Simulations of the Estrogen Receptor Ligand-Binding Domain: Affinity of Natural Ligands and Xenoestrogens

B. Chris Oostenbrink,[†] Jed W. Pitera,[‡] Marola M. H. van Lipzig,[†] John H. N. Meerman,[†] and Wilfred F. van Gunsteren*,[‡]

Division of Chemistry, Department of Pharmacochemistry, Section Molecular Toxicology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands, and Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland

Received July 31, 2000

We have carried out molecular dynamics (MD) simulations and free energy calculations on the α-subtype of the human estrogen receptor ligand-binding domain (ERα LBD) complexed with a number of known agonists and putative xenoestrogens. Our dynamical simulations of ligand−receptor complexes underscore the highly structured nature of the complex and offer some interesting insights into the structure−activity relationship (SAR) for these ligands. With traditional thermodynamic integration (TI) calculations, we calculate relative binding free energies for three known agonists, in good agreement with experimental values. The sheer number of possible xenoestrogenic compounds makes an approach using traditional free energy calculations unfeasible. Instead, we have made use of a single-step perturbation methodology that allows the calculation of relative free energies for a large number of related polyaromatic hydrocarbons (PAHs) from a single simulation. Our results show good (maximum deviation 3.3 kJ mol^{-1}) agreement with experimental data, suggesting the possibility of large-scale xenoestrogen screening in silico to obtain strongly estrogenic compounds for subsequent experimental testing.

Introduction

Estrogens are a family of naturally occurring steroid hormones that exert a physiological effect in the growth, development, and maintenance of a diverse range of tissues. These effects are the result of the activity of the estrogen receptor, of which two subtypes (ERα and ERβ) are currently known.[1,2] The ER is also involved in a range of diseases such as breast cancer, osteoporosis, endometrial cancer, and prostate hypertrophy.[3]

Like other small-molecule hormone receptors, the ER is a member of the nuclear hormone receptor (NR) superfamily. Other members include such species as the thyroid hormone receptor (TR) and retinoic acid receptor (RAR).[4] The NR family shares a common multidomain architecture, consisting minimally of an N-terminal DNA-binding domain (DBD), a ligand-binding domain (LBD), and a C-terminal activation domain. Unlike membrane-bound hormone receptors such as the human growth hormone receptor,[5,6] the nuclear receptors are usually found unbound or associated with heat shock proteins in the cytosol in the absence of ligand. Ligand binding triggers a translocation of the receptor protein to the nucleus, where the DBD of the receptor protein interacts directly with response elements on the DNA, acting to activate or repress transcription.[7]

Following the solution of the TR and RAR LBDs,[8,9] there has been an explosion of high-resolution structural data on nuclear receptor LBDs. Specifically, since 1997, several crystal structures of the LBD of the ER complexed to various agonists and antagonists have become available.[10−12]

Two particularly interesting structural features are present in the known structures of ligand/LBD complexes. First, the ligand occupies a deeply buried, largely hydrophobic cavity within the LBD. In most structures there is no clear path or channel connecting this cavity with the exterior of the protein, in contrast to typical protein−ligand complexes. Second, the structure of an antagonist, raloxifene, bound to the ER LBD shows that a bulky side chain moiety of the antagonist serves to prevent the formation of the active tertiary structure of several helices (particularly helices 11 and 12) known to be crucial for activity.[13,14] Figure 1 shows the overall architecture of the ER LBD and the relationship between the buried ligand and helices 11 and 12.

In addition to the biochemical and physiological importance of the natural ER ligands, in recent years considerable concern over the effects of xenoestrogens, exogenous ER agonists, has emerged. It appears that the ER shows affinity for a remarkably wide range of structurally diverse compounds, such as polyaromatic hydrocarbons (PAHs), phytoestrogens, phthalates, and pesticides.[15−19]

PAHs are products of incomplete combustion of fossil fuels, wood, and other organic matter and as such are ubiquitous in the environment. Oxidative biotransformation by cytochrome P450 introduces hydroxyl groups, which results in structures that can mimic estradiol, the steroidal endogenous ligand of the ER.[20,21] Everyday foodstuffs are another source of estrogen-like compounds. Some foods are known to contain a wide range of phytoestrogens, such as the isoflavonoids genistein and daidzein that show estrogenic activity in vitro and

* To whom correspondence should be addressed. Tel: +41-1-632-5501. Fax: +41-1-632-1039. E-mail: wfvgn@igc.phys.chem.ethz.ch.
[†] Vrije Universiteit Amsterdam.
[‡] Swiss Federal Institute of Technology.

10.1021/jm001045d CCC: $19.00
Published on Web 11/11/2000

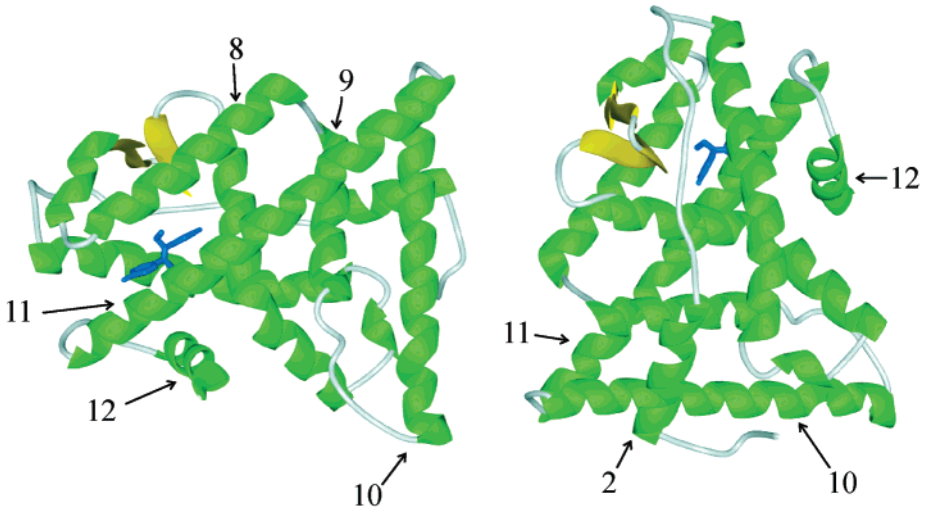


Figure 1. Ribbon model of the ER LBD complexed to **2**, with some secondary structure elements labeled. Helices are shown in green, sheets in yellow, and loops in white. The diethylstilbestrol (**2**) ligand is displayed in blue. Helices mentioned in the text have been labeled with the appropriate number, according to standard nomenclature.

in vivo. The most important human source of these isoflavonoids are soybeans and soybean products.[22]

Given the diverse range of compounds that may bind to the ER LBD and exert an effect on human and animal health, there is considerable medical interest in understanding the details of ligand LBD affinity and developing techniques to predict the affinity of compounds for the ER LBD. Previous computational efforts in this area have concentrated mainly on empirical regression-based (QSAR, CoMFA) approaches.[23−25] In contrast, we decided to explore the ligand/LBD interaction using a combination of molecular dynamics (MD) simulation and detailed free energy calculations. MD simulates the classical dynamics of a microscopic, atomic model of a biomolecular system, including ions and solvent, via the numerical integration of Newton's equations of motion. It provides detailed information on the fluctuations, dynamics, and solvation of the simulated system. In addition, the conformations generated from MD simulations in the appropriate ensemble are used in traditional free energy calculations (free energy perturbation (FEP) or thermodynamic integration (TI)) to determine or predict the *relative* free energies of two species: for example, two ligands, two mutants, or two solutes.[26]

To carry out these calculations, a coupling parameter λ is used to smoothly interpolate a simulated system between a potential energy function that corresponds to one chemical species ("A", with potential energy V_A) and one that corresponds to the alternate species of interest ("B", with potential energy V_B). Equation 1 shows this clearly:

$$V(r,\lambda) = \lambda V_A(r) + (1 - \lambda)V_B(r) \qquad (1)$$

In this way, a nonphysical path is constructed between the two chemical species (or "end points") A and B. Free energy calculations then calculate the work to convert a system from species A to species B along this path defined by λ. Since free energy is a state function, it is path-independent, and the relative free energies of A and B can be determined in this manner. The relative binding free energies discussed in this paper are actually the difference of two relative free energies: the relative free energies of two ligands in solution versus the relative free energies of those two ligands bound to the ER LBD. This is shown more clearly by the thermodynamic cycles used in our calculations (Figure 2).

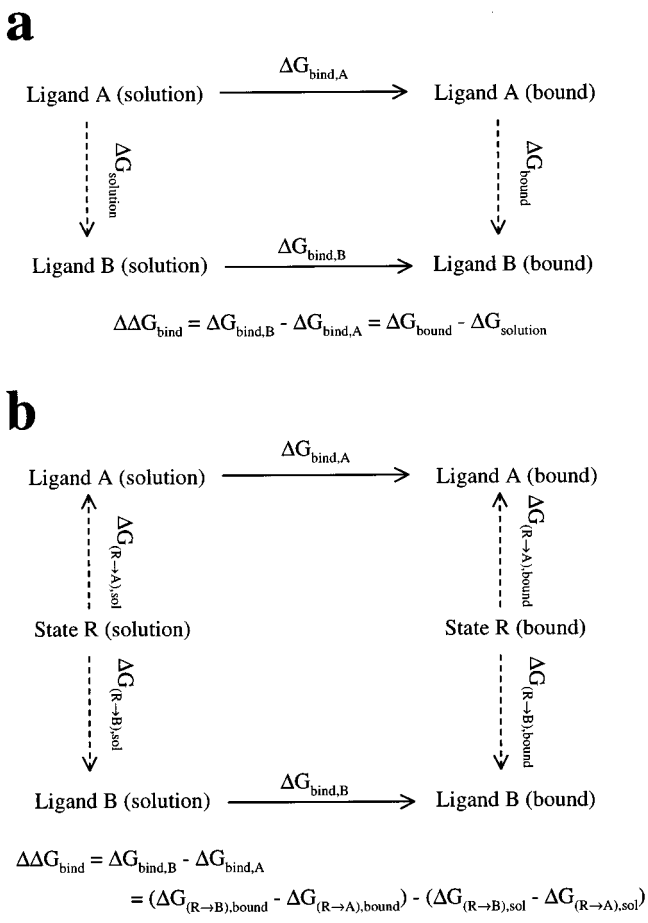


Figure 2. Thermodynamic cycles used in the TI (a) and single-step perturbation (b) free energy calculations. The symbol R denotes a nonphysical reference state chosen to enhance the sampling of configurations relevant for both physical end states A and B.

A drawback of traditional free energy calculations is their ability to compare only two species in a series of

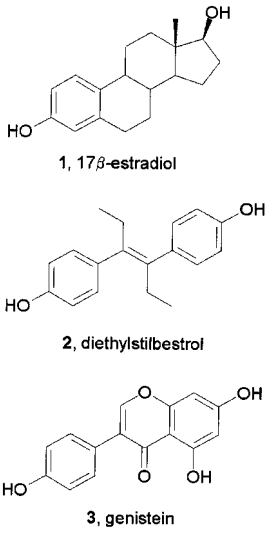


Figure 3. Ligands studied by MD and TI calculations: 17β-estradiol (**1**, E2), diethylstilbestrol (**2**, DES), and genistein (**3**, GEN).

simulations: the two end states of the path, above.[27] Given that these calculations typically take days to weeks of computer time on modern workstations, alternative approaches are needed to enhance the practical utility of these calculations. In our case, there are hundreds if not thousands of possible xenoestrogen ligands that are of interest. To begin to address this problem, we have explored the use of both traditional free energy calculations (TI[28]) and single-step perturbation techniques[29,30] to determine or predict the affinity of a number of ER ligands.

Rigorous TI calculations were used to calculate relative free energies of binding for the three compounds shown in Figure 3: 17β-estradiol (E2, **1**), the synthetic estrogen diethylstilbestrol (DES, **2**), and genistein (GEN, **3**). These three compounds are structurally very distinct, providing a significant challenge for accurate free energy calculations. In addition, a single-step perturbation approach was used to obtain relative free energies of binding for a range of structurally similar PAHs and phytoestrogens (Figure 4) from a single simulation. To complement our free energy calculations, we have also carried out extensive (500 ps−1 ns) simulations of ligands **1**−**3** bound to the ER LBD.

System Setup

Ligands. For the different ligands that were studied, the parameters were chosen based on similar building blocks in the GROMOS force field, version 43A1.[31] Ligand parameters and topologies are described fully in the Supporting Information for this paper. For the TI calculations between **1**−**3** a so-called dual-topology[32] approach was chosen, after initial studies proved single-topology methods to be inappropriate for these complex changes.[33] In general, part of the diethylstilbestrol (**2**) was gradually changed into dummy (noninteracting) atoms while at the same time nonbonded interactions of dummy estradiol (**1**) or genistein (**3**) atoms were increased from zero to the appropriate values. In each calculation the two ligand species always have a single phenolic aromatic ring in common. The dual topologies used are shown in Figure 5. During the TI calculations,

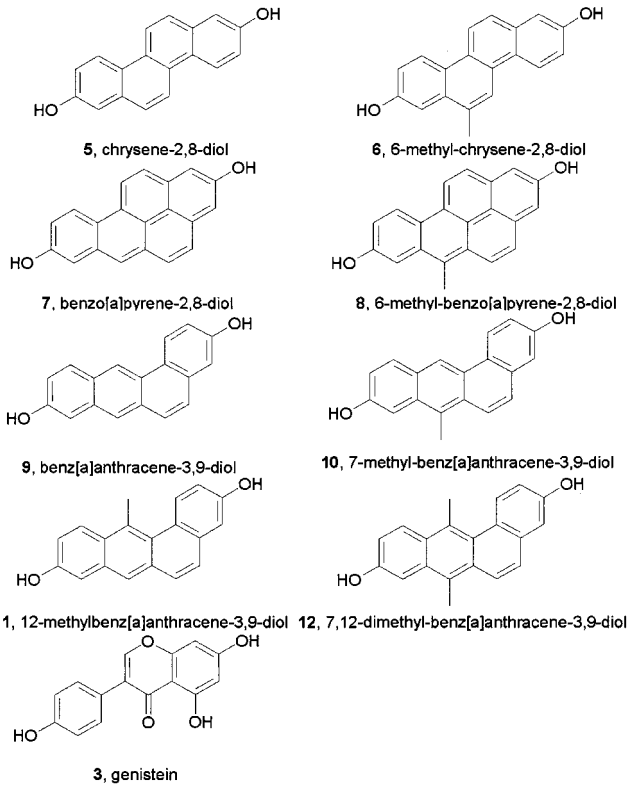


Figure 4. Compounds studied in the single-step perturbation calculations: chrysene-2,8-diol (**5**, CHN), 6-methylchrysene-2,8-diol (**6**, MCHN), benzo[*a*]pyrene-2,8-diol (**7**, BAP), 6-methylbenzo[*a*]pyrene-2,8-diol (**8**, MBAP), benz[*a*]anthracene-3,9-diol (**9**, BAA), 7-methylbenz[*a*]anthracene-3,9-diol (**10**, 7MBA), 12-methylbenz[*a*]anthracene-3,9-diol (**11**, 12MBA), 7,12-dimethylbenz[*a*]anthracene-3,9-diol (**12**, DMBA), and genistein (**3**, GEN).

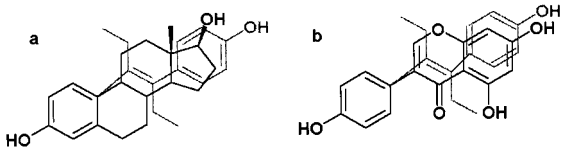
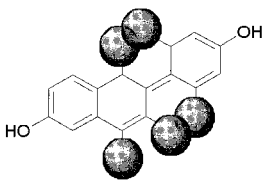


Figure 5. Representations of the dual-topology ligand used for TI calculations between DES (**2**) and E2 (**1**) (a) and between DES (**2**) and GEN (**3**) (b). During the calculation, the DES-specific atoms (light gray) are turned into dummy atoms while the E2 or GEN atoms (black) come into existence. The phenolic ring (black) is common to both end states and does not change.



R

Figure 6. Soft-core reference state (**R**) used for the single-step perturbation calculations. Soft van der Waals interaction sites are drawn as spheres.

soft-core nonbonded interactions[34] were used for all perturbed atoms with an α-parameter of 0.5.

The single-step perturbations were carried out from the reference state (**R**; see Figure 6) by perturbing the nonbonded interactions of atoms and soft-cores to the appropriate values for the dihydroxylated PAHs **5**−**12** as well as genistein (**3**), all shown in Figure 4. For

Table 1. System Composition for Solvent and Protein Simulations[a]

	solvent simulations		protein simulations	
number of:	molecules	atoms	molecules	atoms
protein			1	2439
ligand	1	30/30[b]	1	30
ions			6	6
water	1239/1247[b]	3717/3741[b]	15824	47472

[a] A single molecule of the simulated ER LBD consists of 245 amino acids. [b] TI and single-step perturbation simulations, respectively.

Table 2. Tautomeric Configurations of Histidine Residues in the Simulated ER LBD[a]

	histidine											
	356	373	377	398	474	476	488	501	513	516	524	547
tautomer	T	P	P	P	T	P	T	P	P	T	T	T

[a] T indicates protonation of the imidazole ring at $N_{\epsilon 2}$, and P indicates protonation at $N_{\delta 1}$.

compounds **3**, **11**, and **12** bond and angle parameters were also changed as necessary. In **3**, the bond between atoms 12 and 13, bond types 15 (1) and 26 (2), have both been tested (see Supporting Information). The soft-core atoms in the reference state were simulated with an α-parameter of 1.51.[35]

Simulations. Simulations of the ligands free in solution were carried out in periodic rectangular boxes containing 1239 simple point charge (SPC)[36] water molecules for the TIs and 1247 SPC water molecules for the single-step perturbations. System sizes are summarized in Table 1. All protein simulations were carried out using as initial structure the crystallographic structure of the LBD of the α-subtype receptor with DES as the ligand, entry 3ERD[11] from the RCSB Protein Data Bank,[37] residues Ser 305 to Leu 549. In this structure, the side chains of 18 residues were not resolved, whereas for Asp 321 and Arg 363 two different conformations were proposed. The missing side chains, mainly members of surface loops, were introduced in the structure using the program PROSSC and standard side chain configurations from the GROMOS96 package.[31] For Asp 321 and Arg 363 the B configurations in the structure file were selected, based on local geometries. The terminal amino acids, Ser 305 and Leu 549, were treated as charged.

In the GROMOS force field (version 43A1) aliphatic hydrogen atoms are treated as united atoms together with the carbon atom to which they are attached. The coordinates of polar hydrogen atoms (bound to nitrogen or oxygen) and aromatic hydrogen atoms were generated using the program PROGCH. The charges of the ionizable groups were chosen to correspond to a pH of 7, resulting in a net charge of −6e. The histidine residues were assigned tautomeric configurations as specified in Table 2, based on the local environment of these residues.

The protein was placed at the center of a periodic rectangular box (7.52 × 7.55 × 8.96 nm^3) which was filled with 15 824 SPC water molecules and 6 Na^+ ions to obtain a net charge of zero. In this procedure, the distance between water oxygen atoms and non-hydrogen protein atoms was a minimum of 0.23 nm. To relax the configuration of the solvent, a steepest descent minimization was carried out in which the protein and ligand atoms were positionally restrained to their initial positions using a harmonic interaction with a force constant of 25 10^4 kJ mol^{-1} nm^{-2}.

Table 3. Equilibration Protocol for MD Simulations

step	length (ps)	temp (K)	τ_T (ps)	K_{r0} (kJ mol^{-1} nm^{-1})	ensemble	τ_p (ps)
1	2	50	0.01	2.5×10^4	NVT	
2	5	100	0.01	1.0×10^4	NVT	
3	5	150	0.01	5.0×10^3	NVT	
4	5	200	0.01	1.0×10^3	NVT	
5	5	250	0.01	1.0×10^2	NVT	
6	5	300	0.01	1.0×10	NVT	
7	10	300	0.1	0.0	NPT	0.5

Protocol and Methods

All simulations were carried out using the GROMOS96 biomolecular simulation package.[31,38] Individual MD simulations were equilibrated according to the scheme in Table 3. In a total of 37 ps the temperature was increased stepwise from 50 K (velocities randomly initialized from a Maxwell−Boltzmann distribution) to 300 K, using first-order coupling to a temperature bath[39] with temperature relaxation time τ_T. The solute and water molecules were separately coupled to the temperature bath. Protein and ligand atoms were positionally restrained to the crystal structure coordinates using a harmonic interaction with force constant K_{r0}. After 27 ps the simulation was carried out at a constant pressure of 1 atm, again using a first-order coupling to a pressure bath[39] and pressure relaxation time of $\tau_p = 0.5$ ps. An estimated value of 45.75×10^{-5} (kJ mol^{-1} nm^{-3})$^{-1}$ was used for the isothermal compressibility.[31]

Bond lengths were constrained to ideal values, using the SHAKE algorithm,[40] allowing for a time step of 2 fs. Nonbonded interactions were calculated using a twin-range cutoff. Short-range interactions (within 0.8 nm) were evaluated every time step from a charge-group pair list that was updated every 5 time steps. Longer-range interactions (within 1.4 nm) were evaluated every 5 time steps as well and kept constant between updates. A reaction field correction,[41] using an effective dielectric (ϵ) of 54.0,[42] was used to approximate electrostatic interactions outside the outer 1.4-nm cutoff.

The relative free energies of **1**−**3** were determined in three different environments: in vacuum, in solution, and in the protein. Conversion of the ligands was performed using 11 equally spaced λ-values. At each λ-value one simulation of 20 ps was carried out, consisting of 10 ps of equilibration followed by 10 ps of data collection over which the average $\langle \partial H/\partial\lambda \rangle_\lambda$ was calculated. For the conversion of **2** into **3** in the solvent three additional simulations at $\lambda = 0.05$, 0.85, and 0.95 were introduced to smooth the integrand. Integrations over λ were carried out using trapezoidal integration. For each environment, λ-values were changed in both the forward ($\lambda = 0 \rightarrow \lambda = 1$) and backward ($\lambda = 1 \rightarrow \lambda = 0$) directions. For the protein simulations a Maxwell−Boltzmann distribution of the velocities corresponding to a temperature of 300 K was used to initiate an equilibration starting from step 7 of Table 3.

Extensive simulations of the end states ($\lambda = 0$ and $\lambda = 1$) were carried out under the conditions of equilibration step 7 for the DES (**2**; 1 ns), E2 (**1**; 500 ps) and GEN (**3**; 500 ps) LBD complexes. Typically, 10 ps of simulation for a protein−ligand complex required roughly 20 h of calculation on a dual-processor 450-MHz Pentium II PC.

The single-step perturbation methodology makes use of the fact that the FEP formula for calculating the free energy difference between two states defined by $\lambda = 1$ and $\lambda = 0$ is:

$$\Delta G = G(1) - G(0) = -kT \ln \langle e^{-(H(1)-H(0))/kT} \rangle_0 \qquad (2)$$

where the brackets represent an ensemble average over the configurations corresponding to state $\lambda = 0$. No simulation of state $\lambda = 1$ is actually required to calculate the free energy

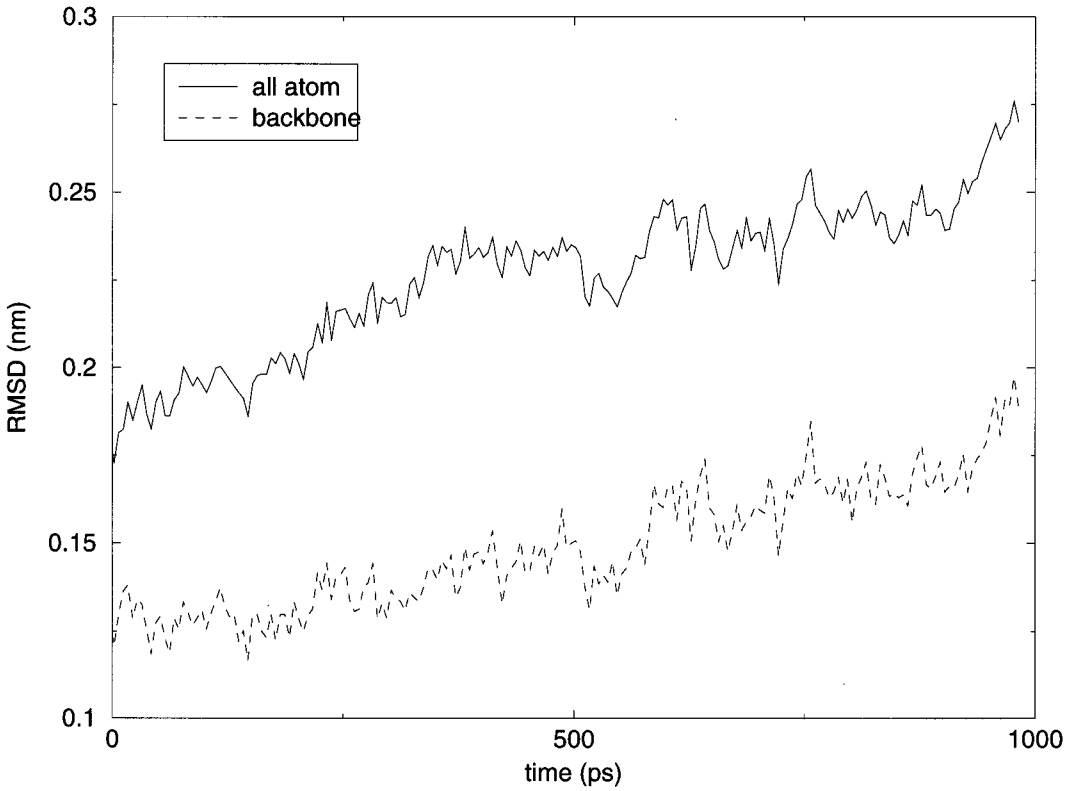


Figure 7. All-atom (solid line) and backbone (CA, N, C, O) atom (dashed line) root-mean-square positional deviations from the initial crystal structure for the 1-ns DES/ER LBD simulation.

difference. In principle, this means that a single simulation, for example of a protein−ligand complex, can be used to predict the relative free energies of any other ligand bound to that protein. In practice, statistical considerations and limited sampling times require that the two states $\lambda = 0$ and $\lambda = 1$ be reasonably similar to achieve accurate results. Traditional FEPs work around this limitation by actually performing a number of perturbations between states smoothly interpolated between those $\lambda = 0$ and $\lambda = 1$ end points. In the single-step methodology, however, care is taken to create a nonphysical reference state (R) that samples configurations relevant to a large number of interesting ligand end states.[29,43] Typically the ligand is "soft" − that is, the normal van der Waals interaction is adjusted to allow occasional overlaps of the ligand with the atoms of its surroundings. This permits the nonphysical ligand−receptor complex to sample configurations that are favorable for real ligands of widely varying shapes and sizes. A single extensive simulation of this soft, nonphysical ligand is carried out, and the perturbation formula above is used to calculate the free energy difference between the nonphysical ligand and two or more real ligands. Since the free energy is a state function, it is then possible to determine the relative free energies of two real ligands (1, 0) from their free energies relative to the nonphysical reference state (R):

$$\Delta\Delta G_{1\to 0} = \Delta G_{1\to R} - \Delta G_{0\to R} \qquad (3)$$

This is shown schematically in Figure 2, which details the thermodynamic cycles used in our calculations.

For the single-step perturbations a 1-ns trajectory of the soft-core reference state in vacuum and 2 ns in the solvent were obtained under the conditions of equilibration step 7 (Table 3). Coordinates were saved every 100 time steps (0.2 ps) for the subsequent perturbation calculations. The complex of the soft-core reference state and the ER LBD was built from the coordinates of the DES end state simulation at 470 ps, to minimize the required equilibration time. The DES was replaced by the soft-core reference state in the binding cavity, and velocities were randomly reassigned according to a Maxwell−Boltzmann distribution at 300 K. After 30 ps of equilibration a 1-ns trajectory of configurations was generated under the conditions of equilibration step 7 (Table 3), but under constant volume rather than constant pressure. The coordinates of the **R**−ER LBD complex were saved every 50 time steps for later analysis.

Using the single-step perturbation formalism, free energy differences between the compounds in Figure 4 and the soft-core reference state were calculated from these trajectories. To test the convergence of these calculations, the simulations of the soft-core reference state in water and the protein were divided into 2 (water; 1 ns each) or 5 (protein; 200 ps each) separate subtrajectories and the results were compared across the subsimulations.

Results

Molecular Dynamics. All three protein−ligand complex simulations (DES, E2, GEN) were largely similar in behavior. Visual inspection showed no large-scale changes in the protein's secondary or tertiary structure over the course of each simulation. On the basis of this observation, the longest trajectory (DES, 1 ns) was selected and examined in detail. Various analysis programs, again from the GROMOS software package, were used to characterize the structure and dynamics of the DES/ER LBD complex over the course of the simulation.

Atom-positional root-mean-square deviations (rmsds) are one way to compare different conformations of a protein. The rmsds for both the backbone atoms (CA, N, C, O) and all atoms (excluding the ligand and ions) of the trajectory coordinates versus the initial structure are shown in Figure 7. The overall values are reasonable (all atoms up to 0.275 nm, backbone up to 0.2 nm), though their general upward trend over a nanosecond suggests that the protein is diverging somewhat from its starting structure. To eliminate the contributions of

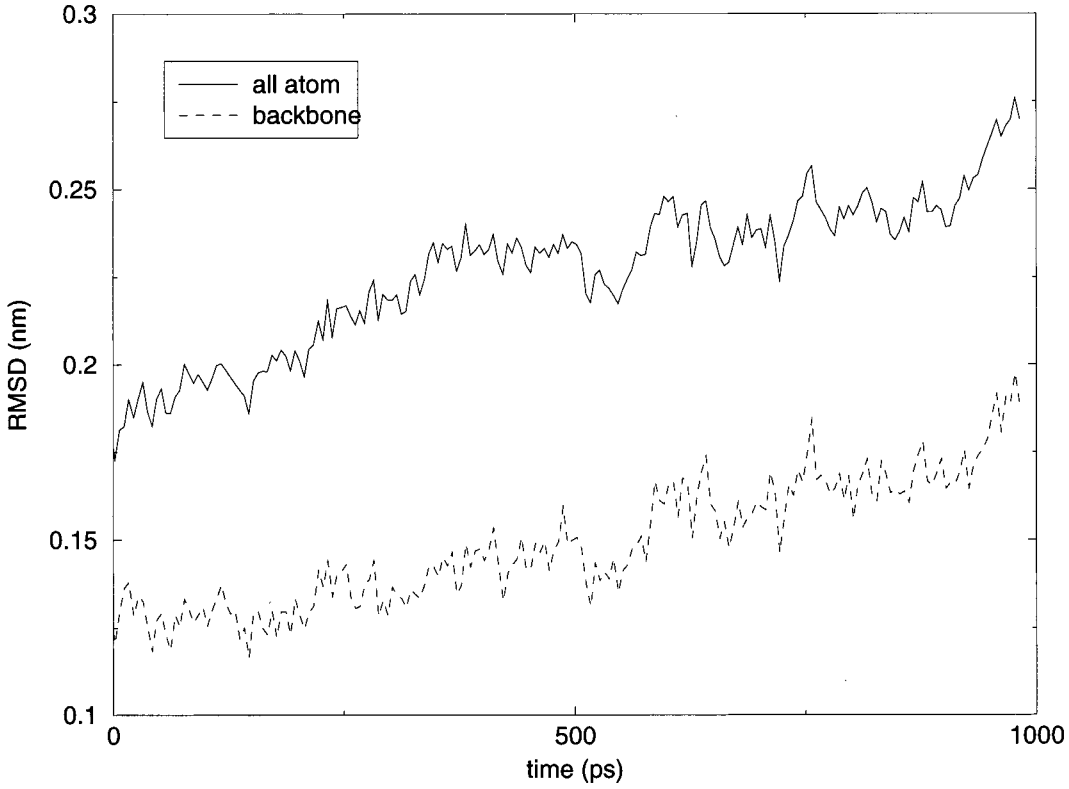


Figure 8. Backbone (CA, N, C, O) atom root-mean-square positional deviations from the initial crystal structure for all helices (solid line) and helix 12 (dashed line) of the ER LBD in the DES/ER LBD simulation.

flexible loop regions, backbone rmsds were calculated for all of the helices of the protein and for helix 12 only. These results are displayed in Figure 8. The value for all of the helices shows that the protein architecture is well-preserved, though there is a slight deformation toward the end of the calculation. In contrast, the structure of helix 12 is very stable, showing little or no disruption of the secondary structure in this functionally critical region.

The fluctuations of atoms about their mean position − somewhat analogous to crystallographic *B*-factors − give an indication of regions of the protein that are more or less flexible. Figure 9 compares the backbone atom-positional mean-square fluctuations for residues of the protein from the simulation with the experimental *B*-factors reported with the X-ray structure. Unsurprisingly, there is good qualitative agreement with the experimental data, which show that flexibility is concentrated in the surface loops of the protein, while the helices remain relatively well-ordered. The stretch of residues from 455 to 475 is a clear feature in both sets of *B*-factors and corresponds to the surface loop between helices 9 and 10 as well as the N-terminus of helix 10 itself (Figure 1). Interestingly, this region contains a highly charged stretch (470−473) of two glutamic acids, a lysine, and an aspartic acid.[12] The side chains of the first three of these residues were not resolved in the original structure, while the aspartic acid was solved in two possible conformations. The strong charge−charge interactions between these side chains, as well as their solvent-exposed position, may account for the observed dynamic behavior.

In the initial setup of the simulation, two additional (noncrystallographic) waters were placed in voids inside the ligand-binding cavity. The ligand-binding cavity is shown in schematic form in Figure 10. Since the waters were placed on purely geometric rather than energetic criteria, they shared a hydrophobic pocket with one of the ethyl groups (C8′−C9′ in the X-ray structure) of DES. This pocket is next to the histidine (His 524) that anchors one end of the ligand and is formed by Gly 420, Ile 424, the aliphatic parts of Lys 520, and Gly 521. Clearly, the hydrophobic character of this pocket makes it an unfavorable location for water to be found. The first water escapes the pocket after roughly 50 ps of postequilibration dynamics, while the second has left by 300 ps. Both waters follow the same pathway, past His 524 toward the charged moiety of Lys 520 and out into the bulk solvent. No other waters enter the cavity over the entire course of the simulation, and the crystallographic waters found in the hydrogen-bonded cluster at the other end of the ligand (near Glu 353 and Arg 394) remain trapped next to the ligand over the entire simulation.

The void volume near the ethyl group of DES allows it to fluctuate once the water escapes. Analysis of the vinyl−ethyl torsion (CY2−CY1−CA1−CB1 or CY1−CY2−CA2−CB2; see Supporting Information) shows that this ethyl group has much more flexibility in the binding site compared to the other, more restricted ethyl group. It samples a full range of torsional positions and shows an average torsion of 228 ± 180° versus 108 ± 39° for its more restricted counterpart.

Perhaps the most significant dynamics in the protein−ligand interface occur in the hydrogen bonds between the two DES phenolic hydroxyls and protein side chains. Monitoring the hydrogen bonds at both ends of the ligand over the course of the simulation shows that the interactions of the phenolic hydroxyls at both ends of the ligand differ substantially in character. The distal

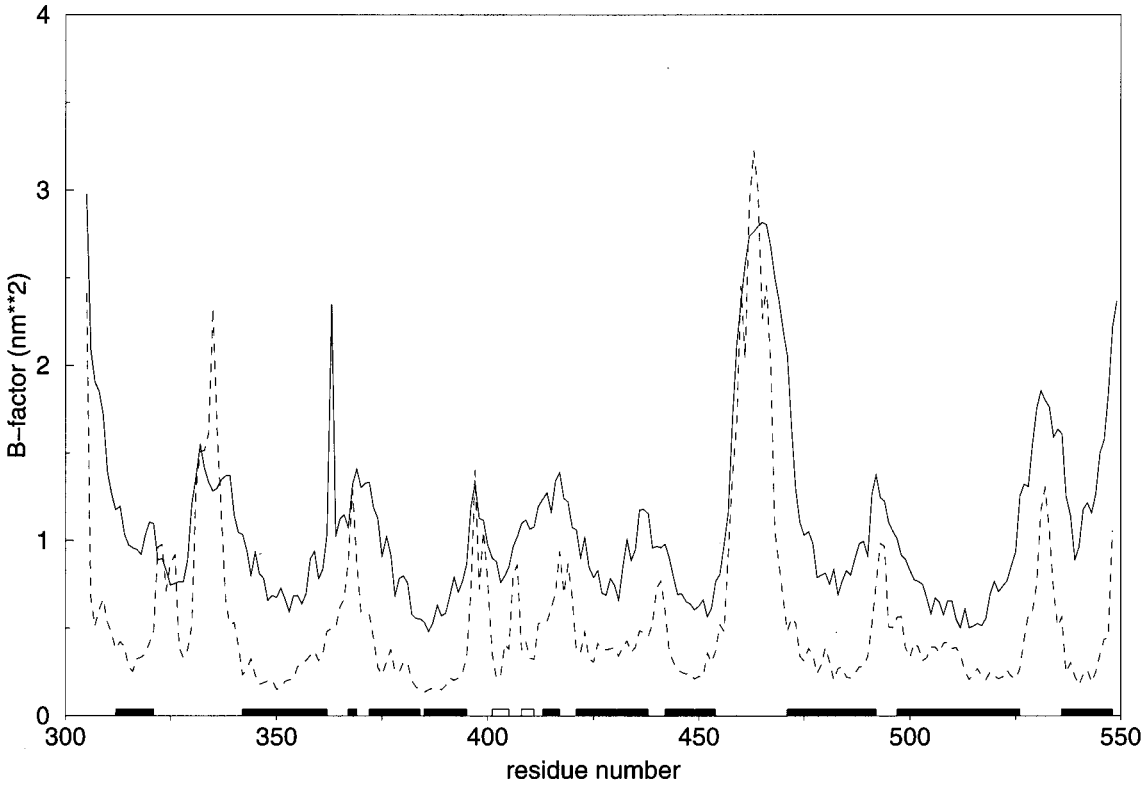


Figure 9. Experimental *B*-factors and corresponding root-mean-square atom-positional fluctuations for the DES/ER LBD X-ray crystal structure (solid line) and DES/ER LBD simulation (dashed line). The values have been averaged over the backbone (CA, N, C, O) heavy atoms of each residue. Positions of secondary structural elements are indicated along the *x*-axis as solid bars (helices) or open bars (sheets). Note that the first helix resolved in the structure is helix H2 in standard nomenclature.

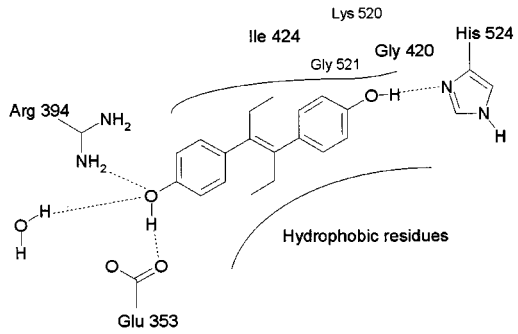


Figure 10. Schematic map of the ligand-binding pocket, drawn for the **2** (DES)/ER LBD complex. Residues and structural features mentioned in the text are indicated.

hydroxyl donates a single hydrogen bond to His 524 which is present during 97% of the simulation. In contrast, the proximal hydroxyl is part of a fluctuating hydrogen bond network that includes Glu 353, Arg 394, and two crystallographic water molecules. While no single hydrogen bond formed by the proximal hydroxyl is present for more than 60% of the trajectory, it is always participating in at least one hydrogen-bonding interaction and an average of 1.5 hydrogen-bonding interactions.

Thermodynamic Integration. Figures 11 and 12 display the value of $\langle \partial H/\partial \lambda \rangle_\lambda$ for the TI calculations as a function of λ. The integrated results can be found in Table 4. The first three columns contain the results for the DES (**2**) to E2 (**1**) conversion, while the second three columns show the DES (**2**) to GEN (**3**) conversion. The hysteresis reported is simply the difference of the forward and backward changes in λ and should only be used as a lower-bound estimate on the uncertainty of the calculated values. The relative free energies are calculated as the appropriate difference from the thermodynamic cycle in Figure 2.

During the conversion of **2** into **1**, His 524 rotates away from the ligand**,** resulting in the loss of a hydrogen bond. After roughly 35 ps of the E2 (**1**) end state simulation, this residue returns, forming a hydrogen bond with the 17β-hydroxyl of the ligand. For the **2** to **3** conversion the loss and re-formation of this hydrogen bond occurs within the duration of the TI simulations. Interestingly, the X-ray structure of the ER LBD/raloxifene complex shows a similar rotation of His 524 away from the ligand, suggesting that our simulation is sampling a realistic fluctuation of the protein binding pocket.[10]

Single-Step Perturbations. The results for the single-step perturbation to different compounds can be found in Table 5. The first column contains the free energy differences between the reference state and the different compounds in a vacuum, the second in the solvent, and third in the protein. The protein values were obtained as averages over five separate 200-ps blocks of the 1-ns simulation. In the fourth and fifth columns we present the calculated free energies of solvation and of binding, relative to the nonphysical reference (R) state. Finally, the available experimental binding free energies are given in the sixth column.

Figure 13 shows a graphical comparison of calculated and experimental free energies of binding, where it should be noted that only the differences between the compounds can be compared. In this figure, the dashed line indicates the experimental relative free energies

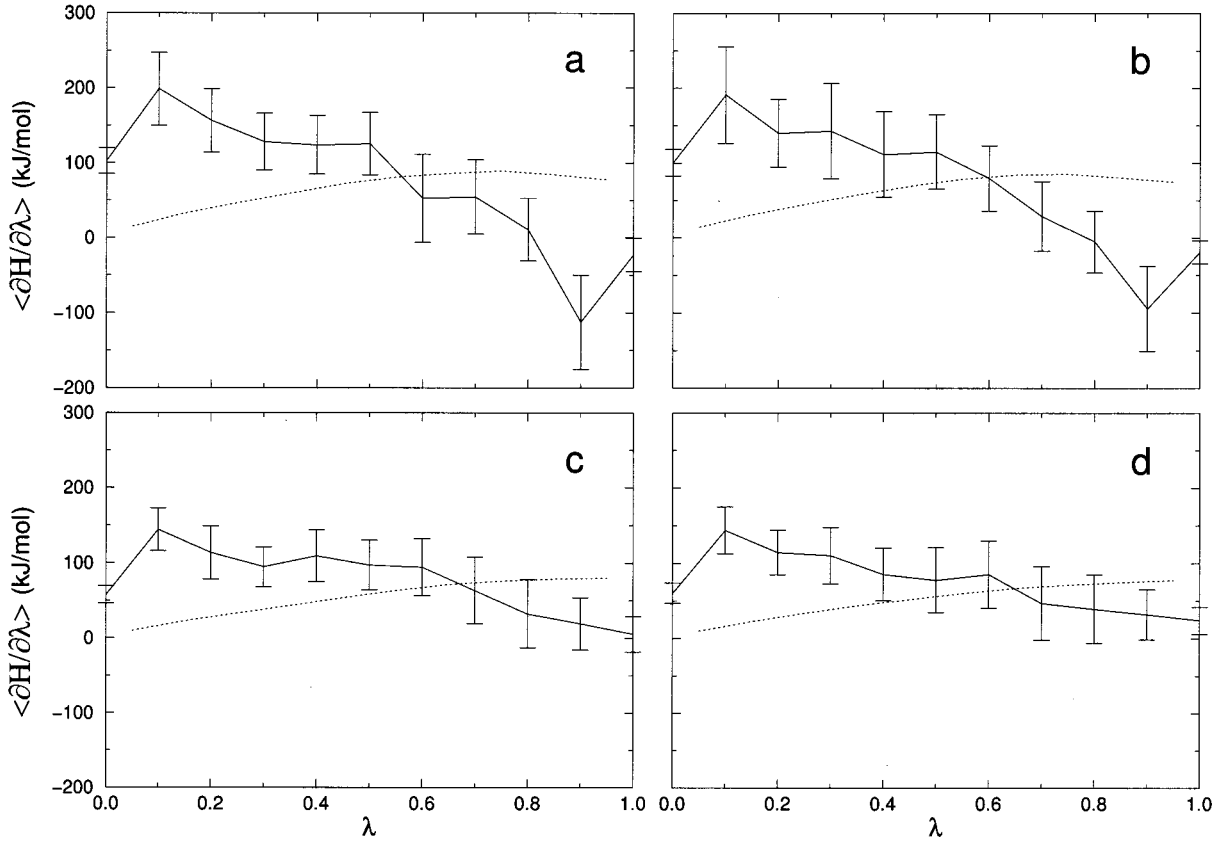


Figure 11. Free energy profiles of the DES (**2**) ↔ E2 (**1**) TI calculation. Each graph shows the value of the integrand $\langle \partial H/\partial \lambda \rangle_\lambda$ at each simulated λ-value. The error bars shown correspond to the standard deviation of $\partial H/\partial \lambda$ at each λ-value. The dashed line indicates the integrated value of the free energy as a function of λ. Separate graphs are shown for the forward and reverse changes of λ in water (a, b) and in the protein (c, d). All values are in kJ mol^{-1}.

relative to an estimated free energy of binding of −57 kJ mol^{-1} for the (nonphysical) reference state.

Discussion

Molecular Dynamics. The MD simulations of the DES/ER LBD, E2/ER LBD, and GEN/ER LBD complexes were all roughly similar. The protein structure was stable over the course of each trajectory, and visual inspection showed that the protein−ligand contacts remained largely unchanged. Detailed analysis was only performed on the DES/ER LBD simulation, as it was the most extensive (1 ns vs 500 ps).

The atom-positional rmsds from the starting structure show reasonable values (all atoms up to 0.275 nm, backbone up to 0.2 nm), particularly in light of the number of residues with side chains missing from the initial structure. The *B*-factors also show good agreement with experiment, particularly in the differentiation of the stable, well-packed helical regions from the more flexible loops between them.

Some dynamics are also seen in the DES itself. In particular, one of the two ethyl groups occupies a slightly larger hydrophobic pocket and shows significant torsional fluctuations. DES derivatives with a slightly larger substituent at this position might make a closer fit with the receptor and show increased affinity. Interestingly, the flexible ethyl group is facing away from helix 12. Raloxifene and other ER antagonists typically have a bulky substituent near the middle of the ligand but on the face that points toward helix 12.[10] It is thought that interactions between this substituent and helix 12 prevent the structural rearrangement seen in the E2/ER LBD complex, thereby blocking estrogenic activity.

Another interesting feature of the protein−ligand interaction observed in the simulation is the character of the hydrogen bond between one of the ligand hydroxyls and the imidazole of His 524. It is formed 97% of the time in our DES complex simulation, versus 40−60% for the other (proximal) hydroxyl interacting with Glu 353, Arg 394, or water molecules in the cavity. Interestingly, it is known from QSAR studies on estrogen-like steroids that the proximal hydroxyl is an absolute requirement for binding, while the distal 17β-hydroxyl (which interacts with His 524) only weakly contributes to binding and might serve merely to improve the solubility of the native hormone.[44] One feature of note is that the proximal hydroxyl acts intermittently both as a donor and as an acceptor, participating in 1.5 hydrogen bonds on average. In comparison, the distal hydroxyl acts only as a hydrogen bond donor. As His 524 could have an alternate protonation state or be doubly protonated, the possibility of it being a hydrogen bond donor cannot be completely excluded from our calculations. Still, it is clear that the distal hydroxyl probably only participates in a single hydrogen-bonding interaction. Thermodynamically, then, the single hydrogen bond formed by the distal hydroxyl might not be sufficient to compensate for the penalty of desolvating this polar substituent.

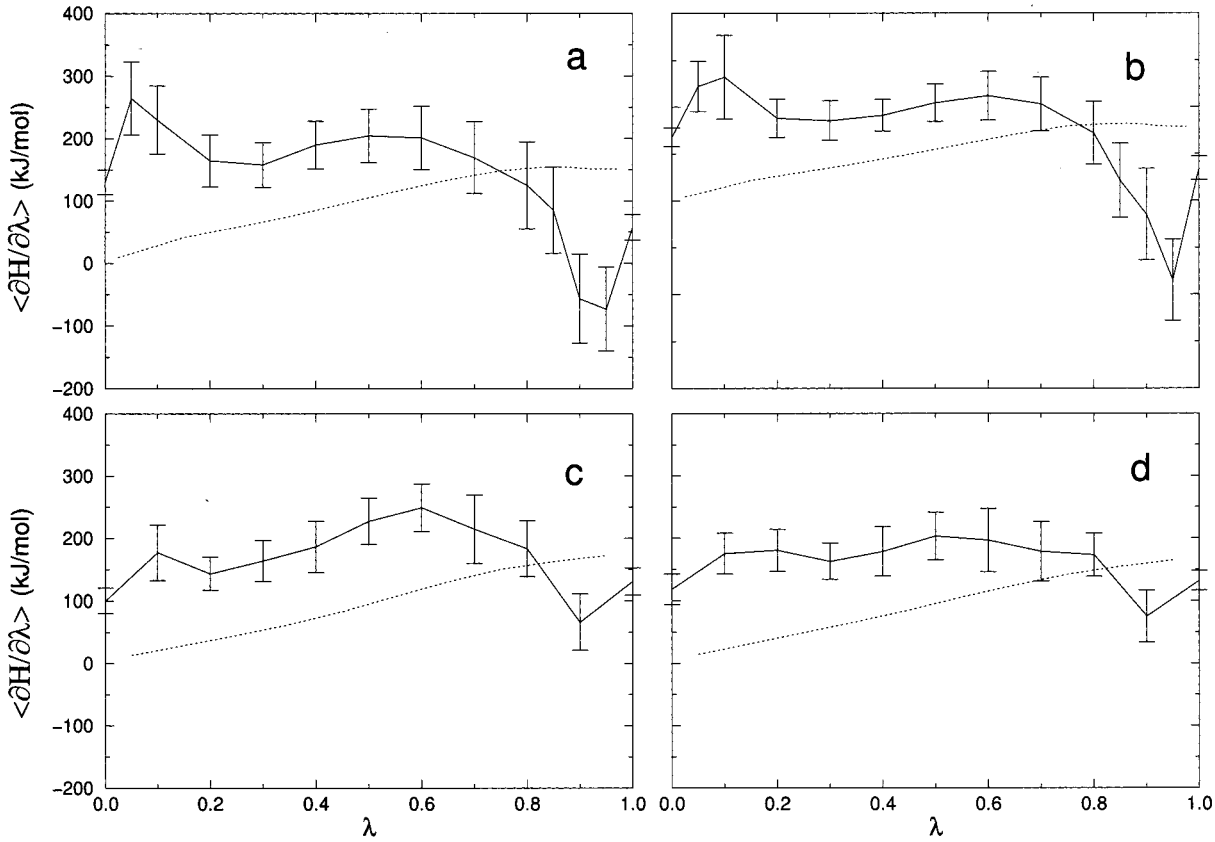


Figure 12. Free energy profiles of the DES (**2**) ↔ GEN (**3**) TI calculation. Each graph shows the value of the integrand $\langle \partial H/\partial \lambda \rangle_\lambda$ at each simulated λ-value. The error bars shown correspond to the standard deviation of $\partial H/\partial \lambda$ at each λ-value. The dashed line indicates the integrated value of the free energy as a function of λ. Separate graphs are shown for the forward and reverse changes of λ in water (a, b) and in the protein (c, d). All values are in kJ mol^{-1}.

Table 4. TI Results (kJ mol^{-1})[a]

	DES ↔ E2			DES ↔ GEN		
TI	forward	backward	hysteresis	forward	backward	hysteresis
vacuum	76.3	76.1	0.2	187.1	186.9	0.2
solvent	79.0	81.6	−2.6	151.5	157.3	−5.8
protein	80.4	78.2	2.2	173.1	165.3	7.8
$\Delta\Delta G_{solv}$	2.8	5.5	−2.7	−35.6	−29.5	−6.0
$\Delta\Delta G_{bind}$	1.4	−3.4	4.8	21.6	8.0	13.6
$\Delta\Delta G_{bind}$	3.8[b]			11.3[b]		
(expt)	0.79[c]			21.69[c]		

[a] The first three columns show the forward and reverse results with hysteresis for DES (**2**) ↔ E2 (**1**). The next three show the data for the DES (**2**) ↔ GEN (**3**) conversion. The reported hysteresis is simply the difference of the corresponding forward and reverse changes in λ. Experimental data are from [b]ref 25 and [c]van Lipzig, M. Unpublished results.

Thermodynamic Integration. On the basis of the course of $\langle \partial H/\partial \lambda \rangle_\lambda$ in Figure 12, three additional λ-values were introduced in the conversion of DES (**2**) into GEN (**3**) in the solvent at $\lambda = 0.05$, 0.85, and 0.95, respectively. From the smooth course of the free energy as a function of λ, it may be concluded that there is no apparent need for additional λ-values in the other calculations. As a test for convergence, an additional 10 ps was simulated for all λ-values in the solvent conversion of DES (**2**) into E2 (**1**), but no significant changes were observed. The hysteresis for individual runs is less than 5% of the calculated value but accumulates in the calculation of relative free energies.

Table 5. Single-Step Perturbation Results (kJ mol^{-1})[a]

compd	ΔG-(vac) (1 ns)	ΔG-(solv) (1 ns)	ΔG-(prot) (1 ns)	$\Delta\Delta G_{sol}$	$\Delta\Delta G_{bind}$	ΔG_{bind} (expt)
5, CHN	6.1	22.6	32.2	16.5	9.6	
6, MCHN	4.7	16.4	22.1	11.7	5.7	
7, BAP	2.1	7.7	12.8	5.6	5.1	
8, MBAP	11.0	25.9	44.7	14.9	18.9	
9, BAA	7.1	23.4	35.0	16.3	11.6	−42.9[b]
10, 7MBA	5.2	16.8	23.7	11.6	6.9	−53.2[b]
11, 12MBA	11.7	26.0	41.3	14.3	15.4	−43.8[b]
12, DMBA	10.0	20.6	30.2	10.6	9.6	−47.3[b]
3, GEN (1)[d]	130.0	162.7	183.0	32.7	20.3	−36.0[c]
3, GEN (2)[d]	145.6	174.8	194.1	29.2	19.3	

[a] Columns 1−3 show the free energy of each ligand relative to the reference state R in vacuum, solution, and the protein, respectively. The relative solvation free energy ($\Delta\Delta G_{sol}$; column 4) is the difference of columns 2 (solution) and 1 (vacuum). The relative binding free energy ($\Delta\Delta G_{bind}$; column 5) is the difference of columns 3 (protein) and 2 (solution). As all values are free energies relative to the nonphysical reference state R, they should not be compared to the experimental data directly. Instead, the relative values between different ligands should be compared. Experimental binding free energies (column 6) are from [b]Morreal, C. E.; Sinha, D. K.; Schneider, S. L.; Bronstein, R. E.; Dwidzik, J. Anitestrogenic Properties of Substituted Benz[*a*]anthracene-3,9-diols. *J. Med. Chem.* **1982**, *25*, 323−326 and [c]van Lipzig, M. Unpublished results. [d] GROMOS96 bond type between aromatic ring systems 15 (0.139 nm; 1) and 26 (0.153 nm; 2), respectively.

If we average over forward and backward changes of λ in the TI calculations, we find a binding free energy difference between **1** and **2** of −1.0 kJ mol^{-1}, whereas we expect a small positive value based on experimental

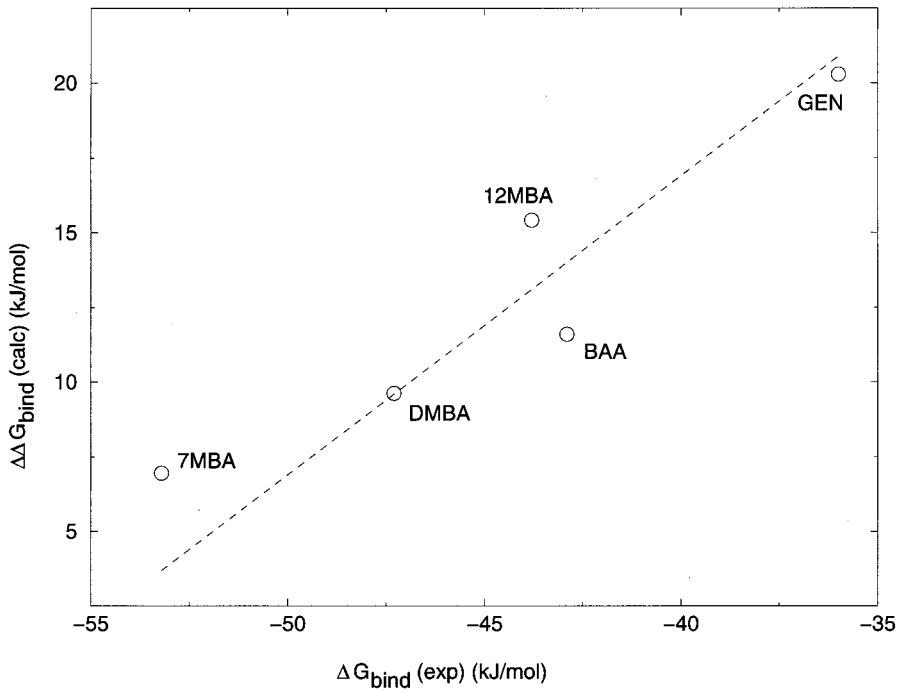


Figure 13. Graphical comparison of experimental and single-step perturbation results for the relative binding free energy of ligands **3** (GEN), **9** (BAA), **10** (7MBA), **11** (12MBA), and **12** (DMBA) to the ER LBD. As the calculated binding free energies (*y*-axis) are relative to the nonphysical reference state R ($\Delta\Delta G_{bind}$) while the experimental values (*x*-axis) are absolute binding free energies (ΔG_{bind}), their values should not be compared directly. Instead, the relative values between different ligands should be compared. The dashed line is the best-fit line of slope 1.0 and provides a guide to the deviation between the experimental and calculated values. All values are in kJ mol^{-1}.

assays. It can be argued, though, that due to the higher symmetry of **2**, a correction is necessary. This symmetry allows the molecule to bind in a number of degenerate binding modes, of which our calculation takes into account only one. This number of different binding modes, n_{bind}, introduces a correction to the free energy of $-k_BT$ ln n_{bind}. Whether two binding modes are degenerate depends on the flexibility of the compound complexed to the protein, a value of n_{bind} of 2 or 4 introduces a correction of 1.7 or 3.5 kJ mol^{-1} in favor of **2**, respectively. For the conversion of **2** into **3** this correction would be required as well, increasing the average value of 14.8 kJ mol^{-1} toward the experimental value of 21.3 kJ mol^{-1}.

As noted, the hydrogen bond between the distal hydroxyl of **2** and His 524 disappears in the conversion to **1** but returns after some simulation time in the E2 (**1**)/ER LBD complex. This may be due to the fact that the distance between hydroxyl groups in **1** is shorter (1.1 nm) than in **2** (1.2 nm), which is comparable to the corresponding distance in **3** (1.2 nm). The fluctuations of His 524 are probably not on the same time scale as the 220 ps in which the conversion has taken place. Another explanation for the loss of this hydrogen bond at intermediate λ-values in both simulations may be the use of the soft-core nonbonded interaction. While the soft-core scaling removes singularities from the nonbonded potential energy terms that have the possibility to cause instabilities in the simulation, it also causes a loss of some short-range structure in these potential energy terms − precisely the region crucial for descriptions of hydrogen bonds.

Single-Step Perturbation. The first and second nanoseconds of the reference state simulation in water showed no significantly different values for the single-step perturbations. As noted, the 1-ns simulation in the protein was divided into five blocks of 200 ps each, and the intrablock results were compared to assess the convergence of the data. The variations between the five different protein runs were larger than the values between the two blocks of the solvent calculation, signifying that in a time span of 200 ps the simulation has not sampled conformational space completely.

The calculated relative free energies of solvation do not correspond with the expectation that GEN should yield a more favorable solvation free energy due to the additional polar groups. Since the soft-core reference state for which the simulations have been carried out is a neutral compound, however, the resulting solvent configurations are probably not particularly favorable for the polar genistein. This issue has been previously noted in other single-step perturbation calculations of solvation free energies that include both neutral and polar species.[35]

It can be argued that, due to the forced coplanar conformation of the genistein in these calculations, an aromatic bond type between the two ring systems (e.g. GROMOS96 type 15) would be more appropriate than an aliphatic bond type (GROMOS96 type 26) as was used in the TI calculations, where the two ring systems were free to rotate. In testing both possibilities, we found that they differ in binding free energy by only 1 kJ mol^{-1}. We based our analysis on the aromatic bond type in view of the coplanar geometry of the reference state.

The comparison of the calculated free energies of binding relative to the reference state to the experimental absolute free energies in Figure 13 shows that the differences between compounds are generally reproduced quite well. On the basis of these figures, one

can estimate the free energy of binding for the reference state to be about −57 kJ mol^{-1}, which would result in a maximum deviation from experimental values of 3.3 kJ mol^{-1} (<1 kcal mol^{-1}) for compound **10** and a mean difference over all five known compounds of 1.7 kJ mol^{-1}.

It is interesting to note that the addition of a methyl to structures **5**, **9**, and **11** to obtain structures **6**, **10**, and **12** is calculated to improve binding free energy by about 4 kJ mol^{-1}, while introducing a methyl group at the same location in structure **7**, yielding **8**, reduces affinity by 13 kJ mol^{-1}. Obviously, the binding cavity offers enough space for the former structures to sample an orientation that allows the methyl to occupy a more favorable position, whereas **8**, the largest compound considered, does not seem to have this freedom of movement.

It is important to emphasize the relative computing time required for the single-step perturbations versus the TI calculations. In a total of 880 ps of simulation, the TI calculations yielded relative free energies for three compounds. In contrast, our 1 ns of reference state simulation allowed us to calculate relative free energies for 10 different compounds. In fact, this set can be extended to a total of 20−30 different species.[45] The single-step perturbation is therefore somewhere between 4 and 6 times less expensive than the TI calculations.

Conclusions

Overall, our foray into the calculation and prediction of the affinities of various ligands for the ER has yielded some exceptionally promising results. We have shown good agreement between experimental results and traditional TI free energy calculations for comparisons between structurally diverse known ER agonists (17β-estradiol (E2), **1**; diethylstilbestrol (DES), **2**; and genistein (GEN), **3**). In addition, we have applied a single-step perturbation methodology to this system which allows us to accurately predict the relative free energies of a large number of xenoestrogen-like compounds from a single simulation. This is particularly important since the large size of the ER LBD/ligand complexes make any sort of MD simulation exceedingly expensive. Given the amazing variety of compounds proposed to bind to the ER, methods such as we have applied may provide a powerful tool for assessing the affinity of putative xenoestrogens in silico prior to subsequent in vivo studies.

Acknowledgment. Financial support for this research was provided by Dr. Hendrik Muller's Vaderlandsch Fonds, Bekker La Bastide Fonds, and Dittmer Fonds.

Supporting Information Available: Molecular topologies and GROMOS96 interaction function parameters for compounds **1**−**12** and the soft reference state R. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Kuiper, G. G. J. M.; Enmark, E.; Pelto-Huikko, M.; Nilsson, S.; Gustafsson, J.Å. Cloning of a novel estrogen receptor expressed in rat prostate and ovary. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5925−5930.

(2) Barkhem, T.; Carlsson, B.; Nilsson, Y.; Enmark, E., Gustafsson, J.Å.; Nilsson, S. Differential Response of Estrogen Receptor α and Estrogen Receptor β to Partial Estrogen Agonists/Antagonists. *Mol. Pharmacol.* **1998**, *54*, 105−112.

(3) Gustafsson, J. Å.; Kuiper, G.; Enmark, E.; Treuter, E.; Rafter, J. Receptor-Mediated Toxicity. *Arch. Toxicol.* **1998**, *20*, 21−28.

(4) Tsai, M.-J.; O'Malley, B. W. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu. Rev. Biochem.* **1994**, *63*, 690−697.

(5) Wells, J. A. Structural and functional basis for hormone-binding and receptor oligomerization. *Curr. Opin. Cell. Biol.* **1994**, *6*, 163−173.

(6) Wells, J. A. Binding in the growth hormone receptor complex. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 1−6.

(7) Parker, M. G.; Arbuckle, N.; Dauvois, S.; Danelian, P.; White, R. Structure and Function of the Estrogen Receptor. *Ann. N. Y. Acad. Sci.* **1993**, *11*, 119−125.

(8) Wagner, R. L.; Apriletti, J. W.; McGrath, M. E.; West, B. L.; Baxter, J. D.; Fletterick, R. J. A structural role for hormone in the thyroid-hormone receptor. *Nature* **1995**, *378*, 690−697.

(9) Renaud, J.; Rochel, N.; Ruff, M.; Vivat, V.; Chambon, P.; Gronemeyer, H. Crystal-Structure of the RAR−Gamma Ligand-Binding Domain bound to all-*trans*-retinoic acid. *Nature* **1995**, *378*, 681−689.

(10) Brzozowski, A. M.; Pike, A. C. W.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engström, O.; Öhman, L.; Greene, G. L.; Gustafsson, J. Å.; Carlquist, M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **1997**, *389*, 753−758.

(11) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. *Cell* **1998**, *95*, 927−937.

(12) Pike, A. C. W.; Brzozowski, A. M.; Hubbard, R. E.; Bonn, T.; Thorsell, A. G.; Gustafsson, J. Å.; Carlquist, M. Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. *EMBO J.* **1999**, *18*, 4608−4618.

(13) Webster, N. J. G.; Green, S.; Tassel, D., Ponglikitmongkol, M.; Chombon, P. The transcriptional activation function located in the hormone-binding domain of the human oestrogen receptor is not encoded in a single exon. *EMBO J.* **1989**, *8*, 1441−1446.

(14) Danielan, P. S.; White, R.; Lees, J. A.; Parker, M. G. Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. *EMBO J.* **1992**, 11, 1025−1033.

(15) Waller, C. L.; Oprea, T. I.; Chau, K.; Park, H. K.; Korach, K. S.; Laws, S. C.; Wiese, T. E.; Kelce, W. R.; Earl Gray Jr., L. Ligand-Based Identification of Environmental Estrogens. *Chem. Res. Toxicol.* **1996**, *9*, 1240−1248.

(16) Soto, A. M.; Sonnenschein, C.; Chung, K. L.; Fernandez, M. F.; Oka, N.; Serrano, F. O. The E-screen assay as a tool to identify estrogens − an update on estrogenic environmental-pollutants. *Environ. Health Perspect.* **1995**, *103*, 113−122.

(17) Breinholt, V.; Larsen, J. C. Detection of Weak Estrogenic Flavonoids Using a Recombinant Yeast Strain and a Modified MCF7 Cell Proliferation Assay. *Chem. Res. Toxicol.* **1998**, *11*, 622−629.

(18) Garner, C. E.; Jefferson, W. N.; Burka, L. T.; Matthews, H. B.; Newbold, R. R. In vitro estrogenicity of the catechol metabolites of selected polychlorinated biphenyls. *Toxicol. Appl. Pharmacol.* **1999**, *154*, 188−197.

(19) Blair, R. M.; Fang, H.; Branham, W. S.; Hass, B. S.; Dial, S. L.; Moland, C. L.; Tong, W.; Shi, L. Perkins, R.; Sheehan, D. M. The Estrogen Receptor Relative Binding Affinities of 188 Natural and Xenochemicals: Structural Diversity of Ligands. *Toxicol. Sci.* **2000**, *54*, 138−153.

(20) Anstead, G. M.; Kym, P. R. Benz[*a*]anthracene diols: Predicted carcinogenicity and structure-estrogen receptor binding affinity relationships. *Steroids* **1995**, *60*, 383−394.

(21) Charles, G. D.; Bartels, M. J.; Zacharewski, T. R.; Gollapudi, B. B.; Freshour, N. L.; Carney, E. W. Activity of Benzo[*a*]pyrene and Its Hydroxylated Metabolites in an Estrogen Receptor-α Reporter Gene Assay. *Toxicol. Sci.* **2000**, *55*, 320−326.

(22) Adlercreutz, H.; Mazur, W. Phyto-estrogens and Western Diseases. *Ann. Med.* **1997**, *29*, 95−120.

(23) Gantchev, T. G.; Ali, H.; Van Lier, J. E. Quantitative Structure−Activity Relationships/Comparative Molecular Field Analysis (QSAR/CoMFA) for Receptor-Binding Properties of Halogenated Estradiol Derivatives. *J. Med. Chem.* **1994**, *37*, 4164−4176.

(24) Bradbury, S. P.; Mekenyan, O. G.; Ankley, G. T. Quantitative Structure−Activity Relationships for Polychlorinated Hydroxybiphenyl Estrogen Receptor Binding Affinity: an assessment of Conformer Flexibility. *Environ. Toxicol. Chem.* **1996**, *15*, 1945−1954.

(25) Tong, W.; Perkins, R.; Xing, L.; Welsh, W. J.; Sheehan, D. M. QSAR Models for Binding of Estrogenic Compounds to Estrogen Receptor α and β subtypes. *Endocrinology* **1997**, *138*, 4022−4025.

(26) Beveridge, D. L.; DiCapua, F. M. Free Energy via Molecular Simulation: Applications to Chemical and Biomolecular Systems. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431−492.
(27) Pitera, J. and Kollman, P. A. Designing an Optimum Guest for a Host using Multi-Molecule Free Energy Calculations: predicting the best ligand for Rebek's "Tennis Ball". *J. Am. Chem. Soc.* **1998**, *120*, 7557−7567.
(28) Straatsma, T. P.; Berendsen, H. J. C.;. Postma, J. P. M. Free energy of hydrophobic hydration: A molecular dynamics study of noble gases in water. *J. Chem. Phys.* **1986,** *85*, 6720−6727.
(29) Liu, H.; Mark, A. E.; van Gunsteren, W. F. Estimating the Relative Free Energy of Different Molecular States with Respect to a Single Reference State. *J. Phys. Chem.* **1996**, *100*, 9485−9494.
(30) Mark, A. E.; Schäfer, H.; Liu, H.; van Gunsteren, W. F. Estimating Relative Free Energies from a Single Simulation of the Initial State. In *Proceedings of the 2nd International Symposium on Algorithms for Macromolecular Models*; Deuflhard, P., Hermans, J., Leimkuhler, B., Mark, A. E., Reich, S., Skeel, R. D., Eds.; Springer-Verlag: Berlin, 1999; pp 149−162.
(31) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, 1996.
(32) Pearlman, D. A. A comparison of alternative approaches to free-energy calculations. *J. Phys. Chem.* **1994**, *98*, 1487−1493.
(33) Pitera, J.; Oostenbrink, C.; van Gunsteren, W. F. Unpublished results.
(34) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(35) Schäfer, H.; van Gunsteren, W. F.; Mark, A. E. Estimating relative free energies from a single ensemble: Hydration free energies. *J. Comput. Chem.* **1999**, *20*, 1604−1617.
(36) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 331−342.
(37) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(38) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(39) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(40) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(41) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(42) Smith, P. E.; van Gunsteren, W. F. Consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300 K. *J. Chem. Phys.* **1994**, 199, 3169−3174.
(43) Mark, A. E.; Xu, Y.; Liu, H.; van Gunsteren, W. F. Rapid nonempirical approaches for estimating relative binding free energies. *Acta Biochim. Pol.* **1995**, *42*, 525−536.
(44) Anstead, G. M.; Carlson, K. E., Katzenellenbogen, J. A. The estradiol pharmacophore: Ligand structure-estrogen receptor binding affinity relationships and a model for the receptor binding site. *Steroids* **1997**, *62*, 268−303.
(45) Oostenbrink, C. Unpublished results.

JM001045D