Biophysical Chemistry 170 (2012) 42–50

Contents lists available at SciVerse ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Insight into estrogen receptor beta–beta and alpha–beta homo- and heterodimerization: A combined molecular dynamics and sequence analysis study

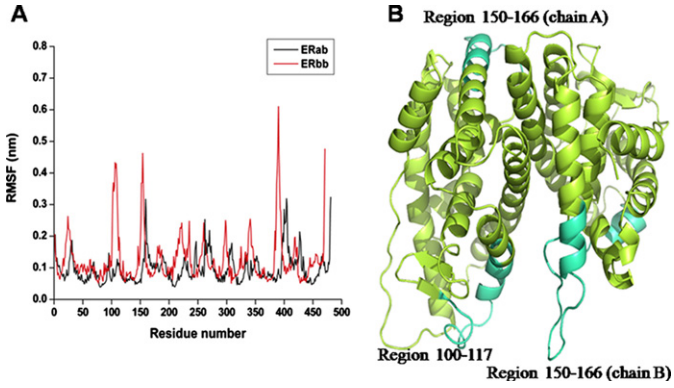
Sandipan Chakraborty [a], Hadassah Willett [a,b], P.K. Biswas [a,*]

[a] Laboratory of Computational Biophysics and Bioengineering, Department of Physics, Tougaloo College, Tougaloo, MS 39174, USA
[b] Department of Biology, Tougaloo College, Tougaloo, MS 39174, USA

HIGHLIGHTS

- ER forms highly stable heterodimer which is more stable than the ERββ homodimers.
- ERα and ERβ share very similar electrostatic potential profile of dimer interface.
- In essential subspace, the homo- and heterodimer dynamics are distinctively different.
- The dimer interface of the LBD in both the ERs is highly conserved.
- Phylogenetic analysis reveals a co-evolutionary signature between the two receptors.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 12 September 2012
Accepted 16 September 2012
Available online 25 September 2012

Keywords:
Estrogen receptor
Homo- and heterodimer
Molecular dynamics
Loop dynamics
Phylogeny
Evolution

ABSTRACT

Biological effects of estrogenic ligands are transduced by two estrogen receptors, ERα and ERβ; they transactivate as dimers. Since ERββ and ERαβ homo- and heterodimers are known to exhibit anti-proliferative effects, we characterized their dimerization interface in atomic details and explored their ligand induced conformational dynamics. ERαβ heterodimer is found to be relatively more stable than the ERββ homodimer and the observed differences are mainly due to loop dynamics. The principal component analysis reveals that, in the essential subspace, the homo- and heterodimer dynamics are distinctively different. The core recognition groove of the dimer interface, formed by helix 9 and helix 10/11, remains unaltered in both homo- and heterodimers. The dimerization surfaces are found to be highly conserved in eukaryotic lineages. Phylogenetic patterns for ERα appear to be very much similar to that of ERβ which signifies that the formation of functional heterodimer is evolutionary selected.

1. Introduction

Biological functions of estrogenic compounds are generally modulated by two estrogen receptors, namely ERα and ERβ. Both the proteins play crucial role in various physiological functions in the reproductive tract, breast, prostate, bone, brain and the cardiovascular system [1,2]. ERβ is expressed mainly in the central nervous system, cardiovascular system, immune system, urogenital tract, gastrointestinal tract, kidneys and lungs [3–7]. ERα, on the other hand, is mainly expressed in mammary glands [8]. It is to be noted that ERβ also shows a considerable expression level in mammary glands [9]. Literature reveals that the biological functions of ERβ are quite distinct from the ERα. ERα is known to induce the transcription of pro-proliferative target genes, whereas several studies indicate that ERβ exerts anti-proliferative and pro-apoptotic [10–12] effects. ERβ is known to act as growth modulator in ERα positive breast cancer cells [13]. Despite their marked differences of biological function, both the ERα and ERβ share a very similar structural topology and a common mechanism of transactivation. Both ERα and ERβ contain five major domains: an N-terminal activation

* Corresponding author. Tel.: +1 601 977 7788; fax: +1 601 977 7898.
E-mail address: pbiswas@tougaloo.edu (P.K. Biswas).

0301-4622/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bpc.2012.09.002

function-1 domain (AF-1), a central DNA-binding domain (DBD), a hinge region, a C-terminal ligand-binding domain (LBD) and a C-terminal transactivation domain (AF2). The LBD is considered to be the key structural domain that controls ligand-dependent regulation of ER signaling. Ligand binding in the LBD induces conformational changes that facilitate receptor dimerization, either as homodimer (ERα/ERα or ERβ/ERβ) or heterodimer (ERα/ERβ), translocation of dimers to the nucleus, and binding with co-regulatory proteins [14]. This supra-molecular assembly finally interacts to the estrogen response element (ERE) in the promoter region of target genes [14]. Receptor dimerization is one of the crucial steps in ER transactivation and hence the physiological responses of a particular ER depend on the activation of a particular dimer pair by the ligand [13]. Homodimerization of ERα accelerates cellular proliferation [13]. On the other hand, the transcriptional activation of ERβ homodimer is thought to be protective against hormone-dependent diseases including breast and prostate cancers [10–12]. Additionally, ERβ is thought to counteract the stimulatory effects of ERα through the formation of functional heterodimer with ERα [15,16]. Experimental evidence suggests that heterodimer induced activation of target genes is markedly different from that induced by homodimers [17,18]. Thus, the activation and agonist induced dynamics of ERαβ heterodimer and ERββ homodimer need to be explored to design more potent ligand that selectively activate the ERαβ heterodimer and ERββ homodimer.

Structural details of ERα/ERβ heterodimer are not known due to non availability of crystal structures. Accordingly, the effect of ligand binding in heterodimers was never been explored. Recently, we have characterized the essential dimer contacts of ERα homodimer in atomic details using *in silico* modeling approach [19]. It is to be noted that the crystal structures of ERβ homodimer in the presence of several bound antagonist are available in Protein Data Bank. But the structural dynamics of ERβ homodimer in the presence of bound agonist was less explored. In this paper, we characterize the ERββ homodimer and ERαβ heterodimer and their dimerization interface in atomic details using *in silico* modeling techniques. We also explore the ligand induced conformational dynamics of both homo- and heterodimers using all atom molecular dynamics simulation. Our simulation reveals a similar activation mechanism of ER homo- and heterodimers in terms of receptor dimerization. We have further explored the sequence similarities and phylogenetic relations between the two ER receptors. The dimerization surface is highly conserved in both the receptors. The observed phylogenetic pattern for ERα parallels with the ERβ signifying co-evolution of both the proteins.

2. Materials and methods

2.1. Modeling the ER homo- and heterodimer

The crystal structure of ERβ LBD homodimer (PDB ID: 1U3S) where each monomer is bound with a highly potent and selective ERβ agonist ligand aryl diphenolic azole (WAY-797) has been considered as ERβ LBD dimer agonist conformation in our study. In this structure, helix 12 is positioned properly to accommodate co-activator proteins. In the crystal structure, a considerable section of residues in the dimerization surface were missing. In chain A, residues 412–419 and in chain B, residues 410–421 were missing. All the missing residues were modelled by using MODELLER 9.9 [20]. No residues were added in the N and C-terminal of the dimer structure. Thus, our final model structure contains 235 residues (263–497) for each chain of ERβ dimer.

Crystal structure does not exist for ERαβ heterodimer. To model one, we primarily superimposed a monomer from ERβ LBD homodimer (PDB ID: 1U3S) on a monomer from the crystal structure of ERα LBD homodimer (PDB ID: 3ERD). This is to emphasize that both 3ERD and 1U3S represent agonist conformations of ERα and ERβ, respectively.

2.2. Molecular dynamics simulations

All simulations were performed using GROMACS [21,22] molecular dynamics code with OPLS [23] force field. The parameters for diethylstilbestrol (DES) and aryl diphenolic azole (797) were developed according to the OPLS force field. Atom definitions, atom types and partial charges were assigned according to the symmetry group analogy in OPLS-AA set [23]. The atomic partial charges are readjusted to maintain the charge neutrality of the whole molecule. The parameters are tested by comparing the GROMACS energy minimized structures with the respective crystal structures. Both the homo- and heterodimers were initially subjected to a short energy minimization *in vacuo* using steepest descent algorithm to remove any bad contacts for added hydrogens. Both the systems were then solvated with SPC explicit water model in a cubic box with periodic boundary conditions. The box dimensions are chosen such that all the protein atoms are at least 1 nm away from the edge of the box. The ionization state of residues was set to be consistent with neutral pH and counter ions were added to neutralize the system. The solvated system was then subjected to a second short energy minimization using steepest descent algorithm to eliminate any bad contacts with added water. After that, a 100 ps position restrained dynamics was carried out where the complex was restrained by restraining forces while the water molecules were allowed to move. It was then followed by 20 ps of NVT simulation at 300 K and 20 ps of NPT simulation to achieve proper equilibration of the simulated system. Final production simulations were performed in the isothermalisobaric (NPT) ensemble at 300 K, using an external bath with a coupling constant of 0.1 ps. The pressure was kept constant (1 bar) by using pressure coupling with the time-constant set to 1 ps. The LINCS [24] algorithm was used to constrain the bond lengths involving hydrogen atoms, allowing the use of 2.0 fs time step. Electrostatic interactions were calculated using the reaction–field method. Van der Waals and Coulomb interactions were truncated at 1.4 nm and the trajectories were stored at every 5 ps.

Structural analysis was carried out by using the in-built tools of GROMACS and the secondary structure assignments were performed with DSSP [25] module integrated with GROMACS. The RMSD matrices were computed on each of the trajectories by the least square fitting of main-chain atoms and the matrices were then processed to extract clusters of similar conformations.

2.3. Principal component analysis (PCA)

Principal component analysis (PCA) was performed on each of the dimer trajectories obtained from the MD simulations in order to obtain an insight into the essential dynamics of the simulated system in reduced subspace. Mass-weighted covariance matrix of atomic positional fluctuations was calculated from each MD trajectory on Cα atoms. Then the essential subspace was defined in terms of the calculated eigenvectors.

2.4. Identification of hotspot residues at the dimer interface

The essential protein–protein interaction interfaces of the homo- and heterodimers were explored by using HotPoint online web-server [26]. The server also predicts hotspot residues that are involved in crucial interactions in the dimer interface.

2.5. Multiple sequence alignments and Phylogenetic analysis

ERα and ERβ protein sequences of different organisms were obtained from NCBI (www.ncbi.nlm.nih.gov) protein sequence database. Redundant sequences were removed by using the CD-HIT program [27] using a sequence identity cut-off of 0.95. Multiple sequence alignment (MSA) was carried out using MUSCLE [28] program. The MSA was then used to draw the phylogenetic tree using PhyML [29]

web-interface and the generated tree was rendered using TreeDyn [30] program.

3. Result and discussions

3.1. Analysis of structural stability of homo- and heterodimers during MD simulation

Molecular dynamics simulations have been performed on both ERββ homodimer and ERαβ heterodimer to provide insights into the stability and structural dynamics of the dimers in presence of bound agonist ligand. Fig. 1 depicts the variation of dynamic parameters (RMSD and radius of gyration, R_g) of the simulated system as a function of time.

A root mean square deviation (RMSD) of the protein over the simulation time has been used to analyze the stability of the simulated system. As evident from Fig. 1A, during first 1 ns of the simulation, both the systems undergo structural readjustment according to its environments and bound ligands and reach the equilibration. RMSD profiles obtained from MD simulation reveal that ERαβ heterodimer is comparably more stable than the ERββ homodimer. Radius of gyration (R_g) is the mass-weighted root mean square distance of a collection of atoms from their common centre of mass and it defines the overall shape and dimensions of the protein. The plot of variation of radius of gyration of each LBD dimer with simulation time is shown in Fig. 1B. As evident from the figure, ERαβ heterodimer is more stable than the ERββ homodimer which is reminiscent with the observation obtained in RMSD analysis. The R_g of ERαβ heterodimer remains stable at 2.27 ± 0.007 nm throughout the simulation period signifying a highly stable dimer complex. On the other hand, for ERββ homodimer, during the initial 4 ns of simulation, the R_g increases from 2.23 to 2.30 nm and in the remaining simulation time it stabilizes at 2.30 ± 0.011 nm. To gain a detailed insight into all the visited conformations during MD simulation for both homo- and heterodimers, we have computed the RMSD matrices and the results are displayed in Fig. 1C and D. ERαβ heterodimer mainly exists in two very similar conformational clusters. The first cluster represents dimer conformations from first 4 ns of simulation. Second one represents conformations of the remaining simulation time. It is to be noted that the conformations between these two clusters are also very similar with an overall RMSD difference of 0.2 nm. In the case of ERββ homodimer, we obtained three conformational clusters from MD simulation. Homodimer conformations from first 4 ns of simulation form the first cluster. In between 4 and 6 ns of simulation time, the dimer undergoes appreciable conformational changes compared to the initial conformation to generate the second cluster. The average RMSD of the ERββ homodimer between the two conformational clusters is ~0.25–0.3 nm. In the remaining simulation time, the dimer undergoes further conformational changes to generate the third cluster with an overall RMSD difference of 0.3–0.4 nm from the initial structure.

3.2. Residue fluctuation analysis

We then analysed the RMSF (root mean square fluctuations) per residues to identify the regions responsible for the observed structural differences in the ER homo- and heterodimer during MD simulation. Also

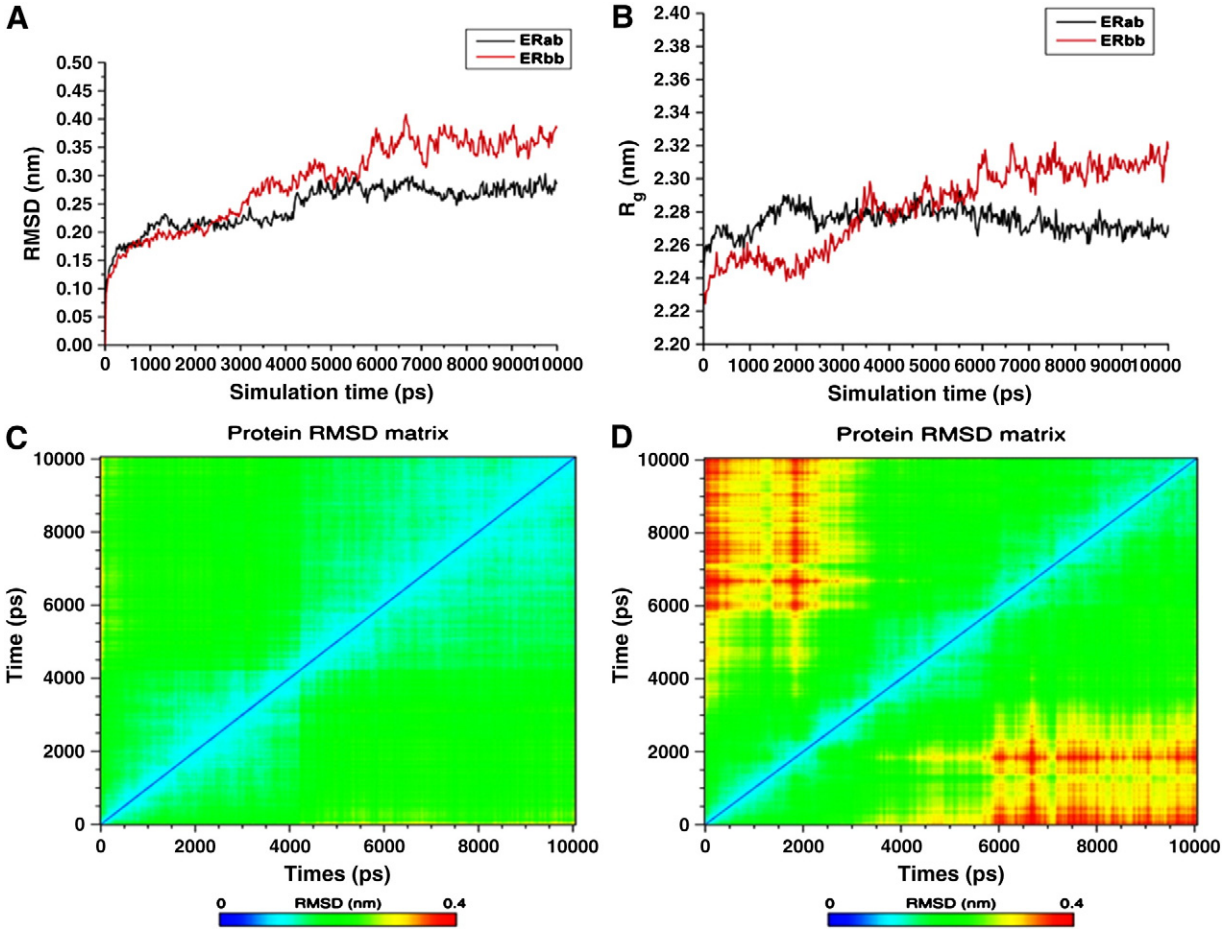


Fig. 1. Variations in C_α-RMSD (A) and radius of gyration R_g (B) of ER dimer with simulation time. Black line represents ERαβ heterodimer while the red line represents ERββ homodimer. RMSD matrices of ERαβ heterodimer (C) and ERββ homodimer (D) computed from MD trajectory.

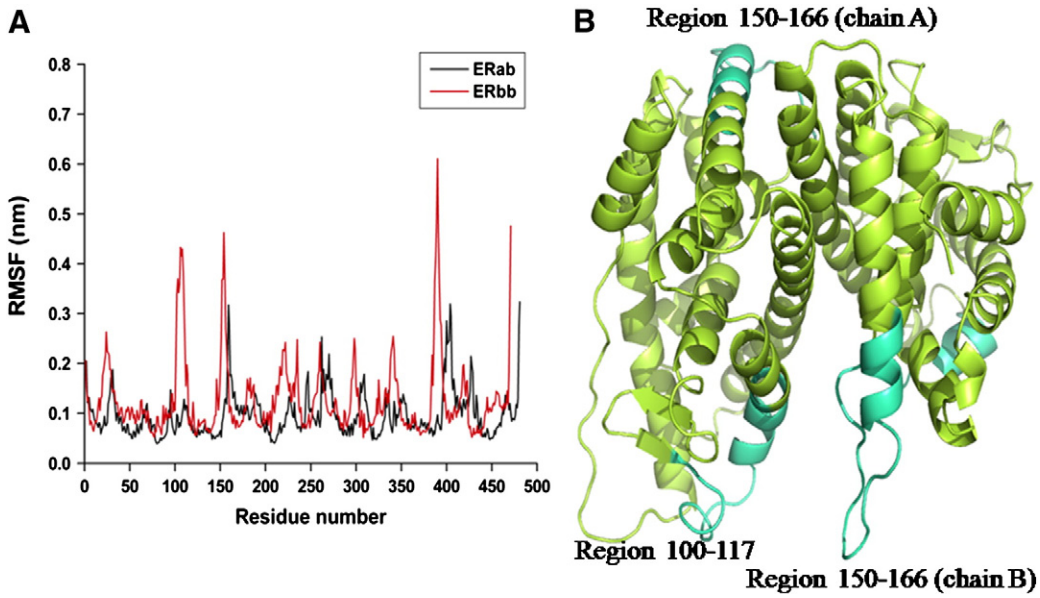


Fig. 2. (A) Variations of RMSF profiles of ER homo- and heterodimer. (B) A cartoon representation of ERββ homodimer. Regions of high fluctuations are colored in cyan.

RMSF helps us to understand the structural basis of the observed high fluctuation in RMSD for ERββ homodimer. Results are summarized in Fig. 2.

In general, the residue fluctuations for ERββ dimer are higher than ERαβ heterodimer. It is to be noted that there is a sequence difference between ERα and ERβ. Also, the size of their LBD is slightly different: ERα LBD comprises of 245 residues while ERβ LBD comprises of 235 residues. Fig. 2A reveals that the RMSF profiles of both homo- and heterodimers are qualitatively very similar though the number of amino acids and the sequences are not the same for the two proteins. We then analyzed the critical differences between the two RMSF profiles to infer the observed high RMSD for ERββ homodimer. We found that the high fluctuations of the homodimer are mainly due to the loop dynamics; no structural uncoiling has been observed in the MD simulation. Particularly, the loop region between the sheet and helix 6 of chain A in ERββ homodimer (from residue 100 to 117) is found to be highly flexible during the MD simulation but the region is not much flexible for ERαβ heterodimer. Another region is of interest is the long loop region between helix 8 and helix 9 (from residue 150 to 166). This region is found to be flexible in both the monomer of the homo- and heterodimer. It is to be noted that more flexibilities have been observed for this region in the homodimer in comparison to the heterodimer during MD simulation. Also, helix 12

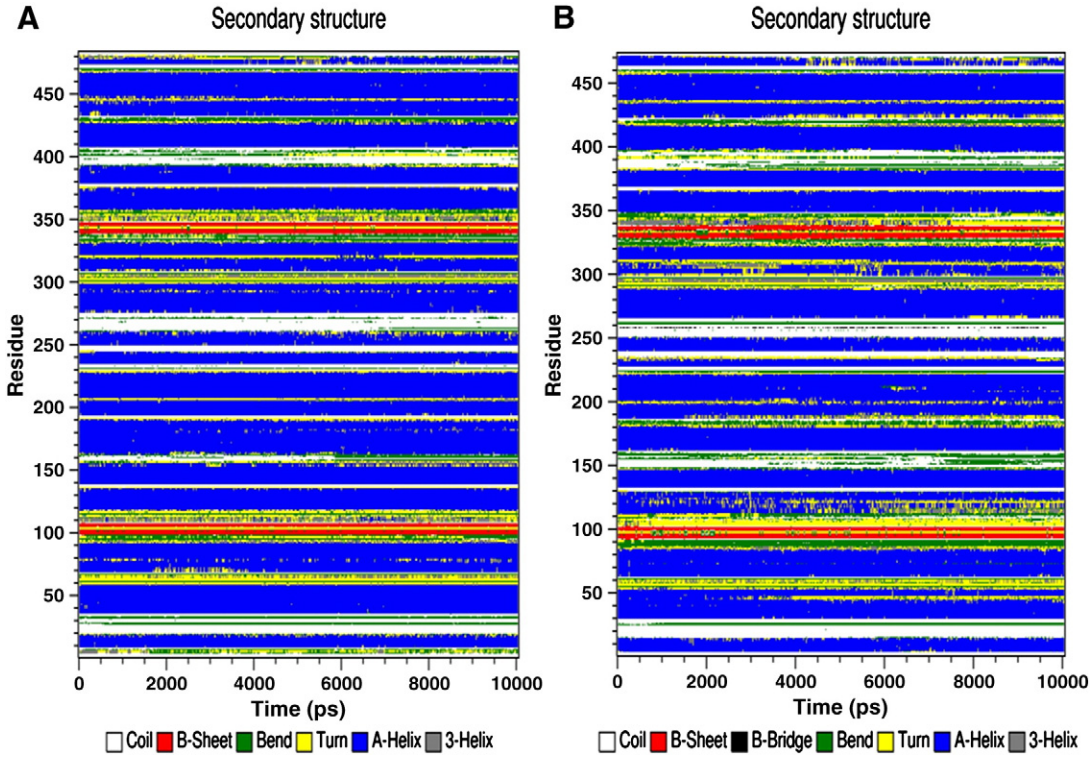


Fig. 3. Secondary structure evolution with simulation time. (A) ERαβ heterodimer. (B) ERββ homodimer.

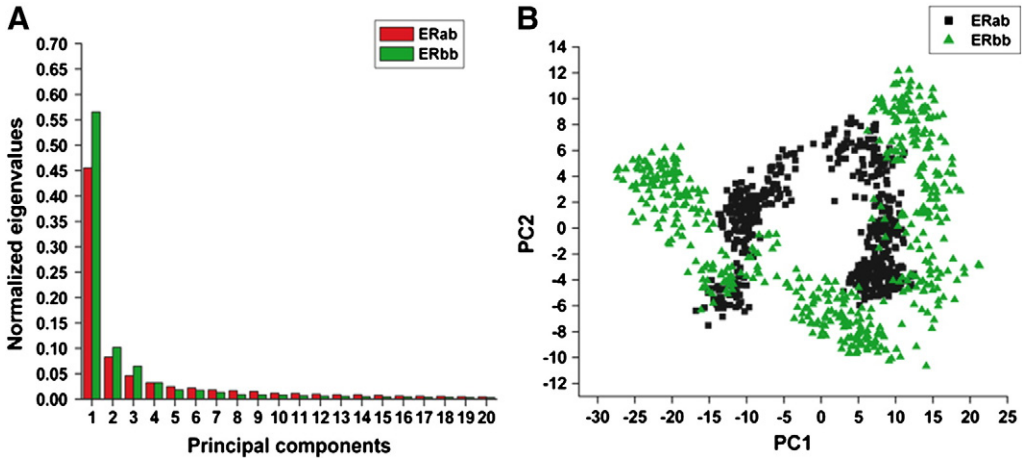


Fig. 4. (A) Normalised eigenvalue distribution for the first 20 eigenvectors obtained from MD trajectory. (B) 2-D projection of the first two principal components (1 and 2) for ERαβ heterodimer (black) while green represents ERββ dimer.

has been found to be more flexible in ERββ homodimer compared to the heterodimer. Recently, we explored the essential dimerization surface between ERαα homodimer and found that core recognition groove is formed by helix 9, helix 10 and 11 [31]. Here we observed that the core dimer interaction groove is highly stable for the both homo- and heterodimers during MD simulation.

3.3. Analysis of secondary structure of the dimers during MD simulation

We have further analyzed the secondary structure evolution of the homo- and heterodimers during the simulation. It is to be noted that the chain B represents ERβ in both the homo- and heterodimers and chain A represents ERα in ERαβ and ERβ in ERββ. The secondary structural profiles of chain B are very much similar in ERαβ and ERββ, as was expected. Alterations in the secondary structures have been observed in chain A. There are more helical content in ERα compared to ERβ. Particularly, for homodimer, the N-terminal portion of helix 6 loses its helicity in later period of the simulation and appears as helical turn. Another region of interest is the loop connecting helix 8 and helix 9. In the later period of simulation this loop length increases and some of the N-terminal residues of helix 9 lose helicity and appear as turn/bend which is in accordance with the RMSF profile for homodimer (Fig. 2). It is to be noted that the general helical topology of both the homo- and heterodimers remains unaltered during the simulation; only some of the linker helical regions attached with the loop region lose their structure during the later period of simulation for the heterodimer (Fig. 3).

3.4. Principal component analysis (PCA)

Recently, we have demonstrated that the essential dynamics of ERαα homodimer is highly dependent on the nature of the bound ligand in the LBD [31]. The essential motion of the ERαα homodimer

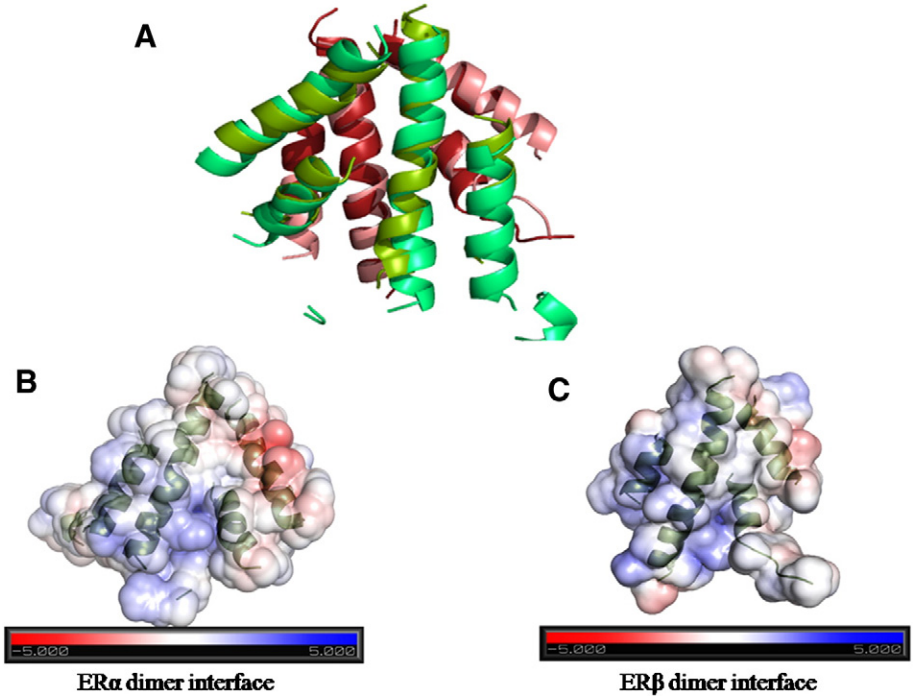


Fig. 5. The essential dimerization interface of ER homo- and heterodimer interface. Green color represents chain A and red color represents chain B. Dark green and salmon represents ERββ homodimer interface, while light green and salmon represent ERαβ heterodimer. B and C represent electrostatic potential of the dimer interface.

with bound SERM (4-OHT) in the LBD is distinctively different from the motion of the dimer when a pure antagonist is bound to the LBD [31]. To further characterize the behavior of ERββ homo and ERαβ heterodimer in the essential subspace, we have performed the principal component analysis (PCA) based on the MD trajectories and the results are displayed in Fig. 4. As is evident from the eigenvalue distribution plot for the homo- and heterodimers (Fig. 4A), first few components are more informative with high normalized eigenvalues while the latter components carry very little information about the dimer dynamics (Fig. 4A). First two PCs can explain 67% of the variance observed in the MD simulation in the case of ERββ homodimer, while 53% variance is explained by the corresponding two PCs for the ERαβ heterodimer. Fig. 4B represents the bi-dimensional projections of the first two PCs. It is evident from the figure that PC1 varies more broadly compared to PC2 in both the dimers. This corroborates well to the eigenvalue distribution plot where PC1 has the highest eigenvalue (Fig. 4A). Interestingly, 2-D projections of PC1 vs. PC2 are very different for the homo compared to the heterodimers. The 2-D projections have been appeared to be in ∩-shaped distribution pattern for ERαβ heterodimer, while for the ERββ homodimer it appears to be in U-shaped patterns. It is to be noted that the distribution pattern for homodimer is also very much similar to the observed 2-D projections for the ERαα homodimer when an antagonist ligand is bound to the LBD and the 2-D projections of first two principal components for the heterodimer is similar with the distribution pattern of ERαα homodimer with bound SERM (4-OHT) to the LBD [31]. Thus, in the essential subspace the ER homo- and heterodimers are quite different.

3.5. Analysis of essential dimer surface

We then analysed the essential dimer surface for both ERββ homodimer and ERαβ heterodimer and results are shown in Fig. 5A.

The dimer interface essentially comprises of helix 7, C-terminal region of helix 8, helix 9 and helix 10/11. It is to be noted that in ER, helix 10 and 11 is a contiguous helix. The dimer interface is much larger for the heterodimer compared to the ERββ homodimer. Particularly, the N-terminal regions of helix 7 and the C-terminal region of helix 8 lose their contribution from the dimer interface for ERββ homodimer. Analysis of the trajectory reveals that the helicity of these two regions diminishes during the later half of the simulation period and the region appears as a highly flexible turn region for ERββ homodimer. It is to be noted that the ER dimerizes in an anti-symmetric orientation where the topology of one monomer is opposite to the other monomer. Electrostatic potential calculation on both ERα and ERβ reveals highly similar electrostatic potential profile on the dimer surface for both the receptors. The dimerization surface contains two distinct regions of opposite electrostatic potential on two opposite edges of the surface for both

Table 1
List of "hotspot" residues of ER homo- and heterodimer interface with their calculated solvent accessible surface area (SASA). Residues those SASA change > 30 Å^2 upon dimerization are marked in bold.

ERββ homodimer						ERαβ heterodimer					
Residue number	Residue	Chain	RelCompASA	RelMonomerASA	ΔASA	Residue number	Residue	Chain	RelCompASA	RelMonomerASA	ΔASA
382	**A**	**A**	**29.47**	**74.14**	**44.67**	426	D	A	3.41	30.32	26.91
407	**N**	**A**	**2.13**	**32.14**	**30.01**	427	M	A	4.33	26.61	22.28
410	**M**	**A**	**21.27**	**69.37**	**48.1**	430	A	A	2.08	22.53	20.45
412	**P**	**A**	**18.7**	**87.24**	**68.54**	434	R	A	37.2	60.27	23.07
431	N	A	49.52	70.65	21.13	**455**	**N**	**A**	**2.43**	**33.33**	**30.9**
434	**T**	**A**	**0.07**	**41.77**	**41.7**	**459**	**Y**	**A**	**18.81**	**71.17**	**52.36**
438	**V**	**A**	**9.83**	**52.89**	**43.06**	476	H	A	12.54	34.57	22.03
449	**Q**	**A**	**38.61**	**97.03**	**58.42**	**483**	**T**	**A**	**0.54**	**47.43**	**46.89**
452	**S**	**A**	**6.2**	**46.92**	**40.72**	487	I	A	28.15	47.06	18.91
453	**M**	**A**	**15.73**	**63.18**	**47.45**	**501**	**H**	**A**	**8.22**	**68.67**	**60.45**
456	**A**	**A**	**0**	**45.82**	**45.82**	**502**	**Q**	**A**	**18.17**	**63.14**	**44.97**
459	**L**	**A**	**1.19**	**43.98**	**42.79**	504	L	A	0	18.62	18.62
460	**M**	**A**	**6.42**	**67.15**	**60.73**	**505**	**A**	**A**	**0.1**	**50**	**49.9**
463	**S**	**A**	**3.94**	**72.92**	**68.98**	508	L	A	0.1	26.69	26.59
466	R	A	7.75	37.07	29.32	**509**	**L**	**A**	**6.22**	**64.07**	**57.85**
467	**H**	**A**	**23.47**	**58.65**	**35.18**	511	L	A	1.56	19.91	18.35
379	M	B	6.22	26.37	20.15	**512**	**S**	**A**	**2.15**	**63.4**	**61.25**
382	**A**	**B**	**26.4**	**63.24**	**36.84**	**513**	**H**	**A**	**9.52**	**62.53**	**53.01**
407	N	B	23.5	46.23	22.73	**515**	**R**	**A**	**12.85**	**52.07**	**39.22**
410	**M**	**B**	**39.34**	**79.18**	**39.84**	516	**H**	**A**	**20.58**	**49.9**	**29.32**
411	Y	B	55.74	71.78	16.04	519	N	A	43.73	65.9	22.17
434	**T**	**B**	**2.42**	**52.73**	**50.31**	549	L	A	97.03	119.52	22.49
438	**V**	**B**	**3.16**	**62.96**	**59.8**	382	A	B	22.23	47.14	24.91
452	**S**	**B**	**2.96**	**67.96**	**65**	386	R	B	28.95	54.23	25.28
453	**M**	**B**	**13.29**	**64.64**	**51.35**	**407**	**N**	**B**	**2.58**	**52.74**	**50.16**
453	L	B	0.34	24.42	24.08	**410**	**M**	**B**	**11.92**	**95.95**	**84.03**
456	**A**	**B**	**0**	**43.95**	**43.95**	**412**	**P**	**B**	**29.95**	**72.01**	**42.06**
459	L	B	0.22	25.42	25.2	**413**	**L**	**B**	**42.35**	**102.37**	**60.02**
460	**M**	**B**	**3.11**	**52.74**	**49.63**	**434**	**T**	**B**	**4.73**	**42.83**	**38.1**
463	**S**	**B**	**0.45**	**61.95**	**61.5**	435	D	B	44.43	59.9	15.47
464	**H**	**B**	**11.05**	**44.82**	**33.77**	**438**	**V**	**B**	**0.1**	**59.2**	**59.1**
466	R	B	18.35	37.6	19.25	**452**	**S**	**B**	**8.33**	**63.72**	**55.39**
467	H	B	26.1	53.49	27.39	455	L	B	0.73	19.72	18.99
470	**N**	**B**	**36.38**	**71.72**	**35.34**	**456**	**A**	**B**	**0.6**	**55.37**	**54.77**
						457	N	B	12.98	34.53	21.55
						459	**L**	**B**	**0.5**	**41.03**	**40.53**
						460	**M**	**B**	**7.82**	**69.44**	**61.62**
						463	**S**	**B**	**2.26**	**59.27**	**57.01**
						464	**H**	**B**	**8.08**	**55.04**	**46.96**
						466	**R**	**B**	**17.67**	**51.12**	**33.45**
						467	H	B	32.52	53.29	20.77
						470	N	B	35.94	58.1	22.16
						471	K	B	20.65	36.12	15.47

the receptors. The regions of negative potential are mainly due to the negatively charged residues of helix 9, while the positive potential of the surface is mostly contributed by helix 11 and helix 7. Due to the anti-symmetric orientation of the dimer surface in the dimer, the regions of positive potential recognize the region of negative potential. The very similar surface electrostatics of both the ER receptors signifies a conserved mechanism of receptor dimerization and also rationalizes the formation of functional heterodimer.

3.6. "Hotspot" residue analysis

We have further analyzed the crucial "hotspot" residues in the dimerization interface of both ERββ homodimer and ERαβ heterodimer. Essential "hotspot residues" have been identified from the changes in solvent accessible surface area (SASA) of the monomer and dimerized ER. Upon dimerization, those residues that undergo a reduction in the SASA value > 15 $Å^2$ have been considered to be involved in the dimer interface. Table 1 lists all the crucial residues involved in the dimerization interface of the both homo- and heterodimers. Residues which undergo changes in SASA by more than 30 $Å^2$ is considered to be the "hotspot" residues. Helix 7 does not contain any hotspot residues for heterodimer but the ERββ only contains a single hotspot residue Ala 382. On the other hand, helix 8 contains three "hotspot" residues for the homodimer but there are two "hotspot" residues for the heterodimer. Residue Asn 407 in the homodimer and residue Asn 455 in the heterodimer play a similar but crucial role in the dimerization process. On the other hand, Met 410 and Pro 412 are crucial for homodimerization while Tyr 459 is essential for the heterodimerization. Most of the crucial residues are clubbed into the core recognition groove formed by helix 9 and helix 10/11. In helix 9, conserved Thr 434 for homodimer and 483 for heterodimer play crucial roles in the dimer interface. In the case of homodimer, Val 438 of helix 9 is also found to be important. Interestingly, in both homo- and heterodimer crucial "hotspot" residues are in general occupied same regions in the dimer interface. In the N-terminal end of helix 10, Ser 452 and Met 453 for homodimer and His 501 and Gln 502 for heterodimer belongs to the same region of the dimer surface and play crucial role in dimer interface. In helix 10, a highly conserved Ala 456 for homodimer and Ala 505 for heterodimer have been found to be "hotspot" residues for the dimer interface. In helix 11, a conserved Ser 463, His 467 for homodimer and His 512, His 516 for heterodimer belong to the same region of the dimer interface and are highly crucial for dimerization. On the other hand, Met 460 and Leu 509 belong to the same region of the helix 10/11 and they are highly crucial for the homo- and heterodimers, respectively. "Hotspot" residues from the other chain (the common ERβ in ERββ and ERαβ) are highly similar in both the homo- and heterodimers. A comparative analysis of the details of the interacting residues has been summarized in Table 1.

Fig. 6. A combined phylogenetic tree for ER alpha and beta across eukaryotic lineages.

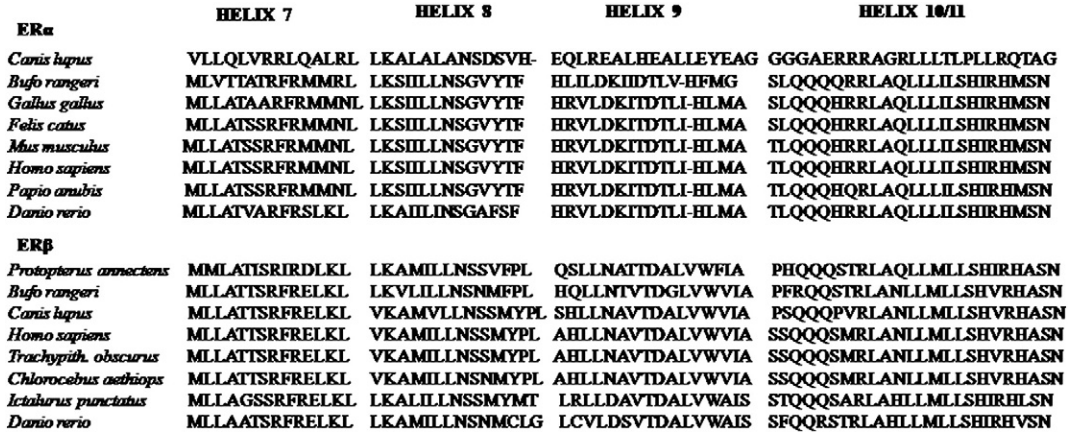


Fig. 7. A combined sequence analysis of ERα and ERβ essential dimer interface.

This conserved interaction profiles of receptor dimerization for both homo- and heterodimer prompted us to study the conservedness of the dimer surface in sequence level across different organisms.

3.7. Sequence and phylogenetic analysis of ER sequences

To infer specific variations in the ER LBD sequences for different species, multiple sequence alignments and phylogenetic analysis have been carried out. Fig. 6 represents a combined phylogenetic analysis of both ERα and ERβ. ER sequences and their phylogenetic distribution patterns found in eukaryotic lineages show an interesting feature. ERα and ERβ sequences are clubbed differently to form two separate clades. The two proteins ERα and ERβ in the same organism are paralogs, due to gene duplication. ERα sequences are orthologs, descendants of a common ancestor, due to speciation, and so are ERβ. Both paralogs and orthologs are homologs. It is to be noted that organisms that are closely related during speciation, their ERα sequences are also closely related and clubbed together within a clade while distantly related homologs are generally distributed in different clades. Sequences for both ERα and ERβ of *Homo sapiens* are highly similar to *Gorilla gorilla*, *Pan paniscus, Pongo pygmaeus*. On the other hand, *Gallus gallus* sequence shows significant similarity with *Coturnix japonica*, *Pseudemys nelson* for both ERα and ERβ proteins. Thus, they are clubbed together within a clade. In general, the observed phylogenetic pattern for ERα is very similar with the ERβ which signifies a common pattern of evolution. It is to be interpreted by the fact that the formation of functional heterodimer can strive the evolutionary process of the both the receptors in a similar way.

ERα and ERβ LBD sequences are moderately conserved throughout the eukaryotic lineages and the variations are observed mainly in the loop regions, devoid of any secondary structural elements. This can be rationalised by the fact that the loop regions are highly flexible and, thus, are capable of adapting to variations. Particularly, sequence variations are observed in several loop regions. These regions are the long N-terminal loop region between helix 1 and helix 2, the connecting loop between helix 8 and helix 9, and the connecting turn between helix 9 and helix 10. Interestingly, residues involved in the dimer surface of the LBD are highly conserved. Even, a combined sequence alignment of both the ERα and ERβ sequences involved in the dimer interface from different species shows that the dimer interface sequences are highly conserved (Fig. 7). "Hotspot residues" identified by in silico analysis are highly conserved and only substitutions to similar amino acids were observed. This signifies that ER LBD dimer interface is less susceptible to mutations and the formation of a functional dimer is an evolutionary selected process. Furthermore, a similar co-evolution pattern observed from phylogenetic analysis and the highly similar dimer contacts imply that the formation of a stable heterodimer is also an evolutionary selected process.

4. Conclusions

Transcriptional activity of estrogen receptors depends on dimerization [14]. It is well documented that both ERα and ERβ form homo- and heterodimers in solution and in vivo [14]. ERα and ERβ are co-expressed in the same cells in mouse mammary gland, in rat cardiac myocytes, and fibroblasts suggest that the heterodimerization probably occurs in vivo [13,17]. Even it has been depicted that ERαβ heterodimers bind to ERE sequences with a specificity and affinity similar to those of the respective homodimers. We observed that in the presence of agonist ligand, ER forms highly stable heterodimer which is more stable than the ERββ homodimers. ERα and ERβ have almost identical LBD structural topology and share highly similar electrostatic potential profile for the dimer interface. Combined sequence analysis reveals that the dimer interface of the LBD of both the ER is highly conserved. Interestingly, phylogenetic analysis reveals a co-evolutionary signature between the receptors. This can be rationalized by the fact that the formation of functional heterodimers is evolutionary selected process which drives the two proteins to evolve similarly. On therapeutic perspectives, structural and mechanistic insights into the effect of ligand binding on ERββ homodimer and ERαβ heterodimer are highly essential to design therapeutics. Considering the fact that ERβ forms heterodimers with ERα and the heterodimer complex inhibits cell proliferation, promoting heterodimerization with ERβ agonist would be an important strategy to design therapeutics for breast cancer.

Acknowledgement

The authors acknowledge financial support from MS-INBRE (USM-GR04015-05-9; NIH/NCRR 5P20RR016476-11) and NIH/NIGMS/8P20GM103476-11 and EPSCoR (EPS-0903787; Sub-contract: 190200-362492-10).

References

[1] N. Heldring, A. Pike, S. Andersson, J. Matthews, G. Cheng, J. Hartman, M. Tujague, A. Ström, E. Treuter, M. Warner, J.A. Gustafsson, Estrogen receptors: how do they signal and what are their targets, Physiological Reviews 87 (2007) 905–931.
[2] B.J. Deroo, K.S. Korach, Estrogen receptors and human disease, The Journal of Clinical Investigation 116 (2006) 561–570.
[3] J. Arts, G.G.J.M. Kuiper, J.M.M.F. Janssen, J.A. Gustafsson, C.W.G.M. Lowik, H.A.P. Pols, J.P.T.M.V. Leeuwen, Differential expression of Estrogen receptors α and β mRNA during differentiation of human osteoblast SV-HFO cells, Endocrinology 138 (1997) 5067–5070.
[4] E. Enmark, M. Pelto-Huikko, K. Grandien, S. Lagercrantz, J. Lagercrantz, G. Fried, M. Nordenskjold, J.A. Gustafsson, Human estrogen receptor β—gene structure, chromosomal localization, expression pattern, Journal of Clinical Endocrinology and Metabolism 82 (1997) 4258–4265.
[5] G.G.J.M. Kuiper, E. Enmark, M. Pelto-Huikko, S. Nilsson, J-Ar. Gustafsson, Cloning of a novel estrogen receptor expressed in rat prostate and ovary, Proceedings of

the National Academy of Sciences of the United States of America 93 (1996) 5925–5930.
[6] G.G.J.M. Kuiper, B. Carlsson, K. Grandien, E. Enmark, J. Haggblad, S. Nilsson, J-Ar Gustafsson, Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors α and β, Endocrinology 3 (1997) 863–870.
[7] G.G.J.M. Kuiper, P.J. Shughrue, M. Pelto-Huikko, I. Merchenthaler, J-Ar. Gustafsson, The estrogen receptor β subtype: a novel mediator of estrogen action in neuroendocrine systems, Frontiers in Neuroendocrinology 19 (1998) 253–286.
[8] Y. Feng, D. Manka, K.U. Wagner, S.A. Khan, Estrogen receptor-alpha expression in the mammary epithelium is required for ductal and alveolar morphogenesis in mice, Proceedings of the National Academy of Sciences of the United States of America 104 (2007) 14718–14723.
[9] J-Ar. Gustafsson, Estrogen receptor β—a new dimension in estrogen mechanism of action, The Journal of Endocrinology 163 (1999) 379–383.
[10] E.C. Chang, J. Frasor, B. Komm, B.S. Katzenellenbogen, Impact of estrogen receptor beta on gene networks regulated by estrogen receptor alpha in breast cancer cells, Endocrinology 147 (2006) 4831–4842.
[11] L.A. Helguero, M.H. Faulds, J.A. Gustafsson, L.A. Haldosen, Estrogen receptors alpha (ERalpha) and beta (ERbeta) differentially regulate proliferation and apoptosis of the normal murine mammary epithelial cell line HC11, Oncogene 24 (2005) 6605–6616.
[12] K. Pettersson, F. Delaunay, J.A. Gustafsson, Estrogen receptor beta acts as a dominant regulator of estrogen signaling, Oncogene 19 (2000) 4970–4978.
[13] E. Powell, E. Shanle, A. Brinkman, J. Li, S. Keles, K.B. Wisinski, W. Huang, W. Xuet, Identification of estrogen receptor dimer selective ligands reveals growth-inhibitory effects on cells that co-express ERa and ERb, PLoS One 7 (2012) e30993, http://dx.doi.org/10.1371/journal.pone.0030993.
[14] K. Pettersson, J.K. Gustafsson, Role of estrogen receptor beta in estrogen action, Annual Review of Physiology 63 (2001) 165–192.
[15] S. Saji, M. Hirose, M. Toi, Clinical significance of estrogen receptor beta in breast cancer, Cancer Chemotherapy and Pharmacology 56 (2005) 21–26.
[16] M.K. Lindberg, S. Moverare, S. Skrtic, H. Gao, K. Dahlman-Wright, J.A. Gustafsson, C. Ohlsson, Estrogen receptor (ER)-beta reduces ERalpha-regulated gene transcription, supporting a "ying yang" relationship between ERalpha and ERbeta in mice, Molecular Endocrinology 17 (2003) 203–208.
[17] K. Pettersson, K. Grandien, G.G. Kuiper, J.A. Gustafsson, Mouse estrogen receptor beta forms estrogen response element-binding heterodimers with estrogen receptor alpha, Molecular Endocrinology 11 (1997) 1486–1496.
[18] G.B. Tremblay, A. Tremblay, F. Labrie, V. Giguere, Dominant activity of activation function 1 (AF-1) and differential stoichiometric requirements for AF-1 and −2 in the estrogen receptor alpha-beta heterodimeric complex, Molecular and Cellular Biology 19 (1999) 1919–1927.
[19] S. Chakraborty, S. Cole, N. Rader, C. King, R.V. Rajnarayanan, P.K. Biswas, In silico design of peptidic inhibitors targeting Estrogen Receptor alpha dimer interface, Molecular Diversity (2012), http://dx.doi.org/10.1007/s11030-012-9378-x.
[20] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, Journal of Molecular Biology 234 (1993) 779–815.
[21] H.J.C. Berendsen, D.V.D. Spoel, R.V. Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Computer Physics Communications 91 (1995) 45–56.
[22] E. Lindahl, B. Hess, D.V.D. Spoel, GROMACS 3.0 a package for molecular simulation and trajectory analysis, Journal of Molecular Modelling 7 (2001) 306–317.
[23] W.L. Jorgensen, T. Rives, Development and testing of the OPLS all-atom force field on conformational energetic and properties of organic liquids, Journal of the American Chemical Society 110 (1988) 1657–1666.
[24] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, Journal of Computational Chemistry 18 (1997) 1463–1472.
[25] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[26] N. Tuncbag, A. Gursoy, O. Keskin, Identification of computational hotspots in protein interfaces: combining solvent accessibility and inter-residue potential improves the accuracy, Bioinformatics 25 (2009) 1513–1520.
[27] W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics 22 (2006) 1658–1659.
[28] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Research 32 (2004) 1792–1797.
[29] S. Guindon, J.F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0, Systems Biology 59 (2010) 307–321.
[30] F. Chevenet, C. Brun, A.L. Banuls, B. Jacq, R. Chisten, TreeDyn: towards dynamic graphics and annotations for analyses of trees, BMC Bioinformatics 7 (2006) 439–447.
[31] S. Chakraborty, P.K. Biswas, Ligand selective conformational changes of Estrogen Receptor alpha homo-dimer: a new insight into agonist action of tamoxifen (2012) (in review).