Role of Cys-298 in specific recognition of glutathione by aldose reductase

Journal:	*Journal of Biomolecular Structure & Dynamics*
Manuscript ID	TBSD-2020-2534
Manuscript Type:	Research Article
Date Submitted by the Author:	29-Oct-2020
Complete List of Authors:	Sekhon, Gurprit; Panjab University Singh, Balvinder; IMTECH, Bioinformatics Singh, Ranvir; Panjab University,
Keywords:	Aldose Reductase, Diabetes, Glutathione, Molecular Dynamics, Polyol Pathway
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
AR_movie.mpg	

Role of Cys-298 in specific recognition of glutathione by aldose reductase

Gurprit Sekhon[1], Balvinder Singh[2*], Ranvir Singh[1*]

[1]Department cum National Centre for Human Genome Studies & Research, Pharmacy Extention Block, Panjab University, Chandigarh 160014, India.

[2]Bioinformatics Center, Institute of Microbial Technology, Council of Scientific and Industrial Research, Sector 39A, Chandigarh 160036, India.

[*]Corresponding authors, Email to: ranvir1@pu.ac.in (Ranvir Singh), bvs@imtech.res.in (Balvinder Singh)

KEYWORDS

Aldose Reductase, Diabetes, Polyol Pathway, Glutathione, Molecular Dynamics

ABSTRACT

Aldose reductase (AR) is an NADPH-dependent oxidoreductase that is well-studied for its role in Diabetes Mellitus. Glutatione conjugated aldehydes are efficiently catalysed by AR. We have employed molecular dynamics simulations to investigate the dynamics of a glutathione analog, γ-glutamyl-S-(1,2-di-carboxyethyl)-cysteinyl-glycine (DCEG), into the binding pocket of AR. Study revealed that backbone nitrogens of Ala-299 and Leu-300 form a tiny pocket gated by thiol group of Cys-298. The glycine moiety of DCEG was able to displace the thiol group of Cys-298 to make hydrogen bond interactions with backbone of Ala-299, Leu-300, and Leu-301. This study provides the details of the dynamic interactions of DCEG in the binding pocket of AR, and shall aid in the design/discovery of differential inhibitors against AR.

INTRODUCTION

Ever since the discovery of polyol pathway in 1956, it has been investigated as one of the mechanisms by which hyperglycemia induced tissue damage leads to chronic diabetic complications [1-2]. The proportion of glucose entering the polyol pathway increases from less than 3% to more than 30% when intracellular conditions change from euglycemic to hyperglycemic [3]. Increased activity of polyol pathway under hyperglycemic conditions is implicated in the etiology of diabetes mellitus (DM) [4-5].

Aldose reductase (AR) catalyzes the conversion of glucose to sorbitol as the first step of polyol pathway. A lot of research effort has been devoted to the development of AR-inhibition-based therapy to manage complications associated with DM [6]. A few small molecules have entered the clinical trials as AR inhibitors (ARIs), but still, there is no FDA approved ARI as a drug. Epelrestat is an ARI which is marketed in Japan and India for acute diabetic neuropathy [7].

Evolutionarily, AR belongs to Aldo-keto reductase (AKR) superfamily [8]. AKRs are NADPH-dependent oxidoreductases that are evolved with broad substrate specificity and generally function as detoxifying enzymes [9]. Like other AKRs, toxic aldehydes generated from lipid peroxidation are preferred physiological substrates of AR [10]. Whereas under hyperglycemic conditions, AR has implications in secondary diabetic complications as a part of polyol pathway; under euglycemic conditions, AR has an equally beneficial role as a part of the cellular detoxification system [11].

AR folds into a TIM barrel structure and three loops at the C-terminal end of the enzyme make a large elliptical binding pocket (*lxbxh* =13x7x12 Å) [12]. The binding pocket of AR can be divided into two sub pockets: a rigid anion-binding pocket conserved among different AKRs that is the site of the catalytic reaction and a flexible specificity pocket to bind different kinds of substrates [13-14].

As a detoxifying enzyme, the binding pocket of AR is evolved to bind a variety of substrates [15]. To facilitate the recognition of several types of substrates, AR primarily recognizes the functional aldehyde group, an oxidized 2-carbon, and part of the side chain that can be polar and uncharged or can be non-polar too. Aldehyde functional group and polar/non-polar side-chain bind into anion binding pocket and specificity pocket, respectively [16].

Enzyme kinetic studies have found several glutathione-conjugated aldehydes to be the preferred substrates for AR as compared to parent free aldehydes (by a factor of 4-1000 fold), leading to the proposition that a separate glutathione binding scaffold is present within the binding pocket of AR [17-19]. Crystal structure of AR complexed with a glutathione analog, γ-glutamyl-S-(1,2-di-carboxyethyl) cysteinyl-glycine (DCEG), and the cofactor NADPH has been determined [20].

DCEG essentially represents glutathione backbone (Tripeptide: 'Glu-Cys-Gly'), with di-carboxyethyl moiety attached to central cysteine. DCEG is a polar molecule with a high degree of flexibility, which makes it an ideal candidate to probe the binding pocket AR for specific interactions. To follow the interactions of glutathione backbone with AR, we carried out a classical molecular dynamics simulation study of the ternary complex of AR, bound with NADPH and DCEG.

We observed two distinct and discrete conformational changes in DCEG with respect to the conformational state observed in crystal structure. Our study revealed that recognition of glutathione by AR is facilitated by a tiny pocket formed by backbone nitrogens of Ala-299 and Leu-300. This tiny pocket is gated by sulfur of Cys-298, which acts as a specificity determinant for recognition of glutathione backbone

MATERIALS AND METHODS

Preparation of MD simulation systems: Starting coordinates to prepare MD simulation systems were taken from the X-ray crystal structure of the AR with bound NADPH and DCEG (PDB ID: 2F2K) [20]. Ambertools19 and Amber18 software packages were used to prepare the initial simulation system and to run MD simulations, respectively [21]. After the removal of crystallographic waters, the protonation states of titratable amino acid residues were assigned at pH 7.0 with the help of h++ web server (http://biophysics.cs.vt.edu/) [22]. The protonation states assigned by h++ server were verified manually. Histidine residues at 41, 46, and 83 positions were protonated at δ-nitrogen, while Histidine residues at 110, 187, 240, 306, and 312 positions were protonated at ε-nitrogen. His-163 was protonated at both δ and ε-nitrogens. His-110 is a critical amino acid residue for the catalytic reaction of AR, and its protonation state assigned by h++ server is consistent with the one reported in the literature [23-24]. All other amino acid residues were kept in their dominant protonation state at the physiological pH.

The protein was modeled with ff14SB force field [25], while NADPH and DCEG were modeled with a general force field available with Amber19, Gaff version 2.0 [26]. The parameters and AMI-BCC charges of the NADPH and DCEG were generated by the antechamber tool of the Amber19 software package [27-28]. The leap module of Amber19 software package was used to set up both simulation systems, with and without DCEG, and to generate the coordinate and topology files. Simulation systems were solvated using the TIP3P water model [29], with a buffer size of 17 Å around the solute. A truncated octahedron box was used for solvation. 19374 and 19163 water molecules were added during the solvation of the simulation system for DCEG and without DCEG, respectively. Both systems were neutralized by adding counter ions. To maintain the physiological conditions, NaCl was kept at 150 mM concentration (total 104 counter ions, 52 each of Na^+ and Cl^-). The resulting coordinate and topology files were saved.

The final simulation system with/without DCEG resulted in a truncated octahedron box with a volume of 696808.20/692268.23 $Å^3$ (corresponding to the side length of 96.73 Å), containing a single chain polypeptide of 315 amino acid residues, one molecule of NADPH, one molecule of DCEG/without DCEG, 53 sodium ions, 52 chloride ions, and 19270/19059 water molecules.

MD simulations were run on an intel cluster using a parallel implementation of PMEMD from Amber18. In the first round of minimization, 5000 steps of steepest descent (SD) were performed with all non-hydrogen atoms of the system restrained at 10 Kcal/mol/$Å^2$. In the second round, another 5000 steps of SD were performed with a restraint weight of 10 Kcal/mol/$Å^2$ on all non-hydrogen atoms of solute only. Another 5000 steps of SD followed the second round of minimization with only carbon and nitrogen atoms of protein backbone restrained at 10 Kcal/mol/$Å^2$. In the final round of minimization, the system was minimized without any restraints by 5,000 steps of SD, followed by 15,000 steps of the conjugant gradient (CG). The minimized simulation systems were taken as the starting coordinates of the MD simulation system.

Equilibration of MD simulation systems: Molecular systems of solvated AR with DCEG and without DCEG, thus prepared were heated and equilibrated in three simulation steps with a total simulation time of 3 ns. In the first step of 300 ps, the system was heated from 10 K to 310 K in the NVT ensemble with a restraint weight of 5.0 Kcal/mol/$Å^2$ on the non-hydrogen atoms of solute (AR, NADPH, and DCEG). In the second step, 300 ps simulation run was performed in the NPT ensemble with restraint weight on the non-hydrogen atoms of solute reduced to 3.0 Kcal/mol/$Å^2$. In the third step, the restraint weight on non-hydrogen atoms of solute was further reduced to 1 Kcal/mol/$Å^2$ for 400 ps in the NPT ensemble.

The final step of equilibration was carried out for 2 ns in the NPT ensemble without any restraints on the system. During equilibration, the temperature of the simulation system was controlled by the Langevin thermostat [30], and the collision frequency for the Langevin dynamics was set to 2 ps^{-1}. The Berendsen barostat [31] was used to control pressure to the default value of 1 bar (~0.987 atm) using the default value of pressure relaxation time (1.0 ps). The density of the systems was equilibrated to 1.00 g/cm^{-3}. Bond lengths involving hydrogen atoms were constrained with SHAKE [32], and an MD time-step of 0.002 ps was used during the equilibration phase of the simulation. Coordinates were saved at 10 ps intervals during the equilibration phase of the simulation.

Production MD simulations: Production run of MD simulations was carried out for a duration of 500 ns each for both the systems. During production MD simulations, the temperature was controlled at 310 K by using the weak-coupling algorithm [31]. The heat bath was loosely coupled to the simulation system with a coupling time set to 10.0 ps. The pressure control parameters, SHAKE parameters, and MD time step remained the same as used in the last phase of equilibration. Coordinates were saved at intervals of 100 ps.

Data analysis: The CPPTRAJ module of Amber19 was used for all types of trajectory analysis tasks [33]. Molecular figures and movies were prepared in PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, and LLC). Graphs were plotted in Gnuplot (http://gnuplot.sourceforge.net/). Center of Mass (CoM) was calculated by a python script downloaded from 'Pymol-script-repo' (https://github.com/Pymol-Scripts/Pymol-script-repo). The molecular visualization program, VMD (http://www.ks.uiuc.edu/ Research/vmd/) was used for visualization of MD simulation trajectories [34].

RESULTS

Molecular dynamics trajectories of AR: Two separate MD simulations of 500 ns each were carried out for holoAR (the binary complex of AR and NADPH), with and without DCEG, respectively (Figure 1). Average values of root-mean-square deviation (RMSD) of the protein backbone atoms (Cα, C, and N) to the minimized structure were calculated to be 3.12 Å and 2.93 Å for MD trajectories without and with DCEG, respectively (Figure 1A).

Average values of radius of gyration (RoG) of the protein were calculated to be 19.57 Å and 19.52 Å for MD trajectories without and with DCEG, respectively (Figure 1B). Average root-mean-square fluctuation (RMSF) values of individual amino acid residues were computed to be 1.09 Å and 1.08 Å for MD trajectories without and with DCEG, respectively (Figure 1C). Besides terminal amino acid residues and loop regions comprising the active site of the enzyme (Figure 1D), the average RMSF was less than 2.0 Å. Loop B showed the highest fluctuation in both MD trajectories. The average RMSF of loop B residues was higher (3.86 Å) for MD trajectory with DCEG as compared to MD trajectory without DCEG (3.23 Å).

Dynamics of γ-glutamyl-S-(1,2-di-carboxyethyl)-cysteinyl-glycine (DCEG): DCEG is an analog of glutathione tripeptide ('Glu-Cys-Gly', with a gamma peptide linkage between Glu and Cys) where the hydrogen atom of the thiol side chain in central cysteine residue is replaced by a di-carboxyethyl group (Figure 2A). Four dihedral angles, τ1-τ4, were selected to follow the movement of three arms of DCEG (Figure 2A). τ1 monitored the movement of the glycine arm of DCEG. τ2 monitored the movement of the glutamate arm of DCEG. τ3 and τ4 monitored the movement of the di-carboxyethyl arm of DCEG (Figure 2A).

During MD run, τ1 showed little variation with an average value of 13.59º (Figure 2B). τ2 changed from an average value of 261º to 65º around 216 ns (Figure 2B). τ3 showed little variation with an average value around -21º. τ4 changed around 79 ns from the average value of 208º and to 299º (Figure 2B). Changes in dihedral angles suggested two conformational transitions in DCEG; first around 79 ns in τ4 and second around 216 ns in τ2 (Figures 2B, Supplementary Information: Figure S1).

To decipher the overall dynamics of DCEG, we carried out dihedral principal component analysis (dPCA) of MD trajectory [35]. Based on four dihedral angles of DCEG (τ1-τ4), four eigenvectors were computed (Supplementary Information: Figures S2, S3, and S4). First two eigenvectors were able to describe the dominant motion of DCEG. A plot of first and second eigenvector resulted in three well-defined clusters which suggested the presence of three conformations of DCEG in MD trajectory (Figure 2C).

To find representative conformations of DCEG, MD trajectory was clustered based on RMSD of DCEG. 2D RMSD plot of DCEG indicated the presence of three clusters in MD trajectory (Figure 2D). Clusters were obtained by using K-means clustering algorithm (Figure 2E, Supplementary Information: Table S1) [36]. The most populated cluster (49% of MD trajectory) was numbered 1 (Orange), second largest cluster (32% of MD trajectory) was numbered 2 (Green) and the least populated cluster (19 % of MD trajectory) was numbered 3 (Skyblue, Figure 2E).

Frame numbers 4397, 989, and 510 represented the centroids of clusters number 1, 2, and 3 respectively (Supplementary Information: Table S1). Centroids were considered representative structures of respective clusters. The superposition of three representative structures of DCEG is shown in figure 2F. First and second conformational changes resulting from the change of τ4 and τ2, respectively, are marked (Figure 2F).

Interactions between DCEG and AR: To look into specific interactions of DCEG with AR, hydrogen bonds were computed between DCEG and AR from MD trajectory. As dynamics of DCEG involved two conformational changes at 79 ns and 216 ns of MD simulation (Figure 2), hydrogen bond occupancy between any two atoms of DCEG and AR was computed for the following periods: 1-5000 ns (total time-interval), 1-79 ns (first time-interval), 79-216 ns (second time-interval), and 216-500 ns (third time-interval) (Supplementary Information: Tables S2, S3, S4, and S5).

A hydrogen bond was considered significant only if it had occupancy of more than 10% in at least one of the four time-intervals defined above. This way we obtained a total of nine hydrogen bond interactions between DCEG and AR (Figure 3A). Atoms of DCEG and AR involved in hydrogen bonding (Figure 3A) are labeled in the crystal structure of AR●NADPH●DCEG (Figure 3B).

The transition from first to second time-interval involved one major change in hydrogen bond interaction between OE2 atom of DCEG and ε-Nitrogen of Trp-111(TRP111@NE1&OE2, Figure 3A). Hydrogen bond occupancy of TRP111@NE1&OE2 interaction changed from 11.39 % for first time-interval to 55.18 % for second time-interval (Figure 3A). This change suggested reorganization of the di-carboxyethyl arm of DCEG into anion binding pocket of AR and is a result of the first conformational change around 79 ns of MD simulation (Figure 2F).

The transition from second to third time-interval involved two changes in hydrogen bond interactions. First change occurred in hydrogen bond interaction between the OE1 atom of DCEG and the side-chain oxygen of Ser302 (SER302@OG&OE1, Figure 3A). Hydrogen bond occupancy of SER302@OG&OE1 interaction changed from 20.51 % for second time-interval to 0 % for third time-interval (Figure 3A). Second change occurred in hydrogen bond interaction

between the OC1 atom of DCEG and ζ-Nitrogen atom of Lys-21 (LYS21@NZ&OC1, Figure 3A). Hydrogen bond occupancy of LYS21@NZ&OC1 interaction changed from 9.84 % to 40 % from second to third time-interval (Figure 3A). These changes in hydrogen bonding interactions between DCEG and AR indicated the movement of the glutamate arm of DCEG away from Ser-302 toward Lys-21 as a result of second conformational change around 216 ns of MD simulation (Figure 2F).

Atoms OC2 and OC3 of DCEG are part of glycine arm of DCEG (Figure 3B). Their hydrogen bond interactions with Ala-299 (ALA299@N&OC2, 22% hydrogen bond occupancy in total time-interval), Leu-300 (LEU300@N&OC2 and LEU300@N&OC3, 53% and 21% hydrogen bond occupancy in total time-interval), and Leu-301 (Leu301@N&OC3, 45% hydrogen bond occupancy in total time-interval) were mostly stable throughout the trajectory as shown in figure 3A. The average distance for each of these hydrogen bond interactions was computed to be 2.8 Å (Figure 3C).

It was observed that the sulfur atom of Cys-298 was regularly attacking backbone nitrogens of Ala-299, and Leu-300. This electrophilic attack seemed responsible for the induction of polarity in backbone nitrogens of Ala-299 and Leu-300 which formed the basis of strong hydrogen bond interactions with OC2 and OC3 atom of DCEG. The whole scenario can be described as Cys-298 acting as a gatekeeper to a tiny pocket made up of backbone nitrogens of Ala-299 and Leu-300 (Figure 3D). Atom OC2 of DCEG was able to enter that pocket multiple times over the MD trajectory as suggested by the distance of sulfur atom of Cys-298 (Blue) and OC2 atom of DCEG (Orange) from the center-of-mass of backbone nitrogen of Ala-299 and Leu-300 (Figure 3E, Supplementary Information: Movie S1).

Dynamics of Loop B of AR: Loop B of AR regulates binding and release of NADPH by a conformational transition between closed to opened state [37]. To follow the dynamics of loop B in both MD trajectories (with and without DCEG), cluster analysis was performed based on RMSD of loop B residues (aa 213-226) (Supplementary Information: Figures S5, S6 and Tables S6, S7). K-means clustering algorithm was used for clustering the trajcetories [36]. Clustering resulted in three conformations of loop B in MD trajectory without DCEG (Figure 4A) as well as in MD trajectory with DCEG (Figure 4B).

Structural superposition of three conformations of loop B revealed interactions that stabilized closed, intermediate, and opened states in MD trajectories, without DCEG (Figure 4C) and with DCEG (Figure 4D). Sulfur-aromatic interaction between sulfur of Cys-298 and indole ring of Trp-219 stabilized closed state of loop B. Intermediate state of loop B was stabilized by π-stacking interaction between indole rings of Trp-219 and Trp-295. The opened state of loop B was stabilized by cation-π interaction between the guanidine group of Arg-293 and the indole ring of Trp-219 (Figures 4C&D).

Distances between interacting residues in closed, intermediate, and opened states were calculated from MD trajectory without DCEG (Blue) as well as for MD trajectory with DCEG (Orange) (Figure 4E). The average distance between Arg-293 and Trp-219 was calculated to be 6.73 Å and 14.17 Å for MD trajectories without and with DCEG, respectively. The average distance between indole ring of Trp-219 and Trp-295 was found to be 6.36 Å and 9.23 Å for MD trajectory without and with DCEG, respectively. The average distance between indole ring of Trp-219 and thiol group of Cys-298 was calculated to be 14.5 Å and 13.8 Å for MD trajectory without and with DCEG, respectively (Figure 4E).

The indole ring of Trp-219 was properly oriented against the guanidine group of Arg-293 only in MD trajectory without DCEG (Figure 4C). This orientation of the indole ring of Trp-219 was not observed in the MD trajectory with DCEG (Figure 4D). Thus opened state of loop B was not observed in MD trajectory with DCEG (Figure 4D). In MD trajectory without DCEG, three states of loop B were stable and transition from closed to intermediate, and intermediate to open state was well-defined (Figure 4C and Supplementary Information: Movie S1), but in MD trajectory with DCEG, loop B seemed to be stuck in an intermediate state and was unable to find the opened state (Figure 4D and Supplementary Information: Movie S1).

DISCUSSION

Diabetes and its complications remain a serious health challenge around the world [38]. Aldose reductase (AR)-inhibition-based therapy is still a promising solution to manage secondary complications associated with diabetes [39]. Under physiological conditions, AR is part of cellular detoxification system, reducing toxic aldehydes generated due to lipid peroxidation. It has been shown that these aldehydes conjugated to glutathione are better substrates to AR and are catalyzed more efficiently by AR [17]. Crystal structure of AR complexed with cofactor NADPH and glutathione analog has been determined [20]. To study the conformational dynamics of DCEG inside the binding pocket of AR, we carried out MD simulations of ternary complex of AR bound to NADPH and DCEG.

During MD simulation, we observed two distinct conformational changes of DCEG with respect to the starting structure. The first conformational change involved rearrangement of the di-carboxyethyl moiety of DCEG that took place around 79 ns of the MD run (Figure 2F). As a result of this DCEG established hydrogen bonding with His-110 and Trp-111 (Figure 3A and 3C). His-110 and Trp-111 are part of the anion-binding pocket of AR [13-14]. Binding of di-carboxyethyl arm of DCEG into anion-binding pocket of AR is supported by the crystal structure of AR bound to NADPH and DCEG [20].

The second conformational change resulted in the movement of the N-terminal glutamate arm of DCEG (Figure 2F). Movement of glutamate arm of DCEG resulted in the loss of a hydrogen bond interaction between OE1 atom of DCEG and side-chain oxygen of Ser302 as well as formation of a hydrogen bond between atoms OC1 of DCEG and ζ-Nitrogen of Lys-21 (Figure 3A and 3C). Hydrogen bond interaction between glutamate moiety of DCEG and Lys-21 of AR reported here (Figure 3A) was not identified in crystal structure of AR bound to NADPH and DCEG [20].

A novel aspect of the dynamics of DCEG was revealed in present study. We observed interplay between atom OC2 of DCEG and sulfur atom of Cys-298 of AR with respect to backbone nitrogens of Ala-299 and Leu-300 (Figure 3D). During the MD run, atom OC2 of DCEG and sulfur of Cys-298 displayed mutually excluding proximity for backbone nitrogens of Ala-299, Leu-300 (Figure 3E, Supplementary Information: Movie S1). The sulfur atom of Cys-298 has to be displaced for access to main chain nitrogens of Ala-299, Leu-300, and Leu-301 (Figure 3D). Atom OC2 of DCEG was able to displace the sulfur of Cys-298 multiple times throughout the MD trajectory (Figure 3E and Supplementary Information: Movie S1). It implied that the movement of thiol group of Cys-298 is critical for specific recognition of glutathione.

Fidarestat is a potent and specific inhibitor of AR with IC50 values 9 nM [40]. Fidarestat is 130 times more selective for AR in comparison to ALR. The crystal structure of fidarestat with AR identified a vital hydrogen bond interaction of the oxygen atom of the carbamoyl group with the main-chain nitrogen of Leu-300 [41]. In aldehdyde reductase, a closely related AKR with high sequence similarity, Leu-300 is replaced by Pro-300, and no such H-bond interaction between Leu-300 and carbamoyl group of fidarestat is possible [42]. It supports our finding that the hydrogen bond between atom OC2 of DCEG and Leu-300 is an important factor for selective recognition of glutathione by AR.

Oxidation of Cys-298 has been implicated in the transformation of native AR into an activated form of AR with decreased substrate specificity and reduction in sensitivity of various ARIs [43-44]. C298S mutant of AR has been shown to reproduce several characteristics of oxidized AR [45]. From the present study, it appears that due to increased electronegativity of oxygen in the hydroxyl group as compared to sulfur atom in the thiol group, the hydroxyl group of Ser-298 should remain attracted to backbone amides of Ala-299, Leu-300, and Leu-301 thus permanently blocking the pocket. Relative positions of the hydroxyl group of Ser-298 and thiol group of Cys-298 with respect to backbone nitrogens of Ala-299, Leu-300, and Leu-301 support this observation (Supplementary Information: Figure S7). It supports our finding that the movement of Cys-298 side-chain is crucial for specific recognition of glutathione by AR.

Loop B residues of AR have one of the highest B-factors in most of the crystal structures of AR [46]. The closed and open conformations of loop B are known [37]. Interactions that stabilize the closed and opened state of loop B [Figure 4C and 4D], as well as change in dihedral angles of key residues (Gly-213, Ser-214, Ser-226) from closed to open state observed in this study (Supplementary Information: Figure S8) are similar to ones previously reported [37].

The present study revealed an intermediate state between the closed and opened state of loop B. Intermediate state is characterized by π-stacking interaction between Trp-219 and Trp-295. We found that the intermediate state is a requisite for the transition of loop B from closed state to opened state. Due to a lack of proper orientation of Trp-219 and Trp-295, loop B was unable to move from closed state to open state in MD trajectory with DCEG. Closed state of loop B is stabilized by a sulfur aromatic interaction between Trp219 and Cys-298. The inability of Trp-219 to move away from Cys-298 (Figure 4E) may be the reason that loops B is stuck in intermediate conformation in MD trajectory with DCEG.

The binding pocket of AR is unusually large and has significant hydrophobicity which allows the enzyme to catalyze a wide variety of aldehydes and ketones. AR is a long-standing drug target for secondary diabetic complications as a part of polyol pathway. The physiological function of AR is to reduce lipid-peroxidation generated aldehydes and their glutathione conjugates to respective alcohols. This study has provided the details of interaction and conformational changes that takes place during recognition of glutathione analog, DCEG, in the binding pocket of AR.

In summary, our study has revealed that DCEG, a glutathione analog is specifically recognized by a tiny pocket made up of loop C residues Cys-298, Ala-299, Leu-300, and Leu301 where Cys-298 acts as a gatekeeper. This knowledge will aid in design of inhibitors which shall preserve the detoxifying activity of aldose reductase.

ACKNOWLEDGEMENTS

Gurprit Sekhon acknowledges Research Fellowship from UGC, Govt. of INDIA.

REFERENCES

1. Brownlee, M. The Pathobiology of Diabetic Complications: A Unifying Mechanism. Diabetes 2005, 54 (6), 1615–1625. https://doi.org/10.2337/diabetes.54.6.1615.
2. Yan, L. Redox Imbalance Stress in Diabetes Mellitus: Role of the Polyol Pathway. Animal Models and Experimental Medicine 2018, 1 (1), 7–13. https://doi.org/10.1002/ame2.12001.
3. González, R. G.; Barnett, P.; Aguayo, J.; Cheng, H.-M.; Chylack, L. T. Direct Measurement of Polyol Pathway Activity in the Ocular Lens. Diabetes 1984, 33 (2), 196–199. https://doi.org/10.2337/diab.33.2.196.

4. Yabe-Nishimura, C. Aldose Reductase in Glucose Toxicity: A Potential Target for the Prevention of Diabetic Complications. Pharmacol Rev 1998, 50 (1), 21–33.

5. Giacco, F.; Brownlee, M. Oxidative Stress and Diabetic Complications. Circ Res 2010, 107 (9), 1058–1070. https://doi.org/10.1161/CIRCRESAHA.110.223545.

6. Gabbay, K. H. Aldose Reductase Inhibition in the Treatment of Diabetic Neuropathy: Where Are We in 2004? Curr Diab Rep 2004, 4 (6), 405–408. https://doi.org/10.1007/s11892-004-0047-z.

7. Hotta, N.; Akanuma, Y.; Kawamori, R.; Matsuoka, K.; Oka, Y.; Shichiri, M.; Toyota, T.; Nakashima, M.; Yoshimura, I.; Sakamoto, N.; Shigeta, Y.; the ADCT Study Group. Long-Term Clinical Effects of Epalrestat, an Aldose Reductase Inhibitor, on Diabetic Peripheral Neuropathy: The 3-Year, Multicenter, Comparative Aldose Reductase Inhibitor-Diabetes Complications Trial. Diabetes Care 2006, 29 (7), 1538–1544. https://doi.org/10.2337/dc05-2370.

8. Hyndman, D.; Bauman, D. R.; Heredia, V. V.; Penning, T. M. The Aldo-Keto Reductase Superfamily Homepage. Chemico-Biological Interactions 2003, 143–144, 621–631. https://doi.org/10.1016/S0009-2797(02)00193-X.

9. Penning, T. M. The Aldo-Keto Reductases (AKRs): Overview. Chemico-Biological Interactions 2015, 234, 236–246. https://doi.org/10.1016/j.cbi.2014.09.024.

10. Petrash, J. M. All in the Family: Aldose Reductase and Closely Related Aldo-Keto Reductases. Cellular and Molecular Life Sciences (CMLS) 2004, 61 (7–8), 737–749. https://doi.org/10.1007/s00018-003-3402-3.

11. Maccari, R.; Ottanà, R. Targeting Aldose Reductase for the Treatment of Diabetes Complications and Inflammatory Diseases: New Insights and Future Directions. J. Med. Chem. 2015, 58 (5), 2047–2067. https://doi.org/10.1021/jm500907a.

12. Wilson, D. K.; Bohren, K. M.; Gabbay, K. H.; Quiocho, F. A. An Unlikely Sugar Substrate Site in the 1.65 A Structure of the Human Aldose Reductase Holoenzyme Implicated in Diabetic Complications. Science 1992, 257 (5066), 81–84. https://doi.org/10.1126/science.1621098.

13. Klebe, G.; Krämer, O.; Sotriffer, C. Strategies for the Design of Inhibitors of Aldose Reductase, an Enzyme Showing Pronounced Induced-Fit Adaptations. CMLS, Cell. Mol. Life Sci. 2004, 61 (7), 783–793. https://doi.org/10.1007/s00018-003-3406-z.

14. Rechlin, C.; Scheer, F.; Terwesten, F.; Wulsdorf, T.; Pol, E.; Fridh, V.; Toth, P.; Diederich, W. E.; Heine, A.; Klebe, G. Price for Opening the Transient Specificity Pocket in Human Aldose Reductase upon Ligand Binding: Structural, Thermodynamic, Kinetic, and Computational Analysis. ACS Chem. Biol. 2017, 12 (5), 1397–1415. https://doi.org/10.1021/acschembio.7b00062.

15. Grimshaw, C. E. Aldose Reductase: Model for a New Paradigm of Enzymic Perfection in Detoxification Catalysts. Biochemistry 1992, 31 (42), 10139–10145. https://doi.org/10.1021/bi00157a001.

16. Bander Jagt, D. L.; Kolb, N. S.; bander Jagt, T. J.; Chino, J.; Martinez, F. J.; Hunsaker, L. A.; Royer, R. E. Substrate Specificity of Human Aldose Reductase: Identification of 4-Hydroxynonenal as an Endogenous Substrate. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1995, 1249 (2), 117–126. https://doi.org/10.1016/0167-4838(95)00021-L.

17. Ramana, K. V.; Dixit, B. L.; Srivastava, S.; Balendiran, G. K.; Srivastava, S. K.; Bhatnagar, A. Selective Recognition of Glutathiolated Aldehydes by Aldose Reductase. Biochemistry 2000, 39 (40), 12172–12180. https://doi.org/10.1021/bi000796e.

18. Vander Jagt, D. L.; Hassebrook, R. K.; Hunsaker, L. A.; Brown, W. M.; Royer, R. E. Metabolism of the 2-Oxoaldehyde Methylglyoxal by Aldose Reductase and by Glyoxalase-I: Roles for Glutathione in Both Enzymes and Implications for Diabetic Complications. Chemico-Biological Interactions 2001, 130–132, 549–562. https://doi.org/10.1016/S0009-2797(00)00298-2.

19. Dixit, B. L.; Balendiran, G. K.; Watowich, S. J.; Srivastava, S.; Ramana, K. V.; Petrash, J. M.; Bhatnagar, A.; Srivastava, S. K. Kinetic and Structural Characterization of the Glutathione-Binding Site of Aldose Reductase. J. Biol. Chem. 2000, 275 (28), 21587–21595. https://doi.org/10.1074/jbc.M909235199.

20. Singh, R.; White, M. A.; Ramana, K. V.; Petrash, J. M.; Watowich, S. J.; Bhatnagar, A.; Srivastava, S. K. Structure of a Glutathione Conjugate Bound to the Active Site of Aldose Reductase. Proteins: Structure, Function, and Bioinformatics 2006, 64 (1), 101–110. https://doi.org/10.1002/prot.20988.

21. DA Case, IY. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, VWD Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman (2018), AMBER 2018, University of California, San Francisco.

22. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating PK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. Nucleic Acids Res 2012, 40 (W1), W537–W541. https://doi.org/10.1093/nar/gks375.

23. Várnai, P.; Warshel, A. Computer Simulation Studies of the Catalytic Mechanism of Human Aldose Reductase. J. Am. Chem. Soc. 2000, 122 (16), 3849–3860. https://doi.org/10.1021/ja994246j.

24. Howard, E. I.; Sanishvili, R.; Cachau, R. E.; Mitschler, A.; Chevrier, B.; Barth, P.; Lamour, V.; Zandt, M. V.; Sibley, E.; Bon, C.; et al. Ultrahigh Resolution Drug Design I: Details of Interactions in Human Aldose Reductase–Inhibitor Complex at 0.66 Å. Proteins: Structure, Function, and Bioinformatics 2004, 55 (4), 792–804. https://doi.org/10.1002/prot.20015.

25. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. J Chem Theory Comput 2015, 11 (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

26. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. Journal of Computational Chemistry 2004, 25 (9), 1157–1174. https://doi.org/10.1002/jcc.20035.

27. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. Journal of Computational Chemistry 2000, 21 (2), 132–146. https://doi.org/10.1002/(SICI)1096-987X(20000130)21:2<132::AID-JCC5>3.0. CO;2-P.

28. Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. Journal of Computational Chemistry 2002, 23 (16), 1623–1641. https://doi.org/10.1002/jcc.10128.

29. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. The Journal of Chemical Physics 1983, 79 (2), 926–935. https://doi.org/10.1063/1.445869.

30. Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin Dynamics of Peptides: The Frictional Dependence of Isomerization Rates of N-Acetylalanyl-N′-Methylamide. Biopolymers 1992, 32 (5), 523–535. https://doi.org/10.1002/bip.360320508.

31. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. J. Chem. Phys. 1984, 81 (8), 3684–3690. https://doi.org/10.1063/1.448118.

32. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. Journal of Computational Physics 1977, 23 (3), 327–341. https://doi.org/10.1016/0021-9991(77)90098-5.

33. Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J. Chem. Theory Comput. 2013, 9 (7), 3084–3095. https://doi.org/10.1021/ct400341p.

34. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. Journal of Molecular Graphics 1996, 14 (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

35. Altis, A.; Nguyen, P. H.; Hegger, R.; Stock, G. Dihedral Angle Principal Component Analysis of Molecular Dynamics Simulations. The Journal of Chemical Physics 2007, 126 (24), 244111. https://doi.org/10.1063/1.2746330.

36. David Arthur and Sergei Vassilvitskii. 2007. K-means++: the advantages of careful seeding. In Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms (SODA '07). Society for Industrial and Applied Mathematics, USA, 1027–1035.

37. Bohren, K. M.; Brownlee, J. M.; Milne, A. C.; Gabbay, K. H.; Harrison, D. H. T. The Structure of Apo R268A Human Aldose Reductase: Hinges and Latches That Control the Kinetic Mechanism. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 2005, 1748 (2), 201–212. https://doi.org/10.1016/j.bbapap.2005.01.006.

38. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, Colagiuri S, Guariguata L, Motala AA, Ogurtsova K, Shaw JE, Bright D, Williams R; IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. Diabetes Res Clin Pract. 2019 Nov;157:107843. doi: 10.1016/j.diabres.2019.107843. Epub 2019 Sep 10. PMID: 31518657.

39. Perfetti, R.; Yepuri, G.; Quadri, N. A.; Ramasamy, R.; Ghannam, A. F.; Shendelman, S. 462-P: AT-001, a Next Generation Aldose Reductase Inhibitor with Improved Selectivity and Specificity, Protects from Cellular Damage Associated with Hyperglycemia. Diabetes 2020, 69 (Supplement 1), 462-P. https://doi.org/10.2337/db20-462-P.

40. Oka, M.; Matsumoto, Y.; Sugiyama, S.; Tsuruta, N.; Matsushima, M. A Potent Aldose Reductase Inhibitor, (2S,4S)-6-Fluoro-2‘,5‘-Dioxospiro[Chroman-4,4‘-Imidazolidine]-2-Carboxamide (Fidarestat): Its Absolute Configuration and Interactions with the Aldose Reductase by X-Ray Crystallography. J. Med. Chem. 2000, 43 (12), 2479–2483. https://doi.org/10.1021/jm990502r.

41. El-Kabbani, O.; Darmanin, C.; Oka, M.; Schulze-Briese, C.; Tomizaki, T.; Hazemann, I.; Mitschler, A.; Podjarny, A. High-Resolution Structures of Human Aldose Reductase Holoenzyme in Complex with Stereoisomers of the Potent Inhibitor Fidarestat: Stereospecific Interaction between the Enzyme and a Cyclic Imide Type Inhibitor. J. Med. Chem. 2004, 47 (18), 4530–4537. https://doi.org/10.1021/jm0497794.

42. El-Kabbani, O.; Carbone, V.; Darmanin, C.; Oka, M.; Mitschler, A.; Podjarny, A.; Schulze-Briese, C.; Chung, R. P.-T. Structure of Aldehyde Reductase Holoenzyme in Complex with the Potent Aldose Reductase Inhibitor Fidarestat: Implications for Inhibitor Binding and Selectivity. J. Med. Chem. 2005, 48 (17), 5536–5542. https://doi.org/10.1021/jm050412o.

43. Grimshaw, C. E.; Bohren, K. M.; Lai, C.-J.; Gabbay, K. H. Human Aldose Reductase: Rate Constants for a Mechanism Including Interconversion of Ternary Complexes by Recombinant Wild-Type Enzyme. Biochemistry 1995, 34 (44), 14356–14365. https://doi.org/10.1021/bi00044a012.

44. Grimshaw, C. E.; Lai, C.-J. Oxidized Aldose Reductase:In VivoFactor, Notin VitroArtifact. Archives of Biochemistry and Biophysics 1996, 327 (1), 89–97. https://doi.org/10.1006/abbi.1996.0096.

45. Balendiran, G. K.; Sawaya, M. R.; Schwarz, F. P.; Ponniah, G.; Cuckovich, R.; Verma, M.; Cascio, D. The Role of Cys-298 in Aldose Reductase Function. J. Biol. Chem. 2011, 286 (8), 6336–6344. https://doi.org/10.1074/jbc.M110.154195.

46. Balendiran, G. K.; Pandian, J. R.; Drake, E.; Vinayak, A.; Verma, M.; Cascio, D. B-Factor Analysis and Conformational Rearrangement of Aldose Reductase. Curr Proteomics 2014, 11 (3), 151–160. https://doi.org/10.2174/15701646110314092216344 4.

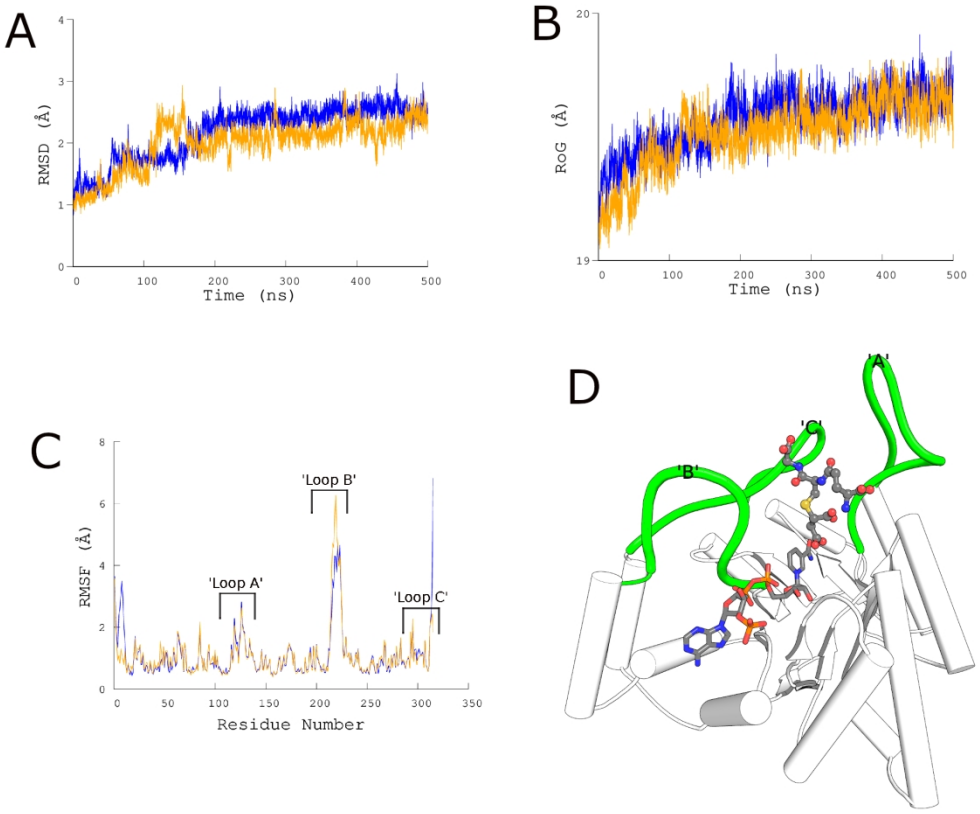

Figure 1. Molecular dynamics trajectories of the ternary complex of AR bound to NADPH, with DCEG (Orange) and without DCEG (Blue). A: Root-mean-square deviation of the protein backbone. B: Radius of gyration. C: Per-residue average root mean square fluctuation. D: Structure of aldose reductase with bound NADPH (stick model) and DCEG (ball-and-stick model). Active site loops (Loop A, aa: 109-135; Loop B, aa: 213-226; and Loop C, aa: 291-315) are colored Green.

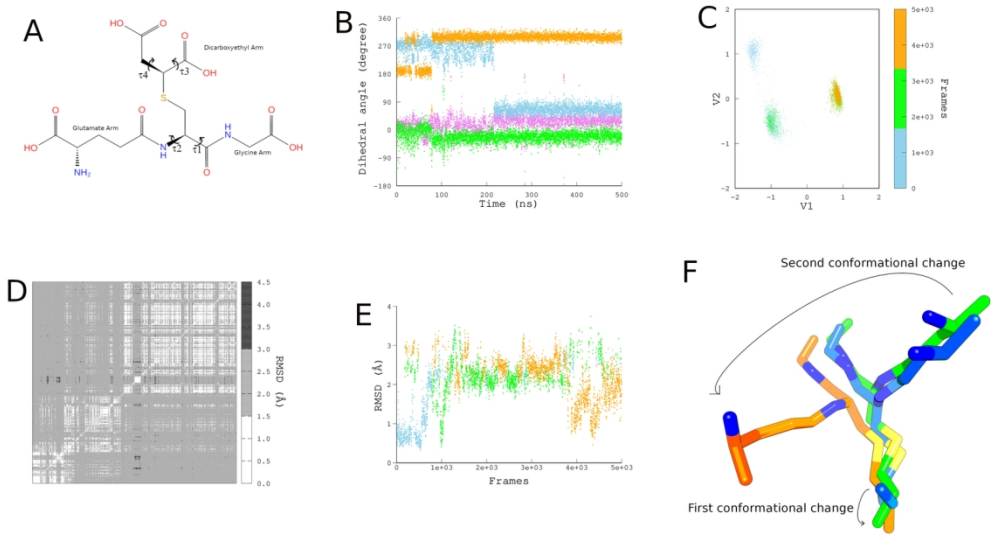

Figure 2. Dynamics of γ-glutamyl-S-(1,2-di-carboxyethyl)-cysteinyl-glycine (DCEG). A: Four dihedrals to monitor the dynamics of three arms of DCEG. B: Dihedrals τ1 (Violet), τ2 (Skyblue), τ3 (Green), and τ4 (Orange) from MD trajectory of AR•NADPH•DCEG. C: Plot of first two eigenvectors, V1 and V2, obtained from dihedral principal component analysis (dPCA) of MD trajectory based on four dihedrals, τ1- τ4. D: 2D RMSD plot of DCEG. E: Clustering of MD trajectory based on RMSD of DCEG. Frames belonging to clusters 1-3 are colored Orange, Green, and Skyblue, respectively. F: Superposition of representative DCEG conformations obtained from clustering (color-code according to Figure 2E) shows conformational changes during MD simulation.

141x77mm (300 x 300 DPI)

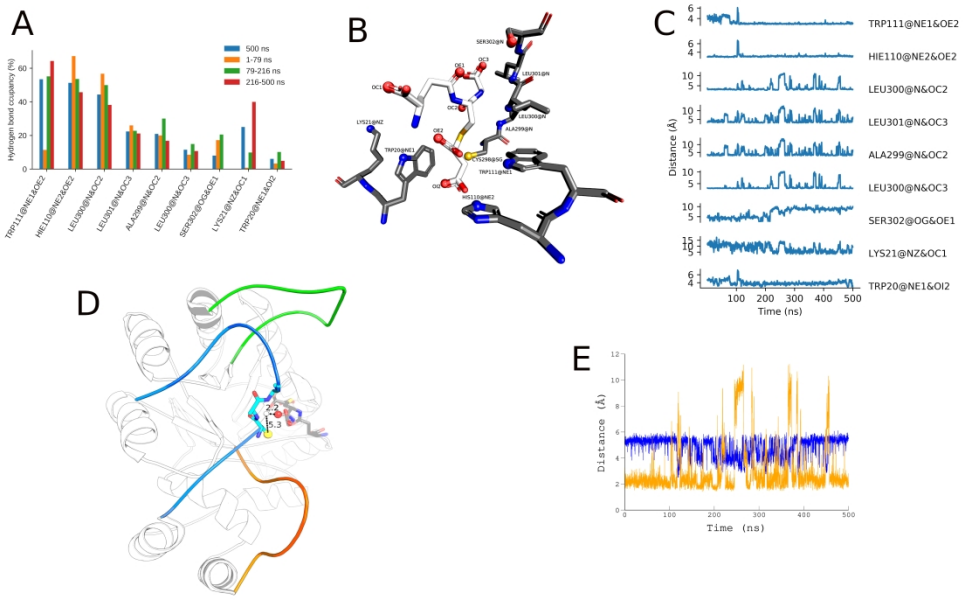

Figure 3. Interactions between DCEG and AR. A: Hydrogen bond interactions between DCEG and AR form the MD trajectory. B: Atoms involved in hydrogen bond interactions are shown as spheres and labeled in the crystal structure of AR•NADPH•DCEG. C: Distances between acceptor and donor atoms of hydrogen bond interactions, as calculated from MD trajectory. D: Distance of center-of-mass of backbone nitrogens of Ala-299 and Leu-300 (White sphere) from sulfur atom of Cys-298 (Yellow sphere) and OC2 atom of DCEG (Red sphere) as calculated in the crystal structure of AR•NADPH•DCEG. E: Distance of center-of-mass of backbone nitrogens of Ala-299 and Leu-300 from Cys-298 side-chain (blue) and OC2 atom of DCEG (Orange) as calculated from MD trajectory.

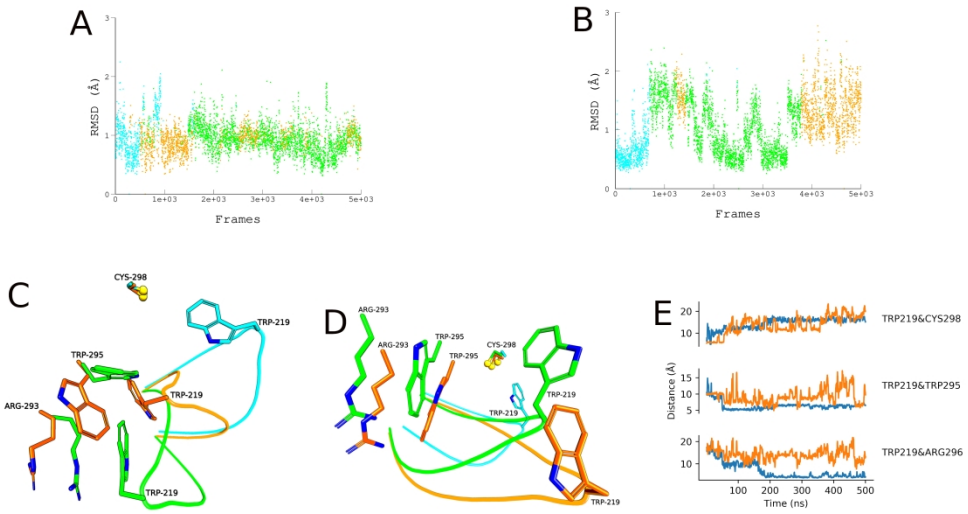

Figure 4. Dynamics of loop B of AR. A&B: Clustering of MD trajectories of the ternary complex (AR•NADPH•DCEG) without DCEG and with DCEG, respectively, based on RMSD of loop B. C&D: Superposition of representative conformations of loop B from MD trajectories without DCEG and with DCEG, respectively, as obtained from clustering. E. Distance between interacting residues that stabilize the closed (indole ring of Trp-219 and sulfur atom of Cys-298), intermediate (indole ring of Trp-219 and indole ring of Trp-295) and opened state (indole ring of Trp-219, and guanidine group of Arg-296) of loop B in MD trajectory without DCEG (Blue) and with DCEG (Orange).

Supplementary Information for

Role of Cys-298 in specific recognition of glutathione by aldose reductase

Gurprit Sekhon[1], Balvinder Singh[2*], Ranvir Singh[1*]

[1]Department cum National Centre for Human Genome Studies & Research, Pharmacy Extention Block, Panjab University, Chandigarh 160014, India.

[2]Bioinformatics Center, Institute of Microbial Technology, Council of Scientific and Industrial Research, Sector 39A, Chandigarh 160036, India.

[*]Corresponding Authors: Email: ranvir1@pu.ac.in (Ranvir Singh) or bvs@imtech.res.in (Balvinder Singh)

This pdf file includes:

- ✓ Figures: S1-S8
- ✓ Tables: S1-S7
- ✓ Description of Movie S1

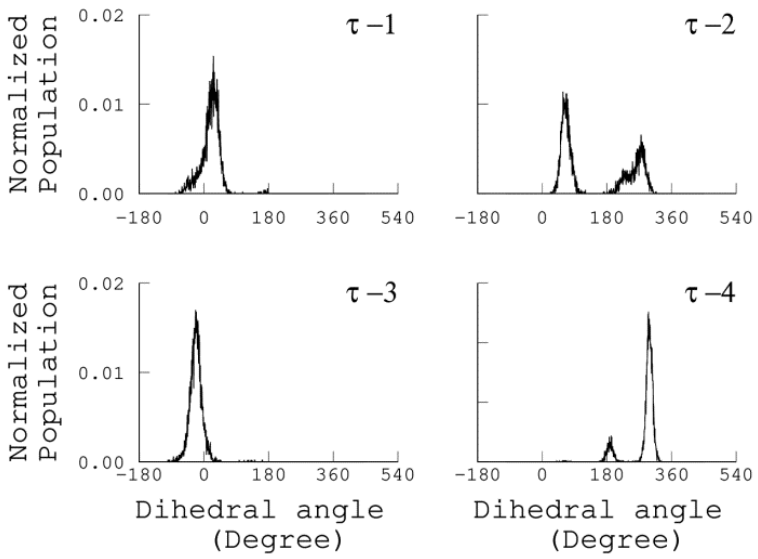

Figure S1. Normalized histograms of four selected dihedral angles of DCEG, τ1- τ4.

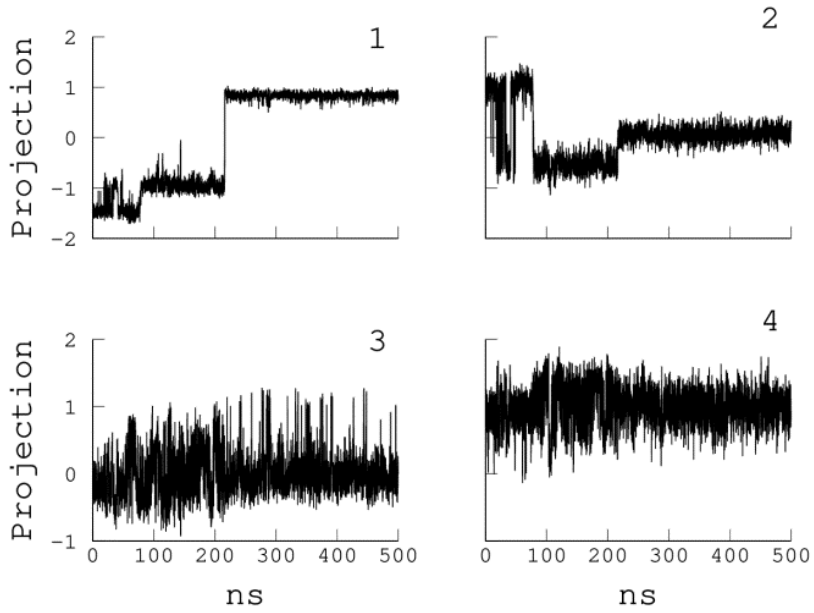

Figure S2. First four eigenvectors obtained from dihedral principal component analysis (dPCA) of MD trajectory.

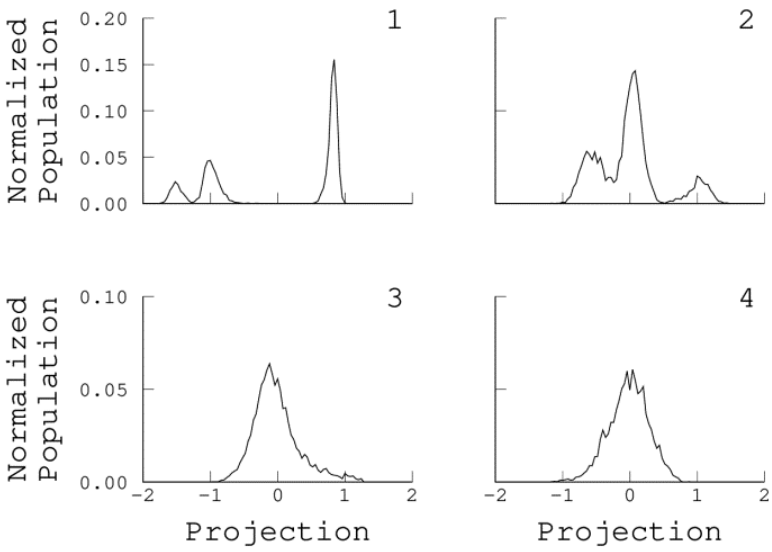

Figure S3. Normalized histograms of four eigenvectors obtained from dihedral principal component analysis (dPCA) of MD trajectory.

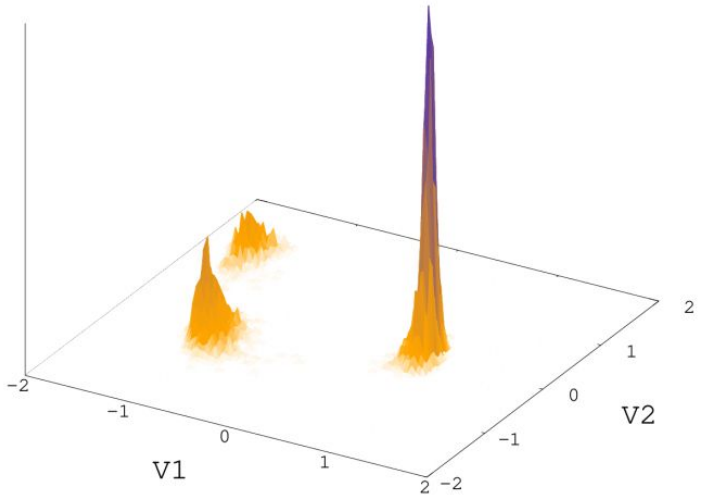

Figure S4. 2D histogram of first eigenvector against second eigenvector.

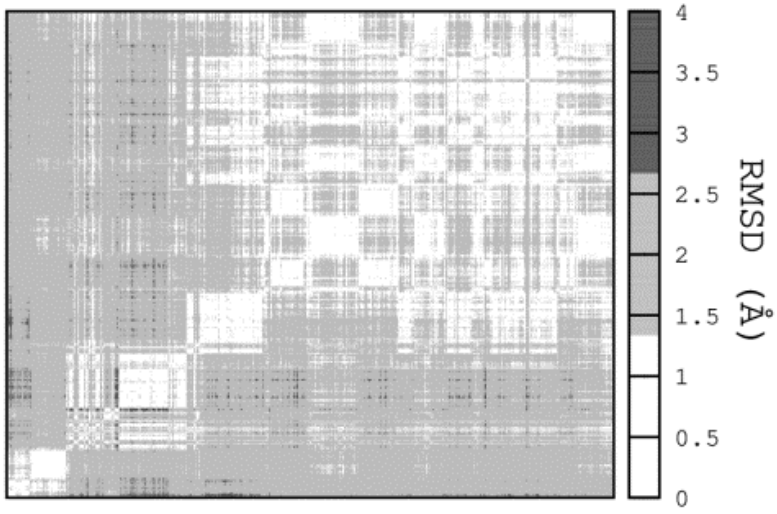

Figure S5. 2D RMSD plot of loop B for MD trajectory of AR bound to NADPH.

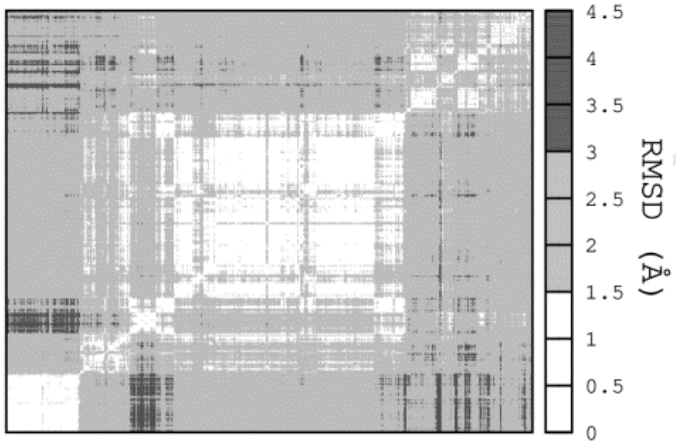

Figure S6. 2D RMSD plot of loop B for MD trajectory of AR bound to NADPH and DCEG.

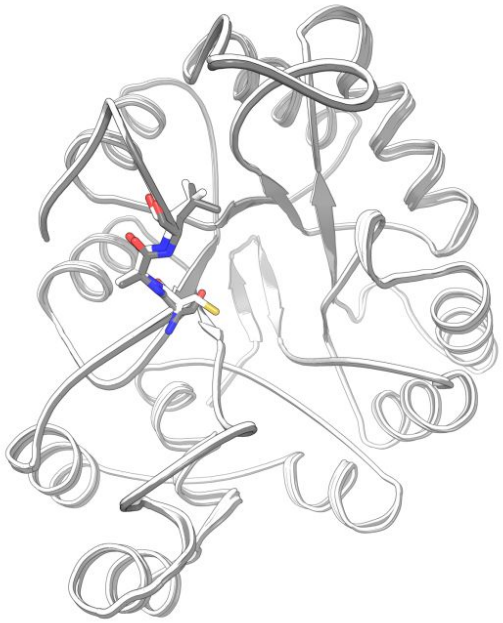

Figure S7. Structural superposition of wild type AR (White) and Cys298Ser mutant of AR (Grey) show relative position of side-chain Sulphur of Cys-298 and side-chain Oxygen of Ser-298. [Atomic coordinates were obtained from RCSB-PDB (PDBID: 3Q65 and 3Q67), "Balendiran, G. K.; Sawaya, M. R.; Schwarz, F. P.; Ponniah, G.; Cuckovich, R.; Verma, M.; Cascio, D. The Role of Cys-298 in Aldose Reductase Function. J. Biol. Chem. 2011, 286 (8), 6336–6344."]

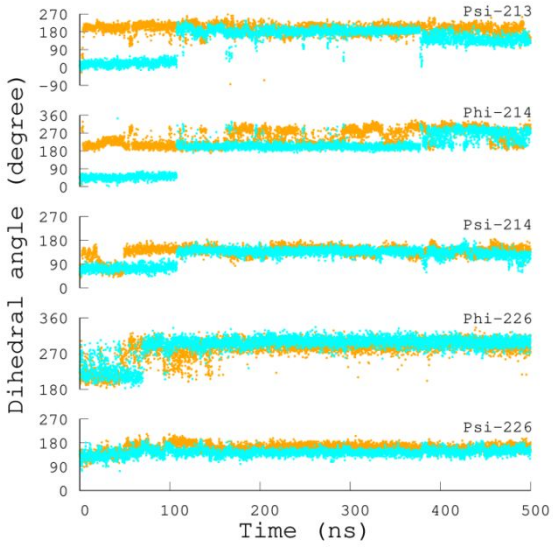

Figure S8. Dihedral angles of key residues (Gly-213, Ser-214 and Ser-226) for MD trajectory of AR bound to NADPH, without DCEG (Orange) and with DCEG (Cyan).

Table S1. Cluster analysis of MD trajectory of AR●NADPH●DCEG on the basis of RMSD of DCEG.

Cluster Number	Frames	Fraction	Average Distance	Standard Deviation	Centroid	Average Cdistance
1	2444	0.489	1.451	0.371	4397	1.350
2	1606	0.321	1.804	0. 487	989	1.312
3	950	0.190	1.574	0.398	510	1.211

[1]Number of Frames: 5000; Number of Clusters: 3; Algorithm: K-means; Maximum Iterations: 1000; DBI: 1.94; pSF: 1013.81; SSR/SST: 0.29.

[2]Average Cdistance: Average distance of every frame from Centroid in a cluster.

[3]All distances are measured in Angstrom (Å).

Table S2. Hydrogen bond analysis between DCEG and AR for total time interval i.e 1-500 ns.

Sr. No.	#Acceptor	DonorH	Donor	Frames	Frac	AvgDist	AvgAng
1	TGG_317@OE2	TRP_111@HE1	TRP_111@NE1	2672	0.53	2.86	158.45
2	TGG_317@OE2	HIE_110@HE2	HIE_110@NE2	2564	0.51	2.86	127.66
3	TGG_317@OC2	LEU_300@H	LEU_300@N	2217	0.44	2.88	153.83
4	TGG_317@OC3	LEU_301@H	LEU_301@N	1118	0.22	2.89	152.73
5	TGG_317@OC2	ALA_299@H	ALA_299@N	1047	0.21	2.88	140.00
6	TGG_317@OC3	LEU_300@H	LEU_300@N	575	0.12	2.90	107.70
7	TGG_317@OC1	LYS_21@HZ1	LYS_21@NZ	425	0.09	2.85	90.77
8	TGG_317@OC1	LYS_21@HZ2	LYS_21@NZ	425	0.09	2.85	81.90
9	TGG_317@OC1	LYS_21@HZ3	LYS_21@NZ	425	0.09	2.85	85.72
10	TGG_317@OE1	SER_302@HG	SER_302@OG	417	0.08	2.72	162.69
11	TGG_317@OC3	ALA_299@H	ALA_299@N	383	0.08	2.87	145.11
12	TGG_317@OT3	LEU_301@H	LEU_301@N	355	0.07	2.91	142.82
13	LEU_301@N	TGG_317@HT3	TGG_317@OT3	355	0.07	2.91	39.87
14	TGG_317@OI2	TRP_20@HE1	TRP_20@NE1	303	0.06	2.88	143.24
15	VAL_297@O	TGG_317@HT3	TGG_317@OT3	210	0.04	2.75	147.97
16	TGG_317@OI2	HIE_110@HE2	HIE_110@NE2	202	0.04	2.90	127.52
17	LYS_21@NZ	TGG_317@HT1	TGG_317@OT1	196	0.04	2.88	29.84
18	TGG_317@OT1	LYS_21@HZ1	LYS_21@NZ	196	0.04	2.88	88.10
19	TGG_317@OT1	LYS_21@HZ2	LYS_21@NZ	196	0.04	2.88	90.49
20	TGG_317@OT1	LYS_21@HZ3	LYS_21@NZ	196	0.04	2.88	80.59
21	ASP_216@OD2	TGG_317@HT3	TGG_317@OT3	165	0.03	2.63	164.14
22	TGG_317@OJ2	TRP_20@HE1	TRP_20@NE1	161	0.03	2.91	135.85
23	TRP_20@NE1	TGG_317@HJ2	TGG_317@OJ2	161	0.03	2.91	55.86
24	ASP_216@OD1	TGG_317@HT3	TGG_317@OT3	156	0.03	2.64	162.80
25	TGG_317@OC1	TRP_20@HE1	TRP_20@NE1	154	0.03	2.88	149.86
26	TGG_317@OC1	SER_302@HG	SER_302@OG	135	0.03	2.77	158.68
27	TGG_317@OC1	LEU_300@H	LEU_300@N	85	0.02	2.88	147.61
28	TGG_317@OC1	ALA_299@H	ALA_299@N	79	0.02	2.84	154.90
29	TGG_317@OC3	SER_302@HG	SER_302@OG	77	0.02	2.77	157.59
30	TGG_317@OT3	LEU_300@H	LEU_300@N	69	0.01	2.93	105.20
31	LEU_300@N	TGG_317@HT3	TGG_317@OT3	69	0.01	2.93	52.79
32	TGG_317@OC3	SER_302@H	SER_302@N	58	0.01	2.90	153.10
33	TGG_317@OC3	LYS_21@HZ2	LYS_21@NZ	48	0.01	2.87	84.36
34	TGG_317@OC3	LYS_21@HZ3	LYS_21@NZ	48	0.01	2.87	100.07
35	TGG_317@OC3	LYS_21@HZ1	LYS_21@NZ	48	0.01	2.87	79.88
36	ALA_299@N	TGG_317@HT3	TGG_317@OT3	32	0.01	2.91	59.16
37	TGG_317@OT3	ALA_299@H	ALA_299@N	32	0.01	2.91	145.75
38	TGG_317@NN1	LYS_21@HZ1	LYS_21@NZ	31	0.01	2.87	86.91

39	TGG_317@NN1	LYS_21@HZ2	LYS_21@NZ	31	0.01	2.87	77.32
40	TGG_317@NN1	LYS_21@HZ3	LYS_21@NZ	31	0.01	2.87	91.43
41	LYS_21@NZ	TGG_317@HN12	TGG_317@NN1	31	0.01	2.87	61.78
42	LYS_21@NZ	TGG_317@HN11	TGG_317@NN1	31	0.01	2.87	56.35
43	SER_302@N	TGG_317@HT3	TGG_317@OT3	24	0.00	2.93	52.15
44	TGG_317@OT3	SER_302@H	SER_302@N	24	0.00	2.93	143.35
45	TGG_317@OI2	TRP_111@HE1	TRP_111@NE1	22	0.00	2.87	154.06
46	ASP_216@O	TGG_317@HT3	TGG_317@OT3	21	0.00	2.72	139.01
47	TGG_317@OT1	LEU_301@H	LEU_301@N	21	0.00	2.92	141.43
48	LEU_301@N	TGG_317@HT1	TGG_317@OT1	21	0.00	2.92	40.95
49	TGG_317@NN1	SER_302@HG	SER_302@OG	20	0.00	2.89	83.68
50	SER_302@OG	TGG_317@HN11	TGG_317@NN1	20	0.00	2.89	91.79
51	SER_302@OG	TGG_317@HN12	TGG_317@NN1	20	0.00	2.89	97.91
52	TRP_111@NE1	TGG_317@HF2	TGG_317@OF2	20	0.00	2.94	68.13
53	TGG_317@OF2	TRP_111@HE1	TRP_111@NE1	20	0.00	2.94	137.97
54	TGG_317@OT1	SER_302@HG	SER_302@OG	18	0.00	2.85	122.54
55	SER_302@OG	TGG_317@HT1	TGG_317@OT1	18	0.00	2.85	74.60
56	TRP_20@NE1	TGG_317@HT1	TGG_317@OT1	18	0.00	2.93	44.18
57	TGG_317@OT1	TRP_20@HE1	TRP_20@NE1	18	0.00	2.93	133.56
58	TGG_317@OC2	LYS_21@HZ3	LYS_21@NZ	14	0.00	2.82	89.44
59	TGG_317@OC2	LYS_21@HZ2	LYS_21@NZ	14	0.00	2.82	86.81
60	TGG_317@OC2	LYS_21@HZ1	LYS_21@NZ	14	0.00	2.82	89.00
61	TGG_317@OC1	LEU_301@H	LEU_301@N	14	0.00	2.92	129.13
62	TGG_317@OT3	LYS_21@HZ2	LYS_21@NZ	12	0.00	2.88	92.28
63	TGG_317@OT3	LYS_21@HZ3	LYS_21@NZ	12	0.00	2.88	75.22
64	LYS_21@NZ	TGG_317@HT3	TGG_317@OT3	12	0.00	2.88	31.36
65	TGG_317@OT3	LYS_21@HZ1	LYS_21@NZ	12	0.00	2.88	92.18
66	ASP_216@OD2	TGG_317@HT1	TGG_317@OT1	8	0.00	2.68	150.96
67	SER_302@OG	TGG_317@HT3	TGG_317@OT3	8	0.00	2.87	73.36
68	TGG_317@OT3	SER_302@HG	SER_302@OG	8	0.00	2.87	106.40
69	ASP_216@OD2	TGG_317@HN12	TGG_317@NN1	7	0.00	2.80	130.24
70	ASP_216@OD2	TGG_317@HN11	TGG_317@NN1	7	0.00	2.80	80.50
71	ASP_216@OD1	TGG_317@HN12	TGG_317@NN1	4	0.00	2.86	123.58
72	ASP_216@OD1	TGG_317@HN11	TGG_317@NN1	4	0.00	2.86	86.21
73	VAL_297@O	TGG_317@HN12	TGG_317@NN1	4	0.00	2.89	84.20
74	VAL_297@O	TGG_317@HN11	TGG_317@NN1	4	0.00	2.89	135.53
75	TGG_317@OC1	SER_302@H	SER_302@N	4	0.00	2.90	145.26
76	TGG_317@OC3	ARG_217@HH21	ARG_217@NH2	3	0.00	2.85	55.53
77	TGG_317@OC3	ARG_217@HH22	ARG_217@NH2	3	0.00	2.85	150.35
78	TGG_317@OT1	ALA_299@H	ALA_299@N	3	0.00	2.91	144.73
79	ALA_299@N	TGG_317@HT1	TGG_317@OT1	3	0.00	2.91	33.90

80	TGG_317@OT1	LEU_300@H	LEU_300@N	3	0.00	2.93	114.08
81	LEU_300@N	TGG_317@HT1	TGG_317@OT1	3	0.00	2.93	18.80
82	TGG_317@OJ2	HIE_110@HE2	HIE_110@NE2	3	0.00	2.94	108.23
83	HIE_110@NE2	TGG_317@HJ2	TGG_317@OJ2	3	0.00	2.94	39.65
84	TGG_317@OT3	TRP_219@HE1	TRP_219@NE1	3	0.00	2.95	124.76
85	TRP_219@NE1	TGG_317@HT3	TGG_317@OT3	3	0.00	2.95	53.69
86	ASP_216@OD1	TGG_317@HT1	TGG_317@OT1	2	0.00	2.72	171.90
87	TGG_317@OC1	SER_214@HG	SER_214@OG	2	0.00	2.72	159.38
88	VAL_297@O	TGG_317@HT1	TGG_317@OT1	2	0.00	2.91	89.81
89	VAL_297@O	TGG_317@HN3	TGG_317@NN3	2	0.00	2.97	93.23
90	TGG_317@OT1	SER_214@HG	SER_214@OG	1	0.00	2.73	166.10
91	SER_214@OG	TGG_317@HT1	TGG_317@OT1	1	0.00	2.73	33.85
92	TGG_317@OE1	LYS_21@HZ3	LYS_21@NZ	1	0.00	2.75	47.65
93	TGG_317@OE1	LYS_21@HZ1	LYS_21@NZ	1	0.00	2.75	165.88
94	TGG_317@OE1	LYS_21@HZ2	LYS_21@NZ	1	0.00	2.75	58.68
95	TGG_317@OC3	ARG_217@HH12	ARG_217@NH1	1	0.00	2.75	172.03
96	TGG_317@OC3	ARG_217@HH11	ARG_217@NH1	1	0.00	2.75	42.65
97	TGG_317@NN1	ALA_299@H	ALA_299@N	1	0.00	2.86	132.67
98	ALA_299@N	TGG_317@HN12	TGG_317@NN1	1	0.00	2.86	68.81
99	ALA_299@N	TGG_317@HN11	TGG_317@NN1	1	0.00	2.86	78.99
100	TGG_317@OT1	SER_302@H	SER_302@N	1	0.00	2.88	131.70
101	SER_302@N	TGG_317@HT1	TGG_317@OT1	1	0.00	2.88	31.18
102	TRP_20@NE1	TGG_317@HN12	TGG_317@NN1	1	0.00	2.90	113.79
103	TRP_20@NE1	TGG_317@HN11	TGG_317@NN1	1	0.00	2.90	35.09
104	TGG_317@NN1	TRP_20@HE1	TRP_20@NE1	1	0.00	2.90	144.78
105	ASP_216@OD2	TGG_317@HN3	TGG_317@NN3	1	0.00	2.93	156.04
106	TGG_317@OC2	LEU_301@H	LEU_301@N	1	0.00	2.93	121.00
107	TGG_317@OF2	HIE_110@HE2	HIE_110@NE2	1	0.00	2.95	124.63
108	HIE_110@NE2	TGG_317@HF2	TGG_317@OF2	1	0.00	2.95	70.56
109	TGG_317@OF2	TYR_209@HH	TYR_209@OH	1	0.00	2.98	59.61
110	TYR_209@OH	TGG_317@HF2	TGG_317@OF2	1	0.00	2.98	97.16
111	VAL_47@O	TGG_317@HJ2	TGG_317@OJ2	1	0.00	3.00	107.23

Table S3. Hydrogen bond interactions between DCEG and AR for first time interval i.e 1-79 ns.

Sr. No.	#Acceptor	DonorH	Donor	Frames	Frac	AvgDist	AvgAng
1	TGG_317@OE2	HIE_110@HE2	HIE_110@NE2	531	0.67	2.85	126.66
2	TGG_317@OC2	LEU_300@H	LEU_300@N	448	0.57	2.88	154.51
3	TGG_317@OC3	LEU_301@H	LEU_301@N	205	0.26	2.88	144.51
4	TGG_317@OC2	ALA_299@H	ALA_299@N	158	0.20	2.89	119.40
5	TGG_317@OE1	SER_302@HG	SER_302@OG	136	0.17	2.73	161.89
6	TGG_317@OE2	TRP_111@HE1	TRP_111@NE1	90	0.11	2.87	156.59
7	LEU_301@N	TGG_317@HT3	TGG_317@OT3	72	0.09	2.90	36.49
8	TGG_317@OT3	LEU_301@H	LEU_301@N	72	0.09	2.90	148.07
9	TGG_317@OC3	LEU_300@H	LEU_300@N	67	0.08	2.91	100.12
10	TGG_317@OC3	ALA_299@H	ALA_299@N	46	0.06	2.85	150.02
11	VAL_297@O	TGG_317@HT3	TGG_317@OT3	34	0.04	2.75	152.01
12	TRP_20@NE1	TGG_317@HJ2	TGG_317@OJ2	27	0.03	2.91	80.64
13	TGG_317@OJ2	TRP_20@HE1	TRP_20@NE1	27	0.03	2.91	106.12
14	TGG_317@OI2	TRP_20@HE1	TRP_20@NE1	26	0.03	2.86	139.73
15	LEU_300@N	TGG_317@HT3	TGG_317@OT3	26	0.03	2.93	53.92
16	TGG_317@OT3	LEU_300@H	LEU_300@N	26	0.03	2.93	104.20
17	TRP_111@NE1	TGG_317@HF2	TGG_317@OF2	17	0.02	2.94	66.36
18	TGG_317@OF2	TRP_111@HE1	TRP_111@NE1	17	0.02	2.94	140.49
19	TGG_317@OC1	SER_302@HG	SER_302@OG	13	0.02	2.81	155.91
20	TGG_317@OC3	SER_302@H	SER_302@N	13	0.02	2.92	143.61
21	TGG_317@OC1	TRP_20@HE1	TRP_20@NE1	6	0.01	2.87	146.73
22	SER_302@OG	TGG_317@HN12	TGG_317@NN1	5	0.01	2.86	73.24
23	SER_302@OG	TGG_317@HN11	TGG_317@NN1	5	0.01	2.86	93.53
24	TGG_317@NN1	SER_302@HG	SER_302@OG	5	0.01	2.86	105.03
25	TGG_317@OC3	LYS_21@HZ3	LYS_21@NZ	3	0.00	2.86	78.84
26	TGG_317@OC3	LYS_21@HZ1	LYS_21@NZ	3	0.00	2.86	85.96
27	TGG_317@OC3	LYS_21@HZ2	LYS_21@NZ	3	0.00	2.86	105.25
28	TGG_317@OT3	SER_302@H	SER_302@N	3	0.00	2.96	146.18
29	SER_302@N	TGG_317@HT3	TGG_317@OT3	3	0.00	2.96	48.77
30	TGG_317@OC3	SER_302@HG	SER_302@OG	2	0.00	2.73	150.93
31	TGG_317@OC1	LYS_21@HZ2	LYS_21@NZ	2	0.00	2.85	110.98
32	TGG_317@OC1	LYS_21@HZ3	LYS_21@NZ	2	0.00	2.85	107.20
33	TGG_317@OC1	LYS_21@HZ1	LYS_21@NZ	2	0.00	2.85	55.65
34	TRP_219@NE1	TGG_317@HT3	TGG_317@OT3	2	0.00	2.93	54.96
35	TGG_317@OT3	TRP_219@HE1	TRP_219@NE1	2	0.00	2.93	122.20
36	TGG_317@OI2	HIE_110@HE2	HIE_110@NE2	1	0.00	2.85	113.95
37	TGG_317@OT3	ALA_299@H	ALA_299@N	1	0.00	2.85	153.74
38	ALA_299@N	TGG_317@HT3	TGG_317@OT3	1	0.00	2.85	43.43

39	TGG_317@OT3	LYS_21@HZ1	LYS_21@NZ	1	0.00	2.87	145.95
40	TGG_317@OT3	LYS_21@HZ2	LYS_21@NZ	1	0.00	2.87	64.81
41	LYS_21@NZ	TGG_317@HT3	TGG_317@OT3	1	0.00	2.87	33.22
42	TGG_317@OT3	LYS_21@HZ3	LYS_21@NZ	1	0.00	2.87	33.98
43	TGG_317@OF2	HIE_110@HE2	HIE_110@NE2	1	0.00	2.95	124.63
44	HIE_110@NE2	TGG_317@HF2	TGG_317@OF2	1	0.00	2.95	70.56
45	TYR_209@OH	TGG_317@HF2	TGG_317@OF2	1	0.00	2.98	97.16
46	TGG_317@OF2	TYR_209@HH	TYR_209@OH	1	0.00	2.98	59.61
47	VAL_47@O	TGG_317@HJ2	TGG_317@OJ2	1	0.00	3.00	107.23

Table S4. Hydrogen bond interactions between DCEG and AR for second time-interval i.e 79-216 ns.

Sr. No.	#Acceptor	DonorH	Donor	Frames	Frac	AvgDist	AvgAng
1	TGG_317@OE2	TRP_111@HE1	TRP_111@NE1	756	0.55	2.86	157.04
2	TGG_317@OE2	HIE_110@HE2	HIE_110@NE2	734	0.54	2.86	127.87
3	TGG_317@OC2	LEU_300@H	LEU_300@N	684	0.50	2.88	154.34
4	TGG_317@OC2	ALA_299@H	ALA_299@N	411	0.30	2.87	144.16
5	TGG_317@OC3	LEU_301@H	LEU_301@N	312	0.23	2.88	156.93
6	TGG_317@OE1	SER_302@HG	SER_302@OG	281	0.21	2.72	163.07
7	TGG_317@OC3	LEU_300@H	LEU_300@N	204	0.15	2.90	104.52
8	TGG_317@OI2	TRP_20@HE1	TRP_20@NE1	139	0.10	2.88	145.93
9	TGG_317@OC3	ALA_299@H	ALA_299@N	117	0.09	2.88	143.98
10	LEU_301@N	TGG_317@HT3	TGG_317@OT3	115	0.08	2.90	37.24
11	TGG_317@OT3	LEU_301@H	LEU_301@N	115	0.08	2.90	135.40
12	TGG_317@OJ2	TRP_20@HE1	TRP_20@NE1	82	0.06	2.91	142.41
13	TRP_20@NE1	TGG_317@HJ2	TGG_317@OJ2	82	0.06	2.91	48.10
14	TGG_317@OC1	SER_302@HG	SER_302@OG	69	0.05	2.77	159.30
15	TGG_317@OI2	HIE_110@HE2	HIE_110@NE2	63	0.05	2.90	125.81
16	VAL_297@O	TGG_317@HT3	TGG_317@OT3	46	0.03	2.74	146.70
17	TGG_317@OC1	LYS_21@HZ1	LYS_21@NZ	45	0.03	2.86	85.22
18	TGG_317@OC1	LYS_21@HZ3	LYS_21@NZ	45	0.03	2.86	91.25
19	TGG_317@OC1	LYS_21@HZ2	LYS_21@NZ	45	0.03	2.86	82.34
20	TGG_317@OI2	TRP_111@HE1	TRP_111@NE1	22	0.02	2.87	154.06
21	TGG_317@OT1	LYS_21@HZ3	LYS_21@NZ	21	0.02	2.88	83.89
22	TGG_317@OT1	LYS_21@HZ1	LYS_21@NZ	21	0.02	2.88	78.64
23	TGG_317@OT1	LYS_21@HZ2	LYS_21@NZ	21	0.02	2.88	97.43
24	LYS_21@NZ	TGG_317@HT1	TGG_317@OT1	21	0.02	2.88	28.86
25	TGG_317@OC1	TRP_20@HE1	TRP_20@NE1	19	0.01	2.89	154.83
26	TGG_317@OT3	SER_302@H	SER_302@N	14	0.01	2.94	142.15
27	SER_302@N	TGG_317@HT3	TGG_317@OT3	14	0.01	2.94	58.76
28	TGG_317@OT1	SER_302@HG	SER_302@OG	12	0.01	2.84	107.56
29	SER_302@OG	TGG_317@HT1	TGG_317@OT1	12	0.01	2.84	99.25
30	ASP_216@O	TGG_317@HT3	TGG_317@OT3	11	0.01	2.68	149.51
31	TGG_317@OT3	LEU_300@H	LEU_300@N	10	0.01	2.91	99.39
32	LEU_300@N	TGG_317@HT3	TGG_317@OT3	10	0.01	2.91	47.74
33	TGG_317@OC3	SER_302@H	SER_302@N	9	0.01	2.92	155.23
34	ASP_216@OD2	TGG_317@HT3	TGG_317@OT3	8	0.01	2.67	158.93
35	TGG_317@OT3	ALA_299@H	ALA_299@N	7	0.01	2.91	155.29
36	ALA_299@N	TGG_317@HT3	TGG_317@OT3	7	0.01	2.91	53.58
37	SER_302@OG	TGG_317@HN11	TGG_317@NN1	7	0.01	2.92	80.59

38	SER_302@OG	TGG_317@HN12	TGG_317@NN1	7	0.01	2.92	123.61
39	TGG_317@NN1	SER_302@HG	SER_302@OG	7	0.01	2.92	77.96
40	TGG_317@OC3	LYS_21@HZ3	LYS_21@NZ	4	0.00	2.92	107.21
41	TGG_317@OC3	LYS_21@HZ1	LYS_21@NZ	4	0.00	2.92	85.32
42	TGG_317@OC3	LYS_21@HZ2	LYS_21@NZ	4	0.00	2.92	83.05
43	TGG_317@OT3	LYS_21@HZ1	LYS_21@NZ	3	0.00	2.87	95.28
44	TGG_317@OT3	LYS_21@HZ2	LYS_21@NZ	3	0.00	2.87	93.95
45	LYS_21@NZ	TGG_317@HT3	TGG_317@OT3	3	0.00	2.87	26.08
46	TGG_317@OT3	LYS_21@HZ3	LYS_21@NZ	3	0.00	2.87	62.59
47	ASP_216@OD1	TGG_317@HT3	TGG_317@OT3	2	0.00	2.65	152.06
48	TGG_317@OC3	SER_302@HG	SER_302@OG	2	0.00	2.84	114.82
49	TGG_317@OT3	SER_302@HG	SER_302@OG	2	0.00	2.90	66.50
50	SER_302@OG	TGG_317@HT3	TGG_317@OT3	2	0.00	2.90	79.45
51	TGG_317@OF2	TRP_111@HE1	TRP_111@NE1	2	0.00	2.94	121.58
52	TRP_111@NE1	TGG_317@HF2	TGG_317@OF2	2	0.00	2.94	77.98
53	HIE_110@NE2	TGG_317@HJ2	TGG_317@OJ2	1	0.00	2.87	26.02
54	TGG_317@OJ2	HIE_110@HE2	HIE_110@NE2	1	0.00	2.87	125.82
55	TGG_317@OT1	TRP_20@HE1	TRP_20@NE1	1	0.00	2.91	113.60
56	TRP_20@NE1	TGG_317@HT1	TGG_317@OT1	1	0.00	2.91	38.02
57	TGG_317@OC2	LEU_301@H	LEU_301@N	1	0.00	2.93	121.00
58	LYS_21@NZ	TGG_317@HN12	TGG_317@NN1	1	0.00	2.93	97.00
59	LYS_21@NZ	TGG_317@HN11	TGG_317@NN1	1	0.00	2.93	41.69
60	TGG_317@NN1	LYS_21@HZ3	LYS_21@NZ	1	0.00	2.93	97.20
61	TGG_317@NN1	LYS_21@HZ2	LYS_21@NZ	1	0.00	2.93	116.99
62	TGG_317@NN1	LYS_21@HZ1	LYS_21@NZ	1	0.00	2.93	65.36
63	TRP_219@NE1	TGG_317@HT3	TGG_317@OT3	1	0.00	2.99	51.14
64	TGG_317@OT3	TRP_219@HE1	TRP_219@NE1	1	0.00	2.99	129.88

Table S5. Hydrogen bond interactions between DCEG and AR for third time-interval i.e 216-500 ns

Sr. No.	#Acceptor	DonorH	Donor	Frames	Frac	AvgDist	AvgAng
1	TGG_317@OE2	TRP_111@HE1	TRP_111@NE1	1826	0.64	2.87	159.13
2	TGG_317@OE2	HIE_110@HE2	HIE_110@NE2	1299	0.46	2.87	127.96
3	TGG_317@OC2	LEU_300@H	LEU_300@N	1085	0.38	2.88	153.22
4	TGG_317@OC3	LEU_301@H	LEU_301@N	601	0.21	2.89	153.35
5	TGG_317@OC2	ALA_299@H	ALA_299@N	478	0.17	2.89	143.23
6	TGG_317@OC1	LYS_21@HZ1	LYS_21@NZ	378	0.13	2.85	91.62
7	TGG_317@OC1	LYS_21@HZ3	LYS_21@NZ	378	0.13	2.85	84.95
8	TGG_317@OC1	LYS_21@HZ2	LYS_21@NZ	378	0.13	2.85	81.69
9	TGG_317@OC3	LEU_300@H	LEU_300@N	304	0.11	2.90	111.52
10	TGG_317@OC3	ALA_299@H	ALA_299@N	220	0.08	2.86	144.68
11	TGG_317@OT1	LYS_21@HZ1	LYS_21@NZ	175	0.06	2.88	89.23
12	TGG_317@OT1	LYS_21@HZ3	LYS_21@NZ	175	0.06	2.88	80.19
13	TGG_317@OT1	LYS_21@HZ2	LYS_21@NZ	175	0.06	2.88	89.66
14	LYS_21@NZ	TGG_317@HT1	TGG_317@OT1	175	0.06	2.88	29.96
15	TGG_317@OT3	LEU_301@H	LEU_301@N	168	0.06	2.91	145.66
16	LEU_301@N	TGG_317@HT3	TGG_317@OT3	168	0.06	2.91	43.12
17	ASP_216@OD2	TGG_317@HT3	TGG_317@OT3	157	0.06	2.63	164.41
18	ASP_216@OD1	TGG_317@HT3	TGG_317@OT3	154	0.05	2.64	162.94
19	TGG_317@OI2	TRP_20@HE1	TRP_20@NE1	138	0.05	2.88	141.19
20	TGG_317@OI2	HIE_110@HE2	HIE_110@NE2	138	0.05	2.90	128.39
21	VAL_297@O	TGG_317@HT3	TGG_317@OT3	130	0.05	2.76	147.36
22	TGG_317@OC1	TRP_20@HE1	TRP_20@NE1	129	0.05	2.87	149.28
23	TGG_317@OC1	LEU_300@H	LEU_300@N	85	0.03	2.88	147.61
24	TGG_317@OC1	ALA_299@H	ALA_299@N	79	0.03	2.84	154.90
25	TGG_317@OC3	SER_302@HG	SER_302@OG	73	0.03	2.76	158.95
26	TGG_317@OC1	SER_302@HG	SER_302@OG	53	0.02	2.76	158.56
27	TRP_20@NE1	TGG_317@HJ2	TGG_317@OJ2	52	0.02	2.90	55.24
28	TGG_317@OJ2	TRP_20@HE1	TRP_20@NE1	52	0.02	2.90	140.96
29	TGG_317@OC3	LYS_21@HZ3	LYS_21@NZ	41	0.01	2.87	100.92
30	TGG_317@OC3	LYS_21@HZ2	LYS_21@NZ	41	0.01	2.87	82.96
31	TGG_317@OC3	LYS_21@HZ1	LYS_21@NZ	41	0.01	2.87	78.90
32	TGG_317@OC3	SER_302@H	SER_302@N	36	0.01	2.89	155.99
33	TGG_317@OT3	LEU_300@H	LEU_300@N	33	0.01	2.94	107.75
34	LEU_300@N	TGG_317@HT3	TGG_317@OT3	33	0.01	2.94	53.43
35	LYS_21@NZ	TGG_317@HN11	TGG_317@NN1	30	0.01	2.86	56.84
36	TGG_317@NN1	LYS_21@HZ1	LYS_21@NZ	30	0.01	2.86	87.63
37	TGG_317@NN1	LYS_21@HZ2	LYS_21@NZ	30	0.01	2.86	76.00

38	TGG_317@NN1	LYS_21@HZ3	LYS_21@NZ	30	0.01	2.86	91.24
39	LYS_21@NZ	TGG_317@HN12	TGG_317@NN1	30	0.01	2.86	60.60
40	ALA_299@N	TGG_317@HT3	TGG_317@OT3	24	0.01	2.92	61.44
41	TGG_317@OT3	ALA_299@H	ALA_299@N	24	0.01	2.92	142.64
42	TGG_317@OT1	LEU_301@H	LEU_301@N	21	0.01	2.92	141.43
43	LEU_301@N	TGG_317@HT1	TGG_317@OT1	21	0.01	2.92	40.95
44	TRP_20@NE1	TGG_317@HT1	TGG_317@OT1	17	0.01	2.93	44.54
45	TGG_317@OT1	TRP_20@HE1	TRP_20@NE1	17	0.01	2.93	134.73
46	TGG_317@OC2	LYS_21@HZ2	LYS_21@NZ	14	0.00	2.82	86.81
47	TGG_317@OC2	LYS_21@HZ3	LYS_21@NZ	14	0.00	2.82	89.44
48	TGG_317@OC2	LYS_21@HZ1	LYS_21@NZ	14	0.00	2.82	89.00
49	TGG_317@OC1	LEU_301@H	LEU_301@N	14	0.00	2.92	129.13
50	ASP_216@O	TGG_317@HT3	TGG_317@OT3	10	0.00	2.77	127.46
51	ASP_216@OD2	TGG_317@HT1	TGG_317@OT1	8	0.00	2.68	150.96
52	TGG_317@OT3	LYS_21@HZ1	LYS_21@NZ	8	0.00	2.88	84.29
53	TGG_317@OT3	LYS_21@HZ3	LYS_21@NZ	8	0.00	2.88	85.11
54	LYS_21@NZ	TGG_317@HT3	TGG_317@OT3	8	0.00	2.88	33.11
55	TGG_317@OT3	LYS_21@HZ2	LYS_21@NZ	8	0.00	2.88	95.09
56	TGG_317@NN1	SER_302@HG	SER_302@OG	8	0.00	2.89	75.33
57	SER_302@OG	TGG_317@HN12	TGG_317@NN1	8	0.00	2.89	90.85
58	SER_302@OG	TGG_317@HN11	TGG_317@NN1	8	0.00	2.89	100.49
59	ASP_216@OD2	TGG_317@HN11	TGG_317@NN1	7	0.00	2.80	80.50
60	ASP_216@OD2	TGG_317@HN12	TGG_317@NN1	7	0.00	2.80	130.24
61	TGG_317@OT3	SER_302@H	SER_302@N	7	0.00	2.90	144.53
62	SER_302@N	TGG_317@HT3	TGG_317@OT3	7	0.00	2.90	40.37
63	TGG_317@OT3	SER_302@HG	SER_302@OG	6	0.00	2.86	119.70
64	SER_302@OG	TGG_317@HT3	TGG_317@OT3	6	0.00	2.86	71.33
65	TGG_317@OT1	SER_302@HG	SER_302@OG	6	0.00	2.87	152.49
66	SER_302@OG	TGG_317@HT1	TGG_317@OT1	6	0.00	2.87	25.28
67	ASP_216@OD1	TGG_317@HN12	TGG_317@NN1	4	0.00	2.86	123.58
68	ASP_216@OD1	TGG_317@HN11	TGG_317@NN1	4	0.00	2.86	86.21
69	VAL_297@O	TGG_317@HN11	TGG_317@NN1	4	0.00	2.89	135.53
70	VAL_297@O	TGG_317@HN12	TGG_317@NN1	4	0.00	2.89	84.20
71	TGG_317@OC1	SER_302@H	SER_302@N	4	0.00	2.90	145.26
72	TGG_317@OC3	ARG_217@HH21	ARG_217@NH2	3	0.00	2.85	55.53
73	TGG_317@OC3	ARG_217@HH22	ARG_217@NH2	3	0.00	2.85	150.35
74	ALA_299@N	TGG_317@HT1	TGG_317@OT1	3	0.00	2.91	33.90
75	TGG_317@OT1	ALA_299@H	ALA_299@N	3	0.00	2.91	144.73
76	LEU_300@N	TGG_317@HT1	TGG_317@OT1	3	0.00	2.93	18.80
77	TGG_317@OT1	LEU_300@H	LEU_300@N	3	0.00	2.93	114.08
78	ASP_216@OD1	TGG_317@HT1	TGG_317@OT1	2	0.00	2.72	171.90

79	TGG_317@OC1	SER_214@HG	SER_214@OG	2	0.00	2.72	159.38
80	VAL_297@O	TGG_317@HT1	TGG_317@OT1	2	0.00	2.91	89.81
81	HIE_110@NE2	TGG_317@HJ2	TGG_317@OJ2	2	0.00	2.97	46.46
82	TGG_317@OJ2	HIE_110@HE2	HIE_110@NE2	2	0.00	2.97	99.44
83	VAL_297@O	TGG_317@HN3	TGG_317@NN3	2	0.00	2.97	93.23
84	SER_214@OG	TGG_317@HT1	TGG_317@OT1	1	0.00	2.73	33.85
85	TGG_317@OT1	SER_214@HG	SER_214@OG	1	0.00	2.73	166.10
86	TGG_317@OE1	LYS_21@HZ3	LYS_21@NZ	1	0.00	2.75	47.65
87	TGG_317@OE1	LYS_21@HZ2	LYS_21@NZ	1	0.00	2.75	58.68
88	TGG_317@OE1	LYS_21@HZ1	LYS_21@NZ	1	0.00	2.75	165.88
89	TGG_317@OC3	ARG_217@HH12	ARG_217@NH1	1	0.00	2.75	172.03
90	TGG_317@OC3	ARG_217@HH11	ARG_217@NH1	1	0.00	2.75	42.65
91	ALA_299@N	TGG_317@HN11	TGG_317@NN1	1	0.00	2.86	78.99
92	ALA_299@N	TGG_317@HN12	TGG_317@NN1	1	0.00	2.86	68.81
93	TGG_317@NN1	ALA_299@H	ALA_299@N	1	0.00	2.86	132.67
94	SER_302@N	TGG_317@HT1	TGG_317@OT1	1	0.00	2.88	31.18
95	TGG_317@OT1	SER_302@H	SER_302@N	1	0.00	2.88	131.70
96	TRP_20@NE1	TGG_317@HN12	TGG_317@NN1	1	0.00	2.90	113.79
97	TGG_317@NN1	TRP_20@HE1	TRP_20@NE1	1	0.00	2.90	144.78
98	TRP_20@NE1	TGG_317@HN11	TGG_317@NN1	1	0.00	2.90	35.09
99	ASP_216@OD2	TGG_317@HN3	TGG_317@NN3	1	0.00	2.93	156.04
100	TGG_317@OF2	TRP_111@HE1	TRP_111@NE1	1	0.00	2.97	127.89
101	TRP_111@NE1	TGG_317@HF2	TGG_317@OF2	1	0.00	2.97	78.46

Table S6. Cluster analysis of MD trajectory of AR●NADPH (Without DCEG) based on RMSD of loop B.

Cluster Number	Frames	Fraction	Average Distance	Standard Deviation	Centroid	Average Cdistance
1	3046	0.61	1.2	0.34	4195	1.34
2	1315	0.26	1.17	0. 35	607	1.46
3	639	0.13	1.33	0.44	282	1.48

[1]Number of Frames: 5000; Number of Clusters: 3; Algorithm: K-means; Maximum Iterations: 1000; DBI: 1.31; pSF: 1606.13; SSR/SST: 0.39.

[2]Average Cdistance: Average distance of every frame from Centroid in a cluster.

[3]All distances are measured in Angstrom (Å).

Table S7. Cluster analysis of MD trajectory of AR●NADPH●DCEG (With DCEG) based on

RMSD of loop B.

Cluster Number	Frames	Fraction	Average Distance	Standard Deviation	Centroid	Average Cdistance
1	2914	0.58	1.44	0.49	2511	1.83
2	1376	0.26	1.79	0. 51	4656	2.05
3	710	0.14	0.93	0.41	301	2.27

[1]Number of Frames: 5000; Number of Clusters: 3; Algorithm: K-means; Maximum Iterations: 1000; DBI: 1.21; pSF: 2043.59; SSR/SST: 0.45.

[2]Average Cdistance: Average distance of every frame from Centroid in a cluster.

[3]All distances are measured in Angstrom (Å).

Movie S1. Recognition of OC2 atom of DCEG (Red) by specificity pocket formed by Cys-298, Ala-299 and Leu-300, and, Closed, opened and intermediate states in dynamics of loop B of AR [Video Format: MPEG-4].