

Application of Q2MM to stereoselective reactions

Sten O. Nilsson Lill,[§] Aaron Forbes,[†] Patrick Donoghue,[†] Vincenzo Verdolino,[†] Olaf Wiest,^{*†} Patrik Rydberg,[‡] and Per-Ola Norrby^{*§}

Department of Chemistry, University of Gothenburg, SE-412 96 Göteborg, Sweden, Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana, 46556-5670, USA, and Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark

Email: owiest@nd.edu, pon@chem.gu.se

Abstract: Q2MM is a method designed to allow application of molecular mechanics calculations to transition states in chemical reactions. It is one of the few methods available that allow determination of a complete set of low-energy transition states for medium-sized systems, and thereby gives a unique opportunity to investigate kinetic selectivity, in particular stereoselectivity. The current review will give an outline of the procedure, an overview of the types of reactions that have been studied using this method, and summarize the factors affecting the accuracy of the results.

Introduction

One of the major concerns in organic synthesis is the issue of selectivity. A large part of all development and application of synthetic methods is to find reaction conditions that maximize the selectivity for and yield of the desired product. In recent years, huge advances in computational methods and hardware have enabled development of modeling tools that can analyze or even predict reaction selectivity [1]. Tools for evaluating relative reactivity can roughly be divided into two classes (Figure 1): Descriptor-based, where a relationship is derived by fitting observed reactivities to known or calculated properties of reagents, or direct modeling and evaluation of transition states. The former type, generally termed QSAR or QSRR (quantitative structure-activity or structure-reactivity relationship), requires access to experimental results for substrates of a type fairly similar to those to be predicted, and can give excellent results [2]. QSAR-approaches frequently utilize descriptors derived from the structure of the reactants, but can also employ data from approximate transition state models, on the assumption that the QSAR procedure tends to compensate for

[§] University of Gothenburg

[†] University of Notre Dame

[‡] University of Copenhagen

systematic errors in the underlying data. More accurate transition state models can be used directly through transition state theory, and therefore does not have to rely on pre-determined experimental data. To this latter class belong mechanistic investigations by quantum mechanical (QM) methods that, if run at an appropriate level, can yield very accurate relative reaction barriers. However, QM methods are inherently slow. They can be efficient and accurate if the goal is to elucidate the selectivity-determining interactions in a specific system, but errors can become significant if a few transition structures are taken to represent the transition state ensemble in flexible systems. When the goal is prediction, and in particular for virtual screening of reactants, QM methods cannot compete efficiently. The computational evaluation can be accelerated by treating part of the system with much more rapid molecular mechanics (MM), so-called QM/MM methods [3], or by using semi-empirical methods in combination with the previously mentioned QSAR methodology to reduce the inherent errors in this type of calculation, but the electronic structure calculations will still be a bottle-neck with these approaches.

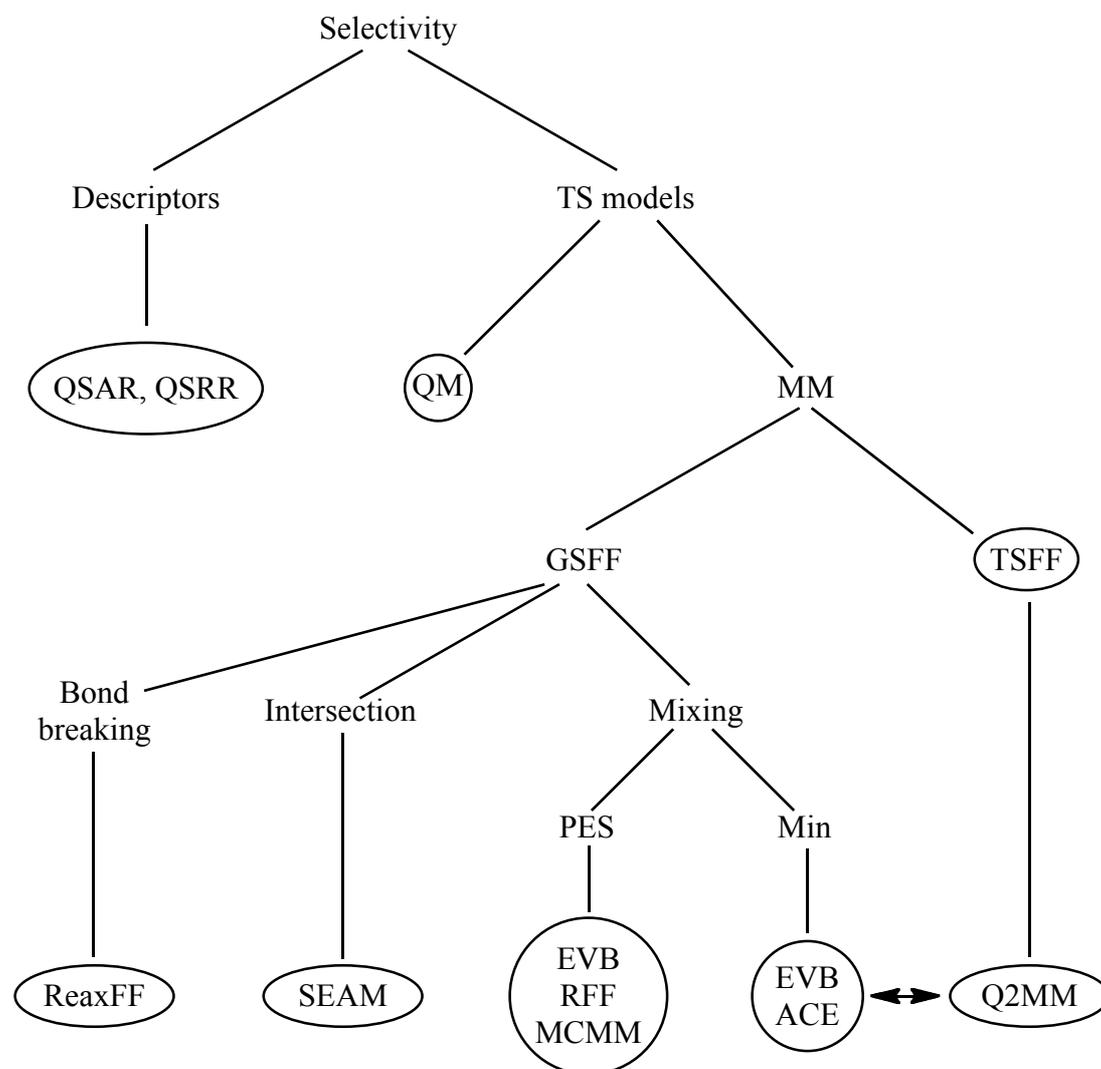


Figure 1 Computational methods for selectivity prediction

Pure MM calculations are several orders of magnitude faster than QM while still delivering a similar accuracy for fully parameterized systems [4, 5], but traditional MM cannot be applied directly to bond breaking reactions, since the position of the bonds have to be predefined. Several approaches have been developed that allow the use of fast MM calculations in evaluation of reaction selectivities [6]. First of all, an MM force field can be set up to reproduce any type of structure, not just observable molecules. It is possible to redefine all bonding parameters so that the method produces a transition state structure upon energy minimization, a so-called transition state force field (TSFF) [7]. The Q2MM method [8, 9], which is the focus of the current review, is a special case of TSFF. Alternative approaches based on ground state force fields (GSFF) include development of new types of MM that base the bonding on the instantaneous proximity and thus allow formation and breaking of bonds (ReaxFF [10]); approximating transition states as intersections between ground state force fields for reactant and product (SEAM [11]); or mixing ground state force fields to produce a proper potential energy surface (PES), as originally done by Warshel and coworkers in EVB [12, 13], and more recently in, for example, RFF [14] and MCMM [15]. Within EVB, it is also possible to combine ground state force fields by taking a weighted average, a procedure that can be viewed as using partial bonds in MM. The recently introduced ACE [16] utilizes this method to produce what is, in effect, a transition state force field.

All methods utilizing MM are strongly dependent on parameters. In traditional MM, the position and shape of a minimum on the PES is fitted to the structure and vibrations of observable molecules through adjustment of parameters. A varied and complete data set is necessary for development of a reliable and transferable parameter set. However, for transition state force fields, experimental data for the relevant structures is of course unavailable. Early applications used assumed structures and adjusted the required parameters to fit a limited number of experimental selectivities, but this practice was criticized since it resulted in severe overfitting [17]. In some instances, parameters could be fit to QM structures for transition states [7], but as is well known in the MM field, structures are not in themselves sufficient data for derivation of complete force fields. It is necessary to include some data relating to the shape of the PES around the energy minimum, most commonly vibrations [18] or, if available from QM calculations, the underlying Hessian (the second derivatives of the energy with respect to Cartesian displacements) [19]. However, one particular problem arises when trying to utilize QM vibrations in derivation of TS force fields. To be useful in conjunction with standard conformational search tools, the TSFF must represent the TS as an energy minimum, whereas the QM vibrational analysis must result in a negative curvature along the reaction coordinate. In the late 1990's, Norrby introduced a method for modifying QM data for transition states to give positive curvatures along all vibrational modes [8]. In combination with a previously developed program package for fitting MM parameters to a combination of structural and QM-derived vibrational data [20, 21, 22], this procedure forms the Q2MM (read: "quantum to molecular mechanics") method for deriving a TSFF entirely and consistently from QM calculations for small model systems. Thus, Q2MM is one of a very limited number of methods that can be applied

to accurate selectivity predictions and virtual screening without dependence on previously determined experimental results.

In the following, we will give an introduction to the Q2MM method, a historical overview demonstrating the performance of the method, and a detailed tutorial on how to apply Q2MM to selectivity predictions in organic synthesis.

The Q2MM procedure

Q2MM is a procedure for creating a transition state force field entirely from QM data for small model systems [8, 9]. The resulting force field, a Q2MM force field, differs only in accuracy from TSFF derived by other methods [7]. It is used in standard MM software, delivers transition structures and their relative energies by standard energy minimization, and can be utilized in conformational searching or molecular dynamics to yield transition state ensembles. Thus, it is one of the few methods that allow a true energy comparison and Boltzmann averaging over all low energy paths of a chemical reaction.

In the original implementation, Q2MM utilized a previously published method for parameterizing molecular mechanics force fields [20]. In the Q2MM method, the required data is transition structures and Hessian elements calculated by DFT (or other accurate QM methods) for small model systems. It is in fact essential that the model systems are kept small, since the DFT method does not properly account for all vdW interactions, and the systematic error increases with model size. Since MM methods include empirical parameters for non-bonded interaction, the vdW description is usually good, and should not be deteriorated by fitting to less accurate DFT data. When the required data has been gathered, parameters are added to a suitable force field to allow calculation of the transition structures. This procedure involves a choice of atom and bond types to use in the MM calculation (*cf.* tutorial section 1.4). The resulting force field is not meant to be accurate at this stage, but should simply allow a geometry optimization in the program of choice.

In the parameter refinement stage, structures, relative energies of transition state conformations, and Hessian elements are calculated by MM and compared to the QM data, yielding a penalty function that is basically an rms sum of weighted deviations [22]. The MM parameters are then varied by any suitable optimization method until the penalty function is minimized. What differentiates Q2MM from other similar MM parameterization methods [23, 24] is the treatment of Hessian data. Diagonalization of the mass-weighted Hessian yields the normal modes of the system, corresponding to all possible vibrations of the molecule. In the QM data for a transition structure, one normal mode must, by definition, have a negative eigenvalue, whereas in TSFF, the MM structure is supposed to be an energy minimum, and therefore to have only positive eigenvalues of the Hessian. Thus, the QM data must be modified before being used in the parameterization. This is simply done by diagonalization, replacement of the negative eigenvalue with a large positive value (*cf.* tutorial section 2.2), and reconstruction of a Hessian from the eigenvectors and modified eigenvalues

(Figure 2). An exact reproduction of this Hessian will now give an MM force field that produces a minimum also along the reaction coordinate (*i.e.*, the normal mode with the negative eigenvalue), but where the energy response to small deviations perpendicular to the reaction coordinate will exactly reproduce the corresponding QM system.

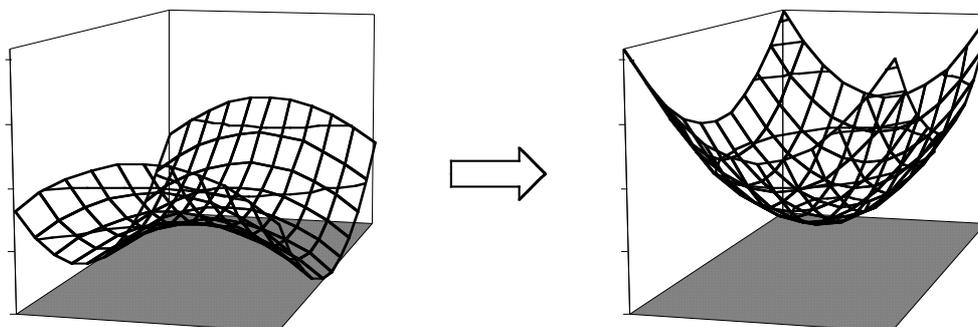


Figure 2 Pictorial representation of the Q2MM Hessian modification

The original implementation of the parameterization method used in Q2MM involves geometry optimization of the model complexes used for optimizing the parameters of interest. This works very well for most cases. However, for systems with very small torsional barriers this can lead to artifacts during the parameter optimization. These occur when small changes in a parameter make the geometry optimization end up in another local minima, and thus a very small parameter change leads to a large change in the penalty function [25]. This large change then makes the derivative of the penalty function with respect to this parameter invalid. The problem can be alleviated by disregarding some data points or freezing torsions during the geometry optimizations [25]. However, to solve this problem permanently we have recently changed the parameter optimization procedure. The most recent implementation does not need to use any geometry optimization, and thus may exclude geometrical data such as bond lengths and angles when computing the penalty function. Instead we compute the first derivative of the energy with respect to the Cartesian coordinates (the gradient). These derivatives have previously been shown to work well for parameter optimizations [19, 23, 26, 27]. Thus, the current implementation uses energies, first and second derivatives of the energy with respect to the Cartesian coordinates as reference data for the parameter optimization, whereas geometries are optional (*cf.* tutorial section 2.2).

Published cases

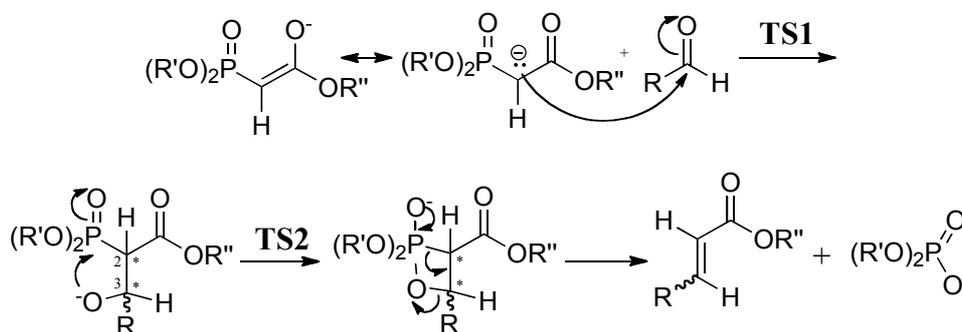
In this section, we will go through all implementations of the Q2MM method that have so far appeared in the literature. Table 1 summarizes these cases.

Table 1 Reaction types studied using Q2MM.

Reaction	Bond formation	References
Horner-Wadsworth-Emmons	C-C	28
Osmium-catalyzed asymmetric dihydroxylation of alkenes	C-O	29, 30, 31
Chiral β -amino alcohol promoted addition of dialkyl zinc to aldehydes	C-C	32
Silver catalyzed hydroamination of alkynes	C-N	33
Cytochrome P450 catalyzed hydroxylation	O-H	25, 34
Rhodium catalyzed asymmetric hydrogenation of enamides	C-H	35, 36

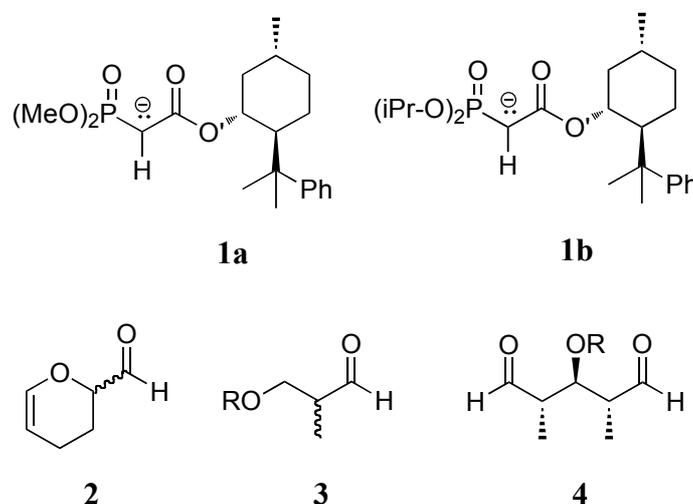
Asymmetric Horner-Wadsworth-Emmons reaction

The first reaction where Q2MM was used for rationalization of product selectivities was the Horner-Wadsworth-Emmons (HWE) reaction (Scheme 1) studied by Norrby, Brandt, and Rein (NBR) [28]. In this Wittig-like reaction, a C-C bond is formed via a nucleophilic attack of a carbanion/enolate on an aldehyde. The negative charge on the reacting carbanion is stabilized via a phosphonate group and an additional electron-withdrawing substituent, typically a carboxylic acid derivative such as an ester.

**Scheme 1** Reaction mechanism of the HWE reaction.

After formation of the carbanion, the initial step of the reaction is the addition to the aldehyde forming an oxyanion. This step is controlled by **TS1**. Since both the aldehyde carbon and the enolate carbon are sp^2 -hybridized, and two new stereogenic carbons are formed in the reaction (Scheme 1), four different diastereoisomeric TS's ($2S3S$, $2S3R$, $2R3R$, $2R3S$) are possible in this step depending on if *re*- or *si*-faces of the reactants are approaching in the TS. Introduction of a stereogenic center ($4R$ or $4S$) in α -position on the R-group of the aldehyde increases the number of possible TS's to be formed to eight. In the second step of the reaction controlled by **TS2**, the oxyanion makes a nucleophilic attack on the phosphonate phosphorus forming a four-membered oxaphosphetane ring. This unstable intermediate will rapidly eliminate forming an alkene and a phosphate ester. The resulting alkene mixture has an (*E*)- to (*Z*)-alkene composition which is determined by the relative ratios of **TS1** and **TS2** for all eight different reaction paths. Thus, there will be a mixture of (*E*)- $4R$ /(*E*)- $4S$ and (*Z*)- $4R$ /(*Z*)- $4S$ isomers. To give enantioselectivity in the HWE-reaction, a chiral

auxiliary has been introduced in the R²-position in the nucleophile. In the NBR-study, two different nucleophiles (**1a** and **1b**) derived from Corey's chiral auxiliary (1*R*,2*S*,5*R*)-8-phenylmenthol were evaluated together with three different aldehydes (**2** - **4**) (Scheme 2).



Scheme 2 Reagents used in the Q2MM study of the HWE reaction.

From the NBR-study, it became evident that **TS1** and **TS2** were close in energy, and thus to be able to accurately predict the reaction selectivity it was necessary to calculate them both for all diastereomeric TS's. Since **TS1** and **TS2** are TS's for different reactions, the energy difference between them is not directly comparable using Q2MM, it contains a systematic but unknown correction (δ). Therefore, δ was varied until an optimal fit to experimental data was achieved. Also, the composition of the starting enolate was investigated, and the best results were obtained with a 92% *E*-/8% *Z*-enolate reactant mixture. For each reaction product *E/Z*-, (*E*)-*R*/(*E*)-*S*-, and (*Z*)-*R*/(*Z*)-*S*-ratios were calculated based on Q2MM-derived relative energies for all TS's (**TS1** and **TS2**) involved (at least 45,000 conformations). In total, 14 selectivities were reported and compared with experimentally determined values. Only in one case was the wrong product isomer predicted and this was for the **1a** + **3** reaction which experimentally showed only a modest *E/Z*-selectivity. Thus, in this case was the energy difference determining the selectivity rather small, making the computational prediction very challenging. Overall, the largest error in relative energies was smaller than 4 kJ mol⁻¹. This accuracy allows for synthetically useful predictions of reaction selectivities.

The main reasons for the selectivity in the different reactions were elucidated and some general rules could be proposed:

- a) The role of the chiral auxiliary is to predestine the face selectivity of the enolate as evidenced from the observation that reaction on the *re*-face of the enolate was blocked in **TS1** and a resulting (*S*)-configuration at the C2-position in all TS's was clearly favored (Figure 3). This efficiently reduces the number of available TS's; more than 90% of the product arises from 2*S*-configured TS's.

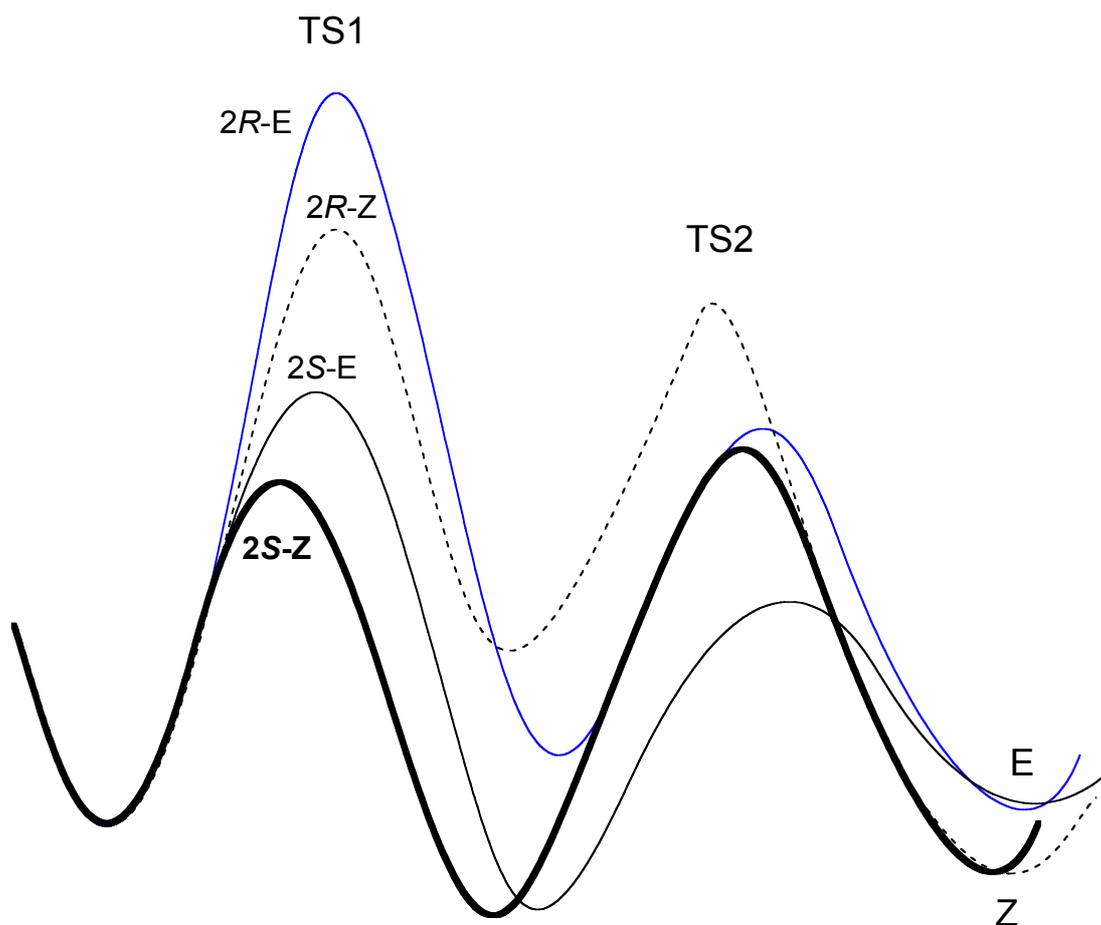


Figure 3 Dominating reaction paths for the four diastereoisomeric TS's in the **1a+2** reaction, with the globally preferred reaction path drawn in bold. *R* and *S* apply to **TS1** and **TS2**, while *E* and *Z* apply to the product.

- b) With the main reaction pathways traced by the TS's with a *2S* configuration, the preferred relative stereochemistry at C3 (*3S* or *3R*) could be controlled by an increased bulk at the phosphonate alkoxy groups (*R'*). The sterically more demanding *iPr* group (**1b**, Scheme 2) was found to give an increased selectivity in reactions with aldehydes **2** or **3** compared to that observed with a methyl group (**1a**). The favored *2S3R* TS's undergo *syn*-elimination to form *E*-alkenes while *2S3S* TS's lead to *Z*-alkenes and thus the *E/Z*-ratio increased with a bulky phosphonate substituent.
- c) The stereochemistry at C3 in the TS's was strongly influenced by the aldehyde stereocenter. Reactions with aldehyde **2**, which has an α -alkoxy function, was found to filter out the *3R4S* and *3S4R* product generating reaction paths and thus to promote the *3R4R* and *3S4S* products following the Felkin-Anh-Eisenstein (FAE) model. Aldehydes **3** and **4** on the other hand, with β -alkoxy functions, filter out the *3R4R* and *3S4S* product generating reaction paths, and instead lower the barriers for the *3R4S* and *3S4R* product generating reactions, in an anti-FAE fashion. This was especially pronounced for the **1a+4** reaction, where the *3R4S* TS generates the *E*-alkene as the major stereoisomer. The reason for the selectivity can be understood by the schematic picture of the product, which shows that the *R*-group of the

aldehyde could cause steric interactions with the ester functionality if positioned to generate a *Z*-alkene. In the NBR-study, aldehyde **4** was the electrophile with the largest side-chain in the α -position. Thus, even without the *iPr* group in the nucleophile the *E/Z*-ratio could be increased.

Osmium-Catalyzed Asymmetric Dihydroxylation

The second reaction that was studied by the Q2MM method was the osmium-catalyzed asymmetric dihydroxylation (AD) reaction developed by Sharpless (Figure 4) [37,38]. Here osmium is coordinated by a chiral nitrogen ligand which induces the enantioselectivity. The first generation of ligands consisted of modified *cinchona* alkaloids, whereas a second generation linked two alkaloids by an aromatic backbone [39].

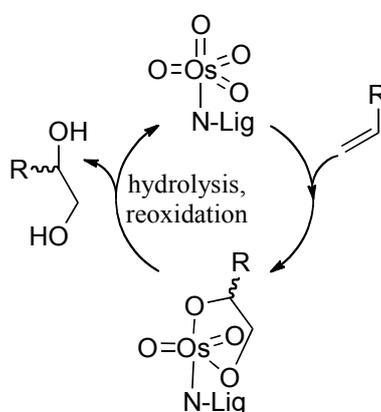


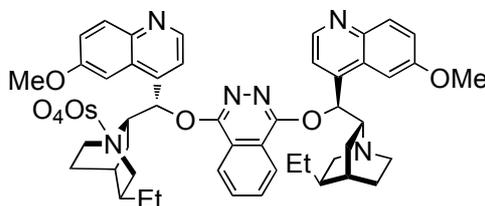
Figure 4 Catalytic cycle for the AD reaction

This reaction has been shown to follow a [3+2] cycloaddition mechanism with a direct formation of an osmium glycolate, which after hydrolysis yields the 1,2-diol. Evidence for this reaction mechanism comes from kinetic isotope effect and DFT-studies [40, 41, 42, 43, 44]. In the three Q2MM-papers on the AD-reaction, a TS-model was built based on this reaction mechanism. In contrast to the HWE study, only one TS-force field was needed. In the TS, the N-ligand was modelled to bind in an apical position to osmium, while the alkene adds to one apical and one equatorial oxygen (Figure 4). In the TS-model generation, it was also assumed that the alkene addition is the rate-limiting step. In total, >100,000 structures were covered in the conformational search.

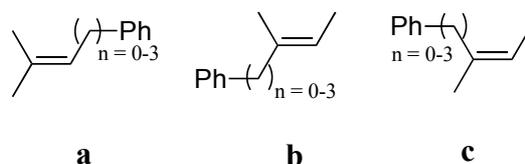
Sharpless and co-workers have developed an empirical mnemonic for the resulting enantioselectivity [39, 45]. The Q2MM studies gave a rationalization for the selectivity based on the possible geometries of the TS's, and allowed a refinement of the mnemonic device [31].

In the first of these three studies [29], both first and second generation ligands were evaluated together with eight different alkenes for comparison to 15 experimentally obtained enantioselectivities. In all cases the correct enantiomer could be predicted, and an excellent correlation between Q2MM-predicted and experimental selectivities was achieved.

In the second paper [30], now with a slightly modified Q2MM-force field, only the second generation (DHQD)₂-PHAL-ligand (Scheme 3) was used to study the dihydroxylation of six different tri-substituted alkenes (Scheme 4, n = 0, 2).



Scheme 3 The second generation catalyst OsO₄·(DHQD)₂PHAL



Scheme 4 Alkene substrates designed for Q2MM and experimental testing.

Again, the correct product enantiomers were predicted in all cases, and good correlation between experimentally determined selectivities and Q2MM-predicted selectivities was presented. The largest error was reported to be smaller than 4 kJ mol⁻¹. The findings were summarized in a pictorial view of how the enantioselectivity is influenced by the osmium coordinating ligand (Figure 5) and an updated mnemonic (Figure 6). The dominating selection point is found in the SE-corner of the mnemonic, where a substituent on the alkene would interfere strongly with the PHAL-linker of the N-ligand. Thus, the alkene will preferably orient its smallest substituent (most often hydrogen) into this position. Second, there are two areas with significant attractive interactions in the mnemonic, SW and NE. In the SW corner, the two quinoline moieties can stabilize the alkene substituent via dispersive van der Waals-interactions. The second alkaloid unit (which does not coordinate to osmium) seems to be especially important, since it can stabilize the alkene substituent with its quinoline unit. This is most effective if the substituent is an aryl group, so that the two aromatic groups can align in a slipped stack arrangement. The other alkaloid unit does not reach far enough to give a strong interaction, but if the aryl group is substituted as in the (DHQD)₂PHAL-ligand, the methyl in the methoxy group can interact with the alkene SW-substituent, especially if this is an aryl group. Also, the methoxy group can interact with one of the oxo-groups in the OsO₄ group via a dipole-dipole interaction. Another important role of the quinoline on the non-coordinating alkaloid is to stabilize the alkene portion via donation of electron density into the forming σ*-CO bonds, in combination with electrostatic interactions between the alkene and the quinoline. In the NE corner of the mnemonic, the substituent can gain stabilization by interacting with the PHAL linker and, if long enough, also with the quinuclidine portion. The NW corner corresponds to a free open space, in which any substituent may fit. If the substituent in this position is long and flexible enough it can interact with the second quinoline ring.

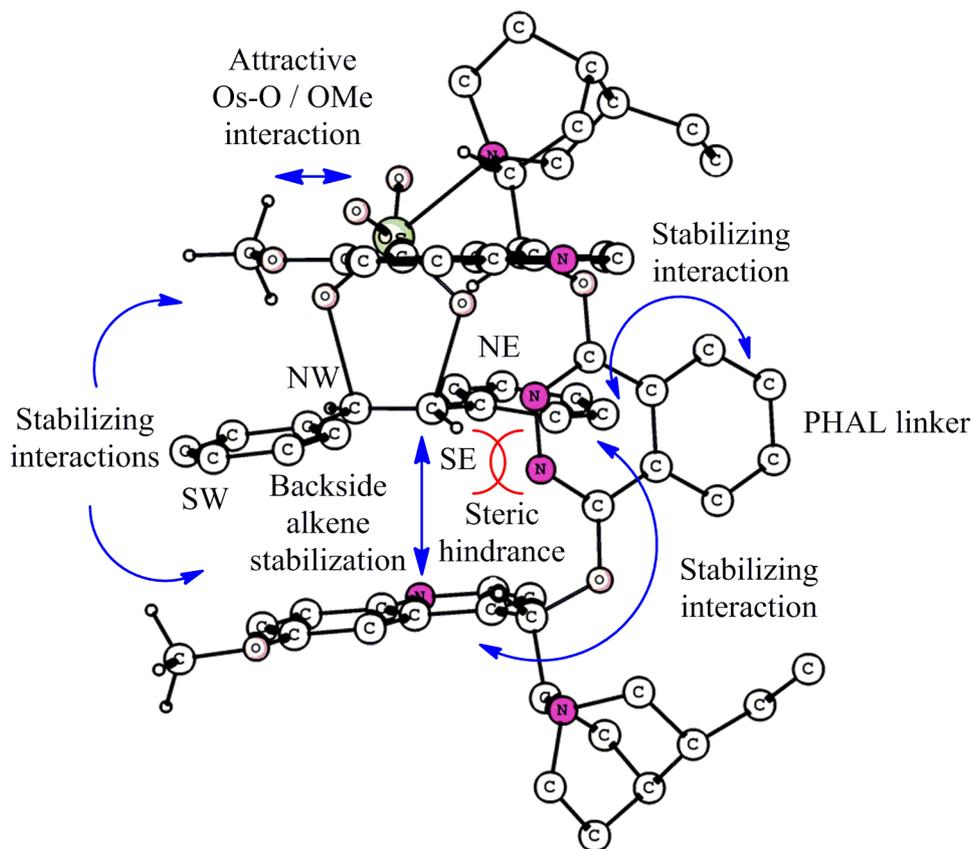


Figure 5 Interactions influencing the AD selectivity. Blue = stabilizing, red = destabilizing.

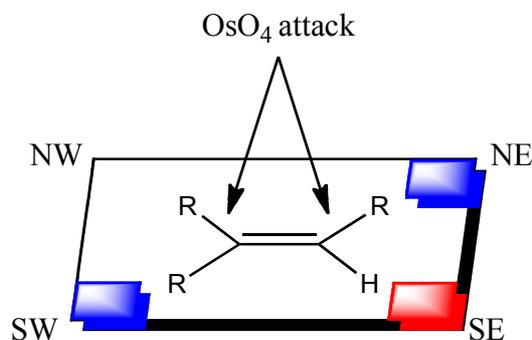


Figure 6 Updated mnemonic. Blue = stabilizing, red = destabilizing interactions.

The updated mnemonic and the scope of the Q2MM-method was further evaluated on another set of six tri-substituted alkenes ($n = 0-3$, Scheme 4) in a third paper [31]. Relative rate constants of the dihydroxylation reaction for different substituents were calculated from isodesmic relationships and compared with experimental results. In agreement with experimental data, the **c**-family among the alkenes (Scheme 4) showed the lowest rates. A comparison between the different alkenes (**a-c**) resulted in a fair linear correlation ($r^2 = 0.77$) between experimental and calculated relative activation barriers. A similar comparison was also presented for experimental and calculated enantioselectivities, with a maximum error of 4 kJ mol^{-1} . The validation of the revised mnemonic was manifested in this more recent paper, and

some limitations of the qualitative power were also pointed out. Especially interesting was the conclusion regarding the SW corner which was found to be sensitive to substrate shape, as a benzyl group was found to preferably go into the NE corner whereas a phenyl group goes to the SW corner.

The AD studies demonstrated for the first time that enantioselectivity could be successfully predicted (as opposed to rationalized). Figure 7 summarizes the ee predictions from the three studies, and also demonstrates the effect of assuming a 3 kJ mol⁻¹ error in the calculations. As can be seen, due to the nonlinear relationship between energy and ee, a small energy error gives a large error in the low ee range, but for development of catalysts, the high ee range is much more interesting. Looking at the overall accuracy, a clear majority of all cases fall within the 3 kJ mol⁻¹ range, and most are actually within 2 kJ mol⁻¹. Only in 2 cases (with low ee) out of 27 is the wrong enantiomer predicted; the error in both of these cases are ≈ 3 kJ mol⁻¹.

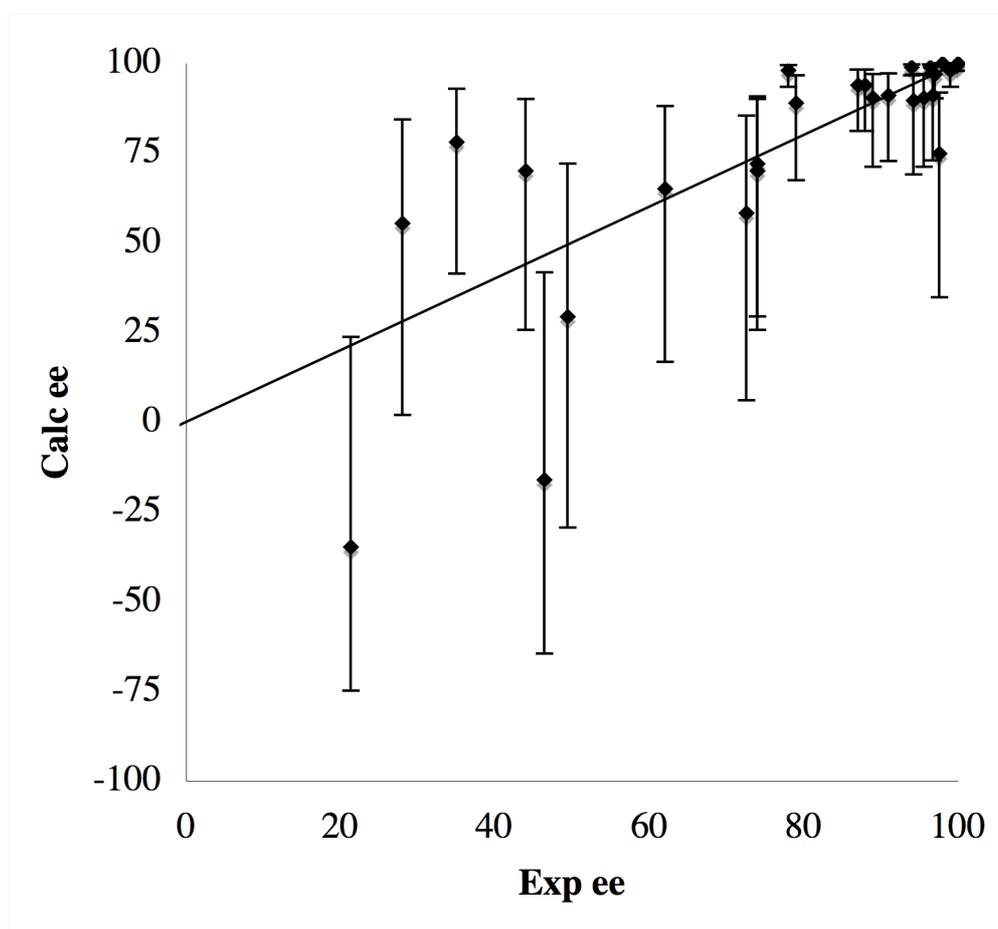
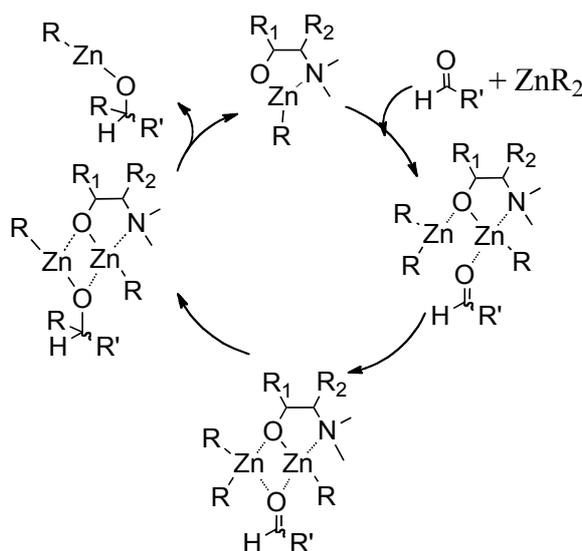


Figure 7 Predicted vs. observed AD enantioselectivity, with 3 kJ mol⁻¹ error bars.

Asymmetric Addition of Dialkyl Zinc to Aldehydes

The enantioselectivity in the chiral β -amino alcohol promoted addition of dialkyl zinc reagents to aldehydes has been investigated using the Q2MM-methodology in a study by Rasmussen and Norrby (RN). In this reaction, the catalyst is believed to be a zinc alkoxide with an additional intramolecular coordination to zinc from the amine

nitrogen (Scheme 5) [46]. The zinc in the catalyst acts as a Lewis acid and coordinates to the carbonyl oxygen of the aldehyde, while the stoichiometric zinc from the reagent coordinates to the alkoxy oxygen. The alkyl group of the zinc reagent is transferred to the carbonyl carbon of the aldehyde under formation of a chiral zinc alkoxide. It is likely that a four-membered (Zn-O)₂ ring is formed before alkyl transfer due to zinc coordination to the oxygen of the aldehyde. This coordination motif has been characterized in a stable pre-product, which after workup yields the desired chiral alcohol and regeneration of the catalyst.



Scheme 5 Catalytic cycle for ligand-promoted addition of dialkyl zinc to aldehydes.

Since the aldehyde can coordinate to the catalytic zinc with the R'-group oriented either *cis* or *trans* to the metal, and in addition attack from either above or below the Zn-plane, this generates four different pre-complexes (Figure 8), in which zinc is also a stereocenter.

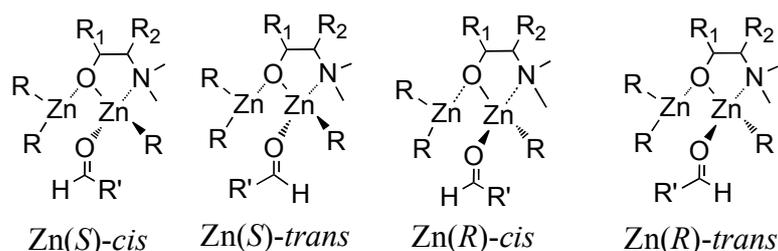


Figure 8 Aldehyde coordination to zinc-alkoxide complexes.

A number of possible mechanisms have appeared in the literature for this reaction. However, work from Noyori's group has prompted a consensus on the reaction mechanism [47, 48, 49, 50]. Two tricyclic TS's can be generated for each aldehyde coordination mode (Figure 8), depending on whether the transferring R-group on the stoichiometric zinc is syn or anti with respect to the amine of the bidentate ligand. This gives an eightfold combination of anti/syn and *cis/trans* arrangements in the TS. The chiral ligand will favour one of the faces of the catalytic zinc over the other, and thus discriminate half of the number of possible TS's. In general, it has been observed that the anti-*trans* TS gives the major product

enantiomer. The minor product enantiomer on the other hand can arise from anti-*cis* and syn-*trans* arrangements in the TS, or from insufficiently effective blockage of the "wrong" Zn face by the chiral ligand [32]. In the RN-study, a Q2MM model based on the tricyclic TS's was developed. Seven different chiral ligands were screened in combination with four different aldehydes (Figure 9). The alkylating reagent was in all cases diethylzinc.

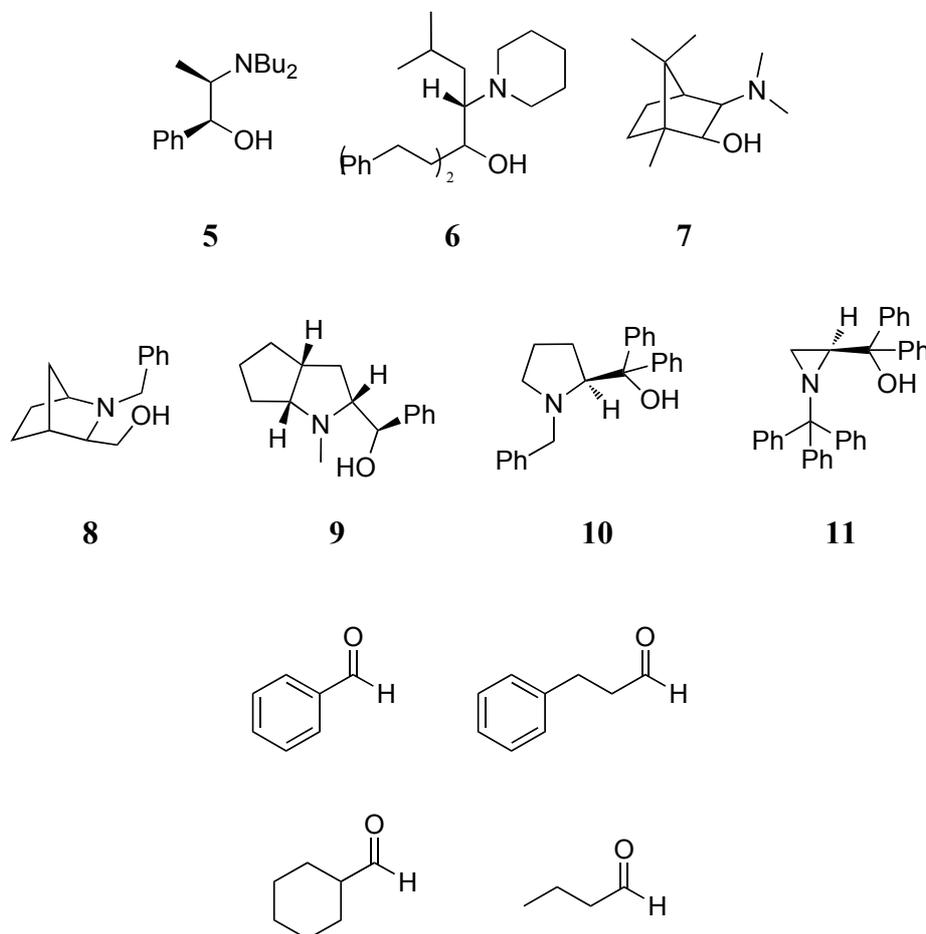


Figure 9 Ligands and aldehydes included in the Q2MM study [32].

The major product enantiomer could, in all cases, be predicted by Q2MM, with the majority of the predictions within 3 kJ mol^{-1} of the experimentally determined enantioselectivities. For all combinations of ligand and aldehyde the anti-*trans* TS was the preferred one. Ligands **5** and **7**, with isomorphic chirality on both the α - and β -carbon in the backbone of the β -amino alcohol, were found to efficiently block one face of the Zn-plane. The chirality on the α -carbon was reported to be more significant due to its vicinal position to the zinc coordinating oxygen. Alternatively, if one of the N-substituents in the ligand was very bulky, as in ligand **11**, similar face selectivity could be achieved. Ligands **10** and **11**, with phenyl substituents on the α -carbon, and importantly a short-chain bridge in between the β -carbon and the nitrogen, were found to block the syn-*trans* pathways. Blocking of the anti-*cis* pathway was found to be the most difficult to control and was strongly dependent on the choice of substrate. For example, aliphatic aldehydes with a large steric bulk close

to the carbonyl moiety experienced high selectivity with respect to this pathway. This was also observed for benzaldehyde.

A generic TS-conformation selectivity model built on the five-membered chelate ring, formed due to the N-Zn coordination, was presented in the RN-study. It was observed that the ring mainly has two conformations (**A** and **B**, Figure 10), interconverting via a ring flip, where conformer **B** is the most stable of the two. Coordination of substrate and alkylating reagent was found to be disfavored in isomer **A** compared to that in conformer **B** due to stronger steric interactions with one of the substituents on the nitrogen in the aminoalkoxide. In isomer **A**, a substituent on the α -carbon would be positioned in an axial position and if the reagent coordinates to the face where the α -substituent is located it was identified that the α -substituent would cause steric interactions with both the zinc atom and the alkyl group. In contrast, for conformer **B** the α -substituent would interact mainly with the zinc atom. By similar reasoning, it was described how a β -substituent would favour conformer **A**, due to stronger steric interactions between the substituent and both the substrate aldehyde and the alkylating reagent in conformer **B**, where the substituent is positioned in the axial position.

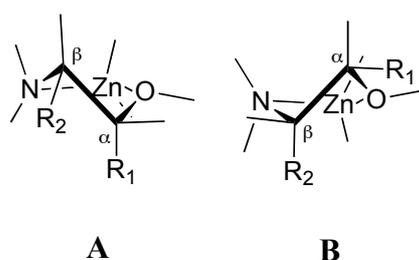
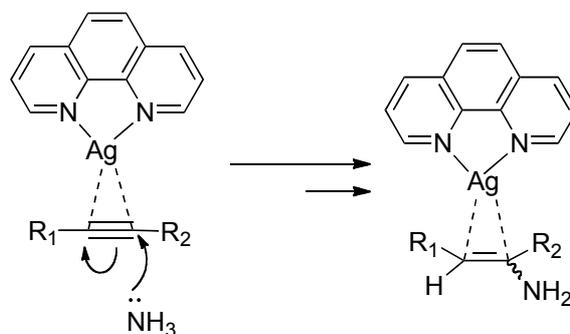


Figure 10 Illustration of interconverting ligand-zinc complexes.

Silver-Catalyzed Hydroamination

Wiest and colleagues have developed a Q2MM force field for the silver-catalyzed hydroamination of alkynes [33]. This reaction is especially interesting if it proceeds intramolecularly to generate novel heterocycles. The active form of the catalyst is a π -complex between Ag(I), the ligand, and the alkyne where the ligand normally is a phenanthroline derivative. The Ag-complexation activates the alkyne for nucleophilic attack by the amine nitrogen, as exemplified in Scheme 6.



Scheme 6 Schematic picture of Ag(I)-catalyzed hydroamination of alkynes

Subsequent proton transfer to the α -carbon will generate the hydroaminated alkyne. In the development of parameters for the Q2MM force field, three different amines, two alkynes, and two silver-ligand complexes were combined (Figure 11). Overall, good agreements between QM and Q2MM structural data were achieved.

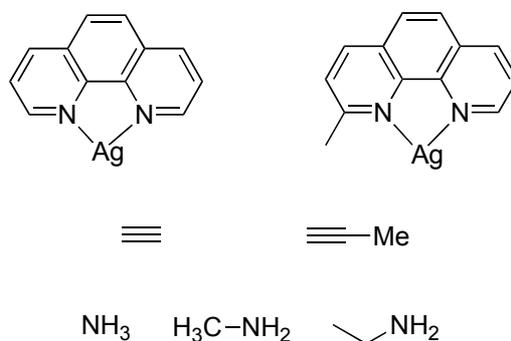
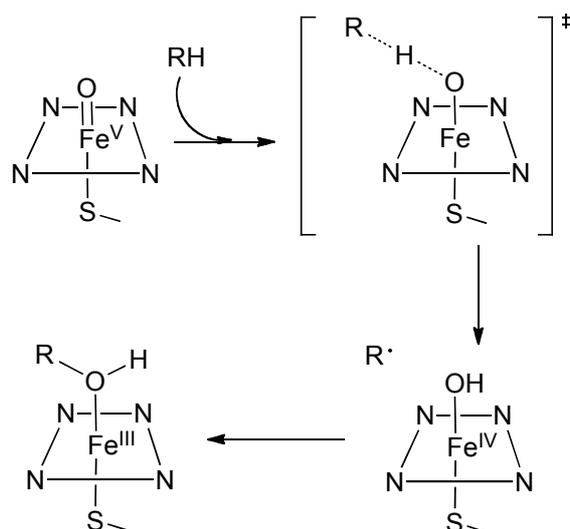


Figure 11 Reactants used for Q2MM development.

P450 Hydrogen Abstraction

A broader application for Q2MM was found by the development of a force field for the hydroxylation reaction catalyzed by the mono-oxygenase cytochrome P450 with a heme group in the active site, or rather the initial hydrogen abstraction which is the selectivity-determining step for the hydroxylation reaction. A very reactive Fe(V) oxo complex formed upon heterolytic O₂ cleavage is electron deficient enough to oxidize even C-H bonds as illustrated in Scheme 7. The oxygen abstracts a hydrogen from the substrate (RH) in the rate limiting step forming a Fe(IV)-OH intermediate and a parent substrate radical. These combine in a second step under formation of a hydroxylated substrate, the so-called rebound mechanism.



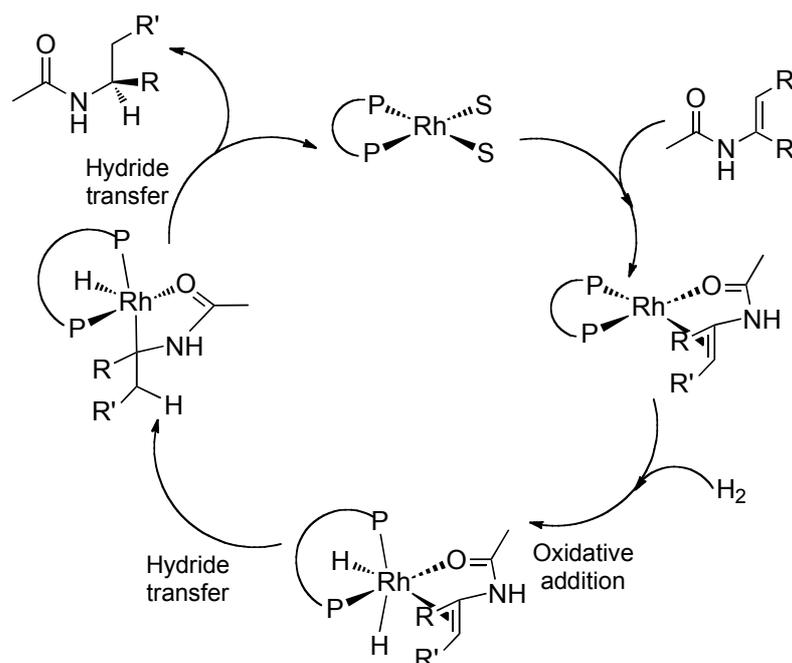
Scheme 7 Hydroxylation catalyzed by cytochrome P450

In the study by Rydberg, Olsen, Norrby, and Ryde, Q2MM parameters were optimized for C(sp³)-H abstraction from 14 compounds [25]. This is an unusually large training set, but the authors wanted to generate a broad and versatile force field, aiming also for molecular dynamics simulations and docking studies. The force field

for the substrate itself was not optimized, rather the standard GAFF force field was used. An excellent correlation coefficient ($r^2 = 0.99$) between QM and Q2MM Hessian elements was presented including more than 180,000 data points. The method was also evaluated based on geometry comparisons with DFT optimized structures, and small RMSD values (*e.g.* 0.010 Å in bond distances) were reported for the training set. The RMSD result for a test set comprising of 10 molecules was of equal magnitude (0.012 Å), indicating that the Q2MM force field is generally applicable for generating TS-structures for the target reaction. However, the structures were found not good enough in comparison to DFT optimized as evaluated by energy criteria. The Q2MM optimized structures were *ca.* 24 kJ mol⁻¹ less stable than DFT optimized structures. Including also the substrate in the parameterization reduced the energy error to *ca.* 9 kJ mol⁻¹, indicating the main source of error. In a later study [34], the designed Q2MM force field was used for transition state docking of the drugs progesterone and flunitrazepam into two different isoforms of human cytochrome P450s.

Asymmetric Hydrogenation of Enamides

The most recent application of Q2MM has been in two studies on the enantioselective Rh(I) catalyzed hydrogenation of enamides yielding novel α -amino acids. Mechanistic studies have shown that the reaction takes place in a multistep fashion [51]. After initial substrate binding to the Rh-catalyst, oxidative addition of H₂ occurs, presumably yielding a dihydride complex, even though this has not yet been observed experimentally. Transfer of one of the hydrides to the enamide is generally the rate-limiting step, although the TS for oxidative addition is close in energy in many cases. The observable alkyl hydride complex then undergoes reductive elimination forming the product (Scheme 8) [51].

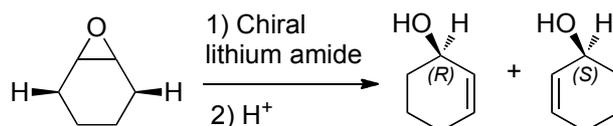


Scheme 8 Catalytic cycle for Rh(I) catalyzed hydrogenation of enamides (S = solvent).

The H₂ addition step to the Rh-complex has been found to occur parallel to the P-Rh-alkene bond, either from below or above. This gives rise to two different alkyl hydride complexes, either the first hydride adds to the β- or α-carbon, respectively. The route of these two pathways has been shown to be dependent on the substrate. For dehydro-α-amino acids the rate limiting hydride addition has occurred on the β-carbon, while for dehydro-β-amino acids it has occurred on the α-carbon. Based on these assumptions, the parameters for a Q2MM force field were generated based on a training set consisting of four different diphosphine ligands operating on one substrate. Evaluation of QM and Q2MM derived structure resulted in very small RMSD values (0.03 Å) for all bonds except for the C-H bond that experienced a slightly larger deviation (0.10 Å). The analysis of relative energies for different diastereoisomeric TS's of the training set revealed good agreement, as low as 1 kJ mol⁻¹ [35]. The derived Q2MM force field was then applied to nine new substrates with two chiral ligands and compared with experimental data. The deviation between calculated and experimentally derived energy differences between diastereoisomeric TS's was in the range of ≈1 kJ mol⁻¹. This energy difference allows for a rapid evaluation and selection of “good” and “bad” ligand-substrate combinations in a screening process. This was demonstrated in a paper by Donoghue, Helquist, Norrby, and Wiest where 14 chiral diphosphine ligands and 7 substrates were screened [36]. The calculated enantioselectivities were compared with experimental results, and the authors found an overall good agreement, with a mean unsigned error (MUE) of *ca.* 3 kJ mol⁻¹. The correlation between experimental and calculated ee was found to be excellent ($r^2 = 0.92$). The largest deviations were isolated to a few excessively hindered ligands, suggesting a change in mechanism for these special cases.

Epoxide Opening Using Chiral Lithium Base

The last example of Q2MM force field applications is the desymmetrization of *meso*-epoxides [52]. In this reaction, in which a chiral lithium amide discriminates between two enantiotopic *syn*- β -protons, the epoxide is transformed into a chiral allylic alcohol (Scheme 9) [53, 54]. In the study by Brandt, Norrby and Andersson [52], a preliminary Q2MM force field was developed and used as an initial screening tool for three chiral bases to generate TS-structures for subsequent DFT optimizations, but no data on the force field or the Q2MM procedure were included.



Scheme 9 Desymmetrization of cyclohexene oxide using a chiral lithium amide, yielding chiral allylic alcohols

Q2MM Tutorial

The current version of Q2MM has been implemented as a package of Unix/Linux scripts and programs written in C. The package has primarily been used with the Schrödinger [55] suite of programs including Maestro, Jaguar, and MacroModel, using the MM3* force field specifically; but in its fundamental form the Q2MM method can be implemented with a range of force fields and using different programs (like Gaussian [56]) to acquire QM data. An example is the P450 hydroxylation force field in AMBER [25, 34]. This tutorial is designed to instruct in the elements necessary for the parameter optimization process regardless of which force field or QM programs are used. Implementation-specific instructions on how to use the Q2MM method are available on the web for MM3* in MacroModel [57] and for AMBER [58], respectively.

This tutorial will be broken up into four sections. The first section discusses the design and development of new parameters for the system being studied. The second section describes the creation of a QM training set. The third section shows how to set up QM reference data so that it can be properly used during the parameter optimization process. The fourth section is a guide to running the optimization of force field parameters.

The basic flow of the Q2MM parameterization process is depicted in Figure 12 [21]. The process starts off by selecting a suitable basic force field and collecting QM reference data. Parameters for the system's unique structural interactions should be estimated within the force field based on QM data. The Q2MM method then goes through an iterative process of parameter refinement, internal validation, and external verification to create a reaction specific TSFF.

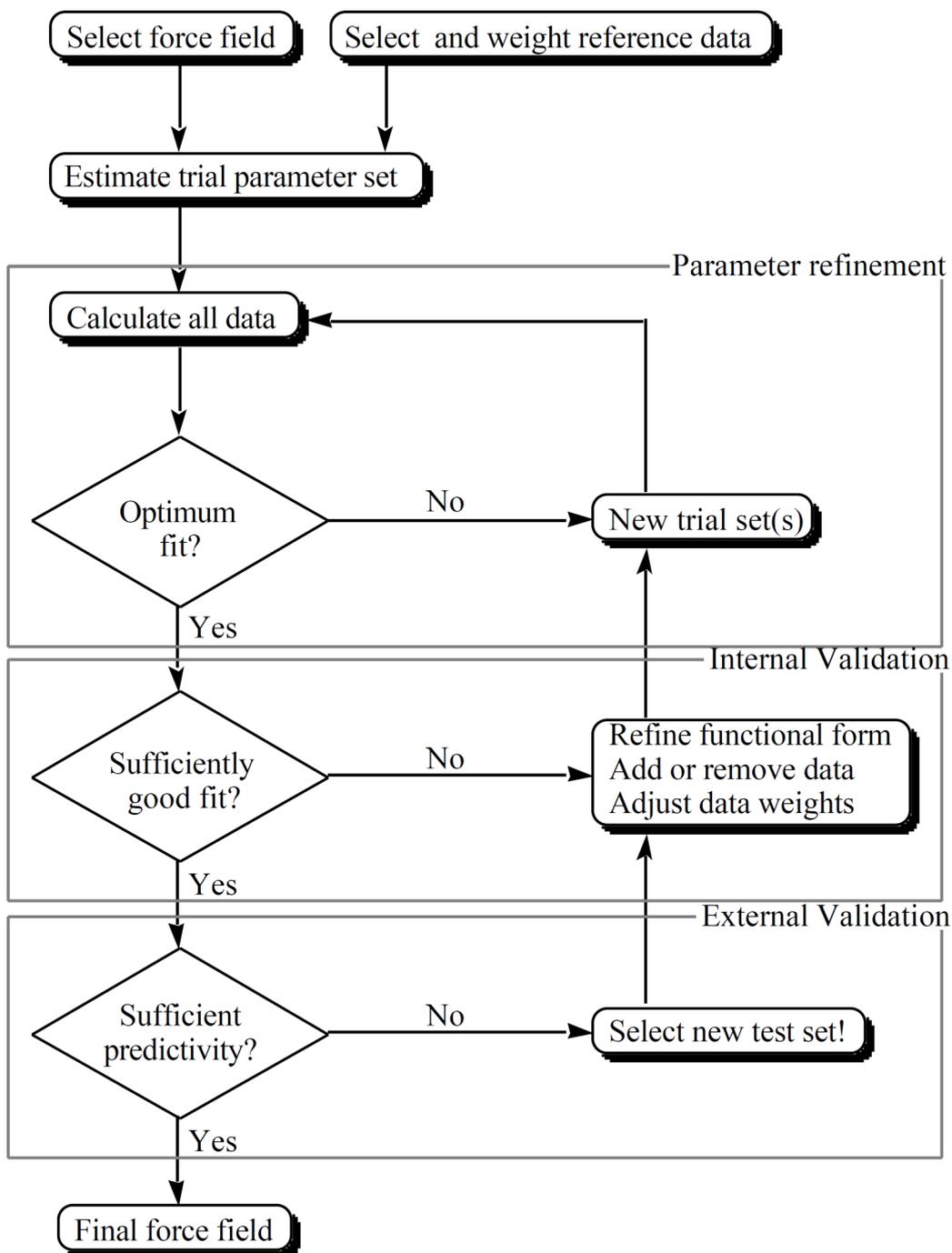


Figure 12 Parameterization Flow Chart.

1 Initial Force Field Development and Design

This section is devoted to the selection of a force field functional form, setup of basic force field parameters, and definition of atom types for the system of interest.

1.1 Choosing a Force Field

The first fundamental choice in using the Q2MM method is choosing an appropriate underlying force field. Parameters within basic force fields are generally derived from empirical data. A variety of basic force fields are available and each one has its pros and cons in their ability to deal with specific chemical systems. The ability of a force field to reproduce empirical data can only be as good as the parameters implemented into them. The chosen force field should already contain suitable parameters for the parts outside the reaction center, to minimize the number of parameters that need to be optimized, and to avoid fitting ground state parameters in a TSFF optimization. Basic force fields have been well scrutinized in the literature in their compatibility with numerous chemical systems [4, 59].

1.2 Molecular Modeling of Transition States

Most force fields are designed to use the parameters to find global energy minima of structures with a conformational search. Transition states, however, are saddle points on the PES defined by negative curvature. However, traditional MM needs the stationary point to be a minimum, and therefore a "transition structure" must be created for which the curvature is positive along all normal modes. This process of creating transition state force fields is well represented in the literature [6-9]. Q2MM specifically modifies the QM Hessian matrix for the transition structure along the normal mode with the large negative frequency to accomplish this. In this way it is possible to create a reaction specific force field for transition states frozen at one specific position of the reaction coordinate, but in all other respects closely mimicking the QM model.

1.3 Force Field Parameters

Each force field will be described by a functional form, similar to Figure 13 but in most cases more complex terms are incorporated [59, 60, 61, 62, 63, 64, 65, 66]. The total energy is described as a sum of bonded and nonbonded terms.

$$E_{\text{tot}} = \sum E_s + \sum E_b + \sum E_t + \sum E_{el} + \sum E_{vdW} + \sum E_{\text{other}}$$

Simple harmonic, diagonal force field:

$$E_s = k_s(l-l_0)^2 \quad E_b = k_b(\theta-\theta_0)^2 \quad E_t = v\cos n\omega$$

$$E_{el} = \frac{q_i q_j}{\epsilon r} \quad E_{vdW} = \frac{A}{r^{12}} - \frac{B}{r^6}$$

Figure 13 Sample functional form for a simple force field.

In Figure 13, l , θ , ω , and r are observables for the structures to be modeled (bond lengths, angles, and interatomic distances), whereas for example k_s and l_0 , the

force constant and reference value for a bond, are parameters to be optimized in the Q2MM refinement. The equation also contains parameters that should *not* be refined, like n , the torsional periodicity (*i.e.*, the number of expected minima when rotating the isolated torsion through 360° ; 2 for a double bond, 3 for many single bonds). In addition, vdW parameters (A and B in Figure 13) are an integral part of the underlying force field and should generally not be adjusted unless a new type of element is added to the force field. Different force fields will have slightly different terms within the functional form, since each specializes in describing certain types of interactions. This is what makes each force field unique in its ability to describe specific systems.

While trying to create a reaction specific force field, the way in which new parameters are set up is something of an art, and a very crucial process of the Q2MM method. The new parameters are what will be manipulated and refined in order for the force field to reproduce structures from QM data, or predict new ones. Q2MM can be used to refine existing parameters already defined by the force field, or for entirely new parameters to be added. For the latter, it is necessary to estimate initial values for all parameters before actual refinement. Interactions that are generally refined by the Q2MM method are bond stretches, angles, torsions, and charges or dipoles (k_s , l_0 , k_e , θ_0 , ν , and q , Figure 13). When modeling a transition state the bonds that are being broken or formed need to be carefully considered. Once these new interactions are described within the force field the Q2MM method will need to extract each of these parameters for optimization.

1.4 Atom types

The basic unit of a force field is the atom type. In designing a force field with new parameters, the basic atom type list may lack elements specific to the structure of choice, or the existing elements may need to be augmented for explicit interactions. Oftentimes there is an option within the force field to add in desired atom types. For example, some transition metals are not easy to describe in a general sense and may not be initially incorporated in the basic atom type list.

Besides just adding in elements it is also possible to add in different ‘types’ of atoms. For example, most available force fields differentiate between sp , sp^2 , and sp^3 carbons, and some further differentiate between, for example, alkene and carbonyl carbons. With different atom type names for the same element, it makes it easier to define geometry within a substructure where it is appropriate. When modeling a transition state in the force field it can be beneficial to use new atom types for the reactive center since they will have significantly different properties. With more specific atom types the process of defining the substructure and its corresponding parameters becomes more simplified. On the other hand, use of existing atom types frequently means that a number of essential parameters (that may not have to be reaction specific) could already be available.

2 The QM Training Set

2.1 *Choosing a training set*

The training set of QM data can be from any level of theory and basis set that is deemed practical and sufficiently accurate for the system being studied. An optimal QM training set should include several structures that highlight key variances in the transition structures being studied. This variance includes significant steric and electronic interactions that may influence the transition state geometry. Transition structures should also be chosen so that there are several sets of diastereomers or conformers. The reproduction of relative energies is crucial in the development of MM parameters; therefore the number of relative energies should be maximized to its practical extent. The required size of the QM training set cannot be specifically defined, but must be determined independently for each reaction being studied. Force field parameters that are developed around one or two QM structures may be very specific and not adequately responsive to changes in steric or electronic environment. While more QM structures are preferable, it is possible to overdefine the system so that the parameters end up not fitting any structures particularly well, but optimization yields no significant improvement. Additionally, an excessive QM training set may be impractical to calculate in a reasonable amount of time. The ideal training set should not be a systematic study of several ligand and substrate combinations, but rather a representative sampling of combinations. The process of identifying and locating the appropriate structures for use in the QM training set is based on chemical knowledge of the system and is not described here. For examples of appropriate training sets consult pertinent publications [25, 28, 29, 32, 33, 35].

We want to make one additional point specifically with respect to nonbonded interactions in QM calculations. Many common methodologies, including HF and several brands of DFT, lack an accurate description of dispersion forces [67, 68]. On the other hand, the attractive dispersion is usually well represented in molecular mechanics force fields. This interaction is frequently critical in describing barriers in asymmetric catalysis [31], since the difference between diastereomeric transition states arises exclusively from nonbonded interactions between side groups. Thus, it is essential that the good vdW description is retained within the force field, and not deteriorated by fitting to methods that are, from this perspective, inferior. This is the main reason why model systems should be kept as small as possible, to avoid significant and erroneous nonbonded interactions between substituents in the QM calculations.

2.2 *QM Generation of the Training Set*

It is important to choose an appropriate QM method and basis set for the system being studied. It is especially important if a transition metal is involved in the system being studied. The application of the force field should also be considered when choosing the level of theory. If the primary use of the force field is to differentiate between diastereomeric transition states for the prediction of enantioselectivity then it will be

necessary to have a level of theory that can accurately determine relative energies, and in particular the distortion cost at the reaction center. This will be the most important consideration when using the Boltzmann distribution to determine the preference for one enantiomer over another based on relative energies of diastereotopic transition states produced by the custom force field. A force field is only as good as the parameters that are programmed into it, so choice of method to calculate QM reference data is crucial.

Required QM information necessary for parameterization are the optimized geometries and energies, QM partial charges and dipole moments, and Hessian elements from frequency calculations. MM force fields are usually developed to produce enthalpies, since they are mostly parameterized towards experimental data. To correspond to this paradigm, QM energies should preferably be converted to enthalpies, or at least corrected for differences in zero point energy. With respect to the electrostatic data, there are many ways to partition QM electron densities into atomic charges or bond dipoles. However, in our experience, charges that are designed to give an accurate representation of the electrostatic field around the molecule, like RESP [69] or CHELPG [70], tend to give good results.

The Hessian data is essential, since this is the only type of data that describes the energy response to small structure changes, and the only data that has a strong correlation to the MM force constants. Without this strong correlation to at least some of the data, parameters become ill-defined, and consequently predictions using that parameter become unreliable. As already discussed, the Hessian data must be modified to correspond to an energy minimum, that is, to have only positive eigenvalues. This requires a diagonalization of the Hessian, preferably in mass-weighted form. Many QM programs simplify this procedure by giving the eigenvalues and normal modes as output. By definition, the Hessian matrix \mathbf{H} can be obtained by the matrix multiplication $\mathbf{H}=\mathbf{X}^T\mathbf{W}\mathbf{X}$, where \mathbf{X} is the orthonormal matrix of all eigenvectors, \mathbf{W} is a diagonal matrix of the eigenvalues, and \mathbf{X}^T is the transpose of \mathbf{X} . The eigenvalue matrix \mathbf{W} should contain one strongly negative value corresponding to the reaction coordinate, six values close to zero for the molecular translations and rotations (unless these have already been removed by projection), and the remainder of the diagonal elements positive. \mathbf{W} is now modified to \mathbf{W}' by replacing the negative value with a large positive value. In atomic units, we usually chose the value "1", which corresponds to a frequency of about 5000 cm^{-1} , that is, a very strong bond. The modified Hessian matrix to be used in the parameterization is now trivially obtained by matrix multiplication: $\mathbf{H}'=\mathbf{X}^T\mathbf{W}'\mathbf{X}$. In the parameterization, each matrix element of the Hessian is then compared to the corresponding element calculated by MM. Since the Hessians are Cartesian, it is of course essential that the MM and QM calculations use identical Cartesian coordinates for the molecules.

3 Preparing Parameter Optimization

3.1 Penalty Function

The optimization of new parameters is accomplished through the use of a penalty function to ensure agreement with the QM reference data. The penalty function is the weighted sum of the squared deviation from QM and MM data points as shown in equation 1.

$$\chi^2 = \sum_i w_i^2 (\hat{y}_i - y_i^\circ)^2 \quad (1)$$

This penalty function, χ^2 , is the weighted sum of the squares of all deviations of the calculated values, \hat{y}_i , from the reference values, y_i° [20]. Each parameter's absolute deviation has a different effect on the penalty function. Changing a bond or an angle by the same value will have drastically different effects on the penalty function. Each parameter, therefore, is assigned a different weight (w_i) to compensate. The weighting factors are defined as the inverse of the tolerance for a given type of data in the penalty function. Through the use of weighting values the penalty function becomes unitless with concerns to the importance and magnitude of each type of data. Norrby and Liljefors used the relative weights of 100:2:1 for bond lengths, angles, and torsional angles, respectively, in their original work [20], meaning that the acceptable average errors were set to 0.01 Å for bond lengths, half a degree for angles, and one degree for torsional angles. Weighting factors for other types of data can also be found in the literature [20, 28, 29, 71, 72, 73].

3.2 Conversion of QM Structures to Parameterization Format

With the QM reference data in hand, it will need to be extracted into a format that can be read by the force field program being used. Whether this is done with a conversion utility or done manually through a graphical interface by switching atom types individually and exporting as a different format does not matter. The important thing during this stage is that the optimized geometry and partial charge information from the QM data remains intact. Pay particular attention to the atom types, bond types and bond lengths after converting to a new format. Bond lengths may be modified by inherent settings within the program and will need to be changed. In particular, as already noted, it is essential that the orientation is retained from QM structures for calculation of Cartesian Hessians. Any structure rotation will completely invalidate the comparison.

3.3 Initial Estimate of Parameters

Parameter refinement works better the closer to the optimum it is started. Many types of data points are sensitive to changes in more than one parameter, and large initial errors may cause unintended changes in the "wrong" parameter. Thus, it is advantageous to set all initial parameters to "reasonable" values. Bond and angle reference parameters should be set to average values from the QM training set. van der Waals parameters must also be well estimated, preferably from ground state experimental data, whereas electrostatic parameters can be set as close as possible to

the QM averages [21]. Force constants, torsional parameters, and cross terms can be more easily manipulated by the automated refinement and don't need to be so rigidly defined before optimization – similar values from the existing force field should be sufficient.

3.4 Tethering

If early on during an optimization certain values drift to unrealistic values it will be prudent to “tether” certain values in order to prevent this for the initial optimization iterations. Tethering adds an additional weight to the penalty function for certain parameters making it mathematically less advantageous for that parameter to be modified. This term is the squared deviation of the parameter from the assumed optimum value. In the majority of cases only weak tethering will need to be applied. Tethering can be very useful during the first couple of iterations and is very easy to implement. The final optimization should always be done without any tethering of parameters.

Tethering certain values is accomplished by implementing harmonic terms of the form $c_i(p_i - p_i^*)^2$ to the penalty function. The p_i term is the current parameter value, p_i^* is the assumed optimal value, and c_i is the tethering force constant. A common practice is to tether bond and angle reference values to the averages found in the QM calculations. The choice of tethering force constant is rather arbitrary. A very large value will in effect freeze the parameter, and this effect could more easily be achieved by simply removing it from the list of parameters to refine. Experimentation is recommended, and in any case, the tethering should be reduced as the refinement is progressing, and removed completely before the force field can be considered converged (*cf.* section 4.3).

4 Parameter Optimization

4.1 Running a Parameter Optimization

With all of the initial parameters compiled, all data points can be calculated and compared to the reference data via the penalty function with suitable weighting factors (equation 1). The task now is varying the parameters in order for the penalty function to reach a minimum. There are a number of different strategies for this [74], but a combination of Simplex and Newton-Raphson (NR) optimizations has proven very effective [20]. The Newton-Raphson method is a very efficient method for finding roots of well-behaved functions, especially with dampening functions like in the Levenberg-Marquadt method [74]. When parameters become too strongly interdependent, or if they are in a parameter space of negative curvature, Newton-Raphson optimization may have difficulty reaching convergence. An alternation between Newton-Raphson and Simplex techniques has been shown to help reduce irregularities in convergence [29, 73]. This combination of methods performs one NR-based refinement of the entire set, selects poorly-defined parameters like those

with a negative second derivative of the penalty function, and performs a Simplex optimization upon the subset [22].

4.2 Initial refinement

Some types of parameters, in particular the nonbondeds, have a strong influence on every other aspect of the force field. There is no real reason to believe that the vdW parameters should be special in transition states, and these are therefore best determined for ground states, and preferably from experimental data since many QM methods give questionable results for these weak interactions [67]. The electrostatic parameters, on the other hand, must be determined specifically from QM calculations for transition states. It is recommended that the electrostatic parameters are adjusted to the closest possible fit with average QM electrostatics for several transition structures, and then left out of the refinement. It is of course necessary to enforce a correct total charge, if the electrostatic scheme is based on point charges.

The second round of optimization involves determining a better “initial guess” for the force constants. One simple method to achieve this is to optimize force constants alone against a reduced penalty function including only Hessian data. Another option involves tethering the bond and angle values, instead of holding them constant. If there are torsions to be optimized, an initial optimization of selected torsion constants can be included in this step as well.

4.3 Setting Convergence Threshold

Optimization of the new parameters is done through the optimization of the penalty function. Convergence of the penalty function is determined through the rate of change in the penalty function. Convergence is defined, in general, as being a difference of less than 0.01% between consecutive optimization steps. This level of conversion is necessary for the final optimization of force field parameters, but not essential for every step of the parameterization. For initial optimization steps, or optimization of subsets of parameters, the penalty function is sufficiently converged at a lower threshold, generally set at 1% difference.

4.4 Refinement Strategies

The initial stages of a parameter refinement should be monitored. In the early stages of refinement, errors are still large and parameters will be more prone to deviate to unnatural values. There are a couple of strategies for curbing these deviations. One method would be to break up the parameters into subsets early on. Similar to the recommendation of optimizing electrostatic parameters before bonds, angles, force constants, and torsions – these different types of parameters can also be broken up further and optimized separately. Torsional parameters and cross-terms can be considered as correction terms that with advantage can be set to very low or zero values initially, and only brought into the parameterization after the more basic bond and angle terms have been reasonably well converged. Already discussed was the option of setting up parameter tethering to further bias the penalty function in concern

to specific parameters of choice. Another thing to be on the look out for are data outliers that are drastically effecting the refinement. Careful scrutiny of outliers is necessary, so that a judgement of whether augmenting or removing the data point will improve the functionality of the force field.

It is also necessary to be aware of potential border effects. In a Q2MM optimization, most of the underlying force field is left intact, and only parameters close to the reaction center are included in the refinement. Sometimes the initial choice of parameters to optimize fails to recognize that certain parameters should have been modified. If so, the effect is frequently to cause some parameters under refinement to get non-physical values. An example would be electrostatics, where an unmodified charge too close to the reaction center might adversely affect charges that are refined.

4.5 Post-parameterization analysis

After the parameterization has converged there are a number of things that should be checked to ensure that the solution found is good, and that it is chemically relevant.

First, the penalty function should have a reasonable size, which is similar to the number of data points if the weighting factors have been set to the inverse of the acceptable error [22]. Second, for a sound solution all the second derivatives of the penalty function with respect to the parameters should be positive, and the second derivatives should be larger than the respective first derivatives.

If either of these conditions is not fulfilled, the optimization has not converged. There can be several reasons for this, but the most common is that the optimization ended up in a local minimum that is not physically correct. To remedy the problems with unphysical parameters there are two solutions. One is to freeze these parameters by setting them to some reasonable value and exclude them from the parameterization. Once the optimization has converged the frozen parameters can be included in the parameterization again. The second solution is most useful if the equilibrium values of bond and angle parameters tend to drift to values far off from the QM values. Then one can use tethering, as already described. A high penalty is then created if these parameter values drift away from the equilibrium value. The tethering should be removed once the optimization has converged, and the optimization should then be restarted without the tethering to assure that an artificial minimum has not been found. If some parameters consistently give unphysical values even when all other parameters seem close to convergence, this is an indication that the data set is incomplete making the parameter badly defined, or that the current functional form is insufficient to describe the interactions in the system. Sometimes the problem can be alleviated by bringing in more parameters from the underlying force field into the refinement process.

If the conditions above are fulfilled the convergence is good. However, this does not imply that the parameters are good, this solution could be a local minimum, or include many unphysical parameters. To ensure that the parameters reproduce the

QM results one should compare the QM geometry of the system to the geometry optimized with the force field to ensure that it is well reproduced. One should also compare the first and second derivatives of the energy (forces and Hessians, respectively) for the QM and MM calculations. Outliers in these comparisons indicate that the parameter optimization could have gone wrong. Due to the nature of force fields, some elements of the Hessian cannot be reproduced, since the functional form does not include the required terms. Such outliers in the Hessian comparison usually have to be ignored.

Finally, force constants and equilibrium values for the parameters should be physically sane. Thus, we expect that the force constants should be of a size reasonable for the force field which is used, and equilibrium values should not be far off from the values found in the QM optimized geometries.

In most problematic cases, when either the convergence is not good, or the parameters are not physically correct, the optimization has ended up in a local minimum. This local minima is then either not the global minima, or the global minima is represented by unphysical parameters. One solution is then to freeze parameters during the initial optimization as described above.

The NR-steps can be analyzed to get a confidence interval for the newly determined parameters. In this step, the first and second derivative of the penalty function with respect to each parameter is determined. The expected change in penalty function upon small single-parameter changes can then be calculated from a simple Taylor expansion. By specifying a specific increase of the penalty function (typically 0.1%) as being a significant disimprovement, it is possible for each parameter to calculate how large a change can be made before resulting in a significant penalty function decrease. This can be reported as a "range" for each parameter, and gives an indication how well that particular parameter is determined by the current data set.

Finally, the ultimate test for a Q2MM force field is in how well it reproduces experimental data. For a true test of predictivity, the data must not have been included in the parameterization. The standard test would be to do full conformational searches for each diastereomeric TS, calculate the final selectivity from a Boltzmann distribution, and compare to experimental selectivities (e.g. Figure 7). As demonstrated in the "History" section, Q2MM has been able to achieve excellent predictivity, in particular in the most recent case, the rhodium-catalyzed asymmetric hydrogenation [36].

Concluding remarks

The Q2MM method described in this overview presents one possible approach to answer the question by Brown and Deeth "Is Enantioselectivity Predictable in Asymmetric Catalysis?" [1] with a clear, though maybe not universal "Yes". Originally starting as a tool to explain the origin of stereoselection [7], computational methods are now accurate enough to be useful as predictive tools for the selection and

design of new ligands for asymmetric catalysis [75]. The wider adoption of Q2MM-based predictive tools has the potential of transforming the area of catalytic asymmetric synthesis from a time- and resource intensive trial-and-error approach to a rational of virtual screening approach where only a small number of ligands need to be experimentally tested to achieve satisfactory results. Such a development would be in analogy to the developments in drug design, where structure-based design and virtual screening methods are now indispensable tools for the medicinal chemist.

As outlined in the review, the predictions made by computational methods can be either improved mnemonics as in the case of the osmium-catalyzed dihydroxylation [31] or quantitatively accurate predictions of e.e. as in the case of the Rh(I) catalyzed hydrogenations of enamides [36]. Other reactions are sure to follow suit in upcoming years. Nevertheless, the upfront effort that needs to be spent for the determination of the mechanism and the parameterization of the TSFF means that reactions to be studied by Q2MM will typically be limited to reactions of significant scientific and/or commercial interest. Furthermore, such reactions will ideally have a single, well-understood stereoselecting transition state across a wide range of ligands and substrates as well as small changes in geometric and electronic structure at the metal center along the reactions pathway.

Application of the Q2MM method to stereoselective reactions can be compared with other methods for similar studies. First, although a direct comparison is difficult to make because different reactions were studied, the correlation coefficient of $R^2=0.96$ between theory and experiment from Q2MM calculations in the best case of the Rh(I) catalyzed hydrogenations [35, 36] is significantly better than the e.e. predictions made for a variety of reactions using QM/MM methods [76], QSPR methods [2, 77, 78, 79, 80, 81] or the model potential originally introduced by Warshel in the EVB program [12, 13], and recently applied to asymmetric catalysis as ACE by Moitessier and coworkers [16, 82], where correlation coefficients were in the range of 0.65 to 0.8. This comparison suggests that carefully parameterized force fields are at the very least of comparable and possibly of superior accuracy than either QM/MM methods or ACE. There are a number of possible reasons. As already discussed earlier, a reasonable sampling of the conformational space is difficult to achieve with QM/MM methods even for small molecules. Furthermore, several of the QM methods used (*e.g.* the popular B3LYP hybrid functional) do not correctly describe dispersive interactions, which are empirically parameterized in force fields. Comparison of force field results with those from ACE and other related methods, which mix empirical force fields for the reactant and product, also supports the intuitive notion that a force field parameterized for the transition state is more accurate than a description of the transition state by ground state force fields. Second, the philosophy behind stereochemical predictions using the Q2MM is quite different from those from either purely QM or QM/MM predictions on one hand and EVB or QSPR calculations on the other. Quantum-based methods are fairly general and can, with sufficient effort, be quite accurate, but they are too slow for either conformational sampling to determine the free energy differences of the diastereoemeric transition structures nor for the screening of large ligand libraries. QSPR- and rule-based mnemonics are useful as qualitative tools, but tend to be case-

specific and not quantitatively accurate. Finally, EVB and related methods are fast and try to balance the accuracy of the results with the generality of the potential expression, many of which are available from empirically parameterized ground state force fields for common molecules. The Q2MM method can therefore be considered a special case of this approach where the reaction specific parameterization to the stereodetermining transition state yields highly accurate results with minimal computational effort.

Acknowledgment

Development of Q2MM is supported by the Swedish Research Council (VR 2007-4276), The Petroleum Research Fund (Grant PRF#47810-AC1) and the National Science Foundation (NSF CHE0833220). SONL gratefully acknowledges financial support from the Sven and Lilly Lawski Foundation. We want to thank AstraZeneca for funding and encouragement.

References

-
- [1] Brown, J.M.; Deeth, R.J. Is Enantioselectivity Predictable in Asymmetric Catalysis? *Angew. Chem. Int. Ed.*, **2009**, *48*, 4476-4479.
 - [2] Lipkowitz, K.B.; Kozlowski, M.C. Understanding Stereoinduction in Catalysis via Computer: New Tools for Asymmetric Synthesis. *Synlett* **2003**, 1547-1565.
 - [3] Ujaque, G.; Maseras, F. Applications of Hybrid DFT/Molecular Mechanics to Homogeneous Catalysis. *Struct. Bond.* **2004**, *112*, 117-149.
 - [4] Gundertofte, K.; Liljefors, T.; Norrby, P.-O.; Pettersson, I. A Comparison of Conformational Energies Calculated by Several Molecular Mechanics Methods. *J. Comput. Chem.* **1996**, *17*, 429-449.
 - [5] Liljefors, T.; Gundertofte, K.; Norrby, P.-O. Pettersson, I. Molecular Mechanics and Comparison of Force Fields. In *Computational Medicinal Chemistry for Drug Discovery*, Bultinck, P.; Tollenaere, J.P.; De Winter, H.; Langenaeker, W. (Eds.), Marcel Dekker: New York, **2004**, 1-28.
 - [6] Jensen, F.; Norrby, P.-O. Transition States from Empirical Force Fields. *Theor. Chem. Acc.* **2003**, *109*, 1-7.
 - [7] Eksterowicz, J.E.; Houk, K.N. Transition-State Modeling with Empirical Force Fields. *Chem. Rev.* **1993**, *93*, 2439-2461
 - [8] Norrby, P.-O. Molecular Mechanics as a Predictive Tool in Asymmetric Catalysis. Chapter 13 in *Transition State Modeling for Catalysis*, Truhlar, D.G.; Morokuma, K. (Eds.), ACS Symposium Series 721, American Chemical Society: Washington, DC, **1999**, 163-172.
 - [9] Norrby, P.-O. Selectivity in asymmetric synthesis from QM-guided molecular mechanics. *J. Mol. Struct. (Theochem)* **2000**, *506*, 9-16.
 - [10] van Duin, A.C.T.; Dasgupta, S.; Lorant, F.; Goddard, W.A., III, ReaxFF: A reactive force field for hydrocarbons. *J. Phys. Chem. A* **2001**, *105*, 9396-9409.

-
- [11] Jensen, F. Using Force Field Methods for Locating Transition Structures. *J. Chem. Phys.* **2003**, *119*, 8804-8808.
- [12] Åqvist, J.; Warshel, A. Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. *Chem. Rev.* **1993**, *93*, 2523-2544
- [13] Warshel, A. *Computer modeling of chemical reactions in enzymes and solutions*. Wiley, New York, **1991**.
- [14] Rappé, A.K.; Pietsch, M.A.; Wiser, D.C.; Hart, J.R.; Bormann, L.M.; Skiff, W.M. RFF, Conceptual Development of a Full Periodic Table Force Field for Studying Reaction Potential Surfaces. *J. Mol. Eng.* **1997**, *7*, 385-400.
- [15] Kim, Y.; Corchado, J.C.; Villa, J.; Xing, J.; Truhlar, D.G. Multiconfiguration molecular mechanics algorithm for potential energy surfaces of chemical reactions. *J. Chem. Phys.* **2000**, *112*, 2718-2735.
- [16] Corbeil, C.R.; Thielges, S.; Schwartzentruber, J.A.; Moitessier, N. Toward a Computational Tool Predicting the Stereochemical Outcome of Asymmetric Reactions: Development and Application of a Rapid and Accurate Program Based on Organic Principles. *Angew. Chem. Int. Ed.*, **2008**, *47*, 2635-2638.
- [17] Menger, F.M.; Sherrod, M.J. Origin of high predictive capabilities in transition-state modeling. *J. Am. Chem. Soc.* **1990**, *112*, 8071-8075.
- [18] Lifson, S.; Warshel, A. *J. Chem. Phys.* **1968**, *49*, 5116-5129.
- [19] Maple, J.R.; Hwang, M.J.; Stockfisch, T.P.; Dinur, U.; Waldman, M.; Ewig, C.S.; Hagler, A.T. Derivation of Class II force fields. 1. Methodology and quantum force field for the alkyl functional group and alkane molecules. *J. Comput. Chem.* **1994**, *15*, 162-182.
- [20] Norrby, P.-O.; Liljefors, T. Automated Molecular Mechanics Parameterization with Simultaneous Utilization of Experimental and Quantum Chemical Data. *J. Comput. Chem.* **1998**, *19*, 1146-1166.
- [21] Norrby, P.-O. Recipe for an organometallic force field. In *Computational Organometallic Chemistry*, Cundari, T. (Ed.), Marcel Dekker: New York, **2001**, 7-37.
- [22] Norrby, P.-O.; Brandt, P. Deriving Force Field Parameters for Coordination Complexes. *Coord. Chem. Rev.* **2001**, *212*, 79-109.
- [23] Hwang, M.J.; Stockfisch, T.P.; Hagler, A.T. Derivation of Class II Force Fields. 2. Derivation and Characterization of a Class II Force Field, CFF93, for the Alkyl Functional Group and Alkane Molecules. *J. Am. Chem. Soc.* **1994**, *116*, 2515-2525.
- [24] Halgren, T.A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490-519.
- [25] 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.
- [26] Villa, A.; Cosentino, U.; Pitea, D.; Moro, G.; Maiocchi, A. Force Fields Parameterization for Gadolinium Complexes Based on Ab Initio Potential Energy Surface Calculations. *J. Phys. Chem. A* **2000**, *104*, 3421-3429.
- [27] Verstraelen, T.; van Neck, D.; Ayers, P.W.; van Speybroeck, V.; Waroquier, M. The Gradient Curves Method: An Improved Strategy for the Derivation of Molecular Mechanics Valence Force Fields from ab Initio Data. *J. Chem. Theory Comput.* **2007**, *3*, 1420-1434.
- [28] Norrby, P.-O.; Brandt, P.; Rein, T. Rationalization of Product Selectivities in Asymmetric Horner-Wadsworth-Emmons Reactions by Use of a New Method for Transition-State Modeling. *J. Org. Chem.* **1999**, *64*, 5845-5852.

-
- [29] Norrby, P.-O.; Rasmussen, T.; Haller, J.; Strassner, T.; Houk, K.N. Rationalizing the Stereoselectivity of Osmium Tetroxide Asymmetric Dihydroxylations with Transition State Modeling Using Quantum Mechanics-Guided Molecular Mechanics. *J. Am. Chem. Soc.* **1999**, *121*, 10186-10192.
- [30] Fristrup, P.; Tanner, D.; Norrby, P.-O. Updating the asymmetric osmium-catalyzed dihydroxylation (AD) mnemonic. Q2MM modeling and new kinetic measurements *Chirality* **2003**, *15*, 360-368.
- [31] Fristrup, P.; Jensen, G. H.; Andersen, M. L. N.; Tanner, D.; Norrby, P.-O. Combining Q2MM Modeling and Kinetic Studies for Refinement of the Osmium-catalyzed Asymmetric Dihydroxylation (AD) Mnemonic. *J. Organomet. Chem.* **2006**, *691*, 2182-2198.
- [32] Rasmussen, T.; Norrby, P.-O. Modeling the Stereoselectivity of the β -Amino Alcohol Promoted Addition of Dialkylzinc to Aldehydes. *J. Am. Chem. Soc.* **2003**, *125*, 5130-5138.
- [33] Donoghue, P.J.; Kieken, E.; Helquist, P.; Wiest, O. Development of a Q2MM Force Field for the Silver(I)-Catalyzed Hydroamination of Alkynes. *Adv. Synth. Catal.* **2007**, *349*, 2647-2654.
- [34] Rydberg, P.; Hansen, S.M.; Kongsted, J.; Norrby, P.-O. Olsen, L.; Ryde, U. Transition-State Docking of Flunitrazepam and Progesterone in Cytochrome P450. *J. Chem. Theory Comput.* **2008**, *4*, 673-681.
- [35] Donoghue, P.J.; Helquist, P.; Norrby, P.-O. Wiest, O. Development of a Q2MM Force Field for the Asymmetric Rhodium Catalyzed Hydrogenation of Enamides. *J. Chem. Theory Comput.* **2008**, *4*, 1313-1323.
- [36] Donoghue, P.J.; Helquist, P.; Norrby, P.-O.; Wiest, O. Prediction of Enantioselectivity in Rhodium Catalyzed Hydrogenations. *J. Am. Chem. Soc.* **2009**, *131*, 410-411.
- [37] Hentges, S.G.; Sharpless, K. B. Asymmetric induction in the reaction of osmium tetroxide with olefins. *J. Am. Chem. Soc.* **1980**, *102*, 4263-4265.
- [38] Johnson, R.A.; Sharpless, K.B. Catalytic asymmetric dihydroxylation – discovery and development. In “*Catalytic asymmetric synthesis*”, Ojima, I. (Ed.), Wiley-VCH: New York, 2nd ed., **2000**, 357-398.
- [39] Sharpless, K.B.; Amberg, W.; Bennani, Y.L.; Crispino, G.A.; Hartung, J.; Jeong, K.-S.; Kwong, H.-L.; Morikawa, K.; Wang Z.-M.; Xu, D.; Zhang, X.-L. The osmium-catalyzed asymmetric dihydroxylation: a new ligand class and a process improvement. *J. Org. Chem.*, **1992**, *57*, 2768-2771.
- [40] Pidun, A.K.; Boehme, C.; Frenking, G. Theory Rules Out a [2 + 2] Addition of Osmium Tetroxide to Olefins as Initial Step of the Dihydroxylation Reaction. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2817-2820.
- [41] Dapprich, S.; Ujaque, G.; Maseras, F.; Lledós, A.; Musaev, D. G.; Morokuma, K. Theory Does Not Support an Osmaoxetane Intermediate in the Osmium-Catalyzed Dihydroxylation of Olefins. *J. Am. Chem. Soc.*, **1996**, *118*, 11660-11661.
- [42] Torrent, M.; Deng, L. Q.; Duran, M.; Sola, M.; Ziegler, T. Density Functional Study of the [2+2]- and [2+3]-Cycloaddition Mechanisms for the Osmium-Catalyzed Dihydroxylation of Olefins. *Organometallics*, **1997**, *16*, 13-19.
- [43] DelMonte, A.J.; Haller, J.; Houk, K.N.; Sharpless, K.B.; Singleton, D.A.; Strassner, T.; Thomas, A.A. Experimental and Theoretical Kinetic Isotope Effects for Asymmetric Dihydroxylation. Evidence Supporting a Rate-Limiting “(3 + 2)” Cycloaddition. *J. Am. Chem. Soc.*, **1997**, *119*, 9907-9908.
- [44] Corey, E.J.; Noe, M.C.; Grogan, M.J. Experimental test of the [3+2]- and [2+2]-cycloaddition pathways for the bis-cinchona alkaloid-OsO₄ catalyzed dihydroxylation

- of olefins by means of $^{12}\text{C}/^{13}\text{C}$ kinetic isotope effects. *Tetrahedron Lett.*, **1996**, *37*, 4899-4902.
- [45] Kolb, H.C.; van Nieuwenzhe, M.S.; Sharpless, K. B. Catalytic Asymmetric Dihydroxylation. *Chem. Rev.*, **1994**, *94*, 2483-2547.
- [46] Oguni, N.; Omi, T. Enantioselective addition of diethylzinc to benzaldehyde catalyzed by a small amount of chiral 2-amino-1-alcohols. *Tetrahedron Lett.*, **1984**, *25*, 2823-2824.
- [47] Kitamura, M.; Okada, S.; Suga, S.; Noyori, R. Enantioselective addition of dialkylzincs to aldehydes promoted by chiral amino alcohols. Mechanism and nonlinear effect. *J. Am. Chem. Soc.*, **1989**, *111*, 4028-4036.
- [48] Yamakawa, M.; Noyori, R. An Ab Initio Molecular Orbital Study on the Amino Alcohol-Promoted Reaction of Dialkylzincs and Aldehydes. *J. Am. Chem. Soc.*, **1995**, *117*, 6327-6335.
- [49] Yamakawa, M.; Noyori, R. Asymmetric Addition of Dimethylzinc to Benzaldehyde Catalyzed by (2S)-3-exo-(Dimethylamino)isobornenol. A Theoretical Study on the Origin of Enantioselection. *Organometallics*, **1999**, *18*, 128-133.
- [50] Kitamura, M.; Oka, H.; Noyori, R. Asymmetric addition of dialkylzincs to benzaldehyde derivatives catalyzed by chiral β -amino alcohols. Evidence for the monomeric alkylzinc aminoalkoxide as catalyst. *Tetrahedron*, **1999**, *55*, 3605-3614.
- [51] Feldgus, S.; Landis, C.R. Large-Scale Computational Modeling of $[\text{Rh}(\text{DuPHOS})]^+$ -Catalyzed Hydrogenation of Prochiral Enamides: Reaction Pathways and the Origin of Enantioselection. *J. Am. Chem. Soc.* **2000**, *122*, 12714-12727.
- [52] Brandt, P.; Norrby, P.-O.; Andersson, P.G. A DFT exploration of the enantioselective rearrangement of cyclohexene oxide to cyclohexenol. *Tetrahedron* **2003**, *59*, 9695-9700
- [53] Nilsson Lill, S.O.; Dinér, P.; Pettersen, D.; Amedjkouh, M.; Ahlberg, P. Development of chiral catalysts for stereoselective synthesis by deprotonations – Experimentation in interplay with computational chemistry *Adv. Quantum Chem.* **2004**, *47*, 1-22.
- [54] Diner, P.; Pettersen, D.; Nilsson Lill, S. O.; Ahlberg, P. Investigation of site selectivity of the stereoselective deprotonation of cyclohexene oxide using kinetic resolution of isotopic enantiomers in natural abundance. *Tetrahedron: Asymmetry*, **2005**, *16*, 2665-2671.
- [55] <http://www.schrodinger.com>
- [56] <http://www.gaussian.com>
- [57] MacroModel implementation: <http://www.nd.edu/~Q2MM>
- [58] AMBER implementation: <http://www.teokem.lu.se/~ulf/Methods/ponparm.html>
- [59] Burkert, U.; Allinger, N. L. *Molecular Mechanics*. ACS Monograph 177: Washington, **1982**.
- [60] Jensen, F. *Introduction to Computational Chemistry*. Wiley: Chichester, **1999**.
- [61] Goodman, J. *Chemical Applications of Molecular Modeling*. Royal Society of Chemistry: Cambridge, UK, **1998**.
- [62] Comba, P.; Hambley, T.W. *Molecular Modeling of Inorganic Compounds*. VCH: New York, **1995**.
- [63] Bowen, J.P.; Allinger, N.L. In *Reviews in Computational Chemistry*, vol. 2, Lipkowitz, K.B.; Boyd, D.B. (Eds.). VCH: New York: **1991**, pp. 81.
- [64] Dinur, U.; Hagler, A.T. In *Reviews in Computational Chemistry*, vol. 2, Lipkowitz, K.B.; Boyd, D.B. (Eds.). VCH: New York, **1991**, pp. 99.

-
- [65] Maple, J.R. In: *Encyclopedia of Computational Chemistry*, vol. 2, Schleyer, P.v.R. (Ed.), Wiley: Chichester, **1998**, pp. 1015.
- [66] Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham, T.E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.* **1995**, *91*, 1-41.
- [67] Nilsson Lill, S. O. Application of Dispersion-Corrected Density Functional Theory. *J. Phys. Chem. A* **2009**, *113*, 10321-10326.
- [68] Sousa, S.F.; Fernandes, P.A.; Ramos, M. J. General Performance of Density Functionals. *J. Phys. Chem. A*, **2007**, *111*, 10439-10452.
- [69] Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: The RESP model. *J. Phys. Chem.* **1993**, *97*, 10269-10280.
- [70] Breneman, C.M.; Wiberg, K.B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **1990**, *11*, 361-373.
- [71] Hagelin, H.; Åkermark, B.; Svensson, M.; Norrby, P.-O. Molecular Mechanics (MM3) Force Field Parameters for Calculations on Palladium Olefin Complexes with Phosphorus Ligands. *Organometallics* **1999**, *18*, 4574-4583.
- [72] Brandt, P.; Norrby, T.; Åkermark, B.; Norrby, P.-O. Molecular Mechanics (MM3) Parameters for Ruthenium(II)-Polypyridyl Complexes. *Inorg. Chem.* **1998**, *37*, 4120-4127.
- [73] Hagelin, H.; Åkermark, B.; Norrby, P.-O. New Molecular Mechanics (MM3*) Force Field Parameters for Calculations on (η^3 -Allyl)palladium Complexes with Nitrogen and Phosphorus. *Organometallics* **1999**, *18*, 2884-2895.
- [74] Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes in C*, 2nd Ed. Cambridge University Press: New York, **1992**.
- [75] Houk, K.N.; Cheong, P.H.Y. Computational prediction of small-molecule catalysts, *Nature* **2008**, *455*, 309-313.
- [76] Schneebeli, S.T.; Hall, M.L.; Breslow, R.; Friesner, R. Quantitative DFT Modeling of the Enantiomeric Excess for Dioxirane-Catalyzed Epoxidations. *J. Am. Chem. Soc.* **2009**, *131*, 3965-3973.
- [77] Ianni, J.C.; Annamalai, V.; Phuan, P.W.; Panada, M.; Kozlowski, M.C. A priori theoretical prediction of selectivity in asymmetric catalysis: Design of chiral catalysts by using quantum molecular interaction fields. *Angew. Chem. Int. Ed.* **2006**, *45*, 5502-5505.
- [78] Kozlowski, M.C.; Dixon, S.L.; Panda, M.; Lauri, G. Quantum mechanical models correlating structure with selectivity: Predicting the enantioselectivity of beta-amino alcohol catalysts in aldehyde alkylation. *J. Am. Chem. Soc.* **2003**, *125*, 6614-6615.
- [79] Kozlowski, M.C.; Panda, M. Computer-aided design of chiral ligands Part I. Database search methods to identify chiral ligand types for asymmetric reactions. *J. Mol. Graph. Mod.* **2002**, *20*, 399-409.
- [80] Kozlowski, M.C.; Panda, M. Computer-Aided Design of Chiral Ligands. Part 2. Functionality Mapping as a Method to Identify Stereocontrol Elements for Asymmetric Reactions. *J. Org. Chem.* **2003**, *68*, 2061-2076.

-
- [81] Kozlowski, M.C.; Waters, S.P.; Skudlarek, J.W.; Evans, C.A. Computer-aided design of chiral ligands. Part III. A novel ligand for asymmetric allylation designed using computational techniques. *Org. Lett.* **2002**, *4*, 4391-4393.
- [82] Moitessier, N.; Henry, C.; Len, C.; Chapleur, Y. Toward a computational tool predicting the stereochemical outcome of asymmetric reactions. 1. Application to sharpless asymmetric dihydroxylation. *J. Org. Chem.* **2002**, *67*, 7275-7282.