CHAPTER 4

To study the structural dynamics and molecular interactions of XPA with the DNA

CHAPTER 4A

To study the structural dynamics and interactions of XPA with the damaged DNA

CHAPTER 4A

To study the structural dynamics and interactions of XPA with the damaged DNA

4A.1 Abstract

Nucleotide Excision Repair (NER) in higher organisms repairs massive DNA abrasions caused by ultra-violet (UV) rays, and various mutagens, where Xeroderma Pigmentosum group A (XPA) protein is known to be involved in the damage recognition step. Any mutations in XPA cause classical Xeroderma Pigmentosum (XP) disease. The extent to which XPA is required in the NER is still unclear. Here, we present a comparative study on the structural and conformational changes in the globular DNA binding domain of XPA_{98-210} in DNA-bound and DNA-free states. Atomistic Molecular Dynamics (MD) simulation was carried out for both XPA_{98-210} systems using AMBER force fields. We observed that XPA_{98-210} in presence of damaged DNA exhibited more structural changes compared to XPA_{98-210} in its free form. When XPA is in contact with DNA, we found marked stability of the complex due to the formation of characteristic longer antiparallel β-sheets consisting mainly of lysine residues.

4A.2 Introduction

Nucleotide Excision Repair (NER), is one of the most predominant repair pathways in higher organisms that repairs, and restores massive DNA distorting lesions caused by harmful ultraviolet (UV) rays, environmental toxins, and various other mutagens [48, 50, 354, 402]. The most common result of these exposures is pyrimidine dimers, such as cyclo-butane pyrimidine dimers (CPD), 6–4 photoproducts, and cisplatin-DNA intra-strand crosslinks [46, 47, 49]. The protein Xeroderma Pigmentosum complementation group A (XPA) plays a vital role in NER by acting as a scaffolding protein, and systematizing the assemblage of other NER core factors around the DNA damage site before damage excision [54, 127, 164].

NER pathway makes use of over thirty different proteins to bring out this multi-step of ‘cut-and-paste’-like repair process [55, 56]. These steps involve: i) DNA lesion recognition, ii) unwinding of the DNA helix around the lesion, iii) 3/ and 5/ dual incision and excision of the lesioned segment, iv) DNA synthesis to fill the gap, and v) ligation

of the newly synthesized strand [5, 49, 57, 81, 86, 402, 403]. The variations in these DNA binding proteins of the NER pathway often result in several genetic disorders, like Xeroderma Pigmentosum (XP), Cockayne syndrome, and Trichothiodystrophy [404, 405], distinguished by extreme sun reactivity, usually coinciding with cancers, developmental retardation, progressive aging, immunological disorders, and neurodegeneration. The mutations in XPA are mostly associated with the Classical XP phenotype [165, 372, 404, 406]. XPA functions in both global genome NER (GG-NER) as well as in transcription-coupled NER (TC-NER) [86, 361], recruited to the damage site by the transcription factor II H (TFIIH) complex that relaxes duplex DNA around the erroneous site by its helicase activity creating the NER bubble [75, 144, 357, 361, 370, 407-411].

The 273 residues (40 KDa) XPA protein comprises a central globular DNA binding domain (98-219), C4 type zinc-finger motif in the N-terminal region, and a shallow basic cleft in the C-terminal region [114-116]. XPA has its N- and C- termini regions highly disordered [134, 144, 412], and oversees a variety of protein interactions [116, 136, 151, 413-415]. Many of the XP symptoms caused as a result of XPA mutations are present in the globular domain [406]. The XPA Zn-containing core (residues 141-176) binds both the DNA junction [136] and the RPA70 domain [125, 127, 369], which protects the undamaged strand. The only determined NMR structure so far of XPA with its globular domain is XPA98-210 (PDB code: 1XPA) [114, 115] and XPA98-208 (PDB code: 1D4U) [116], that present 3 helices spanning between 141-210 residues, followed by a short β-sheet and some imperfect structured loops, including the Zn binding site [115]. A recent study done by Hilton and the team has put forth the DNA binding ability of XPA98-210 to be influenced by lysine residues – K168 and K179 [126].

Even though the DNA binding core of XPA has been studied much, the enigma behind the initial DNA recognition and binding step of XPA in the NER process and recruitment of other proteins to the NER bubble is still persistent. Based on these grounds, the interaction between XPA protein and DNA becomes a priority. So, the main focus of this study is to 1) study the structural and dynamic features that contribute to the stability of the complex structure of XPA (98-210) and DNA; and 2) to determine

the degree of changes in the secondary structure of the protein with regards to DNA interaction.

Here, in this computational study, we have taken experimentally determined 3D structure of XPA (98-210) (PDB ID: 1XPA) (Ikegami et al., 1998), determined using NMR and a DNA decamer containing cis-syn thymine (PDB ID: 1N4E) [416], determined using X-ray crystallography from Protein Data Bank (PDB) [377, 378, 417] and docked them using HEX Dock software 8.0.0 [418]. We obtained the complex structure of XPA_{98-210}-DNA with the highest E-value and submitted it to the DNAproDB server [351] to study their nucleotide-residue interface interaction. We then carried out the conformational dynamics of two XPA (98-210) systems: DNA-bound, DNA-free, and had trajectories analyzed. All the visualization of Molecular data and analyses obtained from the simulation were performed with the UCSF Chimera package [315]. By analyzing the impact of DNA on the conformation of the protein, especially the Zn-motif domain and C-terminus region, we attempted to summarize the binding pattern of XPA_{98-210} to the DNA in NER. We saw the emergence of elongated β-sheets in the case of DNA-bound XPA_{98-210}, where lysine residues K168 and K179 were involved. Helices were directly involved in DNA interaction.

4A.3 Systems and methodology

4A.3.1 Computational model

Out of THE two available NMR structures of XPA: XPA_{98-210} (PDB code: 1XPA), and XPA_{98-208} (PDB code: 1D4U), we selected 1XPA as our initial configuration over 1D4U due to its longer residue length. Zinc metal ion was removed from the 1XPA PDB text format file before proceeding to the docking step. This is because the conformational dynamics and stability of XPA were found to be the same in the presence and absence of Zinc metal ions [418]. We took X-ray crystallography determined nucleic acid decamer (PDB code: 1N4E) from PDB as our starting DNA structure having a cis-syn thymine dimer d[GCTTAATTCG] d [CGAAT ∗T ∗AAGC] with the resolution of 2.5 Å.

4A.3.2 Docking studies

A docking study was performed between XPA98-210 as receptor and DNA as ligand using Hex Dock software 8.0.0, a Fast Fourier Transform (FFT) docking tool. The docked results were scored based on the spherical polar Fourier (SPF) technique, which uses rotational correlations [418] to accelerate the calculation.

HEX Dock Parameters:

Correlation type – Shape + Electro

FFT Mode – 3D

Sampling methods- Range angles

Grid Dimension – 0.6

Receptor range – 180

Ligand Range – 180

Twist range – 360

Distance Range – 40

Steric scan – 16

Final search- 30

Five Stages of Hex Dock Process:

SPF transform -> FFT Steric scan -> FFT final search -> energy refinement -> total dockings.

Once the five-staged docking was completed, it gave a docked structure indicating the corresponding E-values. The higher the negative E-value, the more effective the docking process.

4A.3.3 Inter-molecular nucleotide-residues interaction

To study the nucleotide-residues interaction, and to assess the forces stabilizing the structure of XPA_{98-210} with the DNA, we procured the docked structure generated from HEX-dock with higher binding energy and uploaded it to the DNAproDB server [351]. These interactions were considered if they exhibited any of the following interactions, which is-

i) Major groove residue interaction: at least 2.0 Å2 buried accessible surface area (BASA), 1 hydrogen bond and 1 van-der Waals (vdW) interaction;

ii) Minor groove residue interaction: 1.0 Å2 BASA, 1 hydrogen bond and 1 vdW interaction;

iii) Backbone-residues interaction: 5.0 Å2 BASA, 1 hydrogen bond and 1 vdW interaction.

4A.3.4 Molecular dynamics simulation of XPA_{98-210}- DNA Bound and DNA Free Systems

The molecular dynamics simulation of this study was executed using the Particle Mesh Ewald Molecular Dynamics (PMEMD) [291, 389] module of the AMBER12 [388] software package. XPA_{98-210}-DNA free system was simulated with AMBER ff99SB force field [283] while XPA_{98-210}-DNA bound complex was simulated using AMBER ff99bsc0 force field [420, 421] of the AMBER12 software package. The missing hydrogen atoms were added to all the XPA_{98-210} systems in the Leap module of the AMBER12 package. XPA-Apo was added with 6 Na^+ ions and 7871 water molecules, while XPA-DNA bound system was neutralized with a total of 24 Na^+ ions and 8819 water molecules. For the required hydration of both systems, we used a TIP3P [297] water box with a buffer distance of at least 10 Å around the solute. The method devised by Spector and his team was chosen to address the parameterization of the thymine-thymine dimer [422]. The initial minimization was done by holding the restraints over the solute for 500 steps using the steepest descent algorithm followed by another 500 steps with a conjugate gradient process. The second minimization was carried out without any restraints for another 500 steps. Both minimizations were carried out taking the cut-off of 8Å. Heating dynamics were applied to both the XPA_{98-210}- DNA bound and DNA free complexes with a gradual increase in the temperature from 0-300K.

The whole arrangement was optimized, thermalized, and equilibrated using standard equilibration protocol. All simulations were carried out using periodic boundary conditions under isothermal and isobaric conditions (T = 300 K; P = 1 atm). The Berendsen thermostat [295] was used for controlling the temperature. The shake

[293] algorithm was used to restrain all the bonds at the time step of 2fs. After each MD run, trajectory files generated were analyzed using cpptraj [317]. Comparative analysis of Root Mean Square Deviation (RMSD), Radius of Gyration (Rg), Solvent Accessible Surface Area (SASA), and B-factor was performed for both the XPA_{98-210}- with and without DNA. For the visualization of the results, UCSF Chimera was deployed.

4A.4 Results and discussion

4A.4.1 XPA_{98-210} binding mechanism with the damaged DNA using Hex Dock softwaree 8.0.0.

To understand the proper binding mechanism of XPA_{98-210} with DNA, molecular docking was conducted between XPA_{98-210} and cis-syn DNA decamer having thymine dimer using Hex Dock software 8.0.0. Docking yielded results of top 100 confirmations of XPA_{98-210} bound to DNA. Among these conformations, the best conformation with the highest negative value indicating higher binding energy (-591.91 kcal mol^{-1}) was chosen. We analyzed the obtained docked structure (**Figure 4A.1A**) of XPA_{98-210}-DNA using UCSF Chimera software. We observed the C-terminal cleft and helices in close contact with the lesioned DNA. Surface structure analysis (**Figure 4A.1B**) showed XPA_{98-210} to be tightly bound to the damaged DNA ascertaining its hold over damaged DNA and initiation of the repair process.

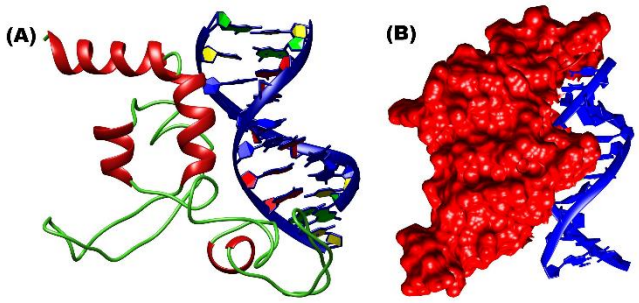

Figure 4A.1. XPA_{98-210}: DNA bound system: (A) Docked Structure obtained from Hex Dock8.0.0 Software, (B) Surface structure of XPA_{98-210} with the DNA showing their closely bound stature.

4A.4.2 XPA$_{98-210}$-DNA interaction study

We studied nucleotide-residues interaction from the results obtained from the DNAproDB server for our complex. A cartoon representation of the immediate interaction of the protein to the DNA surface is shown in **Figure 4A.2A**, where the participating region is highlighted in opaque red color which in this case is the first and third helix, some portions of the C-terminal cleft, a and few regions from the Zn binding loop regions. Buchko *et al*. [368] modeled a 9-n single-stranded DNA oligomer with XPA$_{98-219}$ and observed that the C-terminal cleft region was found at the site of DNA interaction. The polar contact map of the XPA$_{98-210}$-DNA complex is shown in **Figure 4A.2B.** This map in particular describes the occupancy of the residues in each region of the DNA- backbone, and grooves (major and minor). A major portion of the DNA in this polar map is seen to be occupied by the loop region (100-195). Helix 1 covers the major groove while helix 3 occupies the backbone and major groove region of DNA. All the interactions are confined to one side of the map meaning that the protein is bound to one side of the DNA. Helix 3 contains residues Lys183, Leu184, Gln185, Ile186, Val187, Lys188, Arg189, Ser190, Leu191, Glu192, Val193, and Trp194, while helix 1 has residues Tyr116, Leu117, Met118M, Asn119 and His120. The inter-molecular nucleotide-residue interaction between XPA$_{98-210}$ and DNA is illustrated in **Figure 4A.2C**. DNA strand is displayed in the form of the ladder with the interacting residues on either side of the strand. Here, also the squares and circles represent the loop region and helices respectively. XPA$_{98-210}$ interacted with the DNA backbone of strand A with the loop residues Gly195, Cys153, Asp154, Lys157, Arg158, Glu159, Pro160, Glu106, and, Gly109 while the DNA backbone of strand B showed interactions with loop residues Arg158, Leu123, Pro160, Lys163, Asp122, Phe100, and Tyr102. The residues belonging to helix 3 also interacted with the backbone regions of DNA, wherein Val193 and Glu192 showed interactions with strand A, and Glu192, Lys188, Arg189, and Gln185 showed interaction with strand B. Minor groove showed less loop interaction mainly Tyr102 with the strand B and Arg158, Pro160 and Tyr102 with the strand A. Helix 3 residues Glu192 and Arg189 interacted with the nucleotides present in the minor groove of both the strands. Met118 of helix 1 made contact with the nucleotide present within the major groove of strand B. Strand A showed no such interactions. Loop

regions Arg158, Glu106, and Pro124 made the connection to strand A via a major groove. Leu123, Pro124, Asp122, Phe100, and Tyr102 bonded with the nucleotide within the major groove of strand B. It has been observed that the residues 141-176 of XPA in the Zn-containing core region are responsible for the binding of protein to RPA70 as well as DNA junction [58, 125, 127]. We also observed similar findings in our study as shown in **Fig. 4A.2A, 4A.2B,** and **4A.2C**. Each of the interactions seen in **Fig. 4A.2B and 4A.2C** is stabilized and supported by the hydrogen binding, establishment of van-der Waals interactions, and the availability of the buried solvent accessible surface area (BASA).

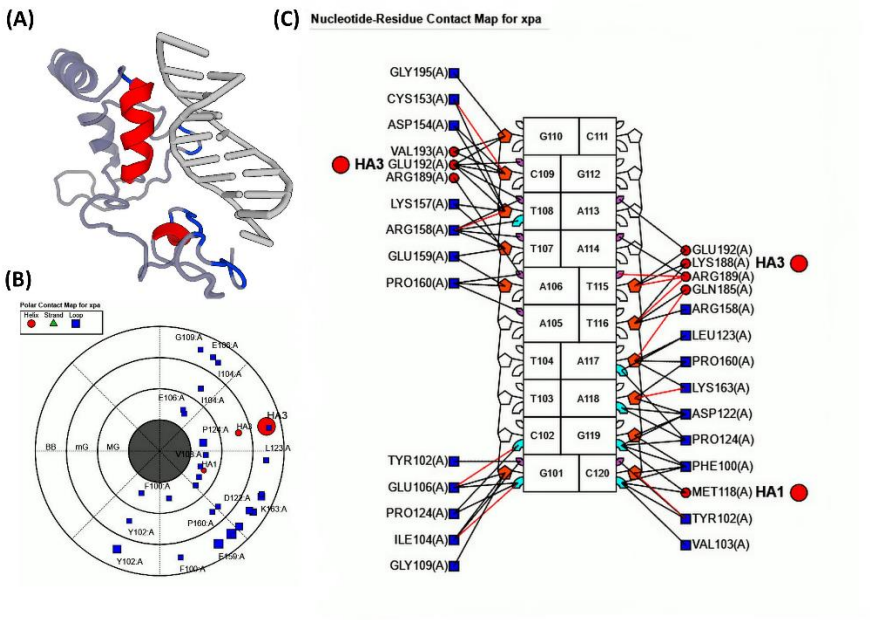

Figure 4A.2. Graphical representation of the interaction between Xeroderma Pigmentosum A (XPA_{98-210}) protein with DNA. The strands are shown in green triangles, loops in blue squares, and helix in red circles. HA1 is helix 1 and HA3 helix 3. (A) Cartoon structure of the XPA_{98-210}-DNA complex showing helices and loop residues in the red opaque color that interacts with the DNA. (B) Polar contact map for XPA_{98-210}-DNA complex. Mostly loop and helices are involved in the interaction with the minor groove (mG), major groove (MG), and backbone regions of the curved helical DNA, with their contribution, restricted to one side of the structure. (C) Nucleotide-residue contact map, showing loop and helix residues interacting with DNA backbone and both grooves.

However, one of the major unusual findings we observed was the involvement of acidic residues Glu and Asp in nucleotide-residue interaction. This might be due to the reason that acidic residues bind to the groove of DNA and weakens the hydrogen bonds between the DNA strands. Glu, in particular, has been known to bind to major grooves disrupting 3 hydrogen bonds thereby destabilizing the DNA helix. Glu, Asp, and Pro residues as a whole decrease the melting temperature of the DNA causing easy strand separation [423]. Another possible explanation for the participation of acidic residues may be aiding in the enhancement of the helicase activity of TFIIH by XPA, after being recruited by the same upon TFIIH-XPA interaction [357].

The detailed features of the hydrogen bonds, BASA, and van der Waals interaction formed between individual nucleotide and protein residues to retain the XPA_{98-210}-DNA complex stable are discussed in **Table 4A.1**. **Table 4A.2A** depicts the values of each bond interaction displayed in the polar contact map of XPA_{98-210}-DNA from **Fig. 4A.2B**. We can observe that protein occupied a major space near the backbone of DNA, especially with the loop region and helix 3. Helix 3 interacted with the backbone regions of DNA, exhibiting 4 hydrogen bonds, and 148 vdW contacts, and the BASA was found to be 206.24 $Å^2$. Arg158 showed higher occupancy in the backbone surface with the BASA score of 31.51 $Å^2$, 2 hydrogen bonds, and 33 vdW interactions. The interaction of major and minor grooves had fewer values. Helix 3 and helix 1 showed bonding with a minor and major groove with the BASA score of 25.66 $Å^2$ and 12.87 $Å^2$ respectively. Polar contact analysis of backbone interaction also showed the residues K183, L184, Q185, I186, V187, K188, R189, S190, L191, E192, V193, and W194 present in helix 3 to have contributed to the contact with 206.24 $Å^2$ BASA, 4 hydrogen bonds, and 148 vdW interactions. Backbone interaction had 7 hydrogens while major and minor grooves had 1 and 2 hydrogen bonds respectively. Helix 1 showed neither hydrogen bonding nor vdW interaction with the major groove. Helix 3 on other hand formed 2 hydrogen bonds and 8 vdW interactions with a minor groove of DNA. Loop residues in both the grooves showed BASA values in the range of 1.56-13.43 $Å^2$, hydrogen bonding, and vdW contributions were less than backbone interactions. **Table 4A.2B** illustrates the contributions of secondary structure element

(SSE), BASA, hydrogen bond, and vdW forces towards the stability of residue-nucleotide interaction in the XPA_{98-210}-DNA complex. Here also, loop resides and helix 3 showed maximum interaction with the DNA through backbone contacts. Major and minor grooves showed comparatively less interaction with the protein. Helix 1 is seen to be involved in the bond formation only with the major groove. Helix 3 contributed to the backbone and minor groove interaction with 4 hydrogen bonds in backbone interaction and with two hydrogen bonds in the minor groove. Loops interaction in all DNA regions, be it backbone or grooves, provided sufficient benefaction towards the establishment of stable protein DNA interaction. All these molecular contacts depicted in **Tables 4A.2A, 4A.2B**, and **4B** contributed greatly to the DNA-protein relationship.

Table 4A.1. Residues involved in β-sheets formation in XPA_{98-210} systems: DNA bound and DNA free during the simulation.

Systems	Residues involved in the β-sheets formation	Time period
XPA-DNA bound	Phe164, Ile165, Val166, Lys167, Lys168, Asp177, Met178, Lys179, Leu180 and Tyr181	Full 40ns
XPA-DNA free	Leu138, Ile139, Thr140, Leu180, Tyr181 and Leu182	Full 40ns
	Val103, Ile104, Cys105, Lys110, Glu111, Phe112, Leu138, Ile139 and Thr140	10-30ns

Table 4A.2. Residues involved in Xeroderma Pigmentosum A (XPA) protein-DNA interaction.

Table 4A.2A. Polar Contact Map of XPA-DNA.

Interactions	Residues	SSE	BASA ($Å^2$)	HB	vdW
Backbone	Y102	Loop	38.91	1	17
	F100		15.61	0	0
	E159		50.29	0	22
	K157		18.78	0	5
	P160		55.04	0	7
	R158		31.04	2	33
	D122		28.07	0	5
	C153		26.77	1	0
	D154		7.95	0	9
	K163		31.51	1	2

	L123		8.21	0	1
	G195		12.07	0	0
	I104		8.64	0	0
	E106		16.54	0	5
	G109		11.66	0	3
	K183, L184, Q185, I186 V187, K188, R189, S190, L191, E192 V193 W194	Helix 3 (HA3)	206.24	4	148
Minor Groove	Y102	Loop	8.05	0	4
	P160		4.72	0	0
	R158		2.32	0	5
	I104		6.86	0	1
	K183, L184, Q185, I186 V187, K188, R189, S190, L191, E192 V193	Helix 3 (HA3)	25.66	2	8
Major Groove	Y102	Loop	0.67	0	4
	F100		8.26	0	2
	R158		0	0	2
	D122		13.43	0	0
	V103		9.1	0	0
	L123		4.18	0	0
	P124		33	0	10
	I104		0.32	1	6
	E106		11.51	1	17
	116Y, 117L, 118M, 119N, 120H	Helix 1 (HA1)	12.87	0	0

Table 4A.2B. Linear Contact Map of XPA-DNA.

Interactions	Residues	SSE	BASA	HB	vdW
Backbone	G195-DG110	Loop	17.44	0	0
	F100-DG119		20.43	0	0
	F100-DC120		7.73	0	0
	Y102-DC120		60.41	1	17
	E106-DC120		22.15	0	5

	G109-DG101		14.54	0	3
	D122-DA118		23.29	0	5
	L123-DA117		29.34	0	1
	C153-DC109		15.65	Q	0
	C153-DT108		16.52	0	0
	C154-DC109		0	0	3
	D154-DC109		4.25	0	6
	R158-DT116		7.72	0	0
	R158-DT108		5.96	2	12
	R158-DT107		7.2	0	21
	K157-DT107		0.95	0	1
	K157-DT108		22.12	0	4
	E159-DA106		44.14	1	16
	P160-DA106		13.36	0	0
	P160-DA118		15.24	0	0
	P160-DA117		31.63	0	7
	K163-DA117		13.91	0	0
	HA3-DG110		10.37	0	2
	HA3-DC109		29.1	0	38
	HA3-DT108	Helix 3 (HA3) Loop Loop Helix 3 (HA3)	1.29	0	4
	HA3-DT115		98.92	1	44
	HA3-DT116		41.3	1	49
	HA3-DA117		14.34	2	11
	R158-DT108		0	0	2
	R158-DT107		0.09	0	3
Minor Groove	P160-DA106	Loop Helix 3 (HA3) Helix 3 (HA3) Loop	8.61	0	0
	P160-DA105		1.43	0	0
	Y102-DG101		4.78	0	0
	Y102-DC120		4.31	0	4
	HA3-DC109		4.38	0	0
	HA3-DT108		6.8	0	2
	HA3-DA106	Helix 3 (HA3) Loop Loop Helix 1 (HA1)	2.26	0	0
	HA3-DA113		2.91	0	0
	HA3-DA114		1.87	0	2
	HA3-DT115		0.21	2	4
	R158-DT108		0	0	2
	E106-DC102		13.85	1	8

Major Groove	E106-DG101	Loop Helix 1 (HA1) Helix 1 (HA1)	2.76	0	9
	P124-DC102		4.24	0	2
	P124-DG119		2.18	0	2
	P124-DA117		2.64	0	0
	P124-DA118		7.91	0	0
	F100-DC120		11.32	0	2
	F100-DG119		6.85	0	0
	Y102-DC120		5.17	0	4
	D122-DA118		4.18	0	0
	D122-DG119		8.63	0	0
	L123-DA117		8.88	0	0
	HA1-DG119		7.06	0	0
	HA1-DC120		4.77	0	0

Table 4A.2C. Nucleotide-residue contact map.

Interactions	Residues	SSE	BASA	HB	vdW
Backbone	Gly195-DG110	Loop	17.44	0	0
	CYS153-DC109		15.65	1	0
	CYS153-DT108		16.52	0	0
	ASP154-DC109		0	0	3
	ASP154-DT109		4.25	0	0
	LYS157-DT108		22.12	0	4
	LYS157-DT107		0.95	0	1
	ARG158-DT108		5.96	2	12
	ARG158-DT107		7.2	0	21
	ARG158-DT116		7.72	0	0
	GLU159-DT107		44.14	0	16
	GLU159-DA106		8.06	0	6
	PRO160-DA106		13.36	0	0
	GLU106-DG101		22.15	0	5
	GLY109-DG101		14.54	0	3
	LEU123-DA117		29.34	0	1
	PRO160-DA117		31.63	0	7
	PRO160-DA118		15.24	0	0
	LYS163-DA117		13.91	0	0
	LYS163-DA118		39.71	1	0
	ASP122-DA118		23.29	0	5

	ASP122-DG119		8.39	0	0
	PHE100-DG119		20.43	0	0
	PHE-DC120		7.73	0	0
	TYR102-DC120		60.41	1	17
	VAL193-DG110	Helix 3 (HA3)	6.57	0	0
	VAL193-DC109		11.3	0	5
	GLU192-DG110		3.8	0	2
	GLU192-DC109		17.8	0	33
	GLU192-DT108		1.29	0	4
	GLU192-DT115		10.21	0	0
	LYS188-DT115		71.64	0	17
	LYS188-DT116		10.27	0	1
	ARG189-DT115		17.07	1	27
	ARG189-DT116		11.97	1	13
	GLN185-DT116		19.06	0	35
	GLN185-DA117		14.34	2	11
Minor Groove	ARG158-DT108	Loop	0	0	2
	ARG158-DT107		0.9	0	3
	PRO160-DA106		8.61	0	0
	PRO160-DA105		1.43	0	0
	TYR102-DG101		4.78	0	0
	TYR102-DC120		4.31		4
	GLU192-DC109	Helix 3 (HA3)	4.38	0	0
	GLU192-DT108		6.37	0	2
	GLU192-DA113		2.91	0	0
	ARG189-DA106		2.26	0	0
	ARG189-DA114		1.87	0	2
	ARG189-DT115		0.21	2	4
Major Groove	ARG158-DT108	Loop	0	0	2
	GLU106-DC102		13.85	1	8
	GLU106-DG101		2.76	0	9
	PRO124-DC102		4.24	0	2
	PRO124-DG101		0	0	6
	LEU123-DA117		8.88	0	0
	PRO124-DA117		2.64	0	0
	PRO124-DA118		7.91	0	0
	PRO124-DG119		2.18	0	2

	ASP122-DA118		4.18	0	0
	ASP122-DG119		8.63	0	0
	PHE100-DG119		6.85	0	0
	PHE100-DC120		11.32	0	2
	TYR102-DC120		5.17	0	4
	MET118-DG119	Helix 1 (HA1)	7.06	0	0
	MET118-DC120		4.77	0	0

SSE- secondary structure element
vdW- van der Waal interactions
HB- hydrogen bonds
BASA- Buried solvent-accessible surface areas
DT- deoxythymidine
DA- deoxyadenosine
DC- deoxycytidine
DG- deoxyguanosine
Helix 3 (HA3)- K183, L184, Q185, I186, V187, K188, R189, S190, L191, E192 V193 W194
Helix 1 (HA1)- Y116, L117, M118, N119, H120

In our study, we found that most of the residues involved in the DNA-protein interaction belonged to the loop region, helix 1 and helix 3 of XPA. The residues of the loop region and helix 1 share a protein binding site for RPA70AB [127]; helix 3 is the protein binding site for DDB1/2 (XPE) [151]. It has been found that the residues of this region upon mutation fail to yield functional XPA protein causing severe XP-A [169, 362, 406]. Mutations of Cys residues that form coordination with Zn metal ion result in protein misfolding adding to the severity of the XP phenotype [166]. The mutations of Y116 [372] and L117 [172] have been known to cause neurological damage in addition to XP. Even though XPA has been reported with much deleterious mutation, XPA has still maintained evolutionary conservation of its genes. XPA studies in mice, Drosophila, Yeast, and humans have shown a high rate of evolutionary conservations of their genes [424, 425]. However, it had never been reported in lower taxa, until recent publications on the presence of XPA gene in Hydra confirmed the conservation of nuclear localization signal (NLS), Zn-finger motif, binding regions to other NER proteins, and 2 Pfam domains [426], agreeing to the evolutionary conservation of XPA genes in all living systems. Most of the lysine residues that have been recognized as important residues in DNA binding are conserved well in both human XPA [149, 427] as well as in Yeast [129]. Few exceptions remain wherein yeast residues Thr239, Phe262 and Gln266 are replaced by Lys151, Trp175, and Lys179 in humans.

4A.4.3 MD simulation study on the XPA$_{98-210}$: DNA free and DNA bound state.

In this study, we probed the stability of the Xeroderma Pigmentosum protein A (XPA$_{98-210}$) in the explicit solvent concerning the presence and absence of DNA investigating its structural and conformational changes as a function of time. We first checked the systems for the attainment of the equilibration stage. Next, we took the equilibrated model structure of both the XPA$_{98-210}$ systems-DNA bound and DNA-free as a reference structure for the RMSD analysis. The comparative RMSD graph for the XPA$_{98-210}$-DNA bound and DNA-free is shown in **Figure 4A.3A**. DNA bound XPA$_{98-210}$ demonstrated a higher RMSD value of 4.5Å while DNA unbound XPA$_{98-210}$ presented the RMSD value of 3.1 Å ± 0.5. Hu *et al*., 2014 conducted a similar comparative study on XPA in Zn-bound, Ni-bound, and Apo forms, where they observed that RMSDs between XPA-Zn bound and Apo form had the least difference compared to what XPA exhibited when bound to Nickel. They too had obtained stability of XPA: Zn bound and Apo systems at around 3 Å.

To quantify the compactness and the mass-weighted spatial distribution of XPA$_{98-210}$ in DNA bound and unbound cases, we calculated the radius of gyration (Rg). **Figure 4A.3B** shows the radius of gyration for both the systems as a function of time. XPA$_{98-210}$ system stayed mostly steady with the least variation due to the loss of the Zinc metal chelation. In contrast, the XPA$_{98-210}$-DNA bound complex has shown a significant amount of fluctuation after 30 ns. XPA$_{98-210}$ in presence of DNA expressed higher Rg values of approximately 17.5 Å till 30 ns, after which it started oscillating. Rg value of XPA$_{98-210}$ in Apo form was indicative of being much stable at the range of 16.5 Å.

Solvent accessible surface area (SASA) analysis for the DNA bound and DNA unbound XPA$_{98-210}$ was assessed as well. From **Figure 4A.3C**, we reckon that majority of the residues in both systems had better access to the solvent provided during the whole simulation process. DNA-bound XPA has a higher SASA value of almost 10000-12000 Å^2 explaining its larger surface area exposed to the solvent, whereas XPA$_{98-210}$ in its Apo form has a lower surface area accessible to the solvent within the range of 8000 Å^2 probably due to the initial structure being NMR derived that was already in the solution form.

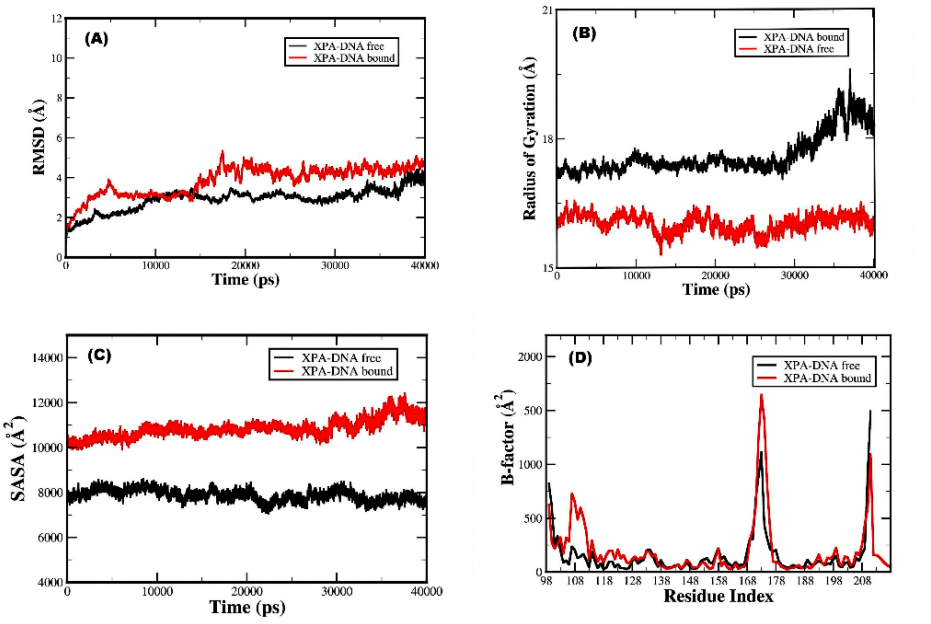

Figure 4A.3 Comparative MD analyses of XPA$_{98-210}$- DNA free and DNA bound form. (A) RMSD values for XPA$_{98-210}$-Apo and XPA$_{98-210}$-DNA complexes relative to the starting structure during MD simulation. (B) The radius of Gyration (Rg) values of XPA$_{98-210}$-Apo and XPA$_{98-210}$-DNA complex. (C) Solvent accessible surface area (SASA) values for XPA$_{98-210}$-Apo and XPA$_{98-210}$-DNA complex. (D) B-factor distribution for XPA$_{98-210}$-Apo and XPA$_{98-210}$-DNA complex.

Additionally, we plotted the B-factor values for the backbone atoms for the XPA$_{98-210}$ with and without the DNA to inspect the fluctuation of all atoms about their average positions and to gain insight into understanding the flexibility and regions of XPA$_{98-210}$. The backbone Cα atoms in XPA$_{98-210}$ taken from MD simulation were taken into account to estimate the B-factor values and were respectively plotted against their residue number as shown in **Figure 4A.3D**. Both the systems expressed similar fluctuations throughout the simulation (**Fig. 4A.3D**) followed by rapid decline indicating the relaxation in the structure. XPA$_{98-210}$-DNA bound complex comparatively expressed higher and wider peaks than apoprotein, which portrays the change in the structure of XPA$_{98-210}$ when in contact with the DNA. XPA$_{98-210}$-DNA bound complex showed remarkable higher peaks at the region 98-118 which is the Zinc binding motif. Peaks were seen in the region of 168-179 where the length of β-sheets saw extensions with loops in between which counter proves its stability in the complex formation with DNA, as mentioned in the literature provided by Hilton et al. [126].

We took comparative snapshots of both the XPA_{98-210} systems –DNA bound and DNA free at different intervals of the simulation period and were shown in **Figure 4A.4.** DNA-bound XPA_{98-210} displayed more conformational variations compared to DNA-free XPA_{98-210}. Over time, helices in both the XPA_{98-210} seemed to have lost part of their helical motif forming turns. The most striking feature which could be easily noticed in the snapshots was the variation in the β-sheets formation between both the systems. Using the UCSF-Chimera, we analyzed the β-sheets forming residues in each snapshot from the obtained trajectories. The resultant observation is shown in **Table 4A.1**. DNA-free XPA_{98-210} in its Apo forms saw β-sheets consistent only with a few residues Leu138, Ile139, Thr140, Leu180, Tyr181, and Leu182, while the residues Val103, Ile104, Cys105, Lys110, Glu111, Phe112, Leu138, Ile139, and Thr140 formed β-sheets in between 10-30ns of simulation, but disappeared during the later stage of simulation. As per the study done by Hilton et al [126], we understand that Lys168 and Lys179 are crucial to DNA binding. Sugitani et al. [54] suggested that since the DNA binding domain of XPA_{98-219} lack major basic residues in its C-terminal region hence isn't a good model to describe the DNA binding capacity of the protein. They also suggested that the DNA binding domain increased beyond 98-219 that is from 98-219 to 98-239, which may lead to a much more concrete model for the DNA-protein interaction. Another piece of evidence provided by Koch et al. [129], upon the study done in XPA homolog Rad14 in yeast highlights the possibility of the existence of XPA in dimeric form, as XPA interacts with numerous proteins and monomeric form would hinder its prime role as a scaffolding protein.

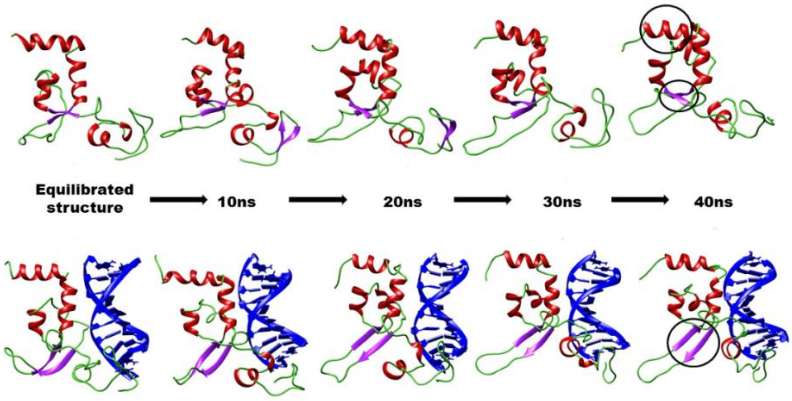

Figure 4A.4 Comparative snapshots for XPA$_{98-210}$- DNA-free and DNA-bound forms relative to the equilibrated structure at 10, 20, 30, and 40 ns respectively.

Figure 4A.5 clearly shows basic residues Lys 168 and Lys179, in addition to Lys167 to be active in the β-sheets formation, which substantiates its role in DNA binding. The importance of the lysine residues has been stressed by Saijo, Takedachi, & Tanaka, [127], who saw those mutations of lysine residues K141A, T142A, K167A, and K179A reduced the binding affinity of XPA to RPA70 protein, affecting the NER mechanism as a whole. To get a better grip on the changes undergone by both XPA$_{98-210}$- with and without DNA, we superimposed the simulated structure of both systems. The acquired results are conveyed in **Figure 4A.6**. It distinctly highlights the appearances of antiparallel β-sheets and helices. DNA is also seen to be more twisted or kinked due to the XPA$_{98-210}$ bound to it.

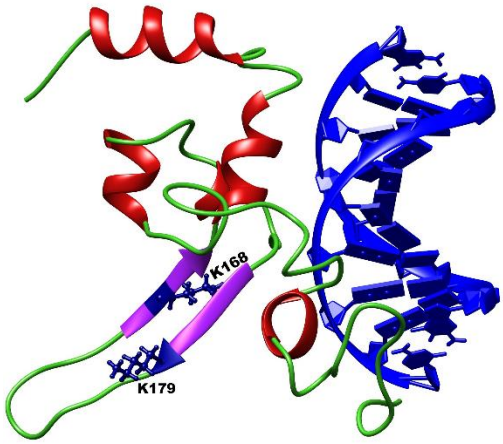

Figure 4A.5 Lysine residues K168 and K179 are depicted in a navy-blue ball and stick form present in the site for β-sheets formation.

Figure 4A.6 Superimposed Structure of XPA_{98-210}: DNA bound and DNA Free obtained at 40ns. XPA_{98-210} –Apo (dark-red) showed short stretches of β-sheets while XPA_{98-210}-DNA bound (XPA_{98-210} displayed in olive-green and DNA in deep-blue color) presented longer stretches of β-sheets in complex with DNA.

To summarize the overall conformational changes undergone by XPA_{98-210} in the presence and absence of DNA, we also determined the probable secondary structure and the score matrix correlating to each of their residues. **Figure 4A.7** explains the most probable secondary structural feature attained by the protein in both systems. **Figure 4A.7A** showed the chances of XPA_{98-210} in absence of DNA to form additional short stretches of 3_{10} α-helices which were seen in the snapshots presented above. XPA_{98-210}-DNA free system showed narrow range amino acids forming β-sheets in the residue index of 103-105, 110-112, 138-140, and 178-180 respectively, which is in agreement with **Table 4A.1**. **Figure 4A.7B** shows the probable secondary structure of XPA_{98-210} in presence of DNA, where it showed a greater possibility of the formation of antiparallel β-sheets from a wider range of residues: 163-184, which fall in the congruence of **Table 4A.1** as well highlighting the importance of the lysine residues. Both **Fig. 4A.7A** and **4A.7B** showed the appearances of antiparallel β-sheets in its N-terminal, 3 helices to its C-terminal, and some turns as described by Ikegami et al. [114, 115], and Buchko et al. [116]. The remaining portion of the protein seemed to have more or less a coherent structure irrespective of its DNA-bound or DNA-free status.

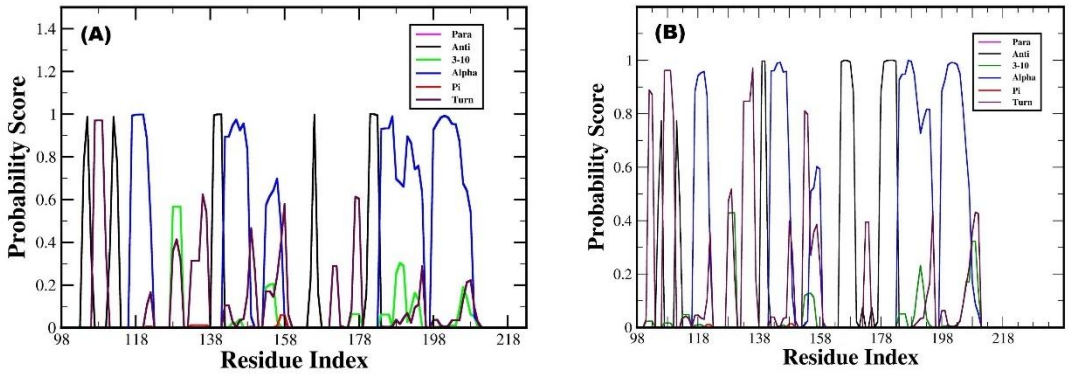

Figure 4A.7 Probable Secondary Structure Graph of XPA_{98-210}: (A) XPA_{98-210}- DNA Free; (B) XPA_{98-210}- DNA Bound system.

Though our obtained structure from docking doesn't exactly agree with the findings reported by Koch et al. [129], where β-sheets were directly involved in the interactions with the DNA. Yet, the commonality is that we saw DNA-bound XPA_{98-210} exhibit consistent β-sheets structure throughout the simulation, even though these β-sheets weren't in close contact with DNA. The discrepancies can be majorly credited to the docking pattern of our choice, which gave us the Zn-region in closeness to DNA rather than expected β-structures. Still, we were able to see the structural changes undergone by XPA in the presence and absence of DNA which agrees with the earlier findings [129]. In addition, we also noticed the bend in DNA upon contact with XPA at the lesion site. Studies by Koch et al. [129] saw the retention of the lesion inside the 13mer duplex and wasn't flipped outside. The 70º kink occurred in the major groove at the site of the Cisplatin-induced lesion. We too obtained bent DNA by 30 ns of MD production. Buchko et al. [368], extensively studied XPA-DNA interaction in which they published the full coverage of XPA by DNA takes place in $3' \rightarrow 5'$ direction rather than $5' \rightarrow 3'$ direction relative to helix 3 of XPA. The study also stressed the importance of the basic cleft of the loop-rich region in DNA binding. In our study, we can observe through the simulation snapshot the structural changes exhibited by all helices, especially the helix 3 which sees an extension of its helical form and again swaying back to its original status. This explains the basic reaction of XPA upon DNA exposure.

The indispensable contribution of XPA to the cause of NER functioning is very appreciative, whether as a scaffold protein [366], or in the early step of lesion

recognition [54, 127, 164]. XPA, encoded by the XPA gene consisting of 6 exons [428] is majorly linked to classical XP phenotype wherein even a partial mutation/loss of the protein would lead to NER complication and the result of which would cause cells to turn cancerous or malfunction [161-163]. Almost all the mutations identified in the XPA gene express mild to severe photosensitivity, extreme neurological aberrations, and mental retardations [54, 165, 174], with one exception so far reported by Takahashi et al. [168], where the patients showed no signs of neurological damage. Deep phenotyping methods have revealed heterogeneity among classical XP due to XPA ([362]. 56 somatic mutations in 121 cases were identified in XPA from the cBio Portal for Cancer Genomics [169]. Computational analysis detected 191 deleterious nsSNPs of XP disease including 7 nsSNPs in XPA [170]. Many novel mutations are still being reported worldwide on NER proteins that have led to irregularity in the repair process. This year's paper documented 4 novel mutations in one of the NER proteins-XPC in 2 unrelated patients [429], with the first patient having a splicing mutation between the border of exon 1-intron and a nonsense mutation c.958 C > T (Q320X) in exon 8, and second patient having deletions of 2 bp in exon 5 and exon 13 of XPC gene leading to frame-shift mutations. A mild variant of an XP-A phenotype too has been discovered where the patients shared ancestral roots in the same geographical area of the Indian subcontinent [160].

So, despite being probed and prodded for many years, XPA still has a very large void to fill due to its unexplored areas. And because XPA holds a strong role in NER (Fadda, 2016), its in-depth pursuit can yield remarkable milestones, whether it is studying its phenotypic outcomes, or venturing into its therapeutic prospects to depreciate the severity of XP phenotypic symptoms or elevate the Pt-based chemotherapeutic response in cancer patients.

4A.5 Conclusion

Our computational study of Xeroderma Pigmentosum Complementation Group A protein in the presence and absence of DNA provided us with considerable insights into the structural and conformational changes the protein undergoes while in contact with the damaged DNA. We studied $XPA_{98\text{-}210}$ in two states- DNA bound and DNA Free

State. From the comparative analysis of both the XPA_{98-210} systems, we found out that XPA_{98-210} in the presence of DNA designated more structural changes in contrast to XPA_{98-210} in its Apo form. In addition to the basic amino acid- lysine 167, 168, and 179, other hydrophobic residues- Phe164, Ile165, Val166, Asp177, Met178, Leu180, and Tyr181 were elaborately entailed in DNA binding and β-sheets formations. Acidic amino acids Asp and Glu too were observed in the DNA interaction which may be due to their ability to weaken the DNA hydrogen bonds, causing DNA instability and ultimately leading to the melting of the DNA wherein it aids TFIIH. We also saw the emergence of the additional helices in both systems. Helices were noticed to be in prominent proximity to the DNA as seen in the surface structure analysis and from the results provided by the DNAproDB server. Over the simulation period, the Zinc motif, the third helix, and β-sheets regions presented the most variations. Trajectory analysis of B-factor, SASA, RMSD, and Radius of Gyration too accorded with the fact that DNA bound XPA_{98-210} to have more changes than Apo XPA_{98-210}. The findings we have reported may not be able to explain the questions on XPA as a whole. Still, it provides a clear picture of how physical, conformational changes occur in XPA in the presence and absence of DNA in an explicit environment.

CHAPTER 4B

To study the binding mechanism of XPA homodimer with the DNA

CHAPTER 4B

To study the binding mechanism of XPA homodimer with the DNA

4B.1 Abstract

Xeroderma pigmentosum complementation group A (XPA) protein is integral to the nucleotide excision repair (NER) pathway as it regulates the recruitment of various NER proteins to the damaged site. Earlier XPA was studied as a monomer but in the recent past, it was reported to exist as a homodimer. Hence, in the present work, we have used a molecular dynamics (MD) approach to study the DNA-binding property of the XPA homodimer. The DNA-protein complex (DPC) was docked in HEX software and was found to be stable during the simulation. DNA-protein interaction (DPI) profiling using the DNAproDB server showed 28 intermolecular hydrogen bonds and 88 hydrophobic contacts. The binding free energy (BFE) analysis done using the molecular mechanics Poisson-Boltzmann (MM-PBSA) algorithm indicated this DPC to have a binding affinity of -62.33 kcal mol^{-1}. The per-residue energy decomposition (PRED) analysis revealed that the residues K213, K217, K221, K222, K224, K236, E225, R228, and R237 belonging to the C-terminal end of the XPA homodimer were involved in the DNA interaction. Furthermore, we also compared the properties of XPA homodimer in DNA bound (B) and unbound (U) states. We observed an increase in the number of interface residues and area, and in the size of accessible surface area (ASA) for the residues of XPA homodimer in the B state in comparison to its U state. The calculation of ASA values for B and U states showed that these changes were due to the partner attraction effect.

4B.2 Introduction

Xeroderma pigmentosum complementation group A (XPA) protein is one of the pivotal members of the nucleotide excision repair (NER) pathway that carries out the repairing of various DNA lesions in higher mammals. XPA has been known to function as both scaffold protein, and primary damage verifier in two sub-pathways of NER: global genome NER (GG-NER) and transcription-coupled NER (TC-NER). Since XPA takes part in both GG-NER and TC-NER, the mutations in XPA can cause an autosomal recessive disorder known as *Xeroderma pigmentosum* (XP), which is characterized by

severe sun-sensitivity, skin, and neurological defects, and is often associated with various types of skin cancers [48, 57, 72, 129, 372, 424, 430].

The damage recognition ability of XPA spans over numerous DNA lesions like cyclobutane pyrimidine dimers (CPD), 6-4 photoproducts (6-4PP), and intra-strand helix distorting cisplatin crosslinks induced by various environmental toxins, harmful sunrays, and antitumor agents. These lesions are repaired and removed by more than 30 different proteins that come together to form multiple protein-protein complexes (PPCs) and DNA-protein complexes (DPCs) during this sequential repair process. All the protein-protein interactions (PPIs) and DNA-protein interactions (DPIs) between various PPCs and DPCs occurring during the NER process are primarily coordinated by XPA, hence the name scaffold protein [48, 57, 72, 129, 142, 164, 222, 360, 368, 370, 372].

The structural composition of XPA has been characterized as 273 residues long protein, consisting of a central globular DNA-binding domain (DBD), which is flanked on either side by disordered N- and C-terminals. It has been observed that the XPA mutation leading to *XP* has been restricted mainly to its DBD region [57, 72, 114-116, 431]. The DBD region of XPA was earlier described between the aa98-219[114-116, 431] but was later redefined between the range aa98-239 after it was found out that the previous DBD lacked a considerable amount of positively charged residues to make a strong DPI with the DNA [126, 58, 176]. The studies conducted by Hilton et al. [126] and Sugitani et al. [58, 176] showed that the redefined DBD consisting mainly of lysine and arginine residues would establish a successful DPI between XPA and DNA. These studies also identified Lys168, Lys179, Lys221, Lys222, Lys223, Lys224, Glu225, and Lys236 as the crucial DNA-binding residues in XPA. In 2019, the 3D structure of redefined DBD of XPA was determined by two different groups using X-ray crystallography (PDB ID: 6J44) [118, 119] and electron microscopy (PDB ID: 6RO4) [74] with the resolution of 2.06 Å and 3.5 Å, respectively. The earlier 3D structures were nuclear magnetic resonance (NMR) structures of the previous DBD region of XPA with PDB IDs: 1XPA [114], and 1D4U [116] with residues described between aa98-208, and aa98-210, respectively.

Over the years, numerous studies have been done to understand PPIs and DPIs of XPA in the context of the NER pathway [48, 57, 72, 123, 129, 355, 357, 370, 372],

however, all these studies were conducted by considering XPA to exist only as a monomer, while newer findings suggested otherwise. Recent reports show that XPA exists in a homodimer form and that it binds to other NER members and DNA in a 2:1 ratio [130, 131]. Similar results were observed in yeast when XPA homolog radiation14 (Rad14) protein recognized 1,2-GG cisplatin lesion and N-(deoxyguanosine-8-yl)-2-acetylaminofluorene (AAF-dG) adduct in a homodimer form and had kinked the DNA by 70˚ [129]. Using computational techniques, we studied both DBDs of XPA (the previous a98-210 and redefined aa98-239) in DNA unbound states and observed that the homodimer model of redefined DBD of XPA was more stable and had higher binding affinity compared to the XPA homodimer model of previously described DBD [432]. One research team reported that XPA homodimer and proliferating cell nuclear antigen (PCNA), a scaffold protein seen only during the ligation step of NER, co-localized to the damaged DNA *in vivo* [83], suggesting that XPA may have a role in other steps of NER besides damage recognition.

Thus, with these recent developments in the study of XPA, where it has now been hypothesized that XPA carries out the function of both scaffold protein as well damage recognition as a homodimer protein and not as a monomer unit. Hence, it becomes important for us to investigate this DPI between the DBD of XPA homodimer and DNA in the context of NER. Therefore, in this study, we have attempted to study the dynamic behavior of redefined DBD of XPA homodimer in DNA bound state using the assisted model building with energy refinement (AMBER) force fields. Furthermore, we have also characterized the DPI and binding free energy (BFE) between XPA homodimer and DNA.

4B.3 Systems and methodology

4B.3.1 Molecular models

Since the 3D structure of redefined DBD of XPA was not determined at the time of commencement of this work, we have used the computationally generated structure of the homodimer model describing the redefined DBD of XPA (aa98-239) for this study. The details for structure modeling and validation of XPA monomer consisting of its redefined DBD (aa98-239), and the protocol for generating a homodimer model of

XPA's DBD as well as the basis on which an ideal homodimer model was selected to represent the DBD region of XPA has been described in our earlier work [400]. Even though the experimental structure of the redefined DBD of XPA (PDB ID: 6J44 and 6RO4) was reported recently [74, 118, 119], we have still carried out our work using the modeled structure [431] (**Figure 4B.1A**) for the following reasons:

I. 3D structures of the redefined XPA determined using X-ray crystallography and electron microscopy are in their monomeric state and do not represent its dimer form (**Figures 4B.1B,** and **4B.1C**).
II. The crystal structure has residues missing between aa169-176, and aa232-239 (**Figure 4B.1B**), while the electron microscope determined structure had residues missing from aa98-103 and aa238-239.
III. Using the University of California, San Francisco (USCF) Chimera software package v.1.12 [315], the root mean square deviation (RMSD) between the crystal structure and our modeled structure was found to be 1.088 Å (**Figure 4B.1D**), while the RMSD between our modeled structure and the electron microscopically determined the structure was found to be 1.129 Å (**Figure 4B.1D**).
IV. From our previous MD simulation study of XPA_{98-239} homodimer [431]. it is evident that our computationally generated structure is stable and can be applied to this study.

The equilibrated structure of the XPA homodimer containing its redefined DBD from our previous study [432] was taken as the starting structure for this study. The 3D structure of NMR determined 13mer DNA (PDB ID: 1LAI) was taken from PDB as the starting structure for our DNA [433].

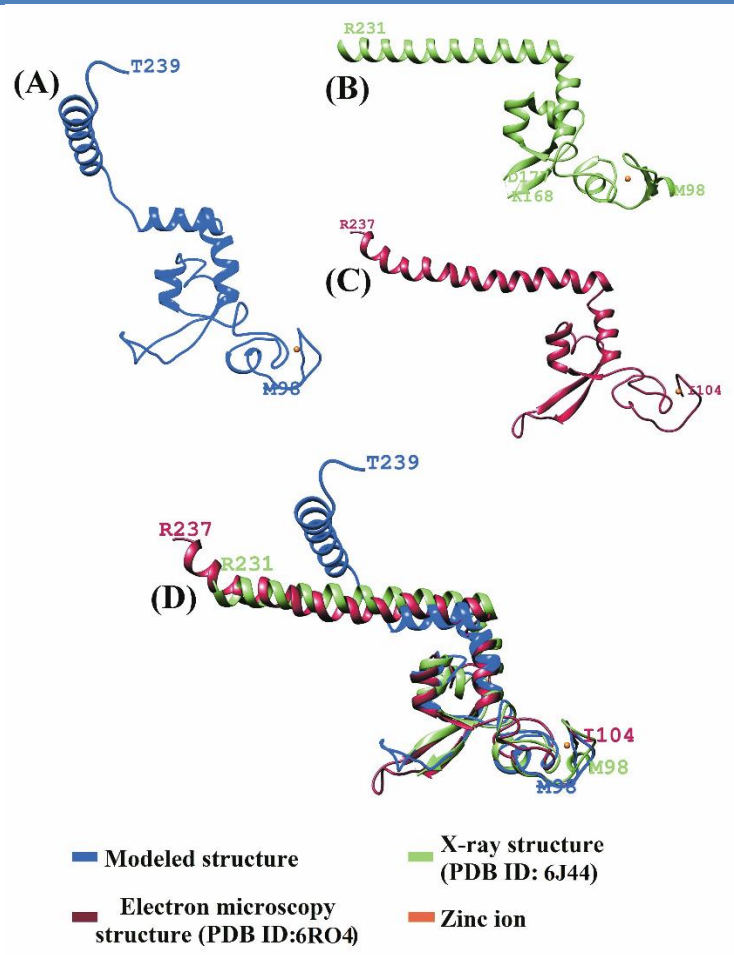

Figure 4B.1 Cartoon representation of the (A) computationally modeled structure, (B) crystal structure, and (C) electron microscopy determined structure of the redefined DBD of XPA; (D) computationally generated redefined DBD of XPA superimposed against its crystal structure and electron microscopically determined structure; RMSD between (A) and (B)=1.088 Å, and RMSD between (A) and (C)= 1.129 Å.

4B.3.2 Preparation of DNA-protein complex

For the construction of our DPC, XPA homodimer was considered as a receptor and DNA as a ligand. The molecular docking of these structures was done using the Hex 8.0.0 software [342, 418, 434], which is a rigid-body fast Fourier transform (FFT) based docking tool. This software uses spherical polar Fourier (SPF) correlations for generating the docked conformers. The parameters used for the docking were the same as given in our earlier work [400]. The five stages of docking employed by Hex software are as follows: SPF transforms → FFT Steric scan → FFT final search → energy refinement → total dockings. The resultant docked structures are ranked based on their

energy values (E-values), wherein the negative value indicates a better-docked conformer.

4B.3.3 Preparation of topology and coordinate file of the complex using Leap

The docked structure of the XPA homodimer-DNA complex having the highest negative energy value was used for LEaP preparation before simulation. The missing hydrogen atoms were added using the xleap module of the AMBER14 software package [242]. The topology and coordinate files for our DPC were generated using Amberff99 force fields, where the protein part was treated with ff99SB [282] and DNA with ff99bsc0 [419, 420]. Zinc AMBER force field (ZAFF) was used for obtaining the metal bonding parameters for Zn ions [434]. Our DPC was kept in the center of a truncated octahedron filled with TIP3P (transferable intermolecular potential with 3 points) water molecules [296]. The system was made neutral by adding charge-balancing counter-ions. We kept 10 Å as the minimum distance between the boundary of the water box and the atoms of the protein. The LEaP parameters optimized for our system have been given in **Table 4B.1**.

Table 4B.1. The system parameterization using Xleap.

Parameters	XPA homodimer-DNA complex
AMBER force field	ff99SB for XPA homodimer, ff99bsc0 for DNA, ZAFF for Zn ions
Initial charge	+24 for DNA, -6 for XPA homodimer
Counter ions added	24 Cl^- ions for DNA, 6 Na^+ ions for XPA homodimer
Final charge	0
Water residues added	20959
Total mass	419804.40 amu
Density	0.903 g cc^{-1}
Volume	771850.83 A^3
Solute vdW bounding box	63.85 86.39 76.51
Solvent unit box	18.77 18.77 18.77
The total bounding box for atom centers	114.57 114.57 114.57
Total vdW box size	88.17 74.39 80.26

4B.3.4 Molecular dynamics simulation

LEaP prepared system consisting of our DPC was subjected to two-step energy minimization with 5000 steps of steepest descent (SD) method followed by 4000 steps

of conjugate gradient (CG) method. The first minimization was done by imposing restraints over solute while no such restraints were added during the second minimization. The energy minimized system was next gradually heated from 1-300 K for 1 ns at NVT conditions after which equilibration of density (0.5 ns) was done with weak restraints, followed by 5 ns of constant pressure equilibration at 300 K at NPT conditions. Finally, we performed a 30 ns MD production run using the Particle Mesh Ewald (PME) algorithm [388, 435] with time step integration of 2 fs. The SHAKE algorithm [292] was used for constraining all hydrogen bonds present in the system. Berendsen thermostat (heat bath=0.5 ps and pressure relaxation= 0.2 ps) was used for maintaining the constant temperature [294], while the time constant for temperature coupling was kept at 2 ps (time constant tautp = 2.0). The cut-off distance for all the Lennard-Jones (LJ) interactions were kept at 8 Å, while the long-range electrostatic interactions were treated with the PME method. The trajectory snapshots were recorded every 500 ps for further analysis.

4B.3.5 MD analysis

All MD trajectory snapshots were analyzed using the CPPTRAJ (PTRAJ in C++) module [316] of the AmberTools 14. The RMSD, radius of gyration (R_g), root mean square fluctuation (RMSF), solvent-accessible surface area (SASA), and intermolecular hydrogen bonding were analyzed from the MD trajectory for our DPC. The same protocol was followed for the trajectory analysis of individual units of DPC, which are DNA, monomer 1, and 2 of the XPA homodimer complex. For the hydrogen bond analyses of our DPC, the cut-off for bond length, and the bond angle between the atoms of hydrogen donor (HD) and hydrogen acceptor (HA) was kept at 3 Å and 135°, respectively. All the graphs for RMSDs, RMSFs, SASAs, and hydrogen bond analyses were plotted using Xmgrace software [436]. Visualization of the MD snapshots and the generation of molecular graphics and images were done using USCF Chimera v.1.12.

4B.3.6 Intermolecular interaction study of XPA homodimer-DNA complex

For the DPI study, we extracted the lowest energy conformer for our DPC from the highly populated cluster using the RMSD clustering algorithm. This conformer was

submitted to DNAproDB server beta v.2 (http://dnaprodb.usc.edu/) [350] for determination of the DPI between XPA homodimer and DNA. DNAproDB server scores DPI based on the presence of at least one interaction among three, which are (i) a buried solvent-accessible surface area (BASA) greater than 0, (ii) one hydrogen bond, or (ii) one van der Waals (vdW) interaction. We also submitted the same structure to the PDBsum server (http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html) [349] to elucidate the PPI between the monomers of this XPA homodimer.

4B.3.7 Free energy analysis of DNA-protein complex

MMPBSA.py script that comes inbuilt with AMBER14 software was used for calculating the binding free energy (BFE) using the molecular mechanics Poisson-Boltzmann surface area continuum solvation (MM-PBSA) algorithm [321, 373, 392]. This particular method is based on the second law of thermodynamics, which calculates BFE for each conformation using the formula:

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

Each ΔG is calculated using the following formula:

$$\Delta G = \Delta E_{int} + \Delta E_{vdW} + \Delta E_{ele} + \Delta G_{PB} + \Delta G_{surf} - T\Delta S \quad (2)$$

Here, the first three terms are components of classical molecular mechanics energy (ΔE_{MM}) representing the internal energy (ΔE_{int}), van der Waals forces (ΔE_{vdW}), and electrostatic energy (ΔE_{ele}). Next two terms, ΔG_{PB} and ΔG_{surf} are polar and non-polar measures of solvation free energy (ΔG_{solv}). The polar energy (ΔG_{PB}) was calculated using the PB solver algorithm [321], while non-polar energy was estimated using LCPO (linear combinations of pairwise overlaps) SASA formula [437]:

$$\Delta G_{surf} = \gamma \times SASA + \beta \quad (3)$$

Where, corresponding values of γ and β were set to 0.00542 kcal mol^{-1} $Å^{-2}$ and 0.92 kcal mol^{-1} [438], respectively. The dielectric constants for solute and solvent were set to 1 and 80. The grid space of 0.5 Å and probe radius of 1.4 Å was used for the calculation.

The last term (TΔS) from the Eq. (2) signifies entropy of the system, and was not considered for our study; hence these energy values represent only the relative binding free energy. Further, the MM-PBSA method was also used to perform per-residue energy decomposition (PRED) analysis on the interface residues of XPA homodimers involved in the DPI with DNA. The decomposition energy of each residue was calculated using this formula:

$$\Delta G = \Delta E_{int} + \Delta E_{vdW} + \Delta E_{ele} + \Delta G_{PB} + \Delta G_{surf} \quad (4)$$

4B.4 Results and discussion

4B.4.1 Molecular docking results

Molecular docking of XPA homodimer with DNA in Hex software gave a total of 100 conformations ranked based on their E-values, out of which the top five conformations have been shown in **Figure 4B.2**. Among these conformers, we chose the top-ranked docked conformer having the lowest E-value (-898.4 kcal mol^{-1}) as our DPC representative for the XPA homodimer-DNA complex. Upon the visualization of this structure in UCSF Chimera, we observed that DNA was in close contact with the fourth helix (aa219-239) of both XPA monomers. This region (aa219-239) has been demonstrated to have a high affinity toward DNA compared to the other regions of XPA 58].

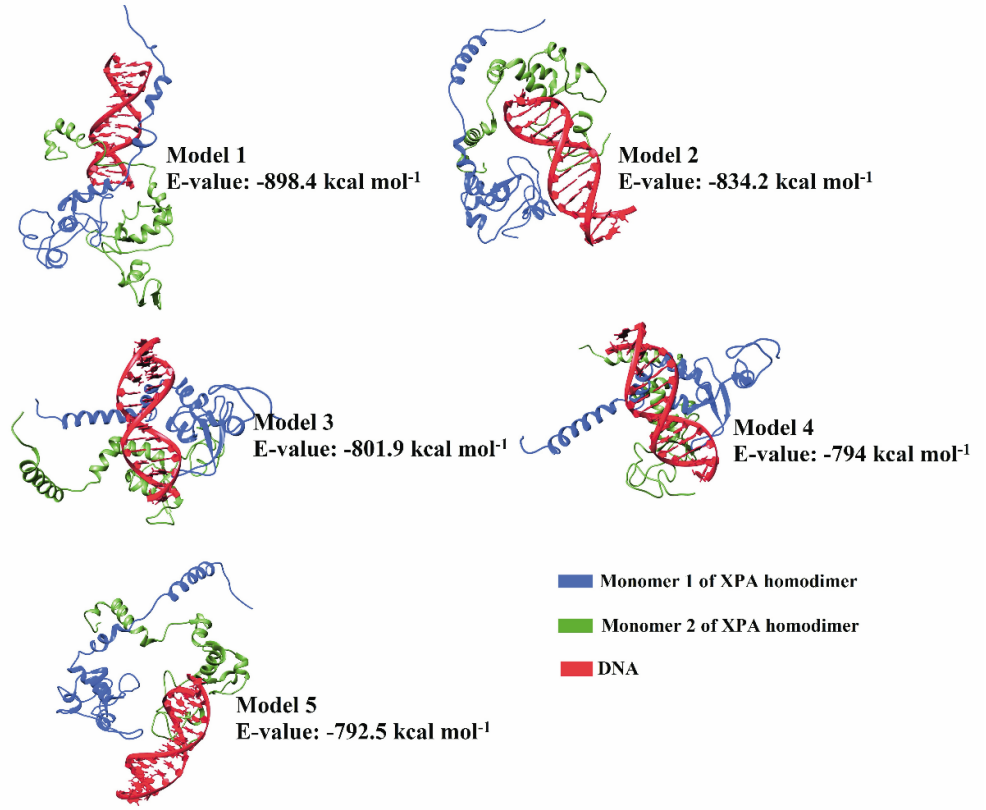

Figure 4B.2 Top five docked conformations of XPA homodimer-DNA complex.

4B.4.2 Molecular dynamics and simulation study of XPA homodimer-DNA complex

The analyses of temperature, pressure, volume, density, and energy, plots showed that our system had successfully reached equilibrium (**Figure 4B.3A-E**). Next, our DPC consisting of XPA homodimer and DNA was simulated for 30 ns in an explicit solvent and the resulting MD trajectories obtained at the end of the simulation have been analyzed accordingly.

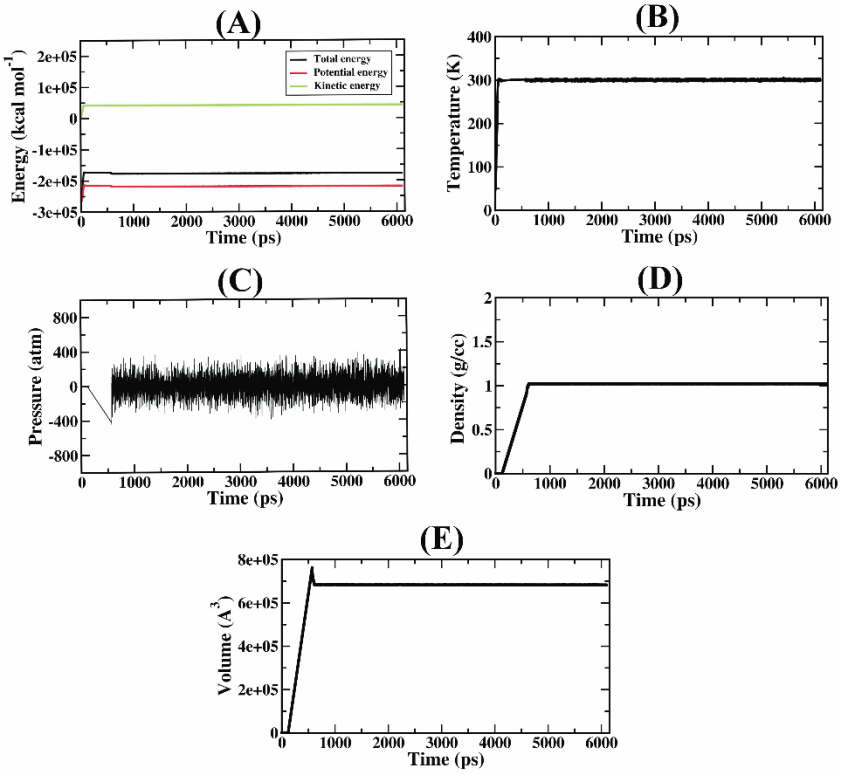

Figure 4B.3. Equilibration plots of (A) energy, (B) temperature, (C) pressure, (D) density, and (E) volume plots for XPA homodimer-DNA complex as a function of time.

4B.4.3 Stability analysis of the XPA homodimer-DNA complex

To assess the structural stability of our system, we examined the all-atoms RMSDs for our complex. The trajectory of our DPC showed convergence with an average RMSD value of 12.69 Å (**Figure 4B.4A**). The RMSD analyses of individual units of our DPC revealed that monomer 2 and DNA were stable throughout the simulation as well with the average RMSD of 5.62 Å and 2.22 Å, respectively. Monomer 1 of XPA homodimer, on the other hand, showed a slight rise in the graph indicating a change in its conformation, after which it attained stability from 15 ns with an average value of 11.7 Å. R_g analysis was done next to study the mass-weighted spatial distribution of the XPA homodimer-DNA complex from the center of mass (**Figure 4B.4B**). The average R_g value for our DPC was found to be 27.75 Å, while R_g values for XPA monomer 1, monomer 2, and DNA were 28.93 Å, 22.24 Å, and 14.52 Å, respectively.

Since SASA is directly associated with the hydrophobic contacts occurring between the biomolecule and solvent and helps in mapping out the accessible surface

area required for probable PPI or DPI interaction, we conducted SASA analysis for our DPC using the rolling ball algorithm.[55] The increase or decrease of SASA value is directly proportional to the exposure of a particular residue to the solvent and can affect the structure of a protein [439]. From **Figure 4B.4C**, we can see that the SASA for the XPA homodimer-DNA complex was found to be high at 25500 Å^2. The SASA values of XPA monomer 1, XPA monomer 2, and DNA were found to be 11600 Å^2, 9800 Å^2, and 4100 Å^2, respectively.

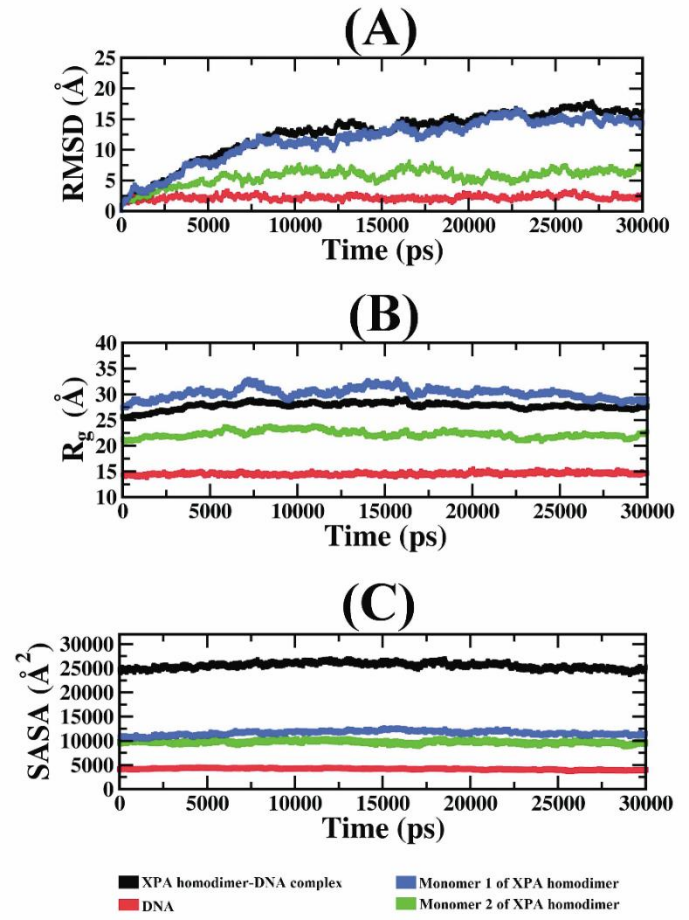

Figure 4B.4. (A) RMSD, (B) R_g, and (C) SASA plot for XPA homodimer-DNA complex as a function of time.

4B.4.3.1 Comparing the RMSD, R_g, and SASA of XPA homodimer in DNA bound and DNA unbound states.

A biomolecule, especially proteins can undergo various changes when it shifts from apo to the complexed state [439]. Hence, to get a better understanding of those changes

occurring in XPA homodimer in DNA bound and DNA unbound states, we have compared the data of DPC from this study with PPC of XPA homodimer in DNA unbound from our earlier work [431]. For convenience, we have used the letter 'B' and 'U' to designate XPA homodimer in DNA bound and unbound states, respectively. We have summarized the average properties like RMSD, R_g, and SASA of both B and U states of XPA homodimer in **Table 2.** Upon the comparison of these features for XPA homodimer in B and U states, we observed that RMSD values for all the components had decreased while R_g had increased in the B state of XPA homodimer as compared to the U state. SASA values were also observed to have increased for all the components of XPA homodimer in B form when compared to its U form. The differences in the RMSDs of proteins in B and U states have been credited to the changes in backbone structure and interface area of DPCs [439].

Table 4B.2. Average properties of XPA homodimer in B and U states.

Average properties	XPA homodimer in B state				XPA homodimer in U state*		
	XPA homodimer-DNA complex	DNA	Monomer 1	Monomer 2	XPA homodimer	Monomer 1	Monomer 2
RMSD	12.69 Å	2.22 Å	11.7 Å	5.62 Å	13 Å	12 Å	12 Å
R_g	27.75 Å	14.52 Å	28.93 Å	22.24 Å	27 Å	26 Å	22 Å
SASA	25500 Å^2	4100 Å^2	11600 Å^2	9800 Å^2	22500 Å^2	11000 Å^2	10500 Å^2

**The computational data for the XPA homodimer in the U state has been taken from our earlier work.*[29]*

4B.4.4 Structural flexibility of DPC

To understand the residue fluctuation of each monomer, we analyzed their RMSF plots. The average RMSF value for monomer 1 and monomer 2 of XPA homodimer was found to be 4.86 Å and 3.66 Å, respectively. We observed residual fluctuations for both monomers to be restricted mainly at their disordered N- and C-terminal regions as compared to the other regions (**Figure 4B.5A** and **4B.5B**). We also observed fluctuations at residues location of aa217-239, which formed the fourth helix in the protein, and as per the findings reported by Lian et al. [118] has been known to undergo a conformational change (bending of the helix) to accommodate DNA substrates.

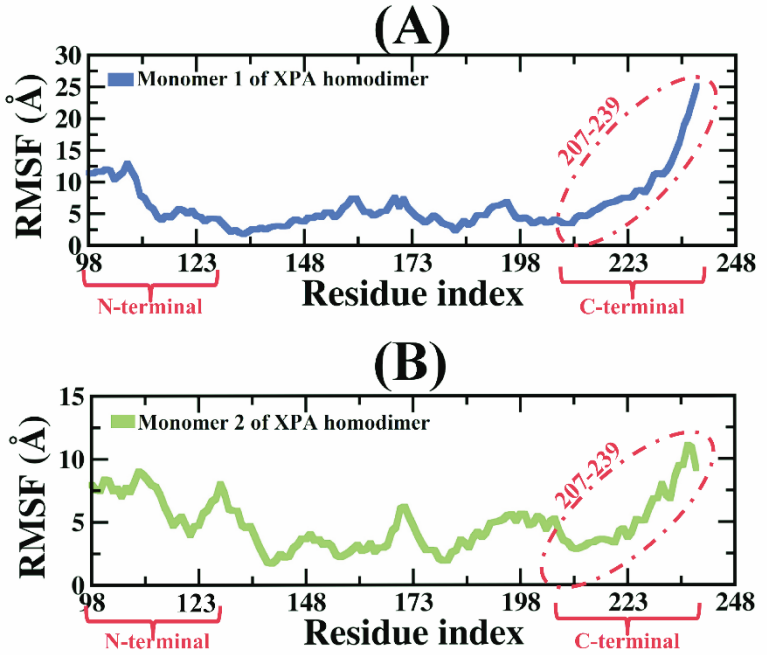

Figure 4B.5. RMSF plots for (A) monomer 1 and (B) monomer 2 of XPA homodimer.

4B.4.5 DNA-protein interaction study of XPA homodimer and DNA

The lowest energy conformer isolated from trajectories was found to be tightly bound (**Figure 4B.6C**). From the cartoon structure of our DPC (**Figure 4B.6D)** as well as from our conformational snapshots (**Figure 4B.7**), we can see that DNA was sandwiched between the two XPA molecules which is similar to Rad14 homodimer's interaction with the DNA, where the monomers of Rad14 had held 1, 2-GG cisplatin and AAF-dG lesion from either side [120]. Here, we also see that DNA was close to the lateral end of XPA (aa210-239), which was newly included in the redefined DBD of XPA. The residues aa210-239 which form a third and fourth helix in the DBD region of XPA along with a basic-loop rich cleft between the two helices were observed to be in close contact with the DNA. Similar results were seen earlier with the DPI between XPA monomer and 9-nt ssDNA, where they found C-terminal helices with the basic cleft to be present at the DNA interaction site [367]. These residues belonging to the C-terminal end have been found to increase the DNA binding affinity of XPA by manifolds and aid in establishing a successful DPI between XPA and DNA [58, 176].

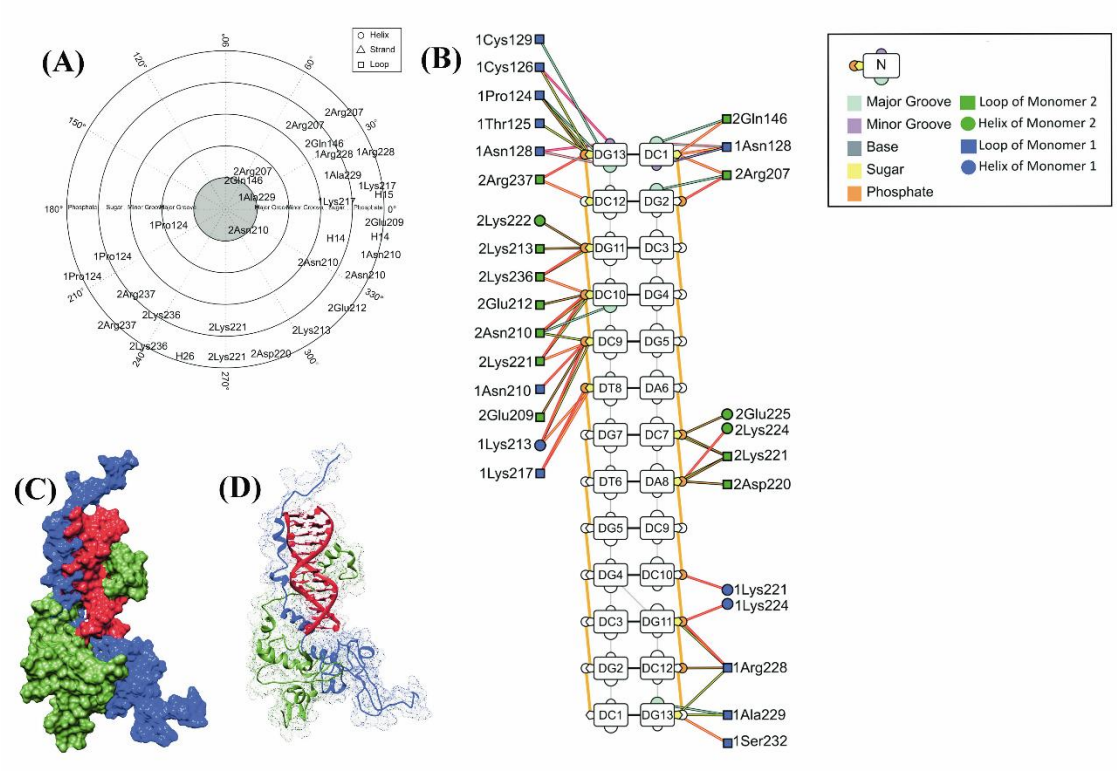

Figure 4B.6. (A) Polar contact map, (B) nucleotide-residue contact map for XPA homodimer-DNA complex, where the numbers 1, and 2 prefixed before residues indicate monomer 1 and 2, respectively. (C) Surface and (D) cartoon representation for the lowest energy conformer of XPA homodimer-DNA complex.

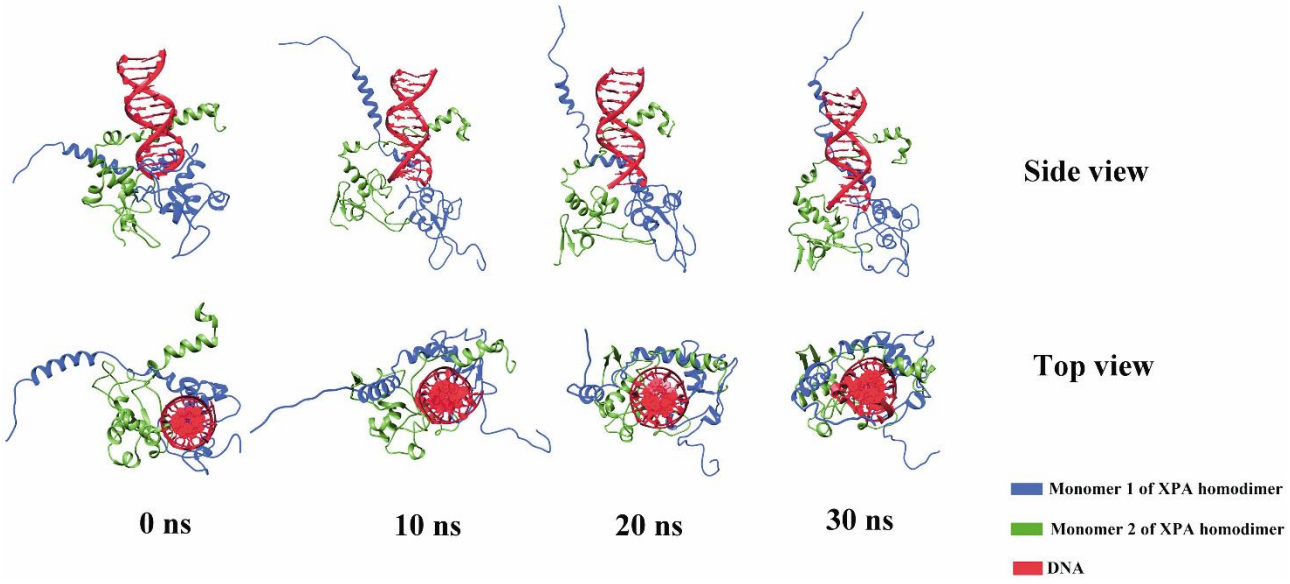

Figure 4B.7. Conformational snapshots of XPA homodimer-DNA complex at different time intervals.

The helical contact map for our DPC (**Figure 4B.6A**) generated using the DNAproDB server shows that the majority of the DPI has taken place through the backbone atoms of DNA (phosphate and sugar) while fewer interactions were with the DNA grooves. This result supports the findings observed by Luscombe and team,[57] where they demonstrated that proteins mostly conduct a DPI with the DNA via backbone interactions, and rarely through the bases. The nucleotide-residue interaction analysis for DPC indicated that both monomers of the XPA homodimer were bound to the DNA (**Figure 4B.6B**) with a significant number of hydrogen bonds and vdW interactions. The hydrogen bond analysis of our DPC from the simulation data also stated the same, where monomer 1 of XPA homodimer had formed 15 hydrogen bonds with DNA (**Figure 4B.8A**), while 13 hydrogen bonds were formed between monomer 2 of XPA homodimer and DNA (**Figure 4B.8B**). The hydrogen bonds were mediated by residues Cys126, Asn128, Asn210, Lys213, Lys217, Lys221, Lys224, Arg228, and Ser232 for monomer 1 (**Table 4B.3**), and residues Gln146, Arg207, Asn210, Lys221, Lys224, Lys236, and Arg237 for monomer 2 of XPA homodimer (**Table 4B.4**). A detailed description of this DPI has been provided in **Tables 4B.5** and **4B.6**. It can also be noted that the majority of the residues involved in DPI with DNA and PPI within the monomers have been known to cause extreme cases of XP phenotypes, making them equally important for DNA interaction as well for the proper functioning of the NER pathway [72].

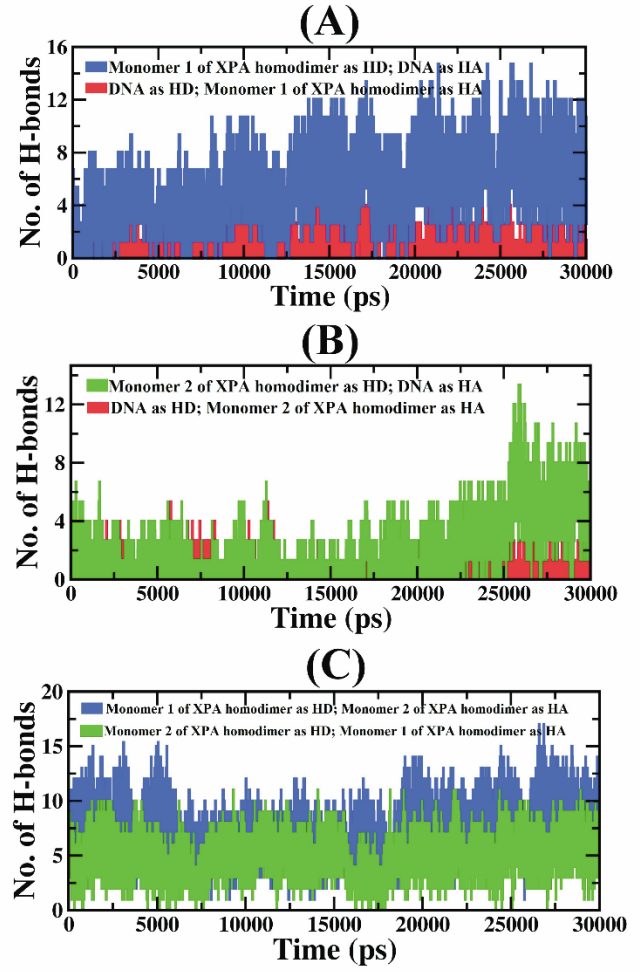

Figure 4B.8. Intermolecular hydrogen bond analyses between (A) monomer 1 and DNA, (B) monomer 2 and DNA, and (C) monomer 1 and monomer 2 of XPA homodimer as a function of time.

Table 4B.3. Intermolecular interactions between DNA and monomer 1 of XPA homodimer.

XPA	Residues	SSE	DNA strand	Nucleotide	DNA moiety	HB count	vdW count	BASA (Å^2)
Monomer 1	Pro124	Loop	Strand 1	DG13	Major groove	0	0	2.39
					Phosphate	0	1	33.43
					Sugar	0	2	33.40
	Thr125	Loop		DG13	Sugar	0	3	15.58
	Cys126	Loop		DG13	Minor groove	1	1	13.90
					Sugar	0	2	20.39
	Asn128	Loop		DG13	Major groove	1	1	7.7
					Minor groove	1	4	10.35
	Cys129	Loop		DG13	Major	0	0	26.29

					groove			
	Asn210	Loop		DC9	Phosphate	1	3	22.59
	Lys213	Helix		DC9	Sugar	0	0	6.37
					Phosphate	1	1	18.82
				DT8	Sugar	1	1	16.89
	Lys217	Loop		DT8	Phosphate	1	3	25.32
					Sugar	1	2	20.43
	Asn128	Loop	Strand 2	DC1	Major groove	1	1	5.20
					Minor groove	0	3	5.77
					Sugar	1	1	23.28
	Lys221	Helix		DC10	Phosphate	1	3	64.08
	Lys224	Loop		DG11	Phosphate	1	1	30.91
	Arg228	Loop		DG11	Phosphate	2	4	30.77
					Sugar	0	1	14.35
				DC12	Phosphate	0	0	26.32
				DG13	Sugar	0	0	49.02
	Ala229	Loop		DG13	Major groove	0	0	3.85
					Sugar	0	0	15.86
	Ser232	Loop		DG13	Sugar	1	4	51.83

XPA: Xeroderma pigmentosum complementation group A; SSE:secondary structure element; vdW: van der Waal interactions; HB: hydrogen bonds; BASA: Buried solvent accessible surface areas; DC: deoxycytidine; DG: deoxyguanosine; DT: deoxythymidine

Table 4B.4. Intermolecular interactions between monomer 2 of XPA homodimer and DNA.

XPA	Residues	SSE	DNA strand	Nucleotide	DNA moiety	HB count	vdW count	BASA (Å^2)
Monomer 2	Glu209	Loop	Strand 1	DC9	Phosphate	0	1	9.39
	Asn210	Loop		DC10	Sugar	0	0	11.62
					Phosphate	1	3	19.78
					Major groove	0	0	12.92
				DC9	Sugar	0	1	12.72
	Glu212	Loop		DC10	Phosphate	0	0	12.9
	Lys213	Loop		DG11	Phosphate	0	0	40.23
	Lys221	Loop		DC9	Sugar	1	1	22.43
				DC10	Phosphate	1	4	27.60
					Sugar	0	0	11.48
	Lys222	Helix		DG11	Phosphate	0	0	19.26
	Lys236	Loop		DC10	Sugar	1	1	16.06
				DG11	Phosphate	1	4	22.15
					Sugar	0	0	18.07
	Arg237	Loop		DC12	Sugar	1	2	41.04
				DG13	Phosphate	1	1	26.64
	Gln146	Loop	Strand 2	DC1	Sugar	2	7	44.15
					Major groove	0	3	31.14

	Arg207	Loop		DG2	Phosphate	2	4	34.46
					Major groove	0	0	10.45
				DC1	Sugar	1	4	29.25
	Asp220	Loop		DA8	Phosphate	0	1	6.357
	Lys221	Loop		DA8	Sugar	0	0	18.14
					Phosphate	0	0	12.37
				DC7	Sugar	0	2	40.57
					Phosphate	0	0	10.89
	Lys224	Loop		DA8	Phosphate	1	4	35.89
	Glu225	Loop		DC7	Phosphate	0	3	49.57

XPA: Xeroderma pigmentosum complementation group A; SSE: secondary structure element; vdW: van der Waal interactions; HB: hydrogen bonds; BASA: Buried solvent accessible surface areas; DT: deoxythymidine; DC: deoxycytidine; DG: deoxyguanosine

Table 4B.5. Intermolecular interactions between monomer 1 of XPA homodimer and DNA.

Table 4B.5A. Intermolecular hydrogen bonds.

Atom name	Residues	Protein chain		Atom name	Nucleotides	DNA strand	Distance (Å)
SG	CYS126	1	<-->	N3	DG13	A	3.39
OD1	ASN128	1	<-->	N1	DG13	A	3.18
OD1	ASN128	1	<-->	N2	DG13	A	3.41
ND2	ASN210	1	<-->	OP1	DC9	A	3.25
NZ	LYS213	1	<-->	O3'	DT8	A	2.91
NZ	LYS213	1	<-->	OP1	DC9	A	2.73
NZ	LYS217	1	<-->	OP1	DT8	A	2.74
NZ	LYS217	1	<-->	O5	DT8	A	3.45
ND2	ASN128	1	<-->	O4'	DC1	B	3.07
OD1	ASN128	1	<-->	N3	DC1	B	3.40
NZ	LYS221	1	<-->	OP2	DC10	B	2.72
NZ	LYS224	1	<-->	OP1	DG11	B	2.86
NH1	ARG228	1	<-->	OP2	DG11	B	2.83
NH2	ARG228	1	<-->	OP2	DG11	B	2.82
OG	SER232	1	<-->	O3'	DG13	B	2.79

Table 4B.5B. Non-bonded contacts.

Atom name	Residues	Protein chain		Atom name	Nucleotides	DNA strand	Distance (Å)
O	PRO124	1	<-->	C2'	DG13	A	3.68
CG	PRO124	1	<-->	OP2	DG13	A	3.40
CG	PRO124	1	<-->	O5'	DG13	A	3.34
CA	THR125	1	<-->	C2'	DG13	A	3.73
C	THR125	1	<-->	O3'	DG13	A	3.45
O	THR125	1	<-->	O3'	DG13	A	3.45
N	CYS126	1	<-->	O3'	DG13	A	3.61
CB	CYS126	1	<-->	O3'	DG13	A	3.49
SG	CYS126	1	<-->	N3	DG13	A	3.39

OD1	ASN128	1	<-->	N1	DG13	A	3.18
OD1	ASN128	1	<-->	C2	DG13	A	3.76
CB	ASN128	1	<-->	N2	DG13	A	3.80
CG	ASN128	1	<-->	N2	DG13	A	3.77
OD1	ASN128	1	<-->	N2	DG13	A	3.41
ND2	ASN210	1	<-->	P	DC9	A	3.89
OD1	ASN210	1	<-->	OP1	DC9	A	3.84
ND2	ASN210	1	<-->	OP2	DC9	A	3.53
NZ	LYS213	1	<-->	O3'	DT8	A	2.91
NZ	LYS213	1	<-->	P	DC9	A	3.49
NZ	LYS217	1	<-->	P	DT8	A	3.73
CE	LYS217	1	<-->	OP1	DT8	A	3.76
NZ	LYS217	1	<-->	OP1	DT8	A	2.74
CE	LYS217	1	<-->	O5'	DT8	A	3.58
NZ	LYS217	1	<-->	O5'	DT8	A	3.45
ND2	ASN128	1	<-->	O5'	DC1	B	3.73
OD1	ASN128	1	<-->	N1	DC1	B	3.80
OD1	ASN128	1	<-->	N3	DC1	B	3.40
OD1	ASN128	1	<-->	C2	DC1	B	3.19
CG	ASN128	1	<-->	O2	DC1	B	3.56
OD1	ASN128	1	<-->	O2	DC1	B	3.08
CD	LYS221	1	<-->	OP2	DC10	B	3.79
CE	LYS221	1	<-->	OP2	DC10	B	3.75
NZ	LYS221	1	<-->	OP2	DC10	B	2.72
NZ	LYS224	1	<-->	OP1	DG11	B	2.86
CZ	ARG228	1	<-->	OP2	DG11	B	3.25
NH1	ARG228	1	<-->	P	DG11	B	3.79
NH1	ARG228	1	<-->	OP2	DG11	B	2.83
NH2	ARG228	1	<-->	OP2	DG11	B	2.82
NH1	ARG228	1	<-->	O5'	DG11	B	3.51
OG	SER232	1	<-->	C1'	DG13	B	3.77
OG	SER232	1	<-->	C3'	DG13	B	3.70
CB	SER232	1	<-->	O3'	DG13	B	2.97
OG	SER232	1	<-->	O3'	DG13	B	2.79

Table 4B.6. Intermolecular interactions between monomer 2 of XPA homodimer and DNA.

Table 4B.6A. Intermolecular hydrogen bonds.

Atom name	Residues	Protein chain		Atom name	Nucleotides	DNA strand	Distance (Å)

ND2	ASN210	2	<-->	OP2	DC10	A	2.72
NZ	LYS221	2	<-->	O3'	DC9	A	3.08
NZ	LYS221	2	<-->	OP1	DC10	A	3.00
NZ	LYS236	2	<-->	O3'	DC10	A	3.28
NZ	LYS236	2	<-->	OP1	DG11	A	2.79
NH2	ARG237	2	<-->	O3'	DC12	A	3.15
NH2	ARG237	2	<-->	OP1	DG13	A	3.48
NE2	GLN146	2	<-->	O4'	DC1	B	3.06
NE2	GLN146	2	<-->	O5'	DC1	B	3.23
NH1	ARG207	2	<-->	O3'	DC1	B	3.28
NH2	ARG207	2	<-->	OP2	DG2	B	2.80
NH1	ARG207	2	<-->	OP2	DG2	B	3.23
NZ	LYS224	2	<-->	OP1	DA8	B	2.76

Table 4B.6B. Non-bonded contacts.

Atom name	Residues	Protein chain		Atom name	Nucleotides	DNA strand	Distance (Å)
O	GLU209	2	<-->	OP2	DC9	A	3.73
CB	ASN210	2	<-->	C2'	DC9	A	3.77
CB	ASN210	2	<-->	OP2	DC10	A	3.34
CG	ASN210	2	<-->	OP2	DC10	A	3.35
ND2	ASN210	2	<-->	OP2	DC10	A	2.72
NZ	LYS221	2	<-->	O3'	DC9	A	3.08
NZ	LYS221	2	<-->	P	DC10	A	3.66
CD	LYS221	2	<-->	OP1	DC10	A	3.76
CE	LYS221	2	<-->	OP1	DC10	A	3.80
NZ	LYS221	2	<-->	OP1	DC10	A	3.00
NZ	LYS236	2	<-->	O3'	DC10	A	3.28
NZ	LYS236	2	<-->	P	DG11	A	3.66
CD	LYS236	2	<-->	OP1	DG11	A	3.72
CE	LYS236	2	<-->	OP1	DG11	A	3.68
NZ	LYS236	2	<-->	OP1	DG11	A	2.79
NH2	ARG237	2	<-->	C4'	DC12	A	3.64
NH2	ARG237	2	<-->	O3'	DC12	A	3.15
NH2	ARG237	2	<-->	OP1	DG13	A	3.48
NE2	GLN146	2	<-->	O5'	DC1	B	3.23
OE1	GLN146	2	<-->	C5'	DC1	B	3.76
NE2	GLN146	2	<-->	C5'	DC1	B	3.37
NE2	GLN146	2	<-->	C4'	DC1	B	3.79
CD	GLN146	2	<-->	O4'	DC1	B	3.72
NE2	GLN146	2	<-->	O4'	DC1	B	3.06
CD	GLN146	2	<-->	C6	DC1	B	3.62
OE1	GLN146	2	<-->	C6	DC1	B	3.15
NE2	GLN146	2	<-->	C6	DC1	B	3.88
OE1	GLN146	2	<-->	C2'	DC1	B	3.62
CZ	ARG207	2	<-->	O3'	DC1	B	3.88
NH1	ARG207	2	<-->	O3'	DC1	B	3.28
NH2	ARG207	2	<-->	C3'	DC1	B	3.87
NH2	ARG207	2	<-->	O3'	DC1	B	3.53

NH2	ARG207	2	<-->	P	DG2	B	3.74
CZ	ARG207	2	<-->	OP2	DG2	B	3.45
NH1	ARG207	2	<-->	OP2	DG2	B	3.23
NH2	ARG207	2	<-->	OP2	DG2	B	2.80
OD2	ASP220	2	<-->	OP1	DA8	B	3.60
CG	LYS221	2	<-->	O3'	DC7	B	3.78
CE	LYS221	2	<-->	O3'	DC7	B	3.40
NZ	LYS224	2	<-->	P	DA8	B	3.76
CD	LYS224	2	<-->	OP2	DA8	B	3.40
CE	LYS224	2	<-->	OP2	DA8	B	3.83
NZ	LYS224	2	<-->	OP2	DA8	B	3.79
CG	GLU225	2	<-->	OP1	DC7	B	3.48
CD	GLU225	2	<-->	OP1	DC7	B	3.72
OE1	GLU225	2	<-->	OP2	DC7	B	3.81

The residues like lysine, arginine, glutamine, asparagine, threonine, and serine, in particular, have been known to possess high DNA binding properties by forming multiple hydrogen bonds and vdW interactions with the DNA backbone.[57-59] From **Tables 4B.5** and **4B.6**, we can also see that many of the residues involved in DPI showed multiple intermolecular interactions with DNA. As seen from our results above, all the DNA binding residues identified by Hilton et al. [126] and Sugitani et al. [58, 176] were found to have participated in the DPI, except for K168 and K179, which may be due to the shorter stretch of DNA taken for this study. The residues K168 and K179 may be seen at the DNA interaction site if a long stretch of DNA is used for the XPA-DNA interaction study.

In addition to DPI profiling, we have also studied the PPI between XPA monomers. The interface statistics for this PPI have been given in **Table 4B.7**. The residues participating in the PPI have been seen in **Figure 4B.8** along with the detailed information on each bond formation in **Table 4B.8**, while the hydrogen bond analysis between two XPA monomers has been given in **Figure 4B.9**. An observation that we have made was that despite an increase in the number of PPI between the two XPA monomers, there was no disulfide bond formed consistent with the previous findings [130, 431].

Table 4B.7. Interface statistics for XPA homodimer.

Protein	No. of interface	Interface area (Å^2)	No. of salt bridges	No. of hydrogen	No. of non-

	residues			bonds	bonded contacts
Monomer 1 of XPA homodimer	17	1093	5	14	112
Monomer 2 of XPA homodimer	18	1102			

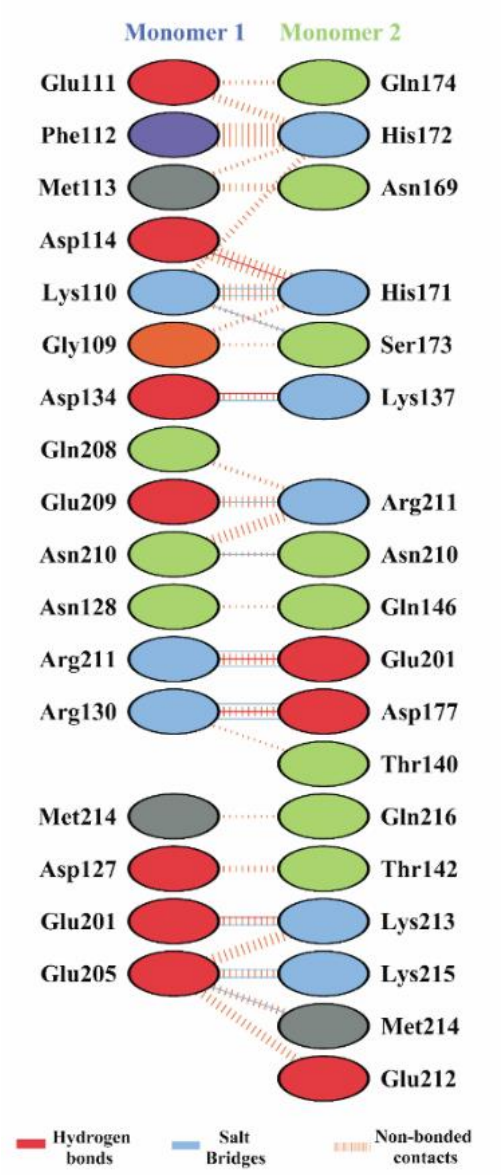

Figure 4B.9. PPI between monomer 1 and monomer 2 of XPA homodimer.

Table 4B.8. Intermolecular interactions between monomer 1 and 2 of XPA homodimer.

Table 4B.8A. Intermolecular hydrogen bonds

Atom name	Residues	Monomer		Atom name	Residues	Monomer	Distance (Å)
N	LYS110	1	<->	O	HIS171	2	2.91
O	LYS110	1	<->	N	SER173	2	2.86
NZ	LYS110	1	<->	ND1	HIS171	2	2.94
NH1	ARG130	1	<->	OD2	ASP177	2	3.01
NH2	ARG130	1	<->	OD2	ASP177	2	2.73
OD1	ASP134	1	<->	NZ	LYS137	2	2.81
OE2	GLN201	1	<->	NZ	LYS213	2	2.97
O	GLN205	1	<->	NZ	LYS215	2	2.81
OE1	GLN205	1	<->	N	MET214	2	2.73
OE1	GLN205	1	<->	N	LYS215	2	2.87
O	GLN209	1	<->	NH1	ARG211	2	2.86
ND2	ASN210	1	<->	O	ASN210	2	3.16
NH1	ARG211	1	<->	OE1	GLN201	2	2.96
NH2	ARG211	1	<->	OE1	GLN201	2	3.04

Table 4B.8B. Salt bridges

Atom name	Residues	Monomer		Atom name	Residues	Monomer	Distance (Å)
NH2	ARG130	1	<->	OD1	ASP177	2	2.73
OD1	ASP134	1	<->	NZ	LYS137	2	2.81
OE1	GLU201	1	<->	NZ	LYS213	2	2.97
NH1	ARG211	1	<->	OE2	GLU201	2	2.96

Table 4B.8C. Non-bonded contacts

Atom name	Residues	Monomer		Atom name	Residues	Monomer	Distance (Å)
CA	GLY109	1	<->	O	HIS171	2	3.68
CA	GLY109	1	<->	OG	SER173	2	3.54
C	GLY109	1	<->	O	HIS171	2	3.70
N	LYS110	1	<->	C	HIS171	2	3.90
N	LYS110	1	<->	O	HIS171	2	2.91
CA	LYS110	1	<->	O	HIS171	2	3.72
O	LYS110	1	<->	O	HIS171	2	3.65
O	LYS110	1	<->	CA	HIS172	2	3.62
O	LYS110	1	<->	ND1	HIS172	2	3.55
O	LYS110	1	<->	C	HIS172	2	3.65
O	LYS110	1	<->	N	SER173	2	2.86
O	LYS110	1	<->	CA	SER173	2	3.82
CB	LYS110	1	<->	O	HIS171	2	3.84
CD	LYS110	1	<->	ND1	HIS171	2	3.73
CE	LYS110	1	<->	ND1	HIS171	2	3.69
NZ	LYS110	1	<->	ND1	HIS171	2	2.94

NZ	LYS110	1	<->	CE1	HIS171	2	3.51
C	GLU111	1	<->	ND1	HIS172	2	3.50
C	GLU111	1	<->	CE1	HIS172	2	3.50
O	GLU111	1	<->	ND1	HIS172	2	3.75
O	GLU111	1	<->	CE1	HIS172	2	3.30
O	GLU111	1	<->	OE1	GLN174	2	3.86
CB	GLU111	1	<->	OE1	GLN174	2	3.75
N	PHE112	1	<->	ND1	HIS172	2	3.25
N	PHE112	1	<->	CE1	HIS172	2	3.51
CA	PHE112	1	<->	CE1	HIS172	2	3.47
C	PHE112	1	<->	CB	HIS172	2	3.84
C	PHE112	1	<->	CG	HIS172	2	3.48
C	PHE112	1	<->	ND1	HIS172	2	3.48
C	PHE112	1	<->	CE1	HIS172	2	3.80
O	PHE112	1	<->	N	HIS172	2	3.85
O	PHE112	1	<->	CA	HIS172	2	3.47
O	PHE112	1	<->	CB	HIS172	2	3.42
O	PHE112	1	<->	CG	HIS172	2	3.52
O	PHE112	1	<->	ND1	HIS172	2	3.61
N	MET113	1	<->	CG	HIS172	2	3.81
N	MET113	1	<->	CD2	HIS172	2	3.80
SD	MET113	1	<->	CG	ASN169	2	3.79
SD	MET113	1	<->	OD1	ASN169	2	3.53
SD	MET113	1	<->	ND2	ASN169	2	3.35
CE	MET113	1	<->	ND2	ASN169	2	3.72
CB	APS114	1	<->	CG	HIS171	2	3.65
CB	APS114	1	<->	ND1	HIS171	2	3.79
CB	APS114	1	<->	CE1	HIS171	2	3.64
CB	APS114	1	<->	NE2	HIS171	2	3.58
CB	APS114	1	<->	CD2	HIS171	2	3.51
CG	APS114	1	<->	CE1	HIS171	2	3.67
CG	APS114	1	<->	NE2	HIS171	2	3.77
OD2	APS114	1	<->	CE1	HIS171	2	3.71
OD2	APS114	1	<->	NE2	HIS171	2	3.53
O	APS127	1	<->	CB	THR142	2	3.89
O	APS127	1	<->	OG1	THR142	2	3.85
O	APS127	1	<->	CG2	THR142	2	3.87
ND2	ASN128	1	<->	NE2	GLN146	2	3.19
O	ARG130	1	<->	CG2	THR140	2	3.81
CZ	ARG130	1	<->	OD2	APS177	2	3.34
NH1	ARG130	1	<->	OD2	APS177	2	3.01
NH2	ARG130	1	<->	CG	APS177	2	3.32
NH2	ARG130	1	<->	OD1	APS177	2	3.03
NH2	ARG130	1	<->	OD2	APS177	2	2.73
CG	APS134	1	<->	NZ	LYS137	2	3.63
OD1	APS134	1	<->	CE	LYS137	2	3.62
OD1	APS134	1	<->	NZ	LYS137	2	2.81
CD	GLU201	1	<->	NZ	LYS213	2	3.61
OE1	GLU201	1	<->	NZ	LYS213	2	3.71
OE2	GLU201	1	<->	CD	LYS213	2	3.40

OE2	GLU201	1	<->	CE	LYS213	2	3.72
OE2	GLU201	1	<->	NZ	LYS213	2	2.97
O	GLU205	1	<->	CE	LYS215	2	3.38
O	GLU205	1	<->	NZ	LYS215	2	2.81
CD	GLU205	1	<->	N	LYS213	2	3.25
CD	GLU205	1	<->	N	MET214	2	3.53
OE1	GLU205	1	<->	CA	GLU212	2	3.72
OE1	GLU205	1	<->	C	GLU212	2	3.74
OE1	GLU205	1	<->	CG	GLU212	2	3.60
OE1	GLU205	1	<->	N	LYS213	2	2.89
OE1	GLU205	1	<->	CA	LYS213	2	3.75
OE1	GLU205	1	<->	C	LYS213	2	3.67
OE1	GLU205	1	<->	N	MET214	2	2.73
OE1	GLU205	1	<->	CA	MET214	2	3.52
OE1	GLU205	1	<->	C	MET214	2	3.60
OE1	GLU205	1	<->	N	LYS215	2	2.87
OE1	GLU205	1	<->	CA	LYS215	2	3.87
OE1	GLU205	1	<->	CB	LYS215	2	3.68
OE2	GLU205	1	<->	CA	GLU212	2	3.39
OE2	GLU205	1	<->	C	GLU212	2	3.43
OE2	GLU205	1	<->	N	LYS213	2	2.91
OE2	GLU205	1	<->	CB	LYS213	2	3.80
C	GLN208	1	<->	O	ARG211	2	3.85
CB	GLN208	1	<->	O	ARG211	2	3.37
C	GLU209	1	<->	CB	ARG211	2	3.66
O	GLU209	1	<->	CB	ARG211	2	3.30
O	GLU209	1	<->	CD	ARG211	2	3.45
O	GLU209	1	<->	NE	ARG211	2	3.61
O	GLU209	1	<->	CZ	ARG211	2	3.45
O	GLU209	1	<->	NH1	ARG211	2	2.86
CA	ASN210	1	<->	CZ	ARG211	2	3.82
CA	ASN210	1	<->	NH1	ARG211	2	3.79
C	ASN210	1	<->	NH1	ARG211	2	3.80
O	ASN210	1	<->	CZ	ARG211	2	3.54
O	ASN210	1	<->	NH1	ARG211	2	3.33
O	ASN210	1	<->	NH2	ARG211	2	3.31
CB	ASN210	1	<->	O	ASN210	2	3.72
ND2	ASN210	1	<->	O	ASN210	2	3.16
CZ	ARG211	1	<->	OE1	GLU201	2	3.44
NH1	ARG211	1	<->	CD	GLU201	2	3.78
NH1	ARG211	1	<->	OE1	GLU201	2	2.96
NH2	ARG211	1	<->	CD	GLU201	2	3.53
NH2	ARG211	1	<->	OE1	GLU201	2	3.04
NH2	ARG211	1	<->	OE2	GLU201	2	3.36
CE	MET214	1	<->	OE1	GLN216	2	3.89

4B.4.5. Partner attraction effect of XPA homodimer with DNA; comparisons of XPA homodimer in B and U states

For the comparative PPI profile study between the monomers of XPA homodimers in B and U states, the data required for XPA homodimers in U state was taken from our previous work [431]. The interface statistics for both states have been summarized in **Table 4B.9**. We observed that the number of interacting residues, hydrogen bonds, and vdW interactions was increased for the XPA homodimer in the B state, except for salt bridges which had decreased when the XPA homodimer underwent complexation from U to B state. There was also an increase in the size of their interface area as seen in **Table 4B.9**. These types of differences for proteins upon change from U to form are either due to partner attraction or partner accommodation effect [439, 440]. Both the terms were coined by Chakravarty et al. [440], which explain the reaction of the interface residues of one partner (protein) upon binding with another partner (protein/DNA/RNA). In the case of the former, the interface atoms shift their position upon contact with the binding partner, increasing their accessible surface area (ASA), while for the latter, the interface residues readjust themselves to accommodate the partner, thereby decreasing their ASA [440, 441].

Table 4B.9. Comparative interface statistics of XPA homodimer in B and U states.

Average properties	XPA homodimer in B state		XPA homodimer in U state*		Status concerning DNA bound (B) state
	Monomer 1	Monomer 2	Monomer 1	Monomer 2	
No. of interface residues	17	18	9	13	Increased
Interface area (Å^2)	1093	1102	536	461	Increased
No. of salt bridges	5		6		Decreased
No. of hydrogen bonds	14		9		Increased
No. of non-bonded contacts	112		65		Increased

**The computational data for the XPA homodimer in DNA unbound (U) state has been taken from our earlier work.*[29]*

The partner attraction or partner accommodation effects are usually calculated by estimating the ASA values for the interface atoms common to protein in both U and B states as described earlier[56,58] using the following formula:

$$\Delta ASA = [ASA(B) - ASA(U)] \quad (5)$$

$$\delta ASA = \Delta ASA/ASA(B) \quad (6)$$

where ASA(B) and ASA(U) are the solvent-accessible surface area for the common interface residues in the B and U states. δASA is the difference in ASA values (ΔASA) of a protein in U and B states relative to the total ASA(B). If the δASA value is positive, the interaction is termed a partner attraction effect, and if the δASA value is negative, the interaction is termed a partner accommodation effect.

We have used the Proteins, Interfaces, Structures and Assemblies (PISA) server (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) [442], which works on Lee and Richards algorithm [443] to calculate the ASA values for XPA homodimer in B and U values, with the probe radius of 1.4 Å [438]. We have taken the lowest energy conformer for the protein structure of XPA homodimer in U state as described earlier [431] for this comparative work. The obtained ASA(B), ASA(U), ΔASA, and δASA values for XPA homodimer have been given in **Table 4B.10A** and **4B.105B**, while the ΔASA and δASA values have been plotted in the graph as shown in **Figure 4B.10A-D**. We observed an increase in the ASA values for the entire DNA interacting residues of XPA homodimer indicating the partner attraction effect. In **Figure 4B.11**, we have shown some examples of the residues from both monomers of XPA homodimers in the B and U states exhibiting positive δASA.

Table 4B.10. ΔASA and δASA values of XPA homodimer in B and U states

Table 4B.10A. ΔASA and δASA values of monomer 1 of XPA homodimer in B and U states.

Residues	ASA of monomer 1 of XPA homodimer in B state (Å^2)*	ASA of monomer 1 of XPA homodimer in U state (Å^2)*	ΔASA = [ASA(B) - ASA(U)] (Å^2)	δASA = ΔASA/ASA(B)	δASA status
PRO124	50.01	42.37	7.64	0.15	positive
THR125	62.47	32.29	30.18	0.48	positive

CYS126	55.94	20.03	35.91	0.64	positive
ASN128	67.2	49.32	17.88	0.26	positive
CYS129	55.32	29.24	26.08	0.47	positive
ASN210	35.4	1.59	33.81	0.95	positive
LYS213	158.43	70.83	87.6	0.55	positive
LYS217	137.69	109.96	27.73	0.2	positive
LYS221	140.83	84.64	56.19	0.39	positive
LYS224	110.52	44.27	66.25	0.59	positive
ARG228	139.42	84.5	54.92	0.39	positive
ALA229	60.17	31.45	28.72	0.47	positive
SER232	90.98	79.92	11.06	0.12	positive

Table 4B.10B. ΔASA and δASA values of monomer 2 of XPA homodimer in B and U states.

Residues	ASA of monomer 2 of XPA homodimer in B state (Å^2) *	ASA of monomer 2 of XPA homodimer in U state (Å^2)*	ΔASA = [ASA(B) - ASA(U)] (Å^2)	δASA = ASA/ASA(B)	δASA status
GLN146	98.43	82.43	16	0.16	positive
ARG207	115.49	87.42	28.07	0.24	positive
GLU209	126.46	75.71	50.75	0.4	positive
ASN210	87.89	24.08	63.81	0.72	positive
GLU212	105.48	56.32	49.16	0.46	positive
LYS213	131.57	77.35	54.22	0.41	positive
ASP220	87.54	53.5	34.04	0.38	positive
LYS221	125.89	91.23	34.66	0.27	positive
LYS222	86.94	71.95	14.99	0.17	positive
LYS224	121.92	65.1	56.82	0.46	positive
GLU225	80.61	42.71	37.9	0.47	positive
LYS236	147.33	108.75	38.58	0.26	positive
ARG237	186.64	128.23	58.41	0.31	positive

*The computational data for the XPA homodimer in the U state has been taken from our earlier work.[29]

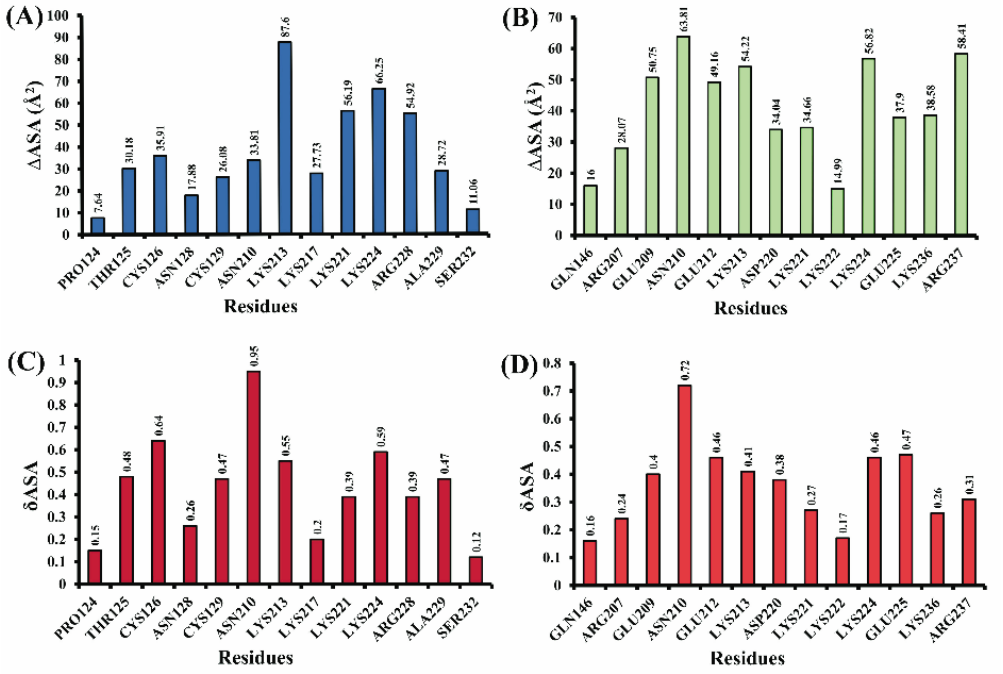

Figure 4B.10. (A) ΔASA and (B) δASA values for the interface residues of monomer 1, and (C) ΔASA and (D) δASA values for the interface residues of monomer 2 of XPA homodimer.

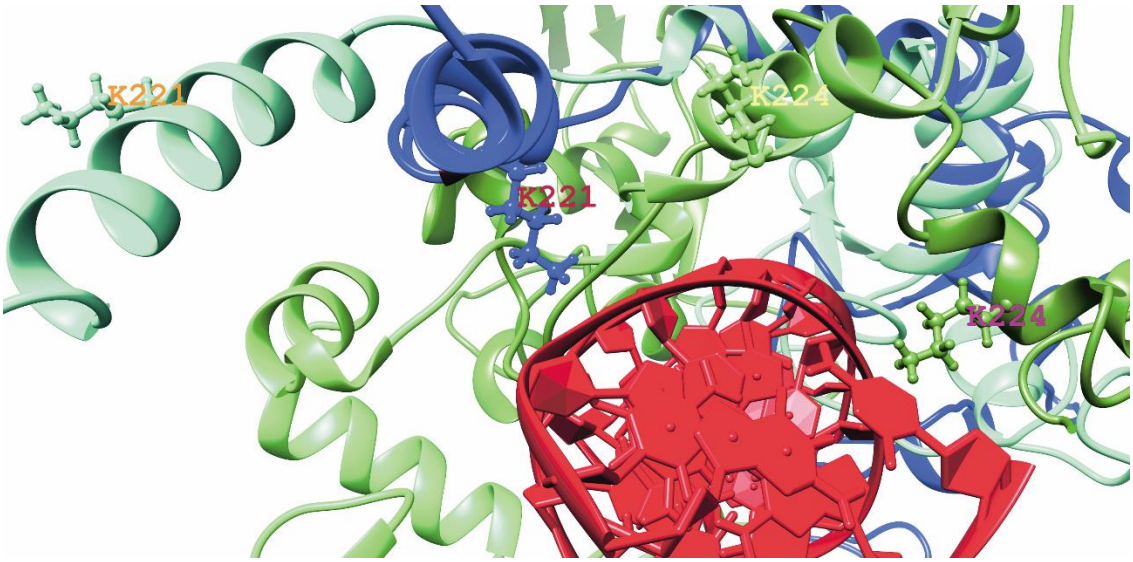

Figure 4B.11. Interface residues (ball and stick form) of XPA homodimer show a partner attraction effect with DNA. The residue K221 (deep blue) and K224 (dark green) from monomers 1 and 2 of XPA homodimer for B state, while K221 (cyan) and K224 (sea green) from monomers 1 and 2 of XPA homodimer for U state. The ASA increased from 84.64 to 140.83 Å^2, and 65.1 to 121.92 Å^2 for the residues K221, and K224, respectively.

4B.4.6 Energetics of DNA binding to XPA homodimer using MM-PBSA approach

The BFE of the XPA homodimer-DNA complex that was determined using the MM-PBSA approach has been summarized in **Table 4B.11**. The BFE value obtained for our DPC was -62.33 kcal mol^{-1}, and this suggested a favorable binding between the XPA homodimer and DNA. The gas-phase energy (ΔE_{MM} =-4816.80 kcal mol^{-1}) was found to be favorable for the binding, with the contributions from electrostatic (ΔE_{ele} = -4742 kcal mol^{-1}) and van der Waals energy (ΔE_{vdW} = -74.75 kcal mol^{-1}). Further, the non-polar solvation energy (ΔG_{surf}) was also seen to be favoring the DPI (-12.87 kcal mol^{-1}). Earlier studies have shown that XPA's binding affinity to DNA in the 2:1 ratio is of a much stronger and dominant type when compared to its binding to DNA in a monomer form.[27] XPA It has also been observed that XPA homodimer equally has a high binding capacity towards other members of the NER family.[10,26,30] It was observed that the redefined DBD domain of XPA (aa98-239) can bind to DNA with equal strength as the full-length protein (aa1-273) and that it has a high preference toward ds/ssDNA junction over ss-DNA and ds-DNA [58, 176].

Table 4B.11. Binding free energy (BFE) analysis for XPA homodimer-DNA complex.

Method	Components	Complex (kcal mol^{-1})		Ligand (kcal mol^{-1})		Receptor (kcal mol^{-1})		ΔG_{bind} (kcal mol^{-1})	
		Energy values	Std. dev. (±)	Energy values	Std. dev. (±)	Energy values	Std. dev. (±)	Energy values	Std. dev. (±)
MM-PBSA	ΔE_{vdW}	-2421.24	20.85	-445.75	6.7	-1900.74	21.2	-74.75	4.21
	ΔE_{ele}	-20464.6	70.26	4628.42	28.1	-20351	69.7	-4742	61.15
	ΔE_{MM}	-22885.83	72.12	4182.67	27.8	-22251.8	64.6	-4816.8	61.24
	ΔG_{PB}	-8573.99	76.2	-6874.17	25.2	-6467.04	77.6	4767.21	60.1
	ΔG_{surf}	179.93	1.09	34.11	0.18	158.61	1.06	-12.87	0.3
	ΔG_{solv}	-8394.07	75.93	-6840.1	25.3	-6308.43	77.8	4754.42	60.8
	PB_{TOT}	-31279.9	42.82	-2657.4	11	-28560.3	41.8	-62.33	5.64

MM-PBSA= molecular mechanics Poisson-Boltzmann surface area; ΔE_{ele} *= electrostatic energy contribution from MM;* ΔE_{vdW} *= van der Waals contribution from MM;* ΔE_{MM} *= total gas phase energy (sum of ele, vdW, and int);* ΔG_{PB} *= polar solvation free energy calculated by PB method;* ΔG_{surf} *= non-polar 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;* $\Delta G_{bind} = \Delta G_{complex} - [\Delta G_{receptor} + \Delta G_{ligand}]$.

4B.4.7 Free energy decomposition

The standard BFE can be partitioned into contributions from various interacting amino acids involved in DPI using the PRED analysis. The per-residue footprint of interface residues of XPA monomer 1 and monomer 2 is shown in **Figure 4B.12**. As seen in **Fig. 4B.12**, we can see that all the interface residues have shown negative energy, which proves their contribution to the DNA binding. Among these, the residues belonging to the C-terminal end of both monomers, especially lysine and arginine residues had made major contributions to the DPI between XPA homodimer and DNA. The residues K213, K217, K221, K224, and R228 from monomer 1 had shown PRED values of -8.69, -5.57, -7.09, -7.97, and -8.88 kcal mol^{-1}, while the residues N210, K213, K221, K222, K224, E225, K236, and R237 of monomer 2 had exhibited higher PRED values of -5.34, -6.24, -8.64, -8.79, -7.35, -4.72, and -7.1 kcal mol^{-1}, respectively. These residues were also involved in intermolecular hydrogen bond formation, suggesting that they are important for DPC formation as well.

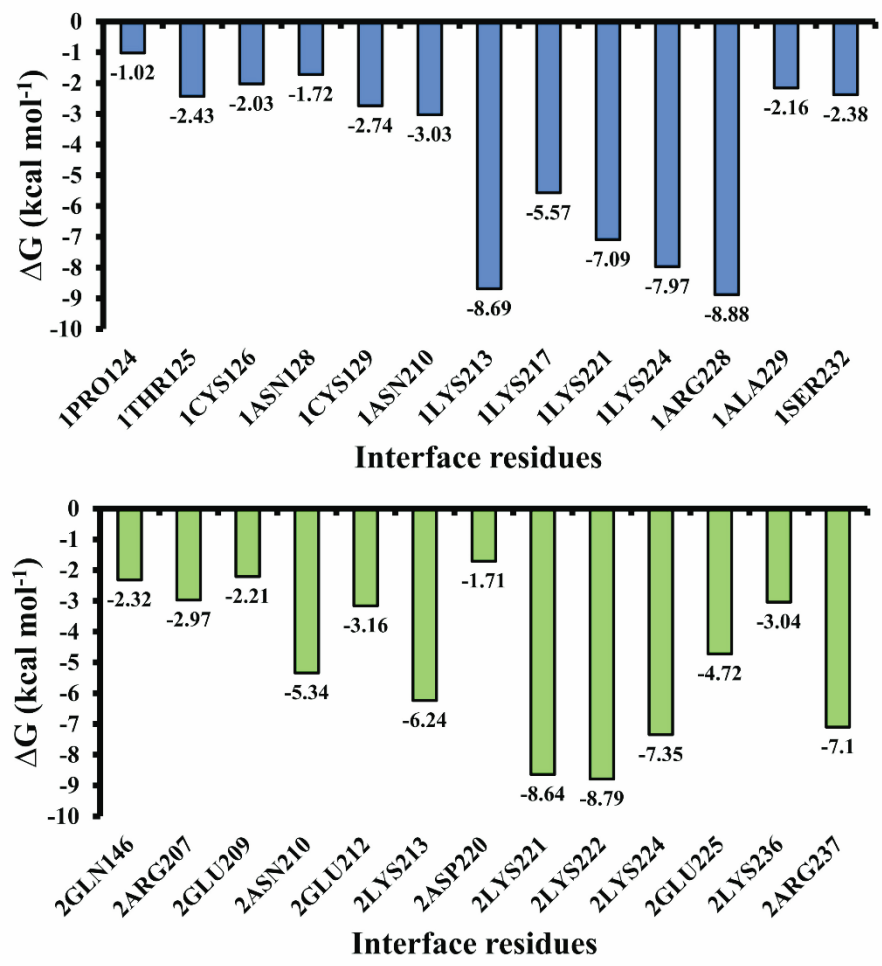

Figure 4B.12. PRED plots for the interface residues of (A) monomer 1 and (B) monomer 2 of XPA homodimer involved in the DPI.

These C-terminal residues that have shown higher decomposition energy here, had earlier exhibited higher chemical shift perturbations (CSP) peaks during NMR analysis upon the titration of DNA substrates into ^{15}N-labelled DBD of XPA. Both single-point mutations of K221E, K222E, R207Q, and R228E, and double mutations of K221E/R228E led to a decrease in the DNA-binding affinity of XPA [176]. It has also been seen that Lys, Arg, Glu, and Asp residues have high chances of being present at the DNA-binding site [444,445]. Furthermore, the truncations of these C-terminal residues have caused a reduction in XPA-DNA interaction, and severe neurological complications in XP patients, indicating DPC between XPA homodimer and DNA is crucial to the smooth execution of NER [371, 176].

4B.5 Conclusion

In this work, we used a computational approach involving molecular docking and MD simulations to understand the DPI between XPA homodimer and DNA at an in-depth molecular level. From the trajectory analyses, we found this DPC to be stable. RMSF analysis showed that most of the residues had participated in the DPI. We also observed the changes in the backbone structure and interface area of DPC, along with the increase in ASA of the interface residues of XPA homodimer due to its partner attraction effect with the DNA. BFE analysis showed that our DPC had a high binding affinity of -62.33 kcal mol^{-1} and that the C-terminal residues from both monomers, K213, K217, K221, K222, K224, K236, E225, R228, and R237, in particular, had played a major role in the DPI. Our results thus provide theoretical details on the DNA-binding function of the DBD of XPA homodimer, which is in agreement with the experimental studies done before.