



# An Efficient Microkinetic Modeling Protocol: Start with Only the Dominant Mechanisms, Adjust All Parameters, and Build the Complete Model Incrementally

Martín Jaraíz,<sup>\*,†,¶,||</sup> José E. Rubio,<sup>†</sup> Lourdes Enríquez,<sup>†</sup> Ruth Pinacho,<sup>†</sup> José L. López-Pérez,<sup>||,§,ⓑ</sup> and Alberto Lesarri<sup>‡,ⓑ</sup>

<sup>†</sup>Department of Electronics, ETSIT, University of Valladolid, Paseo de Belén 15, Valladolid 47011, Spain

<sup>¶</sup>IU CINQUIMA, University of Valladolid, Paseo de Belén 7, 47011 Valladolid, Spain

<sup>||</sup>Department of Pharmacology, Faculty of Medicine, University of Panama, Panama City, Panama

<sup>§</sup>Department of Pharmaceutical Sciences - IBSAL-CIETUS, University of Salamanca, Avda. Campo Charro s/n, Salamanca 37071, Spain

<sup>‡</sup>Department of Physical Chemistry and Inorganic Chemistry, Faculty of Sciences, University of Valladolid, Paseo Belén 7, Valladolid 47011, Spain

## Supporting Information

**ABSTRACT:** In a recent article (ACS Catal. 2018, 8, 11119–11133), a comprehensive catalytic mechanism is proposed to explain the effects of residual water on the reactivity and regioselectivity of tris(pentafluorophenyl)borane catalyst in the ring-opening reaction of 1,2-epoxyoctane by 2-propanol. Using it as a representative example of a common trend followed also by other groups, we show that the heavily under-constrained (*loose*) kinetic modeling approach employed can lead to several pitfalls and propose an alternative, more stringent (*tight*) modeling protocol to avoid them. In addition to providing similar or better accuracy, this approach considerably reduces the DFT parameter calculation time (by a factor of 10 in the present case). We also show an example of how delayed or second-order mechanisms can then be added incrementally to the already built and tested, first-approximation model to achieve a highly predictive and comprehensive microkinetic model. We hope that this simple and robust microkinetic modeling protocol may contribute to the current efforts to establish new, more predictive computational methodologies for synthetic chemistry.

**KEYWORDS:** modeling, mechanism, DFT, kinetic, microkinetic

Microkinetic Modelling Approaches	
<i>Loose</i> (commonly used)	<i>Tight</i> (proposed)
130 reactions (All plausible mechanisms)	8 reactions (only Dominant mechanisms)
149 DFT parameters (only 9 adjusted)	14 DFT parameters (All adjusted)
Can lead to wrong conclusions	<ul style="list-style-type: none"><li>• Same or better accuracy</li><li>• &gt; 10-fold DFT time savings</li><li>• Incremental modelling</li></ul>

## INTRODUCTION

In a recent article published in this journal,<sup>1</sup> Yu et al. proposed a comprehensive catalytic mechanism to explain the effects of residual water on the reactivity and regioselectivity of tris(pentafluorophenyl)borane catalyst in the ring-opening reaction of 1,2-epoxyoctane by 2-propanol. Like other groups,<sup>2–6</sup> they have followed the common procedural approach of starting, right from the beginning, with all the mechanisms and equations that are suspected to play some role and adjust only a reduced set of the many independent parameters involved. Because the system of equations is heavily under-constrained (many parameters, few experimental constraints), such a loose modeling approach can lead, as shown for the present case, to several pitfalls.

The purpose of this article is to show those pitfalls and demonstrate that the inverse strategy, that is, starting with the *minimum number of reactions* that can describe the expected *dominant* mechanisms and *adjusting all* the DFT parameters,

generates a kinetic model that reproduces the experimental data with an accuracy that is comparable to or better than their complex model and uses fewer training experiments. More importantly, this approach reduces the chances of, for example, predicting “multiple catalytically relevant species and multiple catalytic pathways”<sup>1</sup> whose existence is questionable and that result from the large number of reactions that have only been loosely adjusted and tested.

The use of this tighter protocol helps keep under control the mechanisms and equations included because, during the model development process, each equation has proven that its contribution is necessary to explain some experimental feature. Although microkinetic models can also be developed empirically,<sup>7</sup> they are usually obtained through DFT

**Received:** February 3, 2019

**Revised:** April 3, 2019

**Published:** April 17, 2019

calculations. In this case, as a bonus, the reduced number of equations of the tight modeling scheme translates into a reduced number of DFT parameters that need to be calculated and, therefore, in a considerable reduction of time (about 10-fold in the case considered here) and effort (each transition state needs to be dealt with individually). Finally, we also show an example of how a high-performance DFT-based kinetic model was *built incrementally* using this tight modeling protocol.

The rationale behind this approach is that (1) there is no reason for fixing some parameters at the raw values obtained with a specific DFT functional,<sup>8</sup> and (2) each elementary reaction step (single transition state) can best reproduce the intended mechanism if all its parameters (enthalpies and entropies) are set to the appropriate value. Once this basic, initial model has been fine-tuned and understood, additional pathways and equations for second-order effects can be progressively added and adjusted to obtain a more detailed and accurate account of available experimental data.

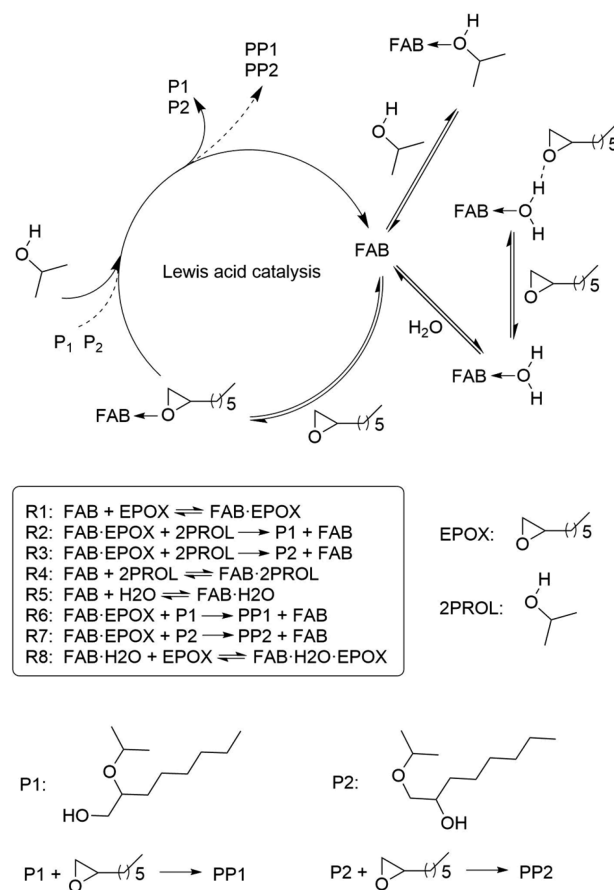
We start with a short summary of the chemical problem and its motivation, as described in ref 1. Tris(pentafluorophenyl)-borane (also known as fluorinated aryl borane or “FAB”) is an excellent Lewis acid catalyst for alkoxylation catalysis with monosubstituted aliphatic epoxides, where FAB achieves significantly higher regioselectivity (80%) than other catalysts (<50%). Even a few percent increase of regioselectivity, achieved through a better understanding of the catalytic mechanism, would have a high impact in the 50 billion dollar per year polyurethane industry, since the primary alcohols in  $\beta$ -alkoxy polyether polyols are more reactive with the isocyanate component of polyurethanes.

A model reaction was used (Scheme 1), consisting of 1,2-epoxyoctane ring-opening with 2-propanol catalyzed by FAB, giving primary-alcohol product P1 and secondary-alcohol product P2. Preliminary experiments demonstrated that residual water at ppm levels dramatically decreased the reaction rate and altered the regioselectivity. The comprehensive catalytic mechanism proposed in ref 1 features competitive binding reactions, traditional Lewis acid catalysis, and nonconventional water-mediated and alcohol-mediated catalysis. The microkinetic model, herein after referred to as loose model, consists of 130 reactions with a total of 149 independent parameters. They used 9 adjustable parameters to influence the rate constants of 29 reactions to achieve optimal agreement between the model output and experimental kinetic data in the training set (7 out of 21 experiments). The model was verified with the other 14 experiments, yielding moderate to good predictions, with some outliers.

## DOMINANT MECHANISMS

Out of their 130 reactions, we have selected 8 reactions (Scheme 1) as the dominant or principal mechanisms. The rationale for the selection of the 8 pathways included in Scheme 1 is as follows. (a) We start with only the traditional Lewis acid catalytic cycles that generate the experimentally observed products P1 and P2 (reactions R1, R2, R3), plus the competitive binding reactions of FAB with the alcohol and water (reactions R4, R5), and compare the simulation results with the experimental data. (b) To improve the agreement simulation experiment we add reactions R6 and R7 because P1 and P2 can react with the epoxide to give second-generation ring-opened products PP1 and PP2. At this point we already

**Scheme 1. Proposed Dominant Mechanisms and the Corresponding Elemental Reaction Steps Implemented in the Tight Model**

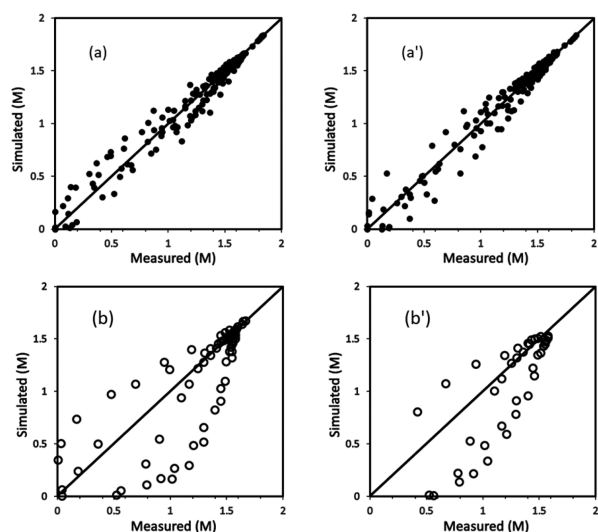


get a fairly good agreement. (c) To improve it further, we added the second-shell binding of the epoxide (reaction R8).

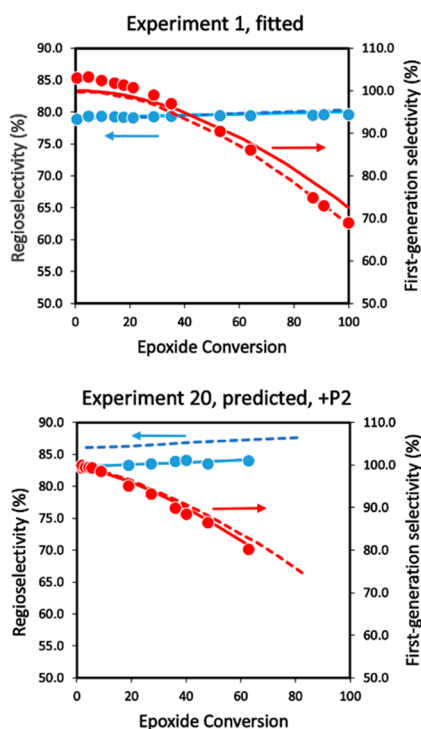
We have implemented this system in the kinetic simulator COPASI,<sup>9</sup> that generates and solves numerically the differential equations corresponding to the reaction system. Once the reaction system has been decided (Scheme 1), the next step is fine-tuning the raw DFT values to reproduce some experimental data.<sup>8</sup> We optimized (fine-tuned) all 14 independent parameters within the ranges proposed in ref 1 (using  $\pm 2$  kcal/mol for unspecified ranges); the details are in the Supporting Information. Typically, this is done with two sets of experiments (runs), one set for adjusting the model and the other set for model verification. We have used a set of only 5 experiments for training (runs 1, 2, 11, 14, and 16). Figure 1a is a parity plot for all experiments except the outliers, that are plotted in Figure 1b, and the accuracy can be compared to Figure 1a' and b' corresponding to the results from the loose model of ref 1.

Figure 2 is an example of experimental regioselectivity and mass-balance data versus model output that can be obtained with this tight model and with the loose model. Figure 3 also displays good experiment–simulation agreement for the dependence, on the initial water level, of the initial rate constant and regioselectivity. All the simulations presented here can be carried out with the COPASI input files and instructions provided in the Supporting Information.

Table S1 shows that the optimized parameters do not differ significantly from those calculated at the B3LYP-D3BJ/6-

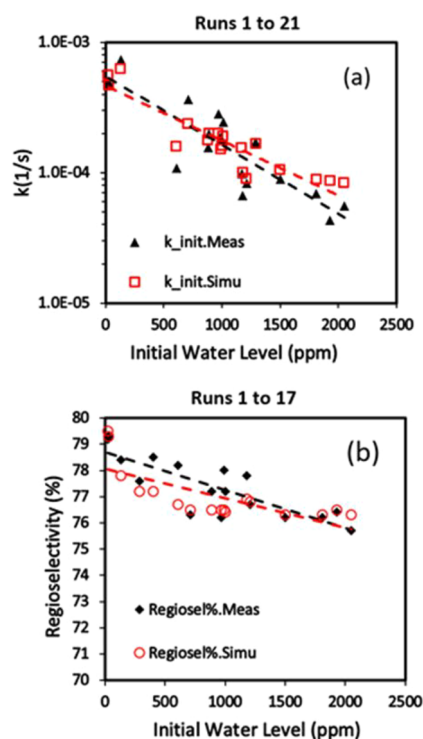


**Figure 1.** (a) Tight model epoxide parity plots for all experiments except the outliers, that are plotted in (b). (a') and (b') are the corresponding loose model plots, with data from ref 1.

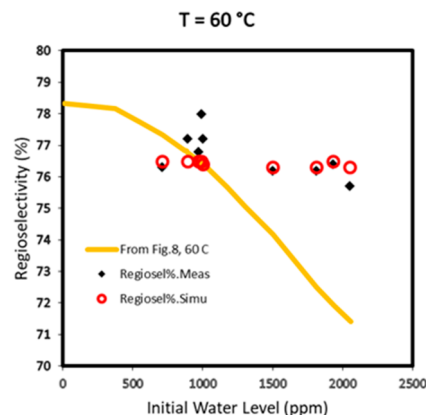


**Figure 2.** Examples of experimental regioselectivity and mass-balance data versus model output for Experiment 1 (used for fitting) and Experiment 20 (predicted). Symbols: experimental data, from ref 1. Solid lines: tight model output. Dashed lines: loose model output, data from ref 1.

31+G(d, p) level of theory, which says that at this level of theory the description of the reaction is rather good. The largest discrepancy between our tight model and the loose model is the strong dependence of regioselectivity on initial water level at, for example, 60 °C reported for the loose model. Our model predicts an almost constant value of 76.3, much closer to the 60 °C experiments (Figure 4). Those experiments can also be seen in Figure 3b.



**Figure 3.** Experimental and simulated data for effective initial rate constants (a) and regioselectivity profiles at 40% epoxide conversion (b) for the FAB-catalyzed ring-opening of 1,2-epoxyoctane with 2-propanol as a function of initial water level. The dotted trend lines serve as guides to the eyes.



**Figure 4.** Overall regioselectivity dependence on the initial water level at  $T = 60$  °C. For the solid line, from ref 1, other initial conditions were fixed at  $[epoxide]_0 = 1.6$  M,  $[2-propanol]_0 = 6.4$  M, and  $[FAB]_0 = 1.6$  mM. Symbols correspond to the experiments at 60 °C and similar other initial conditions.

## DISCUSSION

Therefore, according to our model, the *dominant* effect of water is just the deactivation of a fraction of the catalyst by reversible binding, reactions R5 and R8 in Scheme 1. This model, with only Lewis acid catalysis, accounts for the experimental data with equal or better accuracy than the model with nonconventional catalytic mechanism featuring multiple catalytically relevant species and multiple catalytic pathways. In fact, we also tested the addition of the water-mediated catalytic cycle to the Lewis acid cycle, but the best fit to the experimental data was obtained when the P1 and P2



product contribution from the water cycle was negligible. Only reactions R5 and R8 were necessary to account for the observed dependence on the initial water level.

This observation casts doubts on the validity of some of the conclusions drawn from the loose model. For example, “the effect of increasing water level on lowering the overall regioselectivity can be explained by the increasing accumulated reaction flux through the water-mediated catalytic pathways”,<sup>1</sup> but the same effect is reproduced by our tight model (Figure 3b, dotted lines) without water-mediated catalytic pathways. In addition, “the slightly increasing regioselectivity with conversion can also be explained by more contribution from the water-mediated catalytic pathways with P2 being the HBA that has much higher regioselectivity than the Lewis-acid-catalyzed pathways”.<sup>1</sup> However, as Figure 2 shows, our model, with only Lewis-acid-catalyzed and no water-mediated pathways, reproduces accurately the same experimental observation.

The reliability of the conclusions derived from the flux and speciation analyses conducted “to understand the partitioning of reaction flux through all catalytic pathways in the complex mechanism and their individual contributions to the overall regioselectivity”<sup>1</sup> is also questionable. For example, they extrapolate the effect of initial water content, already away from the experimental data at 2000 ppm as shown in Figure 4, and conclude that “In fact, at a water level of 5000 ppm, more than 78% of the accumulated reaction flux went through water-mediated catalysis”.<sup>1</sup> However, our tight model is closer to the experimental data (Figure 4) and does not have any contribution from a water-mediated catalytic cycle. In addition, if the water-mediated flux is negligible compared to the Lewis-acid catalysis, as our model argues, then the speciation analysis is pointless and the recommendation of “tuning the overall regioselectivity via modulating the contributions of the water-mediated catalytic pathways”<sup>1</sup> has no support.

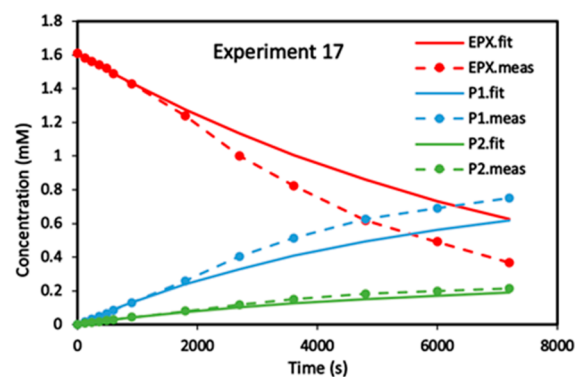
However, our analysis is not meant to be a criticism of the work of ref 1, used only as a representative example of a common trend followed by other groups. It is intended as a constructive contribution to the task of finding the best computational methodologies for microkinetic modeling. The frequently used loose modeling approach, with many mechanisms and very few constraints, can lead to a variety of different but plausible solutions (models). That is, for a given reaction scheme it is possible to obtain different sets of parameter values (optimized with the same set of experimental data) in such a way that each of them provides a model that reproduces the experimental data by conveniently activating some of the many different pathways. In principle, this indetermination could be avoided with sufficiently accurate DFT or ab initio parameter values (narrow adjustment ranges), but for most practical cases, those calculations are currently out of reach.

To avoid those pitfalls, we propose a tight modeling alternative: starting with the minimum number of reactions that can describe the expected dominant mechanisms and adjusting all the DFT parameters. We believe that the predictive character of the resulting basic model stems from the use and calibration of a set of full mechanisms that play a well concerted role, instead of assembling together many reaction equations, that are fitted only in part.

## ■ INCREMENTAL MODELING

In addition, each of the reactions included play a specific role and provide *well-grounded pillars* to continue building the

model with additional mechanisms, toward a full account of all the experimental data. For example, under the conditions of experiment 17, a faster process, not included in the current mechanisms of Scheme 1, shows up after about 1000 s (Figure 5). A similar behavior is exhibited by several other experi-



**Figure 5.** Measured data of experiment 17 reveals that a faster process, not accounted for by the basic, first-approximation model of Scheme 1, begins after an induction period of about 1000 s. Symbols: experimental data from ref 1; solid lines: model output.

ments. This may correspond to the time needed to build up a steady state concentration of an intermediate, like  $\text{FAB}\cdot\text{H}_2\text{O}$ , that feeds a water-mediated catalytic cycle with rates and regioselectivities different from those of the Lewis acid catalytic cycle, as already suggested in ref 1.

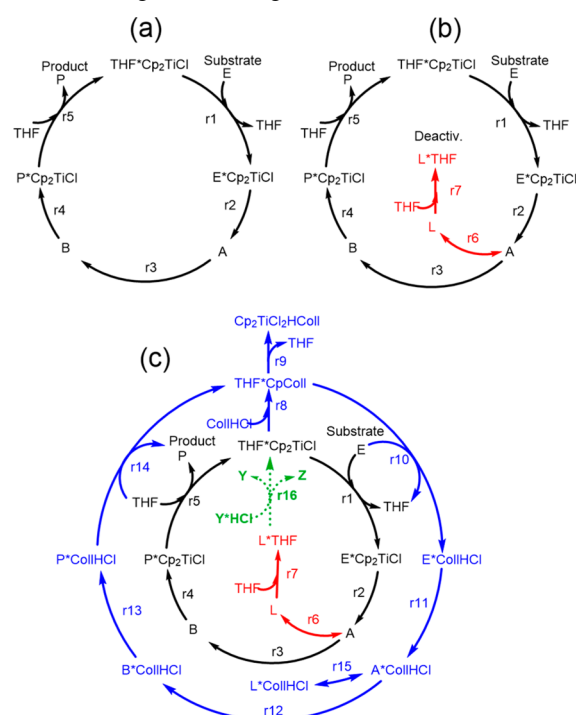
In such cases, one can develop and calibrate a starting model to reproduce the initial stages of the reaction system and then add other mechanisms. In fact, for the training experiments 1, 2, and 11 we used only the data points up to 1000 s. Then, the water-mediated and alcohol-mediated catalytic cycles could be subsequently added, as a refinement of the mechanisms already implemented (and calibrated to a first approximation). This protocol is reminiscent of an experimental counterpart, the reaction progress kinetic analysis (RPKA):<sup>10</sup> “In more complicated cases, these experiments serve as a preliminary analysis alerting us that the reaction under study exhibits complexity beyond the simple reaction network of Scheme 1 and that more work needs to be done.”

In particular, in the example discussed here, it would be advisable to first develop and test a model for experiments at the lowest possible water level, to establish the parameters of the reactions of Scheme 1 that do not include water, and then use a set of experiments with water to refine only the parameters of the reactions with water, as we show below with another example. Overall, to reinforce the strategy of starting with only the dominant mechanisms, it would be desirable to follow an incremental modeling protocol as much as possible, whether through subsequent temporal stages of the system (Figure 5) or by first modeling the system without additives that are not essential and that only modify the system behavior (water), or a combination of both.

To recapitulate, Scheme 2 shows an example of the modeling stages followed to develop a reliable and predictive catalytic kinetic model applying this incremental tight modeling protocol (see ref 8 for a detailed derivation of the model):

**1. Main Catalytic Cycle.** Scheme 2a corresponds to the essential mechanisms for the catalytic opening of an epoxide (E) by titanocene ( $\text{Cp}_2\text{TiCl}$ ) in tetrahydrofuran (THF).

**Scheme 2. Example of Model Development Following the Incremental Tight Modeling Protocol<sup>8</sup>**

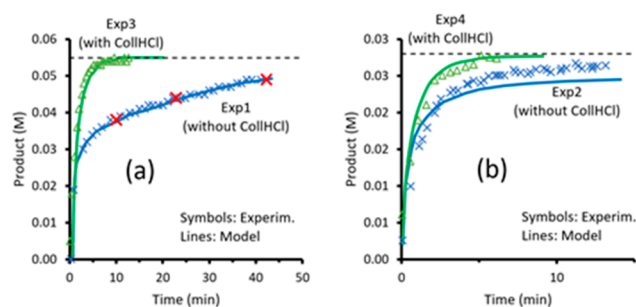


**2. Deactivation Pathway.** In Scheme 2b, an off-cycle pathway (r6, r7) was added to account for the catalyst deactivation observed in some experiments. This mechanism also explains several other experiments performed with different versions of the catalyst.

**3. Mechanism to Suppress Deactivation.** In the next stage, Scheme 2c, r8 through r15 were included to model the effects of an additive, collidine hydrochloride, that is known to prevent catalyst deactivation.

**4. Additional Mechanisms (Chlorohydrins).** Finally, r16 explains why some experiments can also generate chlorohydrins.

Figure 6 shows four product vs time plots from experiments with and without an additive (collidine hydrochloride), at high and low reaction concentrations. The basic catalytic cycle (Scheme 2a) reproduces Exp3 and Exp4 well, because those experiments (with collidine hydrochloride) do not exhibit



**Figure 6.** Experimental data (symbols, from ref 11.) 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 (large symbols) of Exp1 were used throughout the incremental modeling process to adjust the 32-barrier tight model.

deactivation effects. The next level model (Scheme 2b) can only reproduce Exp1 and Exp2, which show deactivation, but cannot predict its suppression by collidine hydrochloride. Finally, the augmented model of Scheme 2c not only reproduces these four experiments but also predicts the effect of using other pyridinium hydrochlorides and many other experimental observations, even quantitatively to a first approximation.<sup>8</sup>

It should be stressed that only three experimental data points (large symbols shown in Exp1) were used for fine-tuning the model. The level of prediction achieved by this incremental modeling protocol, starting from the dominant mechanisms and adjusting all the parameters, attests to the validity of the proposed methodology, as a means to leverage the firm physical bases underlying the DFT calculations.

At first glance, by using only three data points to adjust all the parameters of Scheme 2c, it might seem that we are again applying the loose modeling approach. The difference, though, is that each incremental modeling stage has been built in two steps. For example, for Scheme 2b: (1) using the r1–r5 parameters obtained from Scheme 2a, adjust only r6 and r7; (2) adjust all parameters (r1–r7). Thus, in the last modeling stage (Scheme 2c) we can safely use only three points to refine so many parameters because the system, starting from the dominant mechanisms, has already been settled by the previous modeling steps within the correct local minimum basin. Indeed, this accurate model cannot be obtained by directly trying to adjust Scheme 2c to fit the three data points starting from the as-calculated DFT parameter values.

The value added to the lengthy (typically months) and already widespread DFT calculations by just adding three data points from a single experiment and a few hours to implement the kinetic simulator is noteworthy.

Finally, it is also worth reflecting on the fact that, although only three experimental data points were needed to supplement the theoretical calculations, they were at the same time indispensable to make the predictive power of DFT calculations a reality. This remark has a bearing on current efforts to develop theoretical-only methods that can yield predictions with the accuracy (Figure 6) required for direct use in synthetic chemistry: even if the kinetic parameters of Figure 6 were calculated with, let us say, 0.1 kcal/mol accuracy, it would still be necessary to figure out and calculate also all the possible interaction pathways down to that interaction energy level. Since those accurate and complex calculations are not yet available, more accessible alternatives are being sought in the meantime, like the one that we have discussed here.<sup>12</sup>

## CONCLUSION

In conclusion, in line with the often quoted Einstein's razor ("it can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible, without having to surrender the adequate representation of a single datum of experience")<sup>13</sup> we have proposed an accurate and efficient tight modeling protocol (as an alternative to a loose approach in common use) that can be improved incrementally. We hope that this modeling proposal may contribute to the set of computational tools designed to assist and complement the experimental processing tasks performed in synthetic chemistry.

## ■ ASSOCIATED CONTENT

### § Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acscatal.9b00522](https://doi.org/10.1021/acscatal.9b00522).

Adjusted DFT parameter values, COPASI input files, and instructions (ZIP)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [mjaraiz@ele.uva.es](mailto:mjaraiz@ele.uva.es).

### ORCID

Martín Jaraíz: 0000-0001-6688-3158

José L. López-Pérez: 0000-0002-3728-7602

Alberto Lesarri: 0000-0002-0646-6341

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

Funding by MINECO-FEDER (CTQ2015-68148-C2) and Junta de Castilla y León (VA056G18) is gratefully acknowledged.

## ■ REFERENCES

- (1) Yu, Y.; Zhu, Y.; Bhagat, M. N.; Raghuraman, A.; Hirsekorn, K. F.; Notestein, J. M.; Nguyen, S. T.; Broadbelt, L. J. Mechanism of Regioselective Ring-Opening Reactions of 1,2-Epoxyoctane Catalyzed by Tris(pentafluorophenyl)borane: A Combined Experimental, Density Functional Theory, and Microkinetic Study. *ACS Catal.* **2018**, *8*, 11119–11133.
- (2) Mhadeshwar, A. B.; Vlachos, D. G. Hierarchical, multiscale surface reaction mechanism development: CO and H<sub>2</sub> oxidation, water–gas shift, and preferential oxidation of CO on Rh. *J. Catal.* **2005**, *234*, 48–63.
- (3) Gokhale, A. A.; Dumesic, J. A.; Mavrikakis, M. On the Mechanism of Low-Temperature Water Gas Shift Reaction on Copper. *J. Am. Chem. Soc.* **2008**, *130*, 1402–1414.
- (4) Rush, L. A.; Pringle, P. G.; Harvey, J. N. Computational Kinetics of Cobalt-Catalyzed Alkene Hydroformylation. *Angew. Chem., Int. Ed.* **2014**, *53*, 8672–8676.
- (5) Szlapa, E. N.; Harvey, J. N. Computational Modelling of Selectivity in Cobalt-Catalyzed Propene Hydroformylation. *Chem. - Eur. J.* **2018**, *24*, 17096–17104.
- (6) For a recent review on microkinetic modeling see: Besora, M.; Maseras, F. Microkinetic modeling in homogeneous catalysis. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1372.
- (7) Brezny, A. C.; Landis, C. R. Development of a Comprehensive Microkinetic Model for Rh(bis(diazaphospholane))-Catalyzed Hydroformylation. *ACS Catal.* **2019**, *9*, 2501–2513.
- (8) Jaraíz, M.; Enríquez, L.; Pinacho, R.; Rubio, J. E.; Lesarri, A.; López-Pérez, J. L. A DFT-Based Computational-Experimental Methodology for Synthetic Chemistry: Example of Application to the Catalytic Opening of Epoxides by Titanocene. *J. Org. Chem.* **2017**, *82*, 3760–3766.
- (9) COPASI v 4.24, available free of charge from <http://copasi.org>.
- (10) Blackmond, D. G. Kinetic Profiling of Catalytic Organic Reactions as a Mechanistic Tool. *J. Am. Chem. Soc.* **2015**, *137*, 10852–10866.
- (11) Gansäuer, A.; Behlendorf, M.; von Laufenberg, D.; Fleckhaus, A.; Kube, C.; Sadasivam, D. V.; Flowers, R. A., II *Angew. Chem., Int. Ed.* **2012**, *51*, 4739.
- (12) As an analogy, one can estimate (calculation) the length of a piano's strings, based on the desired frequencies, and then hand-tune (experiment) the stress of each string to get the exact frequencies. Much more complex calculations would be required to also calculate

(theoretical-only approach) the exact stress needed, based on material properties and environmental conditions.

(13) Kozuch, S. Reply to Comment on “Turning Over’ Definitions in Catalytic Cycles”. *ACS Catal.* **2013**, *3* (3), 380.