Computational Study of Evolutionary Selection Pressure on Rainbow Trout Estrogen Receptors

Conrad Shyu[1], Celeste J. Brown[2], F. Marty Ytreberg[1]*

1 Department of Physics, University of Idaho, Moscow, Idaho, United States of America, **2** Department of Biological Sciences, University of Idaho, Moscow, United States of America

Abstract

Molecular dynamics simulations were used to determine the binding affinities between the hormone 17β-estradiol (E2) and different estrogen receptor (ER) isoforms in the rainbow trout, *Oncorhynchus mykiss*. Previous phylogenetic analysis indicates that a whole genome duplication prior to the divergence of ray-finned fish led to two distinct ERβ isoforms, ERβ1 and ERβ2, and the recent whole genome duplication in the ancestral salmonid created two ERα isoforms, ERα1 and ERα2. The objective of our computational studies is to provide insight into the underlying evolutionary pressures on these isoforms. For the ERα subtype our results show that E2 binds preferentially to ERα1 over ERα2. Tests of lineage specific dN/dS ratios indicate that the ligand binding domain of the ERα2 gene is evolving under relaxed selection relative to all other ERα genes. Comparison with the highly conserved DNA binding domain suggests that ERα2 may be undergoing neofunctionalization possibly by binding to another ligand. By contrast, both ERβ1 and ERβ2 bind similarly to E2 and the best fitting model of selection indicates that the ligand binding domain of all ERβ genes are evolving under the same level of purifying selection, comparable to ERα1.

Introduction

Estrogens are essential endogenous hormones that modulate the development and homeostasis of a wide range of target tissues, such as the reproductive tracts, breast and skeletal system [1]. Estrogenic hormones have multi-faceted and wide-ranging effects in vertebrate animals. For estrogens such as 17β-estradiol (E2) to exert their biological effects, they must interact with cellular estrogen receptors (ER). Studies have shown that ERs are part of two distinct estrogenic transduction pathways. One pathway provides a rapid, nongenomic pathway initiated by membrane bound ERs at the cell surface [1–3]. The other pathway provides direct genomic control in which ERs act as transcription factors within the cell nucleus [4,5]. These ERs are members of the nuclear receptor superfamily of ligand-modulated transcription factors [6–8]. There are two different subtypes of these ERs, referred to as α and β, each encoded by a separate gene.

Recently, Nagler et al [7] reported the novel ERα2 and both ERβ isoforms in the rainbow trout, *Oncorhynchus mykiss*, and performed a comprehensive phylogenetic analysis with all other known fish ER gene sequences. Their phylogenetic analysis indicates that the duplication leading to the two ERβ isoforms arose prior to the divergence of the ray finned fish attributable to a whole genome duplication that occurred in the Teleost ancestor (see Figure 1) [9]. The ERα isoforms, on the other hand, appear to have arisen as a result of a second more recent whole genome duplication event that occurred in the salmonid ancestor 25–100

million years ago [10]. These results indicate that the second ERα isozyme that arose during the earlier genome duplication appears to have been lost subsequently, since no other ray finned fish are known to have a second ERα isoform. This also indicates that the expected duplications of ERβ1 and ERβ2 were lost subsequent to the salmonid genome duplication.

The purpose of the study is to employ molecular dynamics simulations to determine the binding affinities between E2 and ERs of the different isoforms in the rainbow trout and to use the results to provide insight into the underlying evolutionary selection pressure on the ERs. Our binding affinity results obtained from insertion and deletion are very similar indicating that our simulations are well converged and that accurate estimates of binding affinities were obtained. Our results show that E2 binds preferentially to ERα1 over ERα2. By contrast, the difference in binding affinity is less significant for the β subtype, i.e., both isoforms bind similarly to E2. We also computed dN/dS ratios for the ER isoforms. These results suggest that the ERα1 gene is evolving under relaxed selection compared to all other salmonid ERα genes.

Results and Discussion

Experimental binding affinity results are not readily available for the four trout ERs due to the difficulty in isolating the different isoforms. Thus, to verify our methodology for estimating ΔG_{bind} for rainbow trout ER-E2 we first performed simulations using
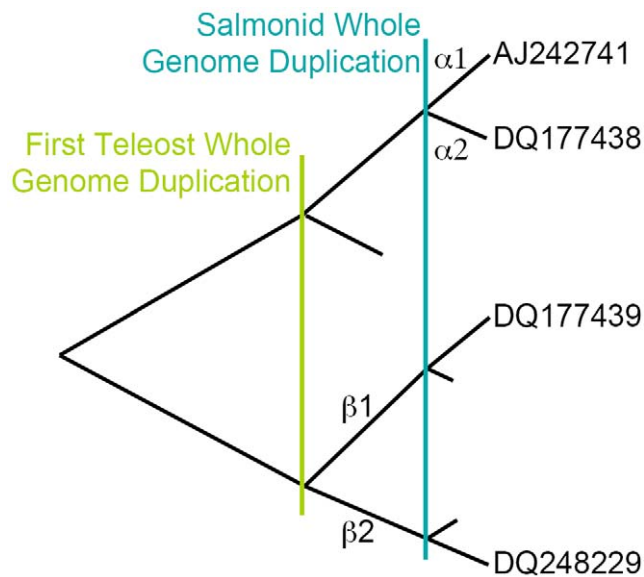
Figure 1. Schematic showing the inferred evolutionary history of the *Oncorhynchus mykiss* estrogen receptors. Vertical bars mark inferred whole genome duplication events; short branches mark inferred duplications that were lost over time.
doi:10.1371/journal.pone.0009392.g001

human ER (PDB: 1QKU) at 300 K and compared the binding affinities to the experimental results. Our computational estimates at 300 K are -63.2 kJ/mol for insertion (when interactions between E2 and its environment are turned on) and -59.6 kJ/mol for deletion (when interactions between E2 and its environment are turned off). Experimental binding affinity for human ER is -52.3 kJ/mol at 300 K [11]. Thus, our human ER binding affinity estimates are within about 10 kJ/mol of

experiment which is within the expected error due to the atomic models [12]. The trout ER simulations followed exactly the same procedure as human, beginning with docking the E2 into the ER. It is important to note that our ER simulations were performed at 277 K to closely mimic the water temperature of rainbow trout natural habitat.

Table 1 shows our binding affinity results from both insertion and deletion. Simulation results from both deletion and insertion of electrostatics and Lennard-Jones interactions provide a rudimentary assessment of the accuracy of our calculations (note that there may be inaccuracies in the atomic models but that is beyond the scope of this study). The fact that both insertion and deletion give very similar results strongly suggests that our simulations are well converged and that accurate estimates of binding affinities have been obtained.

Evolutionary and Functional Analyses

Our results in Table 1 show that the E2 binds preferentially to the ERα1 isoform of the α subtype that has been found in all salmonids. The other isoform ERα2, which appears to have arisen during the recent salmonid whole genome duplication, shares 75.4% sequence identity with the ERα1 and thus a large number of substitutions have accumulated since the initial duplication event. To infer the evolutionary pressures that led to this amount of divergence in both protein sequence and function, we examined the lineage specific differences in dN/dS ratios among the ERα sequences. We used an alignment of the codons in the ligand binding domain for all ERα sequences and a phylogeny inferred from the nucleotide sequence by the neighbor joining method (which did not differ significantly from the tree in [7]). PAML was used to calculate the log likelihood values and dN/dS ratios for each of five hypotheses: a single ratio for all branches, one ratio for all branches except the branch to the rainbow trout ERα2, separate ratios for the two ERα's from rainbow trout and the rest of the tree, separate ratios

Table 1. Estrogen receptor binding affinities for different isoforms obtained at 277 K.

(A) Insertion

	ERα1	ERα2	ERβ1	ERβ2
$\Delta G_{\text{rest}}^{RE}$	-21.5	-27.6	-31.4	-24.2
$\Delta G_{\text{elec}}^{RE} + \Delta G_{\text{LJ}}^{RE}$	-146.7	-126.5	-132.4	-144.1
$\Delta G_{\text{rest}}^{E}$	48.9	49.9	52.4	49.9
$\Delta G_{\text{elec}}^{E} + \Delta G_{\text{LJ}}^{E}$	61.8	—	—	—
$\Delta G_{V_0}^{E}$	-6.8	—	—	—
$\Delta G_{\text{bind}}^{RE}$	**-64.3**	**-49.2**	**-56.4**	**-63.4**

(B) Deletion

	ERα1	ERα2	ERβ1	ERβ2
$\Delta G_{\text{rest}}^{RE}$	-21.5	-27.6	-31.4	-24.2
$\Delta G_{\text{elec}}^{RE} + \Delta G_{\text{LJ}}^{RE}$	-149.2	-125.2	-131.1	-144.7
$\Delta G_{\text{rest}}^{E}$	48.9	49.9	52.4	49.9
$\Delta G_{\text{elec}}^{E} + \Delta G_{\text{LJ}}^{E}$	62.0	—	—	—
$\Delta G_{V_0}^{E}$	-6.8	—	—	—
$\Delta G_{\text{bind}}^{RE}$	**-66.6**	**-47.7**	**-54.9**	**-63.8**

All results are in kJ/mol. The binding affinities ΔG_{bind} were calculated using Eqn 1. Both insertion and deletion directions give very similar results which demonstrates that our simulations are well converged.
doi:10.1371/journal.pone.0009392.t001

for the rainbow trout ERα2, all ERα1 from salmonids and the rest of the tree and the full model where every branch has its own ratio (see Table 2). Using the Aikaike Information Criterion, the model with two ratios, one for the branch to the rainbow trout ERα2 and one for all other branches is the best fitting model. For this model, the dN/dS ratio for all other branches was 0.09 whereas the ratio for the ERα2 branch was 0.30. In all tests, the dN/dS ratio for the ERα2 branch was about three times greater than the other salmonid branches. Therefore, the ERα2 ligand binding domain appears to be evolving under relaxed selection relative to the other salmonid ERα1 ligand binding domains, which is consistent with the decreased affinity of this domain for E2. It is also possible that ERα2 was evolving in a neutral fashion for a short time, but then developed a new function and is now undergoing stronger purifying selection. This possibility could be explored further if more ERα salmonid gene sequences were made available.

Our results show that both ERβ isoforms bind similarly to E2, i.e., the difference between them in binding affinity is small compared to the difference between the ERα isoforms (see Table 1). The two isoforms share only 57.6% sequence identity, having arisen prior to the Teleost radiation, and the difference in their binding affinity might be expected to be greater, given this large degree of divergence. We performed a similar analysis of the dN/dS ratio for these genes by testing the following models: one dN/dS ratio for the whole tree, a dN/dS ratio for each of isoform ERβ1 and ERβ2, dN/dS ratios for each of the two rainbow trout isoforms and for each isoform for all other fish and the full model where every branch has a different dN/dS ratio (Table 2). The best fitting model for this comparison was the single dN/dS ratio (0.07) for the entire tree, indicating that both ERβ isoforms are under the same level of purifying selection. This is also consistent

with our results showing that these two ligand binding domains have similar affinity for E2.

These nuclear ERs have a significant and ubiquitous distribution in the rainbow trout [2,7]. The levels of transcription differ among the four genes with one isoform having higher transcript levels in most tissues than the other isoform. For the ERα isoforms, ERα1 has the higher transcript levels, and for the ERβ isoforms, ERβ2 has the highest transcript levels [7]. While the correlation between reduced transcription levels and binding affinity is clear in the ERα isoforms, there seems to be no such correlation for the ERβ isoforms. These two isoforms share similar binding affinity, and yet, ERβ1 has much lower expression levels than ERβ2 in juvenile rainbow trout. It is possible that both ERα2 and ERβ1 have higher expression levels at other life stages [7]. Given the age of ERβ1 and the equivalent levels of both E2 binding affinity and purifying selection compared with ERβ2, this ER clearly continues to have an important role as an estrogen receptor.

It is not as clear what ERα2's role is as an estrogen receptor. It's reduced affinity for E2, low transcript levels and evidence for relaxed selection suggests that this estrogen receptor may be undergoing subfunctionalization or neofunctionalization. One indication that ERα2 may be undergoing neofunctionalization is that the DNA binding domain of ERα2 does not have the degree of sequence variation that the ligand binding domain has. If the ERα2 was undergoing relaxed selection along it's entire length, the DNA binding domain would also show indications of greater amino acid divergence (Table 2). It appears that ERα2 is not losing its ability to bind to the canonical estrogen receptor element even though it is losing affinity for E2. This suggests that this gene may be undergoing neofunctionalization by binding to some other ligand than E2.

Table 2. Results of fitting evolutionary models for differences in dN/dS ratios.

ERα-LBD	np	$\ln L$	AIC	
H_0: Everyone is equal	59	−5645.4	11409	
H_1: Oα2 ≠ others	60	−5641.7	11403	*
H_2: Oα2 ≠ Oα1 ≠ others	61	−5641.6	11405	
H_3: Oα2 ≠ Sα1 ≠ others	61	−5641.5	11405	
H_{Full}: Everyone is different	115	−5610.3	11451	
ERα-DBD	np	$\ln L$	AIC	
H_0: Everyone is equal	59	−1121.0	2360	*
H_1: Oα2 ≠ others	60	−1120.8	2362	
H_2: Oα2 ≠ Oα1 ≠ others	61	−1120.7	2363	
H_3: Oα2 ≠ Sα1 ≠ others	61	−1120.3	2363	
H_{Full}: Everyone is different	115	−1101.8	2434	
ERβ-LBD	np	$\ln L$	AIC	
H_0: Everyone is equal	77	−7064.6	14283	*
H_1: β2 ≠ β1	78	−7063.9	14284	
H_2: Oβ2 ≠ Oβ1 ≠ β1 ≠ β2	80	−7062.1	14284	
H_{Full}: Everyone is different	151	−7005.9	14314	

LBD and DBD indicate ligand and DNA binding domains, respectively. Oα1, Oα2, Oβ1 and Oβ2 are the *O. mykiss* ERα1, ERα2, ERβ1 and ERβ2 genes, respectively. Sα1 indicates all of the salmonid ERα1 genes. β1 and β2 indicate ERβ1 and ERβ2 from all fish, respectively. np is the number of parameters in the model, $\ln L$ is the log likelihood calculated by PAML, and AIC is the Akaike Information Criterion value [35]. Models labeled with an asterisk are the best fitting models based upon the AIC values.

doi:10.1371/journal.pone.0009392.t002

Summaries

Using molecular dynamics simulations we estimated the binding affinities between the hormone 17β-estradiol (E2) and different estrogen receptor (ER) isoforms in the rainbow trout, *Oncorhynchus mykiss*. Our results show that E2 binds preferentially to ERα1 over ERα2. A recent genome wide duplication event led to two functional ERα isozymes in *O. mykiss*. Our evolutionary and functional analyses along with Nagler's evaluation of transcription levels [7] suggest that the ligand binding domain of ERα2 has been or is currently evolving under relaxed selection relative to ERα1. Low sequence divergence of its highly conserved DNA binding domain suggests that ERα2 is likely undergoing neofunctionalization, in which it continues to recognize the same estrogen receptor element in the DNA but may be binding to a different ligand. For the ERβ subtype both isoforms bind similarly to E2, in keeping with our evolutionary analyses that both isoforms of this subtype are evolving under the same degree of purifying selection.

Materials and Methods

Receptor Structures

The initial coordinates for the estradiol were first extracted from the human ER-E2 complexes (PDB: 1QKU) (Figure 2). The topologies were then generated by the PRODRG server [13] with the options of full charges and no energy minimization. The rainbow trout ER *holo* structures for the E2 binding domain were generated by SWISS-MODEL [14] using human ER as templates (PDB entries 1A52 for ERα's and 3ERT for ERβ's). Sequence identities between trout and human estrogen binding domains are within the range of 75–85%. The estradiol was first docked into the binding pocket of the receptor *holo* structure with AutoDock [15]. In this protocol, the receptor structure is held rigid and the estradiol is free to rotate and explore most probable binding poses using the Lamarckian genetic algorithm. The number of genetic algorithm runs was set to 1,000 with a population size of 5,000 individuals and 5,000,000 generations. The number of evaluations was set to 2,500,000 for each individual in the population to ensure thorough exploration of the search space. The mutation rate was set to 0.02 and crossover 0.8. Two-point crossover was used to generate the offspring at each successive generation. The genetic algorithm automatically preserved the 10 best-fit individuals to the next generation and the 10 least-fit individuals were not used to generate offspring. A total of 1,000 independent docking trials were performed for each of the four ERs. The best binding pose from each trial was collected and ranked based on the scores. These best-fit binding poses were first visually inspected for consistency with human ER and the one with the highest score was then used as the starting structure for the simulations.

Thermodynamic Cycle

To estimate E2-ER binding affinities we note that, since the free energy is a state function, it permits the selection of an arbitrary path connecting the bound and unbound states. Therefore, we decomposed the binding free energy calculation into several steps in which the E2 is annihilated (i.e., decoupled) from its bound state in the receptor complex and then made to reappear in solution to complete the thermodynamic cycle. For brevity, we subsequently define *deletion* to be when interactions between E2 and its environment are turned off and *insertion* to be when these interactions are turned on.

Figure 3 shows the thermodynamic cycle we used to calculate binding affinities (see also Refs [16–20]). Starting with upper right schematic and moving clockwise, the fully interacting E2 (blue) is

Figure 2. Crystal structure of human estrogen receptor binding domain bound to the hormone 17β-estradiol. Similar human ER structures were used as templates to generate structures for trout ERs. Image was rendered using VMD [33].
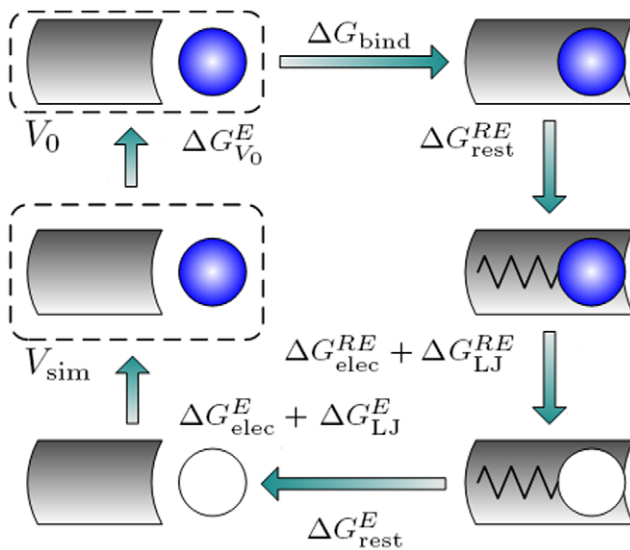doi:10.1371/journal.pone.0009392.g002

Figure 3. Thermodynamic cycle used for calculating binding affinities. Since the free energy is a state function, the calculation of binding affinity is decomposed into several steps [34]. Eqn 1 was used to calculate the binding affinity ΔG_{bind} between the ERs and E2. The gray curved rectangle represents the receptor, the blue circle represents E2 with all interactions turned on, and the white circle indicates that all interactions are turned off. The spring represents the restraints between E2 and receptors.

first restrained in the binding pocket of the receptor. Here RE represents the solvated complex of the E2 and receptor, and $\Delta G_{\text{rest}}^{RE}$ denotes the free energy of restraining the E2 in the binding pocket of the receptor which will depend on the details of restraint. Next, the electrostatic and Lennard-Jones interactions of the E2 are gradually turned off (white) in two separate steps using alchemical simulations. The free energies $\Delta G_{\text{elec}}^{RE}$ and $\Delta G_{\text{LJ}}^{RE}$ are associated with deleting or inserting electrostatic and Lennard-Jones interactions respectively. With the E2 fully decoupled from its environment, the restraint is then removed. The free energy $\Delta G_{\text{rest}}^{E}$ is associated with the removal of this restraint. Next the E2 interactions are turned back on with no receptor present. The free energy $\Delta G_{\text{elec}}^{E}$ and ΔG_{LJ}^{E} are associated with turning on the electrostatic and Lennard-Jones interactions respectively. Finally, we account for the difference between the standard (V_0) and simulation volume (V_{sim}). The binding affinity between the estrogen receptors and E2 is thus the sum of the free energies,

$$-\Delta G_{\text{bind}} = \Delta G_{\text{rest}}^{RE} + \Delta G_{\text{elec}}^{RE} \\ + \Delta G_{\text{LJ}}^{RE} + \Delta G_{\text{rest}}^{E} + \Delta G_{\text{elec}}^{E} + \Delta G_{\text{LJ}}^{E} + \Delta G_{V_0}^{E}. \tag{1}$$

Restraints

To facilitate convergence restraints were applied to restrict the positions of E2 relative to the receptors. Boresch et al [21] and Mobley et al [18] reported that the presence of multiple metastable ligand orientations can cause convergence problems for free energy estimates. The authors further suggested using a restraining potential to keep the ligand in the binding site during the simulation process. With such a restraining potential the ligand is no longer required to sample the entire simulation volume (particularly a problem when ligand is decoupled). Moreover, the

restraint minimizes the detrimental effects of end-point singularities commonly reported in alchemical simulations [17,18,20]. Mobley et al [18] also pointed out that the equilibrium geometry of the restraints is arbitrary and will not affect the asymptotic estimate of the binding free energy. In this work, we judiciously selected anchor atoms from the more rigid alpha helices that form the E2 binding pocket. The restraints included one distance (with the force constant of 1000 kJ/mol/nm^2), two angle (1000 kJ/mol/rad), and three dihedral restraints (1000 kJ/mol/rad^2) that determine the orientation of three carbons in the E2 relative to three α-carbons in the receptors (see Figure 4).

Simulation Protocols

All simulations were performed with the GROMACS 4.0 [12] compiled in single-precision mode at a constant temperature of 277 K in a periodic box with an edge length of approximately 8.2 nm and the default GROMOS-96 43A1 forcefield [22]. The simulation systems each contained approximately 16,500 Simple Point Charge (SPC) water molecules [23]. Short-range interactions were evaluated using a neighbor list of 1.0 nm updated at every 10 steps. Van der Waals interactions used a cutoff with a smoothing function such that the interactions slowly decayed to zero between 0.75 nm and 0.90 nm. A long-range analytical dispersion correction was applied to the energy and pressure to account for the truncation of the Lennard-Jones interactions [24]. Electrostatic interactions were evaluated using the particle mesh Ewald (PME) [25] with a real space cutoff of 1.0 nm, a spline order of 6, a Fourier spacing of 0.1 m, and relative tolerance between long and short range energies of 10^{-6}. All bonds to hydrogen were constrained with LINCS [26] with an order of 12, and a time step of 2 fs was used for dynamics.

For equilibration, the systems were first minimized using 1,000 steps of L-BFGS (Broyden-Fletcher-Goldfarb-Shanno) [27], followed by 1,000 steps of steepest descent minimization. The system was then subject to 1.0 ns of simulation using isothermal molecular dynamics. This was followed by another 1.0 ns of simulation using isothermal-isobaric dynamics with the Berendsen barostat with a time constant of 1.0 ns. For all simulations the temperature was maintained at 277 K using Langevin dynamics [28] with a friction coefficient of 1.0 amu/ps. The coupling time was set to 0.5 ps, and the isothermal compressibility was set to 4.5×10^{-5} bar^{-1}.

After equilibration, production simulations were run with isothermal-isobaric conditions using Langevin dynamics at the temperature of 277 K. The pressure was maintained at 1.0 atm using the Parrinello-Rahman algorithm [29]. The temperature was chosen as it closely resembles the water temperature for the natural habitat of rainbow trout. Energies were recorded every 0.2 ps during production runs, and trajectory snapshots every 1.0 ps. The first 50% of each simulation was discarded for equilibration.

Free Energy Calculations

We used the formula suggested by Boresch et al [21] to analytically calculate the free energy $\Delta G_{\text{rest}}^{E}$ associated with adding the restraints to E2 when decoupled from its environment. We also analytically calculated the free energy $\Delta G_{V_0}^{E}$ that accounts for the difference between the standard (V_0) and simulation volume (V_{sim}) [21].

The free energies $\Delta G_{\text{elec}}^{RE}$, $\Delta G_{\text{LJ}}^{RE}$, $\Delta G_{\text{elec}}^{E}$, and ΔG_{LJ}^{E}, were estimated using the thermodynamic integration (TI) method [16,19,20]. To minimize the numerical integration errors we employed the polynomial regression techniques to calculate free energy difference, instead of trapezoidal quadrature [30]. Separate
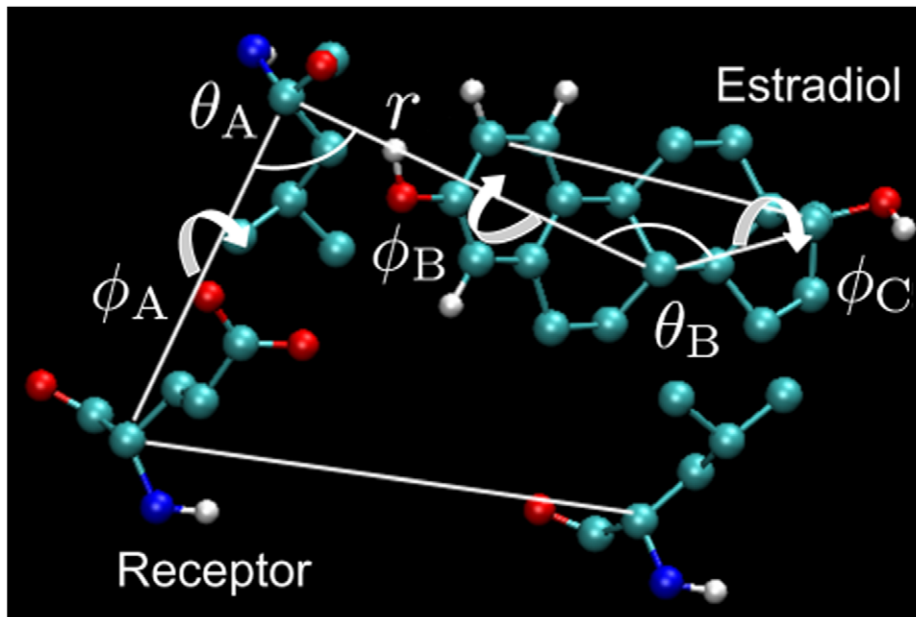
Figure 4. Restraints used for 17β-estradiol. We used six harmonic restraints acting on three anchor atoms in the E2 and three in the receptor. The restraints consisted of one distance (r), two angle (θ_A and θ_B), and three dihedral (ϕ_A, ϕ_B, and ϕ_C) restraints that determine the orientation of three carbons in the E2 relative to three α-carbons in the receptors.
doi:10.1371/journal.pone.0009392.g004

simulations were performed for changes in the Lennard-Jones with 21 values of the scaling parameter, $\lambda = 0.0, 0.05, 0.1 \ldots 0.9, 0.95$, and 1.0, and the electrostatics with 11 λ values, $\lambda = 0.0, 0.024, 0.095, 0.206, 0.345, 0.5, 0.655, 0.794, 0.905, 0.976$, and 1.0. For simulations with only Lennard-Jones, all partial charges were set to zero and the soft-core scaling parameter was set to 0.5. Once the neutral atoms were fully grown in the solvent, the second simulations then computed the free energy associated with the electrostatics with a soft-core scaling parameter of 0.0. This was accomplished by increasing the partial charges from zero to their final values given by the forcefield.

The free energy associated with the restraints, $\Delta G_{\text{rest}}^{RE}$ was calculated using the Bennett acceptance ratio approach [31]. We performed 1.0 ns equilibrium simulation for the estradiol-receptor complex using each of the harmonic restraining potentials with force constants of 0, 25, 40, 60, 90, 150, 200, 300, 450, 700, and 1000 kJ/mol/nm^2 for distance, kJ/mol/rad for angle, and kJ/mol/rad^2 for dihedral restraints. The first 0.5 ns of each simulation was discarded for equilibration and the remaining 0.5 ns was used to compute the free energy differences. No attempt was made to optimize the efficiency of the calculation since our primary objective was to obtain accurate estimates of the restraining free energies.

Evolutionary Analyses

The following sequences were extracted from GenBank: AB037185, AF349412, A133920050, AY727528, AY775183, BD105560, AB190289, AJ487687, AY055725, AF061275, AF253505, AY520443, AJ242741, DQ009007, DQ248228, DQ177438, X89959, TNU7560, AY422089, AF298183, AF136979, AY074780, AB007453, AJ006039, AF253062, AY223902, ORZMER, AY917147, AF326201, AY305026, NM_180966, NM_174862, AB003356, AB070630, AB070901, AB083064, AB117930, AB190290, AF061269, AF136980, AF177465, AF185568, AF298181, AF298182, AF349413, AF349414, AF516874, AJ275911, AJ289883, AJ314602, AJ314603, AJ414566, AJ414567, AJ489523, AJ580050, AY074779, AY211021, AY211022, AY305027, AY307098, AY508959, AY566178, AY770578, AY917148, BC044349, BC086848, DQ177439, DQ248229, TNU75605. The first 30 are ERα sequences and the other 39 are ERβ sequences, and the following analysis was done separately for these two subtypes. The codons were aligned based upon their aligned amino acid sequences, and these alignments were used to infer tree topologies using the neighbor joining method. Then the ligand binding domains were extracted from the alignments. PAML was used to test several codon-based likelihood models that allow for variable dN/dS ratios among lineages based upon the inferred phylogenies and the aligned ligand binding domains [32].

Acknowledgments

Author Contributions

Conceived and designed the experiments: CS CJB FMY. Performed the experiments: CS CJB. Analyzed the data: CS CJB FMY. Contributed reagents/materials/analysis tools: CS CJB. Wrote the paper: CS CJB. Principal investigator: FMY.

References

1. Ruff M, Gangloff M, Wurtz J, Moras D (2000) Estrogen receptor transactivation structure-function relationship in dna and ligand-binding domains of estrogen receptors. Breast Cancer Research. pp 353–359.

2. Nagler J, Krisfalusi M, Cyr D (2000) Quantification of rainbow trout (Oncorhynchus Mykiss) estrogen receptor-alpha messenger rna and its expression in the ovary during the reproductive cycle. J Mole Endorinol 25: 243–251.

3. Pakdel F, Metivier R, Flouriot G, Valotaire Y (2000) Two estrogen receptor (ER) isoforms with different estrogen dependencies are generated from the trout ER gene. Endocrinol 141: 571–580.

4. Fröehlicher M, Liedtke A, Groh K, López-Schier H, Neuhauss S, et al. (2009) Estrogen receptor subtype β2 is involved in neuromast development in zebrafish (Danio rerio) larvae. Dev Biol 330: 32–43.

5. Menuet A, Le Page Y, Torres O, Kern L, Kah O, et al. (2004) Analysis of the estrogen regulation of the zebrafish estrogen receptor (ER) reveals distinct effects of ERα, ERβ1 and ERβ2. J Mol Endo 32: 975–986.

6. Dréan Y, Kern L, Pakdel F, Valotaire Y (1995) Rainbow trout estrogen receptor receptor presents an equal specificity but a differential sensitivity for estrogens than human estrogen receptor. Mole and Cel Endorinol 109: 27–35.

7. Nagler J, Cavileer T, Sullivan J, Cyr D, Rexroad I (2007) The complete nuclear estrogen receptor family in the rainbow trout: discovery of the novel ERα2 and both ERβ isoforms. Gene 392: 164–173.

8. Tsai M, O'Malley B (1994) Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Ann Rev Biochem 63: 451–486.

9. Hoegg S, Birnkmann H, Taylor J, Meyer A (2004) Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J Mol Evol 59: 190–203.

10. Allendorf F, Thorgaard G (1984) Tetraploidy and evolution of salmonid fishes. Evolutionary Genetics of Fishes. New YorkNY: Plenum Press. pp 1–53.

11. Petit F, Valotaire Y, Pakdel F (1995) Differential functional activities of rainbow trout and human estrogen receptors expressed in the yeast Saccharomyces cerevisiae. Euro J Biochem 233: 584–592.

12. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theo and Comput 4: 435–447.

13. Schüttelkopf A, van Aalten D (2004) PRODRG: a tool for high throughput crystallography for protein-ligand complexes. Acta Crystallographics D60: 1355–1363.

14. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22: 195–201.

15. Morris G, Goodsell D, Halliday R, Huey R, Hart W, et al. (1998) Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. J Comput Chem 28: 1639–1662.

16. Kirkwood J (1935) Statistical mechanics of fluid mixtures. J Chem Phys 3: 300–313.

17. Mitchell M, McCammon J (1991) Free energy difference calculations by thermodynamic integration: difficulties in obtaining a precise value. J Comput Chem 12: 271–275.

18. Mobley D, Chodera J, Dill K (2006) On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. J Chem Phys 125: 084902.

19. Shirts M, Pitera J, Swope W, Pande V (2003) Extremely precise free energy calculations of amino acid side chain analogs: comparison of common molecular mechanics force fields for proteins. J Chem Phys 119: 5740.

20. Shirt M, Vijay S (2005) Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. J Chem Phys 122: 144107.

21. Boresch S, Tettinger F, Leitgeb M, Karplus M (2003) Absolute binding free energies: a quantitative approach for their calculation. J Phys Chem B 107: 9535.

22. van Gunsteren W, Billeter S, Eising A, Hünenberger P, Krüger P, et al. (1996) Biomolecular simulation: the GROMOS96 manual and user guide. Zürich, Germany: Vdf Hochschulverlag AG an der ETH Zürich.

23. Berendsen H, Postma J, van Gunsteren W, Hermans J (1981) Interaction models for water in relation to protein hydration. Intermolecular Forces. Dordrecht, Netherlands: Reidel. 331 p.

24. Allen M, Tildesley D (1989) Computer simulation of liquids. New York, NY: Oxford University Press.

25. Darden T, Perera L, Li L, Pedersen L (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. Structure 7: R55–R60.

26. Hess B, Bekker H, Berendsen H, JF (1997) LINCS: a linear constraint solver for molecular simulations. J Compu Chem 18: 1463–1472.

27. Broyden C (1970) The convergence of a class of double-rank minimization algorithms 1. general considerations. IMA J Appl Math 6: 76–90.

28. van Gunsteren W, Berendsen H (1988) A leap-frog algorithm for stochastic dynamics. Mol Simul 1: 173–185.

29. Laio A, Parrinello M (2002) Escaping free-energy minima. Proc Natl Acad Sci USA 99: 12562–12566.

30. Shyu C, Ytreberg F (2009) Reducing bias and uncertainty of free energy estimates by using polynomial regression to analyze thermodynamic integration data. J Comput Chem 30: 2297–2304.

31. Bennett C (1976) Efficient estimation of free energy differences from Monte Carlo data. J Comput Phys 22: 245–268.

32. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24: 1586–1591.

33. Humphrey W, Dalke A, Schulten K (1996) VMD – visual molecular dynamics. J Molec Graphics 14: 33–38.

34. Chipot C, Pohorille A (2007) Free energy calculations. New York, NY: Springer-Verlag.

35. Akaike H (1974) A new look at the statistical model identification. IEEE Transactions on Automatic Control 19: 716–723.