Free energies of ligand binding for structurally diverse compounds

Chris Oostenbrink and Wilfred F. van Gunsteren†

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, Eidgenössische Technische Hochschule–Hönggerberg, CH-8093 Zurich, Switzerland

Edited by Bruce J. Berne, Columbia University, New York, NY, and approved January 14, 2005 (received for review October 6, 2004)

The one-step perturbation approach is an efficient means to calculate many relative free energies from a common reference compound. Combining lessons learned in previous studies, an application of the method is presented that allows for the calculation of relative binding free energies for structurally rather diverse compounds from only a few simulations. Based on the well known statistical–mechanical perturbation formula, the results do not require any empirical parameters, or training sets, only limited knowledge of the binding characteristics of the ligands suffices to design appropriate reference compounds. Depending on the choice of reference compound, relative free energies of binding rigid ligands to the ligand-binding domain of the estrogen receptor can be obtained that show good agreement with the experimental values. The approach presented here can easily be applied to many rigid ligands, and it should be relatively easy to extend the method to account for ligand flexibility. The free-energy calculations can be straightforwardly parallelized, allowing for an efficient means to understand and predict relative binding free energies.

free-energy calculation | molecular dynamics simulation | estrogen receptor | one-step perturbation

The calculation of relative binding free energies for several ligands to a common receptor is of relevance for drug design purposes, and for obtaining a better understanding of the molecular interactions of proteins with small compounds. Despite the increased availability of computational power, calculating free energies from molecular dynamics simulations is still time-consuming, often requiring several extensive simulations to obtain a single relative free energy. The approach in which several free energies can be obtained from a single simulation of a, not necessarily physically meaningful, reference state (1) has been applied successfully in recent years (2–6). The idea behind the method is to simulate a judiciously chosen reference compound R, generating an ensemble of structures that contains conformations representative for several physically relevant compounds. The free-energy difference between any real ligand A and the reference compound can be calculated from the perturbation formula (7)

$$\Delta G_{\mathrm{AR}} = G_{\mathrm{A}} - G_{\mathrm{R}} = -k_{\mathrm{B}}T \ln\langle e^{-(H_{\mathrm{A}}-H_{\mathrm{R}})/k_{\mathrm{B}}T}\rangle_{\mathrm{R}}, \qquad [1]$$

where the angle brackets indicate the ensemble average over the structures generated in a simulation of R. H_{A} and H_{R} are the Hamiltonians for the real compound (A) and the reference compound (R), respectively. Because this expression only involves the difference between the two Hamiltonians, only interactions that differ between compounds A and R need to be reevaluated over the ensemble. k_{B} is the Boltzmann constant, and T is the temperature. As was demonstrated before, one can split up the process of changing the reference compound into a real ligand into insertion of a real ligand and the removal of the reference compound (6). Assuming that the removal of the reference compound is independent of the real ligand that was added, one can estimate the free-energy difference as

$$\Delta G'_{\mathrm{AR}} = G^{*}_{\mathrm{R}} - k_{\mathrm{B}}T \ln\langle e^{-H'_{\mathrm{A}}/k_{\mathrm{B}}T}\rangle_{\mathrm{R}}, \qquad [2]$$

with G^{*}_{R} an offset in the free energy that is independent of the real ligand, A, and H'_{A} involving only the interactions of ligand A with itself and its surroundings, excluding R (6). The relative free energy of binding for two different compounds, B and A, can then be obtained from a thermodynamic cycle (8–10) by taking the difference of the relative free energies of changing the reference compound into compounds B and A when bound to the protein (*complex*) and when free in aqueous solution (*free*) (6):

$$\begin{aligned}\Delta\Delta G^{\mathrm{bind}}_{\mathrm{BA}} &= \Delta G^{\mathrm{bind}}_{\mathrm{B}} - \Delta G^{\mathrm{bind}}_{\mathrm{A}} \\ &= \Delta G_{\mathrm{BA}}(complex) - \Delta G_{\mathrm{BA}}(free) \\ &= [\Delta G'_{\mathrm{BR2}}(complex) - \Delta G'_{\mathrm{AR2}}(complex)] \\ &\quad - [\Delta G'_{\mathrm{BR1}}(free) - \Delta G'_{\mathrm{AR1}}(free)] \\ &= [\Delta G'_{\mathrm{BR2}}(complex) - \Delta G'_{\mathrm{BR1}}(free)] \\ &\quad - [\Delta G'_{\mathrm{AR2}}(complex) - \Delta G'_{\mathrm{AR1}}(free)]. \qquad [3]\end{aligned}$$

Fig. 1 also shows that the reference compounds R1 and R2 in the different surroundings need not be the same.

So far, the one-step perturbation approach was applied to series of compounds with some common structural features that form the basis for the reference compound. The enhanced sampling in the simulation of the reference compound was obtained by describing the non-bonded interaction of selected atoms in the reference compound by a soft interaction (11). The compounds for which the free energy can then be calculated by using expression **1** or **2** were thus far limited by the requirement that their structure should be more or less identical to the structural framework of the reference compound. In the present study, we explore the possibilities to minimize the structural requirements on the real ligands imposed by the structural framework of the reference compound. It will be shown that reasonable relative free energies can be obtained for structurally different compounds without the use of any empirical knowledge. The system of choice is the ligand-binding domain (LBD) of the estrogen receptor (ER). Previous computational approaches to computing binding free energies for this system involve empirical regression-based methods (12–14), docking efforts (15), and free-energy calculations from molecular dynamics simulations (3, 6, 16). A great number of structurally diverse compounds have been shown to interact with the ER LBD (14, 17–24). From structure–activity relationships, the pharmacophore for ligand binding can be described by two hydroxyl groups at a distance of ≈1.2 nm, of which one is bound to an aromatic ring (25). The molecular basis for this pharmacophore was confirmed by three different crystal structures of

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: COU, coumestrol; DES, diethylstilbestrol; ER, estrogen receptor; GEN, genistein; LBD, ligand-binding domain; MBA, 7-methyl-benz[*a*]anthracene-3,9-diol.

†To whom correspondence should be addressed. E-mail: wfvgn@igc.phys.chem.ethz.ch.

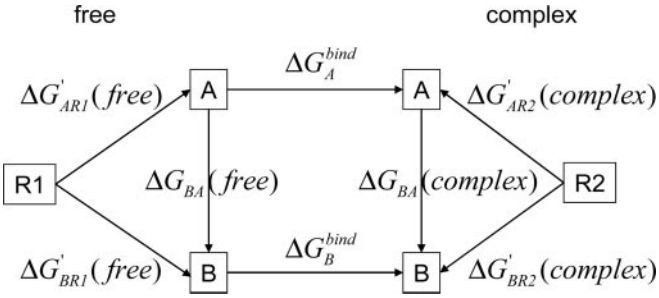

Fig. 1. Thermodynamic cycle to calculate relative free energies of binding for real ligands A and B from simulations of soft reference compounds R1 and R2 in water (*free*) and when bound to the protein (*complex*). The relative free energy of binding, $\Delta\Delta G^{bind}_{BA}$, can be calculated by using formula **3**.

the LBD that have become available (26–28). While the hydroxyl group attached to the aromatic ring was seen to be involved in hydrogen bonds with residues Glu-353 and Arg-394 and an enclosed water molecule, the other hydroxyl group forms a hydrogen bond with residue His-524. In the present study, we limit ourselves to the five ligands in Fig. 2: 17β-estradiol (E2) is the endogenous ligand responsible for the growth, development, and maintenance of a diverse range of tissues; diethylstilbestrol (DES) is a synthetic estrogen; 7-methyl-benz[*a*]anthracene-3,9-diol (MBA) is a possible product of the oxidative biotransformation of a common polyaromatic hydrocarbon (22); and genistein (GEN) and coumestrol (COU) are phytoestrogens that can be found in soy products (GEN) or sprouts (COU) (29). Except for DES, these are rather rigid but structurally distinct compounds.

The ER LBD is characterized by a cavity that is completely closed of from the solvent by helix number 12 (see Fig. 3). This structural feature allowed us to design the three different reference states that are shown in Fig. 4. They were obtained by filling the binding site of the ligand in the protein with a perfect carbon diamond network and the addition of two hydroxyl groups to account for the most important ligand–protein hydrogen bonds. Most of the carbon atoms in these reference compounds are soft atoms, ensuring that the protein does sample a broad conformational ensemble at the reference compound–protein interface but preventing the protein from collapsing into an empty cavity. The reference compounds should be large enough to encompass the entire real ligands, not too coarse to accurately represent the shape of the cavity, and rigid enough to maintain this shape throughout the simulation. As was found earlier when calculating solvation free energies by the one-step

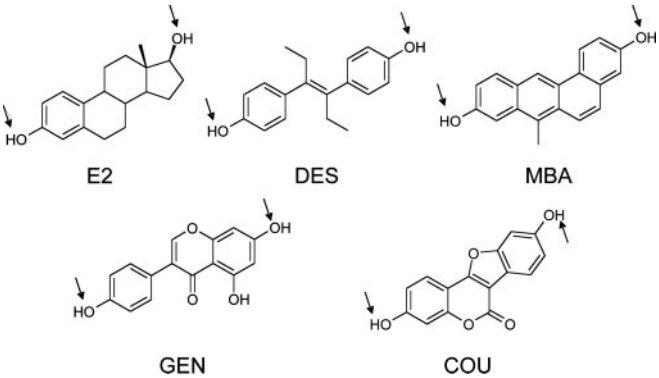

Fig. 2. Ligands known to bind to the LBD of the ER. Arrows indicate the hydroxyl groups that were used for fitting the structures to the reference compounds in Fig. 4.

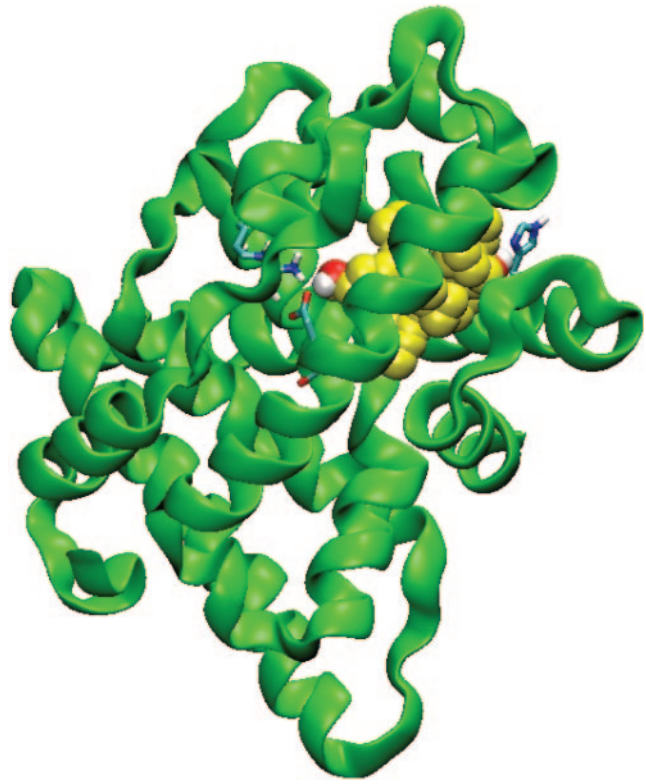

Fig. 3. Ligand-binding domain of the ER in complex with the unphysical reference compound SDM. The reference compound is in yellow, and the protein backbone is in green. Residues Glu-343, Arg-394, and His-524 are represented in cyan.

perturbation approach (4), additional sampling of orientational and translational degrees of freedom of the real ligands improves the convergence of the results significantly. Here, we apply the additional sampling for the first time to larger compounds complexed to a protein. Insertion of the real ligands into conformations from protein-reference compound simulations can also be considered as a ligand docking technique that fully includes the protein conformational flexibility. The calculation of relative binding free energies through formula **2** and an appropriate thermodynamic cycle (Fig. 1 and formula **3**) provides the obtained results with a statistical–dynamical basis, in contrast to docking based on empirical scoring functions.

In the presented approach, the structure of the reference state no longer limits the choice of the real ligands that can be taken into account. The free energy of structurally rather different ligands can be evaluated from simulations of the described reference compounds. In the current implementation, full flex-

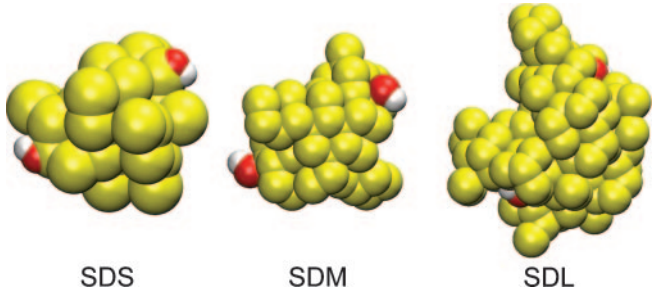

Fig. 4. Unphysical reference compounds based on a perfect carbon diamond structure, designed to fill the LBD of the ER. See Table 1 for details.

Table 1. Properties of reference compounds

Property	SDS	SDM	SDL
Number of atoms	35	67	195
σ, nm	0.4832	0.3361	0.3361
ε, kJ/mol	0.400	0.406	0.406
C–C bond length, nm	0.220	0.153	0.153
r_{min}, nm	0.23	0.23	0.16
% SASA ligand	9.7	21.5	1.9
% hydrogen bonds to			
Water	1	66	44
Glu-353	96	10	52
Arg-394	39	2	0
His-524	98	98	17

Characterization of the three reference compounds shown in Fig. 4. For every compound the number of atoms (including those of the hydroxyl groups), the van der Waals parameters [$\sigma = (C12/C6)^{1/6}$ and $\varepsilon = (C6)^2/(4C12)$] for the carbon atoms, the C–C bond length of the diamond structure, and the minimum distance r_{min} from any carbon atom to any protein atom in the initial configuration is given. The percentage of ligand solvent-accessible surface area (SASA) in an overlay of the real ligands of Fig. 2 with the reference compounds of Fig. 4 is also given, as well as the occurrence of hydrogen bonds involving the two hydroxyl groups of the reference compounds and Glu-353, Arg-394, or a water molecule on the one end or His-524 on the other end of the reference compound. A hydrogen bond is defined by a donor–acceptor distance shorter than 0.25 nm and a donor– hydrogen–acceptor angle >135°.

ibility of the protein degrees of freedom, as well as of the orientational and positional degrees of freedom of rigid ligands within a cavity in the protein or in the solvent, is taken into account. Future research will be aimed at including the intramolecular ligand degrees of freedom as well.

Methods

All simulations were performed by using the GROMOS simulation software package (30, 31). Protein coordinates were taken from a well equilibrated previous simulation of the LBD of the α-subtype of the human ER (3). The soft reference structures for the ligand, called SDS, SDM, and SDL, were created by overlapping a perfect diamond structure built with a particular C–C distance (0.22 nm for SDS, and 0.153 nm for SDM and SDL) with the protein structure and removing all carbon atoms of the diamond structure within a distance r_{min} (0.23 nm for SDS and SDM, and 0.16 nm for SDL) of any protein atom. Two hydroxyl groups from the original ligand (DES) were attached to the closest carbon atom of the soft reference state, resulting in the three reference structures of Fig. 4. Table 1 summarizes the parameters associated with the reference compounds. In Fig. 3, the soft reference compound SDM is displayed when complexed to the ER LBD.

Interaction parameters were taken from the 43A1 GROMOS parameter set (30, 32). Van der Waals parameters for the soft reference states are listed in Table 1. The hydroxyl groups and connected carbon atoms were interacting through Coulomb and a normal van der Waals interaction; all other atoms in the soft reference compounds were neutral soft atoms (11), with a softness parameter value (2), $\alpha_{LJ} = 1.51$. Bonded parameters for the reference compounds involve only C–C bond lengths and standard GROMOS hydroxyl bond lengths and angles. Interaction parameters for the protein and for the real compounds in Fig. 2 are described in ref. 3 (E2, DES, and GEN) or defined by analogy (MBA and COU). Molecular dynamics simulations of the three reference compounds in water as well as in complex with the LBD of the ER in explicit SPC water (33) were performed for 1 ns by using the simulation parameters and protocol described in ref. 3. While the soft reference compounds in the protein maintained their overall structure, they soon became more globular in water. An additional simulation of reference compound SDM in water was performed in which the solute atoms were positionally restrained by using a harmonic potential energy term with force constant 2.5×10^4 kJ·mol^{-1}·nm^{-2}.

Atom coordinates of the reference compounds, when unbound and when bound to the ER LBD, were saved every 0.1 ps from the corresponding simulations. Rigid structures of the real ligands were superimposed onto the reference compound by fitting the hydroxyl atoms indicated in Fig. 2 to the hydroxyl atoms in the soft reference compounds for every of these 10^4 structures. To obtain additional sampling (4), the real ligands were then 10 times rotated around the O–O axis over a random angle obtained from a uniform distribution over the range 0–360°. Subsequently, 10 random displacements of the ligands were performed taken from a uniform distribution over the range −0.05–0.05 nm in all three dimensions. The additional sampling leads to a total of $10 \times 10 \times 10^4 = 10^6$ conformations for a real ligand complexed to the ER LBD or in water (*free*). The interaction energy of the ligand with its surroundings was calculated for every one of these conformations and used in expression **2** to obtain an estimate of the free energy of complex formation. The relative free energy of binding was then obtained by applying formula **3**.

Results and Discussion

In water, reference compounds SDS and SDM soon changed shape toward a more globular form, in which the O–O distance was reduced from 1.18 nm to 0.90 nm. The cavity that was formed by these reference ligands in water no longer had the proper size and shape to fit the real ligands. For this reason, another simulation (SDMres) for reference compound SDM was performed in which its atoms were positionally restrained. In the simulation of SDL in water, the O–O distance was slightly shortened as well but remained at an average length of 1.08 nm. The degree by which the real ligands fit onto the reference compounds is analyzed by calculating the solvent-accessible surface area (SASA) of a structure that was obtained by fitting the ligands onto the reference compound, based on the hydroxyl groups. Table 1 gives the percentage of the SASA of the ligand, averaged over all five ligands and 1,000 reference compound conformations from the simulations in the protein. The highest value was obtained for reference compound SDM, indicating that this reference compound encompasses the ligands only partly.

The overall structure of the protein remained quite stable over the course of the simulations. Backbone atom-positional rms deviations from the crystal structure [PDB ID code 3ERD (27)] go up to 0.2 nm. The occurrence (%) of hydrogen bonds involving the reference compound is indicated in Table 1. Hydrogen bonds with residues that are expected to form hydrogen bonds to the ligands, as well hydrogen bonds with water at the site formed by Glu-353 and Arg-394 are listed. Preliminary simulations with reference compounds that did not involve hydroxyl groups or soft hydroxyl groups showed a complete loss of these hydrogen bonds (data not shown). From Table 1 it becomes clear that reference compound SDS is able to maintain the largest number of hydrogen bonds, whereas SDM and SDL significantly loose hydrogen bonding to the protein, only partly compensated by increased hydrogen bonding with water. As can be seen from Fig. 4, the hydroxyl groups in SDL are quite buried, making it more difficult to form hydrogen bonds. However, Glu-353 remains oriented toward the hydroxyl group while His-524 turns away from the reference compound and reforms a hydrogen bond repeatedly throughout the simulation. Earlier molecular dynamics simulations on the ER in complex with DES showed very similar hydrogen bonding patterns.

The first five rows of Table 2 list the values of $\Delta G'_{AR}$ for several simulations and all real ligands. In brackets are the free-energy

Table 2. Free energies for five real ligands

Free energy	E2	DES	GEN	MBA	COU
1. $\Delta G'_{A,SDMres}$ (*free*)	−98 (0)	−216 (−118)	−5 (93)	−114 (−16)	−79 (19)
2. $\Delta G'_{A,SDL}$ (*free*)	−98 (0)	−177 (−79)	29 (127)	−117 (−19)	−88 (10)
3. $\Delta G'_{A,SDS}$ (*complex*)	−131 (0)	−201 (−70)	−35 (96)	−149 (−18)	−108 (23)
4. $\Delta G'_{A,SDM}$ (*complex*)	−92 (0)	−55 (37)	0 (92)	−123 (−31)	−111 (−19)
5. $\Delta G'_{A,SDL}$ (*complex*)	−127 (0)	−199 (−72)	−19 (108)	−144 (−17)	−92 (35)
$\Delta\Delta G^{bind}$ (3 − 1)	−33 (0)	14 (47)	−30 (3)	−35 (−2)	−29 (4)
$\Delta\Delta G^{bind}$ (3 − 2)	−33 (0)	−24 (9)	−64 (−31)	−32 (1)	−20 (13)
$\Delta\Delta G^{bind}$ (4 − 1)	6 (0)	161 (155)	5 (−1)	−9 (−15)	−32 (−38)
$\Delta\Delta G^{bind}$ (4 − 2)	6 (0)	122 (116)	−29 (−35)	−6 (−12)	−23 (−29)
$\Delta\Delta G^{bind}$ (5 − 1)	−29 (0)	20 (49)	−14 (15)	−30 (−1)	−13 (16)
$\Delta\Delta G^{bind}$ (5 − 2)	−29 (0)	−22 (7)	−48 (−19)	−27 (2)	−4 (25)
ΔG^{bind} (exp)	−51.9 (0)*	−52.5 (−0.6)*	−39.3 (12.6)*	−48.5 (3.4)*	−38.2 (13.7)†

Free energies in kJ/mol obtained for the five real ligands in Fig. 2 from simulations of the three reference compounds of Fig. 4, free and in complex with the ER. Values for SDS and SDM free in water are not listed, because of the globular shape of the reference compounds (see text). The relative free energies of binding ($\Delta\Delta G^{bind}$) are obtained by subtracting the values in the indicated rows of the table. Values in parentheses are relative to E2.

*Ref. 16.

†M. M. H. van Lipzig, private communication.

differences relative to the real ligand E2. Because of the more globular shape of reference compounds SDS and SDM free in water, a fit of the real ligands onto the reference compounds always resulted in structures in which the real ligands were overlapping with water molecules, and no favorable water–ligand configurations were found. For that reason, the values of $\Delta G'_{A,SDS}$ (*free*) and $\Delta G'_{A,SDM}$ (*free*) are not given in Table 2. To determine whether sufficient conformations were sampled to obtain the free-energy values, Table 3 lists the number of conformations that were found to have a H'_A value below $\Delta G'_{AR} + k_BT$. From this table it becomes clear that the most flexible ligand (DES) is poorly represented by a single rigid structure. The lowest number of relevant conformations was consistently found for DES. In line with the finding that the reference compound SDM encompasses the real ligands worst, the number of contributing conformations found in simulations of the SDM (and SDMres) reference compound is very low.

Changing the reference compounds into real ligands seems to be more problematic in water than in the protein, as has been observed before (34). For the two real ligands, MBA and COU, similar results (relative to E2) are obtained for the two reference compounds (SDMres and SDL) in water. The DES and GEN values show larger differences. From previous thermodynamic integration calculations (3), the free-energy difference between E2 and DES in water was found to be −80 kJ/mol, and between E2 and GEN it was 74 kJ/mol. GEN is the compound with the largest number of polar groups. The fact that it has the highest free energy in water relative to E2 indicates that the neutral central part of the reference compound does not represent this polar character very well. The free energies that were obtained from simulations of the protein complexed with reference compounds SDS and SDL relative to E2 are of comparable size, with differences up to 12 kJ/mol. Previous thermodynamic integration calculations on E2, DES, and GEN yield values of comparable size, −79 kJ/mol for DES and 90 kJ/mol for GEN relative to E2 (3). By subtracting the free-energy values obtained from simulations of the reference compounds free in water from the values obtained from simulations when they were complexed to the protein, one can estimate the relative binding free energies for the real compounds (Eq. **3**). These values are reported in the lower part of Table 2 and should be compared with the experimental values at the lowest line taken relative to E2. It is immediately clear that none of the combinations of simulations reproduces the experimentally well established fact that DES has a high affinity for the ER. This discrepancy can undoubtedly be attributed to the lack of flexibility in the treatment of the ligand as being rigid. The ethyl groups are known to fill specific cavities in the binding site of the ER LBD but are not able to adapt to the shape of the protein in the approach involving rigid real ligands taken here. The relative free energy of binding for GEN is reproduced by comparing the protein simulations with SDL and SDS to the liquid simulation with SDM, but the large value that was obtained for the liquid simulation of SDL lowers the relative binding free energy too much. All combinations involving the simulations of SDS and SDL complexed to the protein reproduce the comparatively low free energy of binding for MBA, as well as the higher free energy of binding for COU. Omitting the results for the flexible ligand DES, the combination $\Delta\Delta G^{bind}$ (5 − 1) in Table 2 corresponds well with the experimental values.

No empirical input parameters were used in the calculation of relative binding free energies other than the parameters of the biomolecular force field. The protein flexibility is fully taken into account, as well as a rotational and translational sampling of the (rigid) real ligands. The logical next step would be to also account for the flexibility of the ligands, for instance, by using a representative set of ligand conformations to be fitted onto every protein-reference compound structure. The value of H'_A should include the internal interaction energy as well. Alternatively, one can think of performing an energy minimization for the ligand within the framework of the protein coordinates, or even of performing short simulations involving the intramolecular ligand degrees of freedom within every protein conformation. With the current availability of computational power, the evaluation of the ligand interactions over a large number of previously generated structures by using a standard biomolecular force field is no longer a problem. Even including the enhanced positional and orientational sampling, the computational effort for every real ligand is still significantly less than for

Table 3. Number of structures with H'_A smaller than $\Delta G'_{AR} + k_BT$

Simulation	E2	DES	GEN	MBA	COU
SDMres (*free*)	38	9	60	122	421
SDL (*free*)	4,102	3,779	5,145	6,309	5,275
SDS (*complex*)	1,097	215	759	750	1,504
SDM (*complex*)	8	4	35	52	49
SDL (*complex*)	2,251	169	782	1,827	2,342

Number of configurations involving the real compounds obtained from simulations of the reference compounds that have a value of H'_A lower than $\Delta G'_{AR} + k_BT$. See Eq. **2**. The total number of configurations considered was 10^6.

complete simulations of the protein–ligand complex. Moreover, the free-energy analysis over previously generated configurations of the protein–ligand complex can easily be distributed over a large number of computers.

The advantage of the one-step approach over alternative computational methods is that only a single simulation is required from which relative free energies can be estimated for many ligands. Based on results obtained for one set of compounds, one can easily design and explore alternative ligands, without having to run extensive simulations of these ligands, as would be required by thermodynamic integration calculations (35), or more empirical free energy estimates based on end-point simulations (36, 37). Compared with automated docking approaches combined with empirical scoring functions (15, 38), the one-step approach has the advantage that the protein flexibility is fully taken into account. Instead of allowing for local movement in the active site or using a small set of protein coordinates, this approach naturally selects the most important protein configurations for every individual ligand. In addition, no training set of ligands is required to parameterize the empirical scoring function. A common biomolecular force field describing the interactions between atoms is sufficient to obtain accurate free-energy estimates.

Conclusion

It has been shown that the one-step perturbation approach can be extended to reference compounds that contain little structural information about the ligands in which one is interested. The ligands are fitted into the active site of the protein, taking the full protein flexibility into account as well as an enhanced sampling of the ligand orientation and location. The accuracy of the obtained free energies is limited by the use of rigid real compounds in the current approach. However, inclusion of the ligand flexibility should be possible and is a matter of ongoing research. The approach does not require parameterization of any empirical parameters other than a standard biomolecular force field to describe the interactions. Advantageous is the fact that relative free energies for additional real compounds can easily be calculated *a posteriori*. Moreover, the free-energy calculations from previously generated configurations can be massively parallelized. These advantages make the one-step perturbation approach based on an unphysical reference state without structural limitations to the ligands to which it is being applied an efficient means to understand and predict relative free energies.

This work was supported by the National Center of Competence in Research Structural Biology of the Swiss National Science Foundation.

1. Liu, H. Y., Mark, A. E. & van Gunsteren, W. F. (1996) *J. Phys. Chem.* **100,** 9485–9494.
2. Schäfer, H., van Gunsteren, W. F. & Mark, A. E. (1999) *J. Comput. Chem.* **20,** 1604–1617.
3. Oostenbrink, B. C., Pitera, J. W., van Lipzig, M. M. H., Meerman, J. H. N. & van Gunsteren, W. F. (2000) *J. Med. Chem.* **43,** 4594–4605.
4. Pitera, J. W. & van Gunsteren, W. F. (2001) *J. Phys. Chem. B* **105,** 11264–11274.
5. Oostenbrink, C. & van Gunsteren, W. F. (2003) *J. Comput. Chem.* **24,** 1730–1739.
6. Oostenbrink, C. & van Gunsteren, W. F. (2004) *Proteins* **54,** 237–246.
7. Zwanzig, R. W. (1954) *J. Chem. Phys.* **22,** 1420–1426.
8. Tembe, B. L. & McCammon, J. A. (1984) *Comput. Chem.* **8,** 281–283.
9. Lybrand, T. P., McCammon, J. A. & Wipff, G. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 833–835.
10. Wong, C. F. & McCammon, J. A. (1986) *J. Am. Chem. Soc.* **108,** 3830–3832.
11. Beutler, T. C., Mark, A. E., Van Schaik, R. C., Gerber, P. R. & van Gunsteren, W. F. (1994) *Chem. Phys. Lett.* **222,** 529–539.
12. Gantchev, T. G., Ali, H. & van Lier, J. E. (1994) *J. Med. Chem.* **37,** 4164–4176.
13. Bradbury, S. P., Mekenyan, O. G. & Ankley, G. T. (1996) *Environ. Toxicol. Chem.* **15,** 1945–1954.
14. Tong, W. D., Perkins, R., Xing, L., Welsh, W. J. & Sheehan, D. M. (1997) *Endocrinology* **138,** 4022–4025.
15. Yoon, S. & Welsh, W. J. (2004) *J. Chem. Inf. Comput. Sci.* **44,** 88–96.
16. van Lipzig, M. M. H., ter Laak, A. M., Jongejan, A., Vermeulen, N. P. E., Wamelink, M., Geerke, D. & Meerman, J. H. N. (2004) *J. Med. Chem.* **47,** 1018–1030.
17. Morreal, C. E., Sinha, D. K., Schneider, S. L., Bronstein, E. & Dawidzik, J. (1982) *J. Med. Chem.* **25,** 323–326.
18. Waller, C. L., Oprea, T. I., Chae, K., Park, H. K., Korach, K. S., Laws, S. C., Wiese, T. E., Kelce, W. R. & Earl Gray, L. (1996) *Chem. Res. Toxicol.* **9,** 1240–1248.
19. Kramer, V. J., Helferich, W. G., Bergman, A., KlassonWehler, E. & Giesy, J. P. (1997) *Toxicol. Appl. Pharmacol.* **144,** 363–376.
20. Connor, K., Ramamoorthy, K., Moore, M., Mustain, M., Chen, I., Safe, S., Zacharewski, T., Gillesby, B., Joyeux, A. & Balaguer, P. (1997) *Toxicol. Appl. Pharmacol.* **145,** 111–123.
21. Breinholt, V. & Larsen, J. C. (1998) *Chem. Res. Toxicol.* **11,** 622–629.
22. Charles, G. D., Bartels, M. J., Zacharewski, T. R., Gollapudi, B. B., Freshour, N. L. & Carney, E. W. (2000) *Toxicol. Sci.* **55,** 320–326.
23. Blair, R. M., Fang, H., Branham, W. S., Hass, B. S., Dial, S. L., Moland, C. L., Tong, W. D., Shi, L. M., Perkins, R. & Sheehan, D. M. (2000) *Toxicol. Sci.* **54,** 138–153.
24. Hong, H. X., Tong, W. D., Fang, H., Shi, L. M., Xie, Q., Wu, J., Perkins, R., Walker, J. D., Branham, W. & Sheehan, D. M. (2002) *Environ. Health Perspect.* **110,** 29–36.
25. Anstead, G. M., Carlson, K. E. & Katzenellenbogen, J. A. (1997) *Steroids* **62,** 268–303.
26. Brzozowski, A. M., Pike, A. C. W., Dauter, Z., Hubbard, R. E., Bonn, T., Engstrom, O., Ohman, L., Greene, G. L., Gustafsson, J. A. & Carlquist, M. (1997) *Nature* **389,** 753–758.
27. Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard, D. A. & Greene, G. L. (1998) *Cell* **95,** 927–937.
28. Pike, A. C. W., Brzozowski, A. M., Hubbard, R. E., Bonn, T., Thorsell, A. G., Engstrom, O., Ljunggren, J., Gustafsson, J. K. & Carlquist, M. (1999) *EMBO J.* **18,** 4608–4618.
29. Adlercreutz, H. & Mazur, W. (1997) *Ann. Med.* **29,** 95–120.
30. 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. (1996) *Biomolecular Simulation: The GROMOS96 manual and user guide* (vdf Hochschulverlag AG an der ETH Zürich, Zürich).
31. Scott, W. R. P., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, P., Krüger, P. & van Gunsteren, W. F. (1999) *J. Phys. Chem. A* **103,** 3596–3607.
32. Daura, X., Mark, A. E. & van Gunsteren, W. F. (1998) *J. Comput. Chem.* **19,** 535–547.
33. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F. & Hermans, J. (1981) in *Intermolecular Forces*, ed. Pullman, B. (Reidel, Dordrecht, The Netherlands), pp. 331–342.
34. Mark, A. E., Xu, Y. W., Liu, H. Y. & van Gunsteren, W. F. (1995) *Acta Biochim. Pol.* **42,** 525–535.
35. Beveridge, D. L. & DiCapua, F. M. (1989) *Annu. Rev. Biophys. Biophys. Chem.* **18,** 431–492.
36. Åqvist, J., Medina, C. & Samuelsson, J. E. (1994) *Protein Eng.* **7,** 385–391.
37. Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S. H., Chong, L., Lee, M., Lee, T. S., Duan, Y., Wang, W., *et al.* (2000) *Acc. Chem. Res.* **33,** 889–897.
38. Kuntz, I. D., Meng, E. C. & Shoichet, B. K. (1994) *Acc. Chem. Res.* **27,** 117–123.