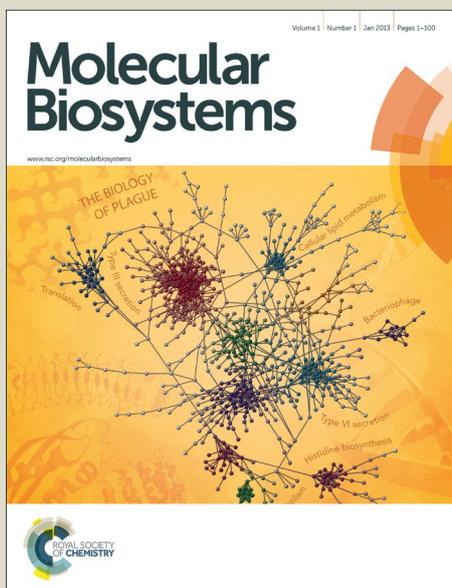


# Molecular BioSystems

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



[www.rsc.org/molecularbiosystems](http://www.rsc.org/molecularbiosystems)

**Structural insights into tumor-specific chaperoning activity of Gamma Synuclein in protecting Estrogen receptor alpha 36 and its role in Tamoxifen resistance in breast cancer**

**Manivel Panneerselvam, Kannan Muthu and Krishna Ramadas\***

Centre for Bioinformatics, School of Life sciences, Pondicherry University, Kalapet, Puducherry-605014, India.

\*E-mail: krishstrucbio@gmail.com

**Abstract**

Gamma synuclein ( $\gamma$ Syn), a tumor-specific molecular chaperone protects Hsp90 client proteins like ER $\alpha$ 36 and stimulates rapid membrane-initiated estrogen signalling in breast cancer cells. However, the structural perspectives of this tumor-specific chaperone function of  $\gamma$ Syn remains unclear. Hence, in this present work, we studied the conformational dynamics of ER $\alpha$ 36 in absence and presence of Hsp90 and  $\gamma$ Syn. Results indicate that in a chaperone-free state, ER $\alpha$ 36 undergoes an inter-domain movement and exposes the hydrophobic patch of residues that are responsible for binding with ubiquitin. However, independent of Hsp90,  $\gamma$ Syn by establishing transient interactions, prevent interdomain movement, unveil co-activator binding groove, masks the ubiquitin-binding residues and maintain 'open' pocket conformation of LBD. By doing so,  $\gamma$ Syn effectively protect ER $\alpha$ 36 from degradation and maintains its functional state as like Hsp90 based chaperoning machinery but independent of ATP. Our studies also shows that the  $\gamma$ Syn protected conformation of ER $\alpha$ 36 can effectively bind with both Estradiol (E2) and 4-hydroxy tamoxifen (4-OHT). Although they exhibit unique binding mode, they maintained the functionally active conformation of ER $\alpha$ 36. Interestingly, the molecular dynamics simulation studies showed that 4-OHT as like  $\gamma$ Syn, prevented the interdomain movements, primes the co-activator binding groove of ER $\alpha$ 36 for complexation with downstream signalling proteins and this mechanism explains its agonist activity and associated anti-estrogen resistance observed in presence of ER $\alpha$ 36. The observed differences in the chaperoning mechanism of  $\gamma$ Syn sheds light on its selectivity over Hsp90 in cancer cells, for promoting rapid protection of crucial oncogenic proteins. Based on our findings, we speculate that the compounds, which can hamper association of  $\gamma$ Syn with ER $\alpha$ 36 and/or can arrest ER $\alpha$ 36 in an ubiquitin binding state, would be promising alternatives for treating ER $\alpha$ 36 expressed breast carcinomas.

**Keywords**

Chaperone function of intrinsically disordered protein; Gamma Synuclein; Synuclein family; Tumor-specific variant of Estrogen receptor- $\alpha$ ; ER $\alpha$ -36; Anti-estrogen signaling; Anti-estrogen resistance; Tamoxifen resistance; Breast cancer.

## 1. Introduction

Heat-shock-protein 90 [Hsp90] class of molecular chaperones are essential to maintain structure and function of various proteins, which are otherwise degraded and proteins such as Estrogen receptors (ER), Androgen receptors (AR), progesterone receptors (PR), human epidermal growth factor receptor 2 (HER-2) and AKT are some of its client proteins. Expression of both Hsp90 and its client proteins are upregulated in several cancers including breast cancer, where the client proteins are protected from degradation by Hsp90, thereby promoting unhindered growth signaling in tumor cells<sup>1</sup>. Abnormal expression of ERs and their splice variants are critical determinants for propagation of estrogen dependent breast cancers<sup>2-4</sup>. Recently, the aberrant expression of a 36kDa splice variant of ER $\alpha$ , ER $\alpha$ 36, is observed to ensure survival and metastasis in both ER $\alpha$ 66 positive and in triple negative (ER $\alpha$ 66 -/HER-2 -/PR-) breast cancer cells and thus identified as the tumor specific splice variant of ER $\alpha$ <sup>5-11</sup>. Unlike its homologues, ER $\alpha$ 36 lacks both transactivation domains (AF-1 and AF-2) of ER $\alpha$ 66, but retains DNA binding domain; hinge region, partial ligand-binding domain and also has a unique stretch of 27 amino acids in its C-terminal region. ER $\alpha$ 36, by residing in cytoplasm and associating with membranes, promotes rapid membrane-initiated estrogen signaling [MIES]<sup>12, 13</sup> wherein it activates MAPK/ERK and PI3K/AKT signaling pathways in cancer cells, hampers apoptosis and promotes their growth and metastasis<sup>14, 15</sup>. ER $\alpha$ 36 is unique in its ability to remain uninhibited even after binding with Tamoxifen, a selective estrogen receptor modulator (SERM), which functions as estradiol and results in activation of the growth stimulating pathways. As a result, it renders cancer cells resistant against anti-tumor activities of Tamoxifen and its derivatives like 4-hydroxy tamoxifen [4-OHT], which are conventionally used to treat ER positive breast cancers<sup>12, 14, 16, 17</sup>. ER $\alpha$ 36, thus identified as a novel tumor specific variant of ER, favors survival and division of both estrogen dependent and independent breast cancers in presence of both Estradiol [E2] and SERMs. However, as like other ER variants, ER $\alpha$ 36 is protected by Hsp90-based chaperone machinery, suggesting that Hsp90 based inhibitors could be used in treating ER $\alpha$ 36 expressing cancer cells. But, an intrinsically disordered protein called Gamma Synuclein [ $\gamma$ Syn] thwarts this strategy by preventing degradation of ER $\alpha$ 36 in spite of the existence of Hsp90 disruptors<sup>18</sup>.

Gamma Synuclein, a member of synuclein family, is abnormally expressed in several cancers particularly in late stages of breast cancers<sup>19-22</sup> and in the ovarian cancer. Its expression correlates with the poor survival of patients with breast cancer<sup>23,24</sup>. At the cellular level,  $\gamma$ Syn by interacting with BubR1, a Mitotic check-point protein, and with microtubules, promotes genetic instability and resistance against microtubule inhibitors like Taxol<sup>25-27</sup>.  $\gamma$ Syn is also reported to exhibit chaperone-like activity and it replace Hsp90-based chaperone machinery under conditions of endoplasmic stress to protect ERs<sup>28, 29</sup>. It also acts independent of Hsp90, and protects several Hsp90 client proteins like Her-2, AKT from degradation<sup>30, 31</sup>. Recently, it is shown that expression of gamma synuclein is

upregulated under conditions of stress developed owing to the inhibition of Hsp90 and under such conditions,  $\gamma$ Syn functions as tumor specific molecular chaperone and protects degradation of ER $\alpha$ 36. Moreover, the chaperoning action of  $\gamma$ Syn is also associated with the promotion of Tamoxifen resistance in breast cancers<sup>18</sup>. As like in Breast cancer, expression of both  $\gamma$ Syn and ERs are elevated in ovarian and endometrial carcinomas and found to promote the growth of blood vessels angiogenic mechanism and they confer metastatic potential to tumor cells<sup>32, 33</sup>. These observations indicate that expression of both  $\gamma$ Syn and ER $\alpha$ 36 are tumor specific and by its chaperoning functions,  $\gamma$ Syn protects ER $\alpha$ 36 and helps unhindered growth and division of tumor cells. Unlike the well-studied molecular chaperone Hsp90,  $\gamma$ Syn is not observed to possess any of the defined characteristics of a molecular chaperone, and so understanding the potential of this typically intrinsically disordered protein to function as a chaperone becomes pivotal for designing strategies to inhibit its oncogenic functions.

Hence, the present study examines the structural and functional impacts of complex formation between  $\gamma$ Syn and ER $\alpha$ 36 using computational methods. By comparing the conformational dynamics of  $\gamma$ Syn bound ER $\alpha$ 36 with that of its chaperone-free state, as well as in the presence of standard chaperone Hsp90, we have proposed possible structural depiction of how  $\gamma$ Syn can protect and promote ER $\alpha$ 36 mediated MIES signalling in breast cancer cells. Moreover, this study also addresses the mechanism behind the binding and activation of ER $\alpha$ 36 by 4-OHT, the derivative of tamoxifen and explains how 4-OHT acts as an agonist mimicking E2. This study puts forward the possibility of alternate strategies to surpass existing SERMS and for better curtailment of ER $\alpha$ 36 and  $\gamma$ Syn involved carcinomas.

## 2. Materials and Methods

### 2.1 Three dimensional structure prediction of ER $\alpha$ 36

Estrogen receptor alpha 36, ER $\alpha$ 36 (312 a.a) is an isoform of human estrogen receptor with DNA binding domain, hinge region and partial ligand-binding domain and a unique C terminal tail region that shares less similarity with its homologues<sup>34</sup>. DBD [6-75] and of ER $\alpha$ 36 shares 100% similarity with the crystal structures of ER $\alpha$ -66 DBD [PDB ID: 1HCQ]<sup>35</sup> and partial LBD [136-312] displays 90% similarity with ER $\alpha$ 66 LBD [PDB ID: 3ERT]<sup>36</sup> respectively which were used as templates for predicting the domains of ER $\alpha$ 36. The structure for hinge region of ER $\alpha$ 36 lacking structural template was predicted using Robetta *ab initio* structure prediction server<sup>37</sup>. Later, the complete structure of ER $\alpha$ 36 was predicted using Modeller9v10<sup>38</sup>. Ten models were generated and model with lowest DOPE score (-28,405.20)<sup>39</sup> was selected as best model and typed with CHARMM force field and minimized using SMART Minimizer protocol in Discovery Studio suite v3(Accelrys Software Inc., Discovery Studio Modeling Environment, Release 3.1.0.11157, San Diego: Accelrys Software Inc.,

2005–2011). Finally, the structure with acceptable backbone phi and psi angles as predicted by Ramachandran plot was selected for further analysis. Reliability of the predicted model was evaluated using the WHAT IF and Molprobrity servers. The WHAT IF server runs several analyses on the predicted model, including the checks for Bond lengths and angles, puckering parameters, side-chain, main chain packing and stereo chemical quality checks. Quality of the model was further ascertained by the Molprobrity score obtained from Molprobrity server<sup>40</sup>. Molprobrity score is a log-weighted combination of clashscore, percentage Ramachandran not favored and percentage bad side-chain rotamers, which is normalized to be on the scale of X-ray evaluation. The results of superimposition were discussed in angstrom units (Å) and the results obtained from the simulation are used in the nanometers (nm)

## 2.2 Molecular dynamics simulation studies on un-liganded chaperone free ER $\alpha$ 36

The conformational dynamics of un-liganded chaperone free ER $\alpha$ 36 was explored using GROMACSv4.6 molecular dynamics suite<sup>41</sup>. Initially, the structure was modified by adding the hydrogen atoms and the topology was built using the AMBER99SB-ILDN force field<sup>42</sup>. The protein was placed in a triclinic box where the edge of the box from the protein fixed with 1.5 nm in all the directions. Here, the SPC216 were used to solvate the system and the total charges were neutralized with addition of chlorine ions. The solvated system were energy minimized by steepest descent algorithm up to the maximum level (50,000 steps) in order to allow for the convergence tolerance of 1000 kJ mol<sup>-1</sup> nm<sup>-1</sup> following which, conjugate gradient algorithm was used with the same steps and convergence tolerance. Further, the system was allowed to equilibrate under NVT and NPT ensembles for a period of 1ns. In both equilibration steps, the V-rescale method was used for temperature coupling with a time constant of 0.1 and the reference temperature was set as 310 K for coupling of both protein and non-protein groups. In both these equilibration runs, the Fast Particle-Mesh Ewald electrostatics [PME] was used to treat long range electrostatics with PME order of 4 and a Fourier spacing of 0.16nm. LINCS method with order of 4 was used to apply constraints on all the bonds, grid method was used for neighbor searching with 5.0ps as the factor to update neighbor list and electrostatic cut off was set to 1.0nm with van der Waals cutoff set at to 1.4nm. The Parrinello–Rahman method was used for pressure coupling with 2.0ps as time constant and 1.0 bar as reference pressure. Finally, the production MD run was carried for of 30ns timescale. The structural transition and conformational orientation of chaperon free and chaperon associated form were explored using PCA and FEL analyses.

## 2.3 Principal components and essential dynamics analysis

Principal components analysis was used to characterize the collective motions in a simulation system and it is applied for discriminating the large-scale collective motions from the random thermal

fluctuations. These large-scale collective motions represent the critical biological motions that determine the functional state of a protein. High-amplitude concerted motions depicted by PCA are captured by Eigen vectors. Initially, the mass-weighted covariance matrix was generated after removing the whole molecule translational and rotational movements by superimposition and least-square fitting onto C $\alpha$  atoms. Here, the protein coordinates were extracted from the 10ps-30ps of well equilibrated trajectory were used to define the cosine content ( $c_i$ ) of the principal component ( $p_i$ ) of covariance matrix, which explains whether the time interval of MD simulation used to extract the sampling of trajectory is sufficient to represent the free energy landscape defined from PCA<sup>43,44</sup>. The covariance matrix was generated and diagonalized using the `g_covar` module and `g_anaeig` module of GROMACS was used to calculate the Eigen values and Eigen vectors. Obtained Eigen values represent the amplitude of motion along the vector and Eigen vectors represent the one single direction of collective motion. PCA in general used as the sensitive measure to understand the trajectory convergence and PCs with value <1.0 were identified to understand whether the trajectory has sampled for the converged free energy landscape<sup>45</sup>. The cosine content value of PCs is used as measure to differentiate the PCs that captured the nonrandom fluctuations form the collective motions. The cosine value of PCs varies between 0 (no cosine) and 1 (perfect cosine) in the total simulation time ( $T$ ):

$$c_i = \frac{2}{T} \left( \int \cos(i\pi t) P_i(t) dt \right)^2 \left( \int P_i^2(t) dt \right)^{-1}$$

Theoretically the first Eigen vector's cosine contribution is the most efficient one to define a protein's characteristic nature in terms of its structural transitions. More frequently the first Eigen vector is observed to have a cosine distribution closer to 1 which depicts the large scale motions in protein dynamics and hence cannot be used to interpret the protein behavior in terms of free energy landscape (FEL). Hence in our studies we have taken then PCs which has less cosine content and used them for further studies. The porcupine plot was then generated using the extreme projections of the selected PCs and the Pymol Modevector module [The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC] helps to visualize the direction of movement represented by each Eigen vectors. Further, the trajectory was split into two halves [10–15ns & 15–30ns] in case of native chaperone-free ER $\alpha$ 36, and Root Mean Square Inner Product, RMSIP was calculated to analyze convergence of subspace which is defined by first ten Eigen vectors.<sup>46</sup> RMSIP offers a confidential measure to understand the convergence of the subspace defined by the selected PCs. Finally, the Eigen vectors with low cosine value are utilized to generate the 2D representation of Free-energy landscape [FEL] and for the same the `g_sham` module of Gromacs was utilized<sup>47</sup>. The inter-domain movements represented by the Eigen vectors were analyzed using DynDom<sup>48</sup>.

#### 2.4. Protein-protein docking study between the ER $\alpha$ 36 with $\gamma$ Syn and Hsp90

The *in silico* predicted structure of  $\gamma$ Syn used in our earlier studies<sup>49</sup>, and the predicted structure of ER $\alpha$ 36, was used for performing HADDOCK<sup>50</sup> based protein-protein docking. The structure for human Hsp90 was predicted using yeast Hsp90 structure [PDB ID: 2CG9] as template and refined by SMART Minimizer in Discovery Studio Suite. The residues of  $\gamma$ Syn C-terminal tail region [90–127], charged linker region [H210–N301] of Hsp90, hinge region [76–132] and LBD [136–312] of ER $\alpha$ 36 were defined as active site residues. HADDOCK docking was carried in three different stages: Rigid body energy minimization step, followed by two cycles of rotational and translational rigid body minimization steps (1000 steps), and semi-flexible simulated annealing (SA) in torsion angle space followed by refinement in an explicit solvent layer (8Å for water). The obtained complexes were then clustered and complexes with least RMSD were selected for further analysis. Total binding energy, total energy and binding surface area were calculated using the default scripts of HADDOCK. PDB ePISA interface analysis ([http://www.ebi.ac.uk/pdbe/prot\\_int/pistart.html](http://www.ebi.ac.uk/pdbe/prot_int/pistart.html)) was used to understand the nature of interface formed by these complexes. The selected structures were then subjected to MD studies and essential component analysis using the same parameters defined in methodology employed for free form of ER $\alpha$ 36. The conserved hot spot residues were analyzed using the MDCONS package<sup>51</sup> and Robetta alanine scanning module<sup>52</sup>. The Gromacs *g\_mmpbsa* tool<sup>53</sup> with solvent probe radius of 1.4 was used for the calculation of binding free energy of  $\gamma$ Syn with ER $\alpha$ 36. Here, the complex coordinates of  $\gamma$ Syn- ER $\alpha$ 36 were extracted from the time period 25-30ns and used for calculating the binding free energy.

### 3. Results

#### 3.1. Structure of ER $\alpha$ 36

The predicted three-dimensional model of ER $\alpha$ 36 has a DNA-binding domain [DBD] (6–75), a hinge region (76–132) and a partial ligand-binding domain [LBD] (133–312). Structure of ER $\alpha$ 36 is predominantly constituted by alpha helices in which DBD is organized as two layered alpha beta sandwich fold, while LBD is folded as an orthogonal bundle [Fig. 1A]. This model has 97% of residues in favored region of Ramachandran plot with good average packing environment evidenced from a value of -1.333 (inferred from WHAT IF server) and has a Molprobrity score of 1.91 indicating that the model has been threaded correctly, and hence reliable for further studies.

**3.1.1. DNA Binding Domain:** The DBD of ER $\alpha$ 36 has two amphipathic helices ( $\alpha$ 1:E30–Q41;  $\alpha$ 2:R61–V75) and two antiparallel beta strands ( $\beta$ 1:S20–H23;  $\beta$ 2:V26–C29). The helices  $\alpha$ 1 and  $\alpha$ 2 form the characteristic “double zinc helix” structural motif in steroid hormone receptors which coordinates two zinc ions with the help of eight cysteine residues. The conserved residues E30, G31, K34, and A35 of the recognition helix  $\alpha$ 1, located near the first zinc finger module form the P-Box elements that are crucial for proper placement of the helix into the major groove of the DNA, helping

in DNA recognition and interaction with palindrome response elements. The residues P49-Q53 form a long exposed loop termed as D-box elements that is known to be involved in dimer formation with DBD of another ER therein mediating the head-to-head binding of ER with their ER responsive elements. These features indicate that DBD of ER $\alpha$ 36 shares similar structural features with ER $\alpha$ 66, and explains how ER $\alpha$ 36 can compete with ER $\alpha$ 66 for binding with DNA-binding elements in estrogen-responsive genes.

**3.1.2. Hinge region:** In ER $\alpha$ 36, the DBD region is connected to the C-terminal ligand-binding domain by a long hinge region, constituted by 56 amino acids (G76–S132), which is intrinsically disordered in nature. In the predicted model, the hinge region adopts a coil-like structure with two small helices:  $\alpha$ 3' (L92–Q97) and  $\alpha$ 4' (A110–R114). Residues G76–R86 and E102–K130 fold to form an irregular U-shaped structure, which positions DBD and LBD at a distance of 47.8Å. Hinge residues, M78–G81, placed C-terminal to the DBD are termed as C terminal extension [CTE]<sup>54</sup> and are crucial in the orientation of DBD and stabilizing its interaction with DNA. In ER $\alpha$ 36, residues G76–G81 offer a site for myristylation<sup>5</sup>, and residues K93, K95, I185, Q202, K129 and K130 can be pivotal for receptor ubiquitination and in binding with chaperones<sup>55,56</sup> Hinge region is also crucial for dictating the spatial orientations of N- and C-terminal regions of ERs in response to binding of Estradiol and other SERMs like Tamoxifen<sup>57</sup>.

**3.1.3. Ligand binding domain:** Ligand binding domain of ER $\alpha$ 36 shares good structural similarity with ER $\alpha$ 66 with an RMSD of 0.9Å. Ligand binding domain of ER $\alpha$ 66 and other nuclear receptors possess 12 helices [H1–H12] and a small beta hairpin [S1/S2] like structure. Of the 12 helices, the spatial orientation of helix H2 remains unsolved, owing to its disordered nature. In general, the helices (H1–H11) in ER $\alpha$ 66 are arranged in a three-layered antiparallel arrangement with a wedge-shaped scaffold, flanked by antiparallel sheet (S1/S2 hairpin) on one side and helix H12 on the other side. Helices H5/6, H9 and H10 form the core layer of this three-layered sandwich and are packed in between helices H1–H4 on one side and by H7, H8, and H11 on the other side<sup>36,58</sup>. However, in the predicted model of ER $\alpha$ 36, LBD lacks the helices H11–12, but its unique C-terminal region forms a long right-handed helix-like structure and is placed similar to helix H10 of other ERs. Here, the core helices H5 [199–209], H6 [211–223], H9 [269–283] and H10 [289–312] are sandwiched by H1 [139–149], H3 [169–190] and H4 [194–196] on one side and partially by helices H7 [239–244], H8 [250–265] on the other side. They are also flanked by S3/S4 hairpin [K228–D238], however due to the absence of H11 and H12; LBD of ER $\alpha$ 36 does not adopt a complete three-layered helical sandwich fold and hence failed to form the wedge-shaped molecular scaffold of standard LBDs [Fig. 1B].

ER $\alpha$ 36 has the conserved hormone-binding region which is formed by the residues of helices H3 [M170, E180 and L181], H6 [W210–R221], H8 [I251 and L255] and by the residues of the loop region connecting helix H7–H8 [M248]. Owing to the absence of helices H11 and H12 in ER $\alpha$ 36, the area of hormone binding site is observed to be 1954.54Å<sup>2</sup> as opposed to 2424.1Å<sup>2</sup> in ER $\alpha$ 66, [Fig. 1C]. The absence of helix H11 also leaves ER $\alpha$ 36 with a partial dimerization interface formed by H8–H10. In ER $\alpha$ 66, the co-activator recruitment groove is formed by residues of helix H3, H4, H5, H6 and H12 and is essential for E2 induced growth signaling in cells by promoting interaction with its cellular partners. This groove is predominantly formed by the residues of helix H12 whose spatial orientation acts as a crucial determinant for propagation of E2-induced growth signaling and exhibits two different conformational orientations in response to the chemical nature of the ligands that bind within the hormone-binding cavity. However, in ER $\alpha$ 36, owing to absence of helix H12, the residues of H3, H4, and H5 alone form the co-activator binding groove, making ER $\alpha$ 36 unresponsive to the type of ligand that binds within the hormone-binding site [Fig. 1D]. Observed conformational differences may offer a functional advantage to ER $\alpha$ 36 favoring its activation by both E2 and Tamoxifen. Due to these structural features, tumor cells would be tuned to express ER $\alpha$ 36 over other variants of ER to confer unhindered activation of Estrogen dependent growth signaling events, and to protect themselves from the inhibitory actions imposed by E2 mimetics such as Tamoxifen.

### 3.2. Conformational changes in chaperone-free ER $\alpha$ 36

Native un-liganded chaperone-free conformation of ER $\alpha$ 36 was subjected to a 30ns molecular dynamics simulation and the obtained RMSD plot affirms the stability attained by ER $\alpha$ 36 and it highlights high fluctuations in the hinge region [Supplementary Fig. 1A-B]. The Radius of gyration plot [Rg] shows a steady decrease from 3nm to 2.4nm implying that structure of ER $\alpha$ 36 undergoes continuous changes and the decrease in distance between DBD and LBD from 50Å to 35Å emphasizes that inter-domain movement might occur in chaperone-free state of ER $\alpha$ 36 [Supplementary Fig. 1C-D]. A Cartesian coordinate based Principal Component Analysis and Free Energy Landscaping was performed to identify and analyze the conformational dynamics of ER $\alpha$ 36. The first ten Eigen vectors obtained from PCA analysis captures 89% of the total motions indicating that these vectors define the essential subspace of the system which capture the dominant motions of the protein. The convergence of this subspace was assayed by calculating the Root Mean Square Inner Product [RMSIP] and a value of 0.66 indicates that the essential subspace spanned by these PCs has converged<sup>59</sup> [Supplementary Tables 1-S2].

In order to visualize the direction of movements captured by the Eigen vectors, porcupine plot was generated using the extreme projections of Principal Component, PC1. The obtained plot displays rotational movement for about 32.7° in helices H3–H6 of LBD and about 84.5° in entire DBD [Fig.

2A]. As a result, the orientation of helix  $\alpha_1$ ,  $\alpha_2$ , P, and D-box elements in DBD have changed, and simultaneously the co-activator groove in LBD changed its direction of projection towards the DBD [Fig. 2B]. Meanwhile, the porcupine plot generated from PC3 captures the movement of DBD towards LBD, and it highlights the presence of rotation in LBD which is similar to that observed in PC1 [Fig. 2C]. Later, the extreme structures of PC1 and PC3 were subjected to DynDom analysis to characterize the domain motions that occurred in the chaperone-unbound state of ER $\alpha$ 36 and the results support the PCA from PC3. These results further confirm the presence of interdomain movements and the hinge bending motions that are instrumental in imparting interdomain movements [Fig. 2D and Supplementary Fig. 2B]. In addition, the fluctuations of residues calculated along first three PCs indicates the presence of huge fluctuations in the DBD and hinge region and confirms that these regions are prominently changing their spatial orientations throughout the simulation period [Supplementary Fig. 1E]. Based on the direction of movement captured by these PCs, it becomes evident that in the chaperone free-state, ER $\alpha$ 36 exhibits rotation in its DBD, following which hinge region undergoes a bending motion causing the DBD to move towards LBD. Meanwhile, the LBD also rotates and as a result, its co-activator groove is projected towards the DBD blocking its role in subsequent signaling.

In coherence to our findings, similar observations were also made in case of ER $\alpha$ 66, where the hinge region is observed to impart changes in the spatial orientation of AF-1 and AF-2 domains in response to the binding of tamoxifen and Estradiol with hormone binding site of ER $\alpha$ 66<sup>57</sup>, indicating that hinge region is pivotal for conferring conformational changes in the structure of ERs in response to their biological requirements.

### 3.2.1. FEL generation and analysis of conformational changes in chaperone free ER $\alpha$ 36

Free energy landscape offers a valuable resource to understand different conformations sampled by the protein at its metastable state. The conformational sampling for ER $\alpha$ 36 was carried out using FEL plot was generated using PCs with a cosine content  $<1$ , as they are considered to capture the large-scale collective motions and not the random thermal fluctuations. In our studies, PC1 obtained for chaperone free ER $\alpha$ 36 trajectory was found to be with a cosine content of 0.76; however, it captures the motions that are in coherence with the experimental findings. Hence, we proposed that the motions represented along this PC are not random, and so we used PC1 along with PC2 and PC3, which had a cosine content of 0.06 and 0.02 respectively, for conformational sampling. The projection of PC1 onto PC2 [Supplementary Fig. 1F] and PC2 onto PC3 [Supplementary Fig. 1G] effectively enumerates the presence of different clusters, that corresponds to the presence of different conformational states for ER $\alpha$ 36.

The FEL plot was then generated using PC1 and PC2 to understand the conformational changes in ER $\alpha$ 36 during the course of the simulation. The 2D FEL plot shows the presence of five energy basins corresponding to different conformations of ER $\alpha$ 36, where the structures clustered at basin 1 represent the initial metastable state, those at basin 3 represent an intermediate state and structures in basin 5 represent the final metastable state [Fig. 3]. Comparison of native starting conformation with structures retrieved from basin 1 displays rotation of both DBD and LBD about 43.9° and 21.1° respectively, and comparison with structure retrieved from basin 5 highlights the presence of interdomain movement in ER $\alpha$ 36, which has caused the DBD and LBD to come closer by a distance of 33.8Å, which is 15.8Å less than that observed in the un-simulated structure of ER $\alpha$ 36. The results of DynDom ascertains this observation which identifies DBD as the moving domain, which rotated by about 146.8° and translated about -23.4Å towards LBD and the residues 81-120 of hinge region were identified to be responsible for hinge bending motion that effected this domain movement [Supplementary Fig. 2A-B]. Taken together the conformations sampled at different metastable states enumerates changes in the spatial orientation of the two domains in ER $\alpha$ 36 and the role of hinge residues in determining these spatial changes, which is similar to that observed in case of in ER $\alpha$ 66<sup>57</sup>.

### 3.2.2. Biological implications of predicted chaperone free form of ER $\alpha$ 36

From the results of MD simulation, PCA and FEL analysis, it is evident that in the chaperone-free state of ER $\alpha$ 36, both DBD and LBD undergoes rotation, followed by their displacement, the co-activator groove of LBD is projected towards DBD, masking it from participating in complex formation with its downstream cellular partners. Unlike its homologues, ER $\alpha$ 36 has a modified LBD, making it susceptible for activation by both agonist and antagonist. So the choice of agonist/antagonist binding of ER $\alpha$ 36 does not have any influence over the functional state of co-activator groove but serves only as a tool for receptor activation. Hence, upon activation, the co-activator binding groove of ER $\alpha$ 36 gains functional significance in deciding its oncogenic activities by participating in complex formation with its downstream signaling partners. In such a scenario, the observed changes in the conformation of LBD and co-activator groove makes it apparent that these changes represent the functionally arrested conformation of ER $\alpha$ 36, while it was usually discussed in respect to the conformation of hormone binding site in other isoforms of ERs.

In general, chaperone-free unliganded ERs in their functionally arrested form are prone to be degraded by poly-ubiquitination where the solvent exposed lysine residues in hinge region of ERs are recognized by proteins like CHIP<sup>56</sup>. The predicted model also exists in its functionally arrested state and , the presence of a solvent exposed hydrophobic patch of ubiquitin binding residues K91, K93, K129, K130 of hinge region, I185, and Q202 of LBD underscores this conformation to be susceptible for ubiquitin mediated degradation [Supplementary Fig. 2C]. These observations confirm that the

conformational dynamics of ER $\alpha$ 36 at its functionally arrested state can generate signals for its degradation as like other ERs, and this observation is quite first of its kind in depicting the structural features of chaperone-free state of estrogen receptors, which could be used as a forerunner for experimental validation.

### 3.3. Insights into the chaperoning mechanism of $\gamma$ Syn

With the understanding of ER $\alpha$ 36 conformations in its chaperone-free state, we then proceeded with *in silico* protein-protein interaction studies to understand how this protein is protected by  $\gamma$ Syn even in the absence of chaperone, Hsp90. Initially, the association between ER $\alpha$ 36 and Hsp90 was investigated to characterize the structural features involved in stabilizing ER $\alpha$ 36 with a chaperone. In general, Hsp90 based chaperone machinery is utilized by breast cancer cells to promote hormone mediated growth signals by rendering ERs in high-affinity ligand binding state, which is disrupted by Hsp90 inhibitors like Geldanamycin, Tanespimycin (17-AAG) and Radicicol<sup>60</sup>. However, Hsp90 targeted inhibition fails to induce degradation of ER $\alpha$ 36 as it is protected by the chaperone-like activity of  $\gamma$ Syn. Interestingly,  $\gamma$ Syn did not possess any of the defined characteristics of molecular chaperone and till now no experimental studies have been carried out to characterize the mechanism underlying its chaperone activity. Hence, to gain structural insights, we performed a HADDOCK based docking and MD dynamics study of ER $\alpha$ 36 with  $\gamma$ Syn and compared with that to ER $\alpha$ 36-Hsp90 complex to understand the conformational changes imposed on ER $\alpha$ 36 on binding with  $\gamma$ Syn.

#### 3.3.1. Association of Hsp90 with ER $\alpha$ 36

In the predicted ER $\alpha$ 36-Hsp90 complex, charged linker region (H210-N301) of Hsp90 interacts with the residues of helices H3, H5, H6, H8, H9 and H7-H8 loop of LBD in ER $\alpha$ 36 [Supplementary Fig. 3A]. The predicted complex has an interface area of 1270.5Å<sup>2</sup> with binding free energy of -79953.8kcal/mol [Table S3 and S4]. Although this charged linker region is intrinsically disordered in nature, it binds strongly with ER $\alpha$ 36 and its dynamics was studied using a 10ns MD simulation. The calculated RMSD and RMSF plots indicate the stability of Hsp90 and ER $\alpha$ 36 upon complexation. In comparison with chaperone-free state, LBD of ER $\alpha$ 36 displays less fluctuation than DBD and hinge region [Supplementary Fig. 3B, D-E]. The Radius of gyration decreases steadily and the distance between DBD and LBD of ER $\alpha$ 36 also reduces from 50Å to 28Å, affirming that in presence of Hsp90, ER $\alpha$ 36 undergoes distinctive conformational changes [Supplementary Fig. 3C and Fig. 4A-B]. Comparison of complexes retrieved at 0<sup>th</sup> and 10<sup>th</sup>ns shows an overall backbone RMSD of 17.6Å [Fig. 4C] and its comparison with chaperone-free ER $\alpha$ 36 displays difference in spatial orientation of DBD and LBD and displays movement of DBD resulting in decrease in their inter-domain space from 47.8Å to 27.8Å in presence of Hsp90 [Fig. 4D]. This is due to the hinge bending motion in ER $\alpha$ 36 initiated by the residues I82, R83, G88 and G108-L118 in hinge region as evidenced by DynDom

analysis. However, unlike the chaperone-free state, in the presence of Hsp90, the DBD does not mask the LBD co-activator groove, which leaves it exposed for favoring interactions with downstream kinases suggesting that Hsp90 induces a different folding pattern in ER $\alpha$ 36 [Fig. 4E].

Structural overlay of both these proteins in post-simulated complexes shows that upon complex formation, the charged linker region of Hsp90 is positioned over the hormone-binding pocket of LBD and the ATP binding domain of Hsp90 is pulled closer to LBD of ER $\alpha$ 36. This structural arrangement could be the initial conformation attained by ER $\alpha$ 36-Hsp90 complex required to maintain the architecture of hormone binding site in ER $\alpha$ 36 through its ATP-dependent chaperoning mechanism. Taken together, these results suggest how Hsp90 can protect hormone binding site of ER $\alpha$ 36, as well as protect the projection of DBD towards LBD, therein helping in maintaining the coactivator groove to participate in interactions with the downstream partners of ER $\alpha$ 36. Our results are first of its kind in substantiating that these features form the essential prerequisite for Hsp90 chaperone machinery and represent the possible chaperoning conformation of Hsp90 and ER $\alpha$ 36. Although the absence of other co-chaperones limits the understanding of complete picture of Hsp90 chaperoning mechanism, these results are significant in revealing the basic structural characteristics imposed by Hsp90 on ER $\alpha$ 36.

### 3.3.2. Interaction of $\gamma$ Syn with ER $\alpha$ 36

We further used the association of ER $\alpha$ 36 with Hsp90 and their complex dynamics to understand the role of  $\gamma$ Syn as a chaperone to ER $\alpha$ 36. The docked  $\gamma$ Syn-ER $\alpha$ 36 complex has a binding free energy of -43325.4 kcal/mol with an interface area of 1016.7Å<sup>2</sup>. In the predicted complex, the C-terminal tail of  $\gamma$ Syn establishes non-bonded interactions with residues of hinge and helices H3, H4, H5, H9-H10 loop, H10 of ER $\alpha$ 36 [Supplementary Table 3 and 4], but unlike Hsp90,  $\gamma$ Syn does not seal the hormone binding site of ER $\alpha$ 36 [Fig. 5]. Results of MD simulation indicates that the complex has an average MMPBSA binding free energy of -3020.64 kcal/mol, supported by conserved interactions between the residues E106, V107, S109, M113, W119, K129, K130, N131, L133, I185, D196, L197, T198, L199, H200, Q202, V203, R295 of ER $\alpha$ 36 and K60, N64, V95, R96, E98, Q107, E116, E117, V118, E120, E121, Q123, D127 of  $\gamma$ Syn. RMSD plot indicates that  $\gamma$ Syn induces vast structural variations in DBD and hinge region of ER $\alpha$ 36 [Supplementary Fig. 4A-B & 4D-E] as observed in Hsp90. However, the calculated Rg values and the obtained inter-domain distance in ER $\alpha$ 36 maintained within the range 45–48Å, indicates a striking difference in folding pattern of ER $\alpha$ 36 when compared with its chaperone-free and Hsp90-associated states [Supplementary Fig. 4C and 4F]. These results are further strengthened by PCA, where the porcupine plot generated using extreme projections of principal component, PC2 displays differences in the direction of movement in DBD of

ER $\alpha$ 36. In presence of  $\gamma$ Syn, the DBD and hinge region of ER $\alpha$ 36 are displaced towards the  $\gamma$ Syn C-terminal region and LBD of ER $\alpha$ 36 displays a rotation movement however, the co-activator groove remain projected towards  $\gamma$ Syn and not towards the DBD as seen in its chaperone-free state. [Fig. 6D]. These conformational changes were consistent in the structures retrieved from four different basin of FEL generated using the PC2 and PC3 and the results confirms the absence of interdomain movement in ER $\alpha$ 36 when associated with  $\gamma$ Syn [Supplementary Fig. 5].

Later, the structures representing the final metastable state of ER $\alpha$ 36- $\gamma$ Syn complex were compared with that of their native starting conformation which shows vast changes in the structure of  $\gamma$ Syn. The helical propensity of residues in the helices N [G14-E20; K23-V26] and C [56E-V66; V66-V71; N76-K81] of  $\gamma$ Syn is lost after complexation with ER $\alpha$ 36 and moreover, helix C of  $\gamma$ Syn is rotated which interchanges the orientation of N and C helices. Concomitantly, the inter-helical distance between N and C helices of  $\gamma$ Syn is also reduced from 17.0Å to 3.2Å and this has placed the C-helix closer to the DBD of ER $\alpha$ 36 with a distance of 7.7Å [Fig. 6B]. Owing to these changes, the interface area between ER $\alpha$ 36 and  $\gamma$ Syn was decreased from 1016.7Å<sup>2</sup> to 898.3Å<sup>2</sup> with a binding free energy of -3020.644 +/- 158.836 kJ/mol, which reinforces that the observed conformational changes in  $\gamma$ Syn is stable and has not altered its binding abilities with ER $\alpha$ 36. Moreover, intrinsically disordered proteins are more efficient in establishing transient interactions with their partners which is characterized by the presence of smaller interface area lesser than 1000Å<sup>2</sup>, and by the less hydrophobic nature of interfaces. The interface between  $\gamma$ Syn and ER $\alpha$ 36 also shows similar properties with less interface area and is enriched with charged residues indicating that  $\gamma$ Syn can establish transient interactions with its partners. Moreover, the ability of intrinsically disordered proteins to establish transient interactions favors them to switch between the bound and unbound states with their partners and this feature helps these proteins to induce unfolding and refolding of their substrates. Interestingly, similar structural changes were noticed in ER $\alpha$ 36- $\gamma$ Syn complex, indicating that the observed binding of  $\gamma$ Syn is benefitted out of transient interactions and can effectively help in folding ER $\alpha$ 36. Moreover, these transient interactions facilitate rapid binding and release of  $\gamma$ Syn, which as in other chaperones does not require the ATP for inducing proper folding of ER $\alpha$ 36<sup>61-65</sup>. This feature might explain the selectivity of  $\gamma$ Syn over Hsp90 in tumor cells for delivering rapid protection of oncogenic proteins.

Structural comparison of ER $\alpha$ 36- $\gamma$ Syn complex of pre- and post-simulated conformations shows that ER $\alpha$ 36 deviates with an RMSD of 26.15Å and exhibits huge differences in the orientation of its DBD. The DBD in post-simulated complex is rotated about 40.9° and displaced by 39.3Å towards  $\gamma$ Syn and is stabilized by the interactions established between residues T9, Y11, Y18, Y22, Y24, V26, W27 of DBD with the residues A55-T76 of  $\gamma$ Syn C-helix. Similarly, the residues G103, A110-A116 and S121 of hinge region also establish contacts with residues in the tail region of  $\gamma$ Syn. These interactions pegs

both DBD and the hinge region of ER $\alpha$ 36 with  $\gamma$ Syn, thereby distancing LBD by 56.3Å which is much larger than the chaperone free ER $\alpha$ 36 model (33.8Å), indicating that in presence of  $\gamma$ Syn, the movement of DBD towards LBD is entirely repressed [Fig. 6E]. In addition, in presence of  $\gamma$ Syn, the LBD of ER $\alpha$ 36 also exhibits rotation of 27.4° but unlike in its chaperone free and Hsp90 associated states, the co-activator groove is not projected towards DBD but remains in interaction with  $\gamma$ Syn. This observation emphasizes that the association of  $\gamma$ Syn has prevented the masking of co-activator groove by DBD and therein it maintained ER $\alpha$ 36 in a functionally active form [Fig. 6A]. In addition, distance between the helices H3, H8 of ER $\alpha$ 36 hormone-binding pocket is also not altered in presence of  $\gamma$ Syn, suggesting that  $\gamma$ Syn also protects the hormone binding pocket of ER $\alpha$ 36 [Fig. 6C]. The ubiquitin binding residues I185, Q202, K129 and K130 of ER $\alpha$ 36 also remain in interaction with the residues of  $\gamma$ Syn thwarting the onset of ER $\alpha$ 36 degradation. Taken together, these results indicate that by associating with hinge and DBD of ER $\alpha$ 36,  $\gamma$ Syn alters the folding pattern in ER $\alpha$ 36, which protects it from ubiquitin-mediated degradation, protects the topology of hormone binding pocket and co-activator binding groove of ER $\alpha$ 36 and hold ER $\alpha$ 36 poised for binding with E2/tamoxifen in breast cancer cells.

### 3.4. Interaction of Estradiol (E2)/4-hydroxytamoxifen (4-OHT) with ER $\alpha$ 36

The association of  $\gamma$ Syn with ER $\alpha$ 36 contributes to resistance against tamoxifen and its derivatives like 4-OHT by shielding functionally active ER $\alpha$ 36 from degradation. In general, binding of both E2 and 4-OHT with ER $\alpha$ 36 imparts similar kind of biological response in breast cancer cells by activating MIES and promotes resistance against SERMs like 4-OHT. These observations makes it interesting to examine how the  $\gamma$ Syn protected ER $\alpha$ 36 can bind both E2 and 4-OHT and promote similar kind of biological responses in tumor cells which is further fortified by the differences seen in the topology of LBD of ER $\alpha$ 36. Hence to elucidate the structural basis of binding of these molecules with  $\gamma$ Syn protected ER $\alpha$ 36; we performed a ligand based docking study of ER $\alpha$ 36 with E2 and 4-OHT.

Results show that both E2 and 4-OHT can bind effectively into the hormone-binding pocket [Supplementary Fig. 6A] with a GLIDE SP score of -6.177kcal/mol and -5.52kcal/mol, respectively which are considerably lesser than that obtained for their binding with ER $\alpha$ 66 [-10.62 kcal/mol for E2 and -10.150 kcal/mol for 4-OHT]. They also exhibit difference in their binding orientation in comparison to the conformation seen in E2 and 4-OHT bound ER $\alpha$ 66 complexes<sup>36, 58</sup>. In case of ER $\alpha$ 36, E2 is placed diagonally between H3 and H6, with the OH groups of rings A and D being docked near helix H6 and H3 and stabilized via a hydrogen bond with T174 and hydrophobic contacts with residues of helices H3, H6, H8 and H7–H8 loop region [Supplementary Fig. 6D]. But, in ER $\alpha$ 66,

E2 is placed between helices H11 and H3 with its ring A placed between H3 and H6, and ring D projected towards H11 [Supplementary Fig. 6B]. As for 4-OHT binding in ER $\alpha$ 36, ring A of 4-OHT is placed near the helix H8 with its OH group forming hydrogen bond with G247 of H7–H8 loop region, while ring B is placed between the helices H3 and H6. This is very much different from that observed in the crystal complex of ER $\alpha$ -66 with 4-OHT, where the rings are projected towards helix H11 favoring high affinity binding and antagonist activity against E2 [Supplementary Fig. 6C & 6E]. These results indicate that the absence of helix H11 in ER $\alpha$ 36, which is crucial in stabilizing interaction of both E2 and 4-OHT in ER $\alpha$ 66, can be instrumental in imposing differences in their binding orientation which is marked primarily by the absence of hydrogen bonds with key residues E180 and R221 in ER $\alpha$ 36. In addition, the absence of helix H12 can also facilitate ER $\alpha$ 36 to annul the antagonistic effects of 4-OHT turning it into an agonist.

In general, although the biological effects of E2 will remain same with all variants of ER; it is necessary to study how 4-OHT can exert agonist activity by stabilizing ER $\alpha$ 36 and promoting growth signaling cascades in tumor cells. Hence, we subjected ER $\alpha$ 36-4-OHT complex onto MD simulation studies for a time scale of 30ns. The obtained RMSD and RMSF plots ascertain that 4-OHT binding can effectively stabilize ER $\alpha$ 36 and the presence of fluctuations in hinge region is similar to the conformations obtained in chaperone-free/bound forms [Supplementary Fig. 7A-B]. Interestingly, the comparison of pre-simulated complexes with post-simulated structures indicates the absence of inter-domain movements in 4-OHT bound ER $\alpha$ 36 which is evident from the observed inter-domain distance of 49.3Å between its DBD and LBD domains maintained throughout the simulation period [Supplementary Fig. 7D] and is also evident from the obtained Rg plot [Supplementary Fig. 7C]. Analysis of the porcupine plot generated using the extreme projections of PC1 reveals movement of DBD away from LBD and rotation in LBD, but this rotation does not orient the co-activator groove towards DBD [Fig. 7A-B]. The analysis of the structures obtained from the FEL generated using PC2 and PC3 also shows the presence of similar changes in the structure of 4-OHT bound ER $\alpha$ 36 [Fig. 7C-D]. Interestingly, these results indicate the similarity between the conformational changes imposed on ER $\alpha$ 36 by  $\gamma$ Syn and 4-OHT, wherein both prevents inter-domain movements and maintains the co-activator groove exposed. Significantly, the  $\gamma$ Syn protects both hormone-binding pocket and co-activator groove, while 4-OHT after binding with this  $\gamma$ Syn-protected ER $\alpha$ 36, keeps the co-activator groove poised for interacting with its downstream kinases [Fig. 7B, E-F]. This observation significantly explains how 4-OHT can effectively bind with and stabilize ER $\alpha$ 36 and can potentially invoke rapid activation of MIES growth signaling.

#### 4. Discussion

In the present study, we have made an attempt to understand the structural changes imposed by  $\gamma$ Syn on ER $\alpha$ 36 and initially, we studied the dynamics of ER $\alpha$ 36 in a chaperone-free state. ER $\alpha$ 36 was originally identified as a tumor specific variant of ER $\alpha$ 66, differing by its unique 27 amino acids at C-terminal and lacking both the activation function domains (AF-1 & AF-2). It was shown to possess a DNA binding domain (DBD) followed by a hinge region that terminates with a partial ligand binding domain (LBD). Our predicted model shows that DBD has retained the necessary structural features of an ER DBD domain, and the hinge region in our ER $\alpha$ 36 predicted model is intrinsically-disordered and holds the conventional residues for myristylation and ubiquitination. As like in ER $\alpha$ 66, hinge region of ER $\alpha$ 36 dictates the spatial orientation of DBD and LBD domains, which is evident from our MD simulation studies. The predicted ER $\alpha$ 36 adopts a unique folding in its chaperone-free state, where its DBD and LBD were rotated and as a result, the co-activator groove in LBD has changed its projection towards DBD. This rotation was accompanied by the interdomain movement facilitated by bending motion of hinge residues, however, did not induced any interactions between DBD and LBD. These structural changes efficiently mask the co-activator groove from interacting with its cellular partners obstructing its downstream signaling. Moreover, this typical folding of ER $\alpha$ 36 exposes its residues that are crucial for interaction with E3-ubiquitin ligases like BRAC1 and CHIP, suggesting that the observed conformation can send signals for promoting degradation of ER $\alpha$ 36 in its native form. Thus we explain the structural elements underpinning the degradation of free ER $\alpha$ 36 observed in absence of chaperones.

However, ER $\alpha$ 36 exhibits different folding pattern based on the predicted models of ER $\alpha$ 36 on association with  $\gamma$ Syn and with Hsp90, and their MD simulation analysis. Initially, we found that ER $\alpha$ 36 exhibits interdomain movement in the presence of Hsp90 which upon binding with LBD induces rotation in both domains of ER $\alpha$ 36, though it does not change the projection of ER $\alpha$ 36 co-activator groove. The charged linker region of Hsp90 lids over the hormone-binding pocket in ER $\alpha$ 36 and the pulling of LBD towards its N-terminal domain may facilitate its ATP dependent chaperoning of ER $\alpha$ 36. Unlike Hsp90,  $\gamma$ Syn does not bind with LBD; rather it associates with hinge region by establishing transient interactions, blocks the interdomain movement, and retained both co-activator groove and hormone-binding pocket of LBD in an open state for facilitating their interactions with E2/E2 mimetics and downstream regulators. Also, by interacting with ubiquitin binding residues,  $\gamma$ Syn protects ER $\alpha$ 36 from degradation which explains how  $\gamma$ Syn, independent of Hsp90, can stabilize and protect ER $\alpha$ 36. Moreover, as noticed in other members of intrinsically disordered chaperone proteins,  $\gamma$ Syn can execute as a chaperone without the additional requirements for ATP, and it explains its selectivity over ATP based chaperoning mechanism of Hsp90 for delivering rapid protection to its oncogenic client proteins like ER $\alpha$ 36.

Also, the results of molecular docking and MD simulations affirm that both E2 and 4-OHT can effectively bind with a  $\gamma$ Syn-protected ER $\alpha$ 36. The binding mode of E2 and 4-OHT differed when compared to that of ER $\alpha$ 66, and the absence of helix H11 and H12 significantly imposed changes in their orientation. However, 4-OHT is able to stabilize ER $\alpha$ 36 by imparting changes similar to  $\gamma$ Syn, explaining how 4-OHT can act as an agonist in ER $\alpha$ 36-expressing tumor cells and its ability to keep the co-activator groove in open state unmasked by DBD, facilitating rapid activation of mitogenic estrogen signaling in breast cancer cells. These findings suggest that  $\gamma$ Syn, being an intrinsically disordered protein favors transient interactions, which facilitates rapid association and dissociation with ER $\alpha$ 36 therein stabilizing and maintaining hormone-binding pocket of ER $\alpha$ 36 in an open state, thus performing its role as a molecular chaperone at a faster rate over other conventional chaperones. The binding of E2/4-OHT also stabilizes ER $\alpha$ 36 and promotes interaction with downstream growth promoters in MIES signaling.

Our results thus provide the first line snapshots of structural changes that take place in ER $\alpha$ 36 in absence and presence of  $\gamma$ Syn and explain how, an unstructured  $\gamma$ Syn can effectively protect and stabilize the structure of ER $\alpha$ 36, independent of Hsp90 and promote rapid activation of MIES favoring kinase signaling in tumor cells. Observed efficiency of  $\gamma$ Syn over Hsp90 can explain its choice in tumor cells to get selectively up-regulated under conditions of endoplasmic stress induced by Hsp90 disruptors. Our findings also explain how E2 and 4-OHT can bind with and stabilize  $\gamma$ Syn-chaperoned ER $\alpha$ 36, and we have postulated the structural mechanism that underpins ER $\alpha$ 36 mediated anti-estrogen signaling in ER+ breast cancer and anti-estrogen resistance in ER- breast cancer fostered by agonist activity of tamoxifen [Fig. 8].

#### **Conflict of interest**

**The authors declare no conflicts of interest.**

**Author contributions:** MP carried out entire work and wrote the manuscript. KM helped in PCA and FEL analysis. Dr. RK guided the entire work. The authors declare no conflict of interest.

#### **Acknowledgements**

RK thanks CSIR, Government of India (37(1610)/13/EMR-II) and UGC Research award (No. F.30-1/2013(SA-II)/RA-2012-14-OB-TAM-1399) and KM thanks UGC Rajiv Gandhi Fellowship (No.F. 14-2(SC)/2009(SA-III)) for supporting research. Authors also thank DBT, Government of India, for supporting research works in Centre for Bioinformatics and IPLS program in Pondicherry University.

#### **References**

1. L. Whitesell and S. L. Lindquist, *Nature reviews. Cancer*, 2005, 5, 761-772.
2. C. J. Gruber, W. Tschugguel, C. Schneeberger and J. C. Huber, *The New England journal of medicine*, 2002, 346, 340-352.

3. S. Cheng and S. P. Balk, *Cancer treatment and research*, 2003, 115, 293-318.
4. S. E. Taylor, P. L. Martin-Hirsch and F. L. Martin, *Cancer Letters*, 2010, 288, 133-148.
5. Z. Wang, X. Zhang, P. Shen, B. W. Loggie, Y. Chang and T. F. Deuel, *Biochemical and biophysical research communications*, 2005, 336, 1023-1027.
6. S. Vranic, Z. Gatalica, H. Deng, S. Frkovic-Grazio, L. M. Lee, O. Gurjeva and Z. Y. Wang, *Journal of clinical pathology*, 2011, 64, 54-57.
7. X. T. Zhang, L. G. Kang, L. Ding, S. Vranic, Z. Gatalica and Z. Y. Wang, *Oncogene*, 2011, 30, 770-780.
8. J. Zhang, G. Li, Z. Li, X. Yu, Y. Zheng, K. Jin, H. Wang, Y. Gong, X. Sun, X. Teng, J. Cao and L. Teng, *Steroids*, 2012, 77, 666-673.
9. L. M. Lee, J. Cao, H. Deng, P. Chen, Z. Gatalica and Z. Y. Wang, *Anticancer research*, 2008, 28, 479-483.
10. W. Hong, X. Zhang, L. Ding, L. Kang and Z.-Y. Wang, *PloS one*, 2012, 7, e30174.
11. K. Wu, S. Huang, M. Zhu, Y. Lu, J. Chen, Y. Wang, Q. Lin, W. Shen, S. Zhang, J. Zhu, Y. E. Shi and Z. Weng, *Medical oncology*, 2013, 30, 612.
12. Z. Wang, X. Zhang, P. Shen, B. W. Loggie, Y. Chang and T. F. Deuel, *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103, 9063-9068.
13. L. Kang, X. Zhang, Y. Xie, Y. Tu, D. Wang, Z. Liu and Z. Y. Wang, *Molecular endocrinology*, 2010, 24, 709-721.
14. S. L. Lin, L. Y. Yan, X. T. Zhang, J. Yuan, M. Li, J. Qiao, Z. Y. Wang and Q. Y. Sun, *PloS one*, 2010, 5, e9013.
15. H. Deng, X. Huang, J. Fan, L. Wang, Q. Xia, X. Yang, Z. Wang and L. Liu, *Oncology reports*, 2010, 24, 171-176.
16. G. Li, J. Zhang, K. Jin, K. He, Y. Zheng, X. Xu, H. Wang, H. Wang, Z. Li, X. Yu, X. Teng, J. Cao and L. Teng, *Molecular oncology*, 2013, 7, 611-624.
17. Y. Gu, T. Chen, E. Lopez, W. Wu, X. Wang, J. Cao and L. Teng, *Journal of translational medicine*, 2014, 12, 16.
18. Y. E. Shi, Y. Chen, R. Dackour, L. Potters, S. Wang, Q. Ding, Z. Wang and Y. E. Liu, *The American journal of pathology*, 2010, 177, 964-973.
19. H. Ji, Y. E. Liu, T. Jia, M. Wang, J. Liu, G. Xiao, B. K. Joseph, C. Rosen and Y. E. Shi, *Cancer research*, 1997, 57, 759-764.
20. T. Jia, Y. E. Liu, J. Liu and Y. E. Shi, *Cancer research*, 1999, 59, 742-747.
21. W. Bruening, B. I. Giasson, A. J. Klein-Szanto, V. M. Lee, J. Q. Trojanowski and A. K. Godwin, *Cancer*, 2000, 88, 2154-2163.
22. K. Wu, Z. Weng, Q. Tao, G. Lin, X. Wu, H. Qian, Y. Zhang, X. Ding, Y. Jiang and Y. E. Shi, *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 2003, 12, 920-925.
23. J. Guo, C. Shou, L. Meng, B. Jiang, B. Dong, L. Yao, Y. Xie, J. Zhang, Y. Chen, D. R. Budman and Y. E. Shi, *International journal of cancer. Journal international du cancer*, 2007, 121, 1296-1305.
24. K. Wu, Z. Quan, Z. Weng, F. Li, Y. Zhang, X. Yao, Y. Chen, D. Budman, I. D. Goldberg and Y. E. Shi, *Breast cancer research and treatment*, 2007, 101, 259-267.
25. H. Zhang, A. Kouadio, D. Cartledge and A. K. Godwin, *Experimental cell research*, 2011, 317, 1330-1339.
26. M. Panneerselvam, K. Muthu, M. Jayaraman, U. Sridharan, P. Jenardhanan and K. Ramadas, *Molecular bioSystems*, 2013, 9, 1470-1488.
27. S. Miao, K. Wu, B. Zhang, Z. Weng, M. Zhu, Y. Lu, R. Krishna and Y. E. Shi, *Molecular cancer therapeutics*, 2014, 13, 699-713.
28. Y. E. Liu, W. Pu, Y. Jiang, D. Shi, R. Dackour and Y. E. Shi, *Oncogene*, 2007, 26, 2115-2125.
29. Y. Jiang, Y. E. Liu, I. D. Goldberg and Y. E. Shi, *Cancer research*, 2004, 64, 4539-4546.
30. Y. Shao, B. Wang, D. Shi, S. Miao, P. Manivel, R. Krishna, Y. Chen and Y. Eric Shi, *Molecular oncology*, 2014, 8, 1521-1531.

31. W. Liang, S. Miao, B. Zhang, S. He, C. Shou, P. Manivel, R. Krishna, Y. Chen and Y. E. Shi, *Oncogene*, 2014, 0.
32. M. N. Singh, H. F. Stringfellow, S. E. Taylor, K. M. Ashton, M. Ahmad, K. R. Abdo, O. M. A. El-Agnaf, P. L. Martin-Hirsch and F. L. Martin, *Molecular Human Reproduction*, 2008, 14, 655-663.
33. K. T. Cheung, S. E. Taylor, Patel, II, A. J. Bentley, H. F. Stringfellow, N. J. Fullwood, P. L. Martin-Hirsch and F. L. Martin, *British journal of medicine and medical research*, 2011, 1, 430-444.
34. J. Rao, X. Jiang, Y. Wang and B. Chen, *The Journal of steroid biochemistry and molecular biology*, 2011, 127, 231-237.
35. J. W. Schwabe, L. Chapman, J. T. Finch and D. Rhodes, *Cell*, 1993, 75, 567-578.
36. A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard and G. L. Greene, *Cell*, 1998, 95, 927-937.
37. D. E. Kim, D. Chivian and D. Baker, *Nucleic acids research*, 2004, 32, W526-531.
38. M. A. Marti-Renom, A. C. Stuart, A. Fiser, R. Sanchez, F. Melo and A. Sali, *Annual review of biophysics and biomolecular structure*, 2000, 29, 291-325.
39. M. Y. Shen and A. Sali, *Protein science : a publication of the Protein Society*, 2006, 15, 2507-2524.
40. V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, *Acta crystallographica. Section D, Biological crystallography*, 2010, 66, 12-21.
41. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, 29, 845-854.
42. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins*, 2010, 78, 1950-1958.
43. B. Hess, *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 2000, 62, 8438-8448.
44. G. G. Maisuradze and D. M. Leitner, *Proteins*, 2007, 67, 569-578.
45. B. Hess, *Physical Review E*, 2002, 65.
46. M. Schulein, *Biochimica et biophysica acta*, 2000, 1543, 239-252.
47. J. Sakon, D. Irwin, D. B. Wilson and P. A. Karplus, *Nature structural biology*, 1997, 4, 810-818.
48. S. Hayward and H. J. Berendsen, *Proteins*, 1998, 30, 144-154.
49. P. Manivel, J. Muthukumaran, M. Kannan and R. Krishna, *Journal of molecular modeling*, 2011, 17, 251-263.
50. C. Dominguez, R. Boelens and A. M. Bonvin, *Journal of the American Chemical Society*, 2003, 125, 1731-1737.
51. S. Abdel-Azeim, E. Chermak, A. Vangone, R. Oliva and L. Cavallo, *BMC bioinformatics*, 2014, 15 Suppl 5, S1.
52. T. Kortemme, D. E. Kim and D. Baker, *Science's STKE : signal transduction knowledge environment*, 2004, 2004, pl2.
53. R. Kumari, R. Kumar and A. Lynn, *Journal of Chemical Information and Modeling*, 2014, 54, 1951-1962.
54. M. M. Montano, V. Muller, A. Trobaugh and B. S. Katzenellenbogen, *Molecular endocrinology*, 1995, 9, 814-825.
55. Y. Ma, S. Fan, C. Hu, Q. Meng, S. A. Fuqua, R. G. Pestell, Y. A. Tomita and E. M. Rosen, *Molecular endocrinology*, 2010, 24, 76-90.
56. N. B. Berry, M. Fan and K. P. Nephew, *Molecular endocrinology*, 2008, 22, 1535-1551.
57. W. Zwart, R. de Leeuw, M. Rondaij, J. Neeffjes, M. A. Mancini and R. Michalides, *Journal of cell science*, 2010, 123, 1253-1261.
58. A. M. Brzozowski, A. C. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G. L. Greene, J. A. Gustafsson and M. Carlquist, *Nature*, 1997, 389, 753-758.
59. A. Amadei, M. A. Ceruso and A. Di Nola, *Proteins*, 1999, 36, 419-424.
60. L. Neckers and P. Workman, *Clinical cancer research : an official journal of the American Association for Cancer Research*, 2012, 18, 64-76.

61. R. van der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, P. Tompa, V. N. Uversky, P. E. Wright and M. M. Babu, *Chemical reviews*, 2014, 114, 6589-6631.
62. P. Tompa and P. Csermely, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 2004, 18, 1169-1175.
63. B. M. Gorovits and P. M. Horowitz, *The Journal of biological chemistry*, 1995, 270, 13057-13062.
64. R. A. Lindner, A. Kapur, M. Mariani, S. J. Titmuss and J. A. Carver, *European journal of biochemistry / FEBS*, 1998, 258, 170-183.
65. T. M. Treweek, A. Rekas, M. J. Walker and J. A. Carver, *Experimental eye research*, 2010, 91, 691-699.

## Figure Legends

**Fig. 1: Predicted Topology of ER $\alpha$ 36.** (A) The predicted structure of ER $\alpha$ 36, shown as cartoon, displays structural features of DNA binding domain (DBD), hinge and ligand binding domain (LBD). (B) Structural comparison of LBD in ER $\alpha$ 36 and tamoxifen bound conformation of ER $\alpha$ 66, displays differences in topology of ER $\alpha$ 36 (C) Displays features of hormone-binding pocket and, (D) Highlights the co-activator binding groove of ER $\alpha$ 36.

**Fig. 2: Domain movements observed in chaperone-free state of ER $\alpha$ 36.** (A) and (C). Porcupine plot depicts the direction of movements observed in the phase space along first and third principal components. Here, the direction of the C $\alpha$  atoms is denoted as arrows and their length corresponds to their Eigen values. (B) and (D) highlight the differences observed in the extreme projections of structures retrieved from PC1 and PC3 respectively.

**Fig. 3: Conformational Dynamics of chaperone-free state of ER $\alpha$ 36.** Displays the two dimensional FEL generated using the Cartesian coordinates of PC1 and PC2. The changes observed in the interdomain space between DBD and LBD were highlighted in the structures retrieved from each energy basin. These changes provide an overview of conformational changes seen in the structure of chaperone-free state of ER $\alpha$ 36.

**Fig. 4: Conformational changes imposed by Hsp90 in ER $\alpha$ 36.** (A) Displays the structure of ER $\alpha$ 36-Hsp90 complex obtained after the MD simulation studies, and highlights the association of Hsp90 linker region over the surface defined by the residues of helices H3, H4, H5 and H6 of ER $\alpha$ 36-LBD. (B) Displays the decrease in the distance between the DBD and LBD of ER $\alpha$ -36 seen in presence of Hsp90. (C) Comparison of structures retrieved at 0th and 10th ns shows the movement of DBD towards LBD and displays differences in orientation of LBD. (D) Structural comparison of native chaperone free and Hsp90 associated ER $\alpha$ 36 displays the difference in the folding pattern of ER $\alpha$ 36. (E) Shows the projection of co-activator groove that remains unmasked by ER $\alpha$ 36 -DBD in presence of Hsp90.

**Fig. 5: Intermolecular association of SNCG ( $\gamma$ Syn) with ER $\alpha$ 36.** The predicted association of SNCG with ER $\alpha$ 36 displays the binding of SNCG C-terminal tail region with hinge region of ER $\alpha$ 36 and displays the hydrogen bond interaction between their residues.

**Fig. 6: Snapshots of changes seen in the structure of ER $\alpha$ 36 in presence of SNCG.** (A) Comparison of structures retrieved from basin1 and basin 4 of FEL shows changes seen in the structure of ER $\alpha$ 36-DBD and LBD. (B) Displays structural changes seen in SNCG. (C) Depicts that in

presence of SNCG, hormone-binding pocket of ER $\alpha$ 36 is unaltered. (D) Displays the porcupine plot, here, the DBD of ER $\alpha$ 36 moves towards SNCG, while the co-activator groove does not show any movement owing to its interaction with SNCG. (E) Displays the conformation of ER $\alpha$ 36 seen in absence (grey) and in presence of SNCG (blue).

**Fig. 7: Snapshots of changes seen in 4-OHT bound ER $\alpha$ 36 complex.** (A) Comparison of extreme projections of PC2 depicting displacement of DBD and rotation of LBD in last frame (light blue) and here helix H3-H6 is denoted in red color and DBD is in blue color. (B) Porcupine plot generated for PC2. (C) FEL generated using PC2 and PC3. (D) Structure retrieved from basin 1 of FEL. (E and F) comparison of ER $\alpha$ 36 in 4-OHT bound ER $\alpha$ 36 (light blue) with the chaperone-free state of ER $\alpha$ 36 (grey) and shows the presence of interdomain space observed in the chaperone-free state.

**Fig. 8:** Structural insights into tumor specific chaperone functions of SNCG ( $\gamma$ Syn).

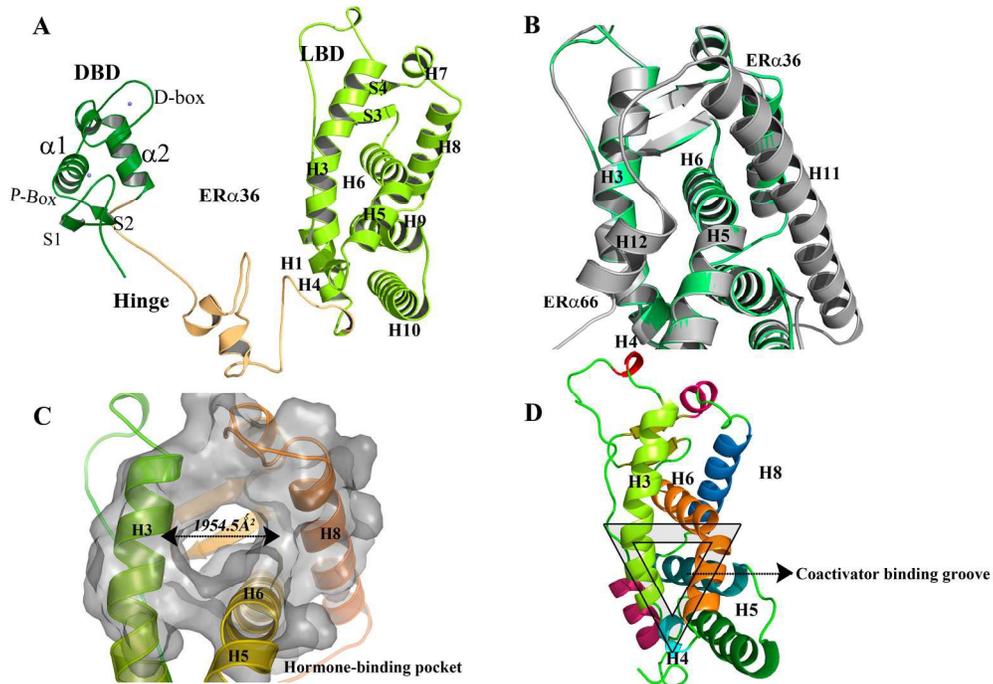


Fig. 1: Predicted Topology of ER $\alpha$ 36. (A) The predicted structure of ER $\alpha$ 36, shown as cartoon, displays structural features of DNA binding domain (DBD), hinge and ligand binding domain (LBD). (B) Structural comparison of LBD in ER $\alpha$ 36 and tamoxifen bound conformation of ER $\alpha$ 66, displays differences in topology of ER $\alpha$ 36 (C) Displays features of hormone-binding pocket and, (D) Highlights the co-activator binding groove of ER $\alpha$ 36.

190x129mm (300 x 300 DPI)

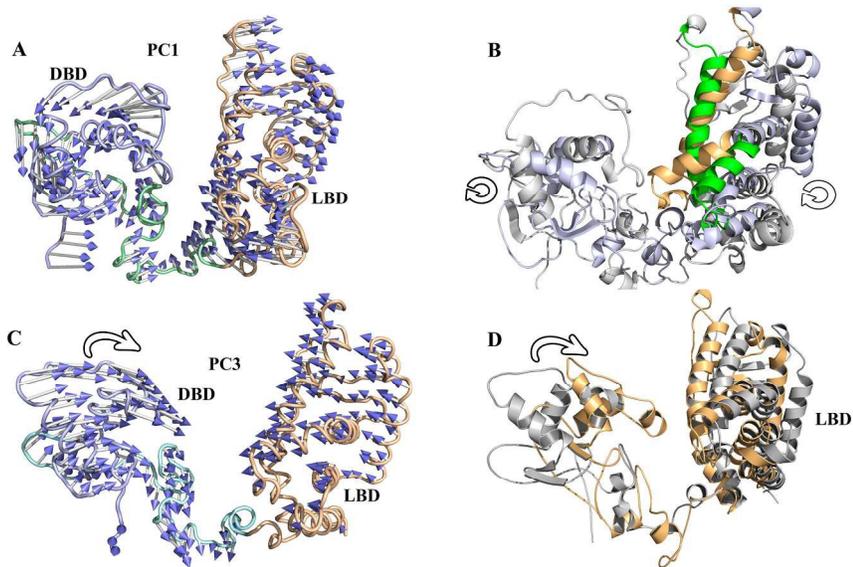


Fig. 2: Domain movements observed in chaperone-free state of ER $\alpha$ 36. (A) and (C). Porcupine plot depicts the direction of movements observed in the phase space along first and third principal components. Here, the direction of the Ca atoms is denoted as arrows and their length corresponds to their Eigen values. (B) and (D) Highlights the differences observed in the extreme projections of structures retrieved from PC1 and PC3 respectively.

235x134mm (300 x 300 DPI)

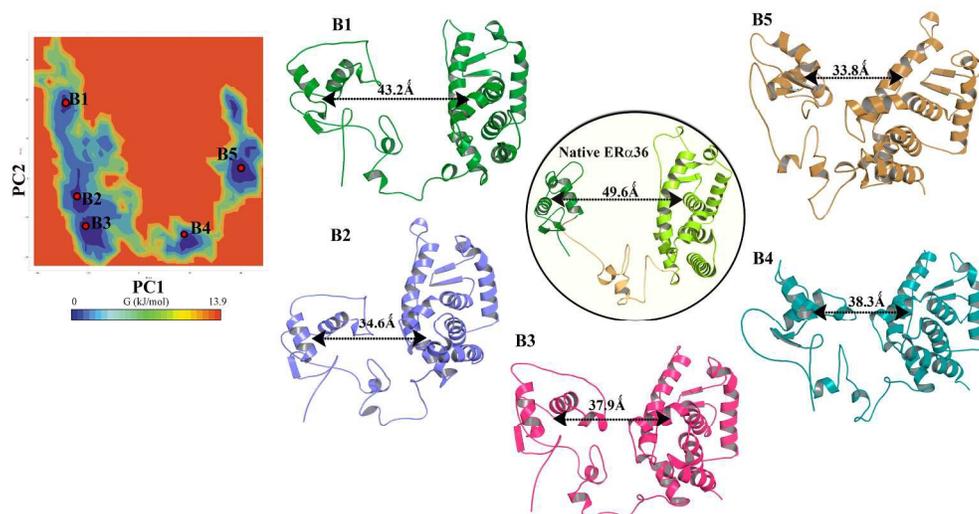


Fig. 3: Conformational Dynamics of chaperone-free state of ERα36. Displays the two dimensional FEL generated using the Cartesian coordinates of PC1 and PC2. The changes observed in the interdomain space between DBD and LBD were highlighted in the structures retrieved from each energy basin. These changes provide an overview of conformational changes seen in the structure of chaperone-free state of ERα36.  
235x124mm (300 x 300 DPI)

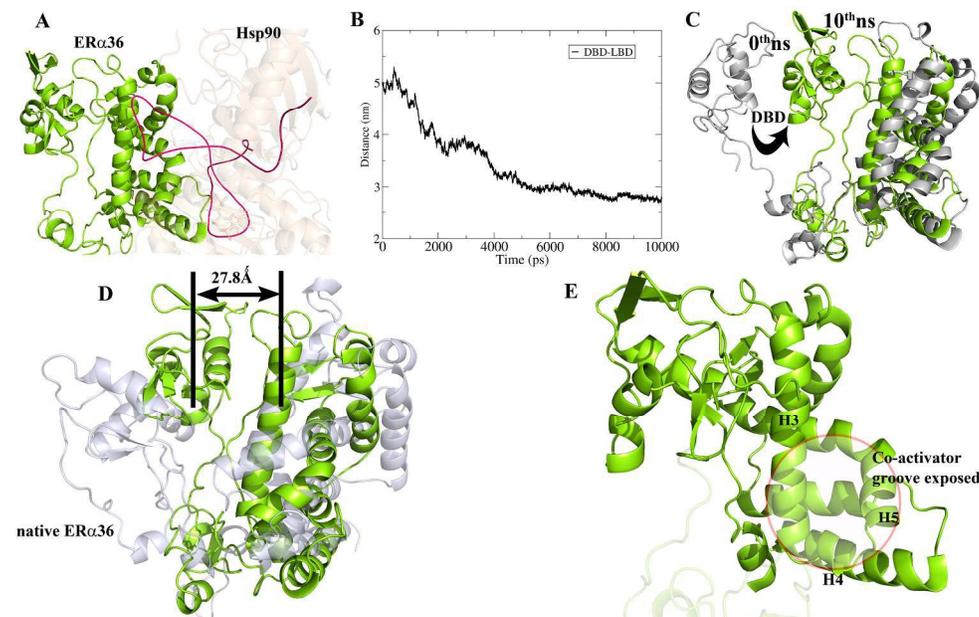


Fig. 4: Conformational changes imposed by Hsp90 in ER $\alpha$ 36. (A) Displays the structure of ER $\alpha$ 36-Hsp90 complex obtained after the MD simulation studies, and highlights the association of Hsp90 linker region over the surface defined by the residues of helices H3, H4, H5 and H6 of ER $\alpha$ 36-LBD. (B) Displays the decrease in the distance between the DBD and LBD of ER $\alpha$ -36 seen in presence of Hsp90. (C) Comparison of structures retrieved at 0th and 10th ns shows the movement of DBD towards LBD and displays differences in orientation of LBD. (D) Structural comparison of native chaperone free and Hsp90 associated ER $\alpha$ 36 displays the difference in the folding pattern of ER $\alpha$ 36. (E) Shows the projection of co-activator groove that remains unmasked by ER $\alpha$ 36 -DBD in presence of Hsp90.

295x182mm (300 x 300 DPI)

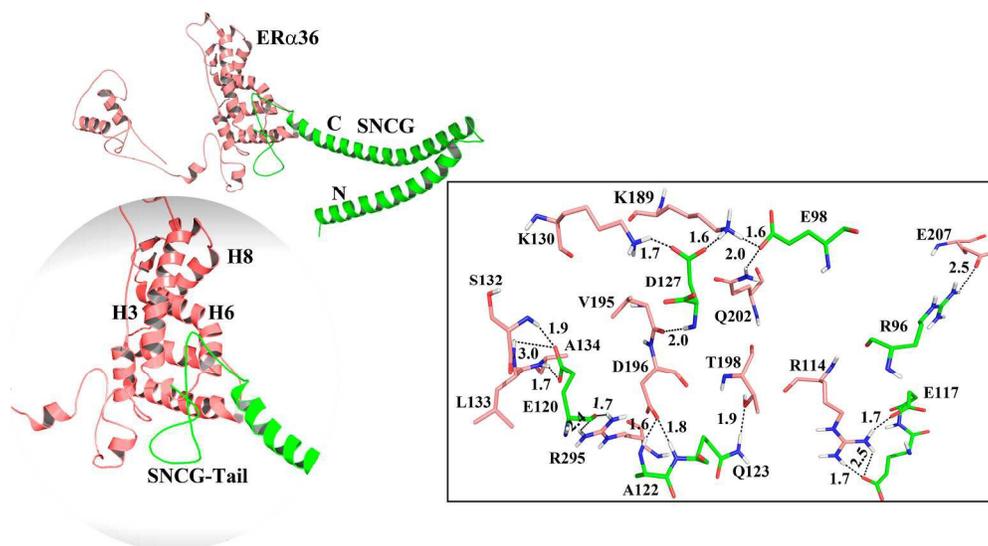


Fig. 5: Intermolecular association of SNCG ( $\gamma$ Syn) with ER $\alpha$ 36. The predicted association of SNCG with ER $\alpha$ 36 displays the binding of SNCG C-terminal tail region with hinge region of ER $\alpha$ 36 and displays the hydrogen bond interaction between their residues.  
235x130mm (300 x 300 DPI)

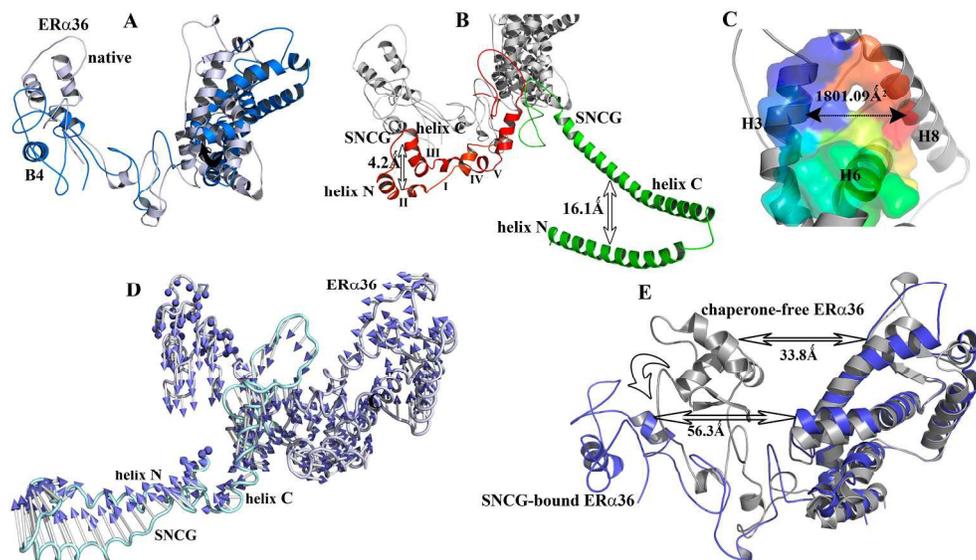


Fig. 6: Snapshots of changes seen in the structure of ER $\alpha$ 36 in presence of SNCG. (A) Comparison of structures retrieved from basin1 and basin 4 of FEL shows changes seen in the structure of ER $\alpha$ 36-DBD and LBD. (B) Displays structural changes seen in SNCG. (C) Depicts that in presence of SNCG, hormone-binding pocket of ER $\alpha$ 36 is unaltered. (D) Displays the porcupine plot, here, the DBD of ER $\alpha$ 36 moves towards SNCG, while the co-activator groove does not show any movement owing to its interaction with SNCG. (E) Displays the conformation of ER $\alpha$ 36 seen in absence (grey) and in presence of SNCG (blue).  
230x130mm (300 x 300 DPI)

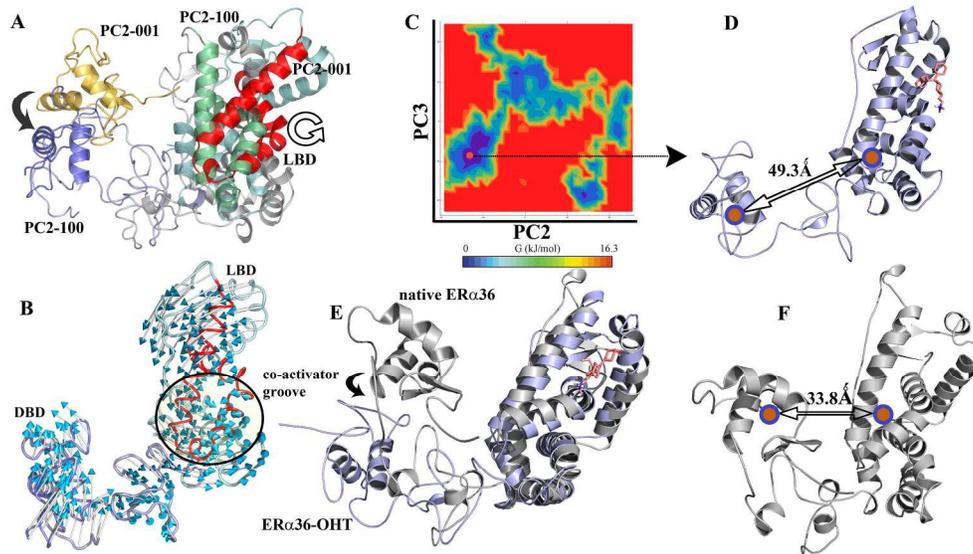


Fig. 7: Snapshots of changes seen in 4-OHT bound ERα36 complex. (A) Comparison of extreme projections of PC2 depicting displacement of DBD and rotation of LBD in last frame (light blue) and here helix H3-H6 is denoted in red color and DBD is in blue color. (B) Porcupine plot generated for PC2. (C) FEL generated using PC2 and PC3. (D) Structure retrieved from basin 1 of FEL. (E and F) comparison of ERα36 in 4-OHT bound ERα36 (light blue) with the chaperone-free state of ERα36 (grey) and shows the presence of interdomain space observed in the chaperone-free state.  
235x132mm (300 x 300 DPI)

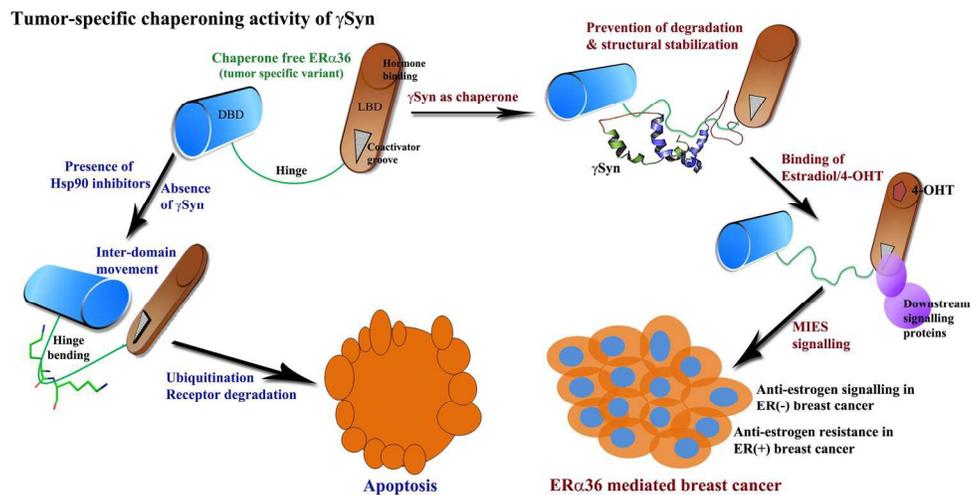


Fig. 8: Structural insights into tumor specific chaperone functions of SNCG ( $\gamma$ Syn).  
199x99mm (300 x 300 DPI)