fs.

5A.3.3 MD analyses

The corresponding MD trajectories (40 ns each) obtained by simulating the three models of the $XPA_{185-226}$-XPE complexes were analyzed using PTRAJ and CPPTRAJ algorithms [316]. Further, we also determined the intermolecular hydrogen bonds formed between $XPA_{185-226}$ and XPE in all the systems. The results were drawn based on the bond occupancy, bond length (cut off= 3 Å), and the bond angle (cut off= 135˚), formed (HA–H–HD) between the hydrogen donor (HD) and the hydrogen acceptor (HA) atoms of both proteins. The molecular graphics, visualization, and analyses of all MD snapshots were performed using UCSF Chimera package v.1.12.

5A.3.4 Determination of the interface residues

The lowest energy conformer of the $XPA_{185-226}$-XPE complex for all three models was

extracted from the highly populated cluster using the RMSD clustering algorithm. The generated structures were then uploaded to the PDBsum server [348] for the determination of the intermolecular interactions of all the models.

5A.3.5. PRED and BFE analysis of $XPA_{185-226}$-XPE complex

The relative binding free energy (BFE) and the per-residue energy decomposition (PRED) of the interface residues of the protein complex in this present study were conducted using the MMPBSA.py script of the AMBER 14 suite based on the MM-PBSA algorithm [321, 322, 372, 373, 374, 375, 452]. The PRED analysis breaks down the energy contributed by each residue of a protein by studying its molecular interactions over all residues in the system/complex. The MM-PBSA analysis was conducted for all three models of $XPA_{185-226}$-XPE using the following components (i) $XPA_{185-226}$ (ligand) (ii) XPE (receptor) and (iii) $XPA_{185-226}$-XPE (complex), by taking full 40 ns MD trajectories into account.

The free energy of $XPA_{ligand} + XPE_{receptor} = XPA - XPE_{complex}$ was calculated using the following equations derived from the second law of thermodynamics, where they were studied in both gas (vacuum) and aqueous environments.

Sub-complex $\Delta G_{binding} = \Delta G_{complex} - [\Delta G_{receptor} + \Delta G_{ligand}]$ (1)

Here, $G_{binding}$ is the total binding free energy; $G_{complex}$, $G_{receptor}$ and G_{ligand} are the relative free energies of the $XPA_{185-226}$-XPE, XPE, and $XPA_{185-226}$, respectively. The free energy (G) for each component can be obtained by adding the sum of changes in the gas-phase molecular mechanics energies (ΔE_{MM}), polar and non-polar solvation energy ($\Delta G_{solvation}$), and conformational entropy ($-T\Delta S$) of the system; see eq 2.

$$\Delta G_{binding} = \Delta E_{MM} + \Delta G_{solvation} - T\Delta S \quad (2)$$

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{vdW} + \Delta E_{ele} \quad (3)$$

$$\Delta G_{solv} = \Delta G_{PB} + \Delta G_{surf} \quad (4)$$

$$\Delta G = \Delta E_{internal} + \Delta E_{vdW} + \Delta E_{ele} + \Delta G_{PB} + \Delta G_{surf} - T\Delta S \quad (5)$$

Where E_{MM} from eq 3 is the differences in the internal energy ($E_{internal}$), van der Waals forces (E_{vdW}), and electrostatic energy (E_{ele}). The solvation free energy, $G_{solvation}$ as seen from eq 4 was calculated as the sum of the polar contribution (G_{PB}) and the non-

polar contribution (G_{surf}). Thus, the free energy of any system can also be explained as the total of molecular mechanics, solvation free energy, and conformational entropies; see eq 5.

Using the PB solver of the AMBER 14 pbsa module [392], the polar solvation energies were determined for the complex. The probe radius and the grid space for our estimation were set to 1.4 Å and 0.5 Å, respectively. We maintained the dielectric constant of the interior (solvent) and the external (solute) as 80 and 1. Using the solvent-accessible surface area (SASA), the non-polar contribution (G_{Surf}) to the solvation free energy was calculated, see eq 6. The non-polar solvation energy for the PB method was obtained by taking $\gamma = 0.00542$ kcal mol^{-1} Å^{-2}) and $\beta = 0.92$ kcal mol^{-1}, respectively.

$$\Delta G_{surf} = \gamma \times SASA + \beta \qquad (6)$$

5A.4 Results and discussion

5A.4.1 Structural modeling and the validation of XPA$_{185\text{-}226}$ peptide, and the selection of an ideal docked model for XPA$_{185\text{-}226}$-XPE complex

The protein structure of XPA$_{185\text{-}226}$ was obtained upon the submission of the protein's FASTA sequence to the Iterative Threading ASSEmbly Refinement (I-TASSER) server. We obtained a total of five models as shown in (**Figure 5A.1**) based on the threading program, out of which the best-ranked structure had a confidence score or C-score of 1.91, a template modeling (TM) score of 0.70±0.12, and a root-mean-square deviation (RMSD) of 2.5±1.9 Å. This structure was selected as our good fit model structure. A RAMPAGE analysis of this top-ranked modeled structure from I-TASSER showed that 100.0% of the residues were in the favored region, 0% were in the allowed region, and the disallowed region (supplementary **Figure 5A.2A**). The protein structure assessment (ProSA) webserver scores the fitness of the modeled structures in the form of a Z-score after analyzing all the atomic coordinates of the candidate model. The ProSA server houses the Z-scores of all the structures that have been experimentally determined using Nuclear Magnetic Resonance (NMR) and X-ray crystallography and have been deposited in PDB. If the Z-score of any modeled structure falls within these Z-scores of deposited PDB structures, then that modeled structure is said to have near-native

structure conformation. The Z-score for our structure, $XPA_{185-226}$ as seen in **Figure 5A.2B** was -4.59, which indicated that our modeled structure is near to the native structure conformation. The final stereochemical quality of our model was analyzed using the MolProbity server, wherein our model was given the overall score of 0.89 (100^{th} percentile). The detailed description scores for other attributes provided by the MolProbity server are shown in **Table 5A.1**. The analyses done by these three protein validation tools confirmed that our modeled structure is indeed a good fit model.

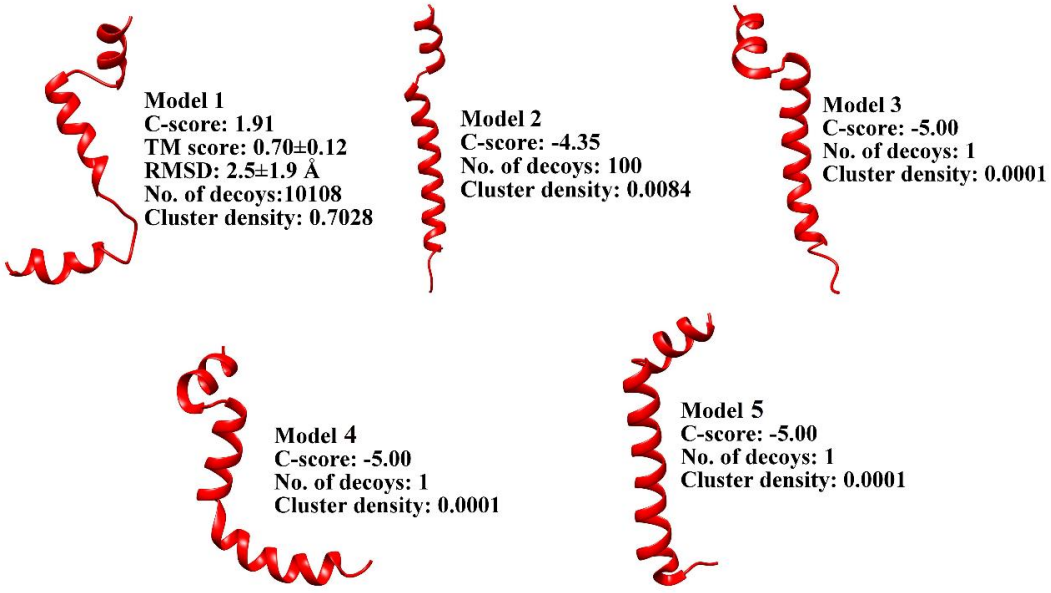

Figure 5A.1. Top five models generated for $XPA_{185-226}$ by I-TASSER server.

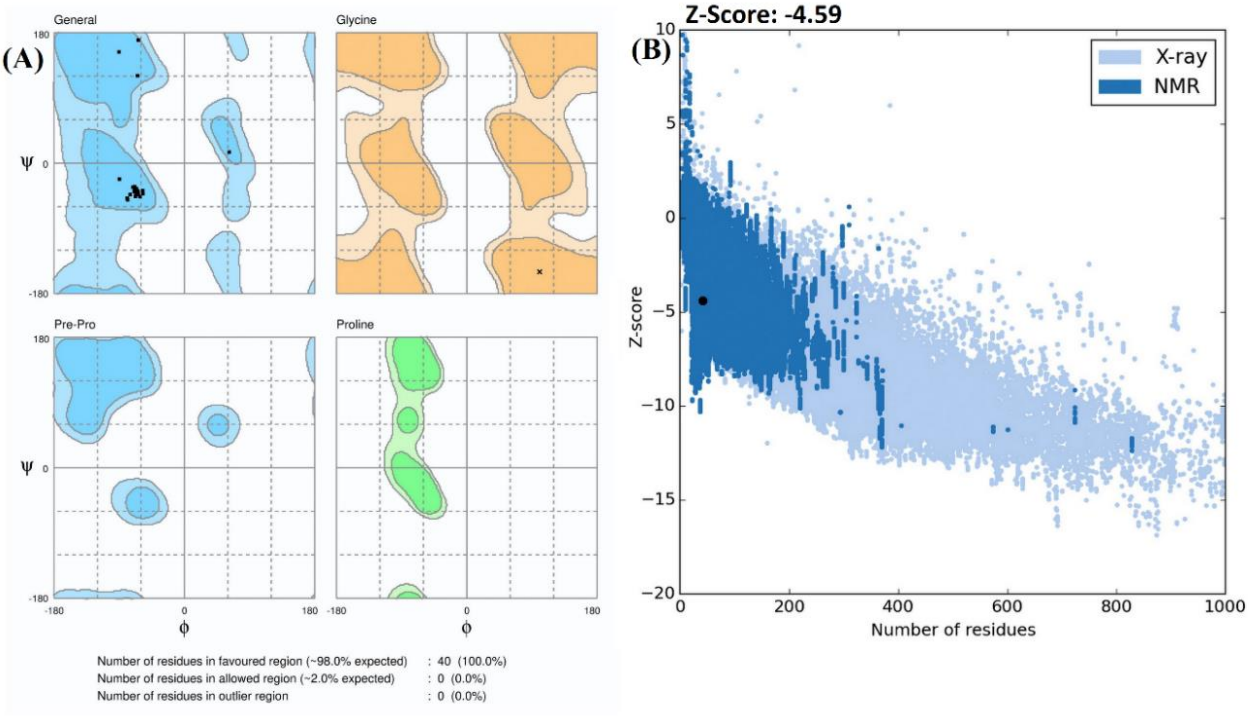

Figure 5A.2. Structure validation of the modeled XPA$_{185-226}$: (A) Ramachandran plot as obtained from Rampage server, and (B) Z-score plot energy plot as determined by the ProSA-web server.

Table 5A.1. MolProbity summary statistics for the modeled XPA$_{185-226}$ protein.

All-Atom contacts	Clashscore, all atoms:	1.36		99th percentile* (N=1784, all resolutions)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	0	0.00%	Goal: <0.3%
	Favored rotamers	38	97.44%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	40	100.00%	Goal: >98%
	MolProbity score^	0.87		100th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	0 / 360	0.00%	Goal: 0%
	Bad angles:	1 / 474	0.21%	Goal: <0.1%

Peptide Omegas	Cis Prolines:	0 / 0	0.00%	Expected: ≤1 per chain, or ≤5%

In the two-column results, the left column gives the raw count, the right column gives the percentage.
* 100th percentile is the best among structures of comparable resolution; the 0th percentile is the worst.

It is an evidentiary fact that XPA's and XPE's PPIs, about one another, determine the NER's outcome which is the removal of the DNA lesions from the genome [151]. The *in vivo/in vitro* binding mechanism of XPA to XPE exclusively studied by Wakasugi and his team were able to decipher the binding site of XPA through it interacted with XPE, yet their interactions at the molecular level are not known. So, to understand the molecular-level PPI between these two proteins, we docked the modeled structure of $XPA_{185-226}$ with XPE using a ClusPro server. The ten model structures that we obtained from ClusPro have been shown in **Figure 5A.3** along with the ranking based on the cluster numbers, center score, and their lowest energy weighted scores. We chose three top-ranked structures for our study, which we named Model 1, 2, and 3 respectively. Model 1 represented the cluster of 181 members, with the cluster center score of -849.6 kcal mol^{-1} and the lowest energy score of -1104.3 kcal mol^{-1}. Model 2 was clustered from 123 cluster members, with the cluster center and the lowest energy score of -952.3 kcal mol^{-1} and -1067.5 kcal mol^{-1}, respectively. Similarly, Model 3 represents the cluster of 111 members, with the cluster center and the lowest energy score of -859.3 kcal mol^{-1} and -1041.45 kcal mol^{-1}, respectively. **Figure 5A.4** shows the docking procedure as well as the surface and the cartoon structure for our complexes. We observed that in all three cases $XPA_{185-226}$ was bound to XPE at the cleft within the residues at the N- terminal and the C-terminal. The surface view of the docked complex shows the closely-bound nature of the complex for all models. As seen from **Fig. 5A.4**, the N-terminal end of $XPA_{185-226}$ in Model 1 and 3 was facing upwards, while the C-terminal end was positioned downwards, but in the case of Model 2, it was just the opposite. Hence, to investigate the proper orientation of XPA's binding characteristics, we performed MD simulation on all three Models.

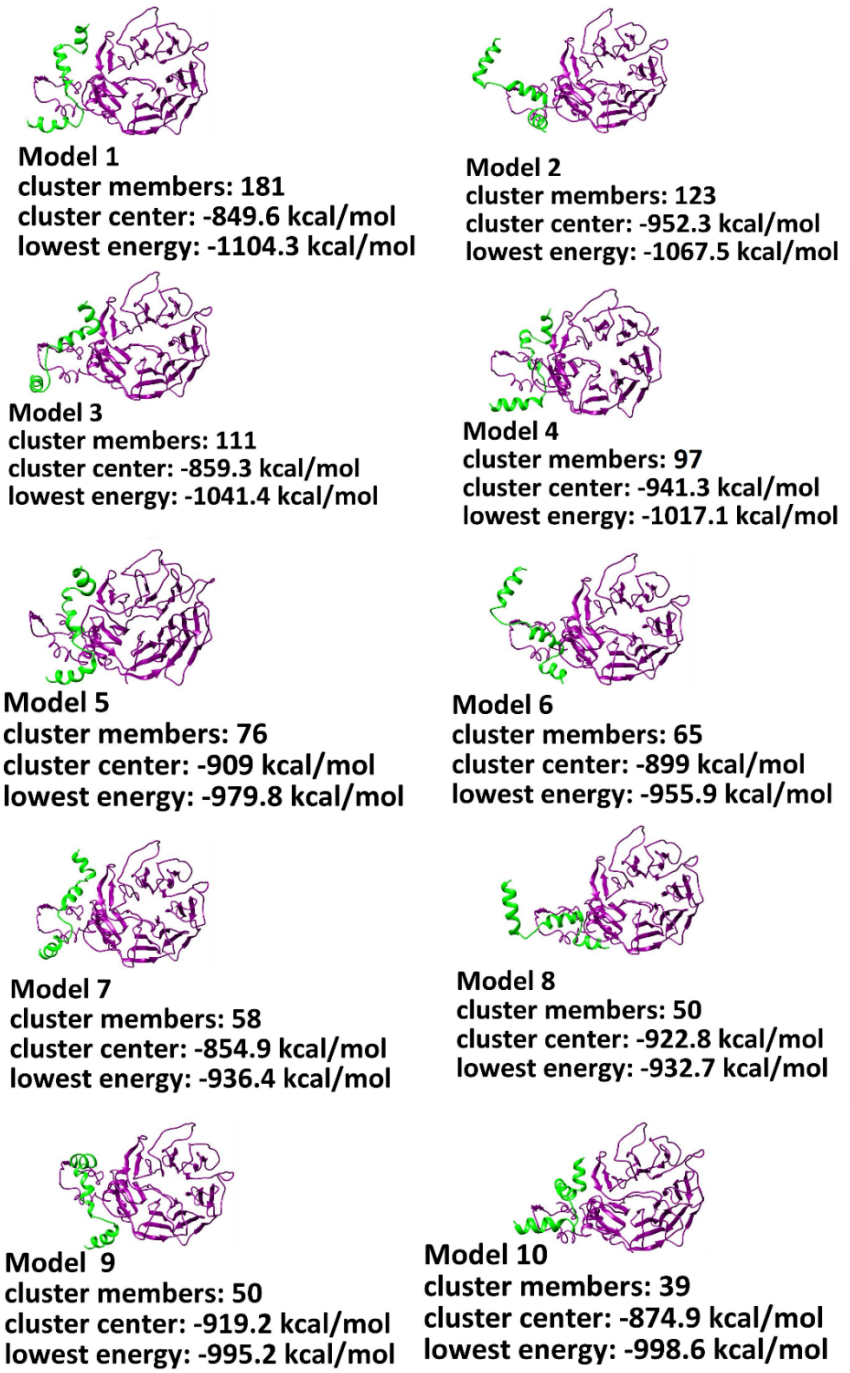

Figure 5A.3. Top ten representative docked models for $XPA_{185-225}$-XPE complex generated by ClusPro server along with their rankings based on the highly populated cluster numbers and their lowest energy weighted scores.

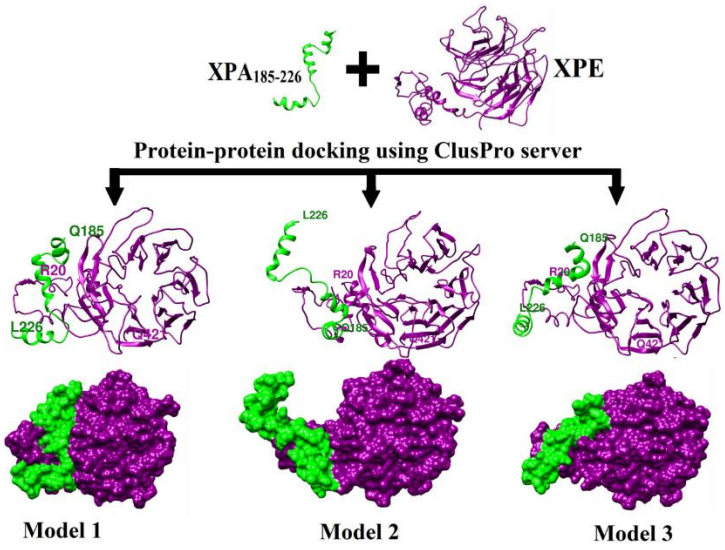

Figure 5A.4. Molecular docking of $XPA_{185-226}$ and XPE.

5A.4.2 Interpretations of the MD trajectories for the $XPA_{185-226}$-XPE complex

The molecular dynamics simulation was administered to probe the structural and conformational changes affecting the PPI between $XPA_{185-226}$ and XPE in an explicit environment. The MD trajectories of 40 ns simulation for all the models were extracted with PTRAJ (short for Process TRAJectory) and CPPTRAJ (a rewrite of PTRAJ in C++) modules as a function of time. The stability of our systems was assessed by measuring the root RMSD of all Cα atoms from their initial coordinates. The RMSD plot for all models of the $XPA_{185-226}$-XPE complex has been given in **Figure 5A.5**. The Cα atoms of model 1 initially fluctuated from 3 Å to 15 Å for a period of 15 ns, after which they started to settle. Model 2 and 2, on the other hand, started settling after 20 ns. Among the three models, Model 3 settled faster compared to the rest.

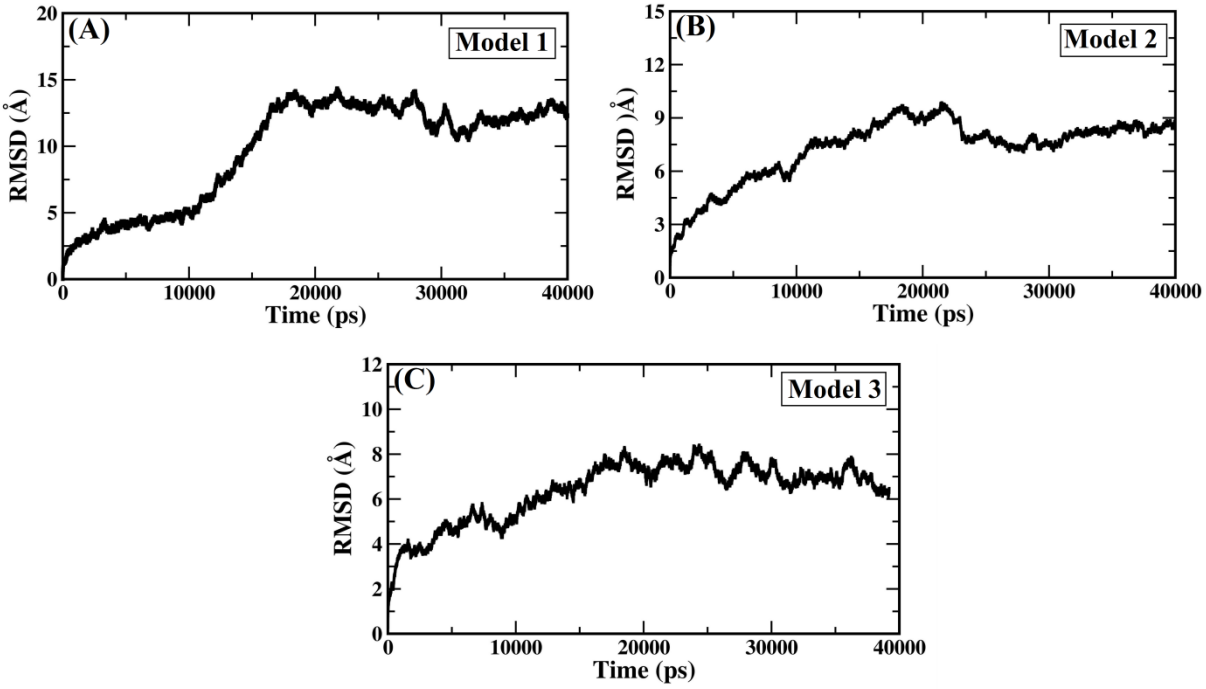

Figure 5A.5. RMSD plot of Cα carbon atoms of XPA$_{185-226}$-XPE complex as a function of time.

Since the solvent-accessible surface area (SASA) of any protein-protein complex (PPC) is often associated with the number of hydrophobic contacts between a solvent and a protein, and since the proteins participating in a PPI are characterized by a larger accessible surface area [453], we also examined the SASA feature for our complex. We found the SASA value for all the models of XPA$_{185-226}$-XPE complexes to be within 24,000-25,000 Å^2 (see **Figure 5A.6**), which suggests a stable interaction.

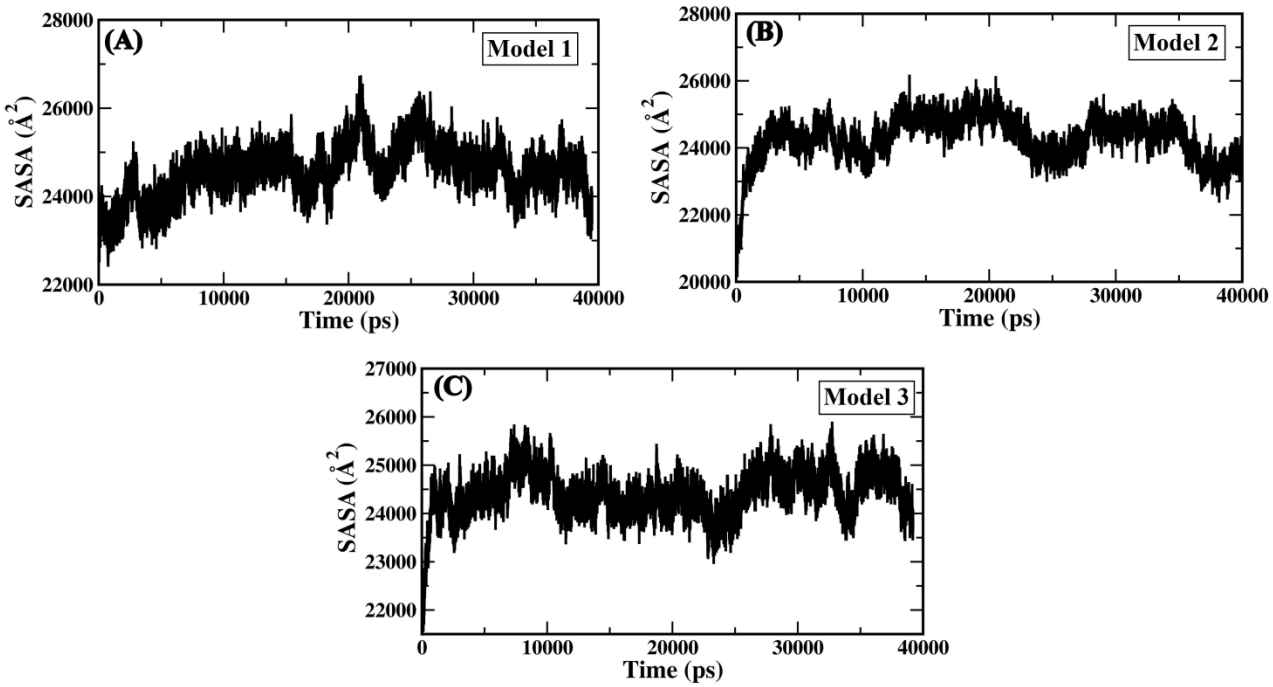

Figure 5A.6. SASA plot for XPA$_{185-226}$-XPE complex as a function of time.

To get a better insight into the RMSD changes, we next extracted the conformational snapshots of our systems at the interval of 10 ns each concerning its equilibrated structure. **Figure 5A.7** shows the trajectory snapshots for all three models of the XPA$_{185-226}$-XPE complex showing its conformational and structural changes at 0, 10, 20, 30, and 40 ns. Here, we observed the XPA$_{185-226}$ peptide of all models to have exhibited secondary structural changes mainly at the helices. One striking feature of XPA as seen in **Fig. 5A.7** was the change in its orientation, where the N-terminal region which was initially positioned upwards in the case of Model 1 and 3 had changed its position by almost 180˚ and was now at the bottom position, meaning that the C-terminal regions were now at the top. But in the case of Model 2, the XPA$_{185-226}$ peptide retained its original position, showing changes only in its secondary structural elements. These changes can be observed in **Fig. 5A.7**, where the end terminal residues of XPA$_{185-226}$ have been labeled to highlight the shift in XPA$_{185-226}$'s orientation throughout the simulation. XP, on the other hand, exhibited the conformational changes restricted only to its C-terminal end, which was stretched a bit as compared to its original position. This may be to place XPA to conduct PPI necessary to regulate NER. The structural changes

on the stretched end of XPE fluctuated between coils and helix, where the helix was the most dominant secondary structural element. The remaining structure of XPE showed no significant changes, maintaining its original β-sheet and the coil form.

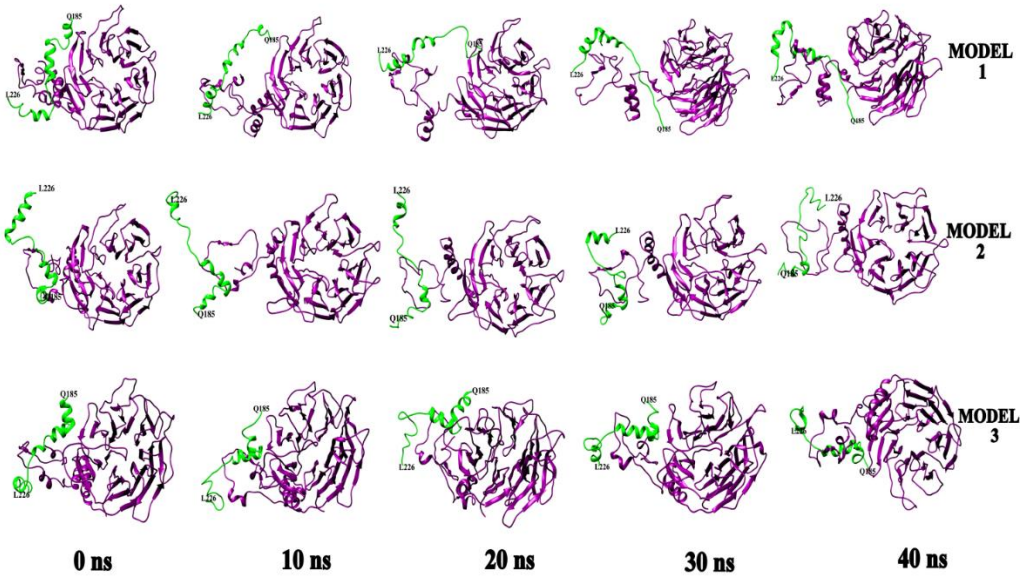

Figure 5A.7. Conformational snapshots of XPA$_{185-226}$-XPE complex at different time intervals.

To gain more insights into the binding between XPA and XPE, we also docked the redefined DBD region of XPA (aa98-239) with XPE to see whether they showed any similarity with these models of XPA-XPE complex or not. We conducted an MD simulation for the same, during which we could observe that the XPA$_{98-239}$ was bound to XPE in the same fashion as model 1 of the XPA$_{185-226}$-XPE complex from the above scenario. The C-terminal residues of XPA$_{98-239}$ were seen to be at the top position while the N-terminal residues were at the bottom. This orientation of XPA$_{98-239}$ was retained throughout the simulation process, which further means that the conformational changes undergone by XPA$_{185-226}$ in the case of Model 1 and 3 with regards to their orientation are indeed concurrent with findings seen from the binding position of Model 2's XPA$_{185-226}$ and DBD region of XPA to XPE (N-terminal end towards bottom and C-terminal end at the top).

Next, we set to analyze the intermolecular hydrogen between XPA and XPE since it is a known fact that the hydrogen bonds provide molecular stability to the PPIs.

The hydrogen bond analysis of the XPA$_{185-226}$-XPE systems for the 40 ns simulation period is provided in **Figure 5A.8**. Model 1 and 2 exhibited a total of 20 intermolecular hydrogen bonds, with an average of 15-17 hydrogen bonds, while Model 3 had a total of 15 hydrogen bonds with an average of 13-14 hydrogen bonds that were all within the ideal bond length and bond angles as demonstrated earlier [454]. The hydrogen bond occupancies of this complex are accordingly given in **Table 5A.2A**, **5A.2B**, and **5A.2C**.

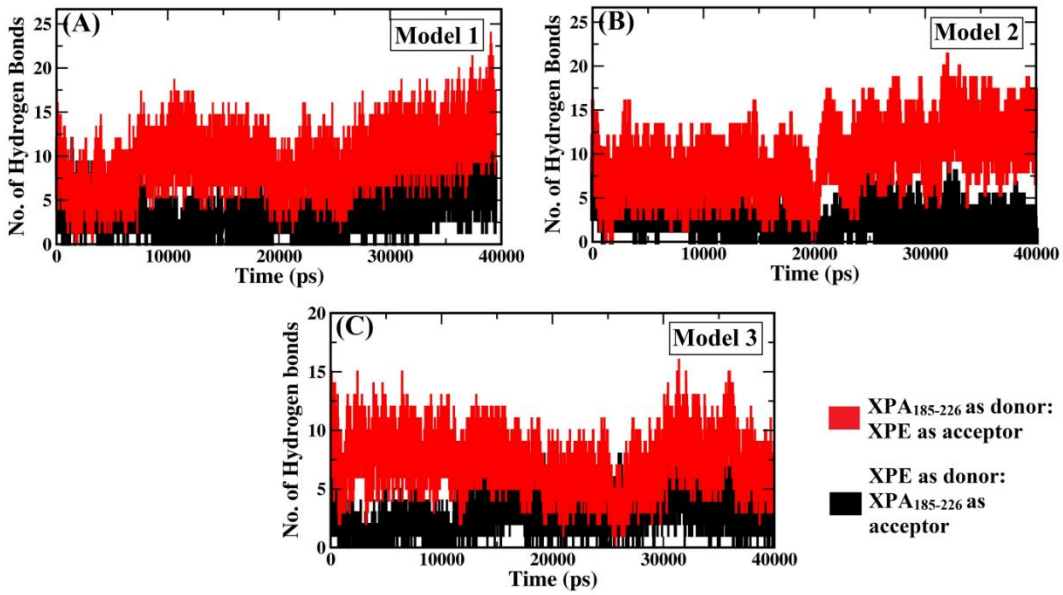

Figure 5A.8. Intermolecular hydrogen bond analyses of XPA$_{185-226}$-XPE complex as a function of time.

Table 5A.2A. Intermolecular hydrogen bond occupancy for Model 1 of XPA$_{185-226}$-XPE complex.

	Donor	Acceptor	Fractions	Bond average distance (Å)	Bond average angle
XPA$_{185-226}$ as a donor: XPE as	GLN_208@NE2	ASP_51@O	0.2819	2.8272	159.378
	GLN_208@NE2	CYS_48@O	0.2362	2.8358	159.641
	ASN_210@N	GLY_56@O	0.1907	2.881	160.999
	GLN_216@NE2	GLY_56@O	0.1717	2.862	159.825
	ARG_189@NH1	ASP_321@OD1	0.1216	2.802	159.66
	ARG_189@NH2	ASP_321@OD2	0.1118	2.8131	157.946
	GLN_216@NE2	GLY_56@O	0.0728	2.8459	160.895
	GLY_195@N	GLN_91@O	0.0626	2.8651	158.379
	GLN_216@NE2	LEU_57@O	0.0609	2.8716	162.427

acceptor	LYS_222@NZ	ASP_51@OD2	0.0548	2.8004	155.791
	LYS_222@NZ	ASP_51@OD2	0.0519	2.8061	155.622
	TRP_194@NE1	SER_93@O	0.0498	2.827	155.412
	LYS_222@NZ	ASP_51@OD2	0.0486	2.798	156.303
	LYS_222@NZ	ASP_51@OD1	0.0469	2.81	154.096
	ARG_189@NH1	ASP_321@OD2	0.0465	2.8171	152.725
	ARG_207@NH1	ASP_49@OD2	0.0456	2.7915	159.582
	LYS_222@NZ	ASP_51@OD1	0.0447	2.8015	155.092
	ASN_210@ND2	GLY_56@O	0.0444	2.8247	152.772
	LYS_222@NZ	ASP_51@OD1	0.0402	2.8075	155.65
	ARG_207@NH2	ASP_49@OD1	0.0358	2.8254	157.202
XPE as Donor: $XPA_{185-226}$ as Acceptor	ARG_20@NH2	GLU_205@OE2	0.2929	2.7985	157.435
	ARG_20@NH2	GLU_198@OE1	0.2642	2.7942	160.114
	ARG_20@NH1	GLU_198@OE2	0.2591	2.7944	159.706
	ARG_20@NE	GLU_205@OE2	0.2573	2.841	156.189
	ARG_20@NH2	GLU_202@OE1	0.2311	2.8127	157.398
	ARG_20@NE	GLU_205@OE1	0.2291	2.8385	156.808
	ARG_20@NH2	GLU_205@OE1	0.227	2.7963	157.549
	ARG_20@NH2	GLU_202@OE2	0.2262	2.8121	157.24
	ARG_20@NH1	GLU_202@OE2	0.2203	2.7943	159.31
	ARG_23@NH1	GLU_205@OE2	0.2182	2.7987	157.13
	ARG_23@NH1	GLU_202@OE2	0.2174	2.7745	159.476
	ARG_23@NH1	GLU_205@OE1	0.2128	2.7941	158.878
	ARG_23@NH1	GLU_209@OE2	0.2018	2.7827	157.197
	ARG_23@NH2	GLU_202@OE1	0.1895	2.8051	159.517
	ARG_20@NH1	GLU_202@OE1	0.1859	2.7926	159.674
	ARG_23@NH1	GLU_205@OE1	0.1726	2.7815	155.178
	ARG_23@NH1	GLU_205@OE2	0.1598	2.8034	157.529
	ARG_23@NH2	GLU_205@OE2	0.1448	2.8076	158.056
	ARG_47@NH2	GLU_202@OE2	0.1402	2.7964	160.25
	GLN_61@NE2	GLN_216@OE1	0.1354	2.8454	160.852
	ARG_23@NH2	GLU_205@OE1	0.1348	2.8	157.639

Table 5A.2B. Intermolecular hydrogen bond occupancy for Model 2 of $XPA_{185-226}$-XPE complex.

	Donor	**Acceptor**	**Fractions**	**Bond average distance (Å)**	**Bond average angle**
	ARG_207@NH1	ASP_49@OD2	0.1904	2.7995	160.4116
	ARG_207@NH2	ASP_49@OD1	0.1865	2.8073	159.3592
	ARG_211@NH2	SER_45@O	0.1289	2.8374	149.009
	ARG_211@NH1	ARG_46@O	0.1177	2.8544	153.8429
	ARG_211@NH1	SER_45@O	0.1045	2.8391	148.7236

$XPA_{185-226}$ as a donor: XPE as acceptor	SER_196@OG	PRO_60@O	0.0982	2.763	160.9462
	LYS_204@NZ	ASP_49@OD2	0.0859	2.8021	156.1736
	ARG_211@NH2	ARG_46@O	0.0713	2.8559	148.9089
	LYS_204@NZ	ASP_49@OD2	0.063	2.8047	157.8416
	ARG_207@NH1	ASP_49@OD1	0.0546	2.8158	154.0108
	ARG_211@NH2	ASP_49@OD2	0.0442	2.8099	155.058
	LYS_204@NZ	ASP_49@OD2	0.0414	2.8033	156.8704
	LYS_204@NZ	SER_50@OG	0.0412	2.8529	155.0998
	ARG_211@NH1	ASP_49@OD2	0.0405	2.8172	155.3912
	ARG_207@NH2	ASP_49@OD2	0.0377	2.8491	149.9096
	LYS_204@NZ	ASP_49@OD1	0.0328	2.8099	152.1185
	LYS_204@NZ	ASP_49@OD1	0.029	2.8101	151.6359
	ARG_211@NH2	ASP_49@OD1	0.0278	2.7993	160.6708
	ARG_211@NH1	ARG_46@HB3	0.0264	2.8005	141.7521
	ARG_211@NH2	ASP_49@OD1	0.0255	2.8112	155.2232
XPE as Donor: $XPA_{185-226}$ as Acceptor	ARG_25@NH2	GLU_202@OE1	0.3855	2.7881	161.3743
	ARG_20@NH2	GLU_202@OE2	0.3633	2.7886	155.4324
	ARG_25@NE	GLU_202@OE2	0.3168	2.8312	160.25
	ARG_20@NE	GLU_198@OE2	0.29	2.8115	155.7706
	ARG_20@NH1	GLU_202@OE2	0.2785	2.8039	153.0602
	GLY_56@N	VAL_193@O	0.2555	2.864	153.6509
	ARG_46@NH2	GLU_205@OE1	0.251	2.798	157.3749
	ARG_20@NH2	GLU_198@OE1	0.2486	2.7958	153.2757
	ARG_46@NH1	GLU_209@OE1	0.2473	2.7884	160.0068
	ARG_20@NH2	GLU_198@OE2	0.2296	2.8008	151.3968
	ARG_20@NE	GLU_198@OE1	0.2218	2.8152	154.4767
	VAL_55@N	VAL_193@O	0.2037	2.8996	157.1968
	SER_24@OG	GLU_198@OE1	0.1874	2.6653	164.1671
	ASN_21@N	GLU_198@OE1	0.1732	2.8402	159.978
	ARG_46@NE	GLU_205@OE2	0.1701	2.8377	158.6113
	SER_24@OG	GLU_198@OE2	0.1622	2.6663	163.4772
	ARG_47@NH1	GLU_205@OE2	0.1566	2.8015	158.6927
	ARG_46@NH2	GLU_205@OE2	0.1464	2.7973	157.1659
	ARG_46@NE	GLU_205@OE1	0.1459	2.8386	155.648
	ASN_21@N	GLU_198@OE2	0.1435	2.8437	159.0869
	ARG_47@NH2	ALA_15@O	0.1414	2.8437	151.5401

Table 5A.2C. Intermolecular hydrogen bond occupancy for Model 3 of $XPA_{185-226}$-XPE complex.

	Donor	**Acceptor**	**Fractions**	**Bond average distance (Å)**	**Bond average angle**
	GLN_185@NE2	PRO_323@O	0.4867	2.8545	161.3575

$XPA_{185-226}$ as a donor: XPE as acceptor	GLN_185@N	LEU_324@O	0.2467	2.902	160.3066
	ASN_210@ND2	SER_50@O	0.2406	2.8595	162.9704
	ASN_210@N	CYS_52@O	0.1434	2.8954	162.6449
	GLN_216@NE2	LEU_53@O	0.1019	2.8487	157.7702
	GLN_208@NE2	ASP_49@OD2	0.0958	2.821	164.0643
	GLN_208@NE2	ASP_49@OD1	0.0914	2.8184	163.2965
	LYS_218@NZ	ASP_51@OD2	0.0782	2.7948	156.7618
	LYS_218@NZ	ASP_51@OD2	0.0761	2.7893	157.1685
	LYS_218@NZ	ASP_51@OD2	0.0663	2.7948	156.9304
	LYS_218@NZ	ASP_51@OD1	0.0634	2.7977	156.2773
	GLN_208@NE2	ASP_49@OD1	0.0572	2.8367	162.2691
	LYS_218@NZ	ASP_51@OD1	0.0563	2.7944	156.6893
	LYS_218@NZ	ASP_51@OD1	0.0444	2.7969	155.6901
	GLN_208@NE2	CYS_48@O	0.0402	2.8541	159.4278
	GLN_185@N	ASP_321@O	0.04	2.8269	152.7492
	SER_196@OG	ASN_378@O	0.04	2.7294	160.0734
	LYS_188@NZ	ASP_321@OD2	0.0365	2.8048	158.0245
	GLN_185@N	ASP_321@OD1	0.0344	2.8084	150.7058
	LYS_188@NZ	ASP_321@OD2	0.0343	2.7977	156.8782
XPE as Donor: $XPA_{185-226}$ as Acceptor	LEU_326@N	GLY_195@O	0.5145	2.8659	158.1377
	ARG_23@NH1	GLU_205@OE2	0.3676	2.7974	159.6708
	ARG_23@NH2	GLU_205@OE2	0.334	2.8044	159.6704
	ARG_47@NH2	GLU_205@OE1	0.2923	2.7809	156.771
	ARG_23@NH1	GLU_205@OE1	0.2788	2.8025	160.396
	ARG_25@NH2	GLU_198@OE1	0.2696	2.8067	159.1859
	ARG_25@NH1	GLU_198@OE2	0.2653	2.7963	158.3053
	ARG_288@NH2	GLU_192@OE1	0.2587	2.7908	160.0791
	ARG_23@NH2	GLU_205@OE1	0.2433	2.8089	161.8708
	ARG_288@NH1	GLU_192@OE2	0.2199	2.8177	161.3264
	ARG_25@NH2	GLU_198@OE2	0.2072	2.8054	158.0281
	ARG_25@NH1	GLU_198@OE1	0.1982	2.8128	158.004
	ARG_47@NE	GLU_205@OE1	0.1449	2.8483	150.4205
	SER_50@N	GLN_208@O	0.1402	2.858	148.878
	LYS_22@NZ	GLU_202@OE2	0.1386	2.7933	157.0229
	ARG_20@NH2	GLU_201@OE1	0.1259	2.7912	160.3325
	LYS_22@NZ	GLU_202@OE1	0.1191	2.7997	157.4864
	TRP_54@N	GLU_25@OE2	0.1139	2.8776	161.1172
	ASN_378@ND2	GLU_198@OE2	0.113	2.8284	159.6941
	ARG_20@NE	GLU_201@OE2	0.1029	2.8394	160.4648

5A.4.3 Protein-protein interface characterization

PPCs/PPIs of any living system are usually characterized by the presence of an interface area having a high number of hydrophobic interactions, and a large solvent-accessible

surface area [327, 393, 455]. The interface residues of any PPC are further characterized by their contact distance, which should be less than 6 Å from its interacting partner proteins [458]. So, to study the PPI between $XPA_{185-226}$ and XPE, we employed the Dimplot program on the PDBsum server to help us identify and characterize the interface residues present in the lowest energy conformer of our protein complex.

The interface statistics for our protein complex are shown in **Table 5A.3**, while the cartoon representation of our lowest energy conformers for all the models can be seen in **Figures 5A.9A 5A.9C,** and **5A.9E**. All the residue interactions for our PPI present within this range (less than 6 Å) can be seen in **Figure 5A.9B, 5A.9D,** and **5A.9F**, and the detailed contributions of each residue required in the formation of our stable PPC are accordingly provided in **Table 5A.4**. We obtained 22 residues of XPE with an interface area of 1171 $Å^2$ to be interacting with 19 residues of $XPA_{185-226}$ having an interface surface area of 1124 $Å^2$. Likewise, Models 2 and 3 had interface residues of 18 and 20 each for XPE with an interface area of 870 $Å^2$ and 1116 $Å^2$, respectively. $XPA_{185-226}$ of models 2 and 3 was observed to have the interface residues of 18 and 22 with the interface area of 867 $Å^2$ and 1074 $Å^2$. The number of hydrogen bonds, salt bridges and hydrophobic contacts as seen in Table 2 was more or less the same for all models.

Table 5A.3 Interface statistics for $XPA_{185-226}$-XPE protein complex.

Models	Protein	No. of interface residues	Interface area ($Å^2$)	No. of salt bridges	No. of hydrogen bonds	No. of non-bonded contacts
Model 1	XPE	22	1171	7	20	147
	$XPA_{185-226}$	19	1124			
Model 2	XPE	18	870	8	17	141
	$XPA_{185-226}$	18	867			
Model 3	XPE	20	1116	9	16	140
	$XPA_{185-226}$	22	1074			

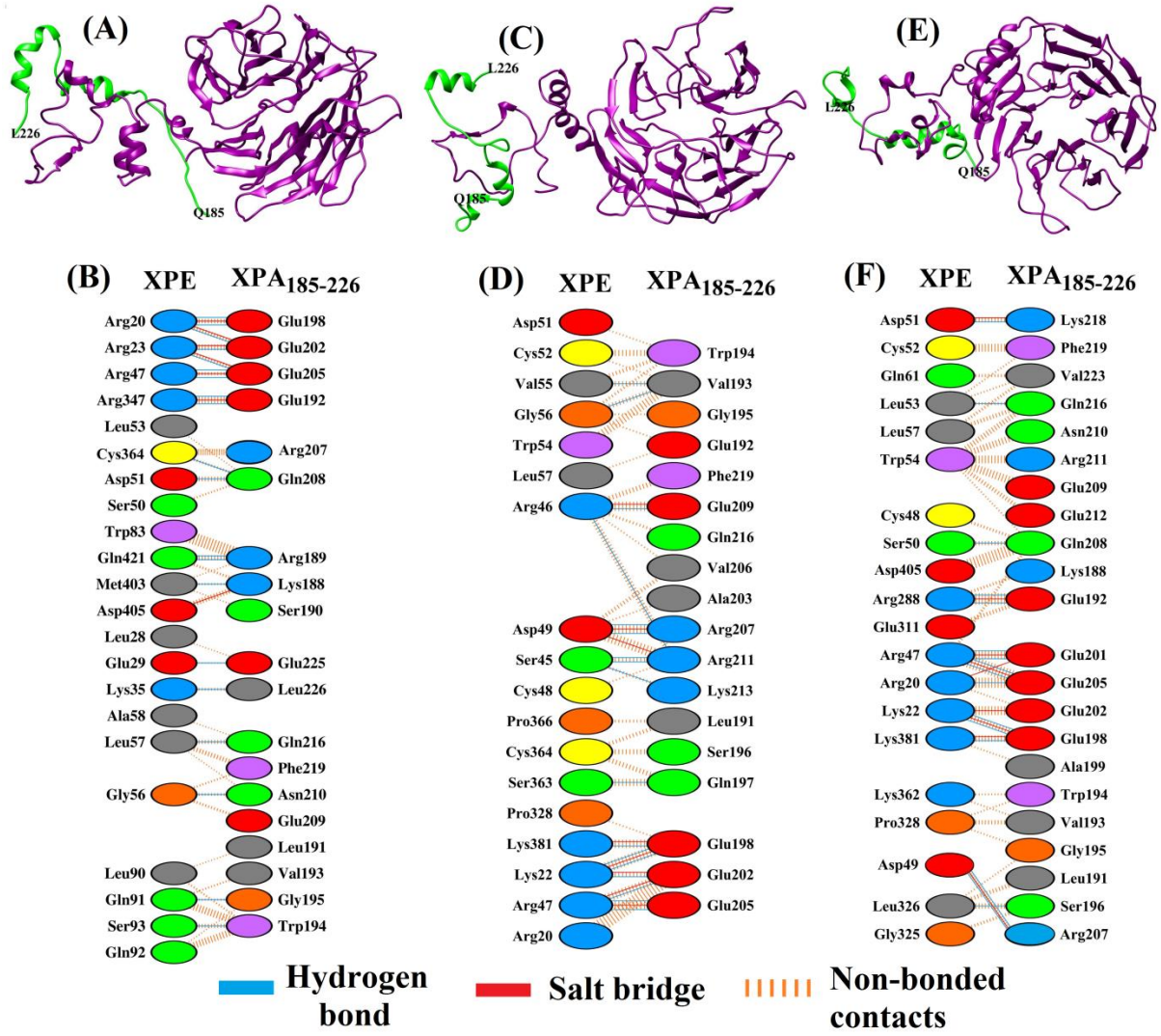

Figure 5A.9. (A) Cartoon representation and (B) Intermolecular interactions between Model 1 of $XPA_{185-226}$-XPE complex. (C) Cartoon representation and (D) Intermolecular interactions between Model 2 of $XPA_{185-226}$-XPE complex. (E) Cartoon representation and (F) Intermolecular interactions between Model 3 of $XPA_{185-226}$-XPE complex.

Table 5A.4. Intermolecular interactions across $XPA_{185-226}$ -XPE interface.
Table 5A.4A. Intermolecular hydrogen bond between $XPA_{185-226}$ and XPE protein complex.

Models	XPE			$XPA_{185-226}$		Hydrogen bond distance (Å)
	Atom name	Residue		Atom name	Residue	
Model 1	SD	MET403	<-->	NZ	LYS188	3.23
	OE1	GLN421	<-->	N	ARG189	3.00
	NE2	GLN421	<-->	O	ARG189	3.17
	NE	ARG347	<-->	OE1	GLU192	2.85
	NH2	ARG347	<-->	OE2	GLU192	2.77
	O	SER 93	<-->	NE1	TRP194	2.79
	O	GLN91	<-->	N	GLY195	2.94

	NH1	ARG20	<-->	OE1	GLU198	2.77
	NH2	ARG20	<-->	OE1	GLU198	2.75
	NH1	ARG20	<-->	OE2	GLU202	2.95
	NH2	ARG23	<-->	OE2	GLU202	2.82
	NH1	ARG23	<-->	OE1	GLU205	2.82
	NE	ARG 47	<-->	OE2	GLU205	3.28
	NH2	ARG 47	<-->	OE2	GLU205	2.86
	O	CYS364	<-->	NE2	GLN208	2.83
	O	ASP51	<-->	NE2	GLN208	2.83
	O	GLY56	<-->	N	ASN210	2.99
	O	LEU57	<-->	NE2	GLN216	3.19
	N	GLU29	<-->	OE1	GLU225	3.27
	NZ	LYS 35	<-->	O	LEU226	2.71
Model 2	N	VAL55	<-->	O	VAL193	3.18
	N	GLY56	<-->	O	VAL193	2.71
	O	SER363	<-->	NE2	GLN197	2.96
	NZ	LYS22	<-->	O	GLU198	3.01
	NZ	LYS22	<-->	OE1	GLU198	2.91
	NZ	LYS381	<-->	OE1	GLU198	2.67
	NH2	ARG47	<-->	OE1	GLU202	2.88
	NZ	LYS22	<-->	OE2	GLU202	2.69
	NH1	ARG47	<-->	OE2	GLU205	3.02
	NH2	ARG47	<-->	OE2	GLU205	2.96
	OD2	ASP49	<-->	NH1	ARG207	2.68
	OD1	ASP49	<-->	NH2	ARG207	2.86
	NH1	ARG46	<-->	OE1	GLU209	2.78
	O	SER45	<-->	NH1	ARG211	3.00
	O	ARG46	<-->	NH1	ARG211	2.68
	O	SER45	<-->	NH2	ARG211	2.93
	O	SER45	<-->	NZ	LYS213	3.29
Model 3	NE	ARG288	<-->	OE1	GLU192	2.89
	NH2	ARG288	<-->	OE2	GLU192	2.69
	N	LEU326	<-->	OG	SER196	3.07
	NZ	LYS22	<-->	O	GLU198	2.92
	NZ	LYS22	<-->	OE1	GLU198	2.79
	NZ	LYS381	<-->	OE2	GLU198	2.67
	NH2	ARG47	<-->	O	GLU201	2.74
	NH1	ARG47	<-->	OE2	GLU201	2.78
	N	ARG20	<-->	OE1	GLU205	2.67
	NE	ARG47	<-->	OE1	GLU205	2.80
	NH2	ARG47	<-->	OE1	GLU205	2.94
	N	SER50	<-->	O	GLN208	3.03
	O	LEU53	<-->	NE2	GLN216	3.00
	OD1	ASP51	<-->	NZ	LYS218	2.91
	OD2	ASP49	<-->	NH1	ARG207	2.68
	OD1	ASP49	<-->	NH2	ARG207	2.86

Table 5A.4B. Intermolecular salt bridge formation between $XPA_{185-226}$ and XPE protein complex.

Models	XPE			$XPA_{185-226}$		Salt bridge distance (Å)
	Atom name	Residue		Atom name	Residue	
Model 1	OD1	ASP405	<-->	NZ	LYS188	2.70
	NH2	ARG347	<-->	OE2	GLU192	2.77
	NH2	ARG20	<-->	OE1	GLU198	2.75
	NH1	ARG20	<-->	OE2	GLU202	2.95
	NH2	ARG23	<-->	OE2	GLU202	2.82
	NH1	ARG23	<-->	OE2	GLU205	2.82
	NH2	ARG47	<-->	OE1	GLU205	2.86
Model 2	NZ	LYS22	<-->	OE1	GLU198	2.91
	NZ	LYS381	<-->	OE1	GLU198	2.67
	NZ	LYS22	<-->	OE2	GLU202	2.69
	NH2	ARG47	<-->	OE1	GLU202	2.88
	NH2	ARG47	<-->	OE1	GLU205	2.96
	OD2	ASP49	<-->	NH1	ARG207	2.68
	NH1	ARG46	<-->	OE1	GLU209	2.78
	OD2	ASP49	<-->	NE	ARG211	3.25
Model 3	NH2	ARG288	<-->	OE2	GLU192	2.69
	NZ	LYS22	<-->	OE1	GLU198	2.79
	NZ	LYS381	<-->	OE2	GLU198	2.67
	NH1	ARG20	<-->	OE2	GLU201	3.95
	NH1	ARG47	<-->	OE2	GLU201	2.78
	NZ	LYS22	<-->	OE2	GLU202	2.97
	NE	ARG47	<-->	OE1	GLU205	2.80
	OD	ASP51	<-->	NZ	LYS218	2.91
	OD2	ASP49	<-->	NH1	ARG207	2.68

Table 5A.4C. Intermolecular non-bonded contacts formed between $XPA_{185-226}$ and XPE protein complex.

Models.	XPE			$XPA_{185-226}$		Non-bonded contacts distance (Å)
	Atom name	Residues		Residues	Atom name	
Model 1	OE1	GLN421	<-->	CA	LYS188	3.79
	OE1	GLN421	<-->	C	LYS 188	3.85
	NE2	GLN421	<-->	CB	LYS188	3.83
	SD	MET403	<-->	CD	LYS188	3.57
	OD2	ASP405	<-->	CD	LYS188	3.38
	CG	ASP405	<-->	CE	LYS188	3.80
	OD1	ASP405	<-->	CE	LYS188	3.81

	OD2	ASP405	<-->	CE	LYS188	2.97
	SD	MET403	<-->	NZ	LYS188	3.23
	CG	ASP405	<-->	NZ	LYS188	3.09
	OD1	ASP405	<-->	NZ	LYS188	2.72
	OD2	ASP405	<-->	NZ	LYS188	2.70
	CD	GLN421	<-->	N	ARG189	3.76
	OE1	GLN421	<-->	N	ARG189	3.00
	OE1	GLN421	<-->	CA	ARG189	3.86
	CE	MET403	<-->	O	ARG189	3.83
	CD	GLN421	<-->	O	ARG189	3.84
	OE1	GLN421	<-->	O	ARG189	3.58
	NE2	GLN421	<-->	O	ARG189	3.17
	OE1	GLN421	<-->	CB	ARG189	3.83
	NE1	TRP83	<-->	CD	ARG189	3.76
	CE2	TRP83	<-->	CD	ARG189	3.67
	CG	TRP83	<-->	NE	ARG189	3.59
	CD1	TRP3	<-->	NE	ARG189	3.65
	CD2	TRP83	<-->	E	ARG189	3.62
	NE1	TRP83	<-->	NE	ARG189	3.62
	CE2	TRP83	<-->	NE	ARG189	3.66
	CB	TRP83	<-->	CZ	ARG189	3.87
	CG	TRP83	<-->	CZ	ARG189	3.36
	CD1	TRP83	<-->	CZ	ARG189	3.23
	NE1	TRP83	<-->	CZ	ARG189	3.58
	CG	TRP83	<-->	NH1	ARG189	3.86
	CD1	TRP83	<-->	NH1	ARG189	3.26
	NE1	TRP83	<-->	NH1	ARG189	3.48
	CB	TRP83	<-->	NH2	ARG189	3.41
	CG	TRP83	<-->	NH2	ARG189	3.44
	CD1	TRP83	<-->	NH2	ARG189	3.61
	CE	MET403	<-->	OG	SER190	3.34
	CD2	LEU90	<-->	O	LEU191	3.47
	NE	ARG347	<-->	CD	GLU192	3.65
	NH2	ARG347	<-->	CD	GLU192	3.36
	CD	ARG347	<-->	OE1	GLU192	3.87
	NE	ARG347	<-->	OE1	GLU192	2.85
	CZ	ARG347	<-->	OE1	GLU192	3.45
	NH2	ARG347	<-->	OE1	GLU192	3.09
	NE	ARG347	<-->	OE2	GLU192	3.62
	CZ	ARG347	<-->	OE2	GLU192	3.61
	NH2	ARG347	<-->	OE2	GLU192	2.77
	CA	GLN91	<-->	O	VAL193	3.46
	CB	GLN91	<-->	O	VAL193	3.16
	O	GLN91	<-->	O	TRP194	3.72
	O	GLN91	<-->	C	TRP194	3.84
	C	GLN91	<-->	CG	TRP194	3.76
	O	GLN91	<-->	CG	TRP194	3.51
	O	LEU90	<-->	CD1	TRP194	3.35

	CA	GLN91	<-->	CD1	TRP194	3.76
	C	GLN91	<-->	CD1	TRP194	3.36
	O	GLN91	<-->	CD1	TRP194	3.54
	N	GLN92	<-->	CD1	TRP194	3.53
	O	SER93	<-->	CD1	TRP194	3.68
	O	GLN91	<-->	CD2	TRP194	3.79
	O	LEU90	<-->	NE1	TRP194	3.65
	C	GLN91	<-->	NE1	TRP194	3.64
	O	GLN91	<-->	NE1	TRP194	3.78
	N	GLN92	<-->	NE1	TRP194	3.35
	CA	GLN92	<-->	NE1	TRP194	3.35
	C	GLN92	<-->	NE1	TRP194	3.25
	O	GLN92	<-->	NE1	TRP194	3.80
	N	SER93	<-->	NE1	TRP194	3.25
	C	SER93	<-->	NE1	TRP194	3.76
	O	SER93	<-->	NE1	TRP194	2.79
	CA	GLN92	<-->	CE2	TRP194	3.71
	C	GLN92	<-->	CE2	TRP194	3.76
	O	SER93	<-->	CE2	TRP194	3.72
	C	GLN92	<-->	CZ2	TRP194	3.88
	O	GLN92	<-->	CZ2	TRP194	3.60
	O	GLN91	<-->	N	GLY195	2.94
	O	GLN91	<-->	CA	GLY195	3.70
	E2	GLN92	<-->	CA	GLY195	3.79
	NH1	ARG20	<-->	CD	GLU198	3.56
	NH2	ARG20	<-->	CD	GLU198	3.80
	CZ	ARG20	<-->	OE1	GLU198	3.24
	NH1	ARG20	<-->	OE1	GLU198	2.77
	NH2	ARG20	<-->	OE1	GLU198	2.75
	NH1	ARG20	<-->	OE2	GLU198	3.62
	NH1	ARG20	<-->	CB	GLU202	3.51
	NH1	ARG20	<-->	CG	GLU202	3.29
	NH1	ARG20	<-->	CD	GLU202	3.44
	NH1	ARG23	<-->	CD	GLU202	3.67
	NH2	ARG23	<-->	CD	GLU202	3.40
	NH2	ARG23	<-->	OE1	GLU202	3.82
	NH1	ARG23	<-->	OE1	GLU202	3.36
	NH2	ARG23	<-->	OE1	GLU202	3.22
	CD	ARG20	<-->	OE2	GLU202	3.55
	CZ	ARG20	<-->	OE2	GLU202	3.82
	NH1	ARG20	<-->	OE2	GLU202	2.95
	CZ	ARG23	<-->	OE2	GLU202	3.60
	NH1	ARG23	<-->	OE2	GLU202	3.55
	NH2	ARG23	<-->	OE2	GLU202	2.82
	NH1	ARG23	<-->	CD	GLU205	3.51
	NH2	ARG 47	<-->	CD	GLU205	3.67
	CD	ARG23	<-->	OE1	GLU205	3.45
	CZ	ARG23	<-->	OE1	GLU205	3.82

	NH1	ARG23	<-->	OE1	GLU205	2.82
	CD	ARG23	<-->	OE2	GLU205	3.71
	NH1	ARG23	<-->	OE2	GLU205	3.65
	NE	ARG47	<-->	OE2	GLU205	3.28
	CZ	ARG47	<-->	OE2	GLU205	3.42
	NH2	ARG47	<-->	OE2	GLU205	2.86
	O	ASP51	<-->	CG	GLN208	3.78
	SG	CYS364	<-->	O	ARG207	3.58
	N	SER50	<-->	CD	GLN208	3.72
	O	ASP51	<-->	CD	GLN208	3.73
	O	CYS364	<-->	OE1	GLN208	3.27
	O	CYS364	<-->	NE2	GLN208	2.83
	N	SER50	<-->	NE2	GLN208	3.85
	N	ASP51	<-->	NE2	GLN208	3.27
	CA	ASP51	<-->	NE2	GLN208	3.75
	C	ASP51	<-->	NE2	GLN208	3.55
	O	ASP51	<-->	NE2	GLN208	2.83
	CB	LEU53	<-->	NE2	GLN208	3.73
	O	GLY56	<-->	CA	GLU209	3.68
	O	GLY56	<-->	C	GLU209	3.82
	O	GLY56	<-->	CB	GLU209	3.57
	O	GLY56	<-->	N	ASN210	2.99
	O	GLY56	<-->	CA	ASN210	3.77
	O	GLY56	<-->	CB	ASN210	3.37
	O	LEU57	<-->	OD1	ASN210	3.88
	CD1	LEU57	<-->	O	GLN216	3.66
	CD1	LEU57	<-->	CB	GLN216	3.78
	O	LEU57	<-->	CG	GLN216	3.83
	CD1	LEU57	<-->	CG	GLN216	3.81
	O	LEU57	<-->	NE2	GLN216	3.19
	O	ALA58	<-->	NE2	GLN216	3.40
	CA	LEU57	<-->	CD1	PHE219	3.69
	CB	LEU57	<-->	CD1	PHE219	3.77
	CD1	LEU57	<-->	CD1	PHE219	3.83
	C	GLY56	<-->	CE1	PHE219	3.65
	N	LEU57	<-->	CE1	PHE219	3.48
	CA	LEU57	<-->	CE1	PHE219	3.62
	CB	LEU57	<-->	CE1	PHE219	3.82
	C	GLY56	<-->	CZ	PHE219	3.80
	O	LEU28	<-->	OE1	GLU225	3.68
	N	GLU29	<-->	OE1	GLU225	3.27
	NZ	LYS35	<-->	C	LEU226	3.75
	CE	LYS35	<-->	O	LEU226	3.60
	NZ	LYS35	<-->	O	LEU226	2.71
	O	CYS364	<-->	CD1	LEU191	3.75
	CG	PRO366	<-->	CD1	LEU191	3.85
	CD 6	PRO36	<-->	CD1	LEU191	3.46
	CA	CYS364	<-->	CD2	LEU191	3.85

Model 2	O	CYS364	<-->	CD2	LEU191	3.77
	CA	GLY56	<-->	O	GLU192	3.56
	N	LEU57	<-->	O	GLU192	3.79
	N	GLY 56	<-->	C	VAL193	3.20
	CA	GLY56	<-->	C	VAL193	3.73
	SG	CYS52	<-->	O	VAL193	3.79
	CB	TRP54	<-->	O	VAL193	3.88
	N	VAL55	<-->	O	VAL193	3.18
	CA	VAL55	<-->	O	VAL193	3.82
	C	VAL55	<-->	O	VAL193	3.54
	CG2	VAL55	<-->	O	VAL193	3.66
	N	GLY56	<-->	O	VAL193	2.71
	CA	GLY56	<-->	O	VAL193	3.70
	CD2	TRP54	<-->	CB	VAL193	3.67
	CE3	TRP54	<-->	CB	VAL193	3.70
	CD2	TRP54	<-->	CG1	VAL193	3.66
	CE2	TRP54	<-->	CG1	VAL193	3.65
	CZ2	TRP54	<-->	CG1	VAL193	3.89
	CZ3	TRP54	<-->	CG2	VAL193	3.86
	N	GLY56	<-->	N	TRP194	3.90
	CG2	VAL55	<-->	CA	TRP194	3.84
	N	GLY56	<-->	C	TRP194	3.83
	CG2	VAL55	<-->	O	TRP194	3.83
	SG	CYS52	<-->	CE2	TRP194	3.73
	CG2	VAL55	<-->	CE3	TRP194	3.83
	SG	CYS52	<-->	CZ2	TRP194	3.58
	CB	CYS52	<-->	CZ3	TRP194	3.76
	O	ASP51	<-->	CH2	TRP194	3.63
	CB	CYS52	<-->	CH2	TRP194	3.60
	SG	CYS52	<-->	CH2	TRP194	3.80
	CA	GLY56	<-->	N	GLY195	3.70
	CB	CYS364	<-->	CA	SER 196	3.79
	SG	CYS364	<-->	CA	SER 196	3.75
	SG	CYS364	<-->	CB	SER196	3.75
	SG	CYS364	<-->	OG	SER196	3.32
	CB	CYS364	<-->	N	GLN197	3.82
	SG	CYS364	<-->	N	GLN197	3.36
	C	SER363	<-->	CB	GLN197	3.61
	O	SER363	<-->	CB	GLN197	3.70
	N	CYS364	<-->	CB	GLN197	3.87
	SG	CYS364	<-->	CB	GLN197	3.78
	O	SER363	<-->	CG	GLN197	3.81
	O	SER363	<-->	CD	GLN197	3.86
	C	SER363	<-->	NE2	GLN197	3.86
	O	SER363	<-->	NE2	GLN197	2.96
	NZ	LYS 22	<-->	C	GLU198	3.66
	NZ	LYS22	<-->	O	GLU198	3.01
	CG	PRO328	<-->	CG	GLU198	3.83

	CE	LYS 22	<-->	CD	GLU198	3.49
	NZ	LYS 22	<-->	CD	GLU198	3.43
	CD	LYS381	<-->	CD	GLU198	3.57
	CE	LYS381	<-->	CD	GLU198	3.48
	NZ	LYS381	<-->	CD	GLU198	3.44
	CD	LYS22	<-->	OE1	GLU198	3.64
	CE	LYS22	<-->	OE1	GLU198	3.12
	NZ	LYS22	<-->	OE1	GLU198	2.91
	CD	LYS381	<-->	OE1	GLU198	3.58
	CE	LYS381	<-->	OE1	GLU198	3.22
	NZ	LYS381	<-->	OE1	GLU198	2.67
	CE	LYS22	<-->	OE2	GLU198	3.24
	NZ	LYS22	<-->	OE2	GLU198	3.61
	CD	LYS381	<-->	OE2	GLU198	3.54
	NE	ARG 47	<-->	O	GLU202	3.62
	CZ	ARG47	<-->	O	GLU202	3.83
	NH2	ARG47	<-->	CB	GLU202	3.54
	NH2	ARG47	<-->	CG	GLU202	3.40
	N	ARG20	<-->	CD	GLU202	3.23
	CA	ARG20	<-->	CD	GLU202	3.62
	C	ARG20	<-->	CD	GLU202	3.86
	O	ARG20	<-->	CD	GLU202	3.89
	NZ	LYS22	<-->	CD	GLU202	3.81
	NH2	ARG47	<-->	CD	GLU202	3.53
	N	ARG20	<-->	OE1	GLU 202	3.22
	CA	ARG20	<-->	OE1	GLU202	3.07
	C	ARG20	<-->	OE1	GLU202	3.33
	O	ARG20	<-->	OE1	GLU202	3.13
	CZ	ARG47	<-->	OE1	GLU202	3.84
	NH2	ARG47	<-->	OE1	GLU202	2.88
	N	AR20	<-->	OE2	GLU202	3.02
	CA	ARG20	<-->	OE2	GLU202	3.63
	C	ARG20	<-->	OE2	GLU202	3.60
	O	ARG20	<-->	OE2	GLU202	3.83
	CE	LYS22	<-->	OE2	GLU202	3.78
	NZ	LYS22	<-->	OE2	GLU202	2.69
	OD1	ASP49	<-->	O	ALA203	3.68
	NH1	ARG47	<-->	CB	GLU205	3.79
	NH1	ARG47	<-->	CG	GLU205	3.55
	NH1	ARG47	<-->	CD	GLU205	3.72
	NH2	ARG 47	<-->	CD	GLU05	3.82
	CZ	ARG47	<-->	OE2	GLU205	3.42
	NH1	ARG47	<-->	OE2	GLU205	3.02
	NH2	ARG47	<-->	OE2	GLU205	2.96
	CB	ASP49	<-->	CG1	VAL206	3.87
	CG	ASP49	<-->	CG1	VAL206	3.73
	OD1	ASP49	<-->	CG1	VAL206	3.64
	NH1	ARG46	<-->	CG2	VAL206	3.55

	OD1	ASP49	<-->	CZ	ARG207	3.84
	OD2	ASP49	<-->	CZ	ARG207	3.32
	CG	ASP49	<-->	NH1	ARG207	3.60
	OD2	ASP49	<-->	NH1	ARG207	2.68
	CG	ASP49	<-->	NH2	ARG207	3.29
	OD1	ASP49	<-->	NH2	ARG 207	2.86
	OD2	ASP49	<-->	NH2	ARG207	3.11
	NH1	ARG46	<-->	CB	GLU209	3.77
	NH1	ARG46	<-->	CD	GLU209	3.61
	CB	ARG46	<-->	OE1	GLU209	3.54
	CG	ARG46	<-->	OE1	GLU209	3.41
	CD	ARG46	<-->	OE1	GLU209	3.48
	CZ	ARG46	<-->	OE1	GLU209	3.75
	NH1	ARG46	<-->	OE1	GLU209	2.78
	CB	ARG 46	<-->	OE2	GLU209	3.77
	OD2	ASP49	<-->	CG	ARG211	3.58
	OD2	ASP49	<-->	CD	ARG 211	3.43
	CB	ASP49	<-->	NE	ARG211	3.19
	CG	ASP49	<-->	NE	ARG211	3.74
	OD2	ASP49	<-->	NE	ARG 211	3.25
	O	SER45	<-->	CZ	ARG 211	3.40
	O	ARG46	<-->	CZ	ARG 211	3.56
	N	ASP49	<-->	CZ	ARG 211	3.82
	CB	ASP49	<-->	CZ	ARG 211	3.29
	O	SER45	<-->	NH1	ARG 211	3.00
	C	ARG46	<-->	NH1	ARG 211	3.71
	O	ARG46	<-->	NH1	ARG 211	2.68
	CB	ARG46	<-->	NH1	ARG 211	3.74
	O	SER45	<-->	NH2	ARG 211	2.93
	O	ARG46	<-->	NH2	ARG 211	3.64
	SG	CYS48	<->	NH2	ARG211	3.67
	N	ASP 49	<-->	NH2	ARG211	3.48
	CB	ASP 49	<-->	NH2	ARG 211	3.43
	O	SER 45	<-->	CE	LYS 213	3.73
	O	SER 45	<-->	NZ	LYS213	3.29
	CB	ARG 46	<-->	OE1	GLN216	3.56
	CD	ARG46	<-->	OE1	GLN216	3.64
	CD	ARG46	<-->	CE2	PHE219	3.79
	CD	ARG46	<-->	CZ	PHE219	3.84
	NE	ARG46	<-->	CZ	PHE219	3.50
	CZ	ARG46	<-->	CZ	PHE219	3.57
Model 3	NH2	ARG 288	<-->	CD	LYS 188	3.79
	NH2	ARG 288	<-->	CE	LYS 188	3.29
	CD1	LEU 326	<-->	C	LEU 191	3.73
	CD1	LEU 326	<-->	O	LEU 191	3.31
	CD2	LEU 326	<-->	O	LEU 191	3.61
	CG	GLU 311	<-->	O	GLU 192	3.30
	CD	GLU 311	<-->	O	GLU 192	3.45

	OE1	GLU 311	<-->	O	GLU 192	3.76
	NE	ARG 288	<-->	CD	GLU 192	3.58
	NH2	ARG 288	<-->	CD	GLU 192	3.46
	CD	ARG 288	<-->	OE1	GLU 192	3.83
	NE	ARG 288	<-->	OE1	GLU 192	2.89
	CZ	ARG 288	<-->	OE1	GLU 192	3.49
	NH2	ARG 288	<-->	OE1	GLU 192	3.34
	NE	ARG 288	<-->	OE2	GLU 192	3.35
	CZ	ARG 288	<-->	OE2	GLU 192	3.41
	NH2	ARG 288	<-->	OE2	GLU 192	2.69
	CG	PRO 328	<-->	C	VAL 193	3.81
	CB	PRO 328	<-->	O	VAL 193	3.56
	CG	PRO 328	<-->	O	VAL 193	3.24
	CB	LYS 362	<-->	CG1	VAL 193	3.80
	CG	PRO 328	<-->	C	TRP 194	3.90
	O	LYS 362	<-->	CH2	TRP 194	3.69
	CG	PRO 328	<-->	N	GLY 195	3.72
	O	LEU 326	<-->	CA	GLY 195	3.61
	O	LEU 326	<-->	N	SER 196	3.70
	CA	GLY 325	<-->	OG	SER 196	3.46
	C	GLY 325	<-->	OG	SER 196	3.74
	N	LEU 326	<-->	OG	SER 196	3.07
	O	LEU 326	<-->	OG	SER 196	3.72
	CB	LEU 326	<-->	OG	SER 196	3.82
	NZ	LYS 22	<-->	C	GLU 198	3.56
	NZ	LYS 22	<-->	O	GLU 198	2.92
	NZ	LYS 22	<-->	CD	GLU 198	3.86
	NZ	LYS 381	<-->	CD	GLU 198	3.34
	CE	LYS 22	<-->	OE1	GLU 198	3.64
	NZ	LYS 22	<-->	OE1	GLU 198	2.79
	CE	LYS 381	<-->	OE1	GLU 198	3.49
	NZ	LYS 381	<-->	OE1	GLU 198	3.19
	CE	LYS 381	<-->	OE2	GLU 198	3.68
	NZ	LYS 381	<-->	OE2	GLU 198	2.67
	NZ	LYS 381	<-->	CB	ALA 199	3.90
	NH2	ARG 47	<-->	C	GLU 201	3.87
	CZ	ARG 47	<-->	O	GLU 201	3.55
	NH1	ARG 47	<-->	O	GLU 201	3.71
	NH2	ARG 47	<-->	O	GLU 201	2.74
	CZ	ARG 47	<-->	OE2	GLU 201	3.59
	NH1	ARG 47	<-->	OE2	GLU 201	2.78
	NH2	ARG 47	<-->	OE2	GLU 201	3.58
	NZ	LYS 22	<-->	CD	GLU 202	3.31
	N	ARG 20	<-->	OE1	GLU 202	3.18
	CD	LYS 22	<-->	OE1	GLU 202	3.71
	CE	LYS 22	<-->	OE1	GLU 202	3.49
	NZ	LYS 22	<-->	OE1	GLU 202	3.19
	CD	LYS 22	<-->	OE2	GLU 202	3.75

	CE	LYS 22	<-->	OE2	GLU 202	3.77
	NZ	LYS 22	<-->	OE2	GLU 202	2.97
	NE	ARG 47	<-->	CB	GLU 205	3.83
	CZ	ARG 47	<-->	CB	GLU 205	3.79
	NH2	ARG 47	<-->	CB	GLU 205	3.61
	NE	ARG 47	<-->	CG	GLU 205	3.47
	N	ARG 20	<-->	CD	GLU 205	3.17
	CA	ARG 20	<-->	CD	GLU 205	3.72
	NE	ARG 47	<-->	CD	GLU 205	3.53
	NH2	ARG 47	<-->	CD	GLU 205	3.64
	N	ARG 20	<-->	OE1	GLU 205	2.67
	CA	ARG 20	<-->	OE1	GLU 205	3.01
	CB	ARG 20	<-->	OE1	GLU 205	3.69
	CG	ARG 20	<-->	OE1	GLU 205	3.28
	CD	ARG 47	<-->	OE1	GLU 205	3.87
	NE	ARG 47	<-->	OE1	GLU 205	2.80
	CZ	ARG 47	<-->	OE1	GLU 205	3.30
	NH2	ARG 47	<-->	OE1	GLU 205	2.94
	N	ARG 20	<-->	OE2	GLU 205	2.94
	CA	ARG 20	<-->	OE2	GLU 205	3.64
	CG	ASP49	<-->	NH2	ARG207	3.29
	OD1	ASP49	<-->	NH2	ARG 207	2.86
	OD2	ASP49	<-->	NH2	ARG207	3.11
	N	SER 50	<-->	O	GLN 208	3.03
	CA	SER 50	<-->	O	GLN 208	3.28
	OG	SER 50	<-->	O	GLN 208	3.85
	CD1	TRP 54	<-->	O	GLN 208	3.73
	OD1	ASP 49	<-->	CG	GLN 208	3.76
	N	ASP 49	<-->	CD	GLN 208	3.48
	CA	ASP 49	<-->	CD	GLN 208	3.53
	CB	ARG 47	<-->	OE1	GLN 208	3.69
	CD	ARG 47	<-->	OE1	GLN 208	3.80
	C	CYS 48	<-->	OE1	GLN 208	3.24
	O	CYS 48	<-->	OE1	GLN 208	3.57
	N	ASP 49	<-->	OE1	GLN 208	3.10
	CA	ASP 49	<-->	OE1	GLN 208	3.58
	N	ASP 49	<-->	NE2	GLN 208	3.59
	CA	ASP 49	<-->	NE2	GLN 208	3.65
	CB	ASP 49	<-->	NE2	GLN 208	3.71
	OD1	ASP 49	<-->	NE2	GLN 208	3.62
	CG	TRP 54	<-->	CA	GLU 209	3.82
	CD1	TRP 54	<-->	CA	GLU 209	3.83
	CB	TRP 54	<-->	CG	GLU 209	3.82
	CG	TRP 54	<-->	CG	GLU 209	3.90
	CD2	TRP 54	<-->	CG	GLU 209	3.86
	CE3	TRP 54	<-->	CG	GLU 209	3.70
	CD2	TRP 54	<-->	N	ASN 210	3.76
	NE1	TRP 54	<-->	N	ASN 210	3.78

	CE2	TRP 54	<-->	N	ASN 210	3.57
	CZ2	TRP 54	<-->	N	ASN 210	3.87
	CZ2	TRP 54	<-->	CB	ASN 210	3.58
	CE3	TRP 54	<-->	N	ARG 211	3.84
	CZ3	TRP 54	<-->	N	ARG 211	3.54
	CH2	TRP 54	<-->	N	ARG 211	3.66
	CZ3	TRP 54	<-->	CA	ARG 211	3.68
	CZ3	TRP 54	<-->	C	ARG 211	3.67
	CH2	TRP 54	<-->	C	ARG 211	3.87
	CZ3	TRP 54	<-->	N	GLU 212	3.79
	CH2	TRP 54	<-->	N	GLU 212	3.50
	CH2	TRP 54	<-->	O	GLU 212	3.58
	O	LEU 53	<-->	CG	GLN 216	3.90
	CD2	LEU 57	<-->	CG	GLN 216	3.74
	CZ3	TRP 54	<-->	CD	GLN 216	3.49
	CZ3	TRP 54	<-->	OE1	GLN 216	3.83
	O	LEU 53	<-->	NE2	GLN 216	3.00
	CE3	TRP 54	<-->	NE2	GLN 216	3.48
	CZ3	TRP 54	<-->	NE2	GLN 216	3.32
	CG	LEU 57	<-->	NE2	GLN 216	3.74
	CD2	LEU 57	<-->	NE2	GLN 216	3.75
	OD1	ASP 51	<-->	CE	LYS 218	.73
	CG	ASP 51	<-->	NZ	LYS 218	3.38
	OD1	ASP 51	<-->	NZ	LYS 218	2.91
	OD2	ASP 51	<-->	NZ	LYS 218	2.98
	CD2	LEU 53	<-->	CB	PHE 219	3.74
	C	CYS 52	<-->	CD2	PHE 219	3.82
	O	CYS 52	<-->	CD2	PHE 219	3.22
	CA	CYS 52	<-->	CE2	PHE 219	3.80
	C	CYS 52	<-->	CE2	PHE 219	3.62
	O	CYS 52	<-->	CE2	PHE 219	3.03
	CA	CYS 52	<-->	CZ	PHE 219	3.84
	NE2	GLN 61	<-->	O	VAL 223	3.13
	NE2	GLN 61	<-->	CB	VAL 223	3.73
	CD2	LEU 53	<-->	CG1	VAL 223	3.78
	CD1	LEU 57	<-->	CG1	VAL 223	3.82
	O	LEU 57	<-->	CG2	VAL 223	3.58

The PPIs of any PPCs can be classified into a transient or permanent interaction based upon their interface area size, the charges carried by the interacting residues [395, 455, 456], the number of hydrogen bonds, and the salt bridge formations [457]. Since the interface size of individual proteins to form the PPC was less than 1500 $Å^2$ for all the models and has the presence of polar charged residues, the interaction between XPA and XPE can be termed as the strong transient PPI (TPPI).

From **Figure 5A.9,** we can see that the most common residues of XPE to be involved in this strong TPPI among all three models were mainly Arg20, Arg47, Asp51, and Leu57. The common residues of $XPA_{185-226}$ responsible for the interaction among three models were Leu191, Gln192, Val193, Trp194, Glu198, Glu202, Glu205, Arg207, Glu209, Gln216, and Phe219. The residue of XPA, R207, which upon mutation to glycine (R207G) had resulted in the inhibition of the XPA-XPE binding in the earlier study [151], was seen to be involved in the PPI with XPE in all three cases. R207 of Model 1's XPA formed a hydrophobic contact with Cys364 of XPE, while R207 of Model 2 and 3 were seen to have interacted with Asp49 of XPE by forming a salt bridge, hydrophobic contact, and hydrogen bond. This means that R207 is very crucial for the binding of XPA and XPE. This particular residue has also been stated to be important for the XPA-DNA interaction [149], and if mutated to other residues, it could lead to severe neurological impairment and cancer [459, 460]. Q208, which reportedly causes classical XP-A phenotype upon mutation [461] was seen to have interacted with four residues of XPE in the case of Model 1 and five residues of XPE in the case of Model 3 but showed no signs of PPI in Model 3. Q208 also has recently been found to be involved in the DNA binding upon the NMR chemical shift perturbation (CSP) assay on the DBD of XPA by Sugitani and the group [176]. The same study also found the weak DNA binding activity of XPA upon the mutation of residue L191V, which in this study was common to all models for conducting the intermolecular interaction with XPE, further affirming its importance to the functioning of XPE in terms of PPI in the working of NER.

Additionally, we had also done a PPI profile study between the DBD of XPA (aa98-239) and XPE to see whether they showed any interactions or not. We extracted the lowest energy conformer of this complex from the highly populated clusters using the RMSD clustering algorithm and then submitted this structure to the PDBsum server for the PPI analysis. From the results obtained, we observed that the residues ranging between aa98-184 were not involved in the PPI between XPA and XPE, and only the residues between aa185-226 were seen to participate in the PPI between the two proteins. The detailed results of these analyses can be found in the material (**Figure 5A.10**, and **Table 5A.5**). This further proves that $XPA_{185-226}$ is the main driving force

behind the binding of XPA and XPE, and their interactions, which is again in the agreement with Wakasugi's findings.

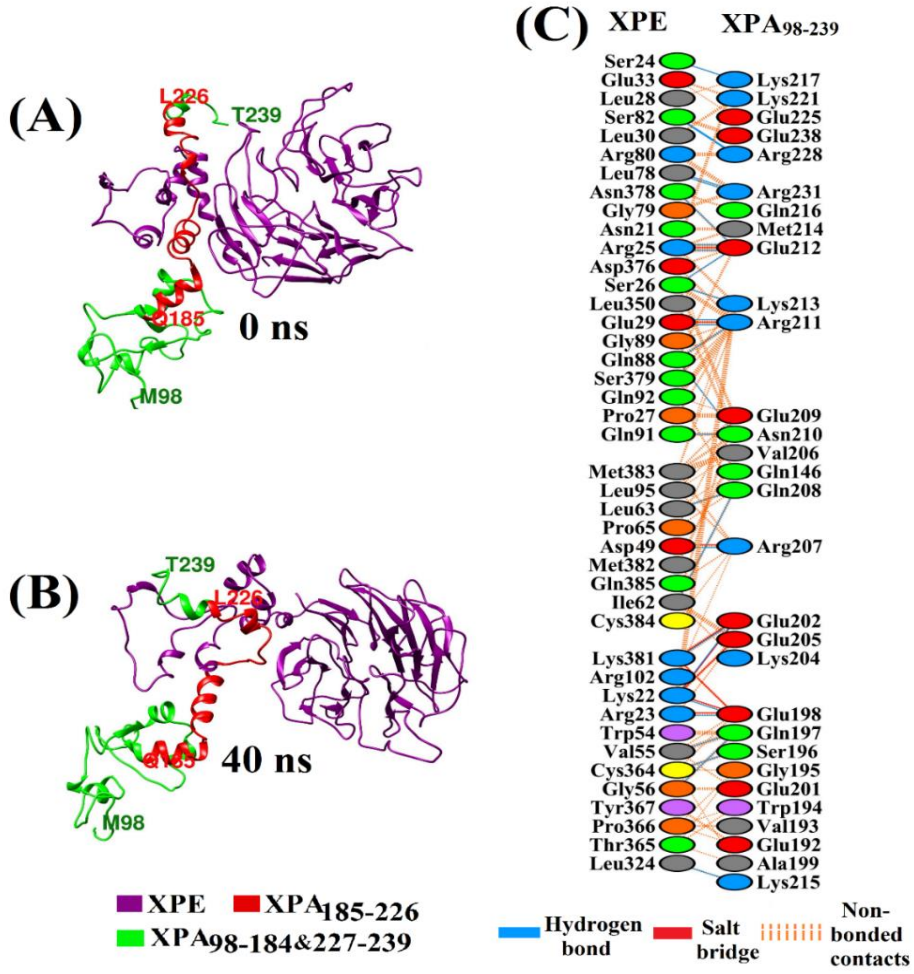

Figure 5A.10. Cartoon representation of the XPA_{98-239}-XPE complex at (A) 0 ns, and (B) 40 ns. (C) Intermolecular interactions between XPA_{98-239}-XPE complex

Table 5A.5A. ***Intermolecular interactions across XPA_{98-239}-XPE interface.***

Protein	No. of interface residues	Interface area (Å^2)	No. of salt bridges	No. of hydrogen bonds	No. of non-bonded contacts
XPE	22	1171	9	26	394
XPA_{98-239}	19	1124			

Table 5A.5B. Intermolecular hydrogen bond between XPA$_{98-239}$ and XPE protein complex.

XPE			XPA$_{98-239}$		Hydrogen bond distance (Å)
Atom name	Residue		Atom name	Residue	
NZ	LYS22	<-->	OE2	GLU198	***2.78***
NZ	LYS22	<-->	OE2	GLU202	2.55
NH1	ARG23	<-->	OE2	GLU198	2.67
O	SER24	<-->	NZ	LYS217	2.74
NE	ARG25	<-->	OE2	GLU12	2.82
NH1	ARG25	<-->	O	GLU212	2.89
OG	SER26	<-->	O	GLU212	2.99
OG	SER26	<-->	NZ	LY213	2.46
OE2	GLU29	<-->	NH1	ARG211	2.72
OE2	GLU29	<-->	NH2	ARG211	2.75
N	VAL55	<-->	OE1	GLN197	2.96
N	GLY56	<-->	OE1	GLN197	2.82
O	ILE62	<-->	NE2	GLN208	2.96
O	LEU63	<-->	NE2	GLN208	2.71
O	LEU78	<-->	NH1	ARG231	2.78
O	LEU78	<-->	NH2	ARG231	2.67
OG	SER82	<-->	NE	ARG228	2.96
OG	SER82	<-->	NH2	ARG228	2.64
O	GLN88	<-->	NH1	ARG211	2.94
NE2	GLN88	<-->	O	ASN210	3.14
NE2	GLN91	<-->	O	ASN210	2.96
OD1	ASP99	<-->	NH1	AR207	2.60
O	LEU324	<-->	NZ	LYS215	2.57
O	CYS364	<-->	OG	SER196	2.73
ND2	ASN378	<-->	OE1	GL212	2.95
NZ	LYS381	<-->	OE1	GLU202	2.54

Table 5A.5C. Intermolecular salt bridge formation between XPA$_{98-239}$ and XPE protein complex.

XPE			XPA$_{98-239}$		Salt bridge distance (Å)
Atom name	Residue		Atom name	Residue	
NZ	LYS22	<-->	OE1	GLU198	2.78
NZ	LYS22	<-->	OE2	GLU202	2.55
NZ	LYS22	<-->	OE1	GLU205	2.71
NH1	ARG23	<-->	OE1	GLU198	2.67
NH2	ARG25	<-->	OE2	GLU212	2.77
OE2	GLU29	<-->	NH1	ARG211	2.72
OD1	ASP99	<-->	NH1	ARG207	2.60
NZ	LYS381	<-->	OE2	GLU198	2.72
NZ	LYS381	<-->	OE1	GLU202	2.54

Table 5A.5D. Intermolecular Non-bonded Contacts formed between XPA_{98-239} and XPE protein complex.

XPE			XPA_{98-239}		Non-bonded contacts (Å)
Atom name	Residue		Atom name	Residue	
C	ASN21	<-->	CB	MET214	3.87
O	ASN21	<-->	CB	MET214	3.26
CG	ASN21	<-->	SD	MET214	3.54
ND2	ASN21	<-->	CG	MET214	3.83
ND2	ASN21	<-->	SD	MET214	3.08
ND2	ASN22	<-->	CE	LYS213	3.68
CA	LYS22	<-->	CB	LYS213	3.66
CA	LYS22	<-->	CG	LYS213	3.89
C	LYS22	<-->	CD	LYS213	3.75
O	LYS22	<-->	CG	LYS213	3.87
O	LYS22	<-->	CD	LYS213	2.96
O	LYS22	<-->	CE	LYS213	30.61
O	LYS22	<-->	NZ	LYS213	3.36
CB	LYS22	<-->	CB	LYS213	3.42
CB	LYS22	<-->	CG	LYS213	3.36
CB	LYS22	<-->	CD	LYS213	3.78
CD	LYS22	<-->	OE2	GLU198	3.64
CE	LYS22	<-->	OE2	GLU198	3.68
CE	LYS22	<-->	OE2	GLU202	3.87
CE	LYS22	<-->	CD	GLU205	3.72
CE	LYS22	<-->	OE1	GLU205	3.39
CE	LYS22	<-->	OE2	GLU205	3.19
CE	LYS22	<-->	CE	LYS213	3.60
NZ	LYS22	<-->	CD	GLU198	3.83
NZ	LYS22	<-->	OE2	GLU198	2.78
NZ	LYS22	<-->	CD	GLU202	3.65
NZ	LYS22	<-->	OE2	GLU202	2.55
NZ	LYS22	<-->	CD	GLU205	3.07
NZ	LYS22	<-->	OE1	GLU205	2.73
NZ	LYS22	<-->	OE2	GLU205	2.71
CG	ARG23	<-->	OE1	GLU205	3.43
CD	ARG23	<-->	OE2	GLU198	3.83
CD	ARG23	<-->	OE1	GLU205	3.28
NH1	ARG23	<-->	CG	GLU198	3.84
NH1	ARG23	<-->	CD	GLU198	3.14
NH1	ARG23	<-->	OE1	GLU198	3.74
NH1	ARG23	<-->	OE2	GLU198	2.67
C	SER24	<-->	NZ	LYS217	3.83

O	SER24	<-->	NZ	LYS217	2.74
CA	ARG25	<-->	NZ	LYS217	3.82
C	ARG25	<-->	O	GLU212	3.48
O	ARG25	<-->	C	GLU212	3.62
O	ARG25	<-->	O	GLU212	2.64
O	ARG25	<-->	CA	LYS213	3.39
O	ARG25	<-->	CD	LYS213	3.61
CB	ARG25	<-->	N	MET214	3.74
CB	ARG25	<-->	CA	MET214	3.81
CB	ARG25	<-->	CD	LYS217	3.74
CG	ARG25	<-->	O	GLU212	3.72
CG	ARG25	<-->	CA	LYS213	3.88
CG	ARG25	<-->	C	LYS213	3.41
CG	ARG25	<-->	O	LYS213	3.62
CG	ARG25	<-->	N	MET214	3.54
CG	ARG25	<-->	CA	MET214	3.62
CG	ARG25	<-->	CB	LYS217	3.88
CG	ARG25	<-->	CD	LYS217	3.71
CD	ARG25	<-->	O	GLU212	3.78
CD	ARG25	<-->	CB	LYS217	3.88
CD	ARG25	<-->	CD	LYS217	3.56
NE	ARG25	<-->	O	GLU212	3.67
NE	ARG25	<-->	CG	GLU212	3.77
NE	ARG25	<-->	CD	GLU212	3.54
NE	ARG25	<-->	OE2	GLU212	2.82
CZ	ARG25	<-->	O	GLU212	3.42
CZ	ARG25	<-->	CB	GLU212	3.43
CZ	ARG25	<-->	CG	GLU212	3.60
CZ	ARG25	<-->	CD	GLU212	3.46
CZ	ARG25	<-->	OE2	GLU212	3.11
NH1	ARG25	<-->	C	GLU212	3.71
NH1	ARG25	<-->	O	GLU212	2.89
NH1	ARG25	<-->	CB	GLU212	3.64
NH2	ARG25	<-->	CB	GLU212	3.23
NH2	ARG25	<-->	CG	GLU212	3.45
NH2	ARG25	<-->	CD	GLU212	2.93
NH2	ARG25	<-->	OE1	GLU212	3.43
NH2	ARG25	<-->	OE2	GLU212	2.77
CA	SER26	<-->	OE1	GLU209	3.64
CA	SER26	<-->	O	GLU212	3.72

C	SER26	<-->	CD	GLU209	3.86
C	SER26	<-->	OE1	GLU209	3.02
C	SER26	<-->	CG	ARG211	3.66
O	SER26	<-->	OE1	GLU209	3.45
O	SER26	<-->	OE2	GLU209	3.80
O	SER26	<-->	CG	ARG211	3.13
O	SER26	<-->	NE	ARG211	3.81
O	SER26	<-->	CZ	ARG211	3.62
O	SER26	<-->	NH2	ARG211	3.57
CB	SER26	<-->	OE1	GLU209	3.12
CB	SER26	<-->	CB	ARG211	3.71
CB	SER26	<-->	NZ	LYS213	3.36
OG	SER26	<-->	C	ARG211	3.83
OG	SER26	<-->	O	ARG211	3.49
OG	SER26	<-->	CB	ARG211	3.21
OG	SER26	<-->	C	GLU212	3.82
OG	SER26	<-->	O	GLU212	2.99
OG	SER26	<-->	CD	LYS213	2.99
OG	SER26	<-->	CE	LYS213	3.25
OG	SER26	<-->	NZ	LYS213	2.46
N	PRO27	<-->	CD	GLU209	3.60
N	PRO27	<-->	OE1	GLU209	2.86
N	PRO27	<-->	OE2	GLU209	3.59
CA	PRO27	<-->	CD	GLU209	3.34
CA	PRO27	<-->	OE1	GLU209	3.10
CA	PRO27	<-->	OE2	GLU209	2.87
CA	PRO27	<-->	NH2	ARG211	3.68
C	PRO27	<-->	NH2	ARG211	3.63
O	PRO27	<-->	NH2	ARG211	3.15
CB	PRO27	<-->	CD	GLU209	3.77
CB	PRO27	<-->	OE1	GLU209	3.62
CB	PRO27	<-->	OE2	GLU209	3.45
CD	PRO27	<-->	OE1	GLU209	3.44
O	LEU28	<-->	NZ	LYS217	3.49
CB	GLU29	<-->	NH1	ARG211	3.63
CB	GLU29	<-->	NH2	ARG211	3.59
CG	GLU29	<-->	NH1	ARG211	3.44
CD	GLU29	<-->	CZ	ARG211	3.83
CD	GLU29	<-->	NH1	ARG211	3.43
CD	GLU29	<-->	NH2	ARG211	3.33
OE2	GLU29	<-->	CZ	ARG211	3.11
OE2	GLU29	<-->	NH1	ARG211	2.72
OE2	GLU29	<-->	NH2	ARG211	2.75
C	LEU30	<-->	NZ	LYS221	3.83
O	LEU30	<-->	NZ	LYS221	3.69

CB	LEU30	<-->	NZ	LYS221	3.78
CG	LEU30	<-->	NZ	LYS221	3.90
CD2	LEU30	<-->	CG	LYS221	3.73
CA	GLU33	<-->	CE	LYS221	3.80
O	LEU30	<-->	NZ	LYS221	3.69
CB	LEU30	<-->	NZ	LYS221	3.78
CG	LEU30	<-->	NZ	LYS221	3.90
CD2	LEU30	<-->	CG	LYS221	3.73
CA	GLU33	<-->	CE	LYS221	3.80
CA	GLU33	<-->	NZ	LYS221	3.75
CB	GLU33	<-->	CE	LYS221	3.62
CB	GLU33	<-->	NZ	LYS221	3.41
CB	GLU33	<-->	OE2	GLU225	3.84
N	TRP54	<-->	OE1	GLN197	3.77
CA	TRP54	<-->	OE1	GLN197	3.83
C	TRP54	<-->	OE1	GLN197	3.77
CB	TRP54	<-->	OE1	GLN197	3.45
CD2	TRP54	<-->	OE1	GLN197	3.88
CE3	TRP54	<-->	CD	GLN197	3.63
CE3	TRP54	<-->	OE1	GLN197	3.20
CE3	TRP54	<-->	NE2	GLN197	3.29
CZ3	TRP4	<-->	NE2	GLN197	3.41
N	VAL55	<-->	OE1	GLN197	2.96
CA	VAL55	<-->	OE1	GLN197	3.61
C	VAL55	<-->	OE1	GLN197	3.58
C	VAL55	<-->	CB	GLU201	3.76
O	VAL55	<-->	CB	GLU201	3.08
CG1	VAL55	<-->	C	GLN197	3.64
CG1	VAL55	<-->	CB	GLN197	3.64
CG1	VAL55	<-->	CG	GLN197	3.76
CG1	VAL55	<-->	CD	GLN197	3.88
CG1	VAL55	<-->	OE1	GLN197	3.23
CG1	VAL55	<-->	N	GLU198	3.63
CG2	VAL56	<-->	O	GLN197	3.84
CG2	VAL55	<-->	CG	GLU198	3.53
N	GLY56	<-->	CG	GLN197	3.54
N	GLY56	<-->	CD	GLN197	3.25
N	GLY56	<-->	OE1	GLN197	2.82
CA	GLY56	<-->	CG	GLN197	3.86
CA	GLY56	<-->	CD	GLN197	3.72
CA	GLY56	<-->	OE1	GLN197	3.69
CA	GLY56	<-->	CB	GLU201	3.73
CA	GLY56	<-->	CG	GLU201	3.70
CA	GLY56	<-->	OE1	GLU201	3.60
C	GLY56	<-->	OE1	GLU201	3.82
O	GLY56	<-->	OE1	GLU201	3.31
C	ILE62	<-->	NE2	GLN208	3.86
O	ILE62	<-->	CB	GLN208	3.51

O	ILE62	<-->	CG	GLN208	3.43
O	ILE62	<-->	CD	GLN208	3.68
O	ILE62	<-->	NE2	GLN208	2.96
CB	ILE62	<-->	CA	GLU208	3.80
CG1	ILE62	<-->	CA	GLU205	3.64
CG1	ILE62	<-->	C	GLU205	3.82
CG1	ILE62	<-->	O	GLU205	3.25
CG1	ILE62	<-->	CG	GLU205	3.53
CG1	ILE62	<-->	CG	GLU209	3.47
CG1	ILE62	<-->	CD	GLU209	3.84
CG1	ILE62	<-->	OE1	GLU209	3.80
CG2	ILE62	<-->	C	LYS204	3.54
CG2	ILE62	<-->	O	LYS204	3.14
CG2	ILE62	<-->	N	GLU205	3.57
CG2	ILE62	<-->	CA	GLU205	3.20
CG2	ILE62	<-->	C	GLU205	3.71
CG2	ILE62	<-->	O	GLU205	3.57
CD1	ILE62	<-->	CG	GLU205	3.86
CD1	ILE62	<-->	OE1	GLU209	3.50
CA	LEU63	<-->	CD	GLN208	3.82
CA	LEU63	<-->	NE	GLN208	3.33
C	LEU63	<-->	NE2	GLN208	3.38
O	LEU63	<-->	CD	GLN208	3.80
O	LEU63	<-->	NE2	GLN208	2.71
CD1	LEU63	<-->	CG	GLN208	3.85
CD2	LEU63	<-->	NE2	GLN146	3.73
CB	PRO63	<-->	CD	GLU209	3.77
CB	PRO65	<-->	OE2	GLN209	3.05
CG	PRO65	<-->	NE2	GLN208	3.72
CD	PRO65	<-->	NE2	ARG208	3.73
C	PRO78	<-->	NH1	ARG231	3.78
C	PRO78	<-->	NH2	ARG231	3.72
O	LEU78	<-->	CZ	ARG231	3.09
O	LEU78	<-->	NH1	ARG231	2.78
O	LEU78	<-->	NH2	ARG231	2.67
CA	GLY79	<-->	NH1	ARG231	3.46
C	GLY79	<-->	CZ	ARG231	3.50
C	GLY79	<-->	NH1	ARG231	3.14
CA	GLY79	<-->	NH1	ARG231	3.46
C	GLY79	<-->	CZ	ARG231	3.50
C	GLY79	<-->	NH1	ARG231	3.14
O	GLY79	<-->	CD	ARG231	3.74
O	GLY79	<-->	NE	ARG231	3.83
O	GLY79	<-->	CZ	ARG231	3.64
O	GLY79	<-->	NH1	ARG231	3.22
N	ARG80	<-->	CZ	ARG231	3.49
N	ARG80	<-->	NH1	ARG231	3.60
N	ARG80	<-->	NH2	ARG231	3.62

CA	ARG80	<-->	NE	ARG231	3.68
CA	ARG80	<-->	CZ	ARG231	3.70
CA	ARG80	<-->	NH2	ARG231	3.88
O	ARG80	<-->	CB	ARG228	3.48
O	ARG80	<-->	CG	ARG228	3.22
O	ARG80	<-->	CD	ARG228	3.59
O	ARG80	<-->	NE	ARG228	3.90
CB	ARG80	<-->	O	ARG228	3.88
CG	ARG80	<-->	O	ARG228	3.22
CG	ARG80	<-->	NE	ARG231	3.78
CG	ARG80	<-->	CZ	ARG231	3.84
CG	ARG80	<-->	NH2	ARG231	3.32
CD	ARG80	<-->	O	ARG228	3.80
NE	ARG80	<-->	O	ARG228	3.52
CB	SER82	<-->	NH2	ARG228	3.72
CB	SER82	<-->	CG	GLU238	3.88
CB	SER82	<-->	CD	GLU238	3.50
CB	SER82	<-->	OE1	GLU238	3.65
CB	SER82	<-->	OE2	GLU238	3.64
OG	SER82	<-->	NE	ARG228	2.96
OG	SER82	<-->	CZ	ARG228	3.24
OG	SER82	<-->	NH2	ARG228	2.64
OG	SER82	<-->	CG	GLU238	3.54
OG	SER82	<-->	CD	GLU238	3.09
OG	SER82	<-->	OE1	GLU238	3.65
OG	SER82	<-->	OE2	GLU238	2.80
C	GLN88	<-->	NH1	ARG211	3.41
O	GLN88	<-->	O	ASN210	3.84
O	GLN88	<-->	CD	ARG211	3.89
O	GLN88	<-->	NH1	ARG211	2.94
CD	GLN88	<-->	N	GLU212	3.69
OE1	GLN88	<-->	N	GLU212	3.62
OE1	GLN88	<-->	CB	GLU212	3.60
NE2	GLN88	<-->	O	ASN210	3.14
NE2	GLN88	<-->	CA	ARG211	3.37
NE2	GLN88	<-->	CB	ARG211	3.90
NE2	GLN88	<-->	CG	ARG211	3.41
NE2	GLN88	<-->	CD	ARG211	3.38
NE2	GLN88	<-->	NH1	ARG211	3.69
NE2	GLN88	<-->	N	GLU212	3.53
N	GLY89	<-->	NH1	ARG211	3.51
CA	GLY89	<-->	NH1	ARG211	3.18
O	GLN91	<-->	CB	ASN210	3.52
CG	GLN91	<-->	O	ASN210	3.46
CD	GLN91	<-->	O	ASN210	3.74
NE2	GLN91	<-->	C	ASN210	3.74
NE2	GLN91	<-->	O	ASN210	2.96
NE2	GLN91	<-->	CG	ASN210	3.75

NE2	GLN91	<-->	OD1	ASN210	3.27
CA	GLN92	<-->	O	GLN208	3.71
O	GLN92	<-->	O	GLN208	3.74
CB	GLN92	<-->	O	GLN208	3.27
CB	GLN92	<-->	NE	ARG211	3.78
CD	GLN92	<-->	NE	ARG211	3.62
CD	GLN92	<-->	CZ	ARG211	3.32
CD	GLN92	<-->	NH1	ARG211	3.71
CD	GLN92	<-->	NH2	ARG211	3.38
OE1	GLN92	<-->	OE2	GLU209	3.51
OE1	GLN92	<-->	NE	ARG211	3.09
OE1	GLN92	<-->	CZ	ARG211	2.98
OE1	GLN92	<-->	NH1	ARG211	3.78
OE1	GLN92	<-->	NH2	ARG211	2.82
NE2	GLN92	<-->	CZ	ARG211	3.54
NE2	GLN92	<-->	NH1	ARG211	3.59
NE2	GLN92	<-->	NH2	ARG211	3.44
CG	LEU95	<-->	ND2	ASN210	3.75
CD1	LEU95	<-->	O	VAL206	3.44
CD1	LEU95	<-->	CA	ARG207	3.55
CD1	LEU95	<-->	C	ARG207	3.80
CD1	LEU95	<-->	O	ARG207	3.46
CD2	LEU95	<-->	ND2	ASN210	3.27
CG	ASP99	<-->	NH1	ARG207	3.04
OD1	ASP99	<-->	NH1	ARG207	2.60
OD2	ASP99	<-->	CG	ARG207	3.84
OD2	ASP99	<-->	CD	ARG207	3.39
OD2	ASP99	<-->	CZ	ARG207	3.86
OD2	ASP99	<-->	NH1	ARG207	2.82
NH2	ARG102	<-->	NH1	ARG207	3.60
C	LEU324	<-->	NZ	LYS215	3.59
O	LEU324	<-->	NZ	LYA215	2.57
CG	LEU350	<-->	ND2	ASN210	3.85
CD1	LEU350	<-->	ND2	ASN210	3.53
CD2	LEU350	<-->	CG	ASN210	3.41
CD2	LEU350	<-->	OD1	ASN210	3.33
CS2	LEU350	<-->	ND2	ASN210	3.18
CA	CYS364	<-->	OG	SER196	3.47
C	CYS364	<-->	OG	SER196	3.30
O	CYS364	<-->	CA	GLY195	3.68
O	CYS364	<-->	N	SER196	3.90
O	CYS364	<-->	CB	SER196	3.74
O	CYS364	<-->	OG	SER196	2.73
CB	CYS364	<-->	OG	SER196	3.61
CB	CYS364	<-->	O	GLU198	3.86
CB	CYS364	<-->	CB	ALA199	3.50
SG	CYS364	<-->	O	GLU198	3.62
CA	THR365	<-->	CA	GLY195	3.79

CG2	THR365	<-->	C	TRP194	3.65
CG2	THR365	<-->	O	TRP194	3.14
CG2	THR365	<-->	CB	ALA199	3.52
O	PRO366	<-->	O	GLU192	3.52
O	PRO366	<-->	CA	VAL193	3.44
O	PRO366	<-->	C	VAL193	3.32
O	PRO366	<-->	O	VAL193	3.57
O	PRO366	<-->	N	TRP194	3.78
CD	PRO366	<-->	CA	GLY195	3.50
CD2	TYR367	<-->	O	GLU192	3.53
CE2	TYR367	<-->	O	GLU192	3.03
CZ	TYR367	<-->	O	GLU192	3.55
OH	TYR367	<-->	O	GLU192	3.88
CB	ASP376	<-->	OD1	ASN210	3.41
CG	ASP376	<-->	C	ARG211	3.66
CG	ASP376	<-->	O	ARG211	3.79
CG	ASP376	<-->	N	GLU212	3.68
OD2	ASP376	<-->	C	ARG211	3.19
OD2	ASP376	<-->	O	ARG211	3.14
OD2	ASP376	<-->	N	GLU212	3.10
OD2	ASP376	<-->	CA	GLU212	3.04
O	ASN378	<-->	NE2	GLN216	3.82
CB	ASN378	<-->	CB	GLU212	3.73
CB	ASN378	<-->	CG	GLU212	3.59
CB	ASN378	<-->	NE2	GLN216	3.43
CG	ASN378	<-->	NE2	GLN216	3.59
ND2	ASN378	<-->	CG	GLU212	3.88
ND2	ASN378	<-->	CD	GLU212	3.81
ND2	ASN378	<-->	OE1	GLU212	2.95
ND2	ASN378	<-->	NE2	GLN216	3.37
CB	SER379	<-->	O	ARG211	3.56
CB	SER379	<-->	N	LYS213	3.49
CB	SER379	<-->	CB	LYS213	3.68
CB	SER379	<-->	CG	LYS213	3.15
CB	SER379	<-->	CD	LYS213	3.81
OG	SER379	<-->	O	ARG211	3.53
CB	LYS381	<-->	OE2	GLU202	3.76
CG	LYS381	<-->	OE2	GLU202	3.21
CD	LYS381	<-->	CD	GLU202	3.78
CD	LYS381	<-->	OE1	GLU202	3.51
CD	LYS381	<-->	OE2	GLU202	3.29
CE	LYS381	<-->	OE1	GLU198	3.77
CE	LYS381	<-->	OE1	GLU202	3.54
CE	LYS381	<-->	OE2	GLU202	3.47
NZ	LYS381	<-->	CG	GLU198	3.46
NZ	LYS381	<-->	CD	GLU198	2.90
NZ	LYS381	<-->	OE1	GLU198	2.72
NZ	LYS381	<-->	OE2	GLU198	3.33

NZ	LYS381	<-->	CD	GLU202	2.94
NZ	LYS381	<-->	OE1	GLU202	2.54
NZ	LYS381	<-->	OE2	GLU202	2.68
C	MET382	<-->	CB	VAL206	3.85
C	MET382	<-->	CG2	VAL206	3.87
O	MET382	<-->	CB	VAL206	2.94
O	MET382	<-->	CG1	VAL206	2.97
O	MET382	<-->	CG2	VAL206	2.87
CA	MET383	<-->	O	VAL206	3.86
CA	MET383	<-->	CB	VAL206	3.30
CA	MET383	<-->	CG2	VAL206	3.60
C	MET383	<-->	O	VAL06	3.86
C	MET383	<-->	CB	VAL206	3.83
C	MET383	<-->	CG2	VAL206	3.63
O	MET383	<-->	C	VAL206	3.21
O	MET383	<-->	O	VAL206	3.06
O	MET383	<-->	CB	VAL206	3.69
O	MET383	<-->	CG2	VAL206	3.44
O	MET383	<-->	N	ARG207	3.46
O	MET383	<-->	CA	ARG207	3.55
O	MET383	<-->	CG	ARG207	3.66
CB	MET383	<-->	O	VAL206	3.85
CB	MET383	<-->	CG	ASN210	3.84
CB	MET383	<-->	ND2	ASN210	3.76
CG	MET383	<-->	O	GLU209	3.77
CG	MET383	<-->	CA	ASN210	3.74
CG	MET383	<-->	CG	ASN210	3.67
CG	MET383	<-->	OD1	ASN210	3.52
SD	MET383	<-->	CB	VAL206	3.54
SD	MET383	<-->	CG1	VAL206	3.40
SD	MET383	<-->	O	GLU209	2.89
CE	MET383	<-->	O	GLU209	3.76
CE	MET383	<-->	N	ARG211	3.55
CE	MET383	<-->	O	ARG211	3.41
CE	MET383	<-->	CE	LYS213	3.85
SG	CYS384	<-->	NH1	ARG207	3.65
N	GLN385	<-->	NH2	ARG207	3.41
CB	GLN385	<-->	NH2	ARG207	3.75

5A.4.4 Decomposing the free energy contributions to the binding free energy of $XPA_{185-226}$ and XPE on a per-residue basis

The estimation of binding free energy (BFE) and its per-residue energy decomposition (PRED) between two or more systems reflects the binding mechanisms of that particular set of systems, whether it is a protein-protein or a protein-ligand or a DNA-protein

system [322, 452]. BFE and PRED studies have been extensively used in deciphering the crucial amino acid residues that contribute greatly towards the PPI, drug-protein interactions [462, 463].

To gain insights into the contribution of the individual amino acid residues toward the overall PPI of the XPA-XPE complex, PRED values were calculated using the MM-PBSA module of the AMBER 14 software package. MM-PBSA uses a continuum solvent approach to determine the binding free energies of a system. The PRED results for all the interface residues present in our complex are given in **Figures 5A.11A** and **5A.11B,** and **Table 5A.6**. R20, R47, and L57 were the common residues to have exhibited the highest PRED values for all three models. The highest energy contribution for $XPA_{185-225}$ were the residues common to all the models: L191, V193, W194, E198, E202, E205, R207, and F219. R207 from XPA and R20 from XPE had the highest energy decomposition rates for the $XPA_{185-226}$-XPE complex among all the models as seen in **Figure 5A.11**. The energy contribution of R207 is fairly reflective of its effect on XPA-XPE binding, and can also be accounted for its role in the DNA binding where R207Q mutant was reported to have lowered the DNA binding efficacy of XPA [176]. The decomposition energy yielded by F219 may also be related to its association with XPE as well as with the DNA, since the protein truncation of XPA at the F219 position decreased the DNA binding activity of XPA [58].

The PRED values of each residue in any PPI/PPC have been known to contribute to the overall BFE of that particular protein complex. Hence, to address the stability of our models, we also calculated the BFE of the $XPA_{185-226}$-XPE complex for all three models. The individual contribution of $XPA_{185-226}$ and XPE towards the complex formation is given in **Table 5A.7**. The tabulated $\Delta\Delta G_{binding}$ using the PB method for Model 1, Model 2, and Model 3 were found to be -48.3718 kcal mol^{-1}, -49.09 kcal mol^{-1}, and -56.51 kcal mol^{-1}, respectively. This BFE can be very well accorded to their electrostatic and polar solvation energies.

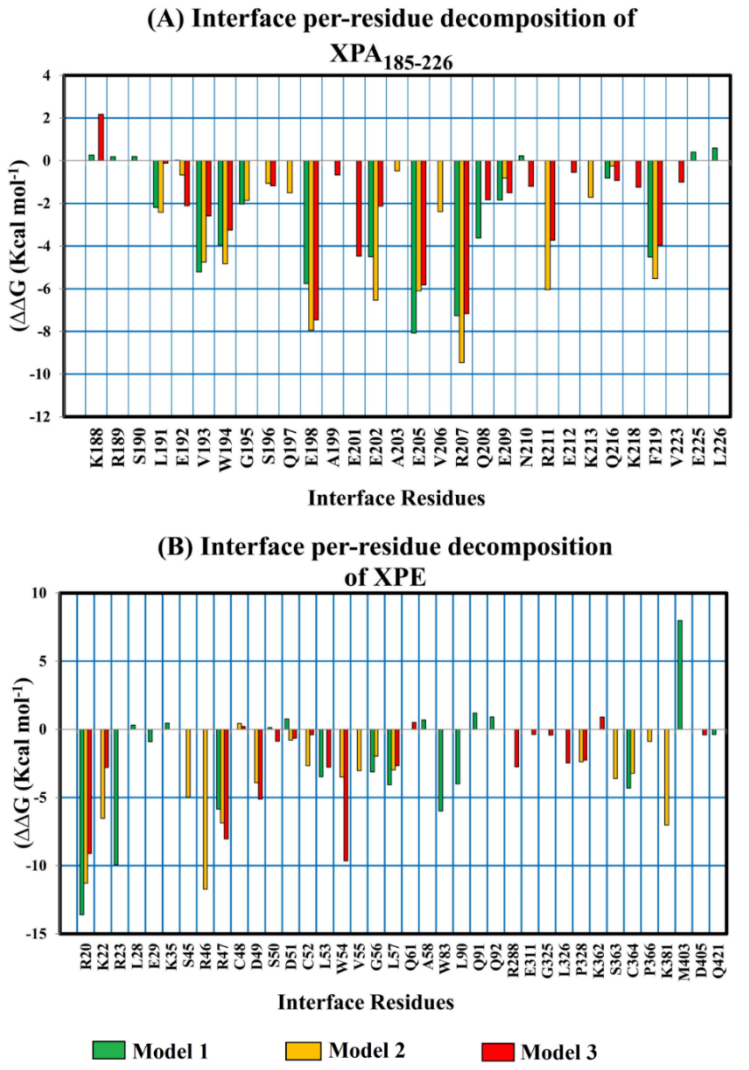

Figure 5A.11. Per-residue energy decomposition (PRED) plots for the interface residues of (A) $XPA_{185-226}$ and (B) XPE.

Table 5A.6. Per-residue energy decomposition (PRED) analysis of the interface residues of $XPA_{185-226}$–XPE complex.

Interface Residues of XPE	Energy contributions (ΔΔG Kcal mol-1)			Interface Residues of $XPA_{185-226}$	Energy contributions (ΔΔG Kcal mol-1)		
	Model 1	Model 2	Model 3		Model 1	Model 2	Model 3
R20	-13.6	-11.3	-9.12	K188	0.26	-	2.17
K22	-	-6.53	-2.8	R189	0.19	-	-
R23	-9.94	-	-	S190	0.2	-	-
L28	0.3	-	-	L191	-2.19	-2.42	-0.12

E29	-0.91	-	-	E192	0.02	-0.67	-2.11
K35	0.45	-	-	V193	-5.21	-4.76	-2.59
S45	-	-4.97	-	W194	-3.96	-4.83	-3.25
R46	-	-11.73	-	G195	-2.03	-1.86	-
R47	-5.86	-6.87	-8.05	S196	-	-1.06	-1.17
C48	-	0.43	0.21	Q197	-	-1.51	-
D49	-	-3.92	-5.12	E198	-5.76	-7.95	-7.46
S50	0.13	-	-0.88	A199	-	-	-0.67
D51	0.77	-0.79	-0.66	E201	-	-	-4.47
C52	-	-2.67	-0.4	E202	-4.5	-6.54	-2.13
L53	-3.49	-	-2.78	A203	-	-0.48	-
W54	-	-3.51	-9.65	E205	-8.08	-6.1	-5.82
V55	-	-3.03	-	V206	-	-2.38	-
G56	-3.12	-1.96	-	R207	-7.27	-9.47	-7.17
L57	-4.08	-2.98	-2.67	Q208	-3.62	-	-1.83
Q61	-	-	0.5	E209	-1.84	-0.81	-1.51
A58	0.69	-	-	N210	0.23	-	-1.2
W83	-5.99	-	-	R211	-	-6.04	-3.73
L90	-4	-	-	E212	-	-	-0.54
Q91	1.2	-	-	K213	-	-1.71	-
Q92	0.9	-	-	Q216	-0.81	-0.25	-0.92
R288	-	-	-2.75	K218	-	-	-1.24
E311	-	-	-0.38	F219	-4.51	-5.52	-3.97
G325	-	-	-0.43	V223	-	-	-1.01
L326	-	-	-2.47	E225	0.4	-	-
P328	-	-2.38	-2.27	L226	0.6	-	-
K362	-	-	0.89				
S363	-	-3.61	-				
C364	-4.31	-3.22	-				
P366	-	-0.89	-				
K381	-	-7.02	-				
M403	7.98	-	-				
D405	-	-	-0.41				
Q421	-0.38	-	-				

Table 5A.7. Binding free energy (BFE) analysis of $XPA_{185-226}$–XPE complex.

Method	Models	Components	$XPA_{185-226}$-XPE protein complex	$XPA_{185-226}$ (kcal mol^{-1})	XPE (kcal mol^{-1})	ΔΔGbind (kcal mol^{-1})

			(kcal mol^{-1})			
MM-PBSA	Model 1	ΔE_{vdW}	-3316.62	-237.46	-3028.03	-50.55
		ΔE_{ele}	-29076.55	-2770.88	-25940.01	-365.65
		ΔE_{MM}	-32393.17	-3008.35	-28968.61	-416.21
		ΔG_{PB}	-6524.36	-1475.01	-5426.65	377.30
		ΔG_{Surf}	173.33	32.01	150.78	-9.46
		ΔG_{solv}	-6351.02	-1442.99	-5275.86	367.84
		PB_{TOT}	-38744.20	-4451.35	-34244.48	**-48.37**
	Model 2	ΔE_{vdW}	-3381.81	-237.17	-3068.72	-75.92
		ΔE_{ele}	-28929.58	-2908.32	-25414.01	-607.25
		ΔE_{MM}	-32311.40	-3145.49	-28482.73	-683.17
		ΔG_{PB}	-6299.89	-1345.77	-5535.91	581.79
		ΔG_{Surf}	1483.35	88.80	1342.26	52.28
		ΔG_{solv}	-4816.54	-1256.97	-4193.65	634.07
		PB_{TOT}	-37127.94	-4402.46	-32676.38	**-49.09**
	Model 3	ΔE_{vdW}	-3325.42	-240.45	-3009.43	-75.53
		ΔE_{ele}	-28844.99	-2865.41	-25348.26	-631.31
		ΔE_{MM}	-32170.42	-3105.87	-28357.70	-706.85
		ΔG_{PB}	-6511.36	-1395.36	-5780.06	664.06
		ΔG_{Surf}	168.38	31.15	150.94	-13.72
		ΔG_{solv}	-6342.97	-1364.20	-5629.11	650.33
		PB_{TOT}	-38513.40	-4470.07	-33986.81	**-56.51**

ΔE_{ele} = electrostatic energy as calculated by the MM force field; ΔE_{vdW} = van der Waals contribution from MM; ΔE_{MM} = total gas-phase energy (sum of ELE, VDW, and INT); ΔG_{PB} = the electrostatic contribution to the polar solvation free energy calculated by PB; ΔG_{Surf} = non-polar contribution to the solvation free energy calculated by an empirical model; ΔG_{sol} = sum of non-polar and polar contributions to solvation; PB_{TOT} = final estimated binding free energy in kcal mol^{-1} calculated from the terms above.

PPIs are vital for all the biological functions, whether it is signaling cascades, molecular switching, hormone-receptor reactions [395, 455, 456], or in this case, successful removal of the DNA lesion, triggered by the ultra-violet radiations and anti-cancer agents. NER mechanism in a higher organism is rigorous and effective due to the numerous transient and permanent PPIs between various proteins, systematized largely by the XPA protein [48, 164]. XPA, in particular, is recruited to the damage site upon its C-terminal interaction with the p8 and p52 subunit of the TFIIH complex [48, 165, 356]. This happens simultaneously with the recruitment of RPA70 protects the undamaged strand and stabilizes the NER bubble, wherein it also interacts with the DBD portion of XPA [125, 127, 142]. The N-terminal interaction of XPA (aa96-114) with the ERCC1 (aa92-119) is responsible for the recruitment of ERCC1/XPF complex, a structure-specific endonuclease to the damage site which incises the lesion at 5$^{/}$ end [144-146].

Apart from these interactions, the accounts of many new interactions of XPA with other proteins have surfaced over the years. Reports of XPA's new binding partners, proliferating cell nuclear antigen (PCNA) [83] that is usually seen only during the ligation stage of NER, and two novel XPA binding proteins (XAB), named XAB1 [155], and XAB2 [157], whose functions are not much known, suggests that XPA may have a much larger role in NER than just a mere damage verifier, and a scaffold protein.

The biochemical data provided by Wakasugi's team [151] on the XPA-DDB2/XPE interaction answered some of the questions surrounding XPA's scaffold nature. They were able to successfully demonstrate that XPA interacted with DDB1/DDB2 complex only through its DDB2 or XPE subunit both *in vivo* and *in vitro* conditions, and not the vice-versa. They further went on to identify the exact binding region on XPA (aa185-226), which helped in establishing the PPI with XPE to conduct the NER process smoothly. Our *in-silico* investigation of the probable PPI of human $XPA_{185-225}$ with XPE using the information provided by Wakasugi et al. research findings showed that this interaction between XPA-XPE is stable. The results that we obtained from the PDBsum server and SASA analysis suggested that this complex formation is of a strong transient PPI type, which is fairly justifiable considering that both proteins, especially XPA are involved in forming various PPCs and PPIs with the repair proteins during the NER process, which is initiated within a fraction of seconds upon the damage detection and has to be quickly tended to. The protein-protein docking study and the analyses of the PPI profiles generated by the PDBsum server for the XPA-XPE complex revealed that among $XPA_{185-226}$ and DBD of XPA (XPA_{98-239}), the former length of the protein is responsible for the binding and interacting with XPE, while the XPA's residues before and after 185-226 in case of XPA_{98-239}-XPE complex were very much farther from XPE to create the intermolecular PPI. The PRED results also showed that XPA's region between 185-226 contains residues with high signatures of binding free energies, especially R207, which contributed greatly to their PPI and PPC formation. The observations made by Wakasugi and his team were also of a similar kind. They saw that the truncations up to 184 amino acids from the N-terminal of XPA showed no effect on the XPA-XPE binding, but even the slightest truncation of a further 41 amino acids ($XPA_{185-226}$) greatly diminished the interaction. Similar binding

properties were obtained with purified DDB heterodimer, consistent with the notion that XPA binds to DDB heterodimer by interacting with DDB2. These results indicate that the amino-acid domain between 185 and 225 is required for the XPA-XPE/DDB2 interaction.

The significance of $XPA_{185-226}$ is that the region that codes for this particular protein segment lie from the terminal part of exon 4 through the whole part of exon 5 till the beginning section of exon 6 in the XPA gene. These exons 4, 5, and 6 are responsible for coding the XPA's residues aa130-273, which is also the partial coding region for the DBD (aa130-239) of XPA. Therefore, any deletion within this section (exon 4-6) or the protein truncation at any position can not only inhibit the PPC formation between XPA and XPE but can also lead to the collapsing of NER [176, 371]. This binding region of XPA on XPE also coincides with the interaction sites for TFIIH [48, 165, 356], and with two NER regulating proteins- ataxia telangiectasia and Rad3-related protein (ATR), which localizes XPA to the damage site by phosphorylating Ser196 in XPA[153], and poly [ADP-ribose] polymerase 1 (PARP-1), which aids XPA in forming the pre-incision complex (PIC) in NER [464]. Our PRED analyses study of the XPA's residues also shed light on the fact that the residues L191, R207, and F219 of XPA which showed higher PRED values for PPI with XPE, have also been mentioned in the literature before for having a role in DNA interaction, PPI with other proteins and in causing XP-A phenotypes upon mutations. The higher energy decomposition values of these residues indicate their importance for XPA's functioning in NER.

Our present study has been able to shine a light on one of the binding partners of XPA in terms of its scaffolding aspects if not all. We hope that with the increasing progress in the study of XPA, the day may come when all the unsolved questions on XPA's role with different NER members (proteins and DNA) and its effect on NER outcomes, and even the newly reported binding partners of XPA may finally be solved. The combination of structural, mechanistic, and dynamics data will prove to be the key element in meeting this goal.

5A.5 Conclusion

In summary, our study has helped us determine the probable binding site of $XPA_{185-226}$

on XPE, and further characterize the interface residues that are responsible for their PPI. The structural and conformational dynamics study for all three models of the $XPA_{185-226}$-XPE complex using all-atom MD simulations suggests that this interaction is stable and also of a strong transient type. The PRED analysis using the MM-PBSA algorithm showed that the binding affinity between two proteins is indeed high, and their intermolecular PPI can be credited to the residues R20, R47, and L57 of XPE and the residues L191, V193, W194, E198, E202, E205, R207, and F219 of XPA, respectively.

CHAPTER 5B

To study the effect of XPA's R207G mutation on its binding affinity with XPE in a dynamic system

CHAPTER 5B

To study the effect of XPA's R207G mutation on its binding affinity with XPE in a dynamic system

5B.1 Abstract

Nucleotide excision repair (NER) is a pivotal DNA repair system in higher organisms, that catalyzes the removal of bulky DNA lesions caused by UV radiation and various other mutagens. The mutations in the NER proteins can cause a major setback to the damage repair process, leading to an autosomal recessive condition like *Xeroderma pigmentosum* (*XP*). The mutations in XPA (*XP* complementation group A) protein, in particular, has the most debilitating effect causing classical *XP* with the occurrences of neurological disorders, and cancers. R207G is one such mutation in XPA that has been highlighted to affect its interaction with XPE (*Xeroderma pigmentosum* complementation group E), damage recognition protein of global genome NER (GG-NER) in both *in Vivo* and *in vitro* conditions. Here, we have studied molecular dynamics and compared the association of Wild type (WT) and R207G mutant XPA with XPE using the AMBERff99 force field. We observed that WT XPA had formed a relatively stable complex with XPE than R207G mutant XPA. The binding free energy (BFE) analyses exhibited that XPE has a relatively higher binding affinity with WT XPA (-58.9 kcal mol^{-1}) as compared to the mutant R207G XPA (+16.44 kcal mol^{-1}). Our findings in this study substantiate the effect of R207G point mutation in XPA on its interaction with XPE at the molecular level.

5B.2 Introduction

DNA being an integral part of any living system is highly susceptible to attacks, which can create havoc in the regular functioning of a cell, and the living being as a whole. Hence, there are certain DNA repair mechanisms: nucleotide excision repair (NER), base excision repair (BER), and mismatch repair (MMR) through which the damages endured by the DNA are repaired. NER is the most dominant DNA repair pathway that repairs helix-distorting damages, induced upon the DNA by ultraviolet radiations, various chemical and environmental mutagens, and anti-tumor therapeutic agents. Any

error or defects in the functioning of NER leads to an autosomal recessive disease known as *Xeroderma pigmentosum* (XP), wherein the individual affected exhibits extreme sensitivity toward sunlight, has neurological defects and is at high risk of developing skin cancers [5, 9, 429, 465, 466].

NER broadly functions through two sub-pathways, global genome NER (GG-NER) and transcription-coupled NER (TC-NER). In higher mammals, especially in humans, this whole process of damage repair is executed by ~30 different proteins, among which the XP complementation group A (XPA), replication protein A70 (RPA), XP complementation group C- human Rad23 homolog (XPC-HHR23B), transcription factor II H (TFIIH) complex, excision-repair cross-complementing group 1 endonuclease (XP complementation group F) (ERCC1/XPF), and XP complementation group G (XPG) are considered to be the NER core proteins. In this multi-step process of NER, right from damage recognition to the synthesis of a new strand, XPA plays an indispensable role as a scaffold protein, where it directs the NER core proteins, and other NER members to the damaged sites so that the damage is properly removed [5, 48, 54, 57, 359].

XPA encoded by the XPA gene (9q22.33) spans 273 residues long (M.W. 40kDa), is constituted of disordered N- and C-terminals, and a central globular DNA binding domain (DBD) ranging from the residues 98-239. Most of the DNA and protein interacting sites of XPA are localized to its DBD, while few interaction sites lie in its disordered regions. XPA has an affinity to recognize an array of DNA distortions, such as 6-4 photoproducts (6-4PP), helix-distorting platinum (Pt) crosslinks, and cyclobutane pyrimidine dimers (CPD), but when XPA is mutated even by a single residue, or even loses partial part of the protein, it loses its ability to bind to these DNA lesions and fails to tether NER core proteins, leading to NER failure. These events, in turn, have a rippling effect in the form of severe XP-A phenotypes, and often culminate in cancerous stages [58, 114-116, 126, 176, 405]. The importance of XPA can be stressed by the fact that since XPA is the common denominator to both GG-NER and TC-NER, the XP-A phenotypes arising due to the deficit/mutant XPA protein are more lethal than the phenotypes arising due to the deficits in other NER proteins [5, 9, 54, 57, 72, 429, 465, 466]. Almost all the XP-A phenotypes have been known to exhibit severe forms of sun

sensitivity, advanced aging, neurological impairment with loss of motor and sensory functions, and high rates of skin cancer [361, 183, 170, 405, 429].

Earlier mutational studies on XPA have shown that the point mutations inflicted upon the protein lead to a decrease in its protein functions as a whole, which can affect the protein's interaction with DNA or other NER proteins. Saijo et al. [127] observed a reduction in the binding affinity of XPA to RPA70 protein upon the mutation of the lysine residues to alanine (K141A, K142A, K167A, and K179A), affecting the damage recognition step of the NER mechanism. The study done by Sugitani et al. [176] showed that the DNA binding activity of XPA was reduced by the point mutations of the residues located at the C-terminal end of XPA (K221E, K222E, W175A, and R228E) and resulting in disrupting the whole NER process. Many such observations have been reported where XPA mutants have led to repercussions in causing severe XP phenotypes, and NER defects. One research group identified 56 somatic mutations of XPA alone in 121 cancer patients [183], while another research has reported 7 single nucleotide polymorphisms (SNPs) among 191 XP SNPs [170]. Amr et al. [361] observed that 31G>T nonsense mutation in the XPA gene led to E111X mutation in the protein, which again saw another severe form of XP-A phenotype. Sethi et al. [160] traced the inheritance of a mild form of XP-A phenotype to, Pakistan, Afghanistan, and some parts of North India. Every so often, a novel mutant of XPA is reported, which shows symptoms of NER defects and XP phenotypes much more severe than existing XPA mutants, hence it becomes important to understand these mutations.

Among many such mutations of XPA protein, one such mutation is R207G which was reported by Wakasugi and team [151]. This mutation was observed to have inhibited the binding and protein-protein interaction (PPI) between XPA and XPE. This PPI between both proteins is generally accorded by 41 residues (aa185-226) of XPA that was identified by Wakasugi et al. [151] by conducting a series of XPA protein truncations of various lengths at the N-terminal end. XPE, also known as the damaged DNA binding protein 2 (DDB2) is a part of the heterodimer unit that comprises DDB1/DDB2 and is responsible for the recognition and verification of DNA lesions in GG-NER [148, 151, 467, 468]. This particular mutation has also been observed to affect the recruitment of XPA to the NER bubble, which in turn leads to incomplete damage

excision. The importance of the residue R207 has also been stressed by findings of Camenisch et al. [149] who reported R207 to be involved in DNA binding, and upon mutation to glutamine or any other residues (R207Q/X) caused skin and neurological complications along with the occurrence of cancer in XP-A patients. The significance of R207's role in modulating the binding affinity of XPA towards other proteins/DNA can also be emphasized by its high peaks that were obtained on conducting the chemical shift perturbation (CSP) assay in its WT form, and its failure to bind DNA upon mutant R207Q form [176]. A recent clinical study also reported a similar observation upon R207X mutation in a seventeen-year-old Brazilian girl, along with some additional complications like variations in the brain white matter, sensorineural hearing loss, steppage gait, and diffuse atrophy, sensory and motor neuropathies, etc [469]. These life-threatening aftermath impacts of R207G/X mutation on XPA emphasize the necessity that needs to be studied extensively.

Recently, we have reported the key residues responsible for the PPI between $XPA_{185-226}$ and XPE using the molecular dynamics (MD) approach. In addition, the per-residue energy decomposition (PRED) analysis using the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method revealed protein complex of $XPA_{185-226}$-XPE has a high binding affinity and that R207 exhibited the highest energy decomposition markers [399]. Even though this result helped us understand the importance of R207 in XPA-XPE binding; however, some questions remained unanswered: (i) causes for inefficient binding/loss of interaction between XPA and XPE? (ii) Effect of point mutation on the binding affinity of XPA with XPE in a dynamic scenario?

Hence, in this study, we have focussed on the effect of R207G point mutation in XPA on its interaction with XPE. We have performed a molecular dynamics simulation study on the WT XPA-XPE and R207G mutant XPA-XPE complex using the Assisted Model Building with Energy Refinement (AMBER) 14 software package [470]. In this study, we have considered the protein sequence of XPA from residue index 185 to 273 as it was reported in the previous study done by Wakasugi [151] that the N-terminal residues (aa1-184) were not involved in XPA-XPE interaction. We investigated the interactions and binding affinity of XPA (WT and R207G mutant) with XPE.

5B.3 Systems and methodology

5B.3.1 Protein systems

The Universal Protein Resource (UniProt) database [378] was used to retrieve the protein sequence of $XPA_{185-273}$ (ID: P23025). The sequence was then submitted to the Iterative Threading ASSEmbly Refinement (I-TASSER) server [448] for the tertiary structure prediction. I-TASSER server is known for its robust technique of determining the 3-D structure of proteins by making use of both homology modeling and protein threading. The server provided us with five models, wherein the best-modeled structure was chosen based on their confidence score or C-score, template modeling (TM) score [449], and a root mean square deviation (RMSD) score. Further validation of the modeled structure was carried out using the protein structure assessment (ProSA) server [382], MolProbity server [450], and RAMPAGE server [383]. In this study, we have used this validated model structure as the 'WT' structure. From this WT model structure, the R207G XPA mutant was constructed, where residue R207 was mutated to glycine using the Rotamer tool of the University of California, San Francisco (UCSF) Chimera software v.1.12 [314]. The XPE structure was obtained using the procedure that we adopted in our earlier work [472].

Next, we docked $XPA_{185-273}$ (WT and R207G mutant) with XPE using the ClusPro web server [124, 447, 471]. The docking procedure and the working mechanism of this docking server, as well as the basis on which the docked complexes are generated, have been described in our earlier study [472]. The number of populated clusters, cluster center, and the lowest energy weighted scores were taken into consideration for the selection of the best-docked models for WT-XPE and R207G-XPE complexes. For convenience, the protein complex of WT $XPA_{185-273}$ and XPE was codenamed 'WT-XPE', while the protein complex of Mut $XPA_{185-273}$ and XPE was named 'R207G-XPE'. $XPA_{185-273}$ with no mutation was referred to as 'WT', and mutated $XPA_{185-273}$ was named as 'R207G'. Similarly, the XPE of the WT-XPE complex was coded as 'XPE_{WC}', while the code 'XPE_{MC}' was given to refer to the XPE of the R207G-XPE complex.

5B.3.2 Setup for MD simulations

WT-XPE and R207G-XPE complex systems were prepared for the simulation using the AMBER 99 force field [282] in the Leap module of the AMBER 14 software package. Both the systems were solvated in an explicit TIP3P (transferable intermolecular potential with 3 points) water model [273] with a buffer distance of 10 Å in a cubic periodic box. Both the systems were neutralized by adding counter ions and replacing the overlapping solvent molecules. The parameterizations used for our systems in Xleap have been shown in **Table 5B.1**. The hydrated, uncharged systems were then next subjected to two-step minimization techniques: steepest descent (SD) and conjugate gradient (CG). The SD method minimized the systems by keeping restraints over the solute, whereas the second step of minimization (CG method) was devoid of such restraints.

Table 5B.1. The system parameterization using Xleap.

Parameters	WT-XPE	Mut-XPE
AMBER force field	ff99SB	ff99SB
Initial charge	-15	-16
Counter-ions added	15 Cl^- ions	16 Cl^- ions
Final charge	0	0
Water residues added	24280	22544
Total mass	461928.218 amu	493103.844 amu
Density	0.840 g/cc	0.843 g/cc
Volume	912761.163 A^3	942816.121 A^3

The MD study was carried out using a standard procedure, which consisted of heating dynamics followed by density, equilibration, and production dynamics. We used minimized systems as our starting structure for subsequent MD steps. The systems were gradually heated from 0-300 K in constant volume (NVT) conditions, after which the density procedure was carried out. The equilibration of our protein systems was conducted in NPT conditions (300 K and 1 atm pressure) for 1 ns. To ensure the successful equilibration of the systems, the temperature, energy, and pressure graphs were plotted and analyzed. Next, we performed a 40 ns MD production run for the equilibrated structures of both the systems using the Particle Mesh Ewald (PME)

algorithm [388, 435] with the time step of 2 fs. We used a cutoff of 8 Å to treat the nonbonding interactions (short-range electrostatic and van der Waals interactions) during the simulation, while the long-range electrostatic interactions were treated with the PME method. All the bonds present in the systems were constrained with the SHAKE algorithm [292]. The pressure and temperature (0.5 ps of heat bath and 0.2 ps of pressure relaxation) were held constant by the Berendsen weak coupling algorithm [294] throughout the simulation process. For each system, the trajectory snapshots were recorded every 10 ps for further analysis.

The MD trajectories for both the protein complexes were analyzed using PTRAJ (short for Process TRAJectory) and CPPTRAJ (a rewrite of PTRAJ in C++) module [316] of AmberTools 14. To assess the convergence of our systems, we studied the RMSDs for WT-XPE and R207G-XPE complexes, wherein the starting structure of MD was used as the reference. We also calculated the root mean square fluctuations (RMSFs) to analyze the flexibility of both protein complexes. In addition, the RMSDs of the amino acids R207 and G207 were also studied to understand their role in maintaining the stability of XPA protein.

5B.3.3 Binding free energy (BFE) analyses for WT-XPE and R207G-XPE systems

Using the MMPBSA.py script under the MM-PBSA module in the AMBER 14 suite [321, 372-375, 472], we have computed the binding free energy (BFE) for the WT/R207G-XPE complexes from their trajectory files generated by 40 ns production run. The free energy analyses are considered important in establishing the binding affinity in the protein-protein, ligand-protein, and DNA-ligand interaction studies. Hence, to gather the differences in the binding affinities of our systems, the MM-PBSA analysis was done for our two systems by considering the following components (i) WT/R207G (ligand), (ii) XPE (receptor), and (iii) WT/R207G-XPE (complex).

$$\Delta G_{bind} = \Delta G_{complex} - [\Delta G_{receptor} + \Delta G_{ligand}] \quad (1)$$

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (2)$$

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{vdW} + \Delta E_{ele} \quad (3)$$

$$\Delta G_{solv} = \Delta G_{PB} + \Delta G_{surf} \quad (4)$$

$$\Delta G_{bind} = \Delta E_{int} + \Delta E_{vdW} + \Delta E_{ele} + \Delta G_{PB} + \Delta G_{surf} - T\Delta S \quad (5)$$

$$\Delta G_{surf} = \gamma \times SASA + \beta \quad (6)$$

The BFE of $WT/R207G_{ligand} - XPE_{receptor} = WT/R207G - XPE_{complex}$ was calculated using the equations (Eqn. 1-5] derived from the second law of thermodynamics, where they were studied in both gas (vacuum) and aqueous environments (**Figure 5B.1**). Here, ΔG_{bind} is the total binding free energy, while ΔG_{ligand}, $\Delta G_{complex}$, and $\Delta G_{receptor}$ correspond to the relative free energies of the WT/R207G, XPE, and WT/R207G-XPE, respectively. Further, the molecular mechanics energies (ΔE_{MM}) are the total of the internal energy (ΔE_{int}), van der Waals forces (ΔE_{vdW}), and electrostatic energy (ΔE_{ele}). The solvation-free energy (ΔG_{solv}), on the other hand, is seen for Eqn. [4] is the sum of the non-polar contribution (ΔG_{surf}) and the polar contribution calculated using the PB solver algorithm [392] of the AMBER 14 pbsa module (ΔG_{PB}). ΔG_{surf} was calculated based on the solvent-accessible surface area using the LCPO (linear combinations of pairwise overlaps) method [437], as shown in Eqn. [6], where γ = 0.00542 kcal mol^{-1} Å-2 and β = 0.92 kcal mol^{-1}. Our calculations were done by taking the dielectric constant of the interior (solvent) and the external (solute) as 80 and 1, and the probe radius and the grid space were set accordingly to 1.4 Å and 0.5 Å.

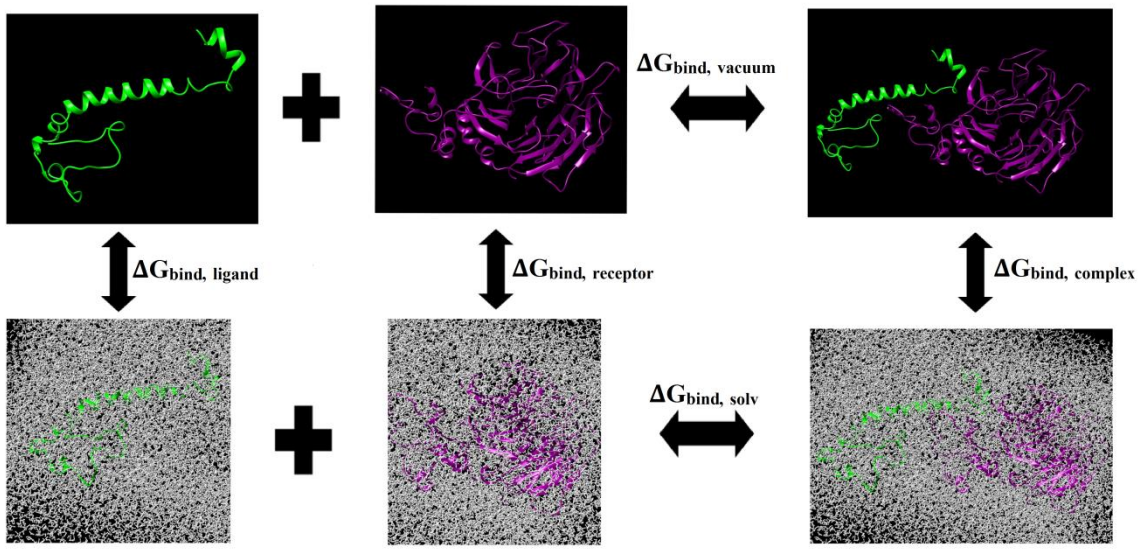

Figure 5B.1. Schematic representation of the thermodynamic cycle used for the calculation of binding free energies (BFE).

5B.4 Results and discussion

5B.4.1 Modeling and the validation of $XPA_{185–273}$, and the selection of the best-docked complex for the $XPA_{WT/R207G}$-XPE complex

We obtained a total of five models (**Figure 5B.2**) for the structure of $XPA_{185–273}$ from the I-TASSER server that uses the homology and threading algorithm. Among these five models, the top-ranked structure (model 1 of **Figure 5B.2**) was chosen as the candidate model for this study. Model 1 had a C-score of -1.84, a TM score of 0.49 ±0.15, and an RMSD of 7.6±4.3 Å. The secondary structure of this model was found to contain helices in the major portion and had coils in its C-terminal end, which was expected due to its disordered nature. The fitness assessment of this model using the RAMPAGE server showed 88.5% of the residues in the favored region, 9.2% residues in the allowed region, and 2.3% residues in the outlier region of the plot (**Figure 5B.3A**). Next, we validated the model's robustness using the ProSA server that assigns a Z-score to the modeled structure. If this score lies within the Z-score range of all the protein structures (NMR and X-ray determined) available in the PDB, then it is considered to be near to the native structure. Our model as seen in **Figure 5B.3B** was assigned the Z-score of -1.59, which designated our model to be near to the native conformation. The stereo-chemical properties analyzed using the MolProbity server showed all atoms clash score of 0.67 (99^{th} percentile), and the MolProbity score of 2.16. The complete list of the features analyzed by the MolProbity server for the modeled $XPA_{185-273}$ has been provided in **Table 5B.2**.

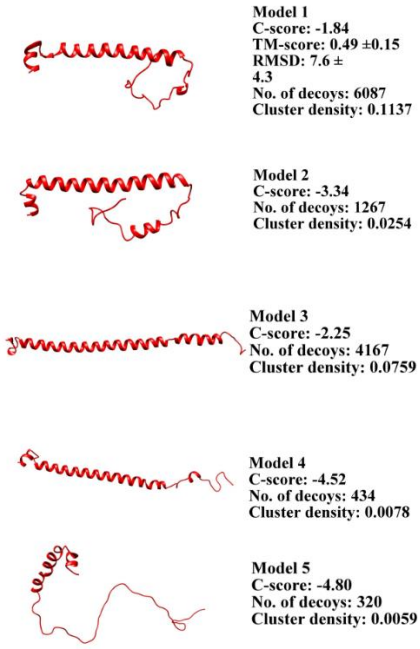

Figure 5B.2. Top five models generated for XPA$_{185-273}$ by the I-TASSER server.

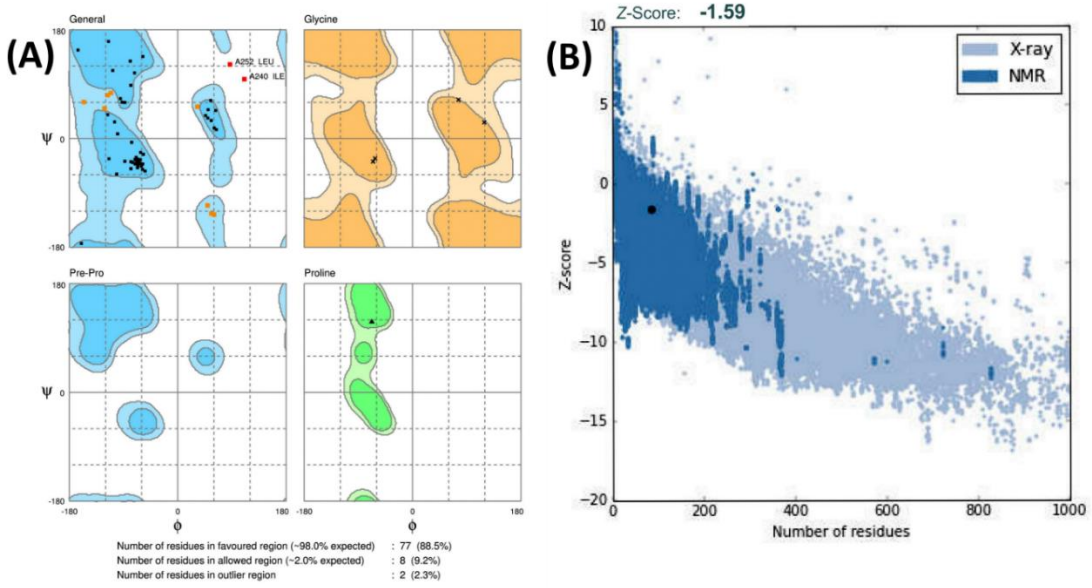

Figure 5B.3. Structure validation of the modeled XPA$_{185-273}$: (A) Ramachandran plot as obtained from the Rampage server, and (B) Z-score plot energy plot as determined by the ProSA-web server.

Table 5B.2. MolProbity summary statistics for the modeled XPA$_{185-273}$ protein.

All-Atom contacts	Clashscore, all atoms:	0.67		99th percentile* (N=1784, all resolutions)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein geometry	Poor rotamers	8	10.00 %	Goal: <0.3%
	Favored rotamers	62	77.50 %	Goal: >98%
	Ramachandran outliers	2	2.30%	Goal: <0.05%
	Ramachandran favored	77	88.50 %	Goal: >98%
	MolProbity score^	2.16		67th percentile* (N=27675, 0Å - 99Å)
	Cβ deviations >0.25Å	4	4.65%	Goal: 0
	Bad bonds:	1 / 737	0.14%	Goal: 0%
	Bad angles:	5 / 984	0.51%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 1	0.00%	Expected: ≤1 per chain, or ≤5%
	Twisted Peptides:	3 / 88	3.41%	Goal: 0

In the two-column results, the left column gives the raw count; the right column gives the percentage.
*100th percentile is the best among structures of comparable resolution; the 0th percentile is the worst.

Since, we know that one of the crucial steps in GG-NER, which is the PPI between XPA and XPE is known to be hindered by the point mutation in XPA (Arg207->Gly207), we sought to draw the comparative analyses between WT XPA and R207G mutant in complex with XPE. R207G mutant was constructed using UCSF Chimera from the validated structure of XPA$_{185-273}$ while retaining the original structure as the WT protein (**Figure 5B.4**). Both the WT and R207G XPA$_{185-273}$ were docked with XPE using the ClusPro web server. The top ten models for the WT-XPE and R207G-XPE complex obtained from the ClusPro server were ranked based on their cluster members, cluster center score, and lowest energy (**Figures 5B.5** and **5B.6**). Our chosen docked

model of the WT-XPE complex (model 1 from **Figure 5B.5**) had 243 members in the cluster, with the lowest energy value of -2143.8 kcal mol^{-1}, and the center score of -1941.3 kcal mol^{-1}. We chose model 1 (**Figure 5B.6**) as our docked model for the R207G-XPE complex, which had 236 members with the center score of -2011.3 kcal mol^{-1}, and the lowest energy weighted score of -2329.3 kcal mol^{-1}. These docked models show that WT and R207G XPA were bound to the cleft region (between N and C-terminal) of XPE in both cases. In our previous study, we had also obtained $XPA_{185-226}$ to be bound to XPE in a similar fashion [472].

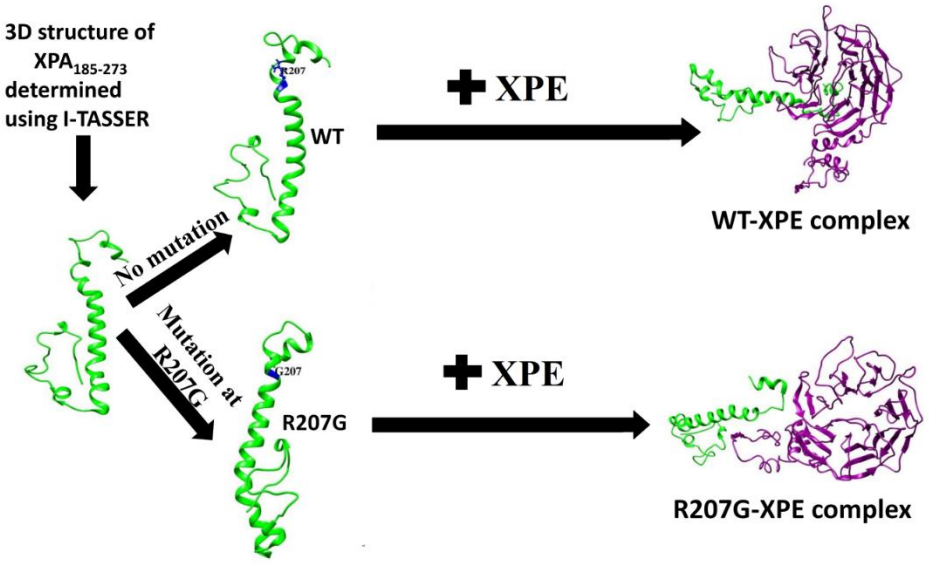

Figure 5B.4. Preparation of the WT and R207G XPA structures, and their molecular docking with XPE.

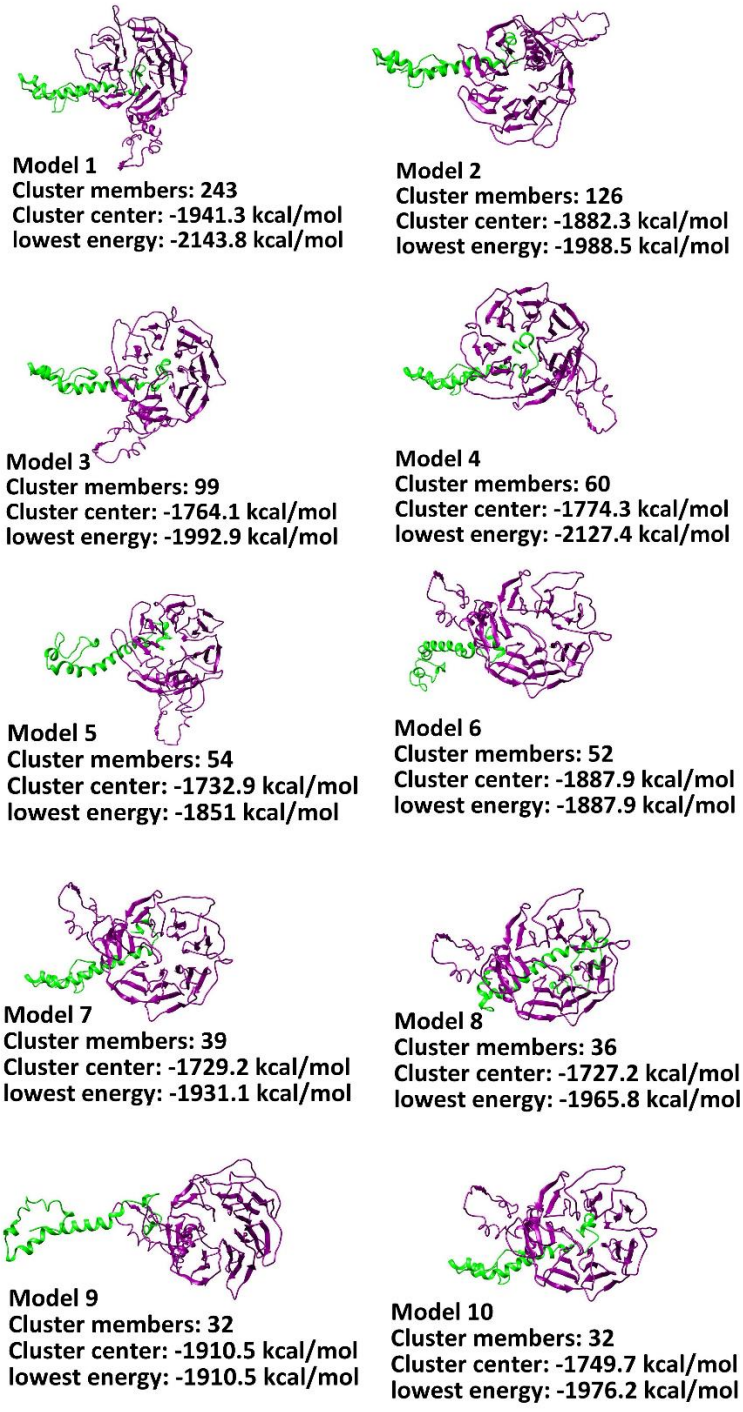

Figure 5B.5. Top ten representative docked models for the WT-XPE complex generated by the ClusPro server along with their rankings based on the highly populated cluster numbers and their lowest energy weighted scores.

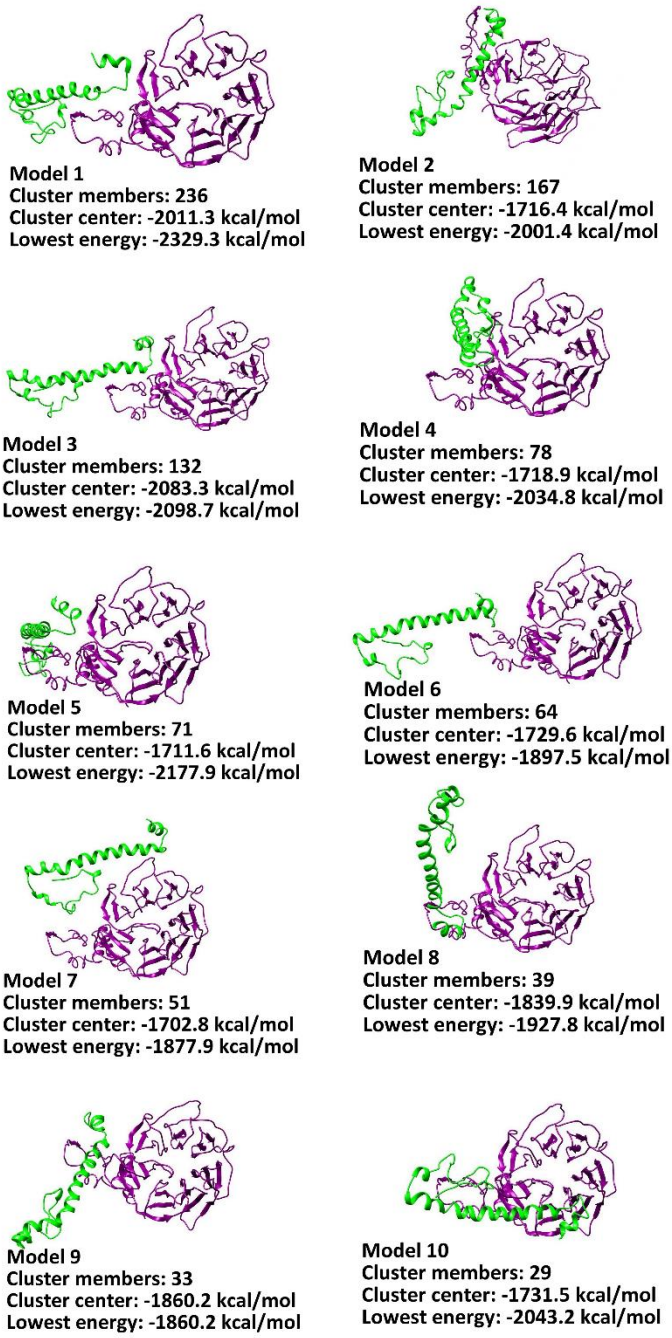

Figure 5B.6. Top ten representative docked models for the R207G-XPE complex generated by the ClusPro server along with their rankings based on the highly populated cluster numbers and their lowest energy weighted scores.

5B.4.2 Analysis of trajectories of WT-XPE and R207G-XPE complexes

MD simulations yield in-depth knowledge about the dynamic behavior of a particular system that is being studied and helps us to understand the changes in their stability and flexibility over the period. The analyses of temperature, energy, and pressure after the equilibration step indicated that our systems being studied here had successfully reached equilibrium, and the plots for which have been shown in **Figures 5B.7A, 5B.7B**, and **5B.7C** for **WT-XPE complex**, and **Figures 5B.7D, 5B.7E** and **5B.7F** for R207G-XPE complex, respectively.

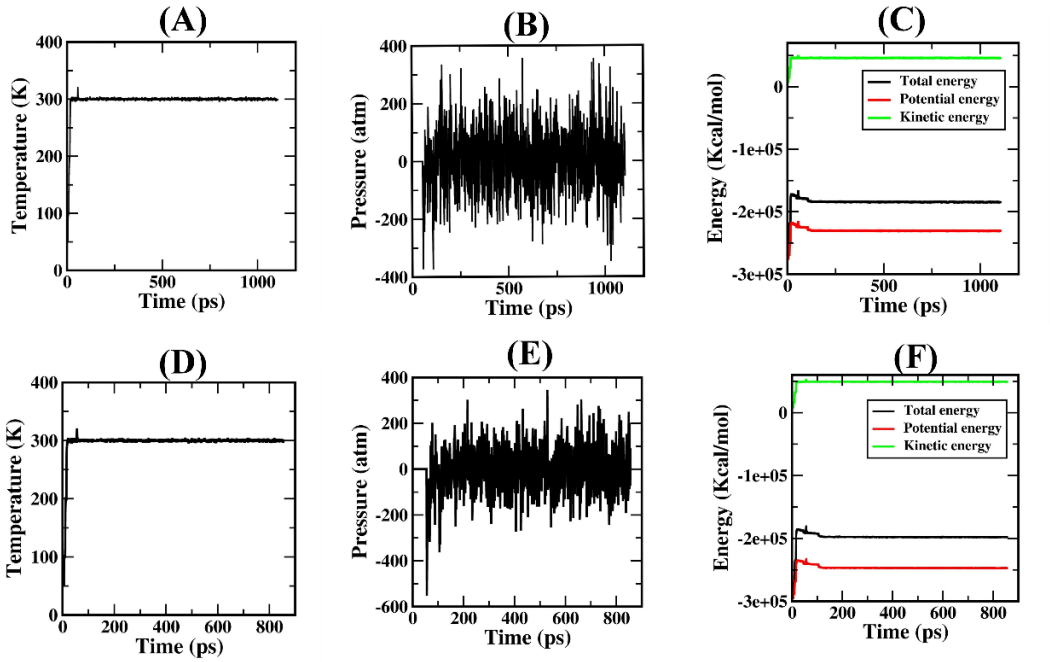

Figure 5B.7. (A) Temperature plot, (B) Pressure, and (B) energy plot for the WT-XPE complex. (D) Temperature plot, (E) Pressure, and (F) energy plot for the R207G-XPE complex.

5B.4.3 Overall stability of the WT-XPE and R207G-XPE systems

In a typical MD simulation, the stability of the system is generally studied by tracking the RMSD of that protein/biological molecule as a function of time. For the two systems studied here, their RMSD values as a function of time have been shown in **Figures 5B.8A**. **Figure 5B.8B** shows a comparative RMSD plot for the WT-XPE complex and R207G-XPE complex, where the WT-XPE system was observed to have converged at

12 ns with an average distance of 8 Å, while the 207G-XPE system showed no signs of convergence and remained unstable throughout the simulation period. We also studied the individual monomeric units of both the heterodimeric complexes (**Figure 5B.8C**), where we observed that the RMSD values converged at about 8 Å and 5 Å after 12 ns for WT and XPE_{WC} of the WT-XPE complex, respectively. On the other hand, the monomeric units of the R207G-XPE complex, R207G, and XPE_{MC} were also seen to be unstable and did not show convergence; this was also reflected in their mutant heterodimer complex form from the RMSD plot. The RMSD analyses of amino acids: WT Arg207 and mutant Gly207 showed that Arg207 was stable throughout the simulation time, while upon its mutation to Gly207, the residue gains instability as seen in the **Figure 5B.8C**, where the RMSD rose from 0.25 Å to 2 Å, after which the mutated residue stabilized. These also tell us that the RMSD values of two systems vary from lower to higher in the following manner, WT-XPE<R207G-XPE. In other words, the complex stability can be written as WT-XPE > R207G-XPE. These differences in their RMS deviations between WT and R207G XPA protein, and in between their respective complexes with XPE explains their ability to become a stable protein/complex, reflecting further on the impact that native and mutated amino acid has on the XPA protein. These RMSD analyses concerning the structural stability of the complexes as shown by both the systems here are in agreement with the binding affinity and the mutagenesis data reported by Wakasugi and team [151]. In our previous study, we also obtained a stable RMSD graph for all of our three docked models of the $XPA_{185-226}$-XPE complex [472].

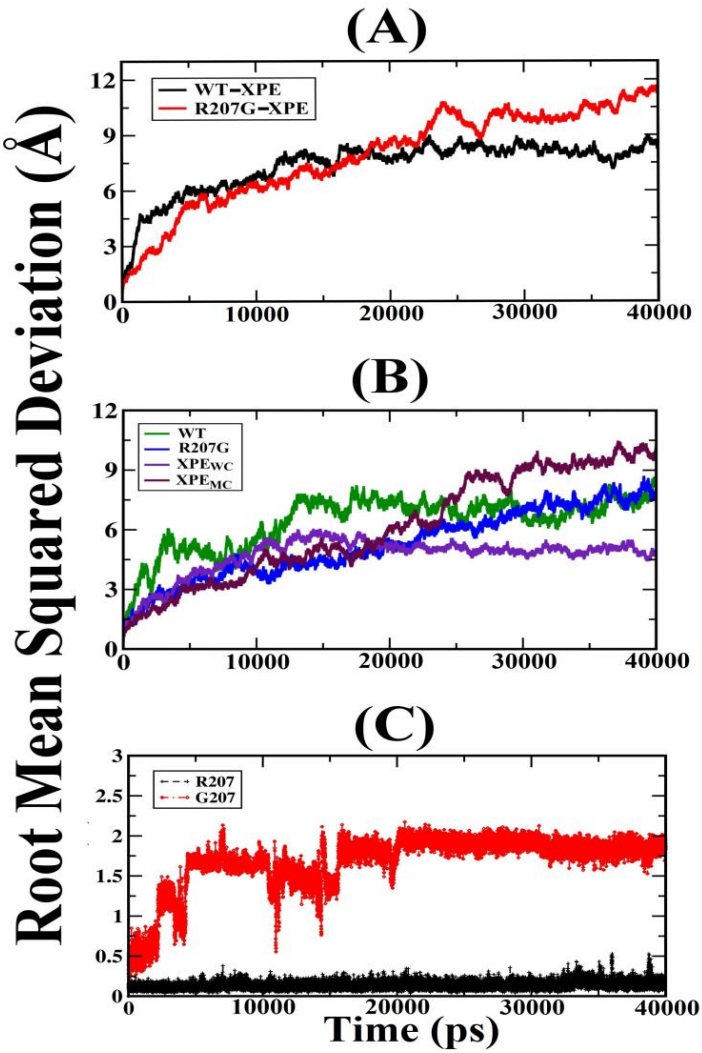

Figure 5B.8. Comparative RMSD analyses for (A) WT-XPE and R207G-XPE complexes, (B) monomeric units of the WT-XPE and R207G-XPE complexes, and (C) WT R207 and mutated G207 amino acids.

5B.4.4 Residue flexibility

The residue flexibility of both the WT and R207G mutant XPA-XPE systems was evaluated using the root mean square fluctuation (RMSF). **Figure 5B.9A** and **5B.9B** show the RMSF values for Cα atoms of individual monomers of both the complexes concerning the time evolution of 40 ns trajectories. For the WT XPA, the residue fluctuations were seen at its disordered region, which falls within the residue length of 219-273, while the fluctuations of the R207G mutant XPA system were observed to be quite higher compared to WT XPA, right from mutated residue G207 till the end part of

the protein. The RMSF comparison of the XPE from both the systems revealed that XPE_{MC} was observed to have endured more residue fluctuations compared to XPE_{WC}, especially at its N-terminal end, which is the binding site of XPE that we had identified in our previous study [472].

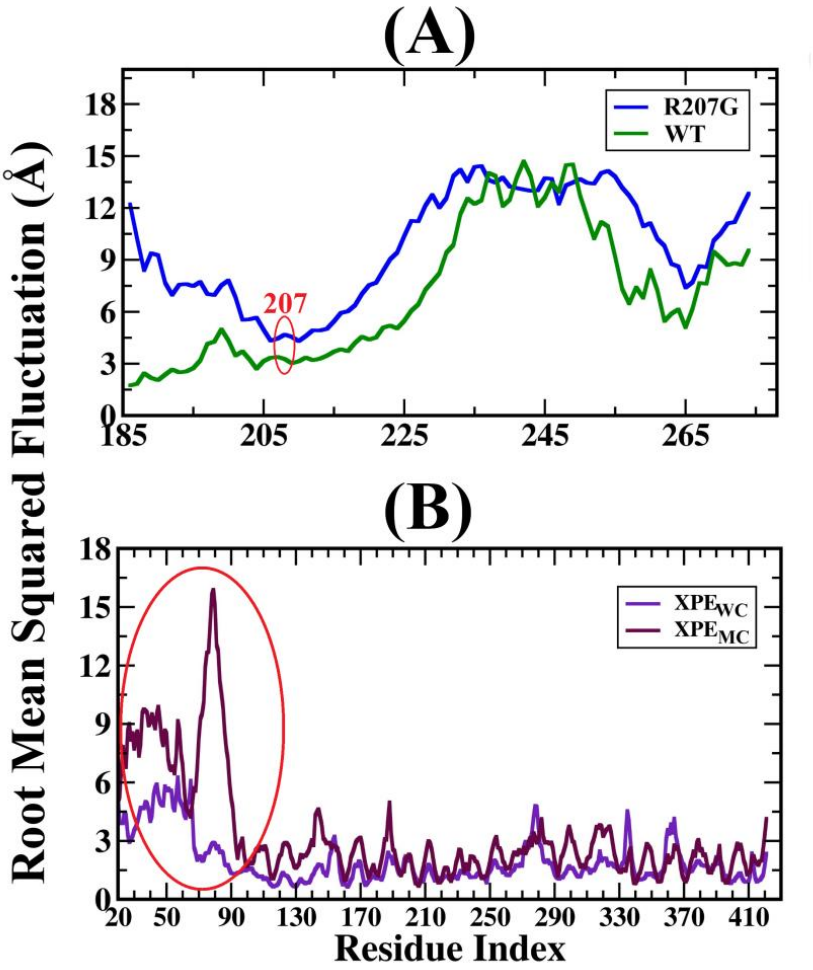

Figure 5B.9. Comparative RMSF analyses for (A) R207G and WT XPA, (B) XPE_{WC} and XPE_{MC}.

5B.4.5 Secondary structural changes

To further strengthen our study, we extracted corresponding snapshots (**Figure 5B.10**) of both the WT-XPE and R207G-XPE complexes at the time interval of 10 ns each to explore the changes in their secondary structural elements, taking their equilibrated structure as the reference. From **Figure 5B.10A**, we can see that the protein complex of WT-XPE was bound throughout the simulation time, while their mutant counterpart complex was bound together till 30 ns, after which they get separated as seen in **Figure**

5B.10B, indicating again that the mutant nature of R207G XPA makes it unfavorable for XPA to bind with XPE. Even when we studied the MD simulation of the $XPA_{185-226}$-XPE complex earlier, this complex was inbound form throughout the 40 ns simulation period, which also points to the fact that XPA in its WT form is true of a stable nature when bound to XPE. Here also, we observed that XPA was bound to XPE with its C-terminal faced upwards and N-terminal end faced downwards, which is again consistent with our previous findings where we had observed that in all three docked models of $XPA_{185-226}$-XPE complexes, the N-terminal end of $XPA_{185-226}$ was at the bottom position and C-terminal end at the top position [472]. XPE, on the other hand, showed no such significant changes in the WT-XPE complex state but exhibited changes in its binding site as seen from the RMSF plot.

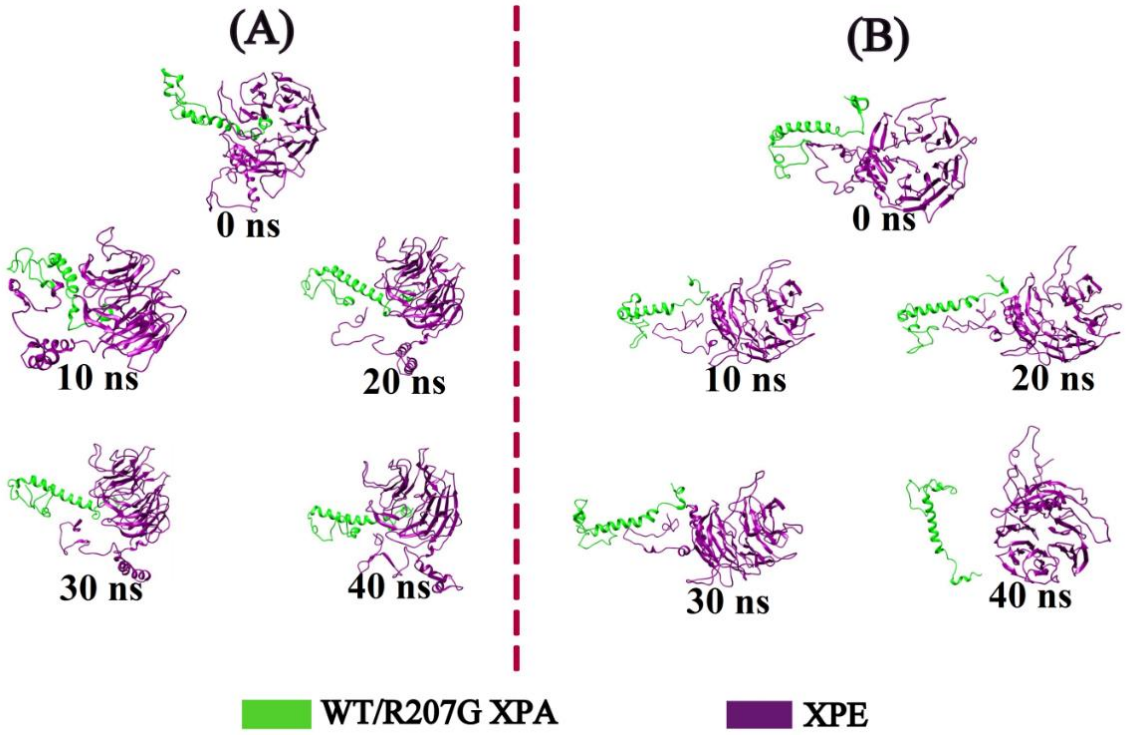

Figure 5B.10. Comparative snapshots for (A) WT-XPE and (B) R207G-XPE complexes.

The experimental study done by Wakasugi [151] showed that R207G mutation in XPA not only reduces the binding activity of XPA with XPE but also fails to ensure the successful removal of CPD damage in a simian vacuolating virus 40 (SV40) transformed human cells. Arginine, in particular, being a positively-charged, polar residue, is usually

involved in providing protein stability by forming hydrogen bonds, and salt bridges with negatively-charged amino acids, like glutamic acid, and aspartic acid. They are usually present in the binding sites of a protein, as a result of which their mutation to any other residue or in this case R207G makes the binding site unavailable for intermolecular interaction with other proteins as reported earlier [474]. In our previous study also, we had observed that R207 had formed multiple intermolecular bonds with XPE, including salt bridges and hydrogen bonds with Asp49 of XPE in all three $XPA_{185-226}$-XPE model complexes. Using the PRED analysis, we had found out R207 had scored the highest energy decomposition values (-7.27 kcal mol^{-1}, -9.47 kcal mol^{-1}, and -7.17 kcal mol^{-1} for models 1, 2, and 3) among the XPA's other interface residues (Leu191, Val193, Trp194, Glu198, Glu202, Glu205, Arg207, and Phe219) for all three docked models of $XPA_{185-226}$-XPE complex [472].

5B.4.5 Binding free energy (BFE) analyses of the WT-XPE and R207G-XPE complexes

The continuum solvent-based approach using the MM-PBSA algorithm was employed in our systems to determine the BFE required for WT/R207G and XPE to form WT/ R207G-XPE complexes. The binding energies of the two systems calculated by the MMPBSA.py script are shown in **Figure 5B.11**, and the detailed description of each component that contributed to their binding affinity has been given in **Table 5B.3**. The total BFE (ΔG_{bind}) for the complexes was found to be -58.9 kcal mol^{-1} and +16.44 kcal mol^{-1} for WT-XPE and R207G-XPE complexes, where the BFE of WT XPA with XPE was observed to be strong when compared to the mutant complex of R207G XPA with XPE. The higher binding affinity between WT XPA and XPE can be accorded to their electrostatic energy and the non-polar solvation free energy contributions, which were relatively low for the R207G-XPE complex. Likewise, we had also obtained similar high BFEs values for $XPA_{185-226}$-XPE complexes earlier (-48.37 kcal mol^{-1}, -49.09 kcal mol^{-1}, and -56.51 kcal mol^{-1} for Models 1, 2, and 3). This suggests that XPA in its wild-type state is the most favored form of XPA for the PPC formation with XPE, while the mutant state of XPA is the least favorable [472].

Apart from the previously known impacts and symptoms of XPA mutation on the human population, XPA mutants have now been linked with newer clinical

complications and disease conditions. A recent study demonstrated the association of polymorphisms in XPA rs1800975 (A23G) to be responsible for the early-onset of pre-eclampsia (PE) [473], myelodysplastic syndrome [474], and also for the occurrences of gastrointestinal stromal tumors [475]. The patients with XPA mutation have also been observed to have a medium risk of developing neuroblastomas, which was not observed before [196], although it was seen in XPD and XPF mutations [469]. Though we have tried to address the effect of one of the XPA's mutations *in silico* using the MD approach, many such XPA mutants are yet to be studied which may help us gain the much need deeper insights at the molecular and structural level. These findings may provide the scientific community with the required tools to combat the chemotherapeutic drug resistance due to enhanced NER expressions, and also build a strategy to tackle bad NER outcomes and their disease phenotypes.

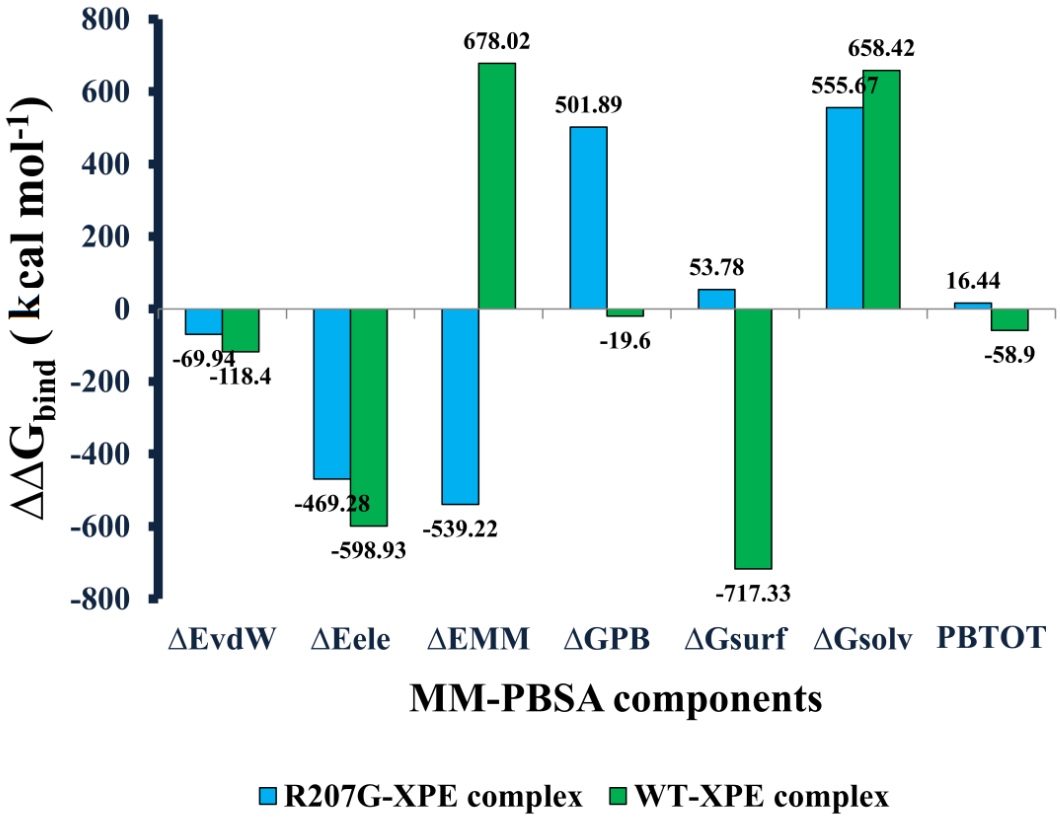

Figure 5B.11. Comparative BFE analyses for the WT-XPE and R207G-XPE complex.

Table 5B.3. Binding free energy (BFE) analyses for the WT-XPE and R207G-XPE complex.

Method	Models	Components	Complex (kcal mol^{-1})	Ligand (kcal mol^{-1})	Receptor (kcal mol^{-1})	$\Delta\Delta G_{bind}$ (kcal mol^{-1})
MM-PBSA	WT-XPE Complex	ΔE_{vdW}	-3642.94	-554.72	-2969.82	-118.4
		ΔE_{ele}	-32448.34	-6060.09	-25789.32	-598.93
		ΔE_{MM}	-6883.86	-2092.36	-5469.52	678.02
		ΔG_{PB}	198.15	54.88	162.87	-19.6
		ΔG_{surf}	-36091.29	-6614.81	-28759.14	-717.33
		ΔG_{solv}	-6685.7	-2037.47	-5306.65	658.42
		PB_{TOT}	-42776.99	-8652.29	-34065.79	**-58.9**
	R207G-XPE complex	ΔE_{vdW}	-3581.63	-552.61	-2959.08	-69.94
		ΔE_{ele}	-32514.87	-6128.04	-25917.5	-469.28
		ΔE_{MM}	-36096.5	-6680.66	-28876.6	-539.22
		ΔG_{PB}	-6767.8	-2109.06	-5160.64	501.89
		ΔG_{surf}	1560.79	215.97	1291.03	53.78
		ΔG_{solv}	-5207	-1893.09	-3869.6	555.67
		PB_{TOT}	-41303.51	-8573.75	-32746.2	**16.44**

ΔE_{vdW} = van der Waals contribution from MM; ΔE_{ele} = electrostatic energy as calculated by the MM force field; ΔE_{MM} = total gas phase energy (sum of ELE, VDW, and INT); ΔG_{PB} = the electrostatic contribution to the polar solvation free energy calculated by PB; ΔG_{Surf} = non-polar contribution to the solvation free energy calculated by an empirical model; ΔG_{solv} = sum of non-polar and polar contributions to solvation; PB_{TOT} = final estimated binding free energy in kcal mol^{-1} calculated from the terms above.

5B.5 Conclusion

In this study, we have demonstrated the effect of the R207G mutant in XPA on its binding interaction with XPE. We found the WT-XPE complex to be relatively stable than the R207G-XPE complex. The atomic fluctuations near the binding site of XPE in R207G XPA were found to be much greater than in WT XPA. From the RMSD analysis, we further noticed the instability in the R207G-XPE complex. The Binding free energy analysis supports the association of XPE with the WT XPA (-58.9 kcal mol^{-1}) to be stronger than with the mutant R207G XPA (+16.44 kcal mol^{-1}). As a whole, we found

R207 to have a significant role in regulating the protein function of XPA that can be affected upon R207G mutation, thereby having a much larger impact on its binding capacity with XPE, and also with the other NER proteins.