LUND UNIVERSITY

Transition-state docking of flunitrazepam and progesterone in cytochrome P450

Rydberg, Patrik; Hansen, Sine Myrup; Kongsted, Jacob; Norrby, Per-Ola; Olsen, Lars; Ryde, Ulf

Published in:
Journal of Chemical Theory and Computation

DOI:
10.1021/ct700313j

2008

Document Version:
Peer reviewed version (aka post-print)

Link to publication

Citation for published version (APA):
Rydberg, P., Hansen, S. M., Kongsted, J., Norrby, P.-O., Olsen, L., & Ryde, U. (2008). Transition-state docking of flunitrazepam and progesterone in cytochrome P450. *Journal of Chemical Theory and Computation*, *4*(4), 673-681. https://doi.org/10.1021/ct700313j

Total number of authors:
6

Creative Commons License:
Unspecified

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 26. Apr. 2024

Transition-state docking of flunitrazepam and progesterone in cytochrome P450

*Patrik Rydberg [ab], Sine Myrup Hansen [a], Jacob Kongsted [b], Per-Ola Norrby [c], Lars Olsen [a] and Ulf Ryde [b]**

Biostructural Research Group, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen Ø, Denmark.

Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, S-221 00 Lund, Sweden

Transition–state docking in CYP

*To whom correspondence should be addressed: phone: (+46) 46 2224502; fax: (+46) 46 2224543; e-mail: Ulf.Ryde@teokem.lu.se. [a] University of Copenhagen. [b] Lund University. [c] Göteborg University.

Abstract. We have developed a method to dock a transition-state structure into the active site of an enzyme. Such an approach is more discriminative than standard docking when looking for substrates of an enzyme, because a transition state has more sterical restrictions than a non-reactive state. We apply this method to the docking of two drugs, progesterone and flunitrazepam, to the active sites of two human cytochromes P450, 2C9 and 3A4, using a molecular mechanics transition-state force field obtained with the Q2MM method. We obtain a qualitative agreement compared to experiments, both for hydrogen atoms bound to the same carbon atom (for which the force-field energies are directly comparable) and for general sites on the drug molecules, if the method is combined with an estimate of the intrinsic reactivity of the various sites. However, the method does not rank all the sites correctly. It is not significantly improved if the proteins are allowed to relax locally or if it is combined with the MM/PBSA approach, which fully accounts for the protein flexibility and explicitly treats solvation and entropy effects. On the other hand our method performs significantly better than standard docking with the GOLD software or predictions of metabolic sites with the MetaSite software.

Keywords. cytochromes P450, docking, MM/PBSA, flunitrazepam, progesterone

Introduction

The effect of drugs is determined not only by their action on their target proteins but also by the concentration of the drug at the target. This concentration is a function of many different factors, e.g. the amount administered, the uptake in the body, the discretion from the body, and the modification and degradation of the drug in the body. Thus, the drug metabolism and pharmacokinetics are very important for the success of a drug candidate and much effort has been devoted to their measurement and prediction.

Of particular interest in this aspect is the group of enzymes called cytochromes P450 (CYPs). They are oxidative liver enzymes that degrade foreign substances in the body. In fact, they metabolize ~90% of the drugs on the market.[1] The active site of these enzymes consists of a haem group, i.e. an iron ion in the centre of a porphyrin ring. Below the ring, the metal is coordinated to a cysteine ligand, whereas the upper site is open to the coordination of a small ligand, like water or O_2. During the catalytic cycle, a highly reactive Fe(V)=O complex (formally) is formed, called compound I. It has the potential to oxidize most chemical groups, by aliphatic hydroxylation, aromatic hydroxylation or epoxidation, dealkylation, N, S, or SO oxidation, etc.

Much effort has been directed to the prediction of CYP reactivity of drug candidates. It has been realized that this consists of two topics, viz. estimation of the intrinsic reactivity of the various groups in the drug and calculation of the steric accessibility of the groups to the oxoferryl group of compound I, whereas dynamic effects during binding seem to be less important.[2] The intrinsic reactivity has been thoroughly studied for the most interesting reaction types, both with density functional theory (DFT) methods[3, 4] and by more approximate, but faster methods.[5, 6] The steric effects have also been studied, e.g. by solvent accessibility, docking, and QSAR methods.[7, 8] In some cases, both types of effects were considered.[7, 9, 10]

However, nobody seems to have employed the fact that for substrates, a transition state must form. Such a state has quite severe sterical restrictions (for aliphatic hydroxylation, the reacting hydrogen

atom on the substrate should be ~1.25 Å from the oxoferryl group with Fe–O–H and O–H–C angles of ~121° and 171°).[11] Thus, docking of the transition state should be more discriminative than a normal docking of the substrate into the active site. Therefore, we in this paper develop a transition-state docking procedure, and test it on two human CYPs.

Docking programs employ an empirical energy function to determine and optimize the interaction energy between the drug candidate and the active site, typically in the form of a molecular mechanics (MM) force field. This is a problem for docking of transition states, because force fields normally are only developed for equilibrium states – for transition states, which are first-order saddle points on the potential surface, special optimization algorithms are needed, which are less robust than those for minima (i.e. they cannot guarantee that a transition state is found from every starting position and they normally require information about the curvature of the potential around the saddle point). Four types of methods have been employed in the optimization of transition states with MM methods.[12, 13] In this paper, we use the simplest approach, Q2MM,[14] in which the transition state is converted to an equilibrium state during the parameterization (the negative frequency is replaced with a high positive frequency). Thereby, any minimization code can be used and a transition state is found from any reasonable structure, which is important for a stable docking code.

We use a general Q2MM force field for the transition state of aliphatic hydroxylation, which we recently developed and tested.[11] It was obtained for a training set of 14 transition-state structures of small model molecules with most chemical groups commonly encountered in drugs, optimized at the DFT level, and it was tested for a set of 10 additional molecules. We use this force field to dock two drugs, progesterone and flunitrazepam, into the active sites of two human CYPs, 3A4 and 2C9. We test whether we can predict which sites of the drugs will be metabolized by the respective enzyme by combining the docked results with a measure of the intrinsic reactivity of each site.[4] We also test if the predictions of the method can be improved by optimizing the geometry of the active site side chains or by including dynamic, solvation, hydrophobic, and entropy effects by the MM/PBSA (molecular

mechanics Poisson–Boltzmann surface area) approach.[15] Finally, we compare the results with two other methods to predict reactivity in the CYPs, standard docking with the GOLD software[16] and CYP reactivity predictions by MetaSite.[7]

Methods

Systems studied. Our calculations are based on the crystal structures of CYP 2C9 and 3A4 (PDB entries 1R9O and 1TQN).[17, 18] The protein part of the system was set up in the same way as in our previous study of the water dynamics in the active-site cavity of these proteins.[19] To describe the transition state, we added an oxygen atom with a bond length of 1.76 Å to the iron ion, opposite to the sulfur atom of the cysteine ligand.

We studied two substrates, progesterone and flunitrazepam (shown in Figure 1). They were described with the general AMBER force field[20] and the transition state was modeled by our recently developed transition-state force-field parameters.[11]

We studied the transition states for hydroxylation of the 2β, 6β, 16α, 16β, 17α, and 21 positions on progesterone, and the 1 and 3 positions on flunitrazepam. The β hydrogen atoms in progesterone are directed towards the viewer in Figure 1.

Transition-state docking. All calculations with the transition-state force field were done with the AMBER software suite, version 9.[21] Molecular mechanics minimizations were run with a distance-dependent dielectric constant of 4*r* and an infinite cut-off for non-bonded interactions. The temperature was kept constant at 300 K using the Berendsen weak-coupling algorithm[22] with a time constant of 1 ps. We also tested a generalized Born (GB) solvation model,[23, 24] but this increased the calculation time too much.

First, we generated a starting position with the reactive hydrogen atom close to the oxoferryl group by manually docking the substrates into the active site in a reasonable conformation. This structure was

refined by a molecular mechanics optimization, in which the protein and the haem group were kept fixed and only the substrate allowed to move. To fully sample all possible conformations, we then did a full conformational analysis of the substrates by systematically rotating (with local software) the four rotable bonds (Fe–O, O–H, H–C, and C17–C20 in progesterone or the bond to the fluorophenyl ring in flunitrazepam) six-fold (eight-fold for the Fe–O bond) with a fixed protein structure, generating 1728 conformers of the substrates. All these structures were optimized with molecular mechanics, keeping the protein and haem group fixed. From this conformational analysis, we took the structure with the lowest energy for further analysis of the binding and estimation of the binding affinities.

For flunitrazepam, there are two possible conformations of the seven-membered ring, as is shown in Figure 2. The two conformations are degenerate in quantum chemical calculations. Therefore, we tested both conformations, but report only the results of the conformation that gives the lowest docked energy (conformation 1 for all calculations, except for the reaction of C1 in CYP 3A4).

In some calculations we studied the effects of the protein flexibility by performing an additional geometry optimization of the structure with the lowest energy, with all side chains around the active site free to move. These side chains were 100–102, 104, 106, 108, 113, 114, 197, 200, 201, 204, 205, 208, 209, 233, 234, 236, 237, 240, 292, 293, 295–297, 299–301, 304, 361, 362, 366, 474, 476, 477, and 479 in 2C9, and 105, 106, 108, 119, 120, 212, 213, 215, 241, 301, 304, 305, 309, 369, and 370 in 3A4.

Binding energies. The binding energy of a ligand (L) to a protein (P) is the free energy of the reaction:

$$P + L \rightarrow PL \qquad (1),$$

where PL is the complex between the protein and the ligand. Therefore, we could estimate the binding by

$$E'_{bind} = E_{PL} - E_P - E_L \qquad (2),$$

where E_{PL}, E_L, and E_P are the MM energies of PL, P, and L. However, for a transition-state, in which

there is a bond between the reactive hydrogen atom and the oxoferryl group of compound I, E_{PL} contains bonded energy terms (one bond, two angles, and six or seven dihedrals) that are not present in either E_L or E_P. A simple and intuitive way to compensate for this is to subtract the energy of these nine or ten interaction terms in a fully optimized small model of the substrate and the haem group, E_{tsbond}.

$$E''_{bind} = E_{PL} - E_P - E_{L,opt} - E_{tsbond} \quad (3)$$

Thereby, all bonded terms appear on both sides in Eqn. 3, making the energies comparable. Moreover, E_{tsbond} provides an estimate of how much the transition state is strained relative to the best possible unstrained structure.

Ideally, the small model for which we calculate E_{tsbond} should be the same model for which the intrinsic reactivity is estimated (see below), because for larger substrates, this energy will already include some steric effects (the large haem group will restrict the approach of some sites to the reactive oxoferryl group). Therefore, we have used the Fe(porphine)(SCH_3)(O)(substrate) model, which was used in our QM calculations.[4]

Strictly speaking, all three species in Eqn. 1 should be studied in their optimum states in aqueous solution. However, as a first approximation, we have kept the geometry of the protein fixed to save time and avoid the risk of ending up in different local minima for the various complexes. On the other hand, we optimize the geometry of the isolated ligand, using an implicit solvent model with a distance-dependent dielectric constant (μ=4r), giving the energy $E_{L,opt}$. Thus, we end up with the estimate:

$$E_{bind} = E_{PL} - E_P - E_{L,opt} - E_{tsbond} \quad (4)$$

To this binding affinity of the transition state, we should add an estimate of the intrinsic reactivity of each site (when comparing different reactive sites), E_{QM}. In our previous paper on aliphatic hydroxylation we found that activation barriers calculated with density functional theory employing the B3LYP functional using a methoxy radical gives a reliable estimate of the reaction barrier (E_{QM}).[4] The E_{QM} values for progesterone were published in that paper and the flunitrazepam data can be found in the supplementary material. Thus, our estimated activation energy for the various reactive sites is

$$E_{\text{est}} = E_{\text{bind}} + E_{\text{QM}} \quad (5)$$

Absolute values of E_{est} do not have any specific meaning, but relative values of different reactive sites of the same drug in one protein should indicate their relative reactivities, a positive value indicating a higher activation barrier and therefore a poorer substrate.

MM/PBSA calculations. The energies in Eqns. 2–4 are pure MM energies and they are obtained with a primitive solvation model. In an attempt to improve these, we have employed the MM/PBSA approach[15] for the docked structure with the lowest energy. In this approach, each of the three (free) energies on the right-hand-side of Eqn. (2) are estimated as a sum of four terms:

$$E = \langle E_{MM} \rangle + \langle G_{Solv} \rangle + \langle G_{np} \rangle - T\langle S_{MM} \rangle \quad (6)$$

where G_{Solv} is the polar solvation energy of the molecule, estimated by the solution of the Poisson–Boltzmann (PB) equation,[15] G_{np} is the non-polar solvation energy (including the cost of making a cavity in the solvent, solvent entropy, the hydrophobic effect, and solvent–solute dispersion and repulsion), estimated form the solvent-accessible surface area of the molecule[25], T is the temperature, S_{MM} is the entropy of the molecule, estimated from a normal-mode analysis of harmonic frequencies calculated at the molecular mechanics (MM) level, and E_{MM} is the MM energy of the molecule, i.e. the sum of the internal energy of the molecule (i.e. the bonded terms, E_{int}) and the electrostatics (E_{es}) and van der Waals interactions (E_{vdW}):

$$E_{MM} = E_{int} + E_{es} + E_{vdW} \quad (7)$$

All the terms in Eqn. (6) are averages of energies obtained from 20 snapshots taken from molecular dynamics (MD) simulations. In order to reduce the time-consumption and to obtain stable energies, the same geometry is normally used for all three reactants (complex, ligand and receptor), i.e. only the *PL* complex is simulated by MD.[26]

The MM/PBSA calculations followed the same protocol and settings as in Ref. 27 , and here we only outline the differences. In the MD simulations, the SHAKE algorithm[28] was not used, and consequently

a smaller time step (0.5 fs) had to be used. However, the duration of the MD simulations for both the equilibration (275 ps) and production runs (200 ps) were the same as in Ref. 27. All MD simulations used periodic boundary conditions with an octahedral box, extending at least 9 Å outside the protein (in total, ~40400 and 46500 atoms in CYP 2C9 and 3A4, respectively). Each MM/PBSA calculation was based on 20 snapshots extracted from the MD production run. The components of the binding energy in Eqn. 6 were obtained using the mm_pbsa module of Amber 8.0.[21] The polar solvation energy was calculated with the Poisson–Boltzmann model, calculated with the DelPhi II software,[29] or with the default Generalized Born model in Amber 8.0 (GB^{OBC})[24] (we use the acronym MM/GBSA for the latter results).

It has previously been shown that the entropic contribution to the free energy obtained within the MM/PBSA method may lead to large fluctuations and thereby a significant standard deviation in the predicted binding affinities.[27] The reason for this is that the protein is truncated and then minimized before the calculation of the frequencies from which the vibrational contribution to the entropy is derived. We have recently formulated an alternative method where a fixed buffer region is introduced into the minimization and frequency calculations,[30] thereby stabilizing the vibrational entropic term. This new method is employed in this paper.

Finally, we have added the same E_{tsbond} correction for the extra bond in the transition-state complex and the E_{QM} correction for the intrinsic reactivity of each site as in the E_{est} estimate to all the MM/PBSA and MM/GBSA energies.

GOLD and MetaSite calculations. For comparison, we also performed a set of standard docking calculations, using the GOLD software,[16] version 3.1. All water molecules except the one bound to the haem iron were deleted from the crystal structure (and also ligands, if any), hydrogen atoms were added, the hydrogen atom on the cysteine sulfur bound to iron was deleted, and a bond between iron and the cysteine sulfur was added. A bond between the iron ion and the remaining water molecule was added

with a distance of 1.62 Å and the water hydrogen atoms were removed. We used a flat-bottomed harmonic distance restraint of 2.4–2.7 Å between this oxygen atom and the hydrogen atom to be abstracted with a spring constant of 50.0 kcal/mol/Å^2 to define the reactive site. This distance was chosen from an optimized complex of methane and compound I, in which it is 2.55 Å. This is similar to the approach used by Vermeulen and co-workers, in which only docked poses for which the reactive carbon atom was within 6 Å of the iron ion were considered.[31]

Finally, we also predicted the reactivity of the various sites with the MetaSite software, version 2.7.5.[7] These calculations employed the default settings and the parameters for the CYP3A4 and CYP2C9 models implemented in the software.

Results and Discussion

Flunitrazepam. Experimentally, it is known that the CYPs convert flunitrazepam to desmethylflunitrazepam and 3-hydroxyflunitrazepam.[32] These metabolites are formed after initial hydrogen abstraction from the aliphatic carbons C1 and C3, respectively, and several different CYP isoforms can catalyze these reactions. For example, CYP2A6 and 3A4 catalyze both reactions, whereas CYP 2B6, 2C9, and 2C19 only cause demethylation.[33]

We have docked the transition states corresponding to the hydroxylation reactions on C1 and C3 into CYP 2C9 and 3A4 and performed a systematic conformation analysis, as described in the Methods section. For the most stable of the 1728 considered conformations, we calculated the estimated binding affinity by Eqn. 5. The results are listed In Table 1. It can be seen that there is a significant correlation between E_{est} and the experimental data. For CYP 3A4, E_{est} is around –85 kJ/mol for both reactions, whereas for CYP 2C9 E_{est} is negative for the 1-demethylation, but positive for the 3-hydroxylation. This is in good agreement with the experimental observation that CYP 2C9 does not catalyze the reaction at C3, whereas CYP 3A4 catalyses both reactions. However, our method predicts that the C1 site should be more reactive than the C3 site in CYP 3A4 (by 4 kJ/mol), although experiments indicate the opposite

(by 6 kJ/mol), so the predictions are not fully quantitative.

Progesterone. It is well-known that steroids in general and progesterone in particular are metabolized by several human CYPs at different positions.[34-36] All the various metabolites of progesterone are alcohols formed by hydroxylation of aliphatic carbons, which makes it a good test case for our force field and transition-state docking procedure. Four different metabolites are produced by CYP 3A4, viz. the 2β-, 6β-, 16α-, and 21-hydroxylated progesterone, whereas CYP 2C9 primarily catalyses the reaction at the 21 position (with minor metabolites from the 6β and 16α positions).

To test our force field, we first studied the E_{PL} energies (i.e. the docked energies) of the 16α and the 16β hydroxylations, because for hydrogen atoms bound to the same carbon atom, the Q2MM energies are directly comparable. Quite satisfactorily, our results show that the E_{PL} energies are 80 and 246 kJ/mol lower for the H16α atom than for the H16β atom in CYP 3A4 and 2C9, respectively, clearly explaining why only the 16α product is observed.[36] These energy differences come from the bonded energy terms (bonds, angles, and dihedrals) and the van der Waals energies, about 50% from each, whereas electrostatics has only a minor influence. This shows that the preference to react at H16α comes from a better fit in the active site.

However, to compare all the other reactive sites, we instead need to study the E_{est} energies. These are listed and compared to experimental V_{max} values in Table 2. It can be seen that most of the results are qualitative correct. E_{est} is large and positive (139–695 kJ/mol) for the two sites (16β and 17) that do not react in any of the two proteins. A similar high value (178 kJ/mol) is observed also for the 2β position in CYP 2C9, the product of which is also not observed. The other sites have lower E_{est} energies in CYP 3A4 (-68 to +136 kJ/mol), and these reactions are also observed experimentally. The experimentally most reactive site in CYP 2C9, C21, has the lowest E_{est}, 82 kJ/mol. However, E_{est} for the non-reactive 2β position (178 kJ/mol) is smaller that of the 6β and 16α positions (335 and 289 kJ/mol), which are observed experimentally, although only as minor metabolites. Moreover, the quantitative correlation

between E_{est} and V_{max} is poor for CYP 3A4.

Improvements to the transition-state docking. The approach used in our transition-state docking method, although similar to other docking approaches, is admittedly quite primitive, ignoring many important contributions to the true free energy of binding.[37] In particular, we keep the protein fixed in both the docking and the energy calculation.

This is the most common approach in docking, and it normally gives good results. However, it does not include the flexibility of the protein, which could be a reason why we cannot rank all sites correctly. A first attempt to include the flexibility of the protein would be to allow some amino acids to relax during the docking.

We tried such an approach by performing a geometry optimization of the amino-acid side chains in the active for the best docked pose of each reactive site, as is described in the Methods section. The results are described in Tables 3 and 4. It can be seen that the estimated activation energies (E_{flex}) all become more favorable as expected, and also much more similar and therefore more realistic. Unfortunately, this means that it becomes harder to discern which sites are metabolized and which are not. For flunitrazepam, flexible transition-state docking performs excellently, pointing out the reactive sites in both proteins. However, for progesterone, the results are worse. E_{flex} clearly and correctly shows that the 16β site is not reactive in any of the proteins (by 77–155 kJ/mol). However, for other non-reactive sites (C17 in both proteins and 2β in CYP 2C9), E_{flex} is equal or more negative than for some of the reactive sites, although the energy difference is not large 0–9 kJ/mol. Moreover, the reactive sites are ranked in a completely erroneous order (but again with an energy difference of only 9–16 kJ/mol). Thus, such a partly flexible method does not lead to any consistent improvement of the transition-state docking.

However, the energy function used in our transition-state docking, Eqns. 2–4, involves many approximations. In particular, many important energy terms are missing e.g. solvation, dynamics,

entropy, hydrophobic effects, etc.[37] We have therefore tried to improve the results of our transition-state docking approach by including all these terms in a well tested and computationally effective way, viz. by the use of the MM/PBSA method.[15] In this approach, the standard MM energy is supplemented by energies for the entropy, polar and non-polar solvation, and all energies are calculated for a number of snapshots taken from a molecular dynamics simulation of the complex, thereby taking into account also dynamics effects and the protein flexibility. The calculations were performed only on the most favorable docked conformation for each site. Naturally, the MM/PBSA method was extended to the transition-state docking by including the E_{QM} correction for the intrinsic reactivity of each site and the E_{tsbond} correction of the additional bond in the transition state. The results of such calculations for the docking of flunitrazepam and progesterone into CYP 3A4 and 2C9 are also included in Tables 3 and 4.

For progesterone in CYP 2C9, MM/PBSA performs rather well, predicting that the three non-reactive sites 2β, 16β, and 17 have higher energies (161–176 kJ/mol) than the two most reactive sites (149–151 kJ/mol). Unfortunately, the least reactive site, 6β, has a higher energy 173 kJ/mol, and the ranking between the 21 and 16α sites is incorrect (by 7 kJ/mol). Likewise, for progesterone in CYP 3A4, MM/PBSA predicts a higher energy for the reactive 21 site (221 kJ/mol) than for one of the non–reactive site (16β with 198 kJ/mol). The MM/GBSA predictions are similar or slightly worse. For flunitrazepam, both MM/PBSA and MM/GBSA are poor, predicting that the C3 site is appreciably less reactive than the C1 site in both enzymes. Thus, the MM/PBSA method does not give any improvement compared to the simpler E_{est} estimate.

The MM/PBSA energies are dominated by a favorable van der Waals energy (~170 kJ/mol) and an unfavorable solvation energy (~180 kJ/mol for PB and ~90 kJ/mol for GB). Interestingly, the electrostatic energy is also unfavorable, so it is not cancelled by the solvation term, as is normally observed. This may explain the poor performance of MM/PBSA for the CYPs and it is probably connected to the fact that the active site is completely hidden in the centre of the enzyme. The difference between the C1 and C3 sites in both enzymes is completely caused by the electrostatic term. The

entropy term is always unfavorable (~80 kJ/mol) and the non-polar solvation is favorable and nearly constant (22 kJ/mol). The internal energy, which comes entirely from the 6–7 extra bonded interactions in the transition state, is unfavorable and small (~10 kJ/mol). It gives an indication of how strained each transition state is. In fact, it is perfectly correlated to V_{max} for progesterone in CYP 3A4 (all six sites are ranked correctly), but the ranking is much worse for CYP 2C9 or for flunitrazepam. The average statistical standard deviation of the MM/PBSA estimates is 6 kJ/mol and slightly lower for MM/GBSA, which also may explain the poor and varying results.

Comparison to other methods. We have seen that transition-state docking is rather successful to in discriminating reactive and non-reactive sites of both flunitrazepam and progesterone in CYP3A4 and 2C9, but it often fails to rank the sites correctly. In order to decide how useful such an approach is, we need to compare its performance with other available methods. Therefore, we have studied the same drugs and enzymes with two other methods. First, we test a standard docking (and scoring) program, GOLD.[16] The second is MetaSite,[7] which is another integrated method to predict the reactivity of each site of a drug molecule against several human CYPs. The results of these methods are shown in Table 3 and 4 for flunitrazepam and progesterone, respectively (full data on the GOLD and MetaSite calculations for all the sites on the substrates are available in the supplementary material).

For flunitrazepam, our method is the only one that gives qualitatively correct results for both enzymes: MetaSite gives a too large difference between 1-demethylation and 3-hydroxylation in 3A4, whereas GOLD does not give any significant difference between the two sites in 2C9.

For progesterone, our method had a qualitative problem with the 2β site in 2C9. However, MetaSite gives the second highest score to the non-reactive 16β position and the lowest score to the reactive 6β position in CYP2C9, and it gives even poorer predictions for CYP3A4 (the two non-reactive sites are ranked as number two and four). GOLD-docking performs somewhat better, but it gives the lowest score to the reactive 6β site in CYP 2C9 and the second lowest score for most reactive site in CYP 3A4.

Thus, transition-state docking gives significantly better results than other available methods.

Conclusions

In this paper, we have developed a method for transition-state docking of putative substrate molecules into the active site of an enzyme, and we have applied it for the prediction of metabolism pattern of two typical drugs (progesterone and flunitrazepam) in two human CYPs (3A4 and 2C9). Transition-state docking is a novel method to increase the discriminative power of a docking procedure by requiring that a productive transition state must form between the drug and the protein. To this end, a MM force field for the substrate is needed and this was developed for the hydroxylation of aliphatic carbon atoms by CYPs in a previous article.[11] This force field gives excellent docked structures of the two drugs in the two enzymes, when it is combined with a systematic search of the conformational space.

However, we also need an estimate of the activation energy of each of the docked structures. To estimate the intrinsic reactivity of each reactive site on the drug we have used DFT calculations with a methoxy radical model of the haem group. Such calculations can be performed within a few hours for most drug-like molecules.[4] A qualitative model,[4] which predicts the activation energies from the chemical environment of the reacting hydrogen atom, gives only slightly worse results (the two estimates differ by 3–18 kJ/mol).

Second, we need to estimate how well the transition state fits into the protein. This means that we should estimate the binding affinity of the drugs to the proteins, with the modification that the transition states involve a partial bond between the protein and the drug (an O–H bond to the oxoferryl group). This is a serious complication, because it means that standard methods of binding affinity cannot be used[37] and that the MM energies of different reactive sites are not comparable.

Depending on the energy of interest, this problem can be solved in different ways. For reactions involving different hydrogen atoms bound to the same carbon atom (e.g. H16α and H16β in progesterone), the force fields contain exactly the same terms and therefore, the MM energies are

directly comparable. Our results show that we can predict the reactivity of such sites successfully and the results can be directly interpreted in contributions from the various MM terms (bonded terms, van der Waals interactions, and electrostatics).

However, for more general applications, we have developed a method to compare activation barriers for different reactive sites in different proteins, combining both the intrinsic reactivity and steric effects, and correcting for the extra bonded terms in the transition state by comparing them to the ideal terms in a small optimized model complex, Eqn. 5. From the results in Tables 1 and 2, it can be seen that the method in all except one case gives qualitatively, but not always quantitatively, correct results (i.e. we can predict which sites are reactive, but we cannot rank them correctly).

Thus, the method works reasonably, but not perfectly. On the other hand, it is significantly better than alternative methods, viz. standard docking and scoring with GOLD[16] and metabolic predictions by MetaSite.[7] We have tried to improve the method by local optimization of the amino-acid side chains in the active site or by including full protein flexibility and improving the energy function with solvation, hydrophobic, and entropy effects, using the MM/PBSA approach.[15] Unfortunately, neither of these tests led to any significant improvement.

There are several possible ways to further improve the approach. First, we could try to use more conformations than the best one with local protein optimization or by MM/PBSA refinement. However, this would be much more expensive in terms of computer resources. Second, the Q2MM method for optimizing transition states essentially fixes the position of the transition state, allowing only for minimal variations in the geometry. This will exaggerate steric effects. Other methods, e.g. SEAM or the empirical valence bond method,[12] can be expected to model variations around the transition state more realistically, but they require special software. We currently work on various approaches to improve the method and we also try to develop similar methods to study other important CYP reactions with this approach.

Acknowledgment. This investigation has been supported by funding from the research school in pharmaceutical science (FLÄK), the Benzon foundation, the Carlsberg foundation, the Novo Nordisk Plus Prize, the Swedish Research Council, the Villum Kann Rasmussen Foundation, and by computer resources of LUNARC at Lund University.

Supporting Information Available. Hydrogen abstraction barriers for C1 and C3 in flunitrazepam calculated with B3LYP. Docking scores for all positions of progesterone. MetaSite scores for all positions in flunitrazepam and progesterone. This material is available free of charge via the internet at http://pubs.acs.org.

References

1. Hodgson, J. ADMET - turning chemicals into drugs. *Nature Biotechnology* **2001,** *19*, 722-726.
2. Rydberg, P. *Theoretical Studies of Cytochrome P450*; Lund University: Lund, 2007.
3. Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. Theoretical perspective on the structure and mechanism of cytochrome P450 enzymes. *Chem. Rev.* **2005,** *105*, 2279-2328.
4. Olsen, L.; Rydberg, P.; Rod, T. H.; Ryde, U. Prediction of activation energies for hydrogen abstraction by cytochrome P450. *Journal of Medicinal Chemistry* **2006,** *49*, 6489-6499.
5. Jones, J. P.; Korzekwa, K. R. Predicting the rates and regioselectivity of reactions mediated by the P450 superfamily. *Cytochrome P450, Pt B* **1996,** *272*, 326-335.
6. de Graaf, C.; Vermeulen, N. P. E.; Feenstra, K. A. Cytochrome P450 in silico: An integrative modeling approach. *Journal of Medicinal Chemistry* **2005,** *48*, 2725-2755.
7. Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. MetaSite: Understanding metabolism in human cytochromes from the perspective of the chemist. *Journal of Medicinal Chemistry* **2005,** *48*, 6970-6979.
8. Zamora, I.; Afzelius, L.; Cruciani, G. Predicting drug metabolism: A site of metabolism prediction tool applied to the cytochrome P4502C9. *Journal of Medicinal Chemistry* **2003,** *46*, 2313-2324.
9. Park, J. Y.; Harris, D. Construction and assessment of models of CYP2E1: Predictions of metabolism from docking, molecular dynamics, and density functional theoretical calculations. *Journal of Medicinal Chemistry* **2003,** *46*, 1645-1660.
10. Afzelius, L.; Arnby, C. H.; Broo, A.; Carlsson, L.; Isaksson, C.; Jurva, U.; Kjellander, B.; Kolmodin, K.; Nilsson, K.; Raubacher, F.; Weidolf, L. State-of-the-art tools for computational site of metabolism predictions: Comparative analysis, mechanistical insights, and future applications. *Drug Metabolism Reviews* **2007,** *39*, 61-86.

11. Rydberg, P.; Olsen, L.; Norrby, P. O.; Ryde, U. General Transition-State Force Field for Cytochrome P450 Hydroxylation. *J. Chem. Theory Comput.* **2007,** *3*, 1765-1773.

12. Jensen, F.; Norrby, P. O. Transition states from empirical force fields. *Theoretical Chemistry Accounts* **2003,** *109*, 1-7.

13. van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A reactive force field for hydrocarbons. *Journal of Physical Chemistry A* **2001,** *105*, 9396-9409.

14. Norrby, P. O. Molecular Mechanics as a Predictive Tool in Asymmetric Catalysis. In *Transition State Modeling for Catalysis*; Truhlar, D. G.; Morokuma, K. Eds.; American Chemical Society: Washington DC., 1999; pp 163-172.

15. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research* **2000,** *33*, 889-897.

16. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology* **1997,** *267*, 727-748.

17. Wester, M. R.; Yano, J. K.; Schoch, G. A.; Yang, C.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human cytochrome P4502C9 complexed with flurbiprofen at 2.0-angstrom resolution. *Journal of Biological Chemistry* **2004,** *279*, 35630-35637.

18. Yano, J. K.; Wester, M. R.; Schoch, G. A.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-angstrom resolution. *Journal of Biological Chemistry* **2004,** *279*, 38091-38094.

19. Rydberg, P.; Rod, T. H.; Olsen, L.; Ryde, U. Dynamics of water molecules in the active-site cavity of human cytochromes P450. *Journal of Physical Chemistry B* **2007,** *111*, 5445-5457.

20. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004,** *25*, 1157-1174.

21. Case, D. A.; Darden, T. A.; Cheatham, I. T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Homak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. AMBER. [9]. 2006. San Fransisco, University of California.

22. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to An External Bath. *Journal of Chemical Physics* **1984,** *81*, 3684-3690.

23. Bashford, D.; Case, D. A. Generalized born models of macromolecular solvation effects. *Annual Review of Physical Chemistry* **2000,** *51*, 129-152.

24. Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins-Structure Function and Bioinformatics* **2004,** *55*, 383-394.

25. Hermann, R. B. Theory of Hydrophobic Bonding .2. Correlation of Hydrocarbon Solubility in Water with Solvent Cavity Surface-Area. *Journal of Physical Chemistry* **1972,** *76*, 2754-&.

26. Swanson, J. M. J.; Henchman, R. H.; McCammon, J. A. Revisiting free energy calculations: A theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophysical Journal* **2004,** *86*, 67-74.

27. Weis, A.; Katebzadeh, K.; Soderhjelm, P.; Nilsson, I.; Ryde, U. Ligand affinities predicted with the MM/PBSA method: Dependence on the simulation method and the force field. *Journal of Medicinal Chemistry* **2006,** *49*, 6596-6606.

28. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of A System with Constraints - Molecular-Dynamics of N-Alkanes. *Journal of Computational Physics* **1977,** *23*, 327-341.

29. Rocchia, W.; Alexov, E.; Honig, B. Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *Journal of Physical*

Chemistry B **2001,** *105*, 6507-6514.

30. Kongsted, J.; Ryde, U. Unpublished results

31. de Graaf, C.; Oostenbrink, C.; Keizers, P. H. J.; van der Wijst, T.; Jongejan, A.; Vemleulen, N. P. E. Catalytic site prediction and virtual screening of cytochrome P450 2D6 substrates by consideration of water and rescoring in automated docking. *Journal of Medicinal Chemistry* **2006,** *49*, 2417-2430.

32. Coller, J. K.; Somogyi, A. A.; Bochner, F. Quantification of flunitrazepam's oxidative metabolites, 3-hydroxyflunitrazepam and desmethylflunitrazepam, in hepatic microsomal incubations by high-performance liquid chromatography. *Journal of Chromatography B* **1998,** *719*, 87-92.

33. Hesse, L. M.; Venkatakrishnan, K.; von Moltke, L. L.; Shader, R. I.; Greenblatt, D. J. CYP3A4 is the major CYP isoform mediating the in vitro hydroxylation and demethylation of flunitrazepam. *Drug Metabolism and Disposition* **2001,** *29*, 133-140.

34. Waxman, D. J.; Lapenson, D. P.; Aoyama, T.; Gelboin, H. V.; Gonzalez, F. J.; Korzekwa, K. Steroid-Hormone Hydroxylase Specificities of 11 Cdna-Expressed Human Cytochrome-P450S. *Archives of Biochemistry and Biophysics* **1991,** *290*, 160-166.

35. Richardson, T. H.; Jung, F.; Griffin, K. J.; Wester, M.; Raucy, J. L.; Kemper, B.; Bornheim, L. M.; Hassett, C.; Omiecinski, C. J.; Johnson, E. F. Universal Approach to the Expression of Human and Rabbit Cytochrome P450S of the 2C Subfamily in Escherichia-Coli. *Archives of Biochemistry and Biophysics* **1995,** *323*, 87-96.

36. Yamazaki, H.; Shimada, T. Progesterone and testosterone hydroxylation by cytochromes P450 2C19, 2C9, and 3A4 in human liver microsomes. *Archives of Biochemistry and Biophysics* **1997,** *346*, 161-169.

37. Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angewandte Chemie-International Edition* **2002,** *41*, 2645-2676.

Table 1. Calculated E_{est} energies (kJ/mol) for the reaction of flunitrazepam in human CYP 2C9 and 3A4. For comparison, experimental V_{max} data are also included (min^{-1}).[34]

	2C9		3A4	
	C1	C3	C1	C3
E_{est}	-27	86	-87	-83
V_{max}	0.38	–	1.54	15.7

Table 2. Calculated E_{est} energies (kJ/mol) for the reaction of progesterone in human CYP 2C9 and 3A4. In addition, experimental V_{max} data is included (min^{-1}).[37]

	2C9						3A4					
	21	16α	6β	2β	16β	17	6β	16α	2β	21	16α	17
E_{est}	82	289	335	178	533	695	136	60	17	-68	139	346
V_{max}	0.51	0.08[a]	0.04	–	–	–	33	8.7	8[a]	1.1	–	–

[a] Estimated from single-point measurements, assuming that the K_M values are equal to those of the 6β hydroxylation.

Table 3. A comparison of the results of transition-state docking with a rigid (E_{est}) or partly flexible (E_{flex}) protein, as well as the MM/GBSA, MM/PBSA, MetaSite and GOLD results for flunitrazepam. Energies are in kJ/mol (a low value indicates a reactive site), MetaSite and Gold scores in arbitrary units (a high value indicates a reactive site), and V_{max} in min^{-1}.

	2C9		3A4	
	C1	C3	C1	C3
E_{est}	-27	86	-87	-83
E_{flex}	-177	-162	-179	-179
MM/PBSA	130	320	174	391
MM/GBSA	82	237	78	219
MetaSite	1.69	1.07	2.13	1.50
GOLD	38.4	37.2	49.9	47.5
V_{max}	0.38	–	1.54	15.7

Table 4. A comparison of the results of transition-state docking with a rigid (E_{est}) or partly flexible (E_{flex}) protein, as well as the MM/GBSA, MM/PBSA, MetaSite and GOLD results for progesterone. Energies are in kJ/mol (a low value indicates a reactive site), MetaSite and Gold scores in arbitrary units (a high value indicates a reactive site), and V_{max} in min^{-1}.

	2C9						3A4					
	21	16α	6β	2β	16β	17	6β	16α	2β	21	16β	17
E_{est}	82	289	335	178	533	695	136	60	17	-68	139	346
E_{flex}	-209	-216	-218	-209	-132	-218	-214	-224	-224	-230	-59	-219
MM/PBSA	151	149	173	161	167	176	196	150	197	221	198	234
MM/GBSA	88	68	110	97	110	110	101	47	107	83	79	113
MetaSite	1.45	1.38	1.25	1.29	1.43	1.33	1.62	1.71	1.49	1.85	1.79	1.70
GOLD	39.1	35.3	3.11	15.4	7.11	19.7	16.2	32.4	35.3	43.7	32.2	-18.8
V_{max}	0.51	0.08[a]	0.04	–	–	–	33	8.7	8[a]	1.1	–	–

[a] Estimated from single-point measurements, assuming that the K_M values are equal to those of the 6β hydroxylation.

Figure 1. The substrates studied, progesterone (a) and flunitrazepam (b), with the considered carbon atoms indicated.

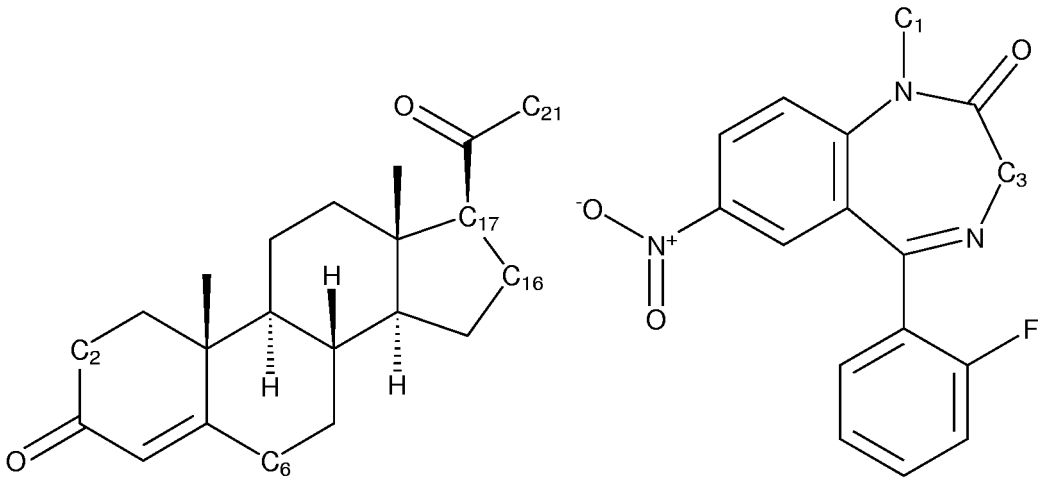

Figure 2. The two flunitrazepam conformations, conformation 1 in color and conformation 2 in black lines.

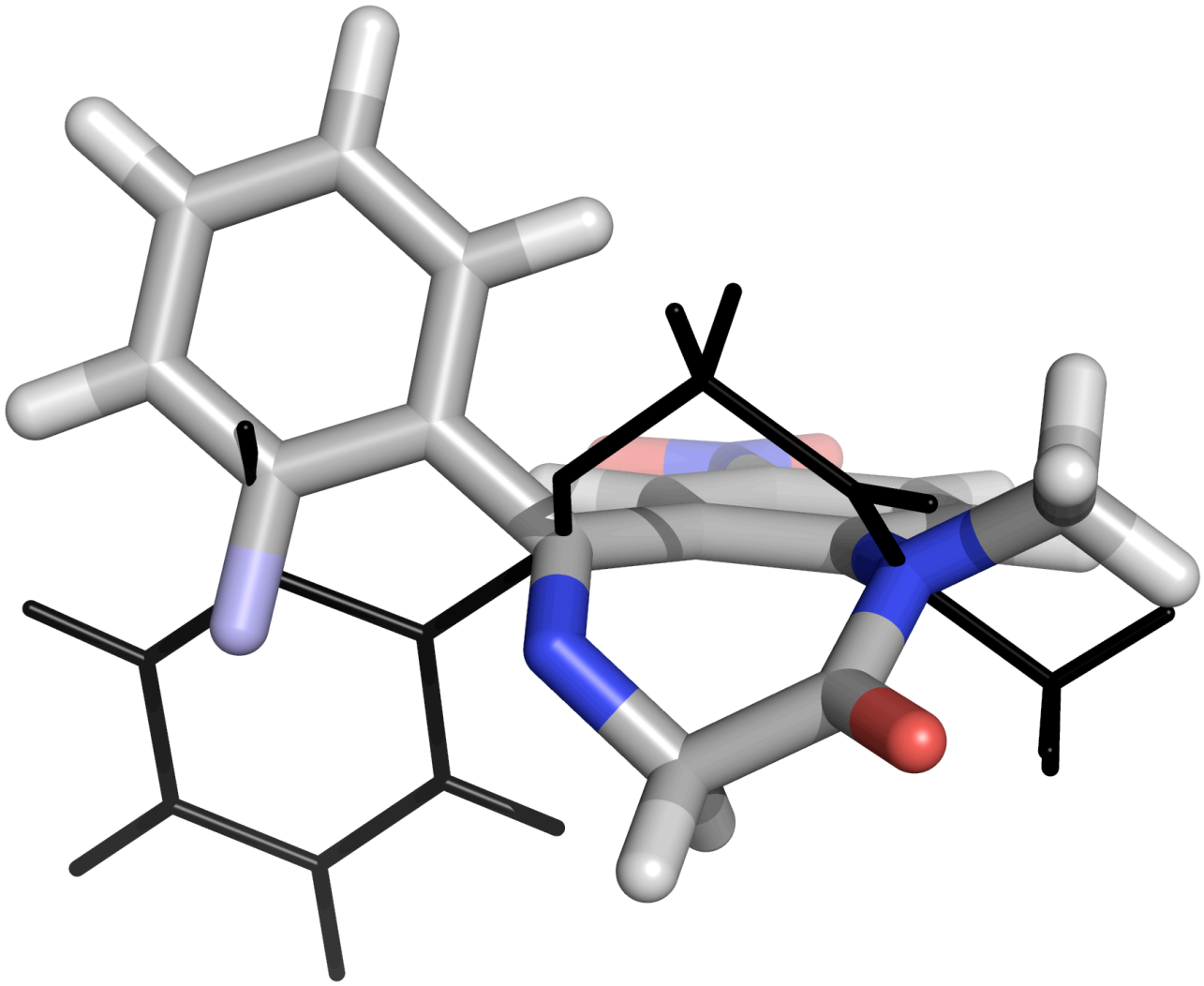

Table of contents graphic