DISCUSSIONS

1

Computational and theoretical approaches for mechanistic understanding: general discussion

Matthias Bauer, Jamie Cadge, David Davies, Derek Durand, Odile Eisenstein, Daniel Ess, Natalie Fey, Simone Gallarati, Mike George, Alexander Hamilton, Jeremy Harvey, Ulrich Hintermair, Alison Hulme, Youichi Ishii, Vidar R. Jensen, Guy Lloyd-Jones, Jennifer Love, Jason Lynam, Stuart Macgregor, Todd Marder, Evert Jan Meijer, Patrick Morgan, Robert Morris, Joseph Mwansa, Aaron Odom, Robin Perutz, Markus Reiher, Laurel Schafer, John Slattery and Tom Young

Please check this proof carefully. Our staff will not read it in detail after you have returned it.

Please send your corrections either as a copy of the proof PDF with electronic notes attached or as a list of corrections. **Do not edit the text within the PDF or send a revised manuscript** as we will not be able to apply your corrections. Corrections at this stage should be minor and not involve extensive changes.

Proof corrections must be returned as a single set of corrections, approved by all co-authors. No further corrections can be made after you have submitted your proof corrections as we will publish your article online as soon as possible after they are received.

Please ensure that:

- The spelling and format of all author names and affiliations are checked carefully. You can check how we have identified the authors' first and last names in the researcher information table on the next page. **Names will be indexed and cited as shown on the proof, so these must be correct.**
- All of the editor's queries are answered.
- Any necessary attachments, such as updated images or ESI files, are provided.

Translation errors can occur during conversion to typesetting systems so you need to read the whole proof. In particular please check tables, equations, numerical data, figures and graphics, and references carefully.

Please return your **final** corrections, where possible within **48 hours** of receipt, by e-mail to: faraday@rsc.org. If you require more time, please notify us by email.

Researcher information

Please check that the researcher information in the table below is correct, including the spelling and formatting of all author names, and that the authors' first, middle and last names have been correctly identified. **Names will be indexed and cited as shown on the proof, so these must be correct.**

If any authors have ORCID or ResearcherID details that are not listed below, please provide these with your proof corrections. Please ensure that the ORCID and ResearcherID details listed below have been assigned to the correct author. Authors should have their own unique ORCID iD and should not use another researcher's, as errors will delay publication.

Please also update your account on our online [manuscript submission system](#) to add your ORCID details, which will then be automatically included in all future submissions. See [here](#) for step-by-step instructions and more information on author identifiers.

First (given) and middle name(s)	Last (family) name(s)	ResearcherID	ORCID iD
Matthias	Bauer		
Jamie	Cadge		
David	Davies		
Derek	Durand		
Odile	Eisenstein		
Daniel	Ess		
Natalie	Fey		
Simone	Gallarati		
Mike	George		
Alexander	Hamilton		
Jeremy	Harvey		
Ulrich	Hintermair		
Alison	Hulme		
Youichi	Ishii		
Vidar R.	Jensen		
Guy	Lloyd-Jones		
Jennifer	Love		
Jason	Lynam		
Stuart	Macgregor		
Todd	Marder		
Evert Jan	Meijer		
Patrick	Morgan		
Robert	Morris		
Joseph	Mwansa		
Aaron	Odom		
Robin	Perutz		
Markus	Reiher		
Laurel	Schafer		
John	Slattery		

Tom	Young		
-----	-------	--	--

Queries for the attention of the authors

Journal: **Faraday Discussions**

Paper: **c9fd90073j**

Title: **Computational and theoretical approaches for mechanistic understanding: general discussion**

For your information: You can cite this article before you receive notification of the page numbers by using the following format: (authors), Faraday Discuss., (year), DOI: 10.1039/c9fd90073j.

Editor's queries are marked on your proof like this **1**, **2**, etc. and for your convenience line numbers are indicated like this 5, 10, 15, ...

Please ensure that all queries are answered when returning your proof corrections so that publication of your article is not delayed.

Query Reference	Query	Remarks
1	[INFO-1] Please carefully check the spelling of all author names. This is important for the correct indexing and future citation of your article. No late corrections can be made. FOR YOUR INFORMATION: You can cite this paper before the page numbers are assigned with: (authors), Faraday Discuss., DOI: 10.1039/c9fd90073j.	

DISCUSSIONS

■ Computational and theoretical approaches for mechanistic understanding: general discussion

Matthias Bauer, Jamie Cadge, David Davies, Derek Durand, Odile Eisenstein, Daniel Ess, Natalie Fey, Simone Gallarati, Mike George, Alexander Hamilton, Jeremy Harvey, Ulrich Hintermair, Alison Hulme, Youichi Ishii, Vidar R. Jensen, Guy Lloyd-Jones, Jennifer Love, Jason Lynam, Stuart Macgregor, Todd Marder, Evert Jan Meijer, Patrick Morgan, Robert Morris, Joseph Mwansa, Aaron Odom, Robin Perutz, Markus Reiher, Laurel Schafer, John Slattery and Tom Young

DOI: 10.1039/C9FD90073J

(600:[600]600) **Simone Gallarati** began a discussion of the paper by Stuart A. Macgregor: Could you comment on the importance of the choice of solvent, experimentally and more importantly computationally? How did this affect your results?

Stuart Macgregor responded: Experimentally, the competition reactions were performed in a 4 : 1 mixture of dichloromethane:methanol. In methanol, the reactions both proceed more rapidly and also go more to full completion, while dichloromethane was introduced to ensure the solubility of all products. We therefore ran calculations incorporating a correction for pure methanol and a second set (in the supporting information) with pure dichloromethane. Importantly, the trends are identical with these two approaches, but with higher computed barriers in dichloromethane (consistent with the slower reactions observed experimentally). The increased barriers are due to the more difficult Cl^- dissociation step (**D** to **E** in the paper). We feel justified in emphasizing the results in methanol as other experimental work has shown similar C–H activation reactions to be promoted upon the addition of just a few equivalents of methanol. This may be due to specific solvation of the Cl^- anion as it dissociates. The computed barriers in methanol are also more consistent with most of the reactions proceeding at a reasonable rate at room temperature experimentally. We also found the balance between intramolecular AMLA/CMD and an external CMD process was solvent dependent: in methanol the former is clearly favoured, whereas in dichloromethane external CMD can become competitive, especially for the more electron-withdrawing substituents.

(601:[601]601) **Alexander Hamilton** asked: Did you see any relationships between the QTAIM results and your transition states parameters and trends?

Stuart Macgregor answered: Yes – firstly, Ru...H-C and C-H...O bond paths were located indicating the presence of both an agostic interaction and H-bonding to acetate; however the associated bond critical point electron densities are low, consistent with these being rather ‘early’ transition states in terms of C-H bond cleavage. A clear trend in BCP electron densities was still seen, however, as a function of the *para*-substituent, with both the agostic and H-bonding character becoming stronger when moving from a *para*-NMe₂ to a *para*-NO₂ substituent. This was backed-up by trends in the NBO and NCI analyses, as well as by variations in the computed distances.

(602:[602]602) **Guy Lloyd-Jones** remarked: Your superb paper powerfully combines experiment and computation to draw insightful conclusions on cyclo-metallation mechanisms, and kinetic/thermodynamic control. The impact of the work is much greater than the sum of its parts. It also highlights an important question regarding semantics: what do you think the literature understands by the term ‘C-H activation’? And specifically, how does this relate to the processes of C-H-interaction (*e.g.* agostic), H-bonding, C-H elongation, and ultimately C-H cleavage? I think that there is a major problem (in the literature) arising from the conflation of C-H cleavage with C-H activation.* For example, you conclude, based on your data (and I agree), that the process is not an ‘aromatic electrophilic substitution’ – however, here we are referring to the mechanism, not the overall process, which is substitution of H by a metal complex ‘ML’ on an aromatic ring, and for some combinations of M, ligand and ionic state, could arguably be correctly referred to based on electronegativity, as ‘aromatic electrophilic substitution’; although of course in your system, Fig 3 in the paper suggests the opposite.

Moreover, there are frequent misunderstandings evident in conclusions drawn in the literature from the measurement of a kinetic isotope effect (Ar-H *versus* Ar-D; *e.g.* by competition), often present as a general ‘add-on’ to a synthetic/catalyst development paper. Indeed, I think that there may be many hundreds of publications on metal catalysed C-H substitutions, which contain incorrect statements (in terms of logic) regarding whether the KIE indicates that C-H activation is, or is not, ‘rate-limiting’ under a catalytic regime. Alarmingly, the refereeing process does not pick up on this, suggesting that the referees could be just as misinformed as the authors. I propose that an essay on the mechanisms and semantics regarding the differences between cleavage and activation, and the conclusions that can (and can not) be drawn from KIE (C-H *versus* C-D) measurements in catalytic reactions would be very helpful to the field.

* Also, to add to the quagmire: is the use of the term ‘C-H bond activation’ (by the general community) not counter-intuitive? surely a C-H bond that is activated remains unbroken? Do we not want to deactivate the bond? The complex might become activated by a specific interaction or event, as a precursor to effecting the C-H cleavage; this might even involve the C-H bond itself – as in your paper.

Stuart Macgregor responded: Thank you for your kind comments. I agree about the distinction between C-H activation and C-H cleavage. C-H activation is

often taken to mean a process that may possibly be comprised of several steps that result in (for our systems) a cyclometallated ligand and substitution of acetic acid by some other ligand. Within that overall C–H activation process, one of the steps will involve a C–H bond cleavage event. The interpretation of that C–H bond cleavage step (*e.g.* through KIE or H/D exchange studies) must take into account the energetics of the preceding and subsequent steps. Thus in our case ‘C–H activation’ involves substrate binding and Cl^- loss, C–H bond cleavage and HOAc/ Cl^- substitution. Thus the hard step in the ‘C–H activation’ process may involve a ligand dissociation in which case no KIE would be expected; nonetheless, ‘C–H activation’ may be rate-limiting. In our study, the C–H cleavage step was often rate-limiting, however the computed transition states showed minimal C–H bond elongation suggesting a minimal KIE; this is simply a case of a very early C–H cleavage transition state, but ‘C–H activation’ is still rate-limiting. We have seen this in other systems computationally and this has been confirmed experimentally.¹ H/D exchange also does not necessarily require the completion of the ‘C–H activation’ process; as we have shown loss of HOAc can be rate-limiting, but kinetically accessible C–H cleavage can permit H/D exchange at the initially formed cyclometallated HOAc adduct. I take the term ‘bond activation’ to mean cleavage of a bond and the formation of (in the case of ‘C–H activation’) a new M–C bond. This then represents a new ‘activated’ form of the original substrate that can now access new reactivity. The conflation of ‘bond activation’ and ‘bond cleavage’ is perhaps unhelpful, but has become the convention in the organometallic community. The formation of an agostic complex is perhaps more in line with your comments, in that the C–H bond remains intact but becomes susceptible to reaction through, for example, deprotonation.

1 A. G. Algarra, W. B. Cross, D. L. Davies, Q. Khamker, S. A. Macgregor, C. L. McMullin, K. Singh, *J. Org. Chem.*, 2014, **79**, 1954–1970.

(603:[603]603) **Vidar R. Jensen** continued the discussion of the paper: Thank you for a very interesting paper. I have a question about the end of this reaction, when you dissociate the carboxylic acid and replace it with a chloride ligand. In this part of the reaction, you did not look for transition states. I think you are correct in assuming that this reaction is not activated. Still, the geometries in Fig. 4 in the paper might indicate significant geometrical changes on dissociation of the carboxylic acid. Could these geometrical changes be taken to indicate that the dissociation is activated after all?

Stuart Macgregor replied: In the absence of steric effects, the $16e^-$ intermediates formed upon HOAc dissociation will adopt a pseudo- C_s geometry in which the cyclometallated ligand lies perpendicular to the plane of the *p*-cymene ring. This is seen in all cases with a substituent either *meta* or *para* to the site of C–H activation. Placing a bulky substituent (*e.g.* NMe_2 or Me) in the *ortho* position results in a steric clash that destabilises the $16e^-$ intermediate, and so makes HOAc loss more difficult.

(604:[604]604) **Youichi Ishii** commented: In your study, a cationic Ru complex is used as the C–H activation centre. I am wondering what change is expected, if one uses a neutral metal centre such as $[\text{CpRu}(\text{R-phpz})]$ instead of $[(p\text{-cymene})$

Ru(R-phpz)]⁺. The donation from C–H or π – π bonds to the metal should be weaker with neutral metal centres, so this could cause changes in some of the kinetic parameters.

Stuart Macgregor replied: For a neutral {CpRu} analogue, I would anticipate the loss of Cl[–] and the k²–k¹-displacement of acetate to be facilitated and hence the overall barrier to drop; indeed we have seen this in calculations comparing the C–H activation of dimethylbenzylamine at {(C₆H₆)Ru(OAc)}⁺ and {CpRu(OAc)} fragments.¹ You may be correct that the interaction of the C–H bond with a neutral metal fragment would be reduced, however as this is already fairly minimal in the present cationic systems any effect would likely be only a minor perturbation on top of the more favourable ligand dissociation steps. C–H activation is likely to be less thermodynamically favoured in the neutral {CpRu} systems.

1 A. I. Poblador-Bahamonde, PhD Thesis Heriot-Watt University 2009

[605:[605]605] **Robert Morris** said: Often when you read a lot of articles about this acetate-assisted C–H bond activation, they often draw the transition state with a CO double bond of acetate directed toward the C–H bond being ‘deprotonated’, but it’s actually the case that the CO bonds are about equal in length so there is a negative charge build-up. I think of this as like dihydrogen coordinating and becoming acidic and then deprotonating. When you look at the two transition states the OC and CO bonds are almost the same in length. They’re often drawn incorrectly.

Stuart Macgregor responded: There is indeed an analogy with the deprotonation of dihydrogen ligands; in AMLA the development of an agostic interaction polarizes the C–H bond and allows for H-bonding to the nearby acetate oxygen and these two interactions work synergistically to break the C–H bond. The balance between the agostic and H-bonding can also vary, however, as we showed in the extreme case of N–H deprotonation of a pyrrole-imine some years ago that proceeds through a similar mechanism.¹ This underpins the AMLA concept in which both the metal and the basic ligand contribute to a varying degree to facilitate the bond cleavage. In fact, in these systems we do find some asymmetry in the C–O distances associated with the k¹-acetate ligand: for the three transition states featured in Figure 7 of our paper the ‘inner’ C–O bond distances are *ca.* 1.32 Å while the outer C–O distance is *ca.* 1.25 Å. This asymmetry may reflect the early geometry of these transition states, in which minimal proton transfer has occurred. As ever, one is on a continuum, but in this system the drawing of a double bond to the pendant oxygen does appear to be appropriate.

1 D. L. Davies, S. M. A. Donald, O. Al-Duaij, J. Fawcett, C. Little and S. A. Macgregor, *Organometallics*, 2006, 25, 5976–5978.

[606:[606]606] **Jeremy Harvey** remarked: I enjoyed your paper, and had a comment and a question.

Comment: The calculated results in Fig 5 in your paper (correlation of computational ΔG and ΔG^{++} with Hammett parameters) seem to match very well with the experimental results in Figs. 2 and 3. Certainly, if my mental arithmetic is

correct, the slopes of the two plots seem to be in very good agreement, and my impression is that the absolute rates also match quite well. Maybe you could plot them together on a combined plot?

The question relates to the barriers for losing acetic acid. You say in the text that you estimate these barriers from the typical rate constant for diffusion-controlled reactions, but it is not clear to me if the results in Fig. 6 are free energies for the resulting estimated barriers, or for the resulting fragments formed. If they are the latter, will the correlations you report in Fig. 5 still be as excellent?

Stuart Macgregor responded: The plots are based on the free energies of the C–H cleavage transition states, **TS(E-F)**, which are always the highest-lying stationary points computed in the profiles for the formation of the *para*- and *meta*-substituted products. While the $16e^-$ intermediate, **F1**, does always lie to lower free energy, in many cases the assumption of a diffusion-controlled ligand addition to these species would place the resultant transition state above **TS(E-F)**. We have not attempted to incorporate this into our model, mainly for the reasons described in the paper –in particular a purely dissociative process may not be realistic. However, if the free energy of **F1** is used as a proxy for the ligand substitution transition state then the correlation in the resultant Hammett plot is much poorer and the previous good agreement between the experimental and computational Hammett plots is lost. This emphasizes the importance of factoring the ligand substitution process in, while highlighting our inability to do this in a reliable fashion with conventional static calculations.

(607:[607]607) **Odile Eisenstein** queried: It is very interesting to see that ΔG of reaction and ΔG of activation vary in opposite directions with the substituent R as illustrated by the opposite slopes in figures 2 and 3 in the paper. Can you comment on these differences?

Stuart Macgregor answered: The thermodynamic trend does suggest a parallel to your work with Robin Perutz on the C–H activation of fluoroarenes, where stronger M–Ar^F bonds are formed with more highly fluorinated substrates (and in particular when F is *ortho* to the site of C–H activation).¹ In other work we have found very little variation in the computed homolytic bond strengths of the *ortho*-C–H bonds in a range of substituted benzoic acids.² This suggests the energetics of the newly formed Ru–aryl bond dominates and the observed and computed thermodynamic trends are therefore explained if this bond is strengthened by electron withdrawing groups. Variation in the new Ru–N bonds formed could also play a role, but as these should be weakened by electron-withdrawing groups (contrary to the thermodynamic trend) it does suggest variation in the Ru–aryl bond strength is the key factor. The dominant factor determining the overall barriers to C–H activation is the ease of dissociation of both the Cl[–] ligand (**D** to **E**) and the subsequent k^2 - k^1 -displacement of acetate that creates a vacant site at the metal centre for C–H bond cleavage to occur. This will be promoted by having a more electron donating substrate and so gives the opposite trend to the thermodynamics. The degree of C–H bond elongation is minimal in the rate-limiting transition states and so any variation in C–H bond strengths will not be important.

- 1 O. Eisenstein, J. Milani and R. N. Perutz, *Chem. Rev.*, 2017, **117**, 8710–8753
2 M. Simonetti, R. Kuniyil, S. A. Macgregor and I. Larrosa, *J. Am. Chem. Soc.*, 2018, **140**, 11836–11847

(608:[608]608) **David Davies** remarked: I have another comment on terminology – these reactions could be described as having electrophilic character, as they react faster with electron donating groups. However, we've learnt the hard way that if you describe something as electrophilic, lots of people assume the mechanism must be electrophilic aromatic substitution, which it's not in this case. We need to be careful with terminology, as you may know what you mean but other people may interpret this in a different way. There is quite a lot of confusion in the literature on this point.

(609:[609]609) **Tom Young** opened a general discussion of the paper by Evert Jan Meijer: Do you include any counter-ions/anions in your simulation box? Do you expect their inclusion to affect your results?

Evert Jan Meijer replied: Presently, we have not included counter-ions/anions in our setup. This is an important point, and including those would render our setup more realistic. On the other hand, when assuming the ions to be at low concentration, we could interpret the present results as those for the situations where counter-ions are at larger distances from the cations. Under such a circumstance, the present setup is a realistic description.

(610:[610]610) **Joseph Mwansa** asked: Is there any difference if you look at other cations, merely due to the difference in cationic radius between lithium and other cations?

Evert Jan Meijer answered: We have considered the effect of the presence Li^+ and K^+ . For the re-orientation step we have not seen a significant effect. The results regarding the hydride transfer are not yet conclusive, and we plan to explore the de-coordination of the cation when tracing the reaction pathway. If de-coordination of the cation would be an important part of the reaction mechanism, one might expect a (substantial) variation among different cations, depending on the strength of the coordination. The latter is correlated to the cationic radius.

(611:[611]611) **John Slattery** remarked: Fascinating paper. Clearly in these fully QM simulations it's not possible to access timescales long enough to see complete sampling of configurational space. How do you decide where to put your metal ions in the starting configurations and how do you know that you've considered enough possibilities to properly model the system?

Evert Jan Meijer responded: We consider specifically the effect on the process when the cation is nearby. Here we assume that this stage and the stage where the cation diffuses from/to the reaction region can decouple. Under this assumption, the stage of diffusion could be addressed in a separate study, possibly using a model that is based on empirical force fields.

(612:[612]612) **Natalie Fey** asked: For these calculations, is there a risk of biasing the reaction pathways by the constraints used, for example towards what you have observed as the non-solvated pathway?

Evert Jan Meijer replied: Indeed, when tracing a reaction pathway using biased (enhanced) sampling methods, one always has to verify if the employed biased sampling methods mimic 'true' representative pathways in an accurate manner. Several 'easy' checks can be done, including a smooth change in observed structural changes or measured forces (when using constraint methods), and verifying if tracing the reverse pathway matches the results of the forward pathway.

(613:[613]613) **Jeremy Harvey** said: I had a question about the dissociation of the CO₂ in the presence of lithium, leading to the product structure shown in Fig. 5b in the paper. You showed this intriguing structure where you have a Ru–C bond. You also said that your potential of mean force isn't quite converged there (Fig. 5a). Could you speculate on whether you think this intermediate is a genuine one, or if it comes from an artefact due to incomplete convergence of the dynamics simulation (your simulation windows are each a few picoseconds long)? This step is a really tough one for constrained MD, as it involves an anion site coordinating lithium which upon dissociation forms a fairly non-polar carbon dioxide molecule. That probably requires increased solvation by the methanol solvent. From experiments in water, the timescale here for ligand exchange at lithium ions is known to be nanoseconds¹ so if you're doing 3 picosecond windows it might be difficult to capture such solvation rearrangement, especially if you are not driving the solvation coordinate. That being said, something like the structure you show might be a genuine intermediate and that would be very interesting!

¹ L. Helm and A. E. Merbach, *Chem. Rev.* 2005, **105**, 1923

Evert Jan Meijer responded: The Ru–C bond appeared as a structure along the hydride-transfer pathway that we explored. This is not a stable structure as the associated free energy profile (Fig 5a) shows. Our present understanding is that it demonstrates that the mechanism for conversion of the formate to CO₂ involves not only a direct hydride transfer, but should include a de-coordination of the cation (Li⁺). To trace such a pathway, we have to revise our biased (enhanced) sampling scheme in order to account for this part of the mechanism. This might indeed be a process that is relatively rare, occurring on a longer timescale, but could well be addressed by using a biased sampling scheme with a suitable reaction coordinate for ion dissociation.

(614:[614]614) **Odile Eisenstein** continued the discussion of the paper: These reactions describe the hydride transfers from formic acid (or formate depending on the conditions) to a Ru (PNP) complex. It was mentioned that the transition state search is not yet finalized in all cases. It would be interesting to know if the CO₂ is already almost neutral (and thus the electrons have been on the hydride before the TS or at the TS) or if it is still anionic at the transition state. It seems that the latter was true because the CO₂ part was significantly bent during the

presentation. Is there a significant difference between the case without the cation and with the cation? If the CO₂ part is still strongly bent, the hydrogen is not yet hydridic. Since no graphical information on the TS is provided for the moment in the manuscript, it would be nice to develop this point.

Evert Jan Meijer replied: Assuming that the question refers to the system with a cation (Li⁺) present and being coordinated to the formate: in our simulation of transferring a hydride towards the Ru complex the analysis of the electronic structure indicates that the CO₂ has not yet been formed and is not neutral, as has been illustrated in the manuscript, using Wannier-center analysis (Fig. 6b). This is in contrast to the system without a cation present.

(615:[615]615) **Ulrich Hintermair** remarked: To move away from the discussion of computational details and come back to the mechanistic implications of your study, I assume you were inspired to look into the re-orientation step of the O-bound formate by the fact that most people draw it this way in their catalytic cycles. However, you could imagine an alternative situation where formic acid transfers the proton and the hydride to the N and the Ru atom respectively in a concerted fashion through a 6-membered cyclic transition state akin to the formation of the same hydride complex from iso-propanol. This would bypass the formation of any O-bound formate complexes within the cycle, and simply park them as off-cycle species. You would still expect to see the experimentally observed cation effects by Lewis acid coordination to the exocyclic O atom, but without having to go through the rearrangement in order to make product. Have you (or anyone else) investigated this possibility? We ourselves are guilty of following this convention as well, but it seems a bit arbitrary to me that with iso-propanol the O-bound alkoxides are regarded as off-cycle species whereas with formic acid everyone draws the O-bound formates in-cycle.

Evert Jan Meijer replied: This is an interesting point. We have not specifically addressed a concerted proton/hydride transfer for the present system, as we assumed the catalyst nitrogen to be protonated. In an earlier study we demonstrated this to be the stable state, even under very basic conditions.¹ However, in general a concerted hydride/proton transfer pathway should be considered as a possible mechanism. We anticipate that the viability of such a pathway is controlled by the pK_a of the N-H moiety and the pH of the solvent. In this context, the role of O-bound species (formate/alkoxide) is indeed an interesting and important question. The O-bound formate species has also been isolated experimentally by Beller and co-workers.²

1 V. Sinha, N. Govindarajan, B. de Bruin, E. J. Meijer, *ACS Catal.*, 2018, **8**, 6908–6913

2 E. Alberico, A. J. J. Lennox, L. K. Vogt, H. Jiao, W. Baumann, H.-J. Drexler, M. Nielsen, A. Spannenberg, M. P. Checinski, H. Junge, M. Beller, *J. Am. Chem. Soc.*, 2016, **138**, 14890–14904

(616:[616]616) **Robin Perutz** opened the discussion of the paper by Daniel Ess: Your results are really nice. I am not surprised that some of the molecules skip the intermediates. Being naïve, my picture of some reactions we've studied looks like your first diagram of the trolley going down the slope and encountering a mogul, the only difference was that I imagined I was on skis! Continuing to be naïve

perhaps, the way to test this, I believe, is to look at a series of reactions with different exothermicities or exoergonicities. For example, if you change to the reaction with dihydrogen instead of methane, it will release considerably more energy and I'd predict that a larger proportion of molecules 'skip the bumps'. Have you tried this? With respect to experiments, when we did laser flash experiments on ruthenium dihydrides and looked at degenerate reactions where you photodissociate H_2 and allow it to react back thermally, we found that some of these reactions occurred at diffusion controlled rates. I was often asked whether there was a dihydrogen complex as an intermediate on the way to the final dihydride product. We never saw one but I imagined that this was a very exothermic reaction with a slope like you described. If there were any bumps on the way, the molecules would ride straight over them with no detectable dihydrogen intermediate.

Daniel Ess responded: Your great ski analogy raises the general question of how the shape of the potential energy surface impacts reaction dynamics, especially non-statistical intermediate skipping. In our experience for organometallic reactions, which is still limited, we have found that dynamical effects can be found in reactions with very shallow well intermediates or highly exothermic reactions. Both of these types of cases have been reported for organic reactions. While we have not reported on this, we have looked at substrates other than methane. As you suspected, the potential energy surface changes and so do the dynamical effects, or lack of dynamical effects.

(617:[617]617) **Aaron Odom** continued the discussion: You talked about carrying energy from one excited state into the ground state of an intermediate and how that can lead to intermediate 'skipping' of a sort. As you go over the transition state, you have certain modes in the molecule that are 'hot'. As you reach the ground state for the next intermediate, different modes likely need to be thermally excited to go up to the next transition state for that intermediate to be skipped. How long does the relaxation of 'hot' modes from the transition state need before the energy is distributed throughout all the normal modes? Is that what sets your time frame for how often you need to sample the energy in your calculations, which I think you did on the order of a quarter of a femtosecond up to a few femtoseconds?

Daniel Ess replied: For organometallic reactions, the timing of intramolecular vibrational relaxation is not well known. However, some experimental estimates suggest that this relaxation is between 2 to 8 picoseconds. This is why our direct dynamics are typically propagated between 1000 to 2000 femtoseconds, which is less than the expected relaxation time.

(618:[618]618) **Tom Young** asked: How classical are your quasiclassical trajectories? For instance, do you include quantum tunnelling?

Daniel Ess responded: From a propagation standpoint, these trajectories are classical. They are only quasiclassical because they include zero-point energy. They do not include tunnelling.

(619:[619]619) **Jeremy Harvey** said: I really enjoyed your paper. In the discussion with Prof. MacGregor, about C–H activation steps, there has been a discussion of the well-known distinction between synchronous and asynchronous reaction steps. Your work raises another important question, that of the distinction between concerted and stepwise reaction steps. In fact, when doing trajectories, in a certain sense this distinction disappears because the timing of reaction events becomes somewhat arbitrary. Your results such as those in Fig. 1 in the paper are a nice example of this. Your are interpreting ‘prompt’ barrier crossing of TS2 (when the blue line heads steeply upwards at short times after crossing TS1) as evidence of ballistic behaviour. However, when looking at trajectories, it is not the case that all reactions occur at the same reaction time. Instead, it is a sort of Poisson process with the probability of barrier crossing of TS2 varying as the total population of ‘surviving’ un-reacted trajectories. So one expects some upward-heading blue lines at short times even if the reaction is stepwise. The question is, how many? Have you considered comparing the number of such trajectories you observe with the expected number at a given time given the estimated rate constant? My suspicion is that such an analysis will lead to a diagnosis of a reaction at the borderline of stepwise behaviour, with a slight excess of ‘early’ barrier crossing, rather than pure ballistic behaviour. We discussed this kind of issue in our paper, *Alkene Hydroboration: Hot Intermediates That React While They Are Cooling*, using a statistical model.¹

1 D. R. Glowacki, C. H. Liang, S. P. Marsden, J. N. Harvey, M. J. Pilling, *J. Am. Chem. Soc.*, 2010, 132, 13621

Daniel Ess replied: Indeed, our work, and your great question/comment, raise the issue for organometallic reactions of the distinction between reactions that are concerted and reactions that are stepwise with an intervening intermediate. IUPAC defines an intermediate as a molecule with a lifetime significantly longer than a molecular vibration, and this often related to having an energy well deeper than RT. To us, this definition is somewhat vague because there is no precise definition of ‘significantly longer’. Does this mean significantly longer to observe by experiment or significantly longer to allow for complete intramolecular vibrational relaxation prior to going over another barrier. The latter scenario with complete relaxation is potentially an arbitrary, but perhaps useful, definition of an intermediate in a statistical framework. This is because, to some extent, the idea of concerted and stepwise mechanisms is based on the assumption of transition-state and other statistical theories. I like your suggestion of looking into possible Poisson processes. We have not done this, and we need to. You may indeed be correct that we will find that for this reaction, it is at the borderline of stepwise behavior.

(620:[620]620) **Alison Hulme** asked: Earlier this summer I heard Suzanne Blum lecture on using single molecule TIRF microscopy to probe the ‘sub-ensemble’ kinetics of metathesis polymerisation reactions.¹ We’ve heard a lot of statements in this discussion meeting such as ‘this is the reaction pathway’, or ‘this is the transition state’. Does molecular dynamics provide a way to explore, computationally, multiple reaction pathways and hence different kinetic profiles for an organometallic process?

1 Q. T. Easter, S. A. Blum, *Acc. Chem. Res.*, 2019, 52, 2244–2255

Daniel Ess answered: Yes, direct dynamics simulations allow us to examine reaction pathway connections beyond those outlined by a static DFT potential energy surface.

(621:[621]621) **Mike George** asked: Considering generally the possible difference between photochemical and thermal initiation, and the difference between stepwise and ballistic, I wonder if it would be possible to affect ligand loss under infrared irradiation to change the pathways or trigger the stepwise process and whether there is there any insight to be gained there?

Daniel Ess responded: It is reasonable to propose that under some conditions that introduction of vibrational excitation can result in changing the reaction mechanism.

(622:[622]622) **Robin Perutz** commented: I'd like to come back as I think we've got a major misunderstanding here. In some photochemical reactions in solution, the reaction intermediate reacts with the solvent. These can take place within a picosecond and are affected by the excess energy provided photochemically. Other reactions take place between the intermediate and a solute in solution. The latter class of reactions cannot take place faster than diffusion control. The diffusion controlled rate constant is *ca.* $10^{10} \text{ M}^{-1}\text{s}^{-1}$. If the concentration is 1 M, which is much higher than normal, the lifetime of the intermediate is *ca.* 10^{-10} s . By that time the reaction intermediate is fully relaxed and equilibrated. Reactions with dissolved hydrogen are slower still because of the low solubility of hydrogen. The femtosecond timescales for molecular dynamics here refer to starting with the transition state and watching its reaction. That's completely different to the nanosecond or millisecond timescales of experiments using laser flash photolysis with a solute reacting in solution. Those are relaxed and occur on a longer timescale.

Daniel Ess replied: I agree. All the reactions we have modelled with direct dynamics are thermal reactions. While we have not reported on laser flash initiated reactions, it is in principle possible to use this type of direct dynamics to examine flash photolysis reactions assuming, as stated in the comment, that the intermediates are fully relaxed and equilibrated.

(623:[623]623) **Mike George** remarked: Framing my previous question in a different way, I wonder if infrared excitation could be used to change the mechanism as I was talking about more generic photochemistry and wonder whether infrared excitation could affect stepwise and ballistic pathways.

(624:[624]624) **Jennifer Love** opened a general discussion: To the comments from Guy and Dai about defining 'activation', what I find useful (regarding Dai's point on electrophilic reactions being misconstrued) is to go back to the classic paper by Labinger and Bercaw published in *Nature*.¹ Here they describe a number of different mechanisms of C–H activation. I find looking at the mechanisms outlined in this paper instructive.

1 J. A. Labinger, J. E. Bercaw, *Nature*, 2002, **417**, 507–514

Guy Lloyd-Jones answered: Thank you Jennifer for the valuable comment and suggestion. The Labinger and Bercaw paper seems like an excellent starting point for definition and direction.

(625:[625]625) **Natalie Fey** addressed Evert Jan Meijer, Daniel Ess and Stuart Macgregor :What are the computational costs of the projects you have described in your contributions? There are experimentalists in the room who might think that all calculations are instantaneous, or at least quite quick. From my experience, the studies you describe have probably had a high computational cost (CPU hours/resources), so perhaps you could give us an indication of the costs involved, along with an assessment of whether this was a good investment compared to what you learned, and in comparison with experimental data?

Evert Jan Meijer answered: The computational costs are significant, but feasible on a medium-scale high-performing computing (HPC) facility, which are typically available at large scale research facilities. The timescale to determine a single reaction free energy is days of continuous use of several tens of computing units on a HPC facility.

As present day HPCs have thousands of computing units, this is a relatively moderate usage. It is therefore a good investment of resources, as novel scientific insight and the potential to contribute to designing novel chemical systems is substantial.

Daniel Ess replied: Sometimes we run direct dynamics on a reaction and we do not learn anything more than what the static DFT surface revealed. In other reactions, we run trajectories and find non-statistical effects or non-IRC motion that is very interesting and well worth the computational expense. These types of direct dynamics simulations require propagating many reactive trajectories, often between 100–200, but sometimes more. Typically, for a single trajectory we run between 1000–1500 steps in both forward and reverse directions, which require at least a force calculation at each step. In addition to running many trajectories, there is also the need to check multiple DFT methods and basis sets. This often results in using millions of CPU hours for examining a single reaction.

Stuart Macgregor replied: Clearly static DFT calculations of the type in our paper require significantly less computational resource than the dynamics studies described by my colleagues. However, a thorough mechanistic study that combines both computation and experiment generally takes a significant time to work through. In my experience, it is very rare that computation aligns with experiment on the first pass, and this may be due to issues concerning computational methodology, or (more likely) the nature of the chemical model adopted in the calculations, or it may reflect a need to reinterpret or reassess the experimental data or to perform additional experimental studies. Both aspects of the project gain from the input of the other, and ideally should run in parallel (and allowing for sufficient time) to maximise this dialogue.

(626:[626]626) **Odile Eisenstein** commented: My comments derive from the talks of Stuart Macgregor and Dan Ess. There was indeed a proposal that there was a continuum between oxidative addition and sigma-bond metathesis. This is

certainly an interesting proposal. But most of the recent results show that assistance by some other chemical function (like in AMLA/CMD) is actually very important and frequent. We have probably not yet discovered all the groups that can assist a C–H activation. The nomenclature will have to include the origin of the chemical group that assists or does the reaction. We may then end up having a better view of this continuum.

Daniel Ess answered: In 2010, our group, and the Periana group, proposed classifying C–H activation reactions along the continuum of electrophilic, ambiphilic, and nucleophilic.¹ These terms refer to how the metal center and ligand covalently activate/cleave a methane C–H bond by covalent donor–acceptor orbital interactions.

1 D. H. Ess, W. A. Goddard, R. A. Periana, *Organometallics*, 2010, **29**, 6459–6472

(700:[700]700) **Simone Gallarati** opened a general discussion of the paper by Jeremy N. Harvey: I found your approach for inclusion of a solid base (sodium carbonate) in a computational thermodynamic scheme based on molecular quantum chemical computations very interesting. Could you highlight the main advantages and limitations of this approach? In the paper you discuss some errors that could arise from it. Could you put this into perspective with respect to the errors that the general DFT approach in the predicted relative Gibbs Free energies can result in (5 kcal mol^{−1})? How does the thermodynamic cycle approach for solid species compare to that? Would you say that the error you get with the thermodynamic cycle approach is comparable to the error you get with this level of DFT computations?

Jeremy Harvey answered: The method itself is described in the paper and in some other replies. Its main advantage is that it allows you to estimate the free energy change for a reaction involving a solid, without having to perform quantum chemical computations on a periodic system, so it can be combined with quantum chemical studies of molecular systems. It can be done provided accurate experimental thermodynamic data is available for that solid, and indeed many simple solids have well-known and accurate Gibbs energies. Its main disadvantage is that it provides no information at all concerning the kinetics of such steps, only on their thermodynamics. This is because the approach does not in any way model the intermediate species formed at the interface between the solid and the reactive phase. In terms of accuracy, the scheme is limited in part by the experimental data. Another issue is that the thermodynamic cycle one constructs needs to contain one or several gas-phase species that have the same overall atomic composition as the solid, and for these species one needs to perform quantum chemical calculations, as well as having experimental data available. In some cases, these species may be quite different in terms of bonding, oxidation states, *etc.*, from the other species modelled. In the present case, we wanted to handle sodium carbonate. We had experimental data for the solid (in the form of its monohydrate), but not for gaseous Na₂CO₃. So instead we used (section 3.1 in the paper) atomic sodium, CO₂, dioxygen, and water molecules as the gas phase systems. An Na atom is obviously rather different than the sodium in solid sodium carbonate in terms of electronic structure. Hence we are a bit

dependent on our chosen quantum chemical method accurately reproducing the ionisation energy of Na, and of similar redox steps. This issue can be side-stepped by carrying out accurate (*e.g.* CCSD(T) with large basis set) calculations for the experimental gas-phase reference species such as Na, and also for more chemically relevant model compounds such as gaseous Na₂CO₃, and using a more complex thermodynamic cycle. We did not do this, not because the calculations are very demanding (the species involved are small), but due to lack of time. This leads to a slightly higher error than necessary in our study. In my view, when done properly, the present approach leads to lower intrinsic errors than does the use of typical DFT methods (maybe 2-3 kcal mol⁻¹?), and in the present case, our estimated error bar is probably a bit larger than that.

One caveat is that error bars for theoretical procedures are often hard to derive – in one sense most quantum chemical computational protocols have an error of zero (repeat the calculation and you get almost exactly the same result), but there is a systematic error for any given species that can be very large. This caveat definitely applies in the present case! It would be good to get statistics on the accuracy of this procedure for a representative set of species, but this would be quite a lot of work.

(701:[701]701) **Jamie Cadge** said: Boronic esters are also of significant synthetic value in the Suzuki-Miyaura cross-coupling. Could a similar approach using these state of the art approaches be used for these more mechanistically challenging substrates as well?

Jeremy Harvey replied: Many boronic esters and indeed other aryl-boron species are indeed used in Suzuki-Miyaura cross-coupling reactions. In some cases, it has been shown (in part through work by Guy Lloyd-Jones and his group – *e.g.* MIDA boronates,¹ but also fluoroboronates, as in our work²) that these other species convert to boronic acids prior to the actual cross-coupling. But in other cases, the species such as boronic esters can react themselves (this is something that Denmark's studies have shown). In principle, we could extend our present work to describe that kind of reactivity, but we haven't done it yet.

1 J. A. Gonzalez, O. M. Ogbu, G. F. Morehouse, N. Rosson, K. N. Houk, A. G. Leach, P. H-Y. Cheong, M. D. Burke, G. C. Lloyd-Jones, *Nature Chemistry* 2016, **8**, 1067

2 M. Butters, J. N. Harvey, J. Jover, A. J. J. Lennox, G. C. Lloyd-Jones, P. M. Murray, *Angew. Chem. Int. Ed.*, 2010, **49**, 5156

(702:[702]702) **Odile Eisenstein** continued the discussion of the paper: Can you describe in more detail how you treated Na₂CO₃ in the solid state? This is not obvious to the general audience. I know that concentration is treated *via* micro-kinetics but I think that this should be also mentioned.

Jeremy Harvey answered: A reaction such as HX + Na₂CO₃ = NaHCO₃ + NaX is hard to describe with molecular quantum chemistry. A full description (energy + kinetics) would require solid-state DFT, a slab model of the surface, and so on. This is not trivial. Our idea here was instead just to calculate the thermodynamics by using experimental reference data for the solids (many simple solids are well characterized) and for corresponding gas-phase fragments. We plan to write up the method in more detail

(703:[703]703) **Jason Lynam** remarked: In the literature, I believe there are many computational papers where they've just used bicarbonate or carbonate anions as the base, and used that as part of the model. What's your feeling on the accuracy of this compared to those, and how accurate it might be? Is the problem the lack of the cation or the inadequate treatment of solvation of those species?

Jeremy Harvey replied: This was one of the motivations of this study; using 'naked' carbonate is not a good choice, if the experimentally-used base is sodium carbonate. In the past, people used to use naked carbonate, or naked protons, or naked hydroxide, to model individual base-induced or acid-induced reactions. This often led to unrealistic energies. Now, if the base (or acid) is present in solution, then in principle a good model of the solvation would be enough to solve the problem, so continuum models would work. However, continuum models don't perform as well for species with a high charge-to-radius ratio, so explicit microsolvation might be needed. If one instead thinks the reaction involves deprotonation at the solid sodium carbonate surface, then a continuum will be less suitable. So I think the 'problem' depends on what the system is.

(704:[704]704) **Guy Lloyd-Jones** commented: Overall you have the challenging but highly rewarding goal to construct an accurate holistic computation model for the catalytic system – how will you go about obtaining the experimental rate data to correlate with each of the different steps? The flip side of this is that in a practical sense, if you calculate what you think the turnover frequency will be, and it does not correlate with the measurements, how do you know which part of the model to adjust? One of the frustrating aspects in the kinetic and computational studies of this reaction, is that there is no single set of 'typical' conditions for a Suzuki-Miyaura reaction: conditions range from *e.g.* single phase protic media at high pH, to biphasic media with a solid base. Indeed, the ability to tune the conditions to the specific R–B and R'–X coupling partners to attain the best process efficiency, is one of the reasons for the very wide-spread application of the reaction.

Jeremy Harvey replied: I will comment first on the issue of 'typical' conditions for the Suzuki-Miyaura reaction. Despite knowing about your study with Alastair Lennox where you reviewed many thousands of experimental reports of Suzuki-Miyaura couplings (reference 28 in Ref. 1), and despite knowing that the diversity of conditions used was very high, I did once suggest to my co-worker Tolga Yaman that a helpful starting point in this project would be to identify one 'typical' experiment as a reference point for our model. It rapidly became clear that such identification would be very difficult, so we instead settled for trying to account for observed reactivity in the Denmark studies.

1 A. J. J. Lennox and G. C. Lloyd-Jones, *J. Am. Chem. Soc.*, 2012, **134**, 7431

Next, about our goal and how we plan to set about it. In principle, computations can be used to predict rate constants as a function of conditions and substituents with enough accuracy to be of some aid in designing new experiments. We had some success along these lines (albeit only for one substrate, by varying *T* and starting material concentrations) in Ref. 2 in the paper. It is fair to say that we are not yet at that level of success for the SM coupling yet, although we plan to

continue our work. Crucial for this effort will indeed be to have reliable experimental data to test our predictions. This is why the mechanistic work by Denmark is so valuable. Some other parts of the system are also quite well known (e.g. oxidative addition to Pd(0) is a reaction where experimental kinetics are available for different ligands). I think at present that computational inaccuracy is a bigger concern than lack of reference experimental data, because even quite approximate experimental procedures (reaction time, amounts of reactants, temperature, solvent) provide some constraints on possible models.

(705:[705]705) **Daniel Ess** asked: On your diagram, with what you've mapped out, which intermediate(s) is the resting state(s) and which transition state(s) control the turnover rate? Do the resting states and turnover limiting transition states fit with the general rate law?

Jeremy Harvey answered: In our rough figure 3 in the paper, the resting state is the Pd(II) adduct ArPdL_nX , and the X to O exchange TS or transmetallation TS is turnover-limiting. However, the barrier to oxidative addition is not all that much lower, so there is the scope for the Pd(0) species to be the resting state and the oxidative addition TS to be turnover-limiting. I'm not aware that there are general rate laws available for catalysis from experiment. The precursors to product formation in Figs. 1 and 2 in the paper, and the TSs, match the observed rate laws in Denmark *et al.*'s studies.

(706:[706]706) **Odile Eisenstein** remarked: There are now publications on how to carry out a proper computational study and hopefully they would be highly useful to the chemical communities (experimental and computational together). I would like to recommend the reading of a number of papers.^{1,2}

1 L. Perrin, K. J. T. Carr, D. McKay, C. L. McMullin, S. A. Macgregor and O. Eisenstein, *Computational Studies in Organometallic Chemistry*, ed. S. A. Macgregor and O. Eisenstein, Springer International Publishing, Cham, 2016, pp. 1–37.

2 J. N. Harvey, F. Himo, F. Maseras, L. Perrin, *ACS Catalysis*, 2019, **9**, 6803–6813.

Most practitioners and most of the readers are aware of the importance of introducing the real chemical systems in the computation. Many of us think that it concerns mostly the nature of the coordinated ligand. This is indeed important and a nice example was developed in Ref. 3. This work describes a situation where the mechanism cannot be modelled unless all substituents on all partners are considered. Probably an rather unusual, extreme situation. However, this is probably not the most difficult situation to model. Jeremy, what is your opinion on the most difficult situations to model? Do you think we are at a level where we can believe the trends and even the results.

3 L. Perrin, E. L. Werkema, O. Eisenstein, R. A. Andersen, *Inorg. Chem.*, 2014, **53**, 6361–6373

Jeremy Harvey replied: There are many things that are difficult to model. As you say, conformationally complex systems pose real challenges. In my view though, one of the hardest things is polar species in polar solvents. These are some of the things we mentioned in the perspective you cite – and which are also present in the book chapter and the paper you mention. Yes, though – I do think

1 that computation yields many good results. It is just that it would be nice if we do
the good things and the less good things better.

5 (707:[707]707) **Todd Marder** asked: You're working entirely with a mono ligand
and monophosphine system, right? So there are still (although, I believe you're
probably right in most cases) questions about whether the active catalyst is
a mono- or bisphosphine complex, and the exact nature of the phosphine. There
are also chelating ones which I think are probably bidentate coordinated
10 throughout, like DPPF which is widely used. This adds complexity to the situa-
tion. You have B(OR)₂ here. The pinacol esters are among the most widely used so
the picture of the transmetallation may vary a bit here. Transmetallation is usually
the fast step. I don't think the esters are converted to boronic acid under the
reaction conditions. It all depends whether there's water present, what the pH is,
15 and so on. Guy Lloyd-Jones has looked at related things in detail.

Jeremy Harvey replied: All of the issues you mention are indeed ones that we
will need to take into account in due course as we continue to develop our model.
About complexity, you are certainly right! I once asked the student working on this
20 project, Tolga Yaman, to identify a 'standard' set of Suzuki-Miyaura reaction
conditions (ligand, base, solvent, temperature, nucleophile, electrophile) so we
could focus on that. It rapidly became clear that there is no 'standard' case...

Briefly, on each of the topics you mention:

25 *We have only considered monophosphines so far, yes (although in other work
on oxidative addition, we have considered chelating diphosphines).

*We consider both 1 : 1 Pd:phosphine species and 1 : 2 species (e.g. species 8
and 3 in Scheme 2 in the paper, respectively).

*Yes DPPF and related phosphines will be well worth looking at!

30 *We have modelled the transmetallation considering only ArB(OH)₂ (and its
conjugate base) as reaction partners. Pinacol boronates, mida boronates, tri-
fluoroborates, and so on are indeed relevant. My impression is that many of these
convert to boronic acids prior to transmetallation (I was involved in some of the
work that Guy Lloyd-Jones has done on this) but I take your point that this does
35 not always happen, depending on conditions. My hope for this project is that we
can work out an understanding of when this does and does not happen, based on
our work.

40 (708:[708]708) **Derek Durand** opened a general discussion of the paper by
Markus Reiher: How does your method of using electrostatic potential to predict
likely protonation sites cope when there are two sites with an equal likelihood of
being protonated?

45 **Markus Reiher** responded: In the current setting, the algorithm would
protonate each of the sites individually and optimize their structure. Afterwards,
the electrostatic potential would be evaluated again for the products and a new
round of protonation steps starts (some following a preceding reduction step that
will increase the likelihood of further protonation). You might argue that there is
then a chance of not finding the doubly protonated product that one would have
50 obtained by simultaneously introducing two protons at the two sites in one step.
However, this can be easily allowed for in the algorithm. In general, we advocate

for a conservative scheme that introduces protonation for all reasonable cases and let the algorithm figure out automatically after structure optimization what the unique species obtained are.

(709:[709]709) **Robert Morris** commented: The electrostatic descriptor is great idea. Often in protonation reactions you have all these kinetic resting states before you get to the final protonation state. Linking the inside to the outside is important, as often metal complexes and metalloenzymes like nitrogenase, hydrogenase and cytochrome-C oxidase get protonated on the outside at a ligand first, then go to the metal through a series of proton hops, ending up in the middle of the complex or active site eventually. For examples of metal complexes reacting with protons *via* a series of hops, see the review by Henderson,¹ articles by Schrock and coworkers (and references therein)² and Bullock and coworkers.³ I think it's quite important, what you're doing.

1 R. A. Henderson, *Angew. Chem. Int. Ed.*, 1996, **35**, 947–967

2 R. A. Kinney, R. L. McNaughton, J. M. Chin, R. R. Schrock, B. M. Hoffman, *Inorg. Chem.*, 2011, **50**, 418–420

3 N. Kumar, J. M. Darmon, C. J. Weiss, M. L. Helm, S. Raugei, R. M. Bullock, *Organometallics*, 2019, **38**, 1391–1396

Markus Reiher replied: Well, a main driver of our algorithmic developments was in fact our work on the hydrogenase mechanism^{1,2} and on that of Schrock dinitrogen fixation,³ which was truly painful as all possible protonation and reduction products, spin and charge states had to be considered manually. In fact, the first presentation of our algorithm⁴ was for the Schrock dinitrogen fixation complex, for which we could map out a very detailed mechanism of protonation and reduction steps involving thousands of intermediates and transition states.

1 A. R. Finkelmann, M. T. Stiebritz, M. Reiher, *Chem. Sci.*, 2014, **5**, 215–221

2 A. R. Finkelmann, M. T. Stiebritz, M. Reiher, *Inorg. Chem.*, 2014, **53**, 11890–11902.

3 S. Schenk, B. Le Guennic, B. Kirchner, M. Reiher, *Inorg. Chem.*, 2008, **47**, 3634–3650.

4 M. Bergeler, G. N. Simm, J. Proppe, M. Reiher, *J. Chem. Theory Comput.*, 2015, **11**, 5712–5722.

(710:[710]710) **Laurel Schafer** asked: Can this model for protons be easily translated to counter-ion effects as well?

Markus Reiher replied: We have not thought about this type of application yet. However, it is a very interesting one. In principle, the electrostatic part can probably be captured by the electrostatic potential, but there is more due to polarization and even covalent effects, which would make a straightforward application of our model more difficult, *i.e.*, high-dimensional.

(711:[711]711) **John Slattery** remarked: I think it's very useful to have automated routines to identify Lewis basic sites on molecules, as you describe in your paper. It would be great to be able to do the reverse as well and have routines that identify potential Lewis acidic sites. Are you planning on doing this within your SCINE project? If so, is electrostatic potential the right approach for doing this? In systems that display halogen bonding, the presence of a 'σ-hole' in the

1 electrostatic potential can be useful in identifying the location and direction of
intermolecular interactions. However, I imagine there will be some limitations in
how far this could be extended to Lewis acids.

5 **Markus Reiher** responded: Naturally, the electrostatic potential could be
useful to highlight potential reactive sites that are Lewis acidic and can be
attacked by Lewis basic sites. As such, it is worth exploiting in a feature vector
for machine learning reactivity. As long as we lack data to exploit such concepts,
10 we rely on structural distortions as an exploration tool to then launch systematic
searches (*e.g.*, a potentially Brønsted acidic site could be discovered by a bond
stretch that delivers an energy pointing towards easy proton abstraction). In the
present case, the Lewis basic sites are attacked by a proton, which, from an
electronic structure point of view, is a grateful counter reactant as it is a struc-
15 tureless point charge without any electrons. With respect to other, structured
Lewis acids (carrying electrons), conclusions about reactivity cannot be expected
to be clear-cut because the interaction with the Lewis basic site is not solely
electrostatic, but will have important covalent contributions that are not
properly caught by the electrostatic picture which, however, is the basis of the
20 electrostatic potential descriptor.

(712:[712]712) **Matthias Bauer** remarked: Your approach is dealing with elec-
trostatic interaction. This can also be weaker than Coulomb interaction. Would
your method also work for weaker interactions than Coulomb ones?

25 **Markus Reiher** answered: Right, it is designed to work well for electrostatic
Coulomb interactions, hence the application to protonation reactions. It could
also be useful for cations (and anions) with an internal structure (*i.e.*, with more
than one atomic nucleus and with electrons), although other contributions (such
30 as covalent contributions) can become more important. For covalent interactions,
descriptors from conceptual DFT will be more useful (such as electronegativities,
softness, Fukui functions, the dual descriptor,...) as targets that could narrow
down the number of reactive sites. These sites would represent a first set that can
then be explicitly probed in an exploration. However, given the inherent issue of
35 being evaluated at equilibrium structures of reactants, it can be anticipated that
structural distortions will be required to activate the reactants. Regarding weak
interactions (*e.g.*, dispersion forces), it will be difficult to propose reactivity
descriptors because it is the nature of these forces that they are cooperative in
order to be strong. Hence, many contacts need to be considered at once and they
40 are often best suited for a reactant pre-arrangement only, rather than for inducing
a reaction.

(713:[713]713) **Jeremy Harvey** remarked: You mention that in one case, you had
45 a 'hidden' protonation site, which only became apparent as an accessible
minimum of the electrostatic potential upon distorting the system. Indeed, in
a lot of cases in chemistry, reactive sites only becomes available after some
'preparation' step where ligands dissociate or molecules distort. Can you describe
how you plan to find such hidden sites, and also perhaps comment on whether
50 this site was discovered in this particular case by an automated process, or did you
have to put in some chemical information to get it to distort?

Markus Reiher responded: This is a general problem of all reactivity descriptors evaluated at equilibrium structures of reactants such as electronegativities, softness, hardness, or the dual descriptor. Although we are currently considering the value of such descriptors from conceptual DFT in a machine learning approach, we hoped for higher reliability of a purely electrostatic criterion (probed by the positive point charge of a proton). In this particular case, we were surprised to find no protonation site on the respective nitrogen atom. It took various exploratory calculations, described in the paper, to understand the reasons. As a consequence, we realized that we also need to consider structural distortions to eventually find all relevant protonation sites. This however, we have not attempted in the paper because it has already been part of an ongoing, further reaching project in our group. That project attempts to exploit structural distortions to uncover reactive sites in a general sense. For instance, decomposition reactions are rather easy to predict (through bond stretches), whereas rearrangement reactions are much more difficult to foresee. For this, collective Cartesian distortions will be required. At the same time, one does not want to create an exploding number of structural distortions to obtain a good guess for a potential reaction coordinate. It is for this reason that we try to exploit information that has already become available at a node in a reaction network. Normal modes from a vibrational analysis, required for free energy calculations, can be exploited as they do not require any additional computations in the first step. In the present case, the normal mode of the amine inversion normal coordinate could be exploited to generate a distorted, *i.e.*, ‘activated’, reactant structure. Additional linear combination of low-frequency modes could be an option to restrict the search space for structure optimizations. And clearly, this is related to the problem of choosing collective variables in metadynamics simulations, which are an efficient realization of the local-elevation algorithm. Eventually, we intend to exploit such structure distortions in such a way that their exploration is significantly less demanding in terms of computer time when compared to molecular dynamics enhanced sampling approaches. It is also clear that there will be no guarantee of having found all relevant reactive sites. Still, there is no alternative to such automated search algorithms. At a much later stage, when a huge amount of data on reaction mechanisms from automated procedures has become available, one can envision machine learning approaches that can discover missing links in new networks based on existing information.

(714:[714]714) **Evert Jan Meijer** commented: Following up on your answer two questions ago – you mentioned that you tried to explore transition states by pushing away from the equilibrium, but if that didn’t work you would look at the molecular dynamics. Have I understood this correctly? You should be able to have a first transition state. Sometimes if the system is very complex, finding a possible pathway is very complex. Did you consider transition path sampling techniques?

Markus Reiher responded: Yes, in fact, we did. Our goal is to have a stable transition state search algorithm that can be driven in an automated fashion. For this we can rely on search algorithms offered in standard quantum chemistry packages, but also on our own developments, which were driven by the desire to have transition state searches that are very stable.^{1,2} In a standard exploration, we target and then optimize transition states from reactants in different

conformations. Clearly, this approach has its limits when molecular size increases and full conformational flexibility is allowed for. What we expect from our exploration algorithm is to detect automatically regions of a potential energy surface that are too rugged so that explicit enumeration of minimum energy pathways becomes pointless. It is such a situation that would divert the exploration algorithm to enhanced sampling approaches and techniques such as transition path sampling. While we have not implemented this yet, it presents no major obstacle as one could switch to established molecular dynamics software (and even integration with our software should not be too complicated). What we then require from our machinery is that information from molecular dynamics can be fed into our kinetic modelling *i.e.*, rates for some chemical transformation step obtained from a molecular dynamics protocol could be combined with those from our quantum chemistry approach to create a consistent phenomenological kinetic model of the overall process. For this purpose, it should also be possible to exploit Markov state models.

1 M. A. Heuer, A. C. Vaucher, M. P. Haag, M. Reiher, *J. Chem. Theory Comput.*, 2018, **14**, 2052–2062.

2 A. C. Vaucher, M. Reiher, *J. Chem. Theory Comput.*, 2018, **14**, 3091–3099

[715:[715]715] **Simone Gallarati** commented: The number of accessible transition state structures for even a single elementary step in an organometallic reaction is often dauntingly large, stemming from the conformational flexibility of the catalyst and substrates. Small energy differences, such as those between diastereomeric TSs, can have profound effects on computed selectivity. Could you comment on the challenges and on the importance in your specific examples of adequately mapping the conformational space? What are the best tools available to the computational chemist to approach this daunting challenge?

Markus Reiher replied: I am afraid that we need to meet this challenge. Obviously, we cannot avoid it. Our SCINE project (www.scine.ethz.ch) tries to do exactly this: Map out all transition states along all paths between conformers of reactants (see Ref. 1 for an example). Currently, we explicitly construct all conformers, and obviously, this breaks down for large flexible molecules. To somehow limit this number, an energy cut-off can be applied. More restrictions will be possible, if for instance, a reaction is local and conformational flexibility in a remote part of the molecule can be neglected. However, this does not solve the principal problem that the larger the molecule, the more discouraging is the conformational freedom, in general. In our specific case here, however, we were lucky that the conformational freedom was restricted due to the rigid structure of the complex lacking aliphatic substituents.

1 G. N. Simm, M. Reiher, *J. Chem. Theory Comput.*, 2017, **13**, 6108–6119.

Jeremy Harvey answered: Simone, this is an excellent point. This is a major challenge that should be taken into account by any researcher. There are numerous examples for which incorrect identification of the best conformer can lead to incorrect chemical conclusions (for two examples, see Ref. 1, where we argue that a previous study² only reached agreement with experiment for one property by using incorrect conformers, and the paper by Perrin, Werkema,

Eisenstein and Andersen,³ where the formidable challenge of identifying the correct conformer for a system involving two Cp'₂Ce centres, with Cp' = 1,2,4-(tBu)₃C₅H₂ was navigated).

1 A. C. Tsipis, A. G. Orpen, J. N. Harvey, *Dalton Trans.* 2005, 2849

2 L. Cavallo, *J. Am. Chem. Soc.*, 2002, **124**, 8965

3 L. Perrin, E. L. Werkema, O. Eisenstein, R. A. Andersen, *Inorg. Chem.* 2014, **53**, 6361–6373

In our study, though we do not dwell on it at too much length, we had to navigate this challenge for several TSs and intermediates. Our general approach is I think similar to that used by many other groups: first, use as much experimental data and chemical intuition as possible to try to pin down which conformers might be relevant, then manually optimize these structures and retain the lowest energy form found. The issues to do with conformational searching are particularly severe in species such as **9** with multiple metals and multiple tris(isopropylphosphine) ligands. Although we do not mention it in detail in the paper, for this species (which we nicknamed 'the Beast' due to its conformational challenges) we also used a more advanced technique whereby we parameterized an approximate molecular mechanics force field, and performed extensive conformational searching at that level. Then, lower-energy structures were refined at the DFT level. This procedure is one we used in the past successfully for some related problems⁴ and which many other groups have used also.

4 M. Burns, S. Essafi, J. R. Bame, S. P. Bull, M. P. Webster, S. Balieu, J. W. Dale, C. P. Butts, J. N. Harvey, V. K. Aggarwal, *Nature*, 2014, **513**, 183

Some groups are starting to use automated potential energy surface exploration tools to find reaction paths at the quantum chemical level. These tools often also perform conformational sampling. Markus Reiher and his group's work presented here is part of the effort to automate conformer searching in this way.

(716:[716]716) **Jason Lynam** addressed Jeremy Harvey and Markus Reiher: Jeremy, can we apply this to 'the beast' to use this as an example? Markus, in your diagram you have a spread of 800 kJ mol⁻¹ so this is a very large difference in energies.

Jeremy Harvey responded: Intermediate **9**, see *e.g.* Fig. 2 in the paper, which we call 'the beast' because of its conformational complexity, is a good example of the conformational problem. We haven't shown all our results on conformational exploration, but if I recall correctly, the different conformers span quite a broad range of energies, many tens of kJ mol⁻¹. I think Markus mentioned that the very large span in his Figure is somewhat specific to that problem, so this 800 kJ mol⁻¹ range is a bit unusual.

Markus Reiher answered: The few compounds high up in energy are reduced species with a negative excess charge. We have not taken into account any sort of environment (solvation, explicit reductant, dielectric continuum model) which creates this huge energy as this was not important for creating the algorithm. However, there is also an advantage to this, as we pointed out in the paper: these energies can be used as bricks to piece together a 'real' reaction energy and so you do not need to specify a specific reductant at this stage. Once you have chosen a reductant, you may calculate its ionization energy that can then be combined

with the electron affinity given in our figure to yield a proper reaction energy of the reduction process (of course, proper environment modelling would be required through, *e.g.*, dielectric continuum embedding). Clearly, in a complete exploration of a mechanism, the choice of acid and reductant must be made and any energy cut-offs, which might be employed to restrict the exploration to steps that are likely to occur at a given temperature, must take all these species into account (*i.e.*, such an energy cut-off would not work for the energies reported in our paper, and obviously, explicit kinetic modelling would also require consideration of all reactants).

(717:[717]717) **Natalie Fey** addressed Markus Reiher and Jeremy Harvey :I wanted to comment on the issue of exploring conformational space. For the part of the computational chemistry community who do not routinely run molecular dynamics, the nightmare scenario is that there is a high energy conformer reacting considerably faster than any of low energy conformations standard systematic and stochastic searches might detect, or indeed that a conformational change itself could contribute to the reaction pathway. Yesterday, we heard about interactions between phenyl rings which helped to transmit selectivity across a complex. This means that conformers need to be sampled to predict selectivity, something in which a lot of effort can be invested, *e.g.* by Norrby's group.

In many studies I have been involved in, we came to the conclusion that doing something about conformers is better than not doing anything about them, but most good catalysts seem to have fairly well-defined conformational preferences, with some additional computational noise from minor changes that are not likely to affect results a lot. In such cases, a full exploration of conformational space may be less important, but it can be difficult to know this without a lot of calculations. It could be interesting to use machine learning/data analysis to develop a descriptor to help researchers determine whether conformational change is likely to matter or not. In organometallic catalysis, for example, 'flapping' phenyl rings do not make a large energetic contribution, whereas isopropyl and cyclohexyl substituents can have a bigger impact, both on ground states and on transition states. It might be possible to mine the calculations used to develop the ligand knowledge bases (described in Ref. 1 for example), where we have looked at a wide range of ligands bound to different metal centres, to quantify the impact of conformational change on organometallic complexes. This will not cover everything, but might provide a useful starting point and I would be interested in further discussions.

1 J. Jover, N. Fey, J. N. Harvey, G. C. Lloyd-Jones, A. G. Orpen, G. J. J. Owen-Smith, P. Murray, D. R. J. Hose, R. Osborne, M. Purdie, *Organometallics*, 2010, 29, 6245–6258

Markus Reiher replied: I can only agree. The whole automated exploration approach is propelled by the necessity to investigate conformational depth sufficiently (see Ref. 28 for an example). I also believe that machine learning could eventually help to predict the severity of the conformational problem at each elementary step of a given problem. However, we first need to generate the data to learn from, before we can exploit such schemes. Owing to the vast nature of chemical reaction space, we must be prepared to have all such models implemented in such a way that they can be applied in a rolling fashion. Only then, new

data can be incorporated to account for an increasing breadth of mapped chemical reaction space.

(718:[718]718) **Patrick Morgan** addressed Jeremy Harvey and Markus Reiher :In terms of the topics that we have been discussing today and throughout the conference, how far away do you believe we are from chemists without expertise in computational theory being able to use these tools to take an organometallic reaction and calculate transition states, reaction pathways and other information of interest?

Markus Reiher replied: This question is very close to my heart and I would like to answer it by referring to some of our past work that points to our vision of the future of computational chemistry for every chemist. One goal of our SCINE project is to dramatically lower the entrance barrier to quantum chemical calculations for non-experts. It is for this reason that we have automated mechanism exploration^{1,2} and interactive exploration^{3–6} within one implementation. For this we had to come up with many new technical innovations that deliver more robust and stable algorithms for the task, too many to be discussed here. Once in final form, this setting shall allow any operator to intuitively interact with the molecular system under consideration without acquiring a huge amount of knowledge in quantum chemistry, molecular dynamics, and kinetics. Eventually, to achieve this goal there must be no necessity for referring to a user's manual or provide a plethora of option in a detailed pull-down menu structure. We would like to have context-dependent menus that will only require operator input if automated procedures cannot take a decision, which is supposed to be a rare event. This also requires us to define a limited number of base (electronic structure) models to be manageable and use machine learning as glue and putty to make up for their deficiencies.^{7–9} It is for this reason that we have been working on uncertainty quantification¹⁰ of electronic structure models that eventually will allow us to overcome the indomitable explosion of quantum chemical acronyms which may be taken as a sign of helpless and unconscious surrender to electronic variability induced by the perfidious complexity of the external potential created by atomic nuclei in molecules. Assuming that there will be no universal and efficient model developed on pre-defined benchmark data that can properly deal with any electron-correlation problem, our models are system focused and improvable in a rolling fashion exploiting uncertainty measures.^{7–9} Having said this, I believe that the complexity of mechanistic explorations of chemical problems can be tamed, without requiring a skilled theoretician.

1 M. Bergeler, G. N. Simm, J. Proppe, M. Reiher, *J. Chem. Theory Comput.*, 2015, **11**, 5712–5722

2 G. N. Simm, M. Reiher, *J. Chem. Theory Comput.*, 2017, **13**, 6108–6119

3 K. H. Marti, M. Reiher, *J. Comput. Chem.*, 2009, **30**, 2010.

4 M. P. Haag, M. Reiher, *Faraday Discuss.*, 2014, **169**, 89.

5 M. P. Haag, A. C. Vauche, M. Bosson, S. Redon, M. Reiher, *ChemPhysChem*, 2015, **15**, 3301.

6 S. Amabilino, L. A. Bratholm, S. J. Bennie, A. C. Vaucher, M. Reiher, D. R. Glowacki, *J. Phys. Chem. A*, 2019, **123**, 4486.

7 G. N. Simm, M. Reiher, *J. Chem. Theory Comput.*, 2018, **14**, 5238.

8 C. Brunken, M. Reiher, arXiv: 1908.10492, 2019

9 J. Proppe, S. Gugler, M. Reiher, Gaussian Process-Based Refinement of Dispersion Corrections, arXiv: 1906.09342, 2019

10 J. Proppe, T. Husch, G. N. Simm, M. Reiher, *Faraday Discuss.*, 2016, **195**, 497.

In the not too far future, I assume that we will have a virtual assistant that carries out the calculations and is steered by natural language. Natural language processing (NLP) is already standard on smartphones and tablet computers. More importantly, algorithms have already become available as open-source software for some time. The first bottleneck therefore is to first provide a program that can process written language and commands. We have implemented such a bot for starting standard quantum chemical calculations and plan to expand on this work. However, once combined with NLP software, a new challenge will be the training of the NLP system. For instance, Jan Unsleber in my group has already done some tests and when you ask such an NLP system to understand the spoken words 'Hartree Fock', it returns the question 'What do you mean by 'Country Folk'?'. From the capabilities of today's commercially available software such as Apple's Siri®, Amazon's Alexa®, and the like, the step to a virtual lab and simulation assistant is not big. Such a system will also be able to act as a translator to rephrase chemical language such that it can be understood by the simulation machine and *vice versa*. For instance, if you would like to change an uncharged nucleophile by an anion in an exploration, which on a piece of paper is not a big issue, the NLP system should ask you about the environment and counter cations to balance the charge and avoid artefacts from excess charges.

Jeremy Harvey added: In the discussions, we have focussed on the progress that still needs to be made. But I'd like to turn this round in my answer, and point out that we already can do what you ask! Compared to ten, twenty years ago, the number of mainly synthetic labs and mainly synthetic researchers who are not experts but regularly use computations to look at reactivity has shot up. Some of the most challenging issues are not easily tractable for non-experts, but many (slightly) more straightforward issues are, and this is reflected in the literature and in this meeting. I feel this is a great development! I am sure that computation will become more and more widespread as a tool to help in looking at reaction mechanisms, although experimental data will remain key in my view.

(719:[719]719) **Jason Lynam** made a final comment: When you start to work with colleagues in other fields it's important to get your head around the terminology so you can communicate properly.

Conflicts of Interest

There are no conflicts to declare.