

# QM/MM calculations on selectivity

Jesús Jover,<sup>a</sup> Feliu Maseras<sup>b,c,\*</sup>

<sup>a</sup>*Inorganic Chemistry Department, Faculty of Chemistry, University of Barcelona, c/Martí i Franquès, 1, 08028 Barcelona, Spain*

<sup>b</sup>*Institute of Chemical Research of Catalonia (ICIQ), Avgda. Països Catalans, 16, 43007 Tarragona, Catalonia, Spain; fmaseras@icIQ.es*

<sup>c</sup>*Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalonia, Spain*

## 1. Introduction

## 2. QM/MM Methods

## 3. Computing Selectivity

## 4. Rhodium-Catalyzed Hydrogenation of Enamides

## 5. Rhodium-Catalyzed Hydroformylation

## 6. Copper-Catalyzed Cyclopropanation

## 7. Conclusions and Perspectives

## References

**Abstract** The application of QM/MM methods to the study of the reaction mechanisms involved in chemo-, regio- and enantio- selective processes has been a very productive area of research in the last two decades. This review summarizes basic general ideas in both QM/MM methods and the computational study of selectivity, and presents selected results on the study of three of the most representative examples of these applications: rhodium-catalyzed hydrogenation, rhodium-catalyzed hydroformylation and copper-catalyzed cyclopropanation.

**Keywords** QM/MM calculations, Selectivity, Hydrogenation, Hydroformylation, Cyclopropanation

## Abbreviations

AMPP	aminophosphane phosphinite
B3LYP	Becke 3-parameter Lee Yang Parr
BINAPHOS	2-(diphenylphosphino)-1,1-binaphthalene-2,2-diylphosphite
BP	Becke Perdew
Box	bis-oxazoline
DIOP	(2,3- <i>O</i> -isopropylidene-2,3-dihydroxy-1,4-bis(diphenylphosphino)butane)
DIPAMP	ethane-1,2-diylbis[(2-methoxyphenyl)phenylphosphane]
DuPHOS	1,2-bis(2,5-dimethylphospholano)benzene

ee	enantiomeric excess
er	enantiomeric ratio
IMOMM	Integrated Molecular Orbital Molecular Mechanics
L-DOPA	L-3,4-dihydroxyphenylalanine
MM	molecular mechanics
ONIOM	our own n-layered integrated molecular orbital molecular mechanics
Q2MM	quantum to molecular mechanics
QM	quantum mechanics
QM/MM	quantum mechanics/molecular mechanics
QSAR	quantitative structure activity relationships
UFF	universal force field
xantphos	4,5-bis(diphenylphosphino)-9,9-dimethylxanthene

## 1 Introduction

The application of computational chemistry to homogeneous catalysis has made large progress in recent decades [1-3]. This progress has been driven mostly by the increase of computer power and the development of new computational algorithms. One of the significant advances from a methodological point of view has been the widespread availability of quantum mechanics / molecular mechanics (QM/MM) methods. These methods had been traditionally developed for biochemical applications [4], but have gained importance in homogeneous catalysis mostly through the expansion of the ONIOM method [5].

QM/MM methods have been a good fit in systems where an accurate QM description is mandatory for a specific region, but a large system is nevertheless required for a proper description of the reactivity. It is not surprising thus that they have been applied widely in enzyme chemistry, where the reaction is usually confined at the active center. But this prescription fits also well to transition metal chemistry, where the metal and its immediate environment often require a QM method, but the steric effects of the potentially bulky ligands can be usually described in a satisfactory way by a force field. Regioselective and enantioselective catalysts occupy a prominent place among the transition metal complexes where the description of the chemistry requires the explicit introduction of the bulky ligands. These bulky ligands often carry the stereogenic centers that define the selectivity of the processes.

The application of QM/MM methods to different branches of chemistry has been the subject of a number of reviews. Some of them have in particular focused on its application to inorganic systems and homogeneous catalysis [6-12]. The use of computational chemistry for the characterization of selectivity has been also the subject of other reviews [13-15]. This review intends to fill a gap because it focuses on the specific application of QM/MM methods to selectivity. After two

sections dealing with a description of QM/MM methods and of computational approaches to selectivity, we will present a summary of results on three representative reactions that have been studied with these methods. The last section will collect conclusions and perspectives.

## 2. QM/MM methods

QM/MM methods are a particular case of multilayer methods. Research on this topic was awarded the Nobel Prize in Chemistry in 2013. Work by Warshel, Levitt and Karplus, the three Nobel laureates, is fundamental in the formulation of the approach and its early developments [16, 17]. A lot of theoretical chemists have contributed to further development of these methods, and among them, one must cite the groups of Thiel [4] and Morokuma [5, 18-20].

The basic idea behind any multilayer method is the use of different methods for the description of different regions of a chemical system. The energy description of QM/MM methods can be rigorously fit to a formula of the type:

$$E_{\text{QM/MM}}(\text{TOT}) = E_{\text{QM}}(\text{QM}) + E_{\text{MM}}(\text{MM}) + E_{\text{QM/MM}}(\text{QM/MM}) \quad (\text{Eq. 1})$$

In equation (1) the labels in subscript correspond to the methodological description and the labels in parentheses to the region involved. The interactions within the QM region are described at the QM level, and the interactions within the MM region are described at the MM level. The specific features of each QM/MM method are defined by the way it manages the interaction between the QM and MM regions. In principle, this interaction can be defined at either the QM or MM level, resulting in equation 2.

$$E_{\text{QM/MM}}(\text{QM/MM}) = E_{\text{QM}}(\text{QM/MM}) + E_{\text{MM}}(\text{QM/MM}) \quad (\text{Eq. 2})$$

The MM description of this interaction has been defined as mechanical embedding, and the QM description as electronic embedding [21]. The mechanical embedding can be roughly assigned to the qualitative concept of steric effects, with the electronic embedding corresponding to electronic effects. Mechanical embedding is easier to implement in a computational code than electronic embedding, and because of this, pure mechanical embedding may be considered as zero order QM/MM.

A method with only mechanical embedding would simplify the formula above to the following expression:

$$E_{\text{QM/MM}}(\text{TOT}) = E_{\text{QM}}(\text{QM}) + E_{\text{MM}}(\text{MM}) + E_{\text{MM}}(\text{QM/MM}) \quad (\text{Eq. 3})$$

Inspection of equation 3 shows that it would correspond to the total MM energy by changing the subscript in the first term. This leads to the subtractive scheme characteristic of the IMOMM and ONIOM methods [5, 18-20]:

$$E_{\text{QM/MM}}(\text{TOT}) = E_{\text{QM}}(\text{QM}) + E_{\text{MM}}(\text{TOT}) - E_{\text{MM}}(\text{QM}) \quad (\text{Eq.4})$$

The formula in equation 4 represented an advance in QM/MM methodology because it defined an approach to compute QM/MM energy relying mostly in previously available QM and MM codes, with a minimum interface to be defined between them. In contrast, previous QM/MM implementations had required the explicit introduction of additional terms in the energy expression.

The use of equation 4 defines a path to compute the QM/MM energy and its derivatives for any system, and most of the results discussed in this review are based in this formula. There are however a few caveats about the limitations of this formula and its practical implementation that deserve comment.

The first issue is the handling of electronic embedding. The formula above assumes its neglect. This is a limitation that can be corrected by the introduction of additional terms in the formula, and this has been indeed implemented also within the ONIOM methodology [22]. Electronic embedding is usually introduced as electrostatic embedding, by placing point charges affecting the QM energy on the positions where the MM atoms are placed. The particular value of these point charges has been heavily discussed in the Literature [21]. On the other hand, the neglect of electronic effects may be viewed as an advantage for the analysis of the results, as it provides an opportunity to compute separately electronic and steric contributions [23]. Pure steric contributions can be computed from the QM/MM calculation. The full QM calculation provides both steric and electronic effects, and the latter can be extracted by comparison with the QM/MM calculation.

The second issue is the handling of connections between the QM and MM regions. If there is a covalent bond across the QM/MM partition, the simple deletion of the MM region would leave a dangling bond in the QM calculation. This QM calculation would not be a realistic model of the QM region in the real system. Two main solutions have been given to this problem, special orbitals or link atoms. The introduction of special orbitals in the border atoms in the QM region is the more sophisticated approach, making easier the transfer of electronic effects from the MM region to the QM region [24, 25]. The results presented in this contribution will not use these special orbitals, which are therefore not discussed here in detail. The alternative approach, used in the reported results, is the introduction in the link atoms [18]. These are extra atoms, usually hydrogen atoms, which are introduced to cap the dangling bonds in the QM regions. In order to avoid the introduction of artificial degrees of freedom in the system, the

placement of these link atoms is completely defined by that of the corresponding atoms in the real system in modern QM/MM implementations.

In summary, there is a large variety of QM/MM methods. Most applications in selectivity problems in transition metal chemistry have been however concentrated in the use of the IMOMM and ONIOM implementations with pure mechanical embedding. This is a simple implementation that allows the introduction of steric effects at a cheap price and with moderate effort in the construction of the input files.

### 3 Computing selectivity

Selectivity is an important concept in chemistry [26]. It can be defined as the preferential outcome of a chemical process over a set of other plausible outcomes. A typical example is the case of enantioselectivity, where two enantiomers are in principle possible, but one of them is preferentially obtained. Enantioselectivity is one of the most interesting and most extreme cases of selectivity, as enantiomeric products have by definition the same energy. But the concept of selectivity is more general. Regioselectivity refers to the preference of one direction or chemical bond making or breaking above other alternative directions. Chemoselectivity refers to the reactivity of a chemical functional group in the presence of others.

The concept of selectivity is often associated to organic chemistry, with its clear definition of stereocenters and functional groups. It can be in general applied to either stoichiometric or catalytic processes. The case of catalytic processes is particularly intriguing, as it is possible to introduce a catalyst that induces the selective formation of only one of the possible products. The selectivity inductor is often more difficult to synthesize than the reactants, and because of this, its introduction as a low-concentration catalyst is desirable. Despite the recent progress in organocatalysis, transition metal catalysis plays still a key role in homogeneous catalysis. Because of this, the prediction of selectivity has been an important goal for homogeneous catalysis, and the application of QM/MM methods to the field.

The translation from the typical free energy profiles obtained from pure QM or QM/MM calculations to selectivity ratio or selectivity excess requires a straightforward application of formal kinetics [27] that will be briefly described in this section.

Let us assume that a reactant A can evolve to two different products B, C through first order kinetics ruled by rate constants  $k_b$ ,  $k_c$ :

$$\frac{d[B]}{dt} = k_b[A] \qquad \frac{d[C]}{dt} = k_c[A] \qquad (\text{Eq. 5})$$

Dividing the two formulas in equation 5, integrating and assuming an initial concentration of zero for both B and C, we reach the expression:

$$\frac{[B]}{[C]} = \frac{k_b}{k_c} \qquad (\text{Eq. 6})$$

Equation 6 states a constant ratio between the concentrations of B and C throughout time, which is defined by the ratio between the rate constants. The rate constants can be related through the Eyring equation to the activation free energies:

$$k_b = \frac{k_B T}{h} \exp\left(\frac{-\Delta G_b^\ddagger}{RT}\right) \qquad k_c = \frac{k_B T}{h} \exp\left(\frac{-\Delta G_c^\ddagger}{RT}\right) \qquad (\text{Eq. 7})$$

Inserting the Eyring equations into equation 6 we obtain the relationship between the concentrations and the activation free energies:

$$\frac{[B]}{[C]} = \exp\left(\frac{\Delta G_c^\ddagger - \Delta G_b^\ddagger}{RT}\right) \qquad (\text{Eq. 8})$$

The expression in equation 8 is strictly valid only for cases where first-order kinetics are present, but it is easy to generalize to other cases. The key to its simplicity is that the paths leading to either B or C start from a common intermediate, but this is one of the usual features of chemical problems where selectivity is considered. Extra species leading to a second-order expression could be introduced in the expression in a straightforward way. The usual complication of multistep processes can be properly handled by using the energy span model [28], which simplifies the kinetic expression.

Equation 8 gives the selectivity ratio. When enantioselectivity is considered, the enantiomeric excess (ee) is usually provided. It can be obtained from the enantioselectivity ratio (er) through the following formula:

$$ee = \frac{er-1}{er+1} \qquad (\text{Eq. 9})$$

The exponential dependence of selectivity ratio (or selectivity excess) with respect to the free energy differences underlines one of the key features of selectivity in chemistry: its strong sensitivity to small energy changes. A free energy difference in the barriers of 1 kcal mol<sup>-1</sup> translates roughly into a selectivity ratio of 5:1 (84:16), a difference of 2 kcal mol<sup>-1</sup> leads to a ratio of 29:1 (97:3), a difference of 3 kcal mol<sup>-1</sup> leads to a ratio of 153:1 (99:1). A difference of 5 kcal mol<sup>-1</sup> would

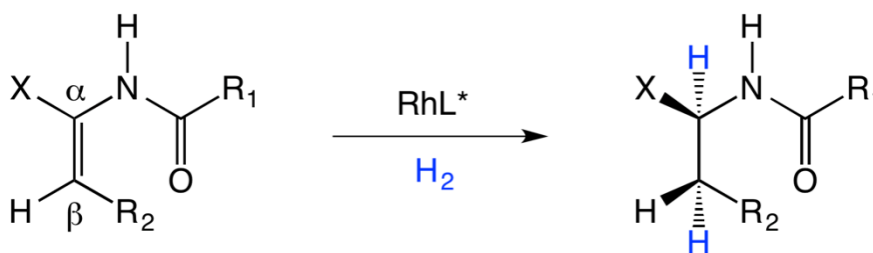
already lead to the non-observation of the minor product, which should only be present in the range of the 0.02%.

The strong sensitivity of computed selectivity to differences in free energy barriers poses a challenge to computational chemistry. Practical reactions are often considered to be efficient when selectivity is above 90%, and this boils down to differences between transition states inferior to 2 kcal mol<sup>-1</sup>. This level of absolute accuracy is difficult to achieve for most modern computational techniques on realistic chemical models. Luckily there is a phenomenon of cancellation of errors which very often favors selectivity calculations. We are not interested in absolute free energy barriers, but in the difference in free energy barriers leading to the different products. And it so happens that most of the absolute error in the computed barriers can be directly transferred from one path to another. This can be better understood with an example on rhodium-catalyzed enantioselective hydrogenation that will be discussed in detail in the following sections. The absolute barrier to the key transition state depends on the interactions between substrate and metal that are difficult to reproduce exactly. But the nature of these interactions is the same in the transition states leading to the *R* and *S* products. The discrimination between the two transition states comes from the steric interactions between the substituents, and these purely steric interactions are more likely to be well reproduced if the force field is sufficiently accurate.

A closely related issue is that the difference of activation free energies has been often replaced in practice by the difference in activation enthalpies or activation potential energies. The validity of this approach is again related to cancelation of errors. The correction necessary to convert the potential energy barrier to free energy barrier for one product is often nearly the same to the correction required to convert the potential energy barrier leading to the alternative product.

## 4 Rhodium-catalyzed Hydrogenation of Enamides

The first homogeneous hydrogenation of alkenes, using  $\text{RuCl}_3$ , was reported in the early 1960s by Halpern, Harrod and James [29]. Later on, a similar reaction was published by Wilkinson using his famous  $[\text{RhCl}(\text{PPh}_3)_3]$  catalyst [30]. These processes, combined to the contemporary development of homogeneous asymmetric reactions by Noyori and Nozaki [31], gave birth to the asymmetric hydrogenation of olefins firstly reported by Knowles and Horner in 1968 [32]. This reaction employed monodentate phosphane ligand ( $\text{L}^*$ ) as sources of chirality and provided poor enantiomeric excesses (Figure 1).

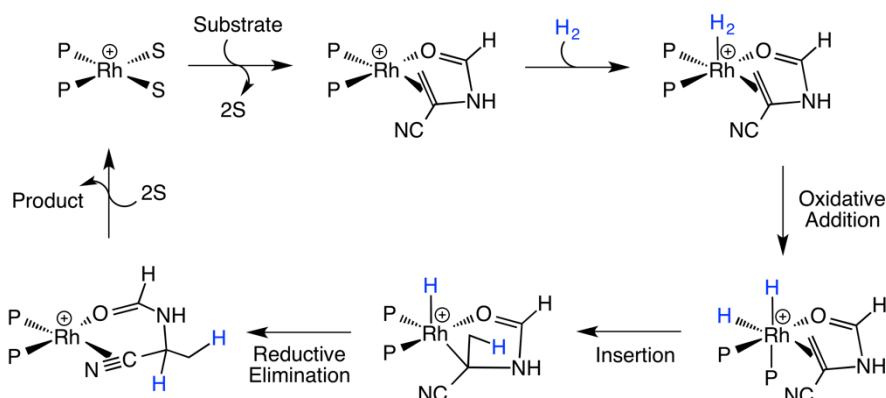


**Figure 1.** First reported asymmetric rhodium-catalyzed hydrogenation of olefins.

The usage of more complex ligands *e.g.* DIOP [33] led to the improvement of these reactions and allowed the obtention of higher enantiomeric excesses. One important application of this methodology was the industrial synthesis of L-DOPA, developed in the group of Knowles, employing the bidentate ligand DIPAMP.

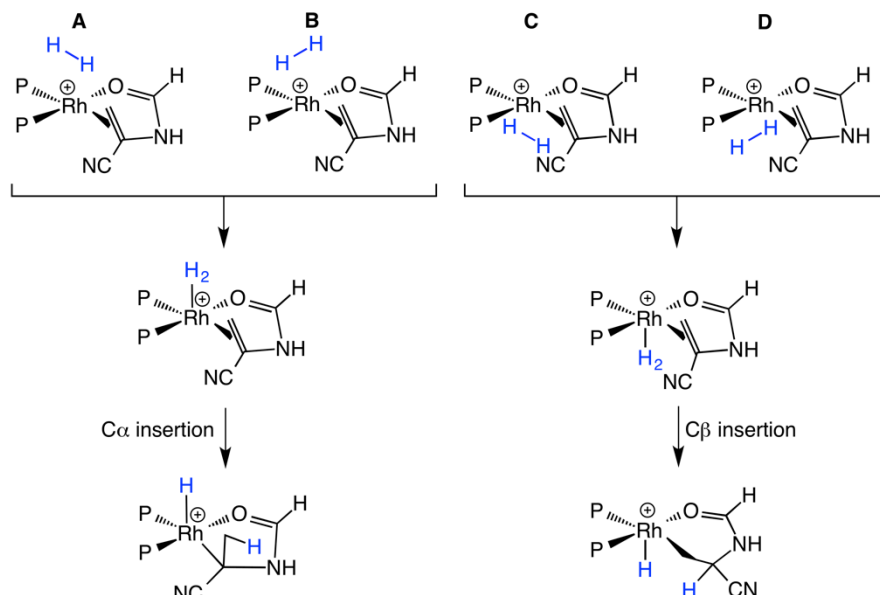
The mechanism of this reaction, fully supported with experimental data, is shown in Figure 2 [34]; this pathway is known as the alkene or the unsaturated mechanism. The catalytic cycle starts by the replacement of the two solvent molecules (S) on the initial rhodium(I) complex by the enamide substrate. Then the hydrogen comes in and the oxidative addition takes place, delivering the Rh(III) dihydride intermediate. After that, one of the hydrides is inserted either in the  $\alpha$  or  $\beta$  position of the olefin (Figure 2 shows the insertion on the latter). Finally the hydrogenated product is obtained by a reductive elimination process. An alternative pathway, called the dihydride mechanism, is also plausible and has been found to be operative in reactions involving highly electron-rich P-ligands [35]. In this case the oxidative addition of the hydrogen happens before the solvent is substituted by the enamide substrate.



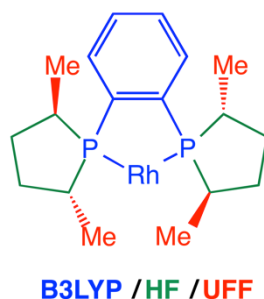


**Figure 2.** Proposed mechanism for the rhodium-catalyzed hydrogenation of enamides.

The first computational studies dealing with the hydrogenation of olefins are among the first reports of computed full catalytic cycles. They date from the late 1980s and were published by Morokuma and coworkers. The full mechanism of the rhodium-catalyzed hydrogenation of ethylene with the Wilkinson catalyst was reported for a model system [36]. A decade later Landis and Feldgus reported a computational study on the hydrogenation of enamides [37]. In order to determine the turnover-limiting step they employed a small model system related to the one shown in Figure 1:  $L^* = \text{PH}_3$ ,  $R_1 = R_2 = \text{H}$  and  $X = \text{CN}$ . Four different pathways (A–D, Figure 3) were found to describe the approach of H<sub>2</sub> to the catalyst/enamide complex. In paths A and C the H<sub>2</sub> moiety comes in parallel to the Rh–olefin bond while in B and D it is parallel to the Rh–O<sub>carbonyl</sub> bond. In all cases the insertion process was computed for the both possible reaction points (C $\alpha$  and C $\beta$ ). For pathways A and B the preferred insertion position is C $\alpha$  whereas C $\beta$  is favored in pathways C and D. The lowest energy pathway was found to be A with the highest step corresponding to the insertion process. Although the model system employed allowed a quite complete study of the whole processes it could not be used to study the enantioselectivity of this hydrogenation reaction. This was fixed later by the same authors in subsequent publications where they included a full chiral diphosphane as DuPHOS in QM/MM ONIOM studies [38, 39]. The system was described with three different layers, as shown in Figure 4 and the enamide substrate was  $\alpha$ -formamidoacrylonitrile, the same employed previously. In this case, and depending on the orientation adopted by the substrate when coordinating the metal center, the number of studied pathways is doubled: A, B, C and D for the pro-*S* and pro-*R* manifolds. Pathway A was found to be the most favored one, with the H<sub>2</sub> oxidative addition step bearing the highest energy barrier.



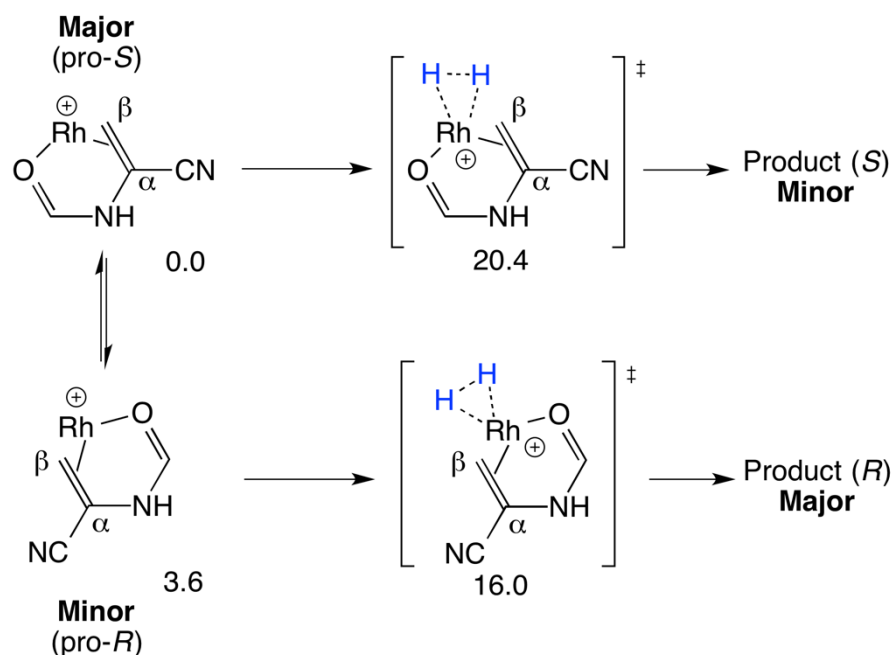
**Figure 3.** Different approaches of H<sub>2</sub> (pathways A-D) to the substrate-catalyst complex in the Rh-catalyzed hydrogenation of enamides, and mechanism for the first insertion of C $\alpha$  and C $\beta$  into the Rh-H bond.



**Figure 4.** Partition scheme describing the [Rh((*R,R*)-Me-DuPHOS)]<sup>+</sup> complex.

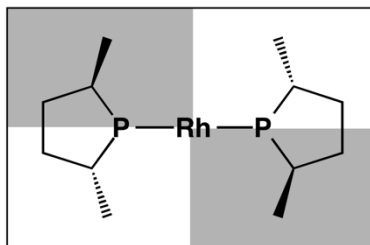
These calculations were able to reproduce the anti “lock-and-key” mechanism observed experimentally; the most stable pro-*S* intermediate displays the highest reaction barrier while the pro-*R* complex, higher in energy, has the lowest barrier to produce the final *R* product, which is the major outcome of the reaction (Figure 5). The energy difference between the computed transition states (4.4 kcal mol<sup>-1</sup>)

corresponds to a theoretical enantiomeric excess of 99.9%, well within the range obtained in experiments [40].



**Figure 5.** Pro-*R* and pro-*S* Rh-complexes and their rate-determining oxidative addition transition states in the asymmetric hydrogenation of  $\alpha$ -formamidoacrylonitrile, relative energies in kcal mol<sup>-1</sup>.

These results could be explained by using a quadrant map of the active catalyst, as originally proposed by Knowles and coworkers (Figure 6) [41]. The pro-*S* enamide complex is more stable because C $\alpha$  can be allocated in the Rh-phosphane plane while in the pro-*R* analog this is not possible, and thus C $\beta$  occupies the position in the metal-ligand plane. In contrast, when H<sub>2</sub> enters the coordination sphere of the metal, in the pro-*S* isomer the CN group is forced to move into a hindered space, which does not happen with the pro-*R* complex. Therefore, the hydrogen activation is more favorable in the latter case and the *R* product is formed much faster. Morokuma and coworkers found similar results using the IMOMM methodology for a related system[42].

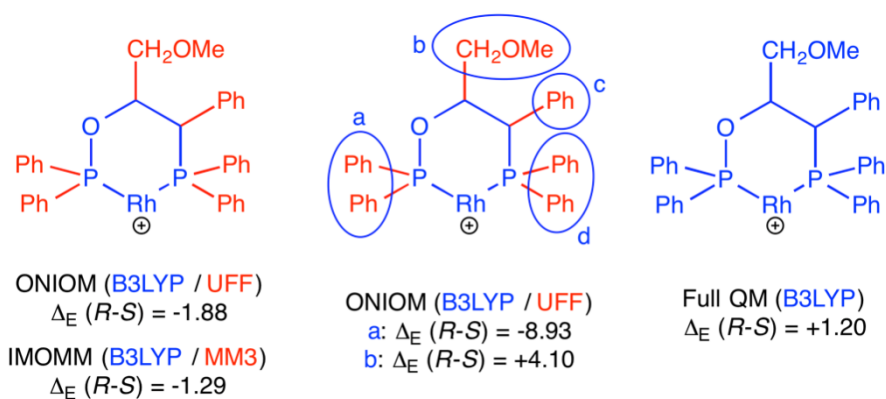


**Figure 6.** Quadrant map of the Rh-diphosphane catalyst.

Subsequent publications pointed out to the fact that A is not always the most favorable pathway and C could arise as the preferred one depending on the electronic features of the substrate. This was first reported separately by Landis and Feldgus using QM/MM calculations [43] and by Wiest and coworkers using full QM calculations [44]. In both cases it was demonstrated that pathway A is the preferred whenever electron-withdrawing groups are present in  $C\alpha$ , favoring the insertion process on this position. In contrast, electron-donating groups can reverse this tendency and end up favoring the insertion of  $C\beta$ , thus following pathway C. This means that the *R/S* selectivity of the final product can be drastically changed depending on the properties of the substrate.

In all these reports, where the ligands have  $C_2$  or  $C_s$  symmetry, the results obtained with QM/MM or full QM calculations provide very similar results. However, Maseras and coworkers found a modification of this trend when using the bidentate phosphane-phosphinite ligand depicted in Figure 7 in the rhodium-catalyzed hydrogenation of methyl-(*N*)-acetylaminoacrylate [45], with which they obtained different ee's depending on the QM/MM partition employed. When the a-d groups were included in the UFF region the energy difference between the lowest transition states delivering the *R* and *S* products was found to be  $-1.88 \text{ kcal mol}^{-1}$ , contradicting the experimental results where a 40% ee in *S* was obtained. The problem remained when the methodology was changed to IMOMM (B3LYP/MM3) calculations. The only way to obtain the right enantiomeric excess was to carry out the calculations with a full QM method, using the B3LYP functional. Single point calculations using different QM/MM partitions were employed to determine the impact of the electronic effect of the ligand substituents on the enantioselectivity. When the phenyl rings of the a region are included in the QM part, the energy difference in favor of the *R* product is strongly enhanced ( $\Delta_E(R-S) = -8.93 \text{ kcal mol}^{-1}$ ). If the d phenyl groups are placed within the QM region the other enantiomer is obtained ( $\Delta_E(R-S) = +4.10 \text{ kcal mol}^{-1}$ ). These results indicate that increasing the basicity on either P-donor atom enhances the formation of one of the two different enantiomers, with the phosphinite favoring the *R* product and the phosphane delivering the *S* product. These observations are in line with what is observed experimentally and theoretically with an (*S*)-BINOL-derived phosphite ligand, where a 99% of ee in *R* was obtained [46]. Full QM calculations on this system indicate that the enantioselectivity depends both on

electronic and steric effects. Four pathways are plausible in this reaction but two of them, the ones where the olefin is in trans to the phosphite, are avoided by electronic effects. One of the two remaining pathways is sterically blocked by the BINOL ligand, leaving just one pathway for the reaction to occur. These results explain the high selectivity found in this reaction.



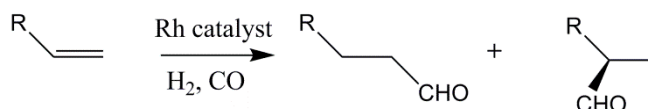
**Figure 7.** QM/MM partitions and methods used for modeling the rhodium active species and the corresponding difference in energy (in kcal mol<sup>-1</sup>) between the lowest Pro-*S* and Pro-*R* transition states in the hydrogenation of methyl-(*N*)-acetylaminoacrylate.

QM/MM and full QM methodologies have demonstrated to be accurate enough to study the enantioselectivities of particular asymmetric hydrogenation reactions, and have helped to unravel the mechanistic complexity of these processes. These methods are, however, still quite expensive to carry out a complete screening of a real catalytic system where many combinations leading to different products may be possible. Other approaches such as QSAR-like approaches based on descriptors [47] and pure MM calculations with a specifically tailored Q2MM force fields have been also productive alternatives [48, 49].

## 5 Rhodium-catalyzed Hydroformylation

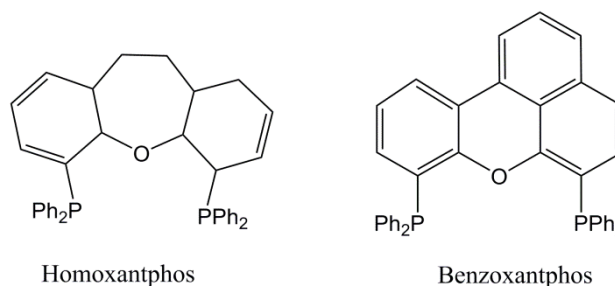
The hydroformylation reaction is the transition metal mediated addition of carbon monoxide and dihydrogen to the double bond of an alkene (Figure 8). It is one of the most important reactions catalyzed by homogeneous transition metal complexes in the industrial production of bulk chemicals [50, 51]. Figure 8 highlights the regioselectivity issue in hydroformylation: two different products are possible when the alkene is not symmetrically substituted. A common case is that where the alkene has a single substituent and two products are possible, one

linear and one branched. The general mechanism of the rhodium-catalyzed hydroformylation and the particular issue of regioselectivity have been examined in a number of theoretical studies with different computational techniques, and reviewed quite recently [52]. Significant progress has been made in the mechanistic understanding, but the topic has not been fully solved for all cases, and it continues to be a matter of study for pure DFT methods [53, 54]. Some significant contributions have been made from the DFT/MM point of view and are reviewed below.



**Figure 8.** The rhodium-catalyzed hydroformylation of alkenes.

Carbó *et al* used a QM/MM IMOMM method [18] to study the origin of regioselectivity in the Rh-diphosphine catalyzed hydroformylation [55]. The study was centered on the rationalization of the experimentally observed dependence of selectivity with the bite angle of a variety of xantphos ligands. It had been observed that ligands with the larger bite angles produced a higher amount of linear product. Because of this, the key transition state in the process, corresponding to olefin insertion into the rhodium-hydride bond, was computed for two different diphosphine ligands, benzoxantphos and homoxantphos (see Figure 9) with extreme cases of natural bite angle, and two different alkenes, propene and styrene. Only the results with propene will be discussed here.



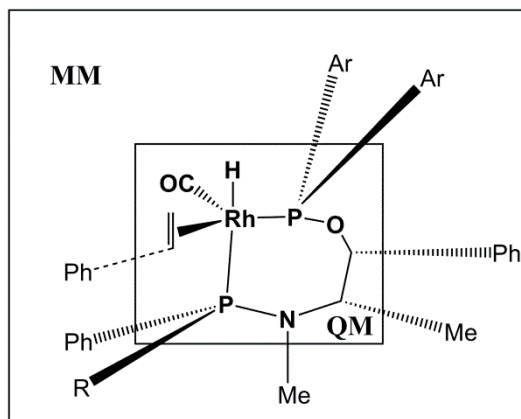
**Figure 9.** The homoxantphos and benzoxantphos diphosphine ligands

A total of eight transition states were computed for each system, four leading to the linear product and four leading to the branched product. The results reproduced satisfactorily the experimental trends. Benzoxantphos produced bite angles around 110 degrees, while the bite angles for homoxantphos were around 100 degrees. The pattern in the relative energy distribution of transition states was the same for both ligands but the relative energy barriers were different. The computed percentages of linear product for propene were 83% for benzoxantphos

vs 73% for homoxantphos systems. This must be compared with experimental values for 1-octene of 98.1% and 89.5%, respectively. Agreement was not perfect but the trend was correctly reproduced.

Calculations allowed moreover the performance of a computational experiment where each phenyl substituent was replaced by hydrogen and maintained the backbone of diphosphine ligands (PH<sub>2</sub> model). By removing the phenyl substituents, the non-bonding effects of the phenyls on regioselectivity were put aside. The PH<sub>2</sub> model produced significantly smaller selectivities than the full model, with the values going down from 83%/73% to 74%/63%. There is an intrinsic selectivity associated to the backbone, but its effect on the overall regioselectivity is exercised through the steric interactions of the phenyls. It is not an orbital effect associated to changes in the electronic properties of the rhodium center, but a steric effect that plays through the substituents at the phosphane. Wider bite angles increase the steric interaction of the diphosphine substituents with branched species, which become more destabilized. One could therefore expect that bulkier groups than phenyl groups would lead to higher linear:branched regioselectivity ratios.

Related systems were reexamined by Zuidema *et al.* [56] a few years later to analyze the rate-determining step in the hydroformylation of 1-octene, catalysed by the rhodium-xantphos catalyst system using a combination of experimentally determined kinetic isotope effects and computational approaches. The focus of this work was to clarify whether alkene coordination or hydride migration is the rate determining step. Both ONIOM(B3LYP:UFF) and B3LYP calculations were carried out on the key catalytic steps, using the real ligand systems. The calculations quantitatively reproduced the energy barrier for CO dissociation. The overall barrier for hydride migration from the resting state was found to be 3.8 kcal mol<sup>-1</sup> higher than the barrier for CO dissociation. This fit well with the experimentally determined trend, confirming the assumptions of the previous work on regioselectivity. The combination of kinetic isotope effects and theoretical studies suggested that the overall barrier for hydride migration, starting from the resting state of the catalyst, determines the activity in the rhodium-xantphos catalysed hydroformylation of 1-octene.



**Figure 10.** QM/MM partition in the AMPP-Rh-styrene complex

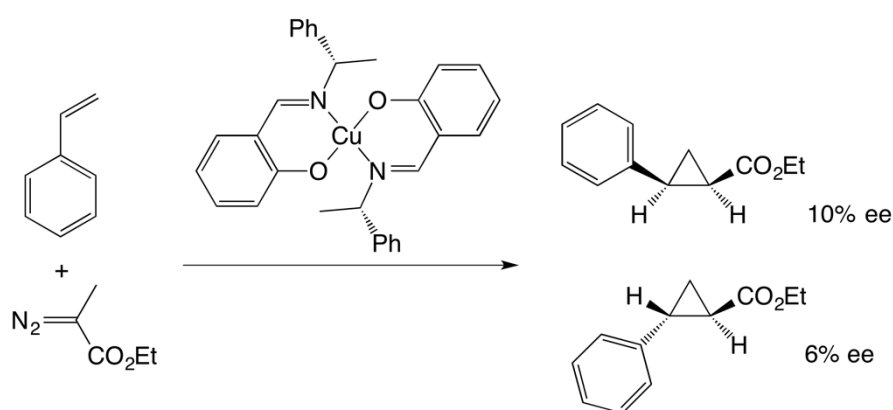
Apart from the issues concerning regioselectivity, there is also the potential for enantioselectivity, as a stereogenic center can be generated in branched products. Carbó *et al.* [57] applied the IMOMM (BP:Sybil) method to analyze the origin of stereinduction in the hydroformylation of styrene by rhodium complexes bearing chiral aminophosphane phosphinite (AMPP, see Figure 10) ligands. The roles of the stereogenic center at the aminophosphane phosphorus atom (NP\*) and of the chirality of the backbone were considered in three experimentally reported cases: 1) P-stereogenic yielding high ee, 2) P-nonstereogenic yielding low ee, and 3) P-stereogenic yielding low ee. Experimentally observed trends for the three studied AMPP ligands were fairly well reproduced and the calculations revealed that the different non-bonding weak-type interactions of styrene with the substituents of the NP\* stereogenic center in an axial position was responsible for stereodifferentiation.

Stereoselectivity issues were also analyzed with the help of IMOMM(BP86:Sybil) calculations by Aguado-Ullate *et al.* [58] in the case of the reaction of styrene where the rhodium catalyst is carrying an unsymmetric bidentate phosphane phosphite ligand such as BINAPHOS. The behavior of the  $[\text{Rh}\{(R,S)\text{-BINAPHOS}\}(\text{CO})_2\text{H}]$  catalyst was studied. The placement of the phosphane moiety in the apical site and the phosphite moiety on the equatorial site was shown to be critical for high enantioselectivity. The axial chirality of the phosphite discriminated one of the competitive equatorial-apical paths, whereas the chirality of the backbone discriminated one of the two enantiomers. QM/MM calculations were also used in the definition of QSAR descriptors for hydroformylation by the same group of research [59].



## 6 Copper-catalyzed Cyclopropanation

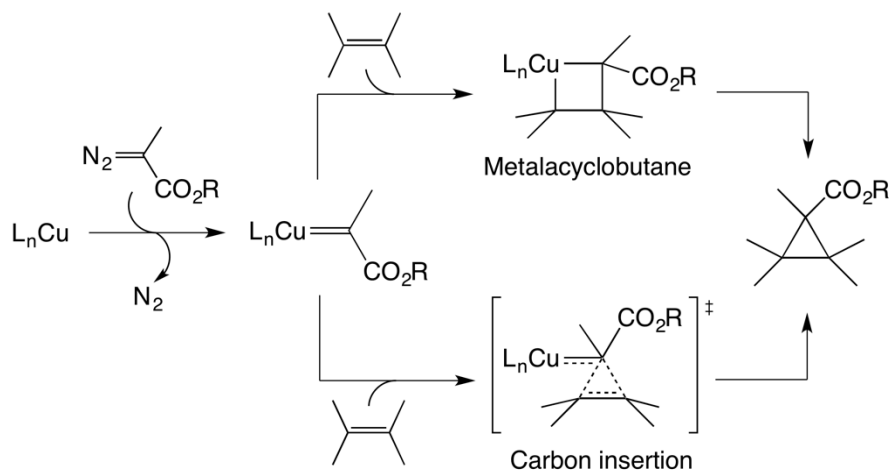
The first asymmetric copper-catalyzed cyclopropanation using a homogeneous catalyst was reported by Noyori in 1966 [31]. This reaction, which allowed the cycloaddition on styrene, was carried out with a chiral Schiff base copper complex and produced poor enantiomeric excess (Figure 11). Further refinement of the chiral ligands produced later much better catalysts, such as the bis-oxazoline (Box) derivatives, able to provide enantioselectivities up to 99%.



**Figure 11.** First reported asymmetric copper-catalyzed cyclopropanation.

Not much is experimentally known about the operating mechanism of this cyclopropanation reaction; the available data indicates that the metal remains as Cu(I) throughout the whole process and that the reaction proceeds through a metal-carbene intermediate [60]. Based on the reactivity observed for analogous catalysts *e.g.* Ni(0) complexes or Zn(II) carbenoid systems, two different mechanistic proposals were suggested. The first one involves the stepwise formation of metalacyclobutane intermediates while the alternative mechanism consists of the concerted direct carbon insertion (Figure 12). A computational study, using DFT calculations, was employed to determine the preferred catalytic cycle for the cyclopropanation reaction of methyl diazoacetate with ethylene and a Cu(I)(Box) complex [61]. At first, a small model system, where the Box ligand was *N,N'*-dimethylmalonaldiimine was used to evaluate all the reaction steps. The results stated that the rate-limiting step was the nitrogen extrusion delivering the copper-carbene intermediate and the computed barriers indicate that the carbene insertion mechanism was preferred over the formation of the metalcycle, with energy requirements of 9.8 and 12.8 kcal mol<sup>-1</sup>, respectively. Nevertheless, the stereoselectivity control comes from the reaction between the alkene and the carbene moieties. In order to check the performance of the computed mechanism

with the experimental observations these results were extended then to a more complex chiral system, where the simple starting diimine was replaced by 2,2'-methylenebis[(4*S*)-methyl-2-oxazoline]. Two different diastereomeric transition states were computed, corresponding to the approach of the alkene to the Si and Re faces of the plane formed by the copper-carbene complex, and the former was favored by more than 1 kcal mol<sup>-1</sup> when the solvent effects were taken into account. This behavior can be related to the steric hindrance that appears between the methyl substituent on the ligand and the ester group when the olefin approaches the copper species through the Si face. The computed energy difference between both transition states showed a good agreement with the enantiomeric excess obtained in the cyclopropanation reaction between ethyl diazoacetate and styrene with a similar Box ligand [62].



**Figure 12.** Proposed mechanisms for the copper-catalyzed cyclopropanation of alkenes.

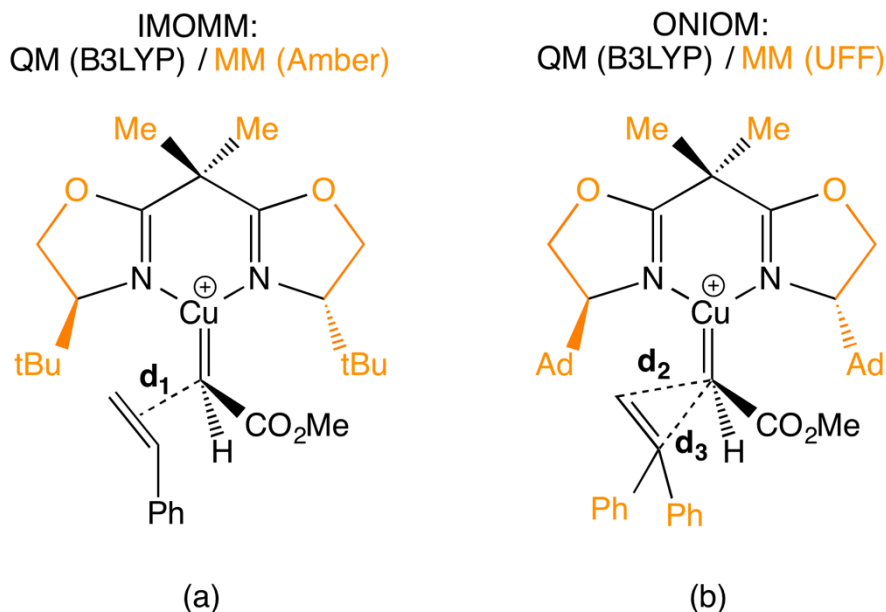
The *cis/trans* selectivity of this reaction was studied replacing the ethylene by 1-propene. Although eight different possible transition states arise from this modification only in four of them the incoming methyl group of the propene lies far from the Box ligand. Thus, four transition states were computed: Re-*cis*, Si-*cis*, Re-*trans* and Si-*trans*, and found to have energy barriers of 12.5, 13.4, 11.3 and 12.2 kcal mol<sup>-1</sup>, respectively, in a nice qualitative agreement with the experiments [62]. In this case the *cis/trans* selectivity seems to be governed by the steric interaction between the substituent on the olefin and the ester group of the carbene.

The enantioselectivity of these cycloaddition reactions is strongly affected by the anion present on the initial Cu(I)(Box) catalyst *e.g.* replacing a triflate by a chloride in the cyclopropanation between ethyldiazoacetate and styrene produces

low yields and enantioselectivities of 8 and 3% for the cis and trans products, respectively; while the triflate systems delivers 92 and 94% [63]. This counterion effect was also studied computationally by the group of García and coworkers employing the same methodology as above [64]. The counterion replacement in the small model system produces different geometries for the intermediates and transition states and thus the relative energy differences along the reaction pathway change substantially; for instance, the barrier of the insertion stage raises more than 3 kcal mol<sup>-1</sup>. In addition, the formation of chloride-bridged dimers seems quite possible, indicating that the whole process should be slower and suffering a remarkable decrease in the catalytic activity. The enantioselectivity reduction observed for the chlorinated system was attributed to the lower steric repulsion between the methyl group of the ligand and the incoming ester substrate in the concerted transition state. The results obtained were clearly in line with the decrease in enantioselectivity observed when chloride is employed as a counterion.

The replacement of ethylene by styrene as a more nucleophilic olefin substrate in the cyclopropanation reactions produced a major issue: the concerted transition state could not be located by classical potential energy surface explorations as reported for the first time by Norrby and coworkers [65]. In this case they studied the reaction between substituted styrenes and ethyldiazoacetate with QM/MM IMOMM using the catalyst shown in Figure 13a. The concerted transition state was not possible to find exploring the potential energy surface because the approximation of the double bond to the carbene produced a monotonic downhill energy profile. Thus, the Gibbs free energy surface was computed using a linear transit scan by optimizing a set of structures with fixed distances between the centroid of the alkene and the carbene carbon ( $d_1$  in Figure 13a). In this way the free energy reached maximum value at  $d_1$  around 2.5 Å; this “transition state” is strongly asynchronous since the carbene-CHPh distance is around 2.75 Å while the carbene-CH<sub>2</sub> distance is much shorter and around 1.95 Å. Even though this might not be the best methodology, the free energies obtained were in a quite good qualitative agreement with the experimental observations. The lowest free energy barrier was found for the alkene approaching the carbene through the Re face, with the phenyl group in a trans arrangement to the ester group.

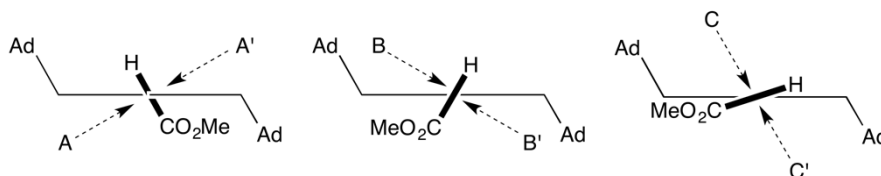
Later on, Drudis-Solé *et al.* studied the cyclopropanation of Ph<sub>2</sub>C=CH<sub>2</sub> using the catalyst shown in Figure 13b with QM/MM ONIOM calculations [66]. In this case a different approximation was employed to locate the concerted transition state; the Gibbs free energy surface was explored with a small model system by using two reaction coordinates ( $x, y$ ) directly related to the two forming C–C bonds ( $d_2$  and  $d_3$  in Figure 13b). The  $x$  coordinate indicates the advance of the reaction while  $y$  corresponds to the synchronicity of the transition state. Using this methodology the concerted transition state was found to lie 5.7 kcal mol<sup>-1</sup> above the reactants with coordinates  $x = 3.6$  Å and  $y = 0.0$ , indicating a completely synchronous process. The transformation of these coordinates back to C–C distances produces values of  $d_2 = d_3 = 3.6$  Å.



**Figure 13.** QM/MM strategies to find the concerted transition state of cyclopropanation in the Gibbs energy surface of styrene (a) and  $\text{Ph}_2\text{C}=\text{CH}_2$  (b) with copper complexes.

Subsequently, they employed the same methodology for building the free energy surface of the full catalytic system using only  $x$  as the reaction coordinate and studying just synchronic pathways ( $y = 0$ ). In this way 24 different reaction pathways were investigated considering three positions of the carbene (A, B and C, Figure 14) which could be attacked through the two faces and with different dihedral angles between the carbene and the oxy and methoxy groups of the ester. In all the cases the R and S products can be obtained. Thus the free energy profile was computed for all the pathways allowing the calculation of the enantiomeric excess for all cases. The use of a QM/MM approach was critical to enhance such a computationally demanding approach.

The most favorable pathway affords a 90% ee for the *S* product, in nice agreement with the 98% ee in *S* found experimentally. The enantioselectivity of this system seems to arise from the relative orientation of the carbene with respect to the bis-oxazoline ligand plane, the steric repulsions that appear between the adamantyl substituent and the ester group and between the phenyl groups of the alkene and both the adamantyl groups and the Box ligand.



**Figure 14.** Possible relative positions of the carbene with respect to the Cu-Box plane in the cyclopropanation of  $\text{Ph}_2\text{C}=\text{CH}_2$ .

García and coworkers employed a different strategy to study the enantioselectivity of cyclopropanation reactions where an enthalpic energy barrier cannot be located. This consists of forcing the system to have one of such barriers by making it less reactive *i.e.* decreasing the nucleophilic nature of the olefin [67]. They studied the cyclopropanation of styrene with the catalyst shown in Figure 13a, with the same level of theory, but in this case they included the phenyl ring of the substrate in the MM part. This replacement increases the barrier of the reaction and thus the concerted transition state could be located without the need of computing the Gibbs free energy surface. The computed enantioselectivity was similar as the one reported above.

## 7 Conclusions and Perspectives

QM/MM methods are already a widely established tool for the calculation of free energy profiles for complex systems in homogeneous catalysis. They are especially suited for mononuclear systems where the explicit introduction of bulky ligands in the calculation is necessary to reproduce the experimental behavior. These are the type of systems which are more frequent in problems involving regioselectivity and enantioselectivity.

The proper characterization of the mechanistic origin of selectivity is a highly desirable target for computational chemistry. Selectivity is the key in many processes of practical interest, and the availability of a reliable predicting tool from computational chemistry is in high demand. QM/MM methods are a very valid tool for the study of reaction mechanisms in selective homogeneous catalysis.

QM/MM methods and computational selective homogeneous catalysis have progressed hand in hand the last two decades, and there is no reason to expect a slowdown in the progress of both areas. The increase in computer power will push further the size of the systems that can be studied with a pure QM description, but there will be always desire for larger and larger systems which require an MM

part. Further developments in QM/MM methodology can be moreover expected from the expansion in the number of available force fields to mainstream codes. On the other hand, the introduction of automated models for the study of large systems will facilitate work on the large homogeneous catalysts often involved in selective processes. The field covered in this review appears in good health and further progress must be expected in coming years.

## References

- [1] Maseras F, Lledós A (2002) (eds) Computational modeling of homogeneous catalysis. Kluwer, Dordrecht.
- [2] Thiel W (2014) *Angew Chem Int Ed* 53:8605.
- [3] Bonney KJ, Schoenebeck F (2014) *Chem Soc Rev* 43:6609.
- [4] Senn HM, Thiel W (2009) *Angew Chem Int Ed* 48:1198.
- [5] Chung LW, Sameera WMC, Ramozzi R, Page AJ, Hatanaka M, Petrova GP, Harris TV, Li X, Ke ZF, Liu FY, Li HB, Ding LN, Morokuma K (2015) *Chem Rev* 115:5678.
- [6] Maseras F (1999) *Top Organomet Chem* 4:165
- [7] Maseras F (2000) *Chem Commun* 1821.
- [8] Maseras F (2001) In: Cundari TR (ed) *Computational organometallic chemistry*. Marcel Dekker, New York
- [9] Ujaque G, Maseras F (2004) *Struct Bond* 112:117.
- [10] Bo C, Maseras F (2008) *Dalton Trans* 2911.
- [11] Ananikov VP, Musaev DG, Morokuma K (2010) *J Mol Cat A* 324:104.
- [12] Sameera WMC, Maseras F (2012) *WIREs Comp Mol Sci* 2:375.
- [13] Balcells, D.; Maseras, F. (2007) *New J Chem* 31:333.
- [14] Brown, J. M.; Deeth, R. J. (2009) *Angew Chem Int Ed* 48:4476.
- [15] Krenske EH, Houk KN (2013) *Acc Chem Res* 46: 979.
- [16] Warshel A, Levitt, M (1976) *J Mol Biol* 103:227.
- [17] Field MJ, Bash PA, Karplus M (1990) *J Comput Chem* 11:700.
- [18] Maseras F, Morokuma K (1995) *J Comput Chem* 16:1170.
- [19] Svensson M, Humbel S, Froese RDJ, Matsubara T, Sieber S, Morokuma K (1996) *J Phys Chem* 100:19357.
- [20] Dapprich S, Komaromi I, Byun KS, Morokuma K, Frisch MJ (1999) *J Mol Struct (THEOCHEM)* 461:1.
- [21] Lin H, Truhlar DG (2007) *Theor Chem Acc* 117:185.
- [22] Vreven T, Morokuma K (2003) *Theor Chem Acc* 109:125.
- [23] Barea G, Lledós A, Maseras F, Jean Y (1998) *Inorg Chem* 37:3321.
- [24] Assfeld X, Rivail JL (1996) *Chem Phys Lett* 263:100.
- [25] Gao JL, Amara P, Alhambra C, Field MJ (1998) *J Phys Chem A* 102:4714.
- [26] Walsh P J, Kozlowski MC (2009) *Fundamentals of Asymmetric Catalysis*, University Science Books, Sausalito.
- [27] Laidler KJ (1987) *Chemical Kinetics*. Prentice Hall, New York.
- [28] Kozuch S, Shaik S (2011) *Acc Chem Res* 44:101.
- [29] Halpern J, Harrod JF, James BR (1961) *J Am Chem Soc* 83:753.

- [30] Young JF, Osborn JA, Jardine FH, Wilkinson G (1965) *Chem Commun* 131.
- [31] Nozaki H, Moriuti S, Takaya H, Noyori R (1966) *Tetrahedron Lett* 5239.
- [32] Knowles WS, Sabacky MJ (1968) *Chem Commun* 1445.
- [33] Dang TP, Kagan HBJ (1971) *Chem Soc Chem Commun* 481.
- [34] Landis CR, Halpern J (1987) *J Am Chem Soc* 109:1746.
- [35] Gridnev ID, Imamoto T (2004) *Acc Chem Res* 37:633.
- [36] Daniel C, Koga N, Han J, Fu XY, Morokuma K (1988) *J Am Chem Soc* 110:3773.
- [37] Landis CR, Hilfenhaus P, Feldgus S (1999) *J Am Chem Soc* 121:8741.
- [38] Feldgus S, Landis CR (2000) *J Am Chem Soc* 122:12714.
- [39] Landis CR, Feldgus S (2000) *Angew Chem Int Ed* 39:2863.
- [40] Burk MJ, Feaster JE, Nugent WA, Harlow RL (1993) *J Am Chem Soc* 115:10125.
- [41] Knowles WS (1983) *Acc Chem Res* 16:106.
- [42] Mori S, Vreven T, Morokuma K (2006) *Chem Asian J* 1:391.
- [43] Feldgus S, Landis CR (2001) *Organometallics* 22:2374.
- [44] Donoghue PJ, Helquist P, Wiest O (2007) *J Org Chem* 72:839.
- [45] Donald SMA, Vidal-Ferran A, Maseras F (2009) *Can J Chem* 87:1273.
- [46] Fernández-Pérez H, Donald SMA, Munslow IJ, Benet-Buchholz J, Maseras F, Vidal-Ferran A (2010) *Chem Eur J* 16:6495.
- [47] Shimizu H, Ishizaki T, Fujiwara T, Saito T (2004) *Tetrahedron Asymmetry* 15:2169.
- [48] Donoghue PJ, Helquist P, Norrby PO, Wiest O (2009) *J Am Chem Soc* 131:410.
- [49] Donoghue PJ, Helquist P, Norrby PO, Wiest O (2008) *J Chem Theory Comput* 4:1313.
- [50] van Leeuwen PWNM, Claver C (2002) (eds) *Rhodium Catalyzed Hydroformylation*, Springer, Heidelberg.
- [51] Franke R, Selent D, Boerner A (2012) *Chem Rev* 112:5675.
- [52] Kégl T (2015) *RSC Advances* 5:4304.
- [53] Schmidt S, Deglmann P, Hofmann P (2014) *ACS Catalysis* 4:3593.
- [54] Jacobs I, de Bruin B, Reek JNH (2015) *ChemCatChem* 7:1708.
- [55] Carbó JJ, Maseras F, Bo C. van Leeuwen PWNM (2001) *J Am Chem Soc* 123:7630.
- [56] Zuidema E, Escorihuela L, Eichelsheim T, Carbó JJ, Bo C, Kamer PJK, van Leeuwen PWNM (2008) *Chem Eur J* 14:1843.
- [57] Carbó JJ, Lledós A, Vogt D, Bo C (2006) *Chem Eur J* 12:1457.
- [58] Aguado-Ullate S, Saureu S, Guasch L, Carbó JJ (2012) *Chem Eur J* 18:995.
- [59] Aguado-Ullate S, Guasch L, Urbano-Cuadrado M, Bo C, Carbó JJ (2012) *Catal Sci Tech* 2:1694.
- [60] Salomon RG, Kochi JK (1973) *J Am Chem Soc* 95:3300.
- [61] Fraile JM, García JI, Martínez-Merino V, Mayoral JA, Salvatella L (2001) *J Am Chem Soc* 123:7616.
- [62] Evans DA, Woerpel KA, Hinman MM, Faul MM (1991) *J Am Chem Soc* 113:726.
- [63] Fraile JM, García JI, Mayoral JA, Tarnai T (1999) *J Mol Catal A* 144:85.
- [64] Fraile JM, García JI, Gil MJ, Martínez-Merino V, Mayoral JA, Salvatella L (2004) *Chem Eur J* 10:758.
- [65] Rasmussen T, Jensen JF, Østergaard N, Tanner D, Ziegler T, Norrby PO (2002) *Chem Eur J* 8:177.
- [66] Drudis-Solé G, Maseras F, Lledós A, Vallribera A, Moreno-Mañas M (2008) *Eur J Org Chem* 5614.
- [67] García JI, Jiménez-Osés G, Martínez-Merino V, Mayoral JA, Pires E, Villalba I (2007) *Chem Eur J* 13:4064.