

MECHGEN: Computer Aided Generation and Reduction of Reaction Mechanisms

András Németh,* Tamás Vidóczy, Károly Héberger,* Zsolt Kúti, and János Wágner

Institute of Chemistry, Chemical Research Center, Hungarian Academy of Sciences, P.O. Box 17,
H-1525 Budapest, Hungary

Received February 20, 2001

The paper describes selection rules implemented in a software generating “possible reaction mechanism”, i.e. a set of elementary reactions chosen from all stoichiometrically possible reactions. The novelty of the approach lies in the fact that the user has to define all species involved (reactants, intermediates, products), and the rules applied with user-set limits reduce the resulting mechanism to a reasonable set of possible elementary reactions. The computer code consists of five parts: (i) definition of species, and introducing its characteristics (structure and thermodynamic data); (ii) definition of the reacting system and generation of all stoichiometrically possible reactions; (iii) reduction of the mechanisms using complexity and thermodynamic constraints based on user-set limits; (iv) calculation of the resulting pathways (routes of the various atoms or groups of atoms transferred from one species to another); and (v) tools to help visualization of the process by finding those elementary processes which realize a given pathway. Reasonable flexibility is ensured for using selection rules based on various criteria with limits set by the user. The various pathways are shown (in a matrix form), which offers an overview of the entire process.

INTRODUCTION

Construction and investigation of reaction mechanisms have been among the main interests in chemical kinetics for a long time and it seems to remain one of the key issues in the future. The traditional way to compile a mechanism is to collect all relevant reactions, but “relevance” is a rather subjective term in this respect. Since the existence of reaction databases relevance can be formulated precisely, e.g. reactions where the reactants are on our list of chosen species, and/or for which (at least some) rate parameters are known. On one hand, such an approach can result in a too extensive mechanism (the interaction between intermediates may generate a large number of reactions); on the other hand, a reaction may be unjustifiably omitted, if there are no data available on that particular reaction. Thus, even for reactions where there are abundant experimental results available (such as combustion, pyrolysis, atmospheric chemistry, etc.) the construction of a reaction mechanism is not a simple task, whereas in a less-studied field such compilation of the mechanism is quite subjective.

For this reason, several groups proposed the application of computers for the generation of reaction mechanisms and analysis of the results.^{1–29} The number of generated reactions increases in a factorial measure with the increasing number of species, and this increase is even more pronounced if large molecules are involved. Therefore, most of these proposals encode expert chemical knowledge in order to avoid including unlikely reactions into the program. The purpose of these rules is to keep the number of reactions within a reasonable limit; however, this approach inherently risks unjustified omission of some relevant reactions.

Due to the large number of relevant publications, we cannot endeavor to survey all the trials to generate and analyze reaction mechanisms by computer here. Below we

summarized the most characteristic investigations and those which are in closest connection to our study.

One of the earliest computer programs is TAMREAC; its use is restricted to organometallic chemistry.^{1–3} “RAIN”, generating reaction networks, was developed for modeling pyrolysis of hydrocarbons.^{4–7} COMSICAT, intended as computer aided prognosis of catalyst efficiency, was written to run on an EC-1050 computer, not in use anymore.⁸ KING is designed for generating kinetic networks.⁹

Characteristic of these programs is that they are not general enough, as far the field of application is concerned, and special expertise is needed to work with them.

Further programs encode expert chemical knowledge in order to avoid unlikely reactions. In the expert system developed by Chinnick et al.¹⁰ generation is constrained by rules relating to initial conditions and kinetic and thermodynamic data. The mechanisms generated so far, describing the combustion of small acyclic hydrocarbons, compare favorably with those produced by human experts.

Chevalier et al. applied rule-oriented programming techniques to model combustion of higher hydrocarbons.¹¹

EROS is also a system based on expert knowledge for prediction of the course and the products of chemical reactions.^{12,13} Central to this approach is a formal handling of reactions based on electron shifting patterns, showing which bonds are broken and which are formed in the reaction.

Considerations based on steady state and partial equilibrium assumptions to reduce the complexity of the chemical reaction schemes are also in use.^{14–16}

EXGAS generates comprehensive primary reaction mechanisms and then proceeds to produce lumped secondary mechanisms describing possible reactions of lumped primary products.^{17,18}

An alternative method to simplify chemical kinetics is Computational Singular Perturbation (CPS).^{19,20} By this method the reaction system is divided into slow and fast reaction groups. Also the method of Intrinsic Low-

* Corresponding authors e-mail: anemeth@chemres.hu (Németh); heberger@chemres.hu (Héberger).

Dimensional Manifolds in Composition Space²¹ and its improvement²² are based on the same basic ideas as CPS.

In these cases, however, knowledge of the kinetics is required to decide which species and reactions may be assumed as less important.

The expert knowledge built in these programs is restricted to specific fields, at least in the present stage of development and does not serve to create new expert knowledge.

MECHEM and ChemNet meet the latter requirement.^{23–27} They carry out mechanism elucidation by knowledge guided search within a combinatorial space to find the simplest plausible mechanisms. Thermodynamics is not incorporated into them, and they were applied mostly to catalytic reactions.

In some measure, also the thermodynamics related to model thermal decompositions is taken into account in CASB.^{28,29}

Based on this review the aim of the present research is to develop a method for creating new knowledge by generation and reduction of reaction mechanisms applicable for a fairly broad field and to encode it in a user-friendly way.

All approaches above can be characterized in that new species may be generated by the algorithm. As a consequence special rules are needed to avoid explosive growth of the number of species considered. Contrary to the general principles applied in the above programs, a different approach is feasible, namely the generation of intermediate species until their reaction produces the desired product. These species must be specified at beginning of the generation.³⁰

In our earlier publication we proceeded along the latter lines, proposing another approach.³¹ The suggested method consists of several steps: (i) compilation of the complete list of species (called species space, containing all starting compounds, intermediates and end products); (ii) computer-aided generation of all stoichiometrically possible reactions within this species space; and (iii) reduction of this set based on mechanistic and thermochemical considerations, resulting in a mechanism with a reasonable number of reactions. Usually, this “reasonable number” is still quite high, thus a convenient tool is offered to follow the pathways of the various atoms or groups of atoms in order to help visualization of the overall reaction.

This method is intended to be used before any kinetic calculation is performed (i.e. no rules based on rate constant values are used). It offers the great advantage of starting from all stoichiometrically possible reactions. The resulting mechanism can be further reduced by kinetic considerations (i.e. further rules based on rate constant values can be applied at a later stage of modeling).

This paper describes the background of selection rules and the outline of the implemented software capable of performing all required steps and calculations.

THEORY

The common problem to most computer codes generating reaction mechanisms is how to avoid the explosive growth of the number of reactions to be included. The majority of the software described earlier limits this number by excluding reactions occurring between secondary species.

The basic idea introduced by some of us earlier³¹ was to define the complete set of reacting species, including reactants, intermediates, and products as the starting point.

Then, the set of all stoichiometrically possible reactions, which can be realized among the species defined, is generated. According to our concept we are interested only in elementary reactions, thus molecularity of a reaction is limited to three, both on the reactant and product sides, fulfilling the requirement of microscopic reversibility. The set of all stoichiometrically possible reactions is generated combinatorially. To each possible combination of maximum three reactants all possible combination of maximum three products is generated, which fulfills the requirement of the conservation of atoms. This set of stoichiometrically possible reactions constitutes the zeroth Possible Mechanism (PM₀). Naturally, this PM₀ may consist of a very large number of reactions for any realistic species set. Therefore, reactions unlikely to occur should be deleted. The following criteria are used in a step-by-step manner to omit reactions, resulting in reduced PM_i's:

Reaction Complexity. Reactions, which cannot be considered as elementary reactions should be omitted.³¹ According to Yoneda³⁰ a reaction is considered to be an elementary reaction, if and only if the number of atom pairs, between which a change in bond multiplicity occurs during the reaction, does not exceed four.³² The number of such pairs of atoms will be denoted by κ . The reactions passing this test constitute PM₁.

Geometric Limitations. We have found that the above definition leaves too many overcomplicated reactions in the PM, which are unlikely to occur. Therefore in addition to defining an upper limit for κ , we introduced the formulation of an obvious geometric constraint (based on a suggestion made by Péter Hajdu).³³ In an elementary reaction the bonding electrons are redistributed simultaneously between the atomic centers. Obviously the number of pairs of atoms between which there was no bonding before the reaction, but a new bond is formed, is very limited, since these atoms usually had not been in the vicinity of each other prior to the reaction, and must, therefore, approach their respective pair. The probability of more than two such pairs approaching simultaneously is negligible, thus we can limit this number to two. Accordingly, we may consider a reaction to be elementary if the number of pairs of atoms, between which there was no bond but a new bond is formed, does not exceed two.³⁴ The number of such pairs of atoms is denoted by χ . The reactions passing this test constitute PM₂.

Thermochemical Considerations Based on the Reaction Enthalpy. Certain sets of reactions can be compared,³⁵ and those being relatively too endothermic can be ruled out from the PM.³⁶ It is important to note that since no kinetic data are considered at this stage of modeling, only reactions with identical reactant sides can be compared (thus avoiding any complications due to unknown concentrations of reactants). Furthermore, since various reaction types (initiation, chain propagation, and chain termination processes) occur at widely different rates in any realistic process, only reactions of similar type should be compared. Accordingly, this reduction will necessarily leave at least one reaction in the PM for each reactant combination and reaction type. The reactions passing this test constitute PM₃.

Thermochemical Considerations Based on the Gibbs Free Energy of the Reaction. In some cases the PM₃ is too large, and it contains unrealistic reactions. To overcome the above limitations we may introduce another constraint. The

driving force of a chemical reaction is described by its free energy change. It would be tempting to delete all reactions having a positive free energy change. However, the experimental error of thermodynamic quantities and their approximation with polynomial description of the heat capacity of the species necessitate defining a limit of tolerance. Hence, not all reactions with positive free energy change can be discarded. In defining the limit of tolerance we must take into consideration the type of the reaction, as well. Investigations showed that different limits should be used for initiation reactions.³⁷ The reactions passing this test constitute PM₄. Since, however, this constraint is valid only for elementary reactions in equilibrium, it should be used only as an approximate guidance if PM₃ still remains too large for analysis.

The PM is generated applying these mechanistic and thermochemical constraints to all stoichiometrically possible reactions. It is, however, not a simple task to present this PM in an easily comprehensible manner. For this purpose, the concept of chemical communication and its use to define various reaction networks had been introduced in our earlier work.³⁸ With the help of associated calculations the pathways of various atoms or groups of atoms can be followed from reactants via intermediates to products. The reactions corresponding to the various pathways can be looked up automatically.

REALIZATION OF THE CONCEPT, THE MECHGEN PACKAGE: STRUCTURE AND FUNCTIONS

A simplified mechanism of methane partial oxidation will be used as an example. Words in *italics* refer to specific terms used by the program.

1. Choice of System. The term *system* is used to designate the list of all species considered in the process, as stated earlier. It includes all reactants, products, and intermediates such as free radicals or excited-state molecules. A *species library* is gradually built up (every new species entered is added to this library), and the species required for the given overall reaction can be selected. The user may also want to build his/her own library, so the program includes adequate entry options to define a new *species*. From a *species library*, a subset called *system* can be constructed. To provide flexibility for the user, each *species* in a *system* can be a reactant and/or a product. Reactants are considered only on the left-hand side of the reactions, while products are considered only on the right-hand side of reactions during combinatorial generation of all stoichiometrically possible reactions.

2. Combinatorial Generation of PM₀. A maximum of three species can be on either side of a reaction. The program automatically generates all possible reactions, i.e., all those which comply with the rule of invariability of the number of atoms. The user may prohibit the generation of trimolecular reactions (to provide adequate flexibility for using the program in modeling condensed phase reactions as well as gas-phase processes).

Since the only requirement is the fulfillment of the invariability of atoms, if



is a possible reaction, then



will be possible also for all possible choices of "A". Naturally, these reactions are not meaningful for a noncatalytic process in solution chemistry, while in the gas phase a vast number of reactions proceed according to this scheme. The user is offered the choice of taking into account or omitting such reactions. It is important to note that elementary reactions in catalysis should not be described by reactions similar to reaction 2, since the reactants as well as the products have to bind to the catalyst, and such a complex must be considered as a separate species.

3. Reduction of the PM. Various attributes of the reaction are calculated in the next step, to make reduction of large mechanisms feasible. Thus κ , χ , ΔH , and ΔG are evaluated for each reaction.

The complexity of the reaction under study, i.e., the number of pairs of atoms among which there is a change in bond multiplicity during the reaction, is calculated first. A simple example explains the concept



In reaction 3 an OH bond is broken, another OH bond is formed, and the bond multiplicity between C and O is changed from 1 to two, so $\kappa = 3$. Thus, even a simple H atom transfer results in $\kappa = 3$, since there is a simultaneous change in bond multiplicity in one of the reactants. The calculation of the complexity number may be rather time-consuming, especially with large molecules, where there are many similar atoms. In case of simple reactions such as (3) it is easy to see where there is a change in bond multiplicity. However, for large molecules, having many C and H atoms, almost all possible combinations of these have to be taken into account during the calculation, and the true complexity of the reaction is the minimal κ value, calculated for all these combinations. Even though symmetry considerations may expedite the calculation, the time required may be considerable. Therefore a limit can be set for these numbers, above which the program will not calculate the exact value of κ , it will just state that κ is larger than the preset limit (usually 4 or 5), saving valuable computer time.

Then the next attribute, χ , is calculated. In reaction 3 $\chi = 1$, since there is only one pair of atoms, between which the OH bond is formed, where there was no bond before the reaction. There is no closed mathematical formula to convert κ into χ or vice versa; therefore, both have to be evaluated independently (naturally $\kappa \geq \chi$ is always true). Though conceptually calculation of χ is simpler, it may be still rather time-consuming, thus similarly to the case above, a limiting value (usually 1 or 2) can be set, above which the program will not calculate the exact value.

For calculation of thermochemical parameters^{39–43} the following equations are used.

The reaction enthalpy is calculated using

$$\Delta H_{r,T} = \sum_i \nu_i \Delta H_{i,T} \quad (4)$$

where ν_i is the stoichiometric coefficient of each species in the reaction (it is positive for the products and negative for reactants) and $\Delta H_{i,T}$ is the formation enthalpy of the i th species at the temperature of the reaction. The user can,

naturally, specify the reaction temperature. The $\Delta H_{i,T}$ is calculated using the polynomial approximation of heat capacity data^{40–43} (see note 39).

As discussed in our earlier publication,³⁶ the omission of reactions based on reaction enthalpy depends on the κ value of the reaction, as well. Therefore the reaction heat corrected for κ is calculated

$$\Delta H_{rc,T} = \Delta H_{r,T} + x\kappa \quad (5)$$

where x is a user-set value (suggested value is $80 \text{ kJ}\cdot\text{mol}^{-1}$ based on empirical experience).

Since this thermochemical selection rule is ineffective in several cases (at least one reaction is left in the possible mechanism from each subgroup), a further attribute, the Gibbs free energy of the reaction ($\Delta G_{r,T}$), can be applied for filtering unnecessary reactions:

$$\Delta G_{r,T} = \sum_i \nu_i \Delta G_{i,T} \quad (6)$$

$\Delta G_{i,T}$ is the free energy of the i th species at the temperature of the reaction, calculated using the polynomial approximation of heat capacity data^{40–43} (see note 39).

The program offers several options to ease calculation of the various characteristics of reactions. The computation of the chemical and thermochemical attributes can be carried out simultaneously or separately. In dealing with large systems, however, it is advisable to proceed first with computation of the chemical attributes. As this is the most time-consuming procedure, optional commands offer several possibilities: computing the constraints only for one reaction, or, if part of the computation had been performed previously, performing only the remaining part. To save computer time, κ values higher than five and χ values higher than two are not computed.

Reducing the number of reactions to be considered temporarily or permanently is called *filtering*. *Filters* are based on the following: species (given by *formula*, *short name* or *identifier*), complexity numbers (κ , χ), molecularity of the reaction on either side,⁴⁴ reaction enthalpy (corrected for κ), Gibbs free energy, and type of reaction (initiation, propagation, termination).

Filters can be used alone (one by one, to see how many reactions remain after application of a certain filter), or in combination, using logical functions, including negation.

Filtering helps focusing on details, digging down to individual reactions of a species, while the whole reaction scheme is always at hand. The results of *filtering* can be made permanent, i.e., it is possible to delete all reactions not complying with the *filtering* rule(s) specified.

Filtering by κ and χ is straightforward (in case κ or χ value of a reaction exceeds the limit, the reaction is omitted); however, *filtering* based on reaction enthalpy requires some consideration. In the subgroups of reactions to be compared (see note 45) when *filtering* is based on reaction enthalpy, the reaction possessing minimal $\Delta H_{rc,T,\min}$ is selected (the most exothermic/least endothermic reaction). Those reactions will remain in the reduced reaction mechanism, which meet the condition

$$\Delta H_{rc,T} \leq \Delta H_{rc,T,\min} + x' \quad (7)$$

where x' is a user-set limit (e.g. $80 \text{ kJ}\cdot\text{mol}^{-1}$).

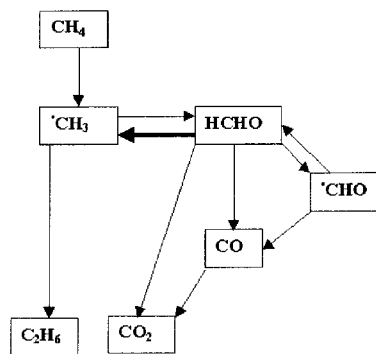


Figure 1. The carbon sequence network of simplified methane combustion. The reaction realizing the heavy arrow is $\text{CH}_2\text{O} + \cdot\text{CHO} \rightarrow \cdot\text{CH}_3 + \text{CO}_2$.

Filtering by the thermochemical attribute, $\Delta G_{r,T}$, means that reactions having free energy exceeding a given value can be omitted from the possible mechanism. The limits of tolerance for each type of reaction (initiation, propagation, and chain termination) can be adjusted separately

$$\Delta G_{r,T} \leq \Delta G_{r,T,\min} + y \quad (8)$$

where y is a user-set limit (e.g. 40 kJ mol^{-1} for chain propagation reactions).

Even though it would be technically possible to *filter* according to the various attributes in any sequence, selection based on reaction enthalpy should be performed as the last *filtering* procedure.⁴⁵

4. Visualization of the PM. The “essence” of chemical information in a mechanism (excluding time-dependent features), embodied in the chemical pathways which describe all the possible routes of transfer of a given atom, or group of atoms between the species is called *sequence*.³¹ This is by far the most complicated and tiresome task to evaluate large mechanisms. After a mechanism has been constructed sequences for each type of atom can be generated. In our example presented below, a sequence of carbon atom is produced (see Figure 1.). This shows all the species, which are involved in the transfer of the C atom. It is the graphical representation of a *one-step communication matrix*,³¹ in which arrows denote the existence of at least one reaction where at least one carbon atom of the species at the start of the arrow is transferred to a carbon atom of the species at the end of the arrow. Reactions corresponding to an arrow can be listed by choosing a “source” species (CH_2O in Figure 1), selecting *Forward* (or *Backward*, defining the direction of transfer), and choosing a “target” species (CH_4 in Figure 1). If the number of species in a sequence is too large, some of them can be viewed by temporarily hiding the unnecessary ones. Then the corresponding matrix will contain only the desired species. This is called a *multistep communication matrix*,³¹ where “+” denotes route(s) through which there exists indirect pathway(s) between two species. A pathway is indirect, if there are reactions which transfer the chosen atom from the starting species to the ending species, without transferring the same atom through any of the species shown in the multistep communication matrix or graph (consequently, a one-step communication matrix cannot show indirect pathways). The program shows the reactions realizing (direct or indirect) pathways by listing appropriate reactions combined with Boolean-algebraic expressions.

Table 1. INPUT: Species Space

| species code | species | restriction |
|--------------|------------------|---------------|
| 10 | methane | reactant only |
| 3 | oxygen | reactant only |
| 11 | formaldehyde | |
| 12 | carbon monoxide | |
| 16 | carbon dioxide | product only |
| 1 | water | product only |
| 13 | ethane | product only |
| 19 | hydrogen atom | |
| 17 | methyl radical | |
| 5 | hydroxyl radical | |
| 18 | oxygen atom | |
| 21 | formyl radical | |

5. Example. Figure 1 is based on a simplified mechanism of methane partial oxidation (where ethane is considered as the only C₂ species, and its consumption is not allowed). The list of *species space* is shown in Table 1.

To illustrate the various steps implemented in the program, let us consider the reaction between oxygen and formaldehyde. Table 2 contains all the stoichiometrically possible reactions, that is all the possible combinations of the species of Table 1, where not more than three species contain one carbon atom + two hydrogen atoms + three oxygen atoms (excluding the identity reaction in which oxygen and formaldehyde would be reformed). The table contains also all the parameters required for the various filtering steps.

Filtering by κ will surely eliminate reaction no. 5 (the numbering of the table is used), and if we employ the suggested limiting value of $\kappa \leq 4$, it will eliminate all reactions except reaction no. 3. Similarly, filtering for χ will surely eliminate reaction no. 5., and employing this filter with $\chi \leq 1$ will eliminate all reactions except reaction no. 3. It is not necessary that the results of filtering by κ and χ are the same; however, this situation is typical.

In this case filtering by ΔH cannot be carried out after filtering by κ and/or χ , since there remains only a single reaction, which cannot be compared with any other reaction. Should we use filtering by ΔH for the whole set of reactions, reactions 1–4 can be compared (all are initiation reactions). Reaction no. 1 can be included in this comparison only because we are dealing with a gas-phase process (in solution the heats of solvation cannot be estimated if the molecularity of the right-hand side of the reaction is not the same for all reactions to be compared). Reaction no. 5 is not an initiation reaction, thus it cannot be included in the comparison. The smallest $\Delta H_{rc,T}$ belongs to reaction no. 2, thus filtering by ΔH will eliminate all reactions having $\Delta H_{rc,T}$ higher than 265 kJ·mol⁻¹. Accordingly, reactions nos. 2, 4, and 5 will remain in the mechanism. This result illustrates the danger of filtering by ΔH before applying the other filters. Unreasonable reactions (e.g. too complex reactions) will force other reactions to be eliminated because of their reaction enthalpy,

Table 3. Generation of the Possible Mechanism

| constraints | no. of the reactions | comment |
|--|----------------------|---|
| no trimolecular reaction | 66 | |
| $\kappa \leq 4; \chi \leq 2$ | 54 | |
| $\kappa \leq 4; \chi \leq 2$ $\Delta H_r \leq 40 \text{ kJ mol}^{-1}$ | 54 | no ΔH_r constraint for initiation reactions |
| $\kappa \leq 4; \chi \leq 2$ $\Delta H_r \leq 40 \text{ kJ mol}^{-1}$ $\Delta G_r \leq 40 \text{ kJ mol}^{-1}$ | 50 | no ΔH_r constraint for initiation reactions |

and when filtering for complexity, they disappear too. In this way reactions proven to occur may be eliminated, consequently filtering by ΔH should be done as the last filtering procedure.

Filtering by ΔG would eliminate reaction no. 3 in case we set too low a limit of tolerance for ΔG . This emphasizes the importance of setting high limits of tolerance at least for the initiation reactions (initiation reactions are characterized by large positive ΔH or ΔG value). It can be seen that all reactions in Table 2, except for reaction no. 3, have a large negative ΔG value; however, the complexity number of these reactions is above the limit set.

The number of reactions generated and remaining after each *filtering* are given in Table 3 and in the sequence network in Figure 1 for the carbon atom. The various pathways of the carbon atom are in agreement with the generally accepted mechanism of methane partial oxidation. There is only one exception, the route from formaldehyde to methyl radical. (The peculiarity of this route lies in the fact that an oxidized intermediate is turned into a pyrolytic intermediate. Usually the opposite reaction is accepted only). There is a single reaction realizing this pathway



which is usually not considered as a possible reaction in such mechanisms. However, it passes all constraints with the adopted limits (see Table 2). This does not mean that this reaction must occur in reality, nor that it is an important reaction. It only means that the possibility of this reaction cannot be ruled out based on the above principles and limits. (It should be noted that employing filtering with $\chi \leq 1$ would lead to the omission of this reaction.)

The power of the method lies in its ability to suggest unorthodox reactions (when employing the filters with due care), which would provide thinking material for the experimentalist.

6. Tips and Tricks. It is strongly recommended to keep the number of species in a system to a minimum. Otherwise combinatorial explosion will cause a large number of reactions generated. Although the generation is quite fast, it

Table 2. Stoichiometrically Possible Reactions between Oxygen and Formaldehyde and Their Characteristic Values Used for Filtering^a

| no. | products | | | κ | χ | $\Delta H_{r,T}$, kJ·mol ⁻¹ | $\Delta H_{rc,T}$, kJ·mol ⁻¹ | $\Delta G_{r,T}$, kJ·mol ⁻¹ | type of reaction |
|-----|-----------------|------------------|------------------|----------|--------|---|--|---|------------------|
| 1 | HO· | HO· | CO | 5 | 2 | 87 | 287 | -142 | initiation |
| 2 | HO· | CO ₂ | H· | 5 | 2 | -15 | 185 | -172 | initiation |
| 3 | HO· | O | ·CHO | 3 | 1 | 455 | 575 | 214 | initiation |
| 4 | CO | O | H ₂ O | 5 | 2 | 16 | 216 | -184 | initiation |
| 5 | CO ₂ | H ₂ O | | >5 | >2 | -527 | na | -489 | molecular |

^a Thermochemical parameters are calculated at $T = 1500 \text{ K}$.

is memory (hard disk space) consuming. For modeling of reactions of small molecules any personal computer is appropriate. However, calculating κ for reactions involving large molecules (ring systems, number of atoms in a species exceeding 20) can be very time-consuming even on extremely fast computers. All visualization tools (sequence networks, pathways, etc.) require the previous calculation of κ (this procedure determines which atoms, or groups of atoms are transferred from one species to the other). Thermodynamic calculations are fairly fast even for large reaction mechanisms (hundreds of thousands of reactions).

At present the computer code is limited to using no more than 25 identical atoms in a species (practically this limits only the number of hydrogens). This limitation can easily be overcome using parts of molecules, which are not participating in a reaction. In such cases substructures can be defined as species A, B, etc. However, such substructures cannot undergo chemical transformation.

If MECHGEN generates too large numbers of reactions, then decomposing the entire reaction set into small(er) subsets might be a reliable alternative.

If a reaction is filtered out although we are sure that it takes place (usually this happens when filtering for κ or χ), definition of a new, intermediate species might be the solution.

It is advised to make the generation and reduction process iteratively. Starting with a small number of species and completing the species space with new intermediates gradually will show interesting details of the mechanism.

The program is available via electronic mail from Károly Héberger <heberger@chemres.hu>.

7. Basic Appearance, Language. The program is written in Borland PASCAL code and runs under DOS to facilitate easy transfer between the various platforms. Its graphical user interface is menu-driven and provides a WINDOWS-like environment. Inputs and outputs are handled via dialogue boxes. All controls are available through standard keyboard and mouse operations.

CONCLUSIONS

A computer code has been developed that can be used to generate and reduce reaction mechanisms. The key feature of the approach is that the user must define the complete species list (including all intermediates and end products). The species structure should be given together with its thermodynamic characteristics. Species library can be built. All stoichiometrically possible reactions are generated. The long reaction list can be reduced using mechanistic and thermochemical constraints (reaction complexities, reaction enthalpy, and Gibbs free energy). Temperature dependence of the thermodynamic quantities has been taken into account in the form of polynomial representation of the heat capacity. Visualization of the reaction pathways has been realized using the concept of reaction sequence and kinetic communication.

ACKNOWLEDGMENT

This work was supported by the Hungarian Research Foundation: OTKA T 029748. We are indebted to the referee Prof. A. V. Zeigarnik for his useful suggestions. A. Németh and K. Héberger also acknowledge the financial support of

DFG and thank Prof. Dr. H. Gg. Wagner and Dr. H. Jander for valuable discussions on the application of the program to evaluate reaction mechanism of soot precursor formation.

REFERENCES AND NOTES

- (1) Barone, R.; Chanon, M.; Green, M. L. H. Possible Mechanisms for Ethylene Dimerization Derived by Use of a New Computer Program. *Organomet. Chem.* **1980**, *185*, 85–93.
- (2) Theodosiou, I.; Barone, R.; Chanon, M. New Computer Program Able to Propose Mechanistic Schemes and Secondary Products for Transformations Catalyzed by Transition Metal Complexes: Possible Mechanisms for Hydroformylation of Ethylene. *J. Mol. Catal.* **1985**, *32*, 27–50.
- (3) Theodosiou, I.; Barone, R.; Chanon, M. Computer Aids for Organometallic Chemistry and Catalysis. *Adv. Organomet. Chem.* **1986**, *26*, 165–216.
- (4) Clymans, P. J.; Froment, G. F. Computer-Generation of Reaction Paths and Rate Equations in the Thermal Cracking of Normal and Branched Paraffins. *Comput. Chem. Eng.* **1984**, *8*, 137–142.
- (5) Baltanas, M. A.; Froment, G. F. Computer Generation of Reaction Networks and Calculation of Product Distributions in the Hydroisomerization and Hydrocracking of Paraffins on Pt-containing Bifunctional Catalysts. *Comput. Chem. Eng.* **1985**, *9*, 71–81.
- (6) Fontain, E.; Reitsam, K. The Generation of Reaction Networks with RAIN. 1. The Reaction Generator. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 96–101.
- (7) Fontain, E. The Generation of Reaction Networks with RAIN. 2. Resonance Structures and Tautomerism. *Tetrahedron Comput. Methodol.* **1990**, *3*, 469–477.
- (8) Likholobov, V. A.; Mayer, L. I.; Bulgakov, N. N.; Fedotov, A. V.; Shteingauer, L. G. Computer-Aided Prognosis of the Catalytic Properties of Metal Atoms in the Gas-Phase Hydrogenation of Carbon Monoxide. In *Homogeneous and Heterogeneous Catalysis*; Yermakov, Yu., Likholobov, V. A., Eds.; Inst. of Catalysis: Novosibirsk, 1986; pp 229–244.
- (9) Di Maio, F. P.; Lignola, P. G. A Kinetic Network Generator. *Chem. Eng. Sci.* **1992**, *47*, 2713–2718.
- (10) Chinnick, S. J.; Baulch, D. L.; Ayscough, P. B. An expert system for hydrocarbon pyrolysis reactions. *Chemometrics Intelligent Laboratory Systems* **1988**, *5*, 39–52.
- (11) Chevalier, C.; Warnatz, J.; Melenk, H. Automatic generation of reaction mechanisms for the description of higher hydrocarbons. *Ber. Bunsenges. Phys. Chem.* **1990**, *94*, 1362–1367.
- (12) Gasteiger, J.; Jochum, C. EROS/A Computer Program for Generating sequences of Reactions. *Topics. Curr. Chem.* **1978**, *74*, 93–126.
- (13) Gasteiger, J.; Hondelmann, U.; Rose, P.; Witzendichler, J. Computer-assisted prediction of degradation of chemicals: Hydrolysis of amides and benzoylphenylureas. *J. Chem. Soc., Perkin Trans. 2* **1995**, 193–204.
- (14) Peters, N.; Williams, F. A. The Asymptotic Structure of stoichiometric methane/air flames. *Combust. Flame* **1987**, *68*, 185–207.
- (15) Peters, N. Reducing Reaction Mechanisms. In *Lecture Notes in Physics, Reducing Kinetic Mechanisms and Asymptotic Approximation for Methane/air Flames*; Smoke, M. D., Ed.; Springer: Berlin, 1991; pp 48–67.
- (16) Peters, N.; Rogg, B. *Reduced Kinetics Mechanisms for Applications in Combustion Systems*; Springer: New York, 1993.
- (17) Come, G. M.; Sacchi, G.; Warth, V.; Michel, V.; Battin-Leclerc, F.; Bizot, P. Computer Aided Design of Gas-phase Free Radical Reaction Mechanism. In *Oxidation Kinetics and Autoignition of Hydrocarbons*; Pilling, M. J., Ed.; Elsevier: 1995; p 308.
- (18) Warth, V.; Stef, N.; Glaude, P. A.; Battin-Leclerc, F.; Scacchi, G.; Côme, G. M. Computer-aided Derivation of Gas-phase Oxidation Mechanisms. *Combust. Flame* **1998**, *114*, 81–102.
- (19) Lam, S. H.; Goussis, D. A. Conventional Asymptotics and Computational Singular Perturbation. In *Lecture Notes in Physics, Reducing Kinetic Mechanisms and Asymptotic Approximation for Methane/air Flames*; Smoke, M. D., Ed.; Springer: Berlin, 1991; pp 193–226.
- (20) Seshadri, K.; Bollig, M.; Peters, N. Numerical and Asymptotic Studies of Structure of Stoichiometric and Lean Premixed Heptane Flames. *Combustion Flame* **1997**, *108*, 518–536.
- (21) Maas, U.; Pope, S. B. Simplifying chemical kinetics: Intrinsic Low-dimensional Manifolds in Composition Space. *Combust. Flame* **1992**, *88*, 239–264.
- (22) Eggels, R. L. G. M.; de Goey, L. P. H. Mathematically Reduced Reaction Mechanisms Applied to Adiabatic Flat Hydrogen/air Flames. *Combust. Flame* **1995**, *100*, 559–569.
- (23) Bruk, L. G.; Gorodskii, S. N.; Zeigarnik, A. V.; Valdes-Perez, R. E.; Temkin, O. N. Mechanism of Oxidative Carbonylation of Phenylacetylene and Methylacetylene: Generation of Hypotheses and Their Experimental Discrimination. *Russ. Chem. Bull.* **1999**, *48*, 873–880.

- (24) Zeigarnik, A. V.; Bruk, L. G.; Temkin, O. N.; Likholobov, V. A.; Maier, L. I. Computer-Assisted Mechanistic Studies. *Usp. Khim.* **1996**, 65, 125–139; [*Russ. Chem. Rev.* **1996**, 6, 117–130].
- (25) Shalgunov, S. I.; Zeigarnik, A. V.; Bruk, L. G.; Temkin, O. N. Computer-assisted Generation of Chemical Reaction Networks: Applications in the Mechanistic Studies of Organometallic Catalysis. *Russ. Chem. Bull.* **1999**, 48, 1876–1874.
- (26) Bruk, L. G.; Gorodskii, S. N.; Zeigarnik, A. V.; Valdes-Perez, R. E.; Temkin, O. N. Oxidative Carbonylation of Phenylacetylene Catalyzed by Pd(II) and Cu(I): Experimental Tests of Forty-one Computer-Generated Mechanistic Hypotheses. *J. Mol. Catal., A: Chemical* **1998**, 130, 29–40.
- (27) Zeigarnik, A. V.; Valdes-Perez, R. E.; Temkin, O. N.; Bruk, L. G.; Shalgunov, S. I. Computer-Aided Mechanism Elucidation of Acetylene Hydrocarboxylation to Acrylic Acid Based on a Novel Union of Empirical and Formal Methods. *Organometallics* **1997**, 16, 3114–3127.
- (28) Porollo, A. A.; Lushnikov, D. E.; Pivina, T. S.; Ivshin V. P.; Computer Representation and Generation of Possible Pathways for Thermal Decomposition Reactions of Organic Compounds. *J. Mol. Struct. (Theochem)* **1997**, 391, 117–124.
- (29) Porollo, A. A.; Lushnikov, D. E.; Pivina, T. S.; Ivshin, V. P.; Zefirov, N. S. Computer simulation of thermal decomposition reactions of alkyl nitrates. *Russ. Chem. Bull.* **1999**, 48, 1845–1857.
- (30) Yoneda, Y. A Computer Program Package for the Analysis, Creation, and Estimation of Generalized Reactions/GRACE. I. Generation of Elementary Reaction Network in Radical Reactions. *Bull. Chem. Soc. Jpn.* **1979**, 52, 8–14.
- (31) Nemes, I.; Vidóczy, T.; Botár, L.; Gál, D. A Possible Construction of a Complex Chemical Reaction Network I. Definitions and Procedure for Construction. *Theor. Chim. Acta* **1977**, 45, 215–223.
- (32) Even though this definition does not consider certain concerted reactions as elementary reactions, it agrees with common chemical considerations very well.
- (33) Héberger, K.; Keszler, Á.; Vidóczy, T.; Gál, D.; Cotarca, L.; Delogu, P. Oxidative decarboxylation of cyclohexane monocarboxylic acid as a degenerate branching chain reaction II. Mechanism. *Ber. Bunsenges. Phys. Chem.* **1994**, 98, 1303–1307.
- (34) There is no strict correspondence between subsets of reactions satisfying different limits in κ and χ . A good example is that most of the reactions fulfilling the condition $\chi \leq 1$ fulfill also the condition $\kappa \leq 4$; however, it is easy to find reactions where $\chi = 1$ and $\kappa > 4$ (in other words $PM_{(\chi \leq 1)}$ is not a subset of $PM_{(\kappa \leq 4)}$).
- (35) The following criteria must be fulfilled when comparing reactions: (i) all the reactants should be the same; (ii) the type of the reaction (e.g. in case of free radical processes initiation, propagation, termination) should be identical; and (iii) the molecularity of the reaction on the product side should be identical. The justification of these requirements is discussed in ref 36.
- (36) Vidóczy, T.; Hajdu, I. P.; Nemes, I.; Gál, D. A Possible Construction of a Complex Chemical Reaction Network IV. Compilation of the Possible Mechanism by Combinatorial Method and its Reduction by Thermochemical Kinetic Treatment. *Oxid. Commun.* **1983**, 4, 203–221.
- (37) Németh, A.; Héberger, K. Generation and Reduction of the Reaction Mechanism Representing the Initial Attack upon Methane in Combustion. *Oxid. Commun.* **1996**, 19, 467–475.
- (38) Nemes, I.; Vidóczy, T.; Gál, D. A Possible Construction of a Complex Chemical Reaction Network III. The Systematization of the Elementary Processes. *Theor. Chim. Acta* **1977**, 46, 243–250.
- (39) Thermochemical data like standard enthalpy and entropy of formation, coefficients of temperature dependence of the heat capacity at constant pressure (C_p) can be entered. The latter one can be given in two polynomial forms at the user's option. One is the third-order polynomial expression recommended⁴⁰ above 300 K: $C_p = a + bT + cT^{-2}$. The other one is the fifth order NASA polynomial with two different sets of coefficients, one below 1000 K, the other between 1000 and 5000 K: $C_p/R = a' + b'T + c'T^2 + d'T^3 + e'T^4$, where R is the gas constant. Thermochemical data of the present species library (above 300 species) were computed by the programs THERM^{41,42} and THERGAS⁴³.
- (40) Stull, D. R.; Westrum, E. F., Jr.; Sinke, G. C. *The Chemical Thermodynamics of Organic Compounds*; John Wiley & Sons: New York, 1969.
- (41) Ritter, E. R.; Bozelli, J. W. THERM: Thermodynamic property estimation for gas-phase radicals and molecules. *Int. J. Chem. Kinet.* **1991**, 23, 767–778.
- (42) Ritter, E. R. THERM: A computer code for estimating thermodynamic properties for species important to combustion and reaction modeling. *J. Chem. Inf. Comput. Sci.* **1991**, 31, 400–408.
- (43) Müller, C.; Michael, V.; Scacchi, G.; Côme, G. M. THERGAS, a computer program for the evaluation of thermochemical data of molecules and free radicals in the gas phase. *J. Chim. Phys.* **1995**, 92, 1154–1178.
- (44) In case a species is defined as “reactant only”, it cannot occur on the right-hand side of a reaction. Similarly, species defined as “product only” cannot be taken into consideration on the left-hand side of a reaction.
- (45) This is the only filter where reactions are compared to each other (all other filtering is based on the reaction in question alone). Should later filtering result in omission of a reaction, which served as the reference reaction in a subgroup (because being the most exothermic/least endothermic within that subgroup), would render omission of reactions based on their reaction enthalpy from that given subgroup invalid. To avoid such ambiguities, filtering based on reaction enthalpy should be performed as the last filtering procedure.

CI010365I