Quantum-mechanical studies of reactions performed by cytochrome P450 enzymes

Rydberg, Patrik; Olsen, Lars; Ryde, Ulf

Published in:
Current Inorganic Chemistry

2012

Link to publication

Citation for published version (APA):

General rights
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

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.
Quantum-mechanical studies of reactions performed by
cytochrome P450 enzymes

Patrik Rydberg, Lars Olsen, Ulf Ryde

a Biostructural Research, Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark

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

E-mail: pry@farma.ku.dk

Abstract

We review density functional theory studies of various types of reactions performed by the cytochrome P450 family of enzymes. We describe the various reactions on equal footing with an emphasis on models to predict sites of metabolism for an arbitrary molecule. The activation barriers range between 0 and 109 kJ/mol, depending more on the atoms surrounding the reactive site than on the type of reaction. Therefore, the intrinsic reactivity can rather well be predicted by simple chemical rules. However, for a full predictive model, the steric effects of the enzyme surrounding the heme group need also to be modeled, which often is harder.

Keywords: cytochromes P450, heme, density functional theory, QM/MM

Introduction

The cytochrome P450 enzymes (CYPs) form a large and important protein superfamily of mono-oxygenases that is found in all types of organisms from bacteria to mammals. They take part in the synthesis and degradation of many physiologically important endogenous compounds, such as steroids, prostaglandins, and fatty acids. However, they also contribute to the degradation of xenobiotic compounds. Thereby, they affect the activation of prodrugs, as well as the bioavailability and degradation of many drugs. In the human genome, there are 57 genes for CYPs, but five of their products, CYP 1A2, 2C9, 2C19, 2D6, and 3A4, account for the oxidation of ~90% of the drugs on the market.¹

Numerous crystal structures of CYPs are available.² They show that the active site is buried inside the protein and is connected to the surface by several channels. At the bottom of the active site, there is a heme group with a central iron ion. Below the plane of the heme group, the sulfur atom of a cysteine (Cys) amino acid coordinates to the iron ion, whereas the site above the heme plane may bind various extraneous ligands during the reaction cycle.

The CYPs catalyze several different types of reactions, e.g. the hydroxylation of saturated C–H
bonds, oxidation of aromatic and alkene carbon atoms, alcohol and aldehyde oxidation, dealkylation, hydroxylation and oxidation of amines, oxidation of heteroatoms, dehalogenation, and desulfurization. The mechanisms of most of these CYP reactions have been extensively studied with both experimental and theoretical methods, and they are illustrated in Figure 1.

**Figure 1.** Reaction mechanisms performed by the CYPs. Intermediates that can exist in multiple electronic states (Fe^{IV} and Fe^{III}) have for simplicity been labeled as Fe^{IV}.

<table>
<thead>
<tr>
<th>Reaction Mechanism</th>
<th>Diagram</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aliphatic hydroxylation and dealkylation</strong></td>
<td><img src="image1.png" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>Dealkylation</strong></td>
<td><img src="image2.png" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>Alkene epoxidation</strong></td>
<td><img src="image3.png" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>Aromatic oxidation</strong></td>
<td><img src="image4.png" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>Hydroxylation of primary and secondary amine nitrogens</strong></td>
<td><img src="image5.png" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>Oxidation of tertiary amine nitrogens</strong></td>
<td><img src="image6.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Oxidation of aromatic nitrogens

Alcohol oxidation

Aldehyde oxidation (A) and deformylation (B)

Desulfurization of phosphors

Oxidation of sulfide sulfurs

Oxidation of sulfoxide sulfurs

Oxidation of aromatic sulphurs
Dehalogenation of halogenated benzenes

<table>
<thead>
<tr>
<th>Dehalogenation of halogenated benzenes</th>
<th><img src="image" alt="Diagram" /></th>
</tr>
</thead>
</table>

\( ^a \) This non-enzymatic step, exemplified by demethylation of an amine, follows upon the aliphatic hydroxylation mechanism.

\( ^b \) The suggested mechanism has been slightly simplified.

\( ^c \) The B mechanism leads to the same product as the A mechanism after rearrangement in water.

In the resting state, the heme site is six-coordinate low-spin Fe\(^{\text{III}}\), with a water molecule as the sixth ligand (1 in Figure 2).\(^{1,3,4}\) The reaction cycle starts with binding of the substrate, which leads to the dissociation of the water molecule and a transition to the high-spin state (2), although the substrate does not coordinate directly to the iron ion. A one-electron reduction transforms the ion to high-spin Fe\(^{\text{II}}\) (3), which subsequently binds \( \text{O}_2 \), giving a Fe\(^{\text{III}}\)–superoxide complex (4). The addition of another electron gives a Fe\(^{\text{III}}\)-peroxo complex (5) that takes up a proton to form a hydroperoxide intermediate, called compound 0 (6). If another proton is added to this complex, the O–O bond breaks and a water molecule dissociates from the site, leaving a highly reactive Fe\(^{\text{V}}\)=O (formally) intermediate, called compound I (7). For most reactions, compound I is assumed to be the catalytically competent intermediate, but in some cases other species have been suggested to be involved, e.g. compound 0 or compound II, which is the one-electron reduced variant of compound I.

**Figure 2.** The CYP reaction cycle, using the hydroxylation of an aliphatic group as a typical example.
Compound I has an interesting electronic structure: Quantum-mechanical (QM) calculations indicate that it is best described as a ferryl state (Fe$^{IV}$=O) with two unpaired electrons with parallel spin in the Fe–O $\pi$ orbitals (giving approximately one unpaired spin each on Fe and O). In addition, there is a third unpaired electron in the porphyrin ring, which is more weakly coupled to the other two unpaired electrons (i.e. compound I is a triradicaloid). This electron can have either parallel or antiparallel spin to the other two electrons (ferro- or antiferromagnetic coupling), giving rise to a quartet and a doublet state, which are energetically essentially degenerate. This gives rise to a two-state reactivity of compound I, which has been much discussed.

The QM studies have shown that the geometric and electronic structure of compound I strongly depend on details in the calculations. For example, if the Cys ligand is modeled by SCH$_3^-$ or a full cysteine residue, the third unpaired spin (outside the Fe–O bond) resides mainly on the S$_{Cys}$ atom. However, if instead the smaller SH$^-$ model is used, the spin moves partly to the porphyrin ring, so that the two groups have about half an unpaired electron each. If the complex is submerged in a continuum solvent with a dielectric constant of 5.7 or if two ammonia groups are included as models of hydrogen bonds to S$_{Cys}$, the spin resides predominantly on the porphyrin ring. The same effect is observed if the surrounding enzyme is included as a point-charge model. Concomitantly to this movement of unpaired spin from S$_{Cys}$ to the porphyrin ring, leading to an
increasing negative charge on S_Cys, the Fe–S bond length decreases, whereas the Fe–O bond length is essentially constant. In view of this sensitivity of the electronic structure and the Fe–S geometry to the details of the calculations, Shaik and coworkers characterized compound I as a chameleon species that changes its nature depending on the external conditions. Recently, there have been quite some interest in using advanced QM methods to study various CYP reactants, in particular compound I. Such calculations show that compound I has a complicated electronic structure, with many low-lying excited states. The results indicate that both pentaradicaloid (with five unpaired electrons) and Fe^V states are nearly degenerate (within 40 kJ/mol) with the triradicaloid ground state. This is important, because both calculations and experiments have implicated these species in the reaction mechanism of the CYPs and indicated that they may lead to reduced barriers compared to the triradicaloid Fe^IV state.

Many different quantum-mechanical (QM) methods have been used to study the geometry, electronic structure, and reactivity of the CYPs, ranging from approximate semiempirical methods to high-level multi-reference and coupled-cluster methods. However, the great majority of the calculations have been performed with density-functional theory (DFT). Calibration calculations have indicated that DFT, in particular with hybrid functionals such as B3LYP, give reliable structures and energies, except for the relative energy of the pentaradicaloid and Fe^V states of compound I and spin populations of this state. Most calculations have been performed for isolated heme models, either in a vacuum or in a homogeneous continuum solvent with a dielectric constant of 4–8. However, it is well-known that many properties, especially the spin populations and excitation energies of compound I are sensitive to the surroundings. Therefore, many studies have been performed with a combination of QM and molecular-mechanics (MM) methods (QM/MM), in which the heme group and the substrate (and possibly a few more active-site residues) are treated at the QM level, whereas the rest of the protein and some explicit water molecules are treated at the MM level. Such an approach is normally assumed to give more accurate results, even if it is more sensitive to details in the set-up of the calculation and it becomes harder to ensure that all states studied belong to the same local minimum of the surrounding protein. Finally, it should be noted that it has recently been observed that dispersion effects, which are missing in standard DFT calculations, can have large effects for structures and energies for metal complexes, in particular for the CYPs. Therefore, dispersion-corrected DFT, such as DFT-D, is recommended for future studies of the CYPs, even if caution is needed when a ligand binds or dissociates from the metal.

In this contribution, we review quantum-mechanical (QM) studies of reaction mechanisms of the CYPs. The focus is on the diversity of reactions catalyzed by the CYPs, with an emphasis on models to predict the reactive sites of an arbitrary molecule. Therefore, we discuss the various reaction types studied by QM methods in separate sections. Unless specified, all QM studies have been performed with DFT using a Fe(porphine)(RS)O model of compound I with R = H or CH_3 (porphine is a porphyrin ring without side chains), possibly using the QM/MM approach. Calculations with faster methods or smaller models of compound I are shortly described in the penultimate section.

### Hydroxylation of aliphatic carbon atoms

**Table XXXdelete.** Exhaustive listing of theoretical studies of aliphatic hydroxylation by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps</th>
<th>Spin</th>
<th>Method</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
</table>

---

---
<table>
<thead>
<tr>
<th>Substance</th>
<th>Reaction Type</th>
<th>Method</th>
<th>Year</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methane</td>
<td>Rebound step</td>
<td>d/q</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>Methane</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>Methane</td>
<td>Full mechanism and KIE’s</td>
<td>d/q</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>Methane</td>
<td>Full mechanism</td>
<td>d</td>
<td>2001</td>
<td></td>
</tr>
<tr>
<td>Methane</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>20 ligands</td>
<td>Hydrogen abstraction barrier</td>
<td>q</td>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>Methane</td>
<td>Barrier and water effect</td>
<td>d/q</td>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>10 substrates</td>
<td>Barriers</td>
<td>d/q</td>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>Methane, isobutane and camphor</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2009</td>
<td></td>
</tr>
<tr>
<td>Methane, isobutane and camphor</td>
<td>Barriers</td>
<td>d</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Ethane</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>Ethane</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>Ethane</td>
<td>Full mechanism, QMMD</td>
<td>d/q</td>
<td>2001</td>
<td></td>
</tr>
<tr>
<td>5 ligands (6 sites)</td>
<td>Barriers</td>
<td>d/q</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>8 ligands (10 sites)</td>
<td>Hydrogen abstraction barrier</td>
<td>d/q</td>
<td>2011</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2002</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Full mechanism and NH—S bonds</td>
<td>d/q</td>
<td>2002</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Full mechanism, electrostatic field effects</td>
<td>d/q</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Full mechanism</td>
<td>d/q/s</td>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>Propene and cyclohexene</td>
<td>Barriers</td>
<td>d/q</td>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Barriers and axial ligand effects</td>
<td>d/q</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Propene and cyclohexene</td>
<td>Barriers</td>
<td>q</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Camphor, propene and cyclohexene</td>
<td>Barriers</td>
<td>d/q</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Isobutane</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>Toluene</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>Camphor</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>Camphor</td>
<td>Hydrogen abstraction step</td>
<td>q</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>Camphor</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>Camphor</td>
<td>Full mechanism</td>
<td>q</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>Camphor</td>
<td>Full mechanism</td>
<td>q</td>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>Camphor</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2006</td>
<td></td>
</tr>
<tr>
<td>trans-2-phenyl-methyl</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>Substrate</td>
<td>Study Details</td>
<td>Spin States</td>
<td>Publication Year</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---------------------------------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Cyclopropane</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>trans-2-phenyl-isopropylcyclopropane</td>
<td>Hydrogen abstraction step</td>
<td>d</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>Methoxyfluorane</td>
<td>Hydrogen abstraction step</td>
<td>d</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>Morpholine</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2009</td>
<td></td>
</tr>
<tr>
<td>Testosterone (4 sites)</td>
<td>Barriers</td>
<td>d/q</td>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>Testosterone</td>
<td>Full mechanism (tunneling effects)</td>
<td>d/q</td>
<td>2009</td>
<td></td>
</tr>
<tr>
<td>N-Palmitoylglycine</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2009</td>
<td></td>
</tr>
<tr>
<td>Nicotine</td>
<td>Full mechanism</td>
<td>d/q</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td>Dimethylsulfide and dimethylsulfoxide</td>
<td>Hydrogen abstraction step</td>
<td>d/q</td>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>24 ligands</td>
<td>Hydrogen abstraction barrier</td>
<td>q</td>
<td>2010</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3.** Substrates for which aliphatic hydroxylation by CYP enzymes has been studied by DFT methods. It is indicated for which substrates the full reaction mechanism or only the optimization of the transition state has been studied. The numbers indicate references. Spin states studied are shown within brackets (d=doublet, q=quartet, s=sextet) and substrates studied by QM/MM are explicitly indicated. If there is more than one possible reaction site, the sites that have been studied are labeled by a *. 
The hydroxylation of \( sp^3 \) hybridized aliphatic carbon atoms occurs through a two-step mechanism. The first step is the abstraction of a hydrogen atom from the substrate carbon atom by the iron-bound oxygen atom in compound I, resulting in a substrate radical intermediate and protonated compound II. This is followed by a rebound step in which the carbon radical binds to the iron-bound oxygen, forming an alcohol, which is weakly bound to the heme Fe\(^{III} \) ion by the hydroxyl group (Figure 1).\(^{66} \) In all reactions studied so far, the activation barrier is higher for the hydrogen-abstraction step than for the rebound step, so many studies have concentrated on the former reaction step.

During the hydrogen-abstraction step, the doublet and quartet spin states of compound I give similar energies and structures, whereas they diverge in the rebound step. For the doublet spin state, there is rarely a barrier for the rebound step, whereas a small barrier is found for most substrates in the quartet spin state. This difference in the rebound step has been used to explain the puzzling experimental results found using radical-clock substrates,\(^{56} \) validating that two competing spin states participate in the CYP reactions (two-state reactivity). This barrier difference for the two spin states occurs because during the rebound step, one electron is transferred from the substrate radical to the heme group. In the doublet spin state this electron is transferred to a low-lying orbital distributed throughout the porphyrin ring and shared with the sulfur atom (usually called \( a_{\text{S}} \)), whereas in the quartet spin state, the electron needs to occupy a high-lying iron \( d \) orbital (\( \sigma^*(d_{z^2}) \)) in order to retain a quartet spin state. This high-lying \( d \) orbital is anti-bonding along the \( O-Fe-S \) axis, which causes the rebound transition state to have a more “open” structure in the quartet spin state. A more detailed description of this rebound step and its implications on the rearrangement of radical clocks can be found in a review by Shaik.\(^{67} \)

The hydrogen abstraction and rebound mechanism has also been studied for pentaradicaloid quartet and sextet spin states.\(^{8,21} \) While these states are higher in energy, their reaction barriers are of similar size as the triradicaloid doublet and quartet spin states, and especially during the rebound step they can potentially contribute to the hydroxylation of aliphatic carbon atoms. In a QM/MM study of cyclohexene hydroxylation, the pentaradicaloid rebound barriers, were shown to be smaller than the barrier for the triradicaloid quartet state,\(^4 \) indicating that the enzyme can potentially tune the stereoselectivity of hydroxylations by populating these states in the protonated compound II intermediate.

This and other mechanisms for the hydroxylation of \( sp^3 \) hybridized aliphatic carbon atoms have been studied extensively during the last decade and have been reviewed several times. Therefore, we refer the reader to two reviews by Shaik and coworkers.\(^3,4 \)

**Figure 4.** Distribution of energy barriers (in kJ/mol) for the mechanisms shown in Figure 1. The energies were obtained in vacuum with the B3LYP method (using PBE geometries for dehalogenation) and large basis sets, and are presented relative to separated compound I and the substrates.
A large number of different substrates have been studied, as is shown in Figure 3. The distribution of hydrogen abstraction barrier energies is quite large, 33–109 kJ/mol (Figure 4). One could imagine that some of the substrates that have high barriers are unlikely to be hydroxylated, but it has been shown that the CYPs can hydroxylate even methane, which is the substrate with the highest hydrogen-abstraction barrier studied so far. The hydrogen-abstraction activation energies are strongly correlated to the C–H bond strength, but the barrier cannot be directly predicted from the bond strength, because the stability of the radical intermediates also contributes (the radical can be stabilized by distributing it over several atoms). The activation energies are mainly governed by the substituents next to the hydroxylated carbon atoms, roughly in the following order: hydrogen (methane) > \( sp^3 \) carbon > acetamide carbonyl > carbonyl next to phenyl = sulfonamide > carboxylic acid = phenyl \( sp \) carbon double bond > aromatic 5-ring = carbon triple bond = conjugated carbon double bond > two \( sp^3 \) carbon atoms > three \( sp^3 \) carbon atoms.

Recently, it has been shown that the addition of a water molecule to the QM system lowers the activation barriers of the hydrogen-abstraction step in model-system calculations if it forms a hydrogen bond to the iron-bound oxygen atom. However, since similar calculations have neither been performed with any other reaction types, nor in a systematic way in QM/MM calculations, we cannot say if this is a general feature of CYP reactions or a specific feature for the abstraction of a hydrogen atom from an aliphatic carbon atom.

Moreover, a recent study by Lonsdale et al. highlighted the importance of dispersion interactions for the CYP reactions. They showed that for the 5-hydroxylation of camphor and the hydroxylation of cyclohexene, the geometries and the activation energies change when adding an empirical dispersion correction to the B3LYP functional (DFT-D). The dispersion correction decreases the distances between the substrate and the heme group. This affects the energies, especially for the transition state, in which the substrate–heme distance is shorter than in the reactant complex, resulting in lower activation barriers. However, while the geometries changed significantly, the
changes of activation energies were less than 8 kJ/mol for the various hydroxylation reactions. They also showed that the dispersion-corrected relative activation energies of hydroxylation and epoxidation for propene and cyclohexene were closer to experimental data.

**Dealkylations**

Table XXX delete . Exhaustive listing of theoretical studies of dealkylation by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>N,N-dimethylaniline</td>
<td>Hydroxylation step</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2006</td>
<td>70</td>
</tr>
<tr>
<td>Dimethylamine, trimethylamine, N-methylaniline, and N,N-dimethylaniline</td>
<td>Hydrogen abstraction barrier</td>
<td>Quartet</td>
<td>Model system</td>
<td>2006</td>
<td>32</td>
</tr>
<tr>
<td>N,N-dimethylaniline</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2007</td>
<td>71</td>
</tr>
<tr>
<td>N,N-dimethylaniline</td>
<td>H-abstraction barrier</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2008</td>
<td>74</td>
</tr>
<tr>
<td>N,N-dimethylaniline and trimethylamine</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2009</td>
<td>72</td>
</tr>
<tr>
<td>N,N-dimethylaniline and N,N-dimethylbenzamide</td>
<td>Hydroxylation step</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2010</td>
<td>73</td>
</tr>
<tr>
<td>N,N-dimethylaniline</td>
<td>Hydrogen abstraction barrier</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2011</td>
<td>41</td>
</tr>
<tr>
<td>Trimethylamine</td>
<td>Hydroxylation step</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2008</td>
<td>63</td>
</tr>
<tr>
<td>Theophylline</td>
<td>Full mechanism</td>
<td>Doublet/quartet?</td>
<td>Model system</td>
<td>2007</td>
<td>74</td>
</tr>
<tr>
<td>N-cyclopropyl-N-methylaniline</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2009</td>
<td>75</td>
</tr>
<tr>
<td>Nicotine</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2010</td>
<td>62</td>
</tr>
<tr>
<td>5-methoxytryptamine</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2010</td>
<td>76</td>
</tr>
<tr>
<td>p-[Cl, NO2, CN]dimethylaniline (3 anilines)</td>
<td>Hydroxylation step</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2007</td>
<td>71</td>
</tr>
<tr>
<td>2,3,4,5,6-pentafluoro-N,N-dimethylaniline</td>
<td>Full mechanism</td>
<td>Doublet</td>
<td>Model system</td>
<td>2010</td>
<td>77</td>
</tr>
<tr>
<td>22 ligands (24 sites)</td>
<td>Hydrogen abstraction barrier</td>
<td>Quartet/doublet</td>
<td>Model system</td>
<td>2010</td>
<td>64,65</td>
</tr>
<tr>
<td>Caffeine (3 sites)</td>
<td>Hydrogen abstraction barrier</td>
<td>q</td>
<td>QM</td>
<td>2009</td>
<td>78</td>
</tr>
</tbody>
</table>

* doublet spin state only computed for some ligands.
**Figure 5.** Substrates for which dealkylation by CYP enzymes has been studied by DFT methods. The Figure is constructed the same way as Figure 3.
Full reaction mechanism

Hydroxylation mechanism

Hydrogen abstraction transition state
When the aliphatic hydroxylation takes place on a carbon atom that is bound to an oxygen or nitrogen atom, the resulting products (gem-diols, carbinols, or carbinolamines) are chemically unstable and break down by a water-mediated non-enzymatic reorganization step. If the reactive carbon atom is not in an alcohol group, the reorganization results in an O- or N-dealkylation (Figure 1). Consequently, from a theoretical perspective, this reaction is identical to the aliphatic hydroxylation, described in the previous section, i.e. with hydrogen-abstraction and rebound steps, of which the first is rate limiting. Because of the importance of understanding these reaction types in drug degradation, a number of theoretical studies have been performed, in several cases including the non-enzymatic reaction, as are summarized in Figure 5.

There is a relatively large variation in the activation energies for different dealkylation reactions (32–80 kJ/mol, cf. Figure 4), but they follow, with few exceptions, quite simple rules, based on the reacting group.

**Figure 6.** Dependence of the activation energies for dealkylation reactions on the structure of the reacting group.

**Epoxidation of alkene carbon atoms**

Table XXX delete. Exhaustive listing of theoretical studies of alkene epoxidation by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethene</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2001</td>
<td>79</td>
</tr>
<tr>
<td>Ethene</td>
<td>Inactivation</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2001</td>
<td>80</td>
</tr>
<tr>
<td>Ethene</td>
<td>Alternative mechanism through synchronous oxygen transfer</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2001</td>
<td>81</td>
</tr>
<tr>
<td>Ethene</td>
<td>Full mechanism and non-compound I oxidants</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2002</td>
<td>82</td>
</tr>
<tr>
<td>Ethene</td>
<td>Full mechanism</td>
<td>quartet</td>
<td>Model system</td>
<td>2003</td>
<td>83</td>
</tr>
<tr>
<td>Ethene and propene</td>
<td>Rebound step</td>
<td>Quartet</td>
<td>Model system</td>
<td>2004</td>
<td>67</td>
</tr>
<tr>
<td>Ethene and styrene</td>
<td>Full</td>
<td>Doublet/quartet</td>
<td>Model</td>
<td>2004</td>
<td>84</td>
</tr>
<tr>
<td>Substrate</td>
<td>Reaction Type</td>
<td>Methodology</td>
<td>Year</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>------------------------------------</td>
<td>-------------------------</td>
<td>-------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Ethene and 2-butene</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>2008</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>7 ligands</td>
<td>Oxidation step</td>
<td>Doublet/quartet</td>
<td>2010</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>2002</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Full mechanism and NH—S bonds</td>
<td>Doublet quartet</td>
<td>2002</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Full mechanism, electrostatic field effects</td>
<td>Doublet/quartet</td>
<td>2004</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Propene</td>
<td>Full mechanism</td>
<td>Doublet/quartet/sextet</td>
<td>2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propene and cyclohexene</td>
<td>Barriers</td>
<td>Doublet/quartet</td>
<td>2006</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Propene and cyclohexene</td>
<td>Barriers</td>
<td>quartet</td>
<td>2010</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Camphor, propene and cyclohexene</td>
<td>Barriers</td>
<td>Doublet/quartet</td>
<td>2010</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Styrene</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 ligands (multiple sites on some ligands)</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>2010</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>5-methylenenylcamphor</td>
<td>Full mechanism and alternative oxidant</td>
<td>Doublet/quartet</td>
<td>2006</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>Oxidation step</td>
<td>Doublet/quartet</td>
<td>2008</td>
<td>89</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 7.** Substrates for which alkene epoxidation by CYP enzymes has been studied by DFT methods. The Figure is constructed the same way as Figure 3.
The epoxidation of alkene double bonds by the CYPs follow a two-step mechanism as shown in Figure 1. The first step involves the formation of a tetrahedral intermediate with a bond between one of the two carbon atoms in the reactive double bond and the iron-bound oxygen atom and a radical at the other atom in the double bond. In the second step, an epoxide is formed by ring closure. The formation of the tetrahedral intermediate is the rate-limiting step.

The epoxidation of alkene double bonds was one of the first CYP reactions studied by DFT methods already in 2001. Since then, a large number of alkenes have been studied, as is summarized in Figure 7. The studies have compared different competing reaction mechanisms (e.g. epoxidation vs. aliphatic hydroxylation), but also the formation of byproducts and the inactivation of the heme group by the formation of a suicidal complex, as described in Figure 8.

The doublet and quartet spin states give similar activation energies for the first reaction step. The ring closure has been shown to be barrierless for the doublet spin state in all studies except one, whereas the quartet spin state gives rise to either a very small or no barrier. Results from these QM calculations have been reviewed several times and we refer the interested reader to two reviews by Shaik and coworkers for a more detailed analysis of the earlier results. Recent QM/MM work on the epoxidation of propene and cyclohexene, performed by two different research groups, show varying effects of including the protein surroundings. Since both groups study the reactions in P450cam, but with different setups and different software, it seems that more extensive investigations are needed to clarify how the protein surroundings affect epoxidation reactions performed by CYPs.

Figure 8. The formation of a suicidal complex during alkene epoxidation.
As is shown in Figure 4, the range of activation energies for alkene epoxidation is quite large (29–82 kJ/mol). Interestingly, both the highest and lowest activation energies are found for molecules with conjugated double bonds (pentadiene and heptatriene), but for different atoms, with the lowest energies arising from the terminal carbon atoms.

As mentioned above, recent work by Lonsdale et al. investigated the effects of dispersion-corrected DFT on the epoxidation of propene and cyclohexene. Also for the epoxidation reactions, the activation energies of the two substrates changed in the same way, indicating that previous work studying relative energies for reactions of a single reaction type with DFT without dispersion corrections are likely to still be reliable, while reactions of different types might be affected differently by including dispersion corrections.

**Oxidations of aromatic carbon atoms**

Table XXX delete . Exhaustive listing of theoretical studies of aromatic oxidation by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>Full mechanism and NIH-shift</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2003</td>
<td>90</td>
</tr>
<tr>
<td>8 ligands</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2003</td>
<td>91</td>
</tr>
<tr>
<td>16 ligands (23 sites)</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2004</td>
<td>92</td>
</tr>
<tr>
<td>Benzene</td>
<td>Full mechanism</td>
<td>Doublet</td>
<td>Model system</td>
<td>2008</td>
<td>93</td>
</tr>
<tr>
<td>11 ligands (17 sites)</td>
<td></td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2011</td>
<td>94</td>
</tr>
<tr>
<td>Toluene</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2007</td>
<td>95</td>
</tr>
<tr>
<td>(S)-N-[1-(3-morpholin-4-ylphenyl)ethyl]-3-phenylacrylamide</td>
<td>Oxidation barrier</td>
<td>Doublet</td>
<td>Model system</td>
<td>2006</td>
<td>96</td>
</tr>
<tr>
<td>(S)-N-[1-(3-morpholin-4-ylphenyl)ethyl]-3-phenylacrylamide</td>
<td>Full mechanism after formation of tetrahedral complex</td>
<td>Doublet</td>
<td>Model system</td>
<td>2007</td>
<td>96</td>
</tr>
<tr>
<td>Dextromethorphan</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>QM/MM</td>
<td>2011</td>
<td>97</td>
</tr>
<tr>
<td>27 ligands (59 sites)</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2010</td>
<td>64,65</td>
</tr>
<tr>
<td>Anisole</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2011</td>
<td>97</td>
</tr>
<tr>
<td>Carbamazepine (2 sites)</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2008</td>
<td>99</td>
</tr>
<tr>
<td>4 substrates (6 sites)</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>QM</td>
<td>2009</td>
<td>78</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>6 ligands (17 sites)</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2008</td>
<td>85</td>
</tr>
</tbody>
</table>

**Figure 9.** Substrates for which aromatic oxidation by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.
The oxidation of aromatic carbon atoms by the CYPs has been studied extensively during the last decade. However, there is still no consensus on the mechanism for the formation of all products. The first step in the oxidation is the formation of a tetrahedral intermediate, as shown in Figure 1. This can be followed by either an NIH-shift involving one of the heme nitrogen atoms or the formation of an epoxide (which is not necessarily stable). The NIH-shift can lead either to the formation of a phenol or a ketone, but there are multiple possible reaction paths and the orientation of the substrate relative to the heme ring can affect the path. For example, face-on and side-on approaches of benzene can lead to different products, and the formation of the tetrahedral intermediate has a lower activation energy for the side-on approach (by 10 kJ/mol for benzene, structures shown in Figure 10). A more detailed discussion of these mechanisms is given in a recent review. Still, all studies so far indicate that the formation of the tetrahedral intermediate is the rate-limiting step. Hence, many studies have focused on this step. As is the case for the alkene epoxidation, the two spin states give similar energies during the formation of the tetrahedral intermediate and depending on the substrate, either the doublet or the quartet spin state may give the lowest barrier. For the various reaction mechanisms that follow the formation of the tetrahedral intermediate, the doublet spin state has given lower barriers in all studies performed until now, viz. NIH-shift, alcohol formation, ketone formation, and epoxide formation.

**Figure 10.** Face-on and side-on approaches.

Recent work by Oláh et al., employing a QM/MM methodology, showed that through steric hindrance the protein surroundings can raise the barrier of aromatic oxidations significantly. They showed that this is the explanation for the preference for O-dealkylation vs. the potential aromatic oxidation in the metabolism of dextromethorphan by CYP2D6.

A vast number of aromatic systems have been studied by QM methods, as is shown in Figure 9, with activation energies ranging from 50 to 94 kJ/mol (Figure 4). The most reactive aromatic carbon atoms are those next to oxygen or nitrogen atoms in aromatic five-ring systems (e.g. in furan, pyrrole, and imidazole). The least reactive aromatic carbon atoms are either those in six-rings containing nitrogen atoms (pyridine or pyrimidine) or with sulfonamide substituent in the ortho position.

**Oxidations of primary and secondary amine nitrogen atoms**

Table XXX delete. Exhaustive listing of theoretical studies of oxidations of primary and secondary amine nitrogen atoms by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
</table>


8 ligands | Direct oxidation barrier | Doublet | Model system | 2010 | 64, 65
---|---|---|---|---|---
1,4-dihydropyridine | Hydrogen-abstraction barrier | Doublet/quartet | Model system | 2010 | 65
Propan-2-amine | Direct oxidation and hydrogen abstraction | Doublet/quartet/sextet | Model system | 2011 | 98

**Figure 11.** Substrates for which hydroxylation of primary or secondary amine nitrogen atoms by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.

**Full mechanism for direct oxidation and hydrogen abstraction mechanisms**

- **Direct oxidation transition state**
  - NH$_2$
  - 64, 98 [d/q/s]

- **Hydrogen abstraction transition state**
  - 65 [d/q]

On the basis of experiments, three possible reaction mechanisms have been suggested for the hydroxylation of primary and secondary amines to hydroxylamine: direct oxygen transfer (addition of the oxygen to the nitrogen lone pair, followed by a rearrangement of the formed N-oxide to hydroxylamine), hydrogen abstraction from the nitrogen followed by a rebound step, and direct insertion of the oxygen into the N–H bond. Either of these reaction mechanisms could potentially be preceded by an electron transfer from the substrate nitrogen atom to the heme. However, so far there is no consistent experimental or theoretical proof, neither for nor against this hypothesis.

The direct insertion mechanism has been shown to be unlikely for primary alkylamines$^{98,99}$ and aniline nitrogen atoms (unpublished DFT data), but other work has suggested that this mechanism should be more probable for nitrogen atoms with a highly delocalized lone pair.$^{100}$ So far, only a single computational study has considered both the direct oxygen-transfer mechanism and the hydrogen abstraction mechanism for the same substrate (cf. Figures 1 and 11),$^{98}$ whereas earlier work only studied one of the two mechanisms.$^{64, 65}$ In the comparative study, the hydroxylation of propan-2-amine was studied in three different spin states and with or without continuum solvation models. The results show that the two mechanisms (direct oxygen transfer and hydrogen abstraction) are highly competitive and that changes in the dielectric continuum constant can change the preference from one reaction mechanism to the other.$^{98}$ Hence, it is likely that an enzyme can tune which mechanism is used depending on substrate properties and interactions with...
both catalytic water molecules and amino acids in the active site. Furthermore, for 1,4-dihydropyridine, our attempts to locate the transition state of the direct oxygen-transfer mechanism failed, owing to a spontaneous hydrogen abstraction from the nitrogen atom. This shows that the hydrogen-abstraction mechanism is more likely for highly conjugated nitrogen atoms in primary and secondary amines, whereas for other types of primary and secondary amines, it is not clear which mechanism occurs.

While the range of activation energies found so far for the hydroxylation of primary and secondary amines is large (0–95 kJ/mol; Figure 4), it is hard to determine the relevance of the computed values because there are multiple possible mechanisms. The hydrogen-abstraction of 1,4-dihydropyridine is spontaneous in the quartet spin state (the energy of the transition state is smaller than the sum of the energies of the separate compound I and substrate). The highest activation energy obtained is for the direct oxidation of methansulfonamide.

**Oxidations of tertiary amine nitrogen atoms**

Table XXX delete . Exhaustive listing of theoretical studies of oxidations of tertiary amine nitrogen atoms by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trimethylamine</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2008</td>
<td>63,72</td>
</tr>
<tr>
<td>N,N-dimethylaniline</td>
<td>Full mechanism</td>
<td>Doublet/quartet/sextet</td>
<td>Model system</td>
<td>2007</td>
<td>72,101</td>
</tr>
<tr>
<td>2,3,4,5,6-pentafluoro-N,N-dimethylaniline</td>
<td>Full mechanism</td>
<td>Doublet</td>
<td>Model system</td>
<td>2010</td>
<td>77</td>
</tr>
<tr>
<td>6 ligands</td>
<td>Oxidation barrier</td>
<td>Doublet</td>
<td>Model system</td>
<td>2010</td>
<td>64</td>
</tr>
</tbody>
</table>

**Figure 12.** Substrates for which oxidation of tertiary amine nitrogen atoms by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.
Only a few studies of the oxidation of tertiary amine nitrogen atoms have been published, all during the latest years.\textsuperscript{63,72,77,101} The mechanism is a direct oxygen transfer (Figure 1) and it occurs in the doublet spin state, which has significantly lower activation energies than the quartet and sextet spin states. The reaction is reversible, as N-oxides of anilines (and possibly also other tertiary amines) can interact with the CYP resting Fe\textsuperscript{III} state and form compound I by transferring the amine oxygen atom back to the heme iron atom.\textsuperscript{102}

While these calculations, together with our recent work on oxidation barriers, form a small data set (Figure 12), the experimentally observed metabolism of tertiary amines is hard to explain. Tertiary amines with only \textit{sp}\textsuperscript{3} carbon substituents can result in both N-oxidation and dealkylations. For example, oxidation of piperidines never results in N-oxides, whereas bicyclic amines such as Zatosetron give the N-oxide as the major product (Figure 13), but these results are hard to reproduce by DFT calculations.\textsuperscript{103} More QM studies of the nitrogen oxidation process are needed to fully understand which tertiary amines lead to experimentally observed N-oxides.

**Figure 13.** The piperidine fragment (left) never reacts by N-oxidation, whereas the bicyclic fragment (right) results in N-oxidation.

Among the substrates that have been studied so far, the activation energies range from 40 to 94 kJ/mol (see Figure 4). Simple alkane-amines give the lowest activation energies (e.g. trimethylamine), whereas acetamide nitrogen atoms give rise to the highest activation energies.\textsuperscript{64,65}
Oxidation of aromatic nitrogen atoms

Table XXX delete . Exhaustive listing of theoretical studies of oxidation of aromatic nitrogen atoms by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 ligands</td>
<td>Oxidation barrier</td>
<td>Doublet</td>
<td>Model system</td>
<td>2010</td>
<td>64</td>
</tr>
</tbody>
</table>

**Figure 14.** Substrates for which oxidation of aromatic nitrogen atoms by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.

**Oxidation transition state**

![Chemical structures](image)

The mechanism for the oxidation of aromatic nitrogen atoms is also a direct transfer of the iron-bound oxo atom of compound I to the substrate nitrogen atom (Figure 1). Until now, the transition states for the nitrogen oxidation of only seven aromatic nitrogen compounds have been studied, viz. four oxidations of nitrogen atoms in six-membered rings and three oxidations of nitrogen atoms in five-membered rings (Figure 14). The calculations were performed for the doublet spin state, assuming that this is the lowest-lying spin state, as have been found for other types of nitrogen oxidations, with the aromatic ring in a side-on geometry (as described in Figure 10). The activation energies are quite high compared to other reaction mechanisms, varying between 71 and 102 kJ/mol (see Figure 4).

In experiments, aromatic N-oxides formed by CYPs have been found for a large number of compounds, primarily with pyridine, pyrimidine, and quinoline fragments. However, no N-oxides of nitrogen atoms in five-membered aromatic rings have been found. Instead, such nitrogen atoms tend to bind directly to the heme iron atom, leading to CYP inhibition.

For four of the seven tested substrates, we can directly compare the oxidation barriers of the nitrogen atoms to those of their neighboring carbon atoms and to experimental data. In pyridine, these carbon atoms give rise to higher activation energies (91 kJ/mol) than the nitrogen atom (71 kJ/mol). For 6-methoxy-quinoline, demethylation of the methoxy group is the major reaction, but both the N-oxide and 6-methoxy-3-quinolinol are formed in small amounts. The activation energy of the N-oxidation is 73 kJ/mol, which is lower than for aromatic carbon atoms in nitrogen-containing six-membered rings (84–93 kJ/mol), explaining why no aromatic oxidation of the carbon atoms near the nitrogen atom is seen. The demethylation has been shown to have a barrier of 68 kJ/mol (for a similar anisole demethylation), explaining why this is the major product. In imidazole
rings, the carbon atom in between the two nitrogen atoms has a much lower activation energy than the nitrogen atoms (55 vs. 86 kJ/mol),\textsuperscript{64} which also is reflected in the metabolites that are formed in experiment: Nitrogen oxidations are never observed, but carbon oxidation is found in several cases.\textsuperscript{113-115} In isothiazole, the oxidation of the sulfur atom gives the lowest activation energy (67 kJ/mol vs. 102 kJ/mol for the nitrogen oxidation) and this is also reflected in experiment, in which the sulfoxide is formed during metabolism of ziprasidone\textsuperscript{116} and perospirone.\textsuperscript{117}

### Oxidation of aliphatic alcohols

The oxidation of alcohols has been studied by Wang et al.\textsuperscript{118} They showed that the formation of a gem-diol, followed by dehydration, which is the expected mechanism (Figure 1, reaction A), gives a reasonable barrier. However, a reversed dual hydrogen-abstraction mechanism (Figure 1, reaction B) has similar energy barriers.

The hydrogen-abstraction step gives activation energies of 43 (49) kJ/mol for the doublet (quartet) spin state. The rebound step was found to be barrierless in both spin states and results in the gem-diol product. The aldehyde product is then formed by a dehydration of the gem-diol. This dehydration on the enzyme model gave large activation energies, 178 and 138 kJ/mol for the doublet and quartet states, respectively. However, the non-enzymatic dehydration mechanism was not considered, so the role of the enzyme in the dehydration is uncertain.

The reversed dual hydrogen-abstraction also starts with a hydrogen abstraction, but from the hydroxyl group, rather than from the carbon atom, leading to an intermediate with an oxo radical. This step is rate-determining, with similar barriers for the doublet and quartet spin states (51 and 53 kJ/mol, respectively). This is followed by a second hydrogen abstraction from the alcohol carbon atom by the iron–hydroxyl group (protonated compound II), leading to the final aldehyde product. The second hydrogen abstraction has a lower barrier in the doublet spin state than in the quartet spin state (41 vs. 61 kJ/mol).

Wang et al.\textsuperscript{118} also showed that in a polar environment (simulated by an implicit solvent model), the reversed dual hydrogen-abstraction mechanism is more favorable than the gem-diol mechanism. This suggests that the choice of mechanism can depend on the polarity of the active site and perhaps different CYP isoforms use different mechanisms for the oxidation of alcohols. It may explain why kinetic isotope effects do not reflect $k_{cat}$ values in CYP 2E1.

Only one additional study of alcohol oxidation has been published, in which the rate-determining step of the gem-diol mechanism for the oxidation of isopropanol was studied.\textsuperscript{64} Due to the small amount of available data, we cannot say much about the variation of activation energies in alcohol oxidation for different substrates.

### Aldehyde oxidation

Table XXX delete . Exhaustive listing of theoretical studies of aldehyde oxidation by CYP enzymes.
**Figure 15.** Substrates for which aldehyde oxidation by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.

**Full reaction mechanism**

\[
\text{CHO} + \text{O} \rightarrow \text{CH}_2 \text{O}
\]

119, 120 [d/q]  120 [d/q]

**Hydrogen abstraction transition state**

\[
\begin{align*}
\text{O} & \quad \text{O} \\
64 \ [d/q] & \quad 120 \ [d/q] \\
120 \ [d/q] & \quad 120 \ [d/q] \\
120 \ [d/q] & \quad 120 \ [d/q]
\end{align*}
\]

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaldehyde</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2006</td>
<td>119</td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2007</td>
<td>120</td>
</tr>
<tr>
<td>30 aldehydes</td>
<td>Hydrogen abstraction</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2007</td>
<td>120</td>
</tr>
<tr>
<td>3 aldehydes</td>
<td>Hydrogen abstraction</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2010</td>
<td>64</td>
</tr>
</tbody>
</table>
Oxidation of aldehydes into carboxylic acids is a common CYP-mediated reaction. There are two major reaction pathways in the oxidation of aldehydes, direct oxidation to a carboxylic acid and deformylation. The direct oxidation pathway is most likely a compound I mediated reaction, whereas the deformylation reaction most likely is mediated by the Fe$^{III}$-peroxo intermediate.\textsuperscript{121-123} Most aldehydes that are CYP substrates undergo direct oxidation to carboxylic acids.\textsuperscript{124} However, branching at the $\alpha$ carbon of the aldehyde seems to induce deformylation and branching at the $\beta$ carbon can also lead to some degree of deformylation.\textsuperscript{121}

Experimental evidence suggests that the direct oxidation to a carboxylic acid follows a hydrogen-abstraction and rebound mechanism (Figure 1),\textsuperscript{119} whereas the deformylation reaction has a branching point after the deformylation at which there are three possible products for the remainder of the substrate, an olefin, alcohol, or inactivation of the CYP by formation of a heme adduct.\textsuperscript{123}

Quantum chemical studies have shown that for the formation of carboxylic acids, the hydrogen-abstraction step is rate-determining as usual. The doublet and quartet spin state give similar energies for all reaction steps, but depending on the substrate, either the doublet or the quartet spin states may give the lowest activation energy.\textsuperscript{119,120} Depending on the cysteine model used, the rebound step can either be barrierless (SH) or have a minor barrier (SCH$_3$).

The range of the activation energies found among 30 studied aldehydes (structures in Figure 15) is quite small, 33–49 kJ/mol (see Figure 4). Although the hydrogen atom is abstracted from a $sp^2$ hybridized carbon, the reactivity seems to follow the same pattern as for aliphatic hydroxylation.\textsuperscript{119} The least reactive aldehyde was $p$-nitro-benzaldehyde and the most reactive one was 2-phenylpropionaldehyde.\textsuperscript{120} A good correlation between the hydrogen-atom bond-dissociation energies of the substrates and their activation energies was also found ($r^2 = 0.75$).

**Figure 16.** The suggested mechanism for the deformylation of lanosterol carboxaldehyde.

Hackett and coworkers have used QM/MM calculations to study several possible mechanisms for the deformylation performed by sterol 14$\alpha$-demethylase.\textsuperscript{125,126} They showed that there are two likely mechanisms, one concerted and one stepwise, of which the stepwise mechanism was energetically more feasible. This mechanism (shown in Figure 16) starts with the terminal peroxo oxygen atom attacking the aldehyde carbon atom in lanosterol carboxaldehyde. This is followed by a cleavage of the C–C bond between the aldehyde carbon atom and the remainder of the substrate, leaving the substrate with a carbanion. After than, peroxo O–O bond is cleaved, which in concert with an electron transfer leads to the formation of compound I. Finally, a hydrogen atom from the substrate is transferred to the compound I oxygen atom, resulting in the formation of a double bond in the substrate.
Hackett et al. have also studied the final catalytic step in cytochrome P450 aromatase. They investigated several different possible reaction mechanisms and suggested that this aromatization and deformylation mechanism is slightly different than the one for lanosterol carboxaldehyde. The first two steps are basically the same, but the O–O peroxo bond is cleaved before formaldehyde is dissociates from the substrate. This is followed by a hydrogen abstraction from the substrate, which directly leads to the observed products.

**Desulphurization of phosphor atoms**

Table XXX delete. Exhaustive listing of theoretical studies of desulfurization of phosphor atoms by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>O,O,O-trimethyl phosphorothioate</td>
<td>Formation of epoxide-like intermediate</td>
<td>Doublet</td>
<td>Model system</td>
<td>2010</td>
<td>64</td>
</tr>
</tbody>
</table>

The CYPs can perform the desulfurization of parathion, chlorpyrifos, and other compounds with a phosphine sulfide group. In this reaction, the sulfur atom is replaced by the oxygen atom of compound I as is shown in Figure 1. However, the details of the mechanism are somewhat unclear. The formation of a triangular P–O–S complex is believed to be the initial step, possibly initiated by the formation of a linear P–S–O intermediate, which then leads to a bifurcation of the pathway, resulting in multiple products whose formation all are spontaneous processes (Figure 1). 127,128 However, our study of O,O,O-trimethyl phosphorothioate indicates that there is no stable linear intermediate during the formation of the triangular complex. Moreover, the formation of the triangular complex is virtually spontaneous: A transition state with an activation energy of 12 kJ/mol was found with a small basis set, but when the energy was recomputed with a larger basis set, the barrier disappeared.64 The steps following the formation of this triangular complex have not been studied yet.

**Oxidation of sulfide sulfur atoms**

Table XXX delete. Exhaustive listing of theoretical studies of oxidations of sulfide sulfur atoms by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 ligands</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2011</td>
<td>129</td>
</tr>
<tr>
<td>Dimethylsulfide</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2003-2008</td>
<td>63,130-132</td>
</tr>
<tr>
<td>Dimethylsulfide</td>
<td>Oxidation barrier</td>
<td>Doublet/quartet</td>
<td>QM/MM</td>
<td>2009</td>
<td>133</td>
</tr>
<tr>
<td>4 different thioanisoles</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2010</td>
<td>134</td>
</tr>
<tr>
<td>8 ligands</td>
<td>Oxidation</td>
<td>Doublet</td>
<td>Model system</td>
<td>2010</td>
<td>64</td>
</tr>
</tbody>
</table>
**Figure 17.** Substrates for which oxidation of sulfide sulfur atoms by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.

**Full reaction mechanism**

The first DFT study of the oxidation of a sulfide sulfur was performed already in 2003, but several studies were required to determine all the details of the mechanism. To date, 15 different sulfide oxidations have been studied (Figure 17). The sulfur oxidation mechanism is a direct transfer of the iron-bound oxygen atom to the substrate sulfur atom giving a sulfoxide.

Calculations using compound I model systems show a clear preference for sulfoxidation in the doublet spin state compared to the quartet spin state. This can be explained by the fact that in the high-spin process, an electron must be shifted to an energetically high-lying orbital to make the reaction possible. However, a recent QM/MM study by de Visser et al. indicates that enzymes (P450\textsubscript{CAM} and P450\textsubscript{BM3}) can change the preference so that the quartet state becomes lower in energy. This change in spin-state preference seems to be induced by the large structural differences of the doublet and quartet transition states, which enables the protein environment to affect the shape of the potential energy surfaces differently for the two spin states.

In a study of eight different substrates, the activation energies ranged from 24 to 49 kJ/mol (Figure 4). The lowest activation energies were found for sulfur compounds with two alkane substituents, whereas conjugation by double-bond substituents (vinyl, phenyl, etc.) gives higher energies, and the highest activation energy was found for dibenzo[b,f]thiophene, in which the sulfur atom is bound to two aromatic rings. Shaik et al. have shown that substituents to the aromatic ring can shift the activation energy in the sulfoxidation of anisoles, and that para substitution with a methoxy group gives the lowest activation energy, whereas para substitution with a nitro group gives the highest activation energy.
Oxidation of sulfoxide sulfur atoms

Table XXX delete . Exhaustive listing of theoretical studies of oxidation of sulfoxide sulfur atoms by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimethylsulfoxide</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2008</td>
<td>63</td>
</tr>
<tr>
<td>7 ligands</td>
<td>Oxidation barrier</td>
<td>Doublet</td>
<td>Model system</td>
<td>2010</td>
<td>64</td>
</tr>
</tbody>
</table>

**Figure 18.** Substrates for which sulfoxide oxidation by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.

**Full reaction mechanism**

![Reaction Mechanism](image)

**Oxidation transition state**

![Transition State](image)

The full reaction mechanism of the oxidation of a sulfoxide to a sulfone has only been studied once, viz. for the oxidation of dimethylsulfoxide. The results are quite similar to the sulfoxidation of dimethyl sulfide, involving a direct transfer of the iron-bound oxo group to the substrate, but with significantly higher barriers compared to the reactant complex. However, if the activation energies are compared to the isolated compound I and substrate, the difference is only 4 kJ/mol. This can probably partly be attributed to the lack of a proper description of dispersion interactions in DFT.

So far the oxidation barrier has been computed for only eight different sulfoxides (Figure 18). The range of activation energies (Figure 4) is only 7 kJ/mol if one excludes 2-(ethylsulfinyl)-1H-benzo[d]imidazole (a model of omeprazole) which has a 12 kJ/mol larger activation energy. The trends are similar to the oxidations of sulfurs to sulfoxides, but the substituent effects are much smaller.

Oxidations of aromatic sulfur atoms

Table XXX delete . Exhaustive listing of theoretical studies of oxidation of aromatic sulfur atoms by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state(s)</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 ligands</td>
<td>Oxidation barrier</td>
<td>Doublet</td>
<td>Model system</td>
<td>2010</td>
<td>64</td>
</tr>
<tr>
<td>Benzothiophene</td>
<td>Oxidation barrier</td>
<td>Doublet</td>
<td>Model system</td>
<td>2011</td>
<td>This work</td>
</tr>
</tbody>
</table>

**Figure 19.** Substrates for which oxidation of aromatic sulfur atoms by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.
The mechanism for the oxidation of aromatic sulfur atoms is the same as for the oxidation of sulfide sulfur atoms, i.e. a direct oxygen transfer. So far, transition states for the oxidation of only four ligands have been studied (Figure 19). All were studied in the doublet state with a side-on geometry (as described in Figure 10).

The activation energies for these substrates are all higher than those of sulfide and sulfoxide compounds, ranging from 57 to 74 kJ/mol (Figure 4). The barriers follow the trend for conjugation effects found for the oxidation of sulfide sulfur compounds.

The barriers computed for the four substrates nicely match experimental data. The oxidation of the sulfur atom of the isothiazole group in ziprasidone and perospirone has a lower predicted barrier (67 kJ/mol) than the neighboring carbon and nitrogen atoms (80 and 102 kJ/mol, respectively), which is in accordance with the fact that sulfur oxide is the only experimentally observed metabolite. On the other hand, the sulfur atom in thiophene has a higher activation energy for oxidation than the carbon atom next to it (67 vs. 69 kJ/mol), which match the fact that the carbon atom is oxidized in clopidogrel, ticlopidine, S-2-[4-(3-methyl-2-thienyl)phenyl]propionic acid, suprofen, tienilic acid, and derivatives of tienilic acid. Experimental data for the metabolism of benzothiophene indicates that substitutions of the ring governs whether the sulfur atom is accessible for oxidation or not, and oxidation of both carbon atoms in the six-membered ring as well as the sulfur atom are observed.

### Dehalogenation of halogenated benzenes

Table XXX delete. Exhaustive listing of theoretical studies of dehalogenation of halogenated benzenes by CYP enzymes.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hexafluorobenzene, pentafluorophenol, Hexachlorobenzene, pentachlorophenol</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2007</td>
<td>143</td>
</tr>
</tbody>
</table>

Figure 20. Substrates for which dehalogenation of halogenated benzenes by CYP enzymes has been studied by QM methods. The Figure is constructed the same way as Figure 3.
The dehalogenation of four perhalogenated benzenes (Figure 20) were studied by Hackett et al.\textsuperscript{143} The dehalogenation of such substrates occurs through multiple steps, the first being an attack by compound I on an aromatic carbon atom, forming a tetrahedral intermediate, just like in the oxidation of aromatic carbon atoms. This is followed by a migration of the halogen atom attached to this aromatic carbon to a neighboring carbon atom. For the chloride compounds, this migration was found to be barrierless, whereas a barrier was found for the fluoride compounds (Figure 1). As in the oxidation of aromatic carbon atoms, the formation of the tetrahedral intermediate is rate-limiting. The lowest activation energy is obtained for either the doublet or quartet spin state depending on the substrate and the activation energies are of a similar size as those for the oxidation of aromatic carbon atoms (see Figure 4).

The formation of the final product requires the release of a halide ion and it was suggested to occur through one-electron transfer from the heme, followed by a spontaneous expulsion of the ion. Optimization of the radical anion of the chlorobenzenes did indeed result in spontaneous expulsion of a chloride ion, but the same did not happen for the fluorobenzenes, indicating that the loss of a fluoride ion involves additional barriers.

The activation energies in Figure 4 are within the range of those for aromatic oxidations. The lowest activation energies (~53 kJ/mol) were found for the attack at the para and ortho positions of pentafluorophenol, whereas the highest activation energy (88 kJ/mol) was found for oxidation at the ipso position of the same compound. For all sites except the ipso one, the barriers for oxidation of the chloro compounds were higher than for the corresponding sites of the fluoro compounds. Consistent with studies of aromatic oxidations,\textsuperscript{64,65} a hydroxyl substituent decreased the barriers for oxidations at the meta and ortho sites for both the fluoro and chloro compounds.

**Special reactions**

Table XXX delete. Exhaustive listing of theoretical studies of reactions by non-standard mechanisms by CYP enzymes.
<table>
<thead>
<tr>
<th>Ligand</th>
<th>Steps studied</th>
<th>Spin state</th>
<th>Methodology</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostaglandin H2</td>
<td>Full mechanism</td>
<td>Doublet/quartet/sextet</td>
<td>Model system</td>
<td>2009</td>
<td>144, 145</td>
</tr>
<tr>
<td>Lanosterol carboxaldehyde</td>
<td>Full mechanism</td>
<td>doublet</td>
<td>QM/MM</td>
<td>2010</td>
<td>125</td>
</tr>
<tr>
<td>Steroid model</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>Model system</td>
<td>2005</td>
<td>126</td>
</tr>
<tr>
<td>trans-2-phenyl-iso-propylcyclopropane</td>
<td>Full mechanism of reduction</td>
<td>Doublet</td>
<td>Model system</td>
<td>2004</td>
<td>146</td>
</tr>
<tr>
<td>Chromopyrrolic Acid</td>
<td>Full mechanism</td>
<td>Doublet/quartet</td>
<td>QM/MM</td>
<td>2008-2009</td>
<td>147, 148</td>
</tr>
<tr>
<td>13-hydroperoxy-9,11,15-octadecatrienoic acid</td>
<td>Full mechanism</td>
<td>Doublet</td>
<td>QM/MM</td>
<td>2011</td>
<td>149</td>
</tr>
</tbody>
</table>

**Figure 21.** The suggested mechanism for the formation of three prostaglandin H$_2$ metabolites illustrated by the model systems used in calculations.$^{144, 145}$

In addition to the standard reactions described above, a number of studies of more specialized CYPs have been performed. Yanai and Mori have studied the formation of two prostaglandin H$_2$ (PGH$_2$) metabolites, prostacyclin (PGI$_2$) and thromboxane A$_2$ (TXA$_2$) by two different CYPs, prostacyclin synthase and thromboxane synthase.$^{144, 145}$ The suggested reaction mechanisms are shown in Figure 21. The isomerization of PGH$_2$ is initiated by the direct binding of one of the two peroxo oxygen atoms to Fe$^{III}$. This is followed by a homolytic O–O bond cleavage and a ring formation, during which the formed oxygen radical attacks the C=C double bond. The product is unstable and a proton-coupled electron transfer leads to the final product. The formation of TXA$_2$ is initiated by the same O–O bond cleavage, followed by the cleavage of a C–C bond. After this, the reaction path diverges and leads to the formation of two different products, TXA$_2$ and 12-L-hydroxy-5,8,10-heptadecatrienoic acid (HHT). HHT is formed through the cleavage of a second C–C bond which results in the formation of not only HHT but also malondialdehyde (MDA), whereas the formation of TXA$_2$ occurs through the formation of two O–C bonds.
The reduction of trans-2-phenyl-iso-propylcyclopropane by P450s has been investigated by Kumar et al. They showed that in addition to the normal radical intermediate, hydrogen abstractions by compound I can lead to cationic intermediates. The formation of this cationic intermediate is shown to originate from an excited state of compound I. This cationic intermediate leads directly to a second hydrogen-atom transfer to the iron–hydroxyl intermediate formed during the first hydrogen abstraction.

**Figure 22.** The reaction of chromopyrrolic acid catalyzed by P450 StaP.

Shaik and coworkers have studied the formation of a C–C bond during the reaction of P450 StaP with chromopyrrolic acid (CPA), shown in Figure 22. Based on experimental data and QM/MM calculations, they suggest that due to the constrained active site, which limits the orientation of CPA relative to the heme group, this reaction does not involve the expected double hydrogen abstraction from the two carbon atoms involved in the bond formation, but instead a proton-coupled electron transfer, in which the proton is shuffled by a Grotthus-like mechanism from CPA to the oxo group of compound I. After the first proton is abstracted from CPA, the C–C bond is formed and the second proton is transferred by a similar mechanism.

**Figure 23.** The mechanisms catalyzed by allene oxide synthase.

The reaction mechanism of allene oxide synthase has been investigated by Cho et al. using QM/MM calculations. They studied the formation of an epoxide from a peroxo group on a substrate and the formation of two different products, as is shown in Figure 23. The epoxide formation is catalyzed by the binding of the peroxo group to heme in the FeIII resting state, which is followed by a scission of the O–O bond of the peroxo group. The remaining oxygen atom in the substrate binds to a carbon atom two bonds away, forming an epoxide with a radical on a neighboring carbon atom. This epoxide can then form two different products, either an epoxide with a vicinal double bond, or an ether formed through the scission of the C–C bond in the epoxide.

**Drug metabolism predictions**

As described in the previous sections, DFT-based methods have routinely been used over the past decade to investigate P450 reactions of many different types. However, these methods are slow, taking several days of CPU time per substrate. If the aim is to predict sites of reactivity for large
sets of compounds, as often is the case during drug development, faster (but also less accurate) methods need to be used.

**Small radical surrogates for compound I**

In the pioneering work by Korzekwa and coworkers, AM1 was used to determine transition states, using \( p \)-nitrosophenoxy radical as a model for compound I in the CYPs to reduce the computational costs. They studied aliphatic hydrogen-abstraction reactions and found that a model including the heat of reaction and the ionization potential of the substrate correlate with the activation energies.\(^{150}\)

This was important, because the time-consuming and error-prone transition-state optimization could then be avoided. This model was later used to predict nitrile toxicity\(^{151}\) and experimental biotransformation rates of halogenated alkanes.\(^{152,153}\) The activation-energy model was later extended to include also aromatic oxidation.\(^{154}\)

In recent years, it has become possible to use computationally more accurate DFT methods. In a predictive study of CYP 2E1 metabolism, Park and Harris noted that the activation energies obtained with the AM1-based activation energy model by Korzekwa and coworkers did not agree with those obtained with DFT using a porphine model.\(^{57}\) Park and Harris also compared the DFT results with transition-state optimizations at the AM1 level with the \( p \)-nitrosophenoxy radical and observed a large, but systematic, deviation from the DFT activation energies. This was followed up in a study in which the activation energies of hydrogen-abstraction reactions for 24 compounds were investigated using a porphine compound I model using DFT.\(^{32}\) This showed that the activation energies of hydrogen-abstraction reactions determined with the radical surrogate models, even those determined at the AM1 level, correlated well with the DFT energies using the porphine system, because the errors are systematic. These models were then applied to calculate the activation energies and, also taking the solvent accessibility into account, used to rationalize the metabolism of progesterone and dextromethorphan (structures shown in Figure 24). Of the 30 possible hydrogen-abstraction reactions in progesterone, those with lowest activation energies, in which the hydrogen atoms are accessible, are indeed also observed experimentally. For example, the hydrogen abstractions in the 2\( \alpha \), 6\( \alpha \), and 14 positions had the lowest activation barriers, being hydrogen abstractions from carbon atoms next to a carbonyl carbon, an alkene double bond, and from a tertiary carbon, respectively. However, the accessibility of the carbon atom in the 14 position is very low, which is probably why this metabolite is not observed. Interestingly, the activation barriers also show a strong preference for the 6\( \alpha \)-position compared to the \( \beta \)-position, in agreement with experiment. For dextromethorphan, the activation energies for the 22 possible hydrogen-abstraction reactions showed that the N-dealkylation reaction is most favorable, in agreement with the fact that this metabolite is observed in CYP 3A4 metabolism. Moreover, the O-dealkylation, as observed in CYP 2D6 metabolism could be rationalized, because it is the most accessible functional group in the compound.

**Figure 24.** The structures of progesterone and dextromethorphan. For progesterone \( \alpha \)-hydrogen atoms are directed away from the viewer and \( \beta \)-hydrogen atoms are directed towards the viewer.
Thus, the fast surrogate models seem to be reasonably accurate for the hydrogen-abstraction reactions, but it was later shown that the correlation for aromatic oxidations using the same radical surrogate models are quite poor. On the other hand, recently it was shown that a radical surrogate models and valence bond modeling at the DFT level can describe aromatic oxidations quite well for a set of substituted benzenes.

Prediction of sites of metabolism for larger sets of compounds.

In a drug-discovery perspective, it is often necessary to screen thousands of compounds, which would be prohibitive even at the AM1 level. In particular, a \( p \)-nitrosophenoxy radical is problematic, because it often requires careful examination of the results and tests of different starting structures to ensure that the correct transition state has been located. Therefore, other approaches have been developed to get a fast and reliable prediction of the activation energy.

Singh et al. calibrated a trend-vector model on a set of 50 AM1 radicals to predict the hydrogen-abstraction activation energy for CYP 3A4. The model was successful for alkane oxidation, but the activation energies for aromatic carbon sites were systematically too large, probably because that reaction does not occur via hydrogen abstraction.

We have shown that the activation energies for the hydroxylation of aliphatic carbon atoms primarily depend on the chemical groups adjacent to the reacting carbon atom. For example, the largest activation energy was observed for hydrogen abstraction from primary carbons (89 kJ/mol, if excluding methane), whereas the smallest one was observed when an amine was adjacent to the reactive carbon (e.g. as in dextromethorphan). This indicates that the activation energy can be predicted by simple chemical-group rules, based on DFT calculations on a set of representative
model compounds. Such an approach has formed the basis for the development of the SMARTCyp approach.\textsuperscript{64,65} In this program, a large number of precomputed activation were compiled into rules consisting of SMARTS patterns and associated energies. Thereby, the site with the lowest activation energy can rapidly be predicted. Currently, 211 activation energy calculations have been performed to form 46 rules. However, since it is a 2D-based method, the fine details, e.g. whether it is the 6\(\alpha\) or 6\(\beta\) position in progesterone that is most reactive, cannot be captured.

Reactivities combined with effect of the protein.

Calculations of the activation energies using small model systems give useful measures for the intrinsic reactivity of a certain reactive site, but they lack information about possible steric and electronic effects from the CYP protein. In the case of CYP 3A4, it has been argued that the protein cavity is flexible and large, and therefore, the substrates may freely orient themselves in a way that the reaction with the lowest activation energy occurs. However, this is certainly not the case for other CYPs, which introduces a new challenge, i.e. determination of the transition state in the presence of the enzyme. This can, of course, be done using QM/MM as discussed in the previous sections, but in a drug-discovery scenario, this is not feasible due to the large computational effort.

A few groups have circumvented this problem by using conventional docking methods, e.g. assuming that the atom in the substrate can react with the enzyme if it is within a certain distance from the Fe ion (typically ~6 Å).\textsuperscript{156-158} Jung et al. used Autodock in connection with the AM1 model of Korzekwa and Jones to predict the site of metabolism for 8 out of 12 CYP 1A2 substrates correctly.\textsuperscript{159} Rydberg et al. followed the same strategy, but used the GOLD docking program and different measure for the reactivity, e.g. using either an AM1 methoxy-radical surrogate model or energy rules derived from precomputed DFT activation energies.\textsuperscript{28} With the precomputed DFT activation energy rules, 90 % of the major metabolites were within top three in rank. Recently, Moors et al. made an extensive study on CYP 2D6, discussing effects of protein ensembles and how to combine docking with the Glide program with reactivities from SMARTCyp to predict CYP 2D6 metabolism. The models using either the docking or the reactivities to predict sites of metabolism were not as successful as using the combination of docking and reactivities, which improved the prediction rates. With the combination, 88% of the predicted metabolites ranked as either 1 or 2 are experimentally observed metabolites.\textsuperscript{160} The quality of the models was shown to depend on the number of protein structures used and the cut-off distance between Fe and the potential site of metabolism. Using multiple protein structures (200–1000) improves the performance and using only the crystal structure is not sufficient, as was also emphasized by Hritz and Oostenbrink.\textsuperscript{161} In another study on CYP 2C9 by Danielson et al.,\textsuperscript{162} in which the combination of docking and reactivities was also successfully employed, it was less clear whether the use of multiple protein structures affects the prediction rate. This emphasizes the complexity of predicting the site of metabolism including the effect from the protein, probably owing to the fact that the CYP enzymes are flexible.

Another fast way to determine the transition state of the reaction in the enzyme is to use transition-state force fields. A force field for hydrogen-abstraction reactions was developed by us\textsuperscript{163} with the Q2MM method\textsuperscript{164} and was used to predict the transition state inside the protein for two CYP 2C9 and 3A4 substrates.\textsuperscript{165} The transition-state structures were reasonable accurate.\textsuperscript{163} However, it was challenging to rank them, owing to the well-known problem of predicting binding free energies.\textsuperscript{166}

A few groups have used a more simple description of the protein than the previous all-atom approaches. Oh et al. used a catalyticphore-based docking method in combination with activation
energies determined by the AM1-based method of Korzekwa,\textsuperscript{150,154} whereas Hasegawa et al. used a somewhat related method,\textsuperscript{167} except that the precomputed activation energies from our work were applied.\textsuperscript{32,85} Both models managed to correctly predict the site of metabolism within top two in rank for about 80\% of the substrates. Several groups have used the solvent-accessible surface area of the various hydrogen atoms as a measure of their intrinsic accessibility.\textsuperscript{32,155} In SMARTCyp, the relative distance to the center of the molecule is used as a measure for the accessibility of the given site.\textsuperscript{64} The SMARTCyp method has been applied to predict the site of metabolism for a set of 361 CYP 3A4 substrates and correctly predicted the metabolic position as either rank one or two in 81\% of the cases.\textsuperscript{65}

Conclusions

In this article, we have reviewed DFT studies of reactions performed by the CYP superfamily of enzymes. The aim of the review has been to describe as many types of reactions as possible on an equal footing, providing information needed to predict sites of reactivity for an arbitrary substrate. It is apparent that some types of reactions, in particular aliphatic hydroxylation, alkene epoxidation, and aromatic oxidation, have been thoroughly studied, whereas several other reactions have been studied only once or twice. Consequently, much work still remain before the CYP mechanisms are fully understood. In particular, many details of the mechanism of aromatic oxidation, hydroxylation of primary and secondary amines, and the desulfurization of phosphor are poorly known. Moreover, the effect of dispersion on the various reaction types needs to be further investigated.

The results have shown that the activation energies go from essentially 0 to 109 kJ/mol. Figure 4 shows that the barrier often depends more on the neighboring groups than on the type of reaction. Therefore, the barrier can often be predicted quite accurately with simple chemical rules.\textsuperscript{32,64,65} Unfortunately, it is much harder to predict steric effects of the surrounding protein, although simple rules based on the solvent-accessible surface area or other related properties often are surprisingly successful.\textsuperscript{64,65,78,155} Therefore, more research are needed before predictive models for sites of metabolism for arbitrary drug candidates become really reliable. Such models would also be improved if more crystal structures of the most important CYP isoforms in complex with different types of compounds were solved, thus shedding light on issues like flexibility and the position of water molecules.

Acknowledgements

This investigation has been supported by grants from the Swedish research council (project 2010-5025) and Lhasa Limited.

Xxx fix journal name which do not get automatic short names: Ref 10

Reference List


103. Ring, B. J.; Parli, C. J.; George, M. C.; Wrighton, S. A. In-Vitro Metabolism of Zatosetron - Interspecies Comparison and Role of Cyp 3A. *Drug Metab. Dispos.* 1994, 22, 352-357.


