Catalysis Letters 28 (1994) 79–87

Human/computer interactive elucidation of reaction mechanisms: application to catalyzed hydrogenolysis of ethane

Raúl E. Valdés-Pérez

Department of Computer Science and
Center for Light Microscope Imaging and Biotechnology,
Carnegie Mellon University, Pittsburgh, PA 15213, USA

Received 13 January 1994; accepted 11 June 1994

The computer program MECHEM is a tool for interactive elucidation of reaction mechanisms. Such programs confront severe combinatorial problems on complex multistep reactions, hence the practicality of these efforts is doubtful. We report the first evidence that MECHEM is successfully applicable to practical experimental problems, in this case, catalyzed hydrogenolysis of ethane.

Keywords: computer-aided elucidation; reaction mechanisms; ethane hydrogenolysis

1. Introduction

The computer program MECHEM, which has been developed over the last five years [1–3], is intended as a tool for the interactive elucidation of reaction mechanisms. Given the observed reaction products and intermediates, and other constraints on plausible mechanisms, the program searches for all the simplest mechanisms that can explain the evidence, i.e., that involve the fewest number of conjectured species and the minimal number of elementary steps.

The various algorithms that make up the program's design have been reported within computational chemistry (the next section makes detailed citations). However, there is no hard evidence that the program is successfully applicable to real problems of experimentalists. A healthy skepticism is warranted, given the long experience with the combinatorial difficulties that confront programs that deal with chemical reactions.

The aim of this letter is to present the first such evidence. MECHEM was applied interactively to find several simplest mechanisms for ethane hydrogenolysis, one of which, in its detail, contains interesting features. For completeness, we make some attempts to evaluate the new mechanism in light of past experiments.

The secondary aim of this letter is to acquaint experimentalists with the program, whose further development is in progress.

2. Overview of the program

The great majority of computer programs for the synthesis (e.g., LHASA [4]) or prediction (e.g., CAMEO [5]) of organic reactions share one principle, whether they work in the synthetic or retrosynthetic direction: a chemical reaction *reactants* → *products* is formulated as (and generated by) mechanistic "operators" that express typical reactions that molecular structures (the reactants) undergo. Since knowledge of the applicability of such operators, as well as their types, is far from perfect, a large search problem ensues which presents combinatorial difficulties in the case of complex reactions. The combinatorics is due to imperfect knowledge: many operators have to be tried, and the recursive nature of the problem leads to a large search space.

The approach underlying MECHEM, on the other hand, is to consider what is the objective "hypothesis space" in an elucidation problem: the set of all mechanisms containing S total species, R total steps, and (as a practical compromise) at most two reactants and two products per elementary step. Given a fixed S and R, one can generate all possible mechanisms and test them against the constraints. This formulation leads to the decision to search for mechanisms by successive increments in S and R. On any specific problem, the first S to try is S_0, the total number of known starting materials (including the catalyst) and observed products and intermediates; no S smaller than S_0 can lead to a satisfactory mechanism, because the presence of all the observed species must be explained.

If the class of mechanisms having S_0 species is searched and finally rejected, then S is set to $S_0 + 1$, which means that one more species is introduced. How does the program introduce species not given to it as input, if it does not use mechanistic operators? MECHEM "conjectures" an unseen species X (literally, a variable) and uses it to search for mechanisms as if it were just another given product or intermediate. If S has reached, say, $S_0 + 4$, then four variables, X_1, X_2, X_3, and X_4 are conjectured. After the program builds a mechanism or partial mechanism, it tries to balance all the steps, thereby determining the molecular formulas for the variables. Then, another algorithm infers possible molecular structures for the new formulas, by exploiting the context of the steps in which the variables take part, together with the assumption that every elementary step involves at most N changes to bonds ($N = 3$ by default, since most proposed elementary reactions involve at most three changes to bonds).

If successful, MECHEM's output is a list of simplest reaction mechanisms that are fully balanced and specified structurally: each step is presented in detail and is balanced, and all "conjectured" species are known structurally (although not yet three-dimensionally). Our hypothesis is that by searching for the simplest mechan-

isms, the inherent combinatorics (due to imperfect knowledge) in all such problems can be tamed sufficiently to achieve useful results on practical problems. Our current applications focus is catalytic chemistry, since catalytic mechanisms are complicated, hence computer-aided methods can have significant impact.

This section is necessarily a brief overview of the program. Further details can be found as follows: initial conception [6], pathway generator [1], structural tests of elementary steps [2], inferring the molecular structures of variables [3], heuristics for taming the combinatorics [7], canonical ordering (and generation) of multi-step mechanisms [8], verifying catalytic action [9], optimal overall stoichiometry [10], and testing overall stoichiometry [11]. These are all rather technical papers; no separate overview paper is yet available.

MECHEM is used interactively. First, the experimentalist (henceforth, the "user") states the starting materials, what species have been observed experimentally, and other constraints on the reaction. The program than reports the simplest mechanisms, which typically prompt the user to object to various aspects of the proposed mechanisms; these objections may lead to rejecting all of the reported mechanisms. The user then draws on his chemical intuition to formulate new constraints, and the program is re-run with these. This interactive process continues until the user is satisfied that none of the reported mechanisms is objectionable. (Some reactions are too complex for the program to handle currently, but many are within its competence.) What to do with this knowledge (e.g., try a different catalyst, design an experiment, carry out a kinetic analysis) is outside the current scope of this research, and is left to the expert or to other techniques.

The rest of this letter examines the reaction of catalyzed ethane hydrogenolysis and some previous proposals regarding its mechanism. MECHEM's output on a small number of constraints is then analyzed in some detail. Again, the aim is to report the first available evidence that the program is not merely an implementation of inapplicable algorithmic inventions, but rather can assist usefully the work of experimentalists.

3. Hydrogenolysis of ethane

The most detailed study of ethane hydrogenolysis was reported by Sinfelt [12,13] about 20 years ago. Goddard et al. [14] proposed a mechanism[#1] for the reaction, as did earlier workers, e.g., Gudkov, Guczi, and Tétényi [15]. In a recent article, Rudd and Dumesic [16] made the following observations:

"Several different mechanisms have been proposed in the literature for ethane hydrogenolysis over supported metals. Although there is disagreement on the detailed nature of individual steps, there is general agreement on the nature and the sequence of these steps. Qualitatively, general aspects of the mechanism can be

[#1] In this paper, the term "mechanism" signifies a list of elementary reaction steps that are fully balanced and with full structural information on all species.

summarized as follows. Hydrogen is adsorbed on the surface in an atomic form. Ethane is adsorbed dissociatively, undergoing cleavage of a C–H bond. Further dehydrogenation of the C_2H_x species occurs, accompanied by the creation of additional bonds between the C_2H_x species and the metal surface. The C–C bond breaks and CH_y species are produced. Hydrogenation of the CH_y species takes place, followed by the desorption of methane."

Rudd and Dumesic proceed to give the following working mechanism (henceforth, RDM) for ethane hydrogenolysis:

1. $H_2 + 2* \leftrightarrow 2H*$
2. ethane $+ 2* \leftrightarrow C_2H_5* + H*$
3. $C_2H_5* + 2* \leftrightarrow C_2H_4* + H*$
4. $C_2H_4* + * \leftrightarrow C_2H_3* + H*$
5. $C_2H_3* + * \leftrightarrow C_2H_2* + H*$
6. $C_2H_4* + 2* \leftrightarrow 2CH_2*$
7. $C_2H_3* + 2* \leftrightarrow CH* + CH_2*$
8. $C_2H_2* + 2* \leftrightarrow 2CH*$
9. $CH* + H* \leftrightarrow CH_2* + *$
10. $CH_2* + H* \leftrightarrow CH_3* + 2*$
11. $CH_3* + H* \leftrightarrow$ methane $+ 2*$

4. Results

We have used MECHEM to see whether simple alternative mechanisms could be found for ethane hydrogenolysis. After several runs in interactive mode (in collaboration with an experimentalist), the following constraints led to an acceptable output:

– The overall stoichiometry is 1(ethane) $+1(H_2) \rightarrow 2$(methane).

– A catalyst reaction site (modelled as $*$) forms one bond. Two reaction sites are modelled as $2*$ (this notation does not by itself imply a bond across the two sites).

– CH_2*–CH_2* and $CH*$–$CH*$ are required intermediates (these do form "bridges" over two reaction sites on the catalyst).

– H_2 is not a product of a step (except as the reverse of initial dissociation).

– Every reaction intermediate is adsorbed on the catalyst, i.e., every intermediate contains $*$.

– No species contains three carbons nor spans three catalyst reaction sites.

– There is a maximum of three bond changes (cleavage or formation) per step. (This constraint is justified by the observation that most proposed elementary reactions involve at most three changes to bonds.)

– $CH_2*–CH_2*$ is a precursor (not necessarily single-step) of $CH*–CH*$.

There is no constraint that carbon is required to be fully coordinated; for example, the mechanism presented below contains an intermediate $CH*$ in which the carbon atom has two free valences.

The program reported [#2] several equally-simple mechanisms, among which the exact mechanism given by Gudkov et al. [15] was included (the RDM version above contains the Gudkov mechanism, but also includes the extra steps 6 and 7 which MECHEM did not find necessary to account for the given constraints). One mechanism reported by the program, shown below, differed substantially from the others mentioned. The intermediate species not given as input to the program are shown within angular brackets for readability.

1. $H_2 + 2* \xrightarrow{1} 2\langle H*\rangle$
2. $2* + \text{ethane} \xrightarrow{1} \langle H*\rangle + \langle + \langle CH_3–CH_2*\rangle$
3. $2* + \langle CH_3–CH_2*\rangle \xrightarrow{1} CH_2*–CH_2* + \langle H*\rangle$
4. $* + CH_2*–CH_2* \xrightarrow{2} \langle H*\rangle + \langle CH_2*–CH*\rangle$
5. $2\langle CH_2*–CH*\rangle \xrightarrow{1} CH_2*–CH_2* + CH*–CH*$
6. $CH*–CH* \xrightarrow{1} 2\langle CH*\rangle$
7. $\langle H*\rangle + \langle CH*\rangle \xrightarrow{4} * + \langle CH_2*\rangle$
8. $2\langle CH_2*\rangle \xrightarrow{2} \langle CH*\rangle + \langle CH_3*\rangle$
9. $\langle H*\rangle + \langle CH_3*\rangle \xrightarrow{2} 2* + \text{methane}$

A set of stoichiometric numbers that generates the required overall stoichiometry is shown over the arrows; these numbers are not to be confused with rate constants. Weighted summation of all the steps according to the indicated numbers gives the stoichiometry $1(\text{ethane}) + 1(H_2) \rightarrow 2(\text{methane})$. Although the symbol $\rightarrow$ has been used, the possible reversibility of any or all steps is not excluded.

5. Discussion

The mechanism found by human/computer interaction (henceforth, HCIM) is of comparable simplicity, exhibits somewhat different chemistry, and seems a plausible account of some catalytic ethane hydrogenolysis processes. This section will

[#2] In its run on the final set of ethane hydrogenolysis constraints, the program takes about ten minutes or so on a high-end workstation.

also examine how well HCIM explains some previous data from isotopic labelling experiments.

Comparison with RDM. An interesting feature of HCIM occurs at step 8, in which one singly-adsorbed CH_2* transfers a hydrogen atom to a second adsorbed CH_2*. This hydrogen transfer of disproportionation gives rise to CH_3*, which proceeds to form methane at step 9, and leaves behind another CH*, which then undergoes hydrogen addition via the previous step 7. Steps 4–6 involve a similar pattern, in which the CH*–CH* product of step 5 goes on to cleave the C–C bond at step 6, while the CH_2*–CH_2* product falls back one step in the mechanism. Here, methane is not formed by the successive hydrogenation of adsorbed hydrocarbon species, contrary to the other mechanisms. Rather, one adsorbed hydrocarbon intermediate hydrogenates another adsorbate. While HCIM contains the same intermediates as RDM, their roles in the reaction differ.

One can examine the hydrogenation phase in HCIM more closely (a similar analysis holds for the dehydrogenation phase):

7. $H* + CH* \rightarrow * + CH_2*$
8. $2(CH_2*) \rightarrow CH* + CH_3*$
9. $H* + CH_3* \rightarrow 2* +$ methane

Stoichiometrically, step 8 could be written as the following two steps in which hydrogen is transferred first to its own site on the catalyst, then on to the second adsorbate:

8a. $* + CH_2* \rightarrow H* + CH*$
8b. $H* + CH_2* \rightarrow * + CH_3*$

It could be argued that steps 8a and 8b are experimentally indistinguishable from step 8. However, step 8a is actually the reverse of step 7: if step 8 in HCIM is replaced by steps 8a and 8b, and the result is simplified (leaving only steps 8b and 9), then the successive hydrogenation steps of the RDM mechanism are obtained. That this can be accomplished is not too surprising because the two mechanisms must generate the same stoichiometry $1(\text{ethane}) + 1(H_2) \rightarrow 2(\text{methane})$. Despite such formalistic manipulations, the chemical implications of the two mechanisms differ somewhat.

The two disproportionation reaction steps 5 and 8 in HCIM imply a relative surface abundance of the reactants CH_2*–CH* and CH_2* (the two disproportionation schemes are independent, in the sense that one or the other could be replaced with a successive-(de)hydrogenation scheme). The likelihood of such abundances is unclear, and is a drawback to this mechanism. In compensation, a point in favor of HCIM is as follows.

Evidence from isotopic labelling. HCIM provides an alternative explanation for the product distributions from ethane/deuterium experiments carried out by Guczi et al. [17], who found high production of CD_4 over a nickel powder catalyst.

HCIM predicts that a significant fraction of the products will consist of CD_4 formed via the following pathway, in which the CD* formed by steps i and ii feeds the (replicated) step iii:

i. $D* + CH* \rightarrow * + CDH*$ (step 7)

ii. $2(CDH*) \rightarrow CD* + CDH_2*$ [50%] (step 8)
$\rightarrow CH* + CD_2H*$ [50%]

iii. $D* + CD* \rightarrow * + CD_2*$ (step 7)

iv. $2(CD_2*) \rightarrow CD* + CD_3*$ (step 8)

v. $D* + CD_3* \rightarrow 2* + CD_4$ (step 9)

On the other hand, the cited proposed mechanisms (RDM in particular) can form CD_4 only by ethane exchange with deuterium, or by methane exchange with deuterium. For example, one route for ethane exchange would begin with these two steps from RDM:

$C_2H_6 + 2* \rightarrow C_2H_5* + H*$ (step 2)

$C_2H_5* + D* \rightarrow C_2DH_5$ (reverse of step 2)

which are followed by adsorption and successive dehydrogenation of C_2DH_5, cleavage of the C–C bond to form CD* and CH*, and successive deuteration of CD* eventually to form desorbed CD_4.

Since HCIM can also undergo ethane exchange in addition to the CD_4-producing pathway above, the two mechanisms differ qualitatively in their CD_4 productivity: the HCIM mechanism predicts more completely-deuterated methane than mechanisms based on successive hydrogenations.

Hydrogen-transfer mechanisms. Previously, it has been proposed (e.g., by Zaera and Somorjai [18]) that hydrogen is transferred between adsorbates, for example, in the context of hydrogenation of ethylene over a single-crystal platinum surface, although they do no provide a detailed pathway. The nine-step pathway reported here could be regarded as a confirmation and elaboration of such proposals (but for hydrogenolysis of ethane into methane), since the mechanism is as simple as other proposed mechanisms.

We have examined in some detail the properties of the reaction mechanism found by human/computer interaction in order to show that such interaction can lead to interesting results of potential value to experimentalists. The goal is not to establish the primacy of this mechanism; disproving it would not change the basic aim of this letter, which is to report initial evidence on the usefulness of a new technique. It would be interesting to follow up by devising and carrying out an experiment to discriminate between the above mechanisms, but these steps are beyond our own expertise.

6. Methodology

No kinetic analysis has been carried out in this letter, since our aim has been to show the value of an interactive technique for consideration of reaction mechanisms. More generally, chemical kinetics has played a minor research role in MECHEM so far [#3], since this project has focused on the reasoning that occurs before a detailed kinetic analysis is warranted. None of this implies any disregard for chemical kinetics: we hope to include some capabilities in a future comprehensive aid for experimentalists. One may regard MECHEM and the work on microkinetic analysis [19] as natural complements. The results here do show, though, that much can be achieved prior to a kinetic analysis.

There arises the natural question of whether interactive elucidation methods will result in a plethora of possible mechanisms. This is a concern in the case of computer-aided synthesis programs, since the user may be left suspecting that the computer program, due to deficient knowledge, will report a majority of impractical synthesis, and hence has the potential to distract rather than to guide. However, the situation in mechanism elucidation is subtly different. If the human/computer interaction leads to a number of simplest reaction mechanisms that cannot be discriminated further based on scientific knowledge or available experimental evidence, then this is a deficiency neither of the program nor of the interactive method, but rather of the current state of scientific knowledge or of experimental technique. Perhaps widespread use of computer tools will serve to explicate these issues, if indeed they are present [#4].

MECHEM is not yet available for external use, partly because it is still under development, and partly because of the complete lack of a user interface that would allow routine autonomous use. Our future plans include taking the steps that will enable independent use, and also implementing the program on distributed workstations to achieve an order-of-magnitude speed-up. In the meantime, the author invites collaborations with interested experimentalists.

7. Conclusion

This letter describes the first available evidence that MECHEM, an interactive aid for elucidation of reaction mechanisms, can be applied successfully (i.e., with interesting results) to practical problems of experimentalists in catalytic chemistry.

[#3] However, the program can test – using nonkinetic algebra – whether a given mechanism can account for individual transitions in concentration levels. This capability was used above, for example, to verify that HCIM accounts for the given overall stoichiometry.

[#4] Our experience so far indicates that the *simplest* reaction mechanisms, in terms of number of steps and conjectured species, tend not to be numerous. However, increments in these two simplicity parameters can result in a sizable number of *non-simplest* mechanisms.

Presenting such evidence is important, since the combinatorial difficulties that have hampered computer programs dealing with complex reactions make it, today, a priori doubtful whether a new program can have any practical success.

Further work on MECHEM is in progress, with the goal being to show that an interaction that combines chemical insight and knowledge with search and calculation, thus drawing on the respective strengths of human and machine, can be a fruitful new approach in catalytic chemistry.

Acknowledgement

The author thanks Professor E.I. Ko for suggesting ethane hydrogenolysis as a test case, for helping with the constraints and inspecting the output of the program, and for commenting on drafts of this paper. This work was supported partly by the W.M. Keck Center for Advanced Training in Computational Biology at Carnegie Mellon University, the University of Pittsburgh, and the Pittsburgh Supercomputing Center, and partly by a Science and Technology Center grant #BIR-8920118 from the National Science Foundation to the Center for Light Microscope Imaging and Biotechnology.

References

[1] E. Corey, A. Long and S. Rubenstein, Science 228 (1985) 408.
[2] W.L. Jorgensen, E.R. Laird, A.J. Gushurst, J.M. Fleischer, S.A. Gothe, H.E. Helson, G.D. Paderes and S. Sinclair, Pure Appl. Chem. 62 (1990) 1921.
[3] R.E. Valdés-Pérez, Tetrahedron Comput. Method. 3 (1990) 277.
[4] R.E. Valdés-Pérez, J. Comput. Chem. 13 (1992) 1079.
[5] R.E. Valdés-Pérez, J. Comput. Chem. 14 (1993) 1454.
[6] R.E. Valdés-Pérez,, J. Comput. Chem., in press.
[7] R.E. Valdés-Pérez, J. Chem. Inform. Comput. Sci., in press.
[8] R.E. Valdés-Pérez, J. Chem. Inform. Comput. Sci. 31 (1991) 554.
[9] R.E. Valdés-Pérez, J. Phys. Chem. 96 (1992) 2394.
[10] R.E. Valdés-Pérez, J. Phys. Chem. 95 (1991) 4918.
[11] R.E. Valdés-Pérez, Deductive assistance for elucidation of reaction pathways, presented at Annual Meeting of American Institute of Chemical Engineers, Chicago, 11–16 November 1990.
[12] J. Sinfelt, Catal. Rev. 3 (1969) 175.
[13] J. Sinfelt, Adv. Catal. 23 (1973) 91.
[14] S. Goddard, M. Amiridis, J. Rekoske, N. Cardona-Martinez and J. Dumesic, J. Catal. 117 (1989) 155.
[15] B. Gudkov, L. Guczi and P. Tétényi, J. Catal. 74 (1982) 207.
[16] D. Rudd and J. Dumesic, Catal. Today 10 (1991) 147.
[17] L. Guczi, B. Gudkov and P. Tétényi, J. Catal. 24 (1972) 187.
[18] F. Zaera and G.A. Somorjai, J. Am. Chem. Soc. 106 (1984) 2288.
[19] J.A. Dumesic, D.F. Rudd, L.M. Aparicio, J.E. Rekoske and A.A. Treviño, *The Microkinetics of Heterogeneous Catalysis* (American Chemistry Society, Washington, 1993).