

# Quantum molecular modelling for efficient chemical synthesis

*Computer Aided Design informs the world around us, but it has taken a long time to have a significant impact on the chemical synthesis industry. Professor Martín Jaraíz at the University of Valladolid in Spain has proposed a novel approach to using a quantum mechanical tool called Density Functional Theory (DFT) in combination with microkinetic modelling to more accurately predict the effect of catalysts and additives on reaction performance. The proposed DFT-based modelling protocol is more computationally efficient and has the same or higher accuracy compared to commonly used protocols.*

The modern world is full of things designed on computers: your car, your toaster, your house and your computer itself. Computer Aided Design (CAD) is an essential tool for creating, modifying and optimising designs in civil and electronic engineering. But what about the chemical engineers creating pharmaceuticals, cosmetics and other synthetic chemicals – can they utilise CAD? Increasingly, the answer to that question is “yes”.

The big difference between designing buildings and chemical reaction mechanisms is the quantum mechanics involved in modelling molecules: the bricks and mortar of the chemical world are atoms with slightly blurry edges, and electrons with probability densities rather than definite locations.

## MICROKINETIC MODELLING

Modelling chemicals and their interactions – known as microkinetic modelling – is tricky, but there are in fact many methods. One of the most successful, Density Functional Theory (DFT), is a utilisation of quantum mechanics to investigate the movement of electrons around multiple atoms. This technology is still developing, and recent progression in DFT software allows synthetic chemists to make complex predictions of how molecules interact.

These analyses give a deeper level of insight into chemical processes than ever before, allowing chemists to choose effective catalysts and predict the effect of additives on a reaction.

Synthetic chemists are interested in producing new molecules and improving current production methods. These reactions, Figure 1A, often include multiple starting materials and products, and a successful computer analysis of a reaction will need to

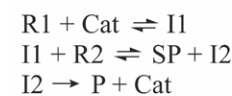
consider all of these, alongside solvents, catalysts, pre-catalysts and side-products. Over the course of a reaction, Figure 1B, these molecules will go through various ‘transition states’ (TS), Figure 1C, all of which also need to be accounted for. If that’s not enough, catalysts – substances which don’t get used up in the reaction – are regenerated, creating cycles, Figure 1D. Taking all of these into account forms a large network or ‘reaction scheme’ of chemicals and transition states, linked by reversible or irreversible reactions.

Typically, researchers using DFT methods start with a suspected full reaction scheme, which may contain more than 100 individual reactions. The energetics of each reaction within the scheme determine how easy it is for that reaction step to proceed. In the standard protocol, the majority of energetic parameters are fixed at the DFT calculated value, leaving a limited number of parameters that can be varied within the error boundaries of the calculation until the simulation matches experimental values. This process can be time-consuming, due to the computational load of calculating a large number of DFT parameters.

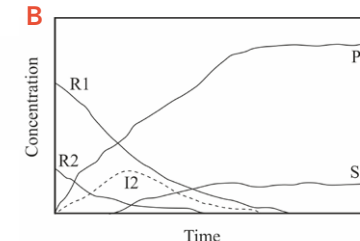
## TIGHTENING UP THE PROCEDURE

Over the last three years, Professor Martín Jaraíz at the University of Valladolid in Spain has been using another method – one which provides not only equally or even more accurate and reliable results, but also requires less experimental input and computation time. Rather than starting with every possible reaction pathway, instead begin with just the dominant mechanisms – Jaraíz calls these ‘loose’ and ‘tight’ modelling approaches. Starting with less information may seem counter-intuitive, but because fewer reactions are being modelled, all the parameters can be adjusted in the DFT process. Once the major mechanisms are in place, important additional

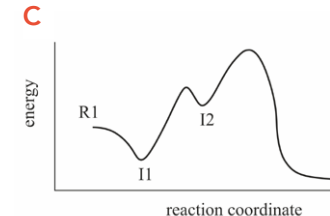
Figure 1: A



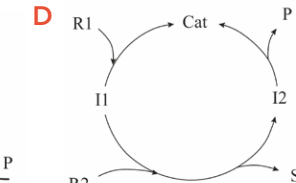
B



C



D



A reaction system (A) determines the time course of conversion of reactants into products (B), that is controlled by the energy path (C) along the reaction pathways (D) of the reaction system.

reactions can be added to fine-tune the model based on experimental data.

Building a model made up of many individual chemical reactions requires specific methods and tools of its own. In the laboratory, Reaction Progress Kinetic Analysis (RPKA) can be used to identify rate laws of the dominant mechanisms – that is, how quickly chemicals are converted into other chemicals. Outside the lab, theoretical approaches like Automated Reaction Path Searches can elucidate the important pathways and transition states. In combination, these can be used to build a reaction scheme of the various reactions that eventually lead to the synthesis of a molecule. The tight method requires initially identifying a minimum number of reactions that can explain the overall reaction.

Once a basic reaction scheme has been proposed, the next step of the DFT process is to consider more deeply

the energetics of each reaction step. If all the reactions in a mechanism are energetically plausible, then the proposed mechanism gets the ‘all clear’. On the other hand, models containing

valuable for explaining experimental spectroscopic observations, such as the slow build-up of particular molecules over time leading to catalyst deactivation.

**Recent progression in the combined use of DFT and microkinetic software allows synthetic chemists to predict the outcome of complex reaction systems.**

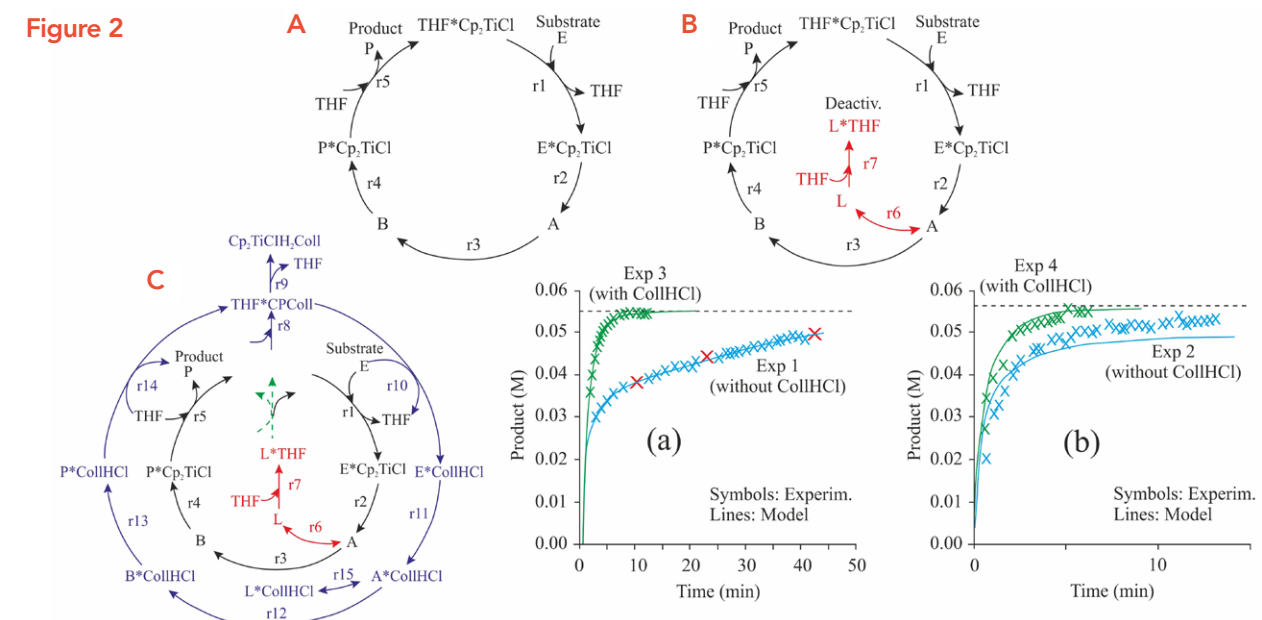
reactions which are energetically unfavourable – reactions which are too tricky – are unlikely to represent reality. It is also important to consider the effect of fluctuating levels of chemicals on reversible reactions and their transition states.

Once each reaction in the initial model has earned its place in the scheme, additional information can be incrementally added to the model for fine-tuning. These details are

## VERIFYING A PROPOSED MODEL

Professor Jaraíz and his team applied this new protocol to a series of case studies. The first of these is a reaction involving titanocene monochloride ( $Cp_2TiCl$ ), a catalyst sometimes used in ring-opening reactions of epoxides. A mild acid called collidine hydrochloride (CollHCl) is often added to this reaction, which prevents the  $Cp_2TiCl$  from being deactivated over the course of the reaction. The resulting reaction is highly efficient and converts

Figure 2



Example of model development following the incremental tight modelling protocol. Experimental data (symbols) and simulation (lines) for two sets of experiments, without CollHCl (Exp1, Exp2) and with CollHCl (Exp3, Exp4), for two initial epoxide concentrations. Only 3 data points (red symbols) of Exp1 were used throughout the incremental modelling process to adjust the 32-barrier tight model.

almost all of the substrate into the desired product, with only a small amount of unwanted side-product. Catalyst deactivation is a very important factor when deciding upon catalytic processes for large-scale industrial chemical synthesis.

Jaraíz and his team began their DFT modelling process with a reaction scheme (Fig. 2A) of five key reactions making up the catalytic cycle. The initial scheme did not take into account catalyst deactivation and the crucial role of CollHCl to suppress deactivation, in order to begin with a simpler reaction system for DFT. Only once the energetics of these reactions were established (i.e. calculated and fine-tuned), Jaraíz modified the reaction scheme by adding steps to account for deactivation (Fig. 2B) and the effect of the CollHCl additive (Fig. 2C).

To test the predictive capability of the procedure, the model was trained on only three experimental data points (Fig. 2A), and the data was collected from a reaction performed without CollHCl. The model was used to predict the results from the synthesis done with

## The use of predictive, DFT-based reaction simulators could have a significant economic impact on the chemical industry.

and without CollHCl. The model was very successful: it could very accurately predict the shape and yields for this (Fig. 2 A,B) and several other reagents.

For another case study, even the as-calculated DFT parameters predicted correct yields, the only difference with the simulation was that it predicted the reaction to occur about 360 times faster than it actually does. After fine-tuning the DFT parameters, the simulator could match the timescale for the reaction as well.

### PROPOSING A RELIABLE MODELLING PROTOCOL

Another case study illustrates the potency of the 'tight' method and how it can provide more reliable results (Fig. 3). For this case study, Jaraíz chose as a representative example a reaction

**Figure 3** Microkinetic Modelling Approaches

<b>Loose</b> (commonly used)	<b>Tight</b> (proposed)
<b>130</b> reactions (All plausible mechanisms)	<b>8</b> reactions (Only dominant mechanisms)
<b>149</b> DFT parameters (Only 9 adjusted)	<b>14</b> DFT parameters (All adjusted)
<b>X Can lead to wrong conclusions</b>	<b>✓ Same or better accuracy</b> <b>✓ &gt;10-fold DFT time savings</b> <b>✓ Incremental modelling</b>

The commonly used modelling approach (loose) includes many DFT parameters but, due to the limited experimental data, can only fine-tune a few of them. The proposed approach (tight) starts with only the main, dominant mechanisms (reactions), adjusts all the DFT parameters, and proceeds incrementally adding mechanisms to build a robust and more reliable full model.

from a recently published reaction scheme, proposed by another research group: the ring opening reaction of 1,2 epoxyoctane by 2 propanol, catalysed by fluorinated aryl borane (FAB).

This reaction is interesting because the presence of unwanted water can cause the regioselectivity of FAB to decrease,

resulting in lower amounts of the desired product. A reaction scheme had been previously proposed by another research group, created using the often used 'loose' modelling approach. To explain the effect of water on the reaction, the scheme contained 130 reactions, resulting in 149 DFT parameters. Given the large number of parameters, only nine were adjusted, and the remainder were set at the calculated DFT values.

For a tight modelling approach, Jaraíz chose eight for the initial scheme of 130. This resulted in a reduction in the number of DFT parameters to 14, all of which could be adjusted. This initial reaction scheme included the catalytic cycle of FAB to convert the substrate into the product, two deactivation reactions involving FAB, and two reactions creating side-products. The simulation

was fine-tuned using only data from five experiments. By this point, the simulation could predict the relationship between the amount of water and both the initial rate and regioselectivity of the reaction with high accuracy.

The biggest discrepancy between the tight and loose models is in the relationship between water level and regioselectivity. The loose model predicted a much stronger relationship than is actually seen in experiments, whereas the tight model provides more accurate values and did not include some unconventional water mediated reactions of the loose model. The tight model deems these steps unnecessary and, thus, reduces the chances of, for example, predicting multiple catalytically relevant species and multiple catalytic pathways whose existence is questionable and that result from the large number of reactions that have only been loosely adjusted and tested.

### CHEMICAL CAD

In summary, the tight protocol proposed by Prof Jaraíz is a highly effective substitute for the previous protocol that can provide more robust and reliable models and also drastically decreases computation time. CAD can increasingly be used for chemical synthesis, and this novel implementation of DFT-based kinetic modelling is another step towards fully predictive CAD for chemical processes.



# Behind the Research

## Prof Martín Jaraíz

**E:** [mjaraiz@ele.uva.es](mailto:mjaraiz@ele.uva.es) **T:** +34 983 423 677 **W:** [https://www.researchgate.net/profile/M\\_Jaraiz](https://www.researchgate.net/profile/M_Jaraiz)

## Research Objectives

Professor Jaraíz's research is focused on the development of atomistic models and simulation tools for materials processing.

## Detail

Department of Electronics, ETSIT, University of Valladolid, Paseo de Belen 15, Valladolid 47011, Spain

### Bio

Professor Martín Jaraíz obtained his Ph.D. from the University of Valladolid in 1981. He has carried out kinetic Monte Carlo for semiconductor process modelling for ATT Bell Laboratories, New Jersey and Synopsys Inc., California. While exploring molecular electronics he became interested on improving reaction kinetics modelling for synthetic chemistry in general.

### Funding

MINECO-FEDER (CTQ2015-68148-C2) and Junta de Castilla y León (VA056G18)

### Collaborators

- Dr José E. Rubio (Univ. Valladolid)
- Dr Lourdes Enríquez (Univ. Valladolid)
- Dr Ruth Pinacho (Univ. Valladolid)
- Prof José L. López-Pérez (Univ. Panamá)
- Prof Alberto Lesarri (Univ. Valladolid)

## References

Jaraíz, M; Rubio JE; Enríquez, L; Pinacho, R; Lopez-Pérez, JL; Lesarri, A. (2019). An Efficient Microkinetic Modeling Protocol: Start with Only the Dominant Mechanisms, Adjust All Parameters, and Build the Complete Model Incrementally. *ACS Catalysis*, 9, 4804-4809. <https://doi.org/10.1021/acscatal.9b00522>

Jaraíz, M; Enríquez, L; Pinacho, R; Rubio JE; Lesarri, A; Lopez-Pérez, JL. (2017). A DFT-Based Computational-Experimental Methodology for Synthetic Chemistry: Example of Application to the Catalytic Opening of Epoxides by Titanocene. *Journal of Organic Chemistry*, 82, 3760-3766. <https://doi.org/10.1021/acs.joc.7b00220>

## Personal Response

**To what extent do industry use predictive techniques like DFT, and in what ways do they stand to benefit from more implementation of these methods?**

**//** Kinetic simulations have been used for decades, especially for industrial applications. However, instead of the lengthy and complex DFT calculations, they commonly use empirical models based on chemical intuition and expertise, and the parameters are adjusted from dedicated experiments. We can expect that the widespread use of predictive computational tools such as DFT-based microkinetic simulators, together with concerted and well-planned interaction between computation and experiment, will contribute to a synthetic chemistry less based on trial and error and more on predictive assessments. In addition, considering the yearly volume of industrial chemical processing that could benefit from this predictive simulation tool, any improvement in this direction would have a huge economic impact. **//**