RSC Advances

Insight into the structural stability of Coumestrol with Human Estrogen Receptor α and β subtypes: A combined approach involving docking and molecular dynamics simulation studies

Journal:	*RSC Advances*
Manuscript ID	RA-ART-07-2015-014745.R1
Article Type:	Paper
Date Submitted by the Author:	12-Sep-2015
Complete List of Authors:	Zafar, Atif; Aligarh Muslim University, Biochemistry Ahmad, Sabahuddin; Jamia Millia Islamia, Computer Science Naseem, Imrana; Aligarh Muslim University, Biochemistry

Insight into the structural stability of Coumestrol with Human Estrogen Receptor α and β subtypes: A combined approach involving docking and molecular dynamics simulation studies

Atif Zafar [1 #], Sabahuddin Ahmad [2 #], Imrana Naseem [1] *

[1] Department of Biochemistry, Faculty of Life Sciences,
Aligarh Muslim University, Aligarh 202002, Uttar Pradesh, India

[2] Department of Computer Science, Faculty of Natural Sciences,
Jamia Millia Islamia, New Delhi 110025, India

[#] Co-first authors, [*]Corresponding author: imrananaseem2009@gmail.com

Abstract

Epidemiological studies suggest that dietary consumption of phytoestrogens is associated with lower risk of breast cancer. Among phytoestrogens, coumestrol employs estrogen receptor (ER) as a target to induce apoptosis in cancer cells. Competitive binding experiments revealed higher affinity of coumestrol for ERβ than for ERα. However, recent evidence demonstrates that apoptotic potential of coumestrol in breast cancer cells requires ERα and not ERβ. It was, therefore, pertinent to enhance our understanding of coumestrol selecting ERα or ERβ subtype. In the present study, we elucidated binding mechanism of coumestrol to ERα and ERβ at molecular level using molecular docking, access channel analysis and molecular dynamics (MD) simulations. MD approach was used to determine the structural stability of coumestrol docked to ERα and ERβ by analysing H-bond, interaction energy, radius of gyration, solvent-accessible surface area, root mean square deviation (RMSD), RMS fluctuation and secondary structure elements. Our results clearly suggest that coumestrol on interaction with ERβ causes an overall destabilization of Apo-ERβ structure whereas the same on interaction with ERα leads to strong substrate binding and increase in Apo-ERα structural stability. Principal component analysis revealed higher strenuous motions of coumestrol-ERβ complex further supporting destabilization of coumestrol-ERβ during the MD run. In conclusion, this is the first report in which *in silico* approaches were implemented to suggest the effect of structural stability on selective binding of coumestrol to ERα and not to ERβ. We expect these findings to provide significant insights into ER-based drug development particularly for receptor mediated mechanism for breast cancer treatment.

Introduction

Breast carcinoma is the most commonly diagnosed female cancer with significant metastatic potential and a leading cause of mortality in women worldwide.[1,2] According to the statistics, over 10 lakh women are newly diagnosed with breast cancer every year worldwide and more than 400,000 cases will die from breast cancer.[3] Therefore, it is necessary to develop novel therapeutic approaches and identify chemotherapeutic candidates for the treatment of breast cancer.

Most types of breast cancer are classified according to the expression of estrogen receptor (ER), progesterone receptor (PR) or human epidermal growth factor receptor (HER2).[4] Genetic and histopathological heterogeneity in different subtype of breast cancer makes it difficult to treat the cancer with existing therapies.[5] Thus, newer successful therapies such as anti-estrogen drugs, aromatase inhibitors or targeting the ER have been widely applied for cancer chemotherapy.[6-9]

Recently, epidemiological studies suggest that intake of phytochemicals (soy) rich diet may result in lower risk of estrogen-dependent cancers, suggesting a potential approach for preventing breast cancer.[10-13] Coumestrol is a plant derived compound that belongs to the class of phytochemical (phytoestrogen), which mimic the biological activity of estrogens by competing with endogenous estrogens for receptor binding sites (ERα and ERβ).[14,15] This helps in decreasing the promotional effects of high levels of estrogens, induction of apoptosis and anti-proliferative effects against breast cancer cells.[16]

In a previous report, competitive binding experiments revealed higher affinity of the coumestrol (phytoestrogen) for ERβ, which is thought to be responsible for its growth inhibitory properties.[17] However, in a recent work, Obiorah *et al.*[18] have shown that the loss of ERβ in MCF-7:5C cells using siRNA did not affect the coumestrol-mediated apoptosis and

growth inhibition compared with cells transfected with the control siRNA. Interestingly, it was also found that knockdown of ERα did prevent the ability of coumestrol to either induce apoptosis or inhibit the growth of the MCF7:5C cells.[18] This suggests that ERα signaling is the initial site for coumestrol to cause growth inhibition and apoptosis in breast cancer cells.

The actual reason that favours coumestrol to select ERα repertoire instead of ERβ in human breast cancer cells is an object of current research but remains incompletely defined. To decipher the coordination between "positive selection" of ERα and "negative selection" of ERβ by coumestrol and to elucidate the molecular mechanism explaining its biological actions, we perform *in silico* experiments to investigate the detailed binding mechanism of coumestrol to ERα and ERβ at the molecular level.

In the present study, molecular docking studies were performed to determine the possible binding modes of coumestrol in human ERα and EBβ. In addition, active site access channel analysis was performed to identify the possible tunnels essential for substrate ingress and egress from the active site to the surface of the protein. Further, molecular dynamics (MD) simulation was used to investigate the binding interaction of coumestrol to ERα and ERβ by analysing the structural aspects of the protein in terms of H-bond, energy, secondary structure, radius of gyration (R_g), solvent-accessible surface area (SASA), root mean square deviation (RMSD) and RMS fluctuation. Our results clearly indicate that coumestrol on interaction with ERβ causes an overall destabilization of Apo-ERβ structure whereas the same on interaction with ERα leads to strong substrate binding and increases the stability of ERα molecule. To the best of our knowledge, this is the first report implementing MD simulation, docking and other *in silico* approaches to unravel the effect of stability on selective binding of coumestrol to ERα and not to ERβ. We expect that our study would be useful to understand the selectivity mechanism of coumestrol and will be highly helpful in ameliorating the future ER-based drug designing approaches.

Materials and methods

Selection of Human ERα and ERβ 3D Structures as Templates

The success of molecular docking protocol in predicting protein-ligand interactions depends on the availability of known 3D structure of the target protein. Hence, it was necessary to make a reasonable decision on the selection of ERα and ERβ 3D structures available in the Protein Data Bank (PDB). In order to choose the representative structure for docking protocol, the search was based on three criteria: (1) receptor bound to endogenous ligand that shares structural similarity with coumestrol; (2) receptor protein contained no mutations or modified residues; and (3) the best possible resolution. Therefore, in this study, X-ray crystal structure of human ERα in complex with E2 (PDB ID: 1G50)[19] and X-ray crystal structure of human ERβ in complex with E2 (PDB ID: 3OLS)[20] were used as target proteins in the docking protocol.

Protein and ligand preparation for Molecular Docking

Ligand Preparation

Chemical structure of coumestrol was saved in SDF format from Pubchem.[21] SDF file of coumestrol was converted into PDB format using Avogadro 1.0.1. Energy minimization and molecular optimization of compound was done using Arguslab 4.0.1.[22] Geometry optimization was carried out using AM1 (Austin Model 1), semi-empirical quantum mechanics force field in Arguslab 4.0.1. The best conformer thus obtained was based on energy minimisation and geometry optimization. The final structure exhibiting the lowest energy was saved in PDB format for input into the docking protocol.

Protein Preparation

The 3D crystal structures of human ERα (PDB ID: 1G50) and ERβ (PDB ID: 3OLS) were retrieved from the RCSB Protein Data Bank in PDB format.[23,24] The retrieved structure files of ERα and ERβ receptors contain 3 and 2 identical chains of the protein, respectively. Before starting the docking protocol, 2 identical chains out of 3 of ERα and 1 out of the 2 chains of ERβ were removed using Swiss-PDB viewer program (SPDBV).[25] All the water molecules and bounded ligand (E2) were removed using SPDBV to form Apo state of ERα and ERβ proteins.

Molecular Docking studies

To investigate the binding interaction between coumestrol and Apo state conformation of ERα (PDB ID: 1G50) and ERβ (PDB ID: 3OLS) systems, molecular docking studies were performed with the standard AutoDock (v4.2) suit incorporated in MGL tools (v1.5.6), using Lamarckian Genetic Algorithm.[26-28] Before starting the docking protocol, the target receptor and ligand were prepared using standard docking protocol and saved into 'PDBQT' format.

In docking calculations, the target-ligand poses so obtained are ranked using an energy based scoring function. To determine the most favourable binding sites of coumestrol in target, blind docking was performed. The input 'grid parameter' files were modified and the grid size was adjusted to X=Y=Z=70 points with 0.375 Å grid spacing to cover the active site region of receptors. Rest all docking parameters were set to default values. After docking, the top pose conformation of docked ligand was saved as complex in 'PDBQT' format, which was later visualized and written to PDB format using Chimera (v1.8.1).[29] Hydrogen bond interactions and its distance between protein and ligand were visualized and measured via PyMOL software (Molecular Graphics System, version 1.5.0.1, Schrodinger.LLC).[30]

Analysis of Access Channels

Program Caver (v3.0)[31] was employed to identify the possible active site access channels necessary for coumestrol to ingress and egress from active site to the surface of ERα and ERβ proteins. The probe radius and the clustering threshold were set to 1.0 and 3.5, respectively. Rest all parameters were set to default values during the calculations. All tunnels of both the estrogen receptors were visualized using PyMOL.[30]

Stability Evaluation by Molecular Dynamics (MD) Simulations

In order to determine the stability of docked complex of coumestrol with ERα and ERβ proteins, MD simulations were performed using GROMACS software package (v4.6.5).[32] In the first step, the PDB file of protein-ligand complex was separated into PDB files of protein and ligand. Protein topology was prepared using PDB of protein with 'pdb2gmx' using GROMOS96 43a1 force field.[33]

It is beyond the scope of GROMACS to parameterize heteroatom groups in PDB files. Therefore, ligand topology was developed using the PRODRG server.[34] Next, 'unit cell' was defined and the system was filled with water. The protein structure was then confined in a cubic box maintaining a minimum of 10 Å distance between any protein atom and walls of the box with periodic boundary conditions (PBC). The resulting system was then solvated by simple point charge (SPC) 216 solvent model.[35] At physiological pH, ERα and ERβ systems were found to have a net charge of -6 and -1, respectively. Therefore, counter ions 6 Na^+ and 1 Na^+ were added to neutralize ERα and ERβ systems, respectively, that replaced water molecules at positions of favourable electrostatic potential.

Next, the system was energy minimized to remove steric clashes introduced during the process. The system was minimized in 50,000 steps using the steepest descent method. After

system relaxation, position restraint dynamics (equilibration run) was applied in the system. Equilibration run was performed for 100 picoseconds (ps) (50,000 steps) in two consecutive steps: NVT (Number of particles, Volume and Temperature) and NPT (Number of particles, Pressure and Temperature). NVT equilibration was performed for 100 ps at a temperature of 300K and a coupling constant of 0.1 ps. After temperature stability, NPT simulation was performed in which the temperature was set to 300K and the pressure was 1 bar, with coupling constants of 0.1 and 2.0 ps, respectively.

Long-range electrostatic interactions were calculated using Particle-Mesh Ewald (PME) method[36] and the cut off distance for short-range van der Waals was set to 1.2 nm. LINCS (Linear Constraint Solver) algorithm[37] was used to constrain all the bond lengths, while water molecules were constrained with SETTLE algorithm.[38] Berendsen coupling scheme was also employed to equilibrate the ensembles during equilibration run.[39] Finally, a 2 nanoseconds (ns) long production simulation (MD run) was performed with a 2 femtoseconds (fs) time step at a pressure of 1 bar and a temperature of 300 K, to confirm the stability of the given systems.

Analysis of structural stability

Results of MD simulations were analysed using standard modules within GROMACS package. Secondary structure database (DSSP)[40] was installed into GROMACS to monitor time-dependent secondary structure fluctuation of Apo and docked form of receptor. The *g_energy* module in the program was used to calculate potential energy, interaction energy and total energy changes in the system. Intermolecular hydrogen bond interactions between ligand and protein were calculated using the GROMACS module *g_hbond*. Radius of gyration (R_g) via *g_gyrate* module, root mean square deviation (RMSD) via *g_rms* module, root mean square fluctuation (RMSF) via *g_rmsf* module and solvent accessible surface area

(SASA) via *g_sas* module were also analysed. Trajectories were stored every 500 ps. All trajectories of simulations were plotted using Gnuplot (v4.6) (http://www.gnuplot.info/).

Essential dynamics analysis

Essential dynamics (ED)[41,42] or principal component analysis (PCA) is a robust tool to filter large-scale concerted motions from the trajectory of MD simulation. In this study, the trajectory of MD simulation was used to determine the strenuous motions of docked complexes of coumestrol with ERα and ERβ proteins. A covariance matrix was built using atomic fluctuations in Cartesian coordinate space. After diagonalization of the covariance matrix, a set of eigenvectors and corresponding eigenvalues are obtained. Eigenvectors of a covariance matrix are called its principal components. The eigenvectors are directions in conformational space and represent the collective motion of atoms along those directions. Eigenvalues are the mean square fluctuations (MSF) of atoms along the corresponding eigenvectors. The first few eigenvectors represent the most biological significant large-scale concerted motions of a protein molecule.[43] In this study, ED analysis was performed on backbone atoms using the trajectory generated in MD simulation. The GROMACS in-built modules *g_covar* and *g_anaeig* was used to perform ED analysis.

Results and discussion

Docking analysis of coumestrol into human ERα and ERβ

In this study, molecular docking approach was used to inspect the possible binding modes of coumestrol in ERα and ERβ. We selected the top binding pose of coumestrol bound to estrogen receptors, based on energy scoring function of AutoDock program. Here, it is important to note that no information is available about the crystal structure of coumestrol-ER complexes in Protein Data Bank. Therefore, we consider it necessary to confirm the

reliability of docked coumestrol complexes for present studies. Since coumestrol shares structural similarity with 17β-estradiol (E2), we believe that coumestrol should bind at the same site as estradiol. Keeping this in mind, docked coumestrol-ERα and coumestrol-ERβ were superimposed on estradiol-ERα and estradiol-ERβ X-ray crystal structures. Superposed images revealed coumestrol present at the same position in the active site of ERα and ERβ as estradiol and also conserves many key contacts as that of estradiol (Fig. 1A and B). These results suggest that coumestrol docking complexes are reliable for the present studies.

The docking summary for coumestrol docked to receptors is listed in Table 1. All the docking experiments clearly demonstrate that coumestrol on binding to ERα exhibit less negative values of binding, intermolecular and van der Waals H-bond desolvation energies as compared to the interaction of coumestrol to ERβ. Consequently, the inhibition constant (Ki) for coumestrol-ERα complex was also higher in comparison to coumestrol-ERβ complex. These results clearly support the findings of Kostelac *et al*.,[17] where it was also demonstrated that coumestrol exhibits stronger affinity with ERβ than ERα.

Using PyMOL, the inter-molecular interactions of docked coumestrol with ERα and ERβ were observed. The docked poses of coumestrol with ERα and ERβ are shown in Fig. 2. It was found that coumestrol on interaction with ERα forms three hydrogen bonds (Arg394, His524 and Leu525) (Fig. 2A) whereas the same on interaction with ERβ forms five hydrogen bonds (Glu305, Leu339, Arg346, Gly472 and His475) (Fig. 2B). This illustration of greater hydrogen bond interaction number in ERβ than in ERα further confirms the greater stability of coumestrol-ERβ complex.

To better understand the interaction of coumestrol with the residues of ER, a plot was drawn using the java based software Ligplot.[44] Ligplot reveals hydrophobic interactions, hydrogen bond interactions and the length of hydrogen bond (Å) between the ligand and interacting residues. As evident from Fig. 3A, the residues of ERα: Arg394, His524 and Leu525 form

hydrogen bonds of length 2.69, 2.95 and 2.83Å, respectively. In the case of ERβ (Fig. 3B), the residues Glu305, Leu339, Arg346, Gly472 and His475 form hydrogen bonds of length 2.99, 3.22, 3.27, 3.20 and 2.80Å, respectively. Further, as described by Chandsawangbhuwana and Baker,[45] our Ligplot results also suggest that the receptors undergo conformational changes to assist in the binding of coumestrol in the steroid binding pocket.

Analysis of access tunnels in Coumestrol-ER complexes

PyMOL visualization (surface view) revealed that the coumestrol does not bind to the surface, but it finds its path all the way inside to the buried active site of estrogen receptors. When the opaqueness of the coumestrol-ER complexes was reduced by 20-30%, the ligand (in pink colour) was found inside the ERα (Fig. 4A) and ERβ (Fig. 4B). This possibly suggests that coumestrol on interaction with ER tries to search a path necessary to enter the receptor molecule.

To explore the possible ingress/egress pathways of coumestrol in the estrogen receptors, a program Caver (v3.0) was used. The statistics of the top ranked pathways/tunnels are summarized in Table 2. In this table, pathways were ranked on the basis of the priority value. Occupancy (%) is the occupancy of snapshots in which at least one pathway with bottleneck radius ≥1.0 Å accounted entire snapshots. The curvature of pathway = L/D, where L and D are the length of the pathway and shortest possible distance between the calculation starting point and pathway ending point, respectively. Throughput = e^{-cost} (cost=L/r^n), where r is the radius of pathway, and n is a non-negative real number. Priority is calculated as a sum of throughputs of all pathways in a given cluster, divided by the total number of snapshots that were analysed.

Caver analysis revealed top four pathway clusters: **a, b, c, d** in Apo-ERα (Fig. 5A); **a', b', c', d'** in coumestrol-ERα (Fig. 5B), and top three pathway clusters: **e, f, g** in Apo-ERβ (Fig. 5C); **e', f', g'** in coumestrol-ERβ (Fig. 5D). Profile (pathway beginning and end section) of top pathway clusters were also depicted via heat maps as shown in Fig. 6. Among these seven pathways, path **e** of ERβ was the shortest one with mean length of 14.805 Å, while the mean length of the shortest path **a** in case of ERα was 16.293 Å. In addition, the "path **a**" is mainly constituted of H3, H4 and H7 residues of ERα, whereas the "path **e**" is located mainly around H3, H5, H6 and B1 of ERβ receptor system. Hence, it signifies that among all the ranked pathways, the "path **a**" and "path **e**" facilitate the ingress/egress of coumestrol to the binding region of ERα and ERβ, respectively.

Stability evaluation by MD simulations: Is the ER complex stable on time frame?

Thermodynamic stability of protein molecule is an important feature that determines the structural and functional stabilization of protein entities.[46] Molecular docking approach helps to determine the potential binding modes of ligand but lacks in providing information about the structural stability of binding modes of ligand in protein complexes.[47] As mentioned earlier, recent investigations on receptors revealed that ERα, but not ERβ, is necessary for coumestrol to cause apoptosis in breast cancer cells.[18] To analyse the specific reason for coumestrol action via ERα signaling, comparative MD simulations were performed to examine and understand the difference in stability of coumestrol with ERα and ERβ on the time scale. The data of simulations were collected for further analysis via different built in modules of GROMACS to understand the stability of coumestrol-ER systems.

Backbone RMSD of ERα and ERβ structures bounded with coumestrol

RMSD is an important technique to calculate main chain "root mean square deviations". This analysis provides the measure of deviations (in nm) of bound ER complexes from the

corresponding starting structure over a period of time, during MD simulation. Using the module "*g_rms*", backbone RMSD was calculated after least square fit to C-alpha. It was found that the RMSD stabilizes for both coumestrol-ER systems at ~0.15 nm after 250 ps (Fig. 7). Later, between 1000 ps and 2000 ps, the structure of coumestrol-ERβ complex has higher deviation and the RMSD reaches to its maximum ~0.25 nm. However, in the case of coumestrol-ERα complex, there is much lower RMSD and it remains ~0.16 nm for longer time period (Fig. 7). The RMSD plot suggested that coumestrol was bound tightly in the active site region of ERα via hydrogen and hydrophobic interactions with the surrounding residues. On the other hand, coumestrol in ERβ system was not bound tightly and moved within the cavity, resulting in continuous rise in the RMSD. This result indicated that ERα bounded with coumestrol remains more stable system than ERβ during the simulated period.

RMSF of residues in ERα and ERβ bounded with coumestrol

To examine the flexibility and local changes in the structure, root-mean-square fluctuation (RMSF) versus the residue number for coumestrol-ERα/β system was investigated. Fluctuation analysis revealed that overall flexibility of ERβ structure increased to a greater extent (0.3 nm) as compared to ERα structure (0.23 nm) on binding of coumestrol, confirming the stability of ERα system (Fig. 8). The ligand binding cavity of ERα and ERβ is nearly identical and formed by residues between H3 and H11 helixes.[48-51] Since loops of a protein play an important role in substrate or drug binding, hence it was necessary to detect the extent of fluctuations in the loops of coumestrol-ER systems. As evident from Fig. 8, slight fluctuations were observed for H3 (341-363), H6 (421-438), H7 (442-455) and H8 (466-492) helixes for ERα system whereas in the case of ERβ system greater extent of fluctuations were seen for the respective (H3:324-348; H6: 394-406; H7: 422-444 and H8: 448-482) helixes. In ER systems, residues of H3, H4 and H5 play a critical role in forming co-activator recruitment site necessary for transcriptional activation of receptor.[52] In case of

ERβ, greater extent of fluctuations were seen for residues in H3, H4 and H5 as compared to ERα system, suggesting diminution of activation of ERβ system after binding of coumestrol. These results support the findings of dynamic communication between androgen and co-activator that establishes androgen receptor (AR) functional potency.[53] Thus, it may be inferred that loop regions of coumestrol-ERα complex had the lowest fluctuation values indicating higher stability of docked ERα system.

Energy analysis of docked complexes: Stability of ERα/β system

To examine the stability of docked ER systems, trajectories obtained via MD simulations were analysed for interaction, potential and total energies with respect to the starting conformation, as a function of time.

The potential energy of a system is a simple measure of system stability. Analysis of trajectories revealed that both the ERα and ERβ molecular systems in the simulation were well equilibrated and remained stable throughout the simulation of 2000 ps (Fig. 9). This also implies that the energy minimization was successful. However, potential energy plots show that the potential energy for bound ERα remains more negative (approximately -500000 kJ/mol) as compared to ERβ (approximately -440000 kJ/mol) (Fig. 9). This result indicates that the bound form of ERα, on an average, was more stable than the bound ERβ system. Similarly, the total energy plots also indicate that the total energy for ERα remains more negative (approximately -410000 kJ/mol) as compared to ERβ (approximately -360000 kJ/mol), confirming the structural stability of bound ERα as compared to ERβ system (Fig. 10).

Interaction energy, which results from the binding of ligand to the active site region of protein, was calculated to measure the stability of ER complexes. The interaction energy plots were drawn as function of time for ERα (Fig. 11A) and ERβ (Fig. 11B). The plot clearly

indicates the unstable nature of ERβ complex with the passage of time. It is evident from Fig. 11 that ERβ has less negative interaction energy (-5 kJ/mol) as compared to ERα (-20 kJ/mol) system till ~1000 ps. Here, it is also important to note that the interaction energy of ERβ is not stable; major fluctuations were perceived from 1000-2000 ps, which correspond to association and dissociation of coumestrol in ligand binding cavity of ERβ system (Fig. 11B). However, in the case of coumestrol-ERα system, the variation in interaction energy was found to be insignificant and was more stable from 500 to 2000 ps (Fig. 11A). This illustration of significant variation in the interaction energy of coumestrol-ERβ system, during simulated period, confirms it to be a non-stable system.

Radius of gyration (R_g) and Solvent accessible surface area analyses (SASA)

To determine the level of structure compactness of ERα and ERβ systems (in complex with coumestrol), radius of gyration analysis was performed for 2000 ps long MD run at 300K. Radius of gyration (Rg) can be defined as the mass weighted root mean square distance of a collection of atoms from their common center of mass. As evident from Fig. 12, the radius of gyration for ERα remained low (~1.74 nm), while the same for ERβ was higher (~1.82 nm). This suggests that coumestrol on binding to ERβ decreases the stability of the receptor.

The result of Rg analysis was found to be supported by solvent accessible surface area analysis (SASA) plot, which measures the solvent accessibility of ERα and ERβ. Accessible surface area (nm^2) of ERα and ERβ (each bound with coumestrol) was analysed using the plot drawn as a function of time for 2000 ps long simulation (Fig. 13). From this plot, it is evident that coumestrol-ERα has lower SASA value (107-125 nm^2) as compared to the coumestrol-ERβ system (125-135 nm^2). Higher SASA indicates that on binding of coumestrol to ERβ, the receptor unfolds and exposes the underlying hydrophobic amino acid residues to the solvent. Lower SASA indicates that coumestrol on binding to ERα leads to

tight packing of the hydrophobic core residues in the binding pocket, making the pocket inaccessible to the solvent molecules and thereby optimizing van der Waals interactions. These interactions, thus, lead to more compact structure of coumestrol-ERα system, thereby supporting the higher stability of ERα protein.

Hydrogen bond analysis

Hydrogen bonds play vital role in molecular recognition and stability of protein structure.[54] Higher number of intermolecular hydrogen bond interactions leads to greater stability of the protein-ligand complex. In the present investigation, hydrogen bond analysis was performed to depict the stability of docked ERα and ERβ systems (Fig. 14).

In the case of ERα, the hydrogen bond interactions reach a maximum of four and remain one or two for most of the time (Fig. 14A); whereas in the case of ERβ the hydrogen bond interaction number reaches (sometimes) two from 1100-1600 ps and remains one in number for much lesser time (Fig. 14B). Therefore, we suggest that lesser H-bonding of coumestrol with ERβ may facilitate its detachment, and hence the ligand can egress the receptor via shortest "path **e**" (Fig. 15).

To get more insight into the hydrogen bonding at major time intervals during the simulated period, trajectories of docked ERα and ERβ were analysed. The PDB(s) were recorded at the interval of 500 ps, 1000 ps, 1500 ps and 2000 ps from simulations of both ERα and ERβ. Ligplot(s) were drawn for ERα (Fig. 16) and ERβ (Fig. 17). This analysis provided valuable evidence in support of constant interaction of binding residues of ERα with coumestrol via hydrogen bonding and hydrophobic contacts throughout the 2000 ps MD simulation.

Ligplot for ERα at 500 ps reveals that there were two hydrogen bonds formed with Arg394 and His524 (Fig. 16A). At 1000 ps, three hydrogen bonds with Ala350, Arg394 and His524

made the complex more stable at this very instant of time (Fig. 16B). At 1500 ps and 2000 ps two hydrogen bonds were formed (at each time interval) with Arg394 and His524 (Fig. 16C), and Ala350 and His524 (Fig. 16D), respectively.

Study of the Ligplot for ERβ at 500 ps reveals that hydrogen bonds were missing at this very instant of time (Fig. 17A). Similarly, at 1000 ps, no hydrogen bond was formed (Fig. 17B). Later, only one hydrogen bonding with the residue Thr299 was evident at both time intervals 1500 ps (Fig. 17C) and 2000 ps (Fig. 17D).

The details of Ligplot(s) for ERα and ERβ are summarised in Tables 3 and 4, respectively. Highlighted residues in Tables 3 and 4 correspond to important residues which appear constantly throughout the simulations for ERα and ERβ, respectively. In ERα, His524 (involved in hydrogen bonding), and Leu384 and Leu387 (involved in hydrophobic contacts) were perceived throughout the simulation. But, in the case of ERβ, only Leu476 (involved in hydrophobic contacts) is seen throughout 2000 ps simulation. These illustrations of greater hydrogen bond interaction number and higher participation in H-bonding with important residues of ERα confirm greater stability of coumestrol-ERα system.

Eigenvectors (ED analysis)

Eigenvectors obtained from ED analysis were used to determine the overall strenuous motions within the two systems: coumestrol-ERα and coumestrol-ERβ. Eigenvector (or principal component) plot (Fig. 18) illustrates the 2D projections of coumestrol-ERα and coumestrol-ERβ systems for 2ns, where snapshots were taken at every 2ps. The level of conformational changes within the docked ERα and ERβ structures can be understood via distribution of dots within the graph. From these projections, it was observed that the clusters of coumestrol-ERβ system covered a greater region of conformational space than that of the coumestrol-ERα system. This suggests that internal strenuous motions of the coumestrol-ERβ

are much greater than the same of coumestrol-ERα, supporting more stability of coumestrol-ERα system during the MD simulation. Fig. 19 shows the superposition of extreme projections on PC1 for coumestrol-ER complexes obtained from MD simulation trajectories. Results clearly suggest that coumestrol bound ERβ complex shows greater conformational variations in H3, H4 and H5 helixes (involved in forming ligand binding cavity and co-activator recruitment site) whereas in case of coumestrol bound ERα complex no evident variations were observed in respective helixes. This further suggests that coumestrol facilitates stable interaction at active site region causing less ERα protein motion. For eigenvector 2, the results are quite similar to eigenvector 1 for the respective coumestrol-ER complexes (data not shown).

Secondary Structure Analysis

Secondary structure elements (helix, sheet, and coil) were analysed to explore the stability of ERα and ERβ systems during MD simulation. Secondary structure elements for ERα (Apo) and ERα (bounded with coumestrol) during 2000 ps simulations are depicted in Fig. 20A and Fig. 20B, respectively. For ERβ (Apo) and ERβ (bounded with coumestrol), the secondary structure elements are illustrated in Fig. 21A and Fig. 21B, respectively.

Important residues in ERα such as His524 have significant contribution in hydrogen binding and thus responsible for overall stability of the protein. Similarly, residues responsible for hydrophobic interactions (Leu384 and Leu387) are also important for structural stability. As evident from secondary structure plot for ERα (bounded with coumestrol) (Fig. 20), the secondary structure elements for the above mentioned residues remain preserved, with a slight or no distortion noticed throughout 2000 ps MD simulation. Important helixes (H3 and H4) are found stable during the course of MD simulation for 2000 ps (Fig. 20B). This shows

that binding of coumestrol to ERα does not have major impact on the elements of secondary structure, and hence the coumestrol-ERα complex remains more stable.

In contrast to ERα, the secondary structural elements of important residues in ERβ were found to be distorted. Higher distortion was observed in important residue Leu476 which is responsible for hydrophobic interactions with coumestrol (Fig. 21B). In the coumestrol-ERβ complex, the residues of important helix (H3) are found distorted from 900-1700 ps (Fig. 21B). Similarly, important helix (H5) is also unstable from 700-1500 ps (Fig. 21B). In the course of simulation the residues of other helixes (H4, H6-H9) show minor distortions. Thus, the secondary structure analysis suggests that coumestrol on interaction with ERβ causes major changes in the secondary structure elements, thereby destabilizing the receptor.

Role of active site water molecule in coumestrol-ERβ system

The crystal structure of ERα (1G50) lacks the active site water molecule, whereas the crystal structure of ERβ (3OLS) at the same site contains a well ordered water molecule which forms H-bond framework with estradiol and important active site residues (Glu305 and Arg346) of ERβ. Therefore, it was important to investigate whether ERβ active site water molecule plays any role in coumestrol-ERβ complex. For this, coumestrol was also docked into the crystal structure of ERβ (3OLS) without removing the active site water molecule. The docking summary for coumestrol docked to ERβ without removing active site water molecule is listed in ESI Table S1. Docking experiments clearly demonstrate that coumestrol on binding to ERβ (without removing water) exhibits insignificant change in the values of binding, intermolecular and van der Waals H-bond desolvation energies (ESI Table S1) as compared to those when active site water molecule was removed from ERβ system (Table 1). Similarly, the inhibition constant (Ki) of coumestrol-ERβ complex with water molecule was comparable to Ki of coumestrol-ERβ complex without active site water molecule (Table 1 and ESI Table

457 S1). Further, we have investigated inter-molecular interactions of coumestrol in ERβ without
458 removing active site water molecule. Interestingly, we found that, like ERβ system (without
459 water) (Fig. 2B), coumestrol forms identical five H-bond interactions with active site residues
460 of ERβ (Glu305, Leu339, Arg346, Gly472 and His475) even in the presence of active site
461 water molecule (ESI, Fig. S1). Ligplot also reveals identical hydrophobic interactions of
462 coumestrol with ERβ with and without active site water molecule (Fig. 3B and ESI, Fig. S2).

463 For thorough analysis of the role of active site water molecule, MD simulation was
464 performed for coumestrol-ERβ complex without removing active site water molecule on the
465 time scale of 2 ns. As described earlier, RMSD is used to calculate main chain "root mean
466 square deviations". It was found that RMSD of coumestrol-ERβ complex with and without
467 active site water molecule crosses ~ 0.2 nm after 1000 ps and follow similar trend throughout
468 the course of MD run (ESI, Fig. S3). This shows that the presence of active site water
469 molecule in coumestrol-ERβ system provides insignificant change in RMSD values obtained
470 without water during the simulated period. Further, RMSF analysis revealed that overall
471 flexibility of coumestrol-ERβ complex was found to be 0.3 nm with and without active site
472 water molecule (Fig. 8B and ESI, Fig. S4). As evident from Fig. 8B and ESI Fig. S4, the
473 extent of fluctuation in important helixes (H3, H4, H5, H6, H7 and H8) were similar in
474 coumestrol-ERβ complex with and without water molecule. These results suggest that active
475 site water molecule plays insignificant role in coumestrol-ERβ complex.

476 To further examine the effect of active site water molecule in coumestrol-ERβ complex,
477 potential and total energies were estimated. Analysis of trajectories revealed that potential
478 energy of coumestrol-ERβ complex with and without active site water remains same
479 (approximately -440000 kJ/mol) (ESI, Fig. S5). Similarly, the total energy plot indicates that
480 the total energy of coumestrol-ERβ complex with and without water remains same
481 (approximately -360000 kJ/mol) (ESI, Fig. S6). Interaction energy plots suggest that

coumestrol-ERβ complex with and without active site water molecule is unstable; major fluctuations were observed from 1000-2000 ps (without water) and 1200-2000 ps (with water), which correspond to association and dissociation of coumestrol in ligand binding cavity of both the systems (Fig. 11B and ESI, Fig. S7). This illustration of similar nature of fluctuation (association and dissociation) in the interaction energy of coumestrol-ERβ complex (with and without water) confirms insignificant role of active site water molecule in coumestrol-ERβ system.

Results of Rg analysis revealed that coumestrol-ERβ complex with and without active site water molecule exhibits similar (~ 1.82 nm) Rg values, whereas the plot of SASA also indicates that coumestrol-ERβ complex with and without water exhibits similar (125-135 nm^2) values during the simulated period (ESI, Fig. S8 and S9). These results further confirm insignificant role of active site water in coumestrol-ERβ complex.

In the present investigation, hydrogen bond analysis was also performed for docked ERβ complex in the presence of active site water molecule. Similar pattern of lesser H-bonding of coumestrol with ERβ (a maximum of two interactions and minimum of one) was observed during 2 ns (ESI, Fig. S10). Further, no change was observed in the hydrogen bonding and hydrophobic interactions at major time intervals of 500 ps, 1000 ps, 1500 ps and 2000 ps for coumestrol-ERβ complex (without removing water) (data not shown).

ED analysis was used to determine the strenuous motions for coumestrol-ERβ complex without removing the active site water molecule. 2D projections plot suggests similar internal strenuous motions with and without active site water molecule in coumestrol-ERβ complex (Fig. 18 and ESI, Fig. S11). From these projections, it was observed that clusters of coumestrol-ERβ complex with active site water molecule covered a greater region of conformational space similar to that observed without active site water. Secondary structure

elements for coumestrol-ERβ complex without removing active site water molecule were also explored during MD simulation (ESI, Fig. S12). Results clearly show that the presence (ESI, Fig. S12) and absence (Fig 21) of water molecule in the active site region of coumestrol-ERβ system resulted in similar major distortions in secondary structure elements of ERβ protein.

In view of the above results, we conclude that active site water molecule in ERβ protein plays insignificant role in coumestrol-ERβ system. Coumestrol on interaction with ERβ continues to destabilize the receptor even in the presence of active site water molecule. Therefore, we suggest that water molecule in ERβ system can be there because of crystallized estradiol ligand and not because of active site.

Conclusion

In the present study, the interactions between coumestrol and the two receptors, ERα and ERβ were analysed using different computational approaches such as molecular docking, MD simulations and active site access channel analysis. Our molecular docking results reflected higher affinity of coumestrol with ERβ than ERα which is in agreement with the competitive binding experiments. However, the induction of apoptosis in breast cancer cells via coumestrol-ERα based mechanism cannot be ruled out. Therefore, our aim was to unravel the differences in substrate specificity and determine the stability of coumestrol-ER complexes on time scale via MD simulations. Our results of structural characteristic features such as RMSD, RMSF, R_g, secondary structure, SASA and H-bonding plots revealed higher stability of coumestrol-ERα complex due to increased hydrogen bond interactions. The results of potential energy, interaction energy and total energy plots were also in consistent with the higher stability of coumestrol-ERα system. ED analysis revealed less internal strenuous motions in coumestrol-ERα complex as compared to coumestrol-ERβ system, supporting high structural stability of ERα complex. Further, the active site access channel analysis

indicates the paths “a” and “e” as the shortest channels for ingress/egress of coumestrol to the binding region of ERα and ERβ, respectively. Finally, simulation results indicated that coumestrol exhibits fewer interactions in the binding pocket of ERβ, and hence may undergo detachment to the extracellular site of receptor via the shortest path “e”. Thus, detailed investigations about structural stability of coumestrol-ER systems will provide crucial insights into the mechanism of action of phytoestrogens and aid in designing of ER-based drugs for breast cancer therapy.

Acknowledgements

Non-financial computational support from the computer labs of Dept. of Physics, Faculty of Science, AMU, Aligarh is greatly acknowledged. The authors would also like to thank the Chairperson, Dept. of Biochemistry, AMU, Aligarh for providing the necessary lab facilities. The financial support to AZ by DST-INSPIRE Junior Research Fellowship from the Govt. of India is acknowledged.

References

1. R. Siegel, D. Naishadham and A. Jemal, *CA-Cancer J. Clin.*, 2013, 63, 11–30.

2. J.G. Brody, R.A. Rudel, K.B. Michels, K.B. Moysich, L. Bernstein, K.R. Attfield and S. Gray, *Cancer*, 2007, 109, 2627–34.

3. A. Jemal, R. Siegel, J. Xu and E. Ward, *CA-Cancer J. Clin.*, 2010, 60, 277–300.

4. C.R. Tate, L.V. Rhodes, H.C. Segar, J.L. Driver, F.N. Pounder, M.E. Burow and B.M. Collins-Burow, *Breast Cancer Res.*, 2012, 14, R79.

5. A.H. Sims, A. Howell, S.J. Howell and R.B. Clarke, *Nat. Clin. Pract. Oncol.,* 2007, 4, 516–25.

6. C.A. Hudis and L. Gianni, *Oncologist*, 2011, 1, 1–11.

7. X. Xu, W. Yang, Y. Li and Y. Wang, *Expert Opin. Drug Discov.*, 2010, 5, 21-31.

8. Y. Wang, Y. Li, J. Ding, Y. Wang and Y. Chang, *Mol. Divers.,* 2008, 12, 93-102.

9. Z. Wang, Y. Li, C. Ai and Y. Wang, *Int. J. Mol. Sci.,* 2010, 11, 3434-58.

10. H. Adlercreutz, *J. Steroid Biochem. Mol. Biol.*, 2002, 83, 113–8.

11. T. Oseni, R. Patel, J. Pyle and V.C. Jordan, *Planta Med.*, 2008, 74, 1656–65.

12. T. Lorand, E. Vigh and J. Garai, *Curr. Med. Chem.*, 2010, 17, 3542–74.

13. M. Hedelin, K.A. Balter, E.T. Chang, R. Bellocco, A. Klint, J.E. Johansson, F. Wiklund, C. Thellenberg-Karlsson, H.O. Adami and H. Gronberg, *Prostate*, 2006, 66, 1512–20.

14. R. de Leeuw, J. Neefjes and R. Michalides, *Int. J. Breast Cancer*, 2011, 2011, 232435.

15. E.A. Musgrove and R.L. Sutherland, *Nat. Rev. Cancer*, 2009, 9, 631–43.

16. Y.H. Lee, H.J. Yuk, K.H. Park and Y.S. Bae, *Food Chem.*, 2013, 141, 381-8.

17. D. Kostelac, G. Rechkemmer and K. Briviba, *J. Agric. Food Chem.*, 2003, 51, 7632–5.

18. I.E. Obiorah, P. Fan and V.C. Jordan, *Cancer Prev. Res.*, 2014, 7, 939-49.

19. S. Eiler, M. Gangloff, S. Duclaud, D. Moras and M. Ruff, *Protein Expr. Purif.*, 2001, 22, 165-73.

20. S. Möcklinghoff, R. Rose, M. Carraz, A. Visser, C. Ottmann and L. Brunsveld, *Chembiochem.*, 2010, 11, 2251-4.

21. E.E. Bolton, Y. Wang, P.A. Thiessen and S.H. Bryant, *Annu. Rep. Comput. Chem.*, 2008, 4, 217–241.

22. C. Peng, P.Y. Ayala, H.B. Schlegel and M.J. Frisch, *J. Comput. Chem.*, 1996, 17, 49–56.

23. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov and P.E. Bourne, *Nucleic Acids Res*., 2000, 28, 235-42.

24. H.M. Berman, T. Battistuz, T.N. Bhat, W.F. Bluhm, P.E. Bourne, K. Burkhardt, Z. Feng, G.L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J.D. Westbrook and C. Zardecki, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2002, 58, 899-907.

25. N. Guex and M.C. Peitsch, *Electrophoresis,* 1997, 18, 2714-23.

26. G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell and A.J. Olson, *J. Comput. Chem.*, 2009, 30, 2785–91.

27. M.F. Sanner, *J. Mol. Graph Model*, 1999, 17, 57-61.

28. G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew and A.J. Olson, *J. Comput. Chem.,* 1998, 19, 1639–1662.

29. E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng and T.E. Ferrin, *J. Comput. Chem.,* 2004, 25, 1605-12.

30. W.L. DeLano, The PyMOL molecular graphics system, DeLano Scientific, Palo Alto, CA, 2002.

31. B. Kozlikova, E. Sebestova, V. Sustr, J. Brezovsky, O. Strnad, L. Daniel, D. Bednar, A. Pavelka, M. Manak, M. Bezdeka, P. Benes, M. Kotry, A. Gora, J. Damborsky and J. Sochor, *Bioinformatics,* 2014, 30, 2684-5.

32. H.J.C. Berendsen, D. van der Spoel and R. van Drunen, *Comput. Phys. Commun.*, 1995, 91, 43-56.

33. W.R.P. Scott, P.H. Hunenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A.E. Torda, T. Huber, P. Kruger and W.F. van Gunsteren, *J. Phys. Chem. A*, 1999, 103, 3596–3607.

34. A.W. Schuttelkopf and D.M. van Aalten, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, 60, 1355–63.

35. H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren and J. Hermans, in *Intermolecular Forces*, ed. B. Pullman, Reidel, Dordrecht, 1981, 331–342.

36. U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee and L.G. Pedersen, *J. Chem. Phys.*, 1995, 103, 8577–8593.

37. B. Hess, H. Bekker, H.J.C. Berendsen and J.G.E.M. Fraaije, *J. Comput. Chem.*, 1997, 18, 1463-1472.

38. S. Miyamoto and P.A. Kollman, *J. Comput. Chem.*, 1992, 13, 952–962.

39. H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola and J.R. Haak, *J. Chem. Phys.*, 1984, 81, 3684–3690.

40. W. Kabsch and C. Sander, *Biopolymers*, 1983, 22, 2577–637.

41. A. Amadei, A.B. Linssen and H.J. Berendsen, *Proteins*, 1993, 17, 412-25.

42. M.A Balsera, W. Wriggers, Y. Oono and K. Schulten, *J. Phys. Chem.*, 1996, 100, 2567–2572.

43. S.Q. Liu, Z.H. Meng, Y.X. Fu and K.Q. Zhang, *J. Mol. Model*, 2010, 16, 17–28.

44. A.C. Wallace, R.A. Laskowski and J.M. Thornton, *Protein Eng.*, 1995, 8, 127–34.

45. C. Chandsawangbhuwana and M.E. Baker, *Steroids*, 2014, 80, 37-43.

46. V.M. Balcão and M.M. Vila, *Adv. Drug Deliv. Rev.*, 2014, doi:10.1016/j.addr. 2014.10.005.

47. S.H. Mahboobi, A.A. Javanpour and M.R. Mofrad, *PLoS One*, 2015, 10, e0112969.

48. A.M. Brzozowski, A.C. Pike, Z. Dauter, R.E. Hubbard, T. Bonn, O. Engström, L. Ohman, G.L. Greene, J.A. Gustafsson and M. Carlquist, *Nature*, 1997, 389, 753-8.

49. A.K. Shiau, D. Barstad, P.M. Loria, L. Cheng, P.J. Kushner, D.A. Agard and G.L. Greene, *Cell*, 1998, 95, 927-37.

50. D.M. Tanenbaum, Y. Wang, S.P. Williams and P.B. Sigler, *Proc. Natl. Acad. Sci. USA*., 1998, 95, 5998-6003.

51. P. Ascenzi, A. Bocedi and M. Marino, *Mol. Aspects Med.*, 2006, 27, 299-402.

52. W. Feng, R.C. Ribeiro, R.L. Wagner, H. Nguyen, J.W. Apriletti, R.J. Fletterick, J.D. Baxter, P.J. Kushner and B.L. West, *Science*, 1998, 280, 1747-9.

53. X. Xu, W. Yang, X. Wang, Y. Li, Y. Wang and C. Ai, *Proteins*, 2011, 79, 1154-71.

54. D.H. Williams, E. Stephens, D.P. O'Brien and M. Zhou, *Angew. Chem. Int. Ed. Engl.*, 2004, 43, 6596–616.

Legends

Fig. 1 Superposition of docked coumestrol-ERα with estradiol-ERα crystal structure (1G50) (A) and docked coumestrol-ERβ with estradiol-ERβ crystal structure (3OLS) (B). Docked coumestrol-ER complexes and estradiol-ER X-ray structures (1G50 and 3OLS) are represented by blue and red colours, respectively.

Fig. 2 Binding modes of coumestrol docked into human ERα (A) and ERβ (B) binding sites. Coumestrol forms 3 and 5 hydrogen bonds with ERα and ERβ residues, respectively.

Fig. 3 Protein-ligand interactions analysed by the program Ligplot. (A) Coumestrol with human ERα and (B) Coumestrol with human ERβ.

Fig. 4 Surface view of bound coumestrol into human ERα (A) and ERβ (B) proteins. Pink colour represents the docked coumestrol in ligand binding domain of ER and found inside the receptor molecule.

Fig. 5 Top ranked collective pathways of Apo-ER and coumestrol-ER systems generated by CAVER (v3.0). (A) Apo-ERα system, (B) ERα-coumestrol system, (C) Apo-ERβ system, and (D) ERβ-coumestrol system. Pathways "a" and "e" represent the shortest channels for ERα and ERβ, respectively.

Fig. 6 CAVER heat maps of Apo-ER and coumestrol-ER systems depicting pathway end section, bottleneck and pathway beginning. (A) a, b, c and d for Apo-ERα and a', b', c' and d' for ERα-coumestrol systems, and (B) e, f and g for Apo-ERβ and e', f' and g' for ERβ-coumestrol systems.

Fig. 7 Backbone RMSD plots of coumestrol-ER systems during MD simulation at 300 K. Black colour RMSD plot indicates coumestrol-ERα complex and red colour RMSD plot indicates coumestrol-ERβ complex.

Fig. 8 RMSF of backbone atoms of Apo-ER and coumestrol-ER systems at 300 K. (A) Apo-ERα plot is shown in red and coumestrol-ERα plot is shown in black, and (B) Apo-ERβ plot is shown in red and coumestrol-ERβ plot is shown in black.

Fig. 9 Potential energy plots of coumestrol-ERα (shown in black) and coumestrol-ERβ (shown in red) complexes during the MD simulations.

Fig. 10 Total energy plots of coumestrol-ERα (shown in black) and coumestrol-ERβ (shown in red) complexes during the MD simulations.

Fig. 11 Interaction energy plots for 2 ns MD simulation of coumestrol with ERα (A) and ERβ (B).

Fig. 12 Variation of gyration radius (R_g) of ER backbone atoms calculated as a function of time for coumestrol-ERα (shown in black) and coumestrol-ERβ (shown in red) systems.

Fig. 13 Solvent accessible surface area (SASA) during 2 ns simulations calculated from trajectory files for coumestrol-ERα (shown in black) and coumestrol-ERβ (shown in red) systems.

Fig. 14 Stability evaluation of coumestrol-ERα (A) and coumestrol-ERβ (B) complexes using intermolecular hydrogen bonding pattern as a function of time.

Fig. 15 Surface view of coumestrol into human ERβ at 2000 ps. Docked coumestrol (blue colour) appears to come out at the surface of ERβ system.

Fig. 16 Binding modes of coumestrol with residues in active site of ERα at different time intervals. (A) binding of coumestrol at 500 ps, (B) binding of coumestrol at 1000 ps, (C) binding of coumestrol at 1500 ps, and (D) binding of coumestrol at 2000 ps.

Fig. 17 Binding modes of coumestrol with residues in active site of ERβ at different time intervals. (A) binding of coumestrol at 500 ps, (B) binding of coumestrol at 1000 ps, (C) binding of coumestrol at 1500 ps and (D) binding of coumestrol at 2000 ps.

Fig. 18 2D projection of the backbone atoms of coumestrol-ERα (shown in red) and coumestrol-ERβ (shown in green) systems over the first two principal components.

Fig. 19 Motion of atoms described by the eigenvector 1 in coumestrol-ERα (A) and coumestrol-ERβ (B) complexes, by superimposing the two extreme projections (green and grey) and average structure (pink). Significant movements in H3, H4 and H5 helixes of coumestrol-ERβ system are represented by black arrows.

Fig. 20 Secondary structure elements changes during the 2 ns MD simulation at 300 K. (A) Apo-ERα and (B) coumestrol-ERα complex. The colour scale at the bottom represents the DSSP classification of each secondary structure element.

Fig. 21 Secondary structure elements changes during the 2 ns MD simulation at 300 K. (A) Apo-ERβ and (B) coumestrol-ERβ complex. The colour scale at the bottom represents the DSSP classification of each secondary structure element.

Tables

Table 1 AutoDock analysis of docked coumestrol with ERα and ERβ proteins

AutoDock Parameter	Coumestrol-ERα	Coumestrol-ERβ
Binding energy (kcal/mol)	-8.58	-8.95
Inhibition constant (nM)	510.37	261.23
Intermolecular energy	-9.18	-9.54
Van der Waals H-bond desolvation energy	-8.83	-8.94
AutoDock refRMS	104.02	35.31
No. of Hydrogen bonds	3	5
Residues involved in hydrogen bonding	Arg394, His524, Leu525	Glu305, Leu339, Arg346, Gly472, His475

Table 2 Characteristics of the pathways of four systems

A. Characteristics of the top 4 ranked pathways of Apo-ERα

Pathway	a	b	c	d
Occurrence (%)	100	100	100	100
Mean bottleneck radius [Å]	1.005	1.005	1.005	1.005
Max bottleneck radius [Å]	1.01	1.01	1.01	1.01
Mean pathway length[Å]	16.293	35.997	39.336	42.934
Mean pathway curvature	1.596	1.563	1.682	1.541
Mean throughput	0.54603	0.24641	0.21872	0.16904
Priority	0.54603	0.24641	0.21872	0.16904

B. Characteristics of the top 4 ranked pathways of Coumestrol-ERα

Pathway	a'	b'	c'	d'
Occurrence (%)	100	100	100	100
Mean bottleneck radius [Å]	1.005	1.005	1.005	1.005
Max bottleneck radius [Å]	1.01	1.01	1.01	1.01
Mean pathway length[Å]	18.936	36.846	41.867	45.958
Mean pathway curvature	1.754	1.630	1.712	1.990
Mean throughput	0.53348	0.23332	0.19554	0.16959
Priority	0.53348	0.23332	0.19554	0.16959

C. Characteristics of the top 3 ranked pathways of Apo-ERβ

Pathway	e	f	g
Occurrence (%)	100	100	100
Mean bottleneck radius [Å]	1.225	1.022	1.084
Max bottleneck radius [Å]	1.23	1.02	1.08
Mean pathway length [Å]	14.805	16.645	18.811
Mean pathway curvature	1.280	1.371	1.303
Mean throughput	0.63541	0.45990	0.37827
Priority	0.63541	0.45990	0.37827

D. Characteristics of the top 3 ranked pathways of Coumestrol-ERβ

Pathway	e'	f'	g'
Occurrence (%)	100	100	100
Mean bottleneck radius [Å]	1.245	1.035	1.074
Max bottleneck radius [Å]	1.24	1.04	1.07
Mean pathway length [Å]	15.177	16.230	16.549
Mean pathway curvature	1.199	1.417	1.201
Mean throughput	0.63402	0.48663	0.40951
Priority	0.63402	0.48663	0.40951

Table 3 Residues of ERα forming H-bonds and hydrophobic contacts with coumestrol at different time intervals

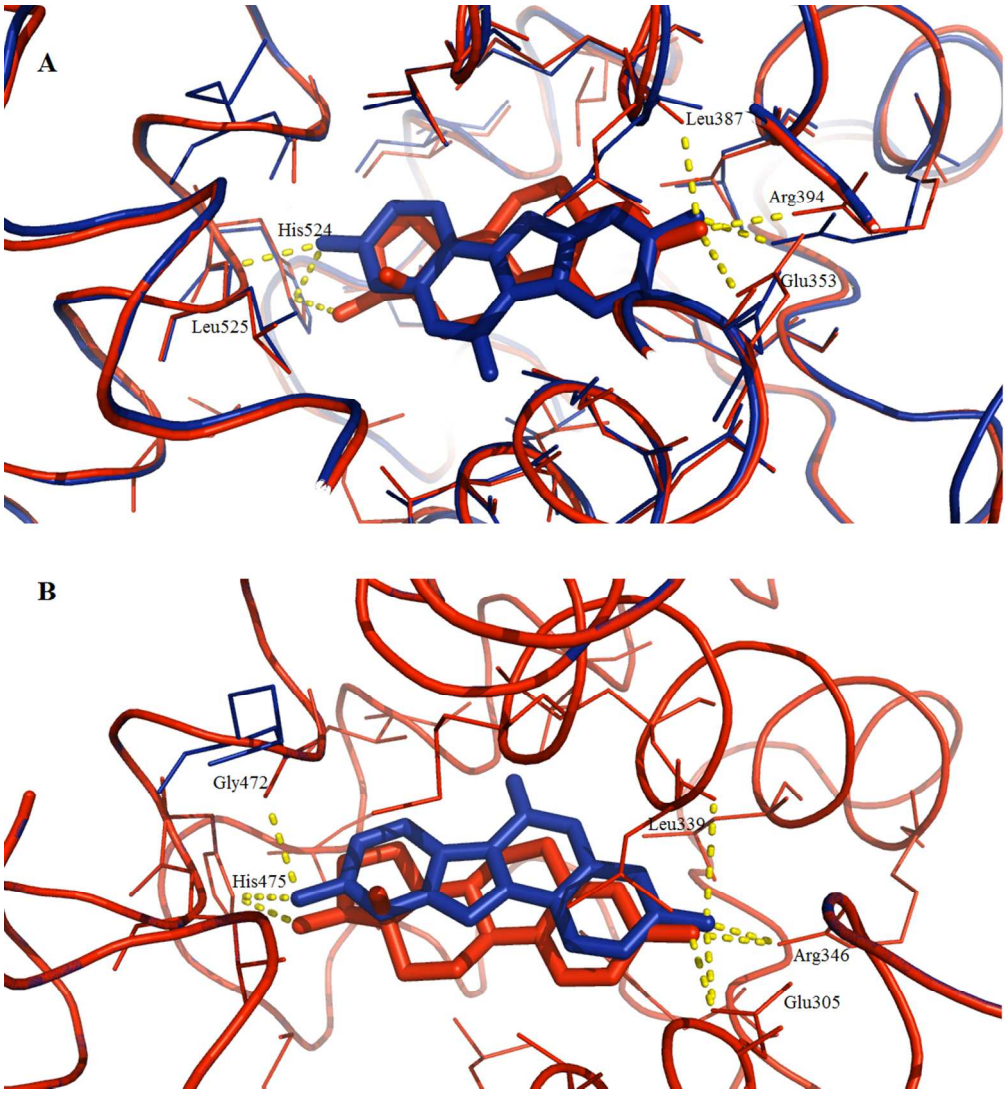
Time period	Residues involved in hydrogen bonding	Residues involved in hydrophobic interactions
0	Arg394, **His524**, Leu525	Ala350, Glu353, **Leu384**, **Leu387**, Gly521
500	Arg394, **His524**	Leu346, Ala350, **Leu384**, **Leu387**, Leu391, Phe404, Met421, Ile424, Leu525
1000	Ala350, Arg394, **His524**	Met343, Leu349, **Leu384**, **Leu387**, Leu391, Phe404, Met421, Leu525
1500	Arg394, **His524**	Met343, Leu346, Ala350, **Leu384**, **Leu387**, Leu391, Leu525
2000	Ala350, **His524**	Leu346, **Leu384**, **Leu387**, Met388, Met421, Leu525

Highlighted residues have constant interaction with coumestrol throughout the MD simulations.

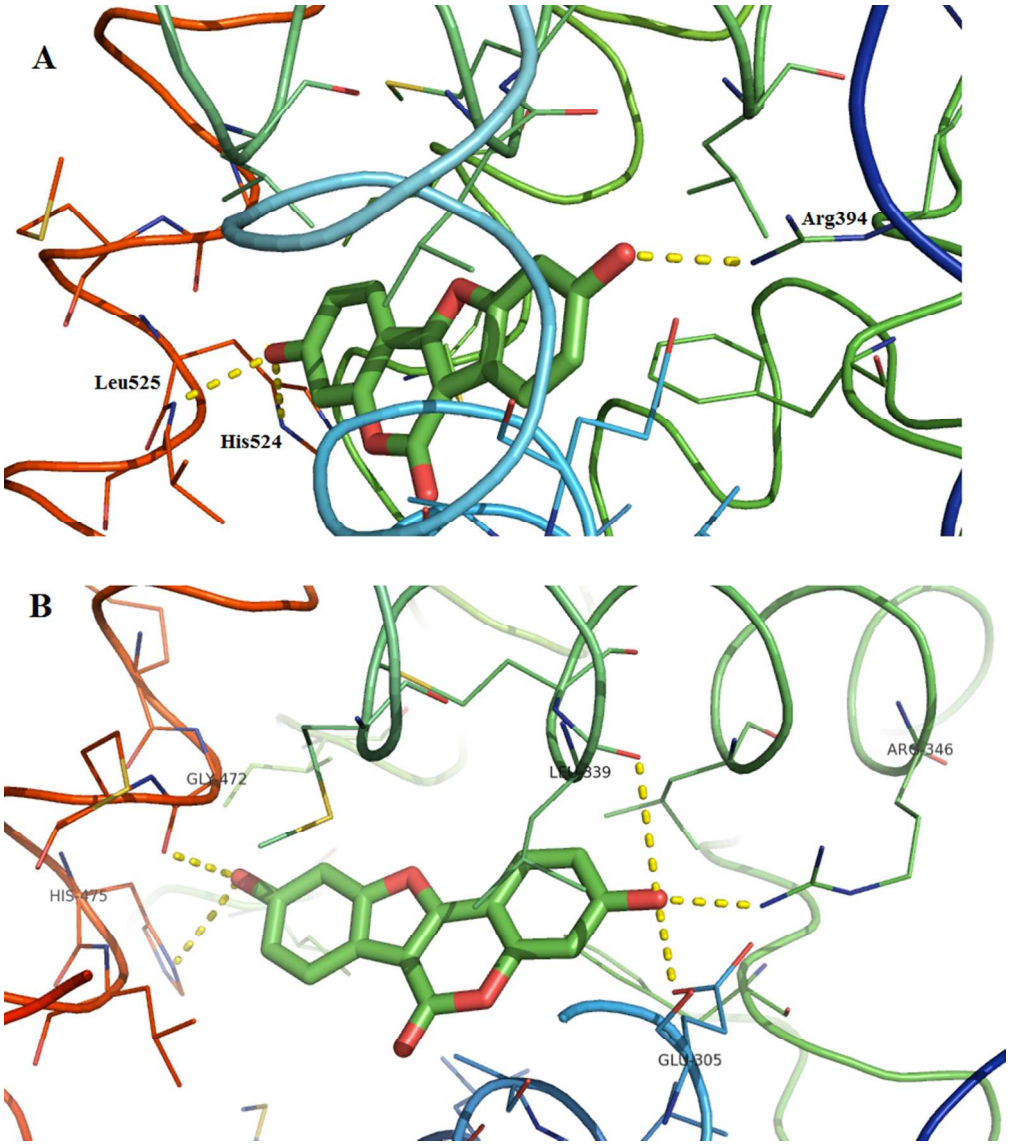
Table 4 Residues of ERβ forming H-bonds and hydrophobic contacts with coumestrol at different time intervals

Time period	Residues involved in hydrogen bonding	Residues involved in hydrophobic interactions
0	Glu305, Leu339, Arg346, Gly472, His475	Met295, Ala302, Met340, Leu343, Ile376, Leu380, **Leu476**
500	-	Met295, Leu298, Leu301, Ala302, Met336, Leu339, Leu343, Ile376, Gly472, His475, **Leu476**
1000	-	Leu298, Leu301, Met340, Leu343, Phe356, Gly472, **Leu476**
1500	Thr299	Leu298, Met336, Met340, Leu343, Gly372, Ile373, Gly472, His475, **Leu476**, Met479
2000	Thr299	Met295, Leu298, Met336, Met340, Phe356, Ile376, His475, **Leu476**

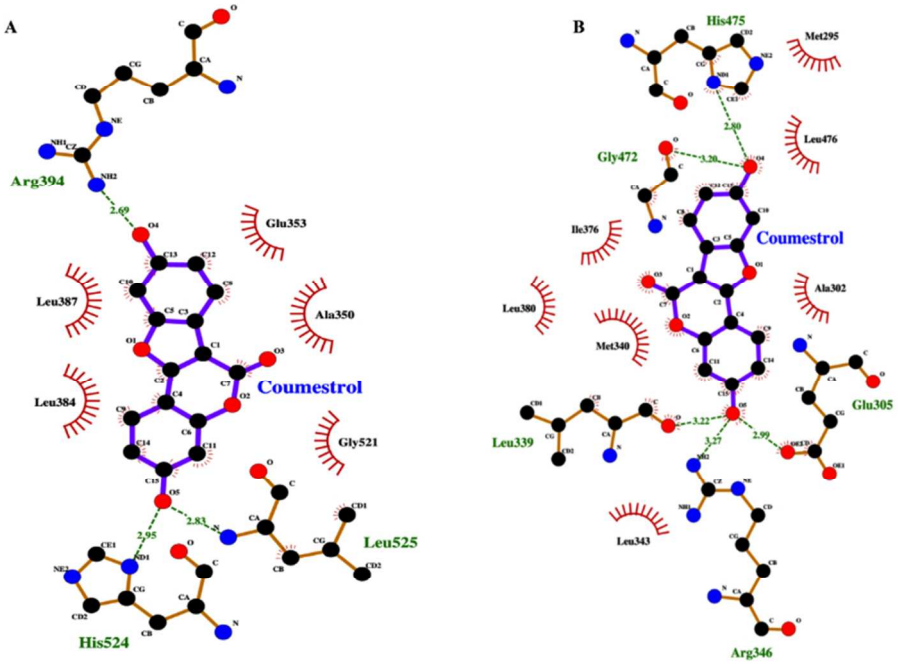
Highlighted residues have constant interaction with coumestrol throughout the MD simulations.



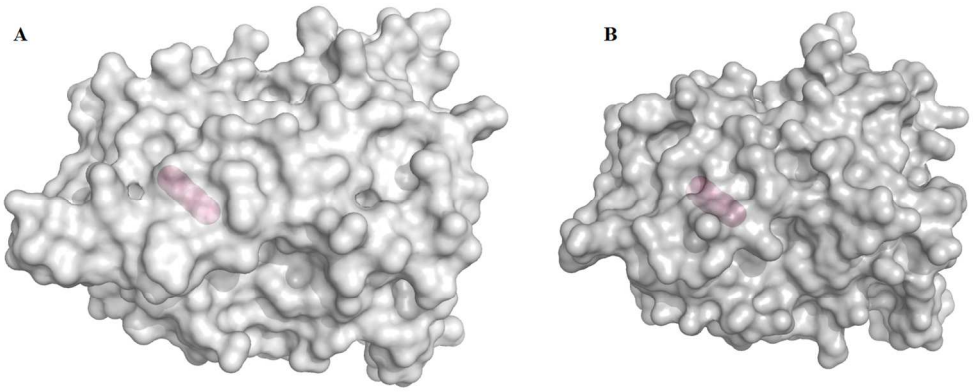
301x325mm (96 x 96 DPI)



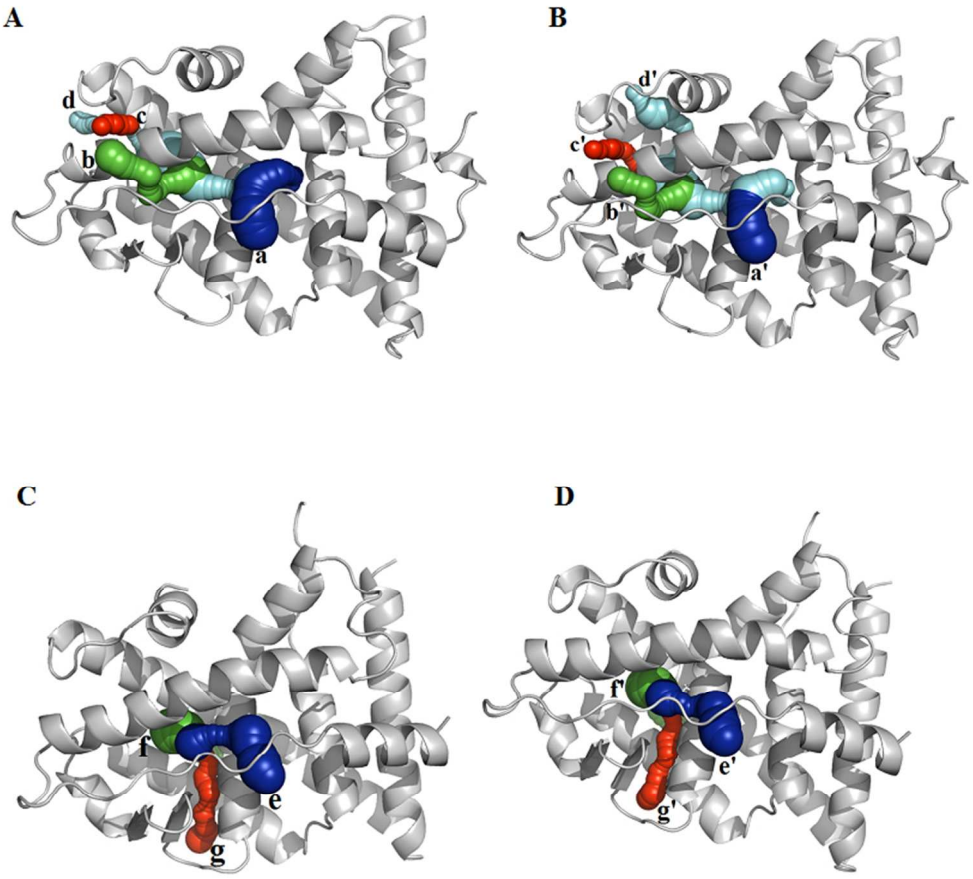
244x277mm (96 x 96 DPI)



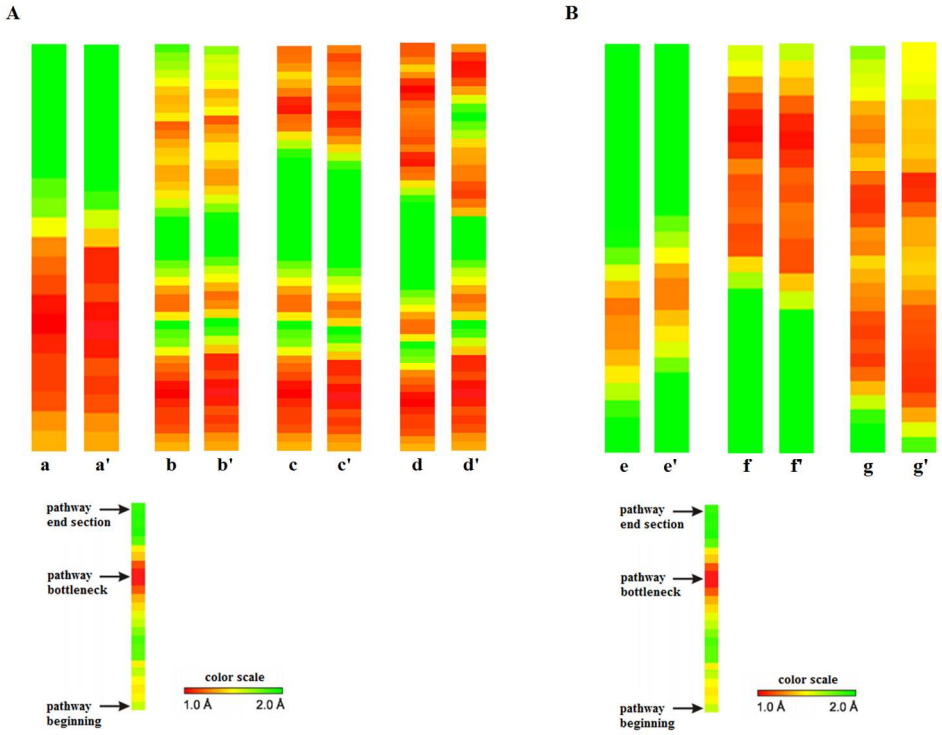
237x162mm (96 x 96 DPI)



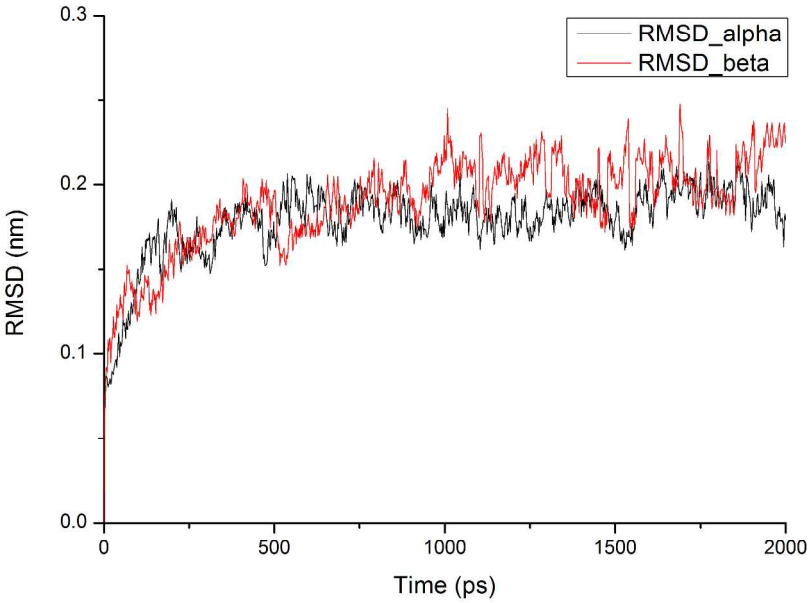
423x166mm (96 x 96 DPI)



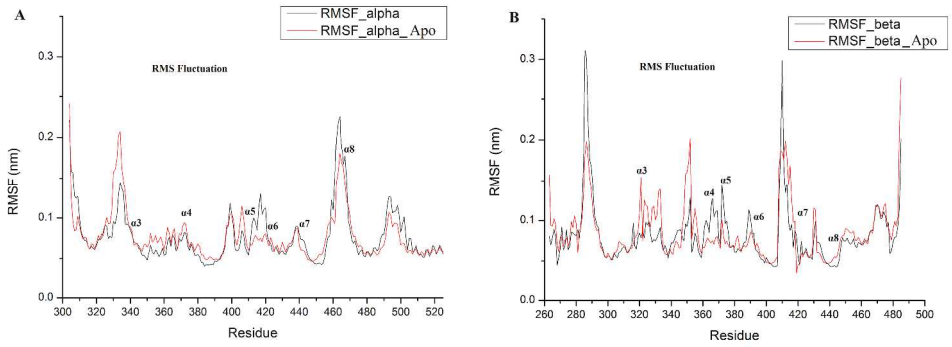
221x203mm (96 x 96 DPI)



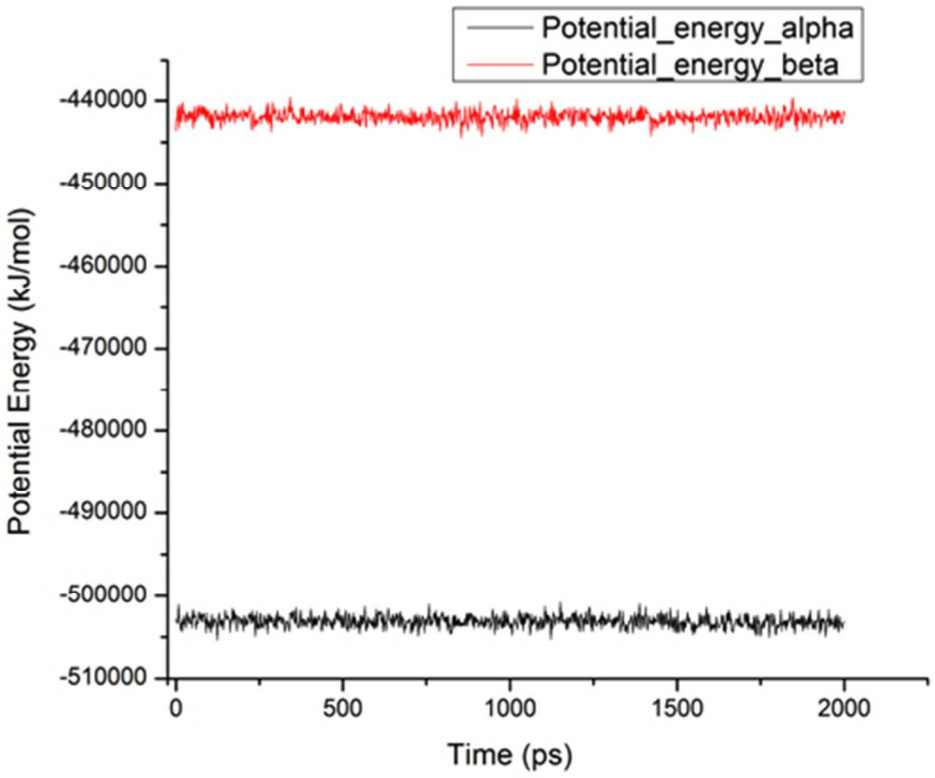
424x324mm (96 x 96 DPI)



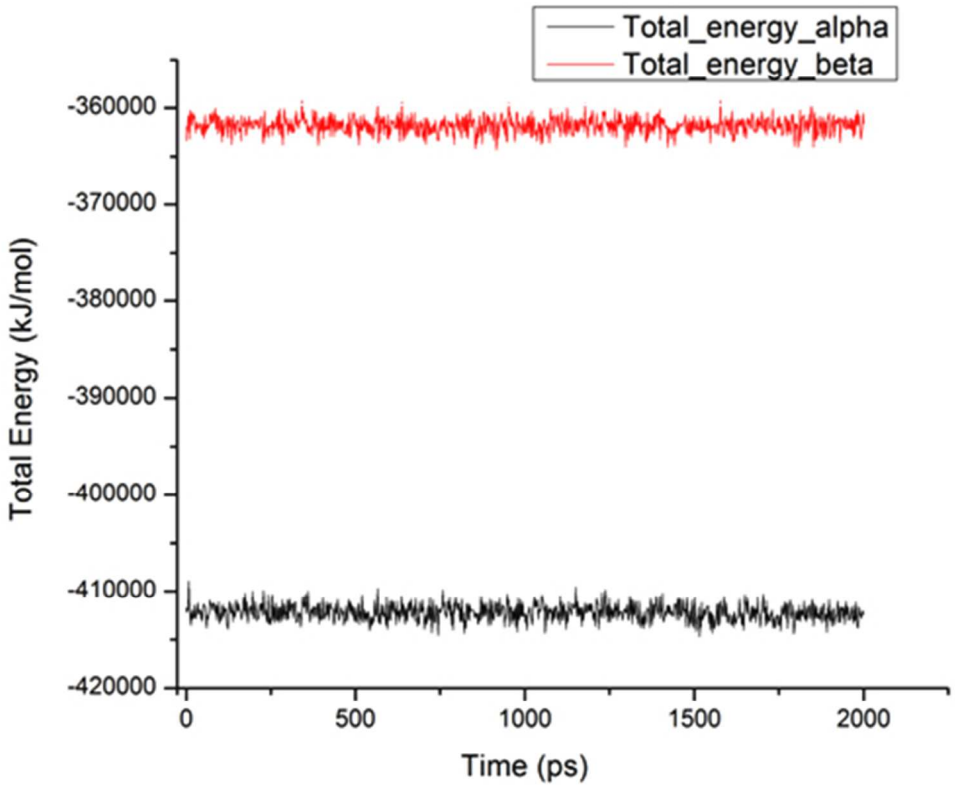
928x656mm (96 x 96 DPI)



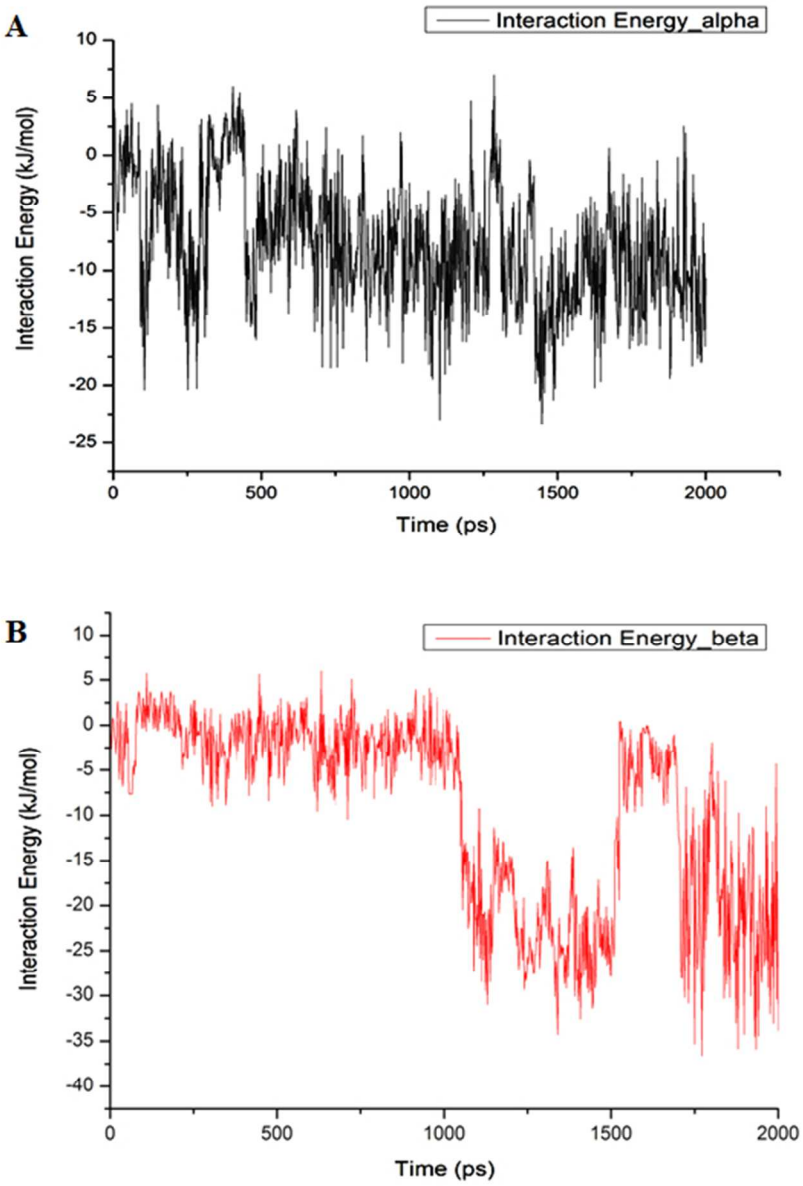
1667x656mm (96 x 96 DPI)



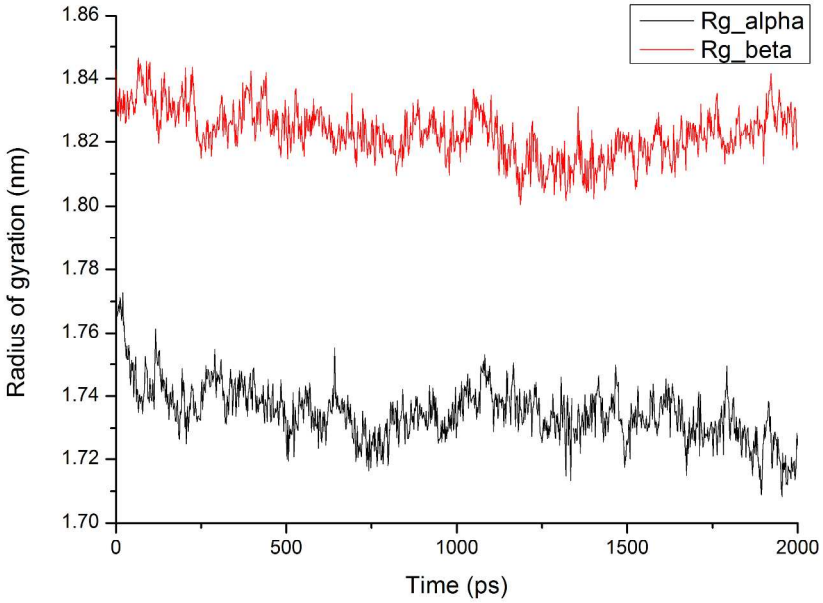
140x113mm (96 x 96 DPI)



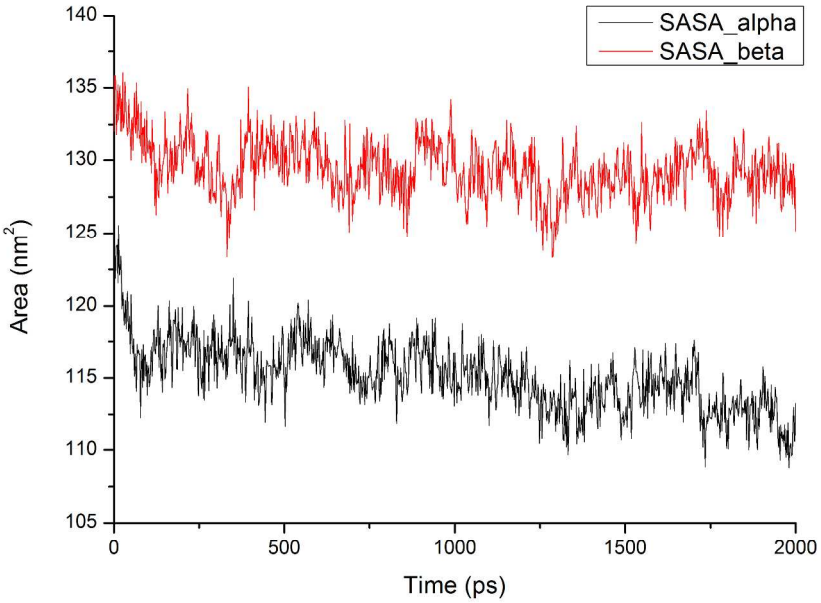
135x112mm (96 x 96 DPI)



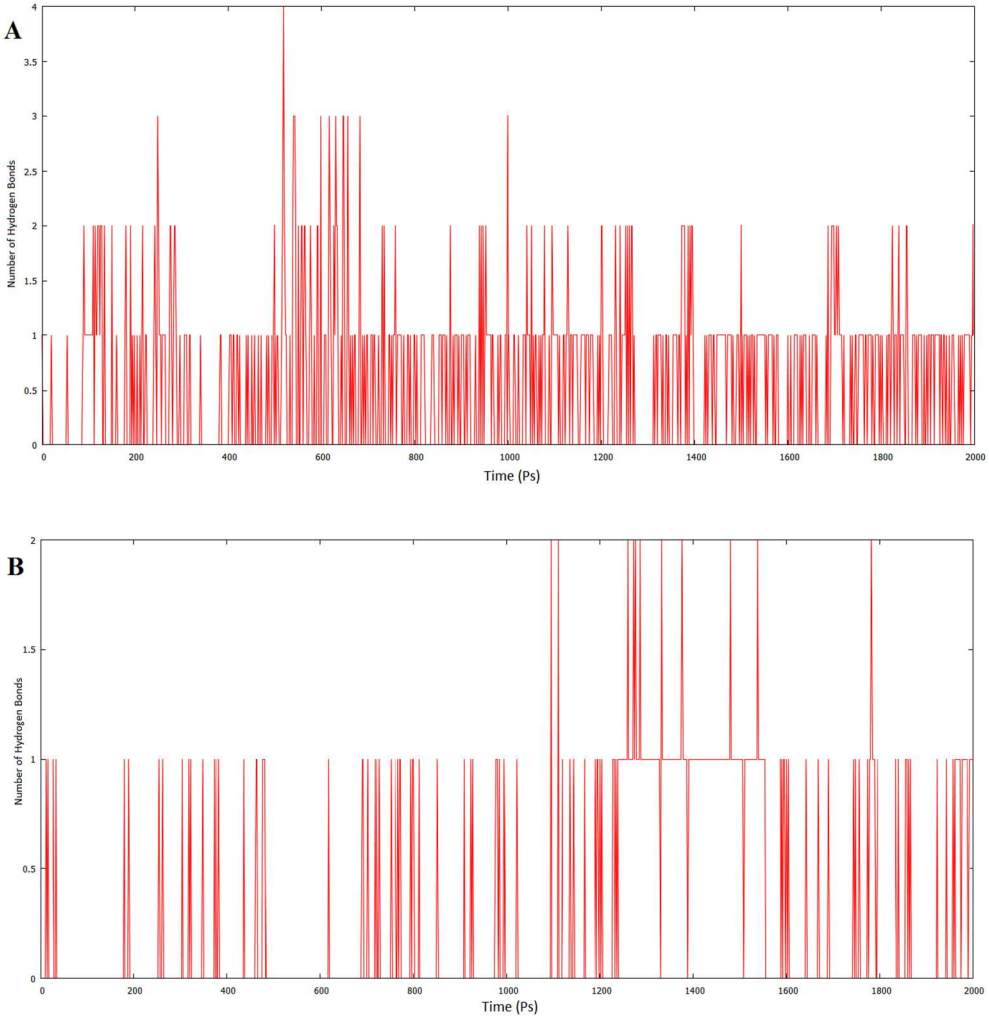
169x207mm (96 x 96 DPI)



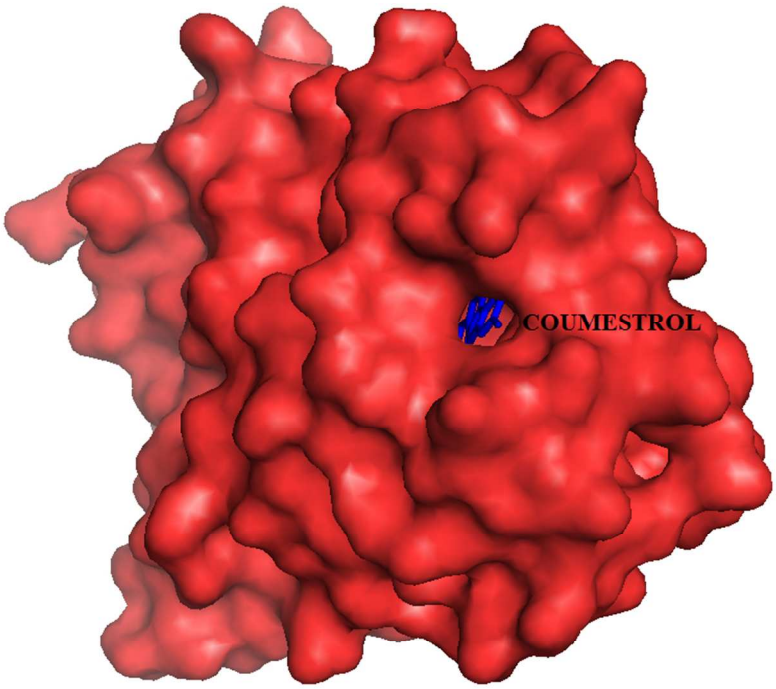
297x210mm (300 x 300 DPI)



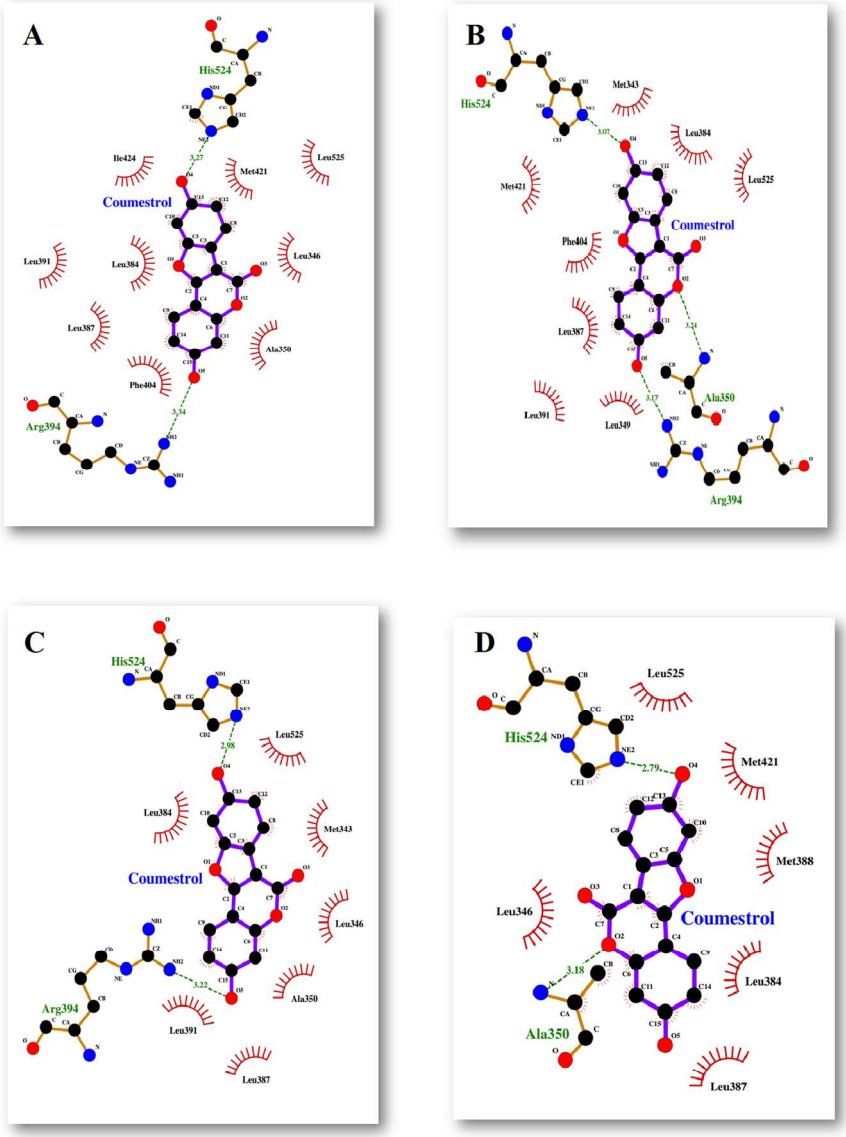
928x656mm (96 x 96 DPI)



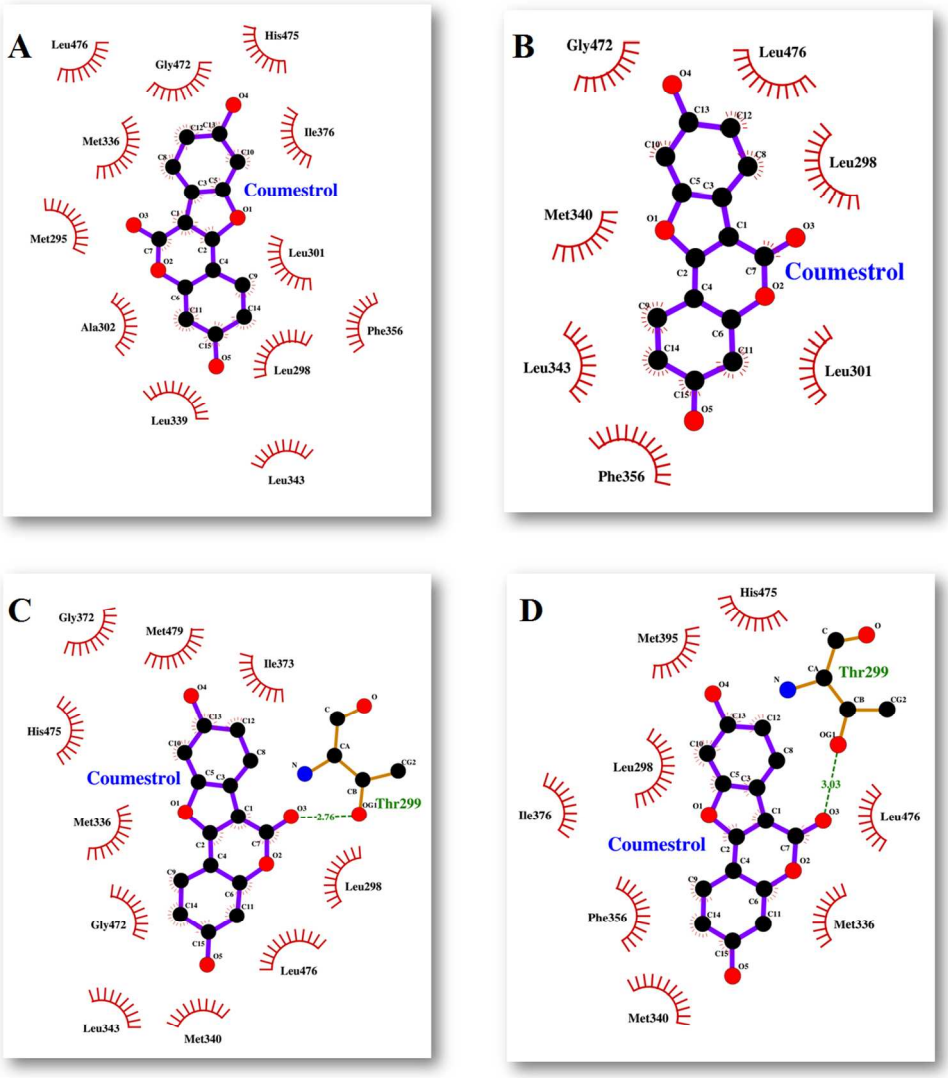
369x383mm (96 x 96 DPI)



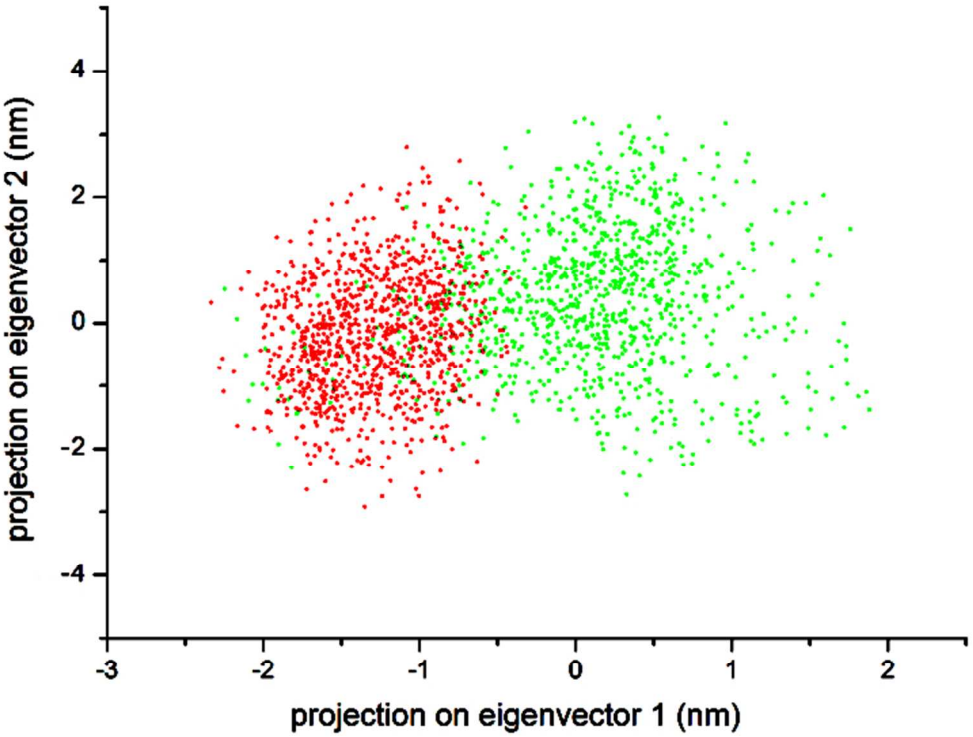
254x188mm (96 x 96 DPI)



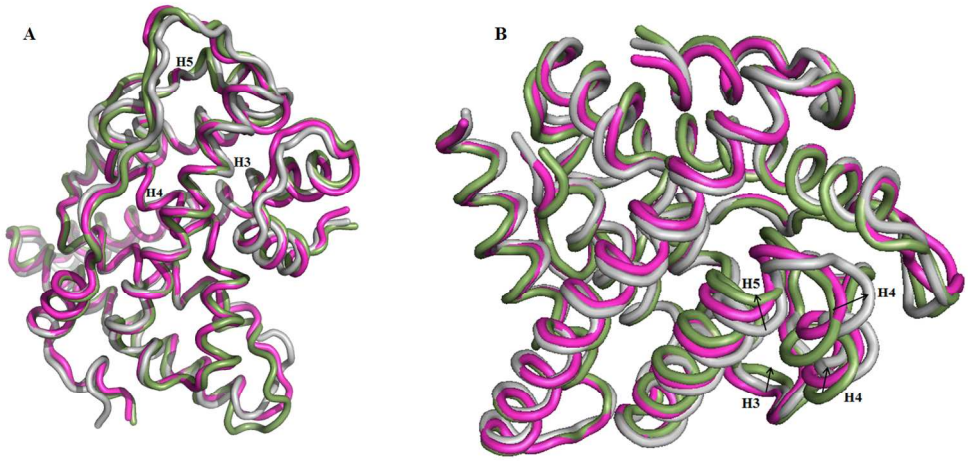
383x515mm (96 x 96 DPI)



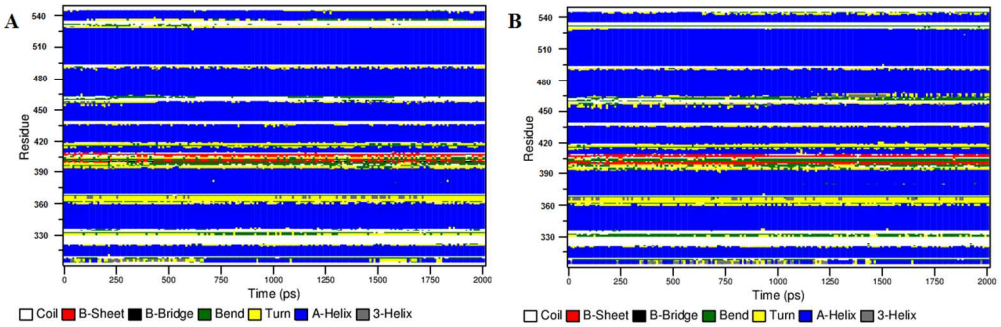
372x424mm (96 x 96 DPI)



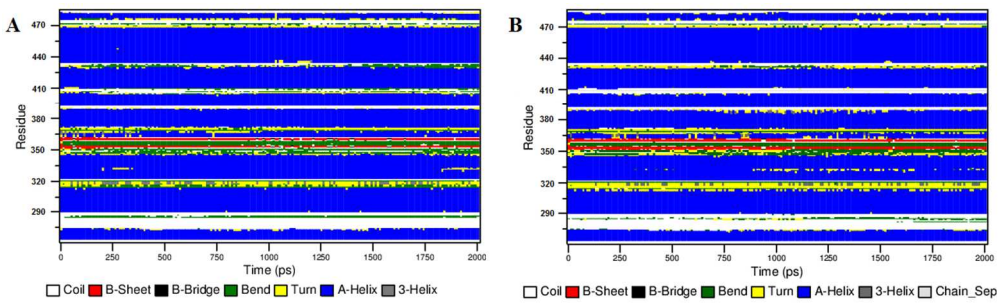
195x153mm (96 x 96 DPI)



412x206mm (96 x 96 DPI)



338x110mm (96 x 96 DPI)



339x102mm (96 x 96 DPI)