Implementation of stereochemistry in automatic kinetic model generation

Nick M. Vandewiele[b], Ruben Van de Vijver[a], Kevin M. Van Geem[a][*], Marie-Françoise Reyniers[a] and Guy B. Marin[a]

[a] Laboratory for Chemical Technology, Universiteit Gent, Technologiepark 914, 9052 Ghent, Belgium
[b] Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract

This work discusses the extension of an existing network generation tool, called Genesys, by accounting for stereochemistry in kinetic models. Genesys was extended by the addition of algorithms for the detection of stereocenters, and for the identification of stereoisomers. To that purpose a so-called 2.5D representation of molecules that accounts for the presence of stereocenters and that keeps track of the associated stereoconfiguration of the stereocenters is introduced. A novel algorithm for the detection of steric relations between substituents allows the automated assignment of rate coefficients to stereoselective reactions. The functionality of the tool is illustrated for the thermal rearrangement of the monoterpenoid 2-pinanol, in which accounting for the stereochemistry is crucial to explain the reactive behavior of 2-pinanol and its products. Good agreement with experimental data was obtained.

Introduction

During the last decades, new applications involving stereochemistry, such as asymmetric synthesis of pharmaceuticals and agrochemicals (Noyori, 2002), and conversion processes of natural products such as terpenes (Corma et al., 2007) emerged for which the characterization of stereochemistry is required to understand the reactivity of the molecules. In this context, computer-aided modeling of these chemical processes requires an adequate representation of stereochemistry in order to capture the essential characteristics of the chemical process. A valuable tool to study chemical processes are kinetic models, which give quantitative insights in the product distribution as a function of reaction conditions. For many chemical processes manual construction of kinetic models is no longer possible because of their complexity. To cope with this, computers were programmed to automatically generate kinetic models (Ratkiewicz and Truong, 2006; Pierucci and Ranzi, 2008; Battin-Leclerc et al., 2011). Many tools have been proposed to automatically create kinetic models (Warth et al., 2000; Vandewiele et al., 2012; Green et al., 2013), but none of them account for stereochemistry. The reason for the lack of support for stereochemistry in the generated kinetic models was and is the underlying representation of molecules in these codes. Molecules are often represented as mathematical graphs, in which the nodes of the graph represent atoms, and the edges represent bonds. By doing so, a plethora of graph-theoretic algorithms and solutions have been proposed and implemented to accomplish tasks such as identification of species uniqueness, substructure searching, greatly facilitating the efforts of extracting knowledge from chemical data by means of computers. Unfortunately, the graph representation also implied that the characterization and manipulation of stereochemical features of molecules and reactions became impossible since graphs only represent the connectivity between atoms, and not the arrangement in the three-dimensional space.

This work discusses the necessary steps to automatically create kinetic models that account for stereochemistry. The functionality of the new tool is illustrated for the thermal rearrangement of the monoterpenes cis- and trans-2-pinanol. A kinetic model is constructed for this chemical process that accounts for the stereochemistry and is compared to experimental data obtained from literature.

Methodology

Genesys (Vandewiele et al., 2012) is a tool for the automatic generation of kinetic models, consisting of a reaction network containing molecules and elementary reactions together with thermodynamic and kinetic data. The reaction network is generated by the iterative application of a limited set of reaction families that convert reactant molecules into product molecules. A pool of species, e.g. the reactants of a chemical process, is used to initiate the network generation and the species list is continuously extended with species that arise as products of elementary reactions.

For kinetic models that want to take into account stereochemical effects, a number of additional aspects need to be considered. First of all, the data structures designed to represent molecules inside the network generation code should allow distinguishing between stereoisomers. Information on stereochemical aspects of molecules needs to be correctly converted from the user interface into an internal molecule representation, and vice versa. Alternatively, if the user specifies a reactant in which the absolute configuration of the stereocenters is not uniquely defined, Genesys identifies these stereocenters through an adequate stereocenter detection algorithm, and exhaustively generates all possible stereoisomers using a so-called stereoisomer generation algorithm.

Secondly, continuous updating of stereocenters in reactant stereoisomers is needed

because of the formation of new stereocenters in product stereoisomers may arise in the course of the network generation. Furthermore, a reaction may lead to multiple distinct stereoisomers. The same stereocenter detection algorithm and stereoisomer generation algorithm used for the exhaustive generation of all stereoisomers corresponding to the reactant structure is now used to handle the newly created products of a reaction.

Finally, reactions may be stereoselective or stereospecific. Stereoselectivity refers to the preferential formation of or the other of two (or more) stereoisomers from a single molecule with a prostereogenic element (Eliel, 1962). The latter refers to an element that can be converted from non-stereocenter to stereocenter in a single step (McNaught and Wilkinson, 1997). Stereospecificity refers to the difference in reaction rates of two stereochemically different molecules, i.e. diastereomers or enantiomers (Eliel, 1962). The stereospecificity of a reaction is sometimes linked to the absolute stereoconfiguration of stereocenters and is explained by the lock-and-key specificity of catalysts such as enzymes. Stereoselectivity depends on the differences in free energies of the respective transition states and is explained by stereoelectronic and steric factors among others.

An algorithm is constructed that identifies the steric relation of substituents in molecules allowing the introduction of stereoselective and stereospecific rate coefficients for a reaction. After a reaction creates the product structures, containing newly created stereocenters with unspecified absolute configurations, the stereoisomer generation algorithm determines the possible stereoisomers. If the user disposes of information on the stereoselectivity or stereospecificity of a reaction family, Genesys allows the assignment of distinct rate coefficients for these reactions, based on the steric relation between the user-specified substituents. Once the steric relation between the designated substituents of the product or reactant stereoisomers is detected, prefixes such as *cis* and *trans* can be assigned to the different reactants or products. For each of the elementary reactions that convert a reactant into a possible stereoisomer a distinct rate coefficient is assigned, which is determined by the steric relation that is detected in the previous step. The rate coefficient that is associated with a particular steric relation is retrieved from the user-defined reaction family definition.

Pinanol rearrangements

The method and implementation of kinetic model generation incorporating stereochemistry is demonstrated and validated for the thermal rearrangements of the monoterpenoid 2-pinanol, shown in Figure 1.This compound belongs to the terpenoids, a highly diverse and functional class of biomass-derived isoprene oligomers. 2-pinanol is heavily used in the fragrance, flavor and pharmaceutical industry

(Mercier and Chabardes, 1995; Nowicki, 2000; Swift, 2004; Corma et al., 2007). The thermal rearrangement of 2-pinanol is an important industrial production route for linalool, also depicted in Figure 1, used in perfumes and as a precursor for vitamins A and E (Mercier and Chabardes, 1995).
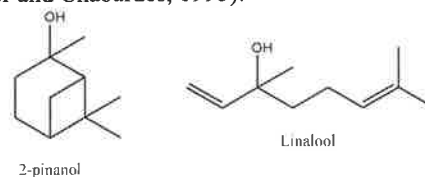

Figure 1: Structures of 2-pinanol and linalool.

There are four stereoisomers of 2-pinanol. They relate to each other as two pairs of diastereomers, depending on the spatial arrangement of the dimethyl bridge relative to the hydroxyl group. Each pair of diastereomers relates to each other as enantiomers. Previous work (Vandewiele et al., 2011) showed that significant differences in pyrolysis reactivity and selectivity could be observed between the two diastereomeric pairs. Furthermore, distinguishing between the enantiomers of linalool is important because each enantiomer has distinct odor characteristics and thresholds (Ohloff and Klein, 1962; Fritsch and Schieberle, 2005). The structures of the product species arising during the pyrolysis of 2-pinanol are shown in Figure 2.
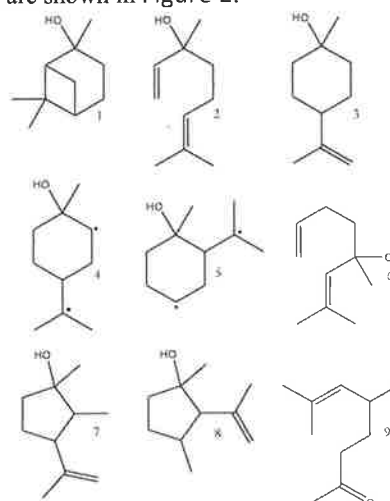

Figure 2: Chemical structures of compounds involved in the thermal rearrangement of 2-pinanol: 1: 2-pinanol, 2: linalool, 3: ϑ-terpineol, 4: 1,4-disubstituted cyclohexanol biradical, 5: 1,2-disubstituted cyclohexanol biradical, 6: isolinalool, 7: plinol, 8: isoplinol, 9: 5,7-dimethyloct-6-ene-2-one.

The procedure of the automated kinetic model construction proceeds as follows: 2-pinanol is specified as reactant using a species identifier such as an InChI without specifying the absolute configurations of the stereocenters. The stereoisomer generation algorithm generates four distinct structures. The reaction pathways that are relevant for the thermal decomposition of 2-pinanol were described in detail in

previous work (Vandewiele et al., 2011). The chemical knowledge on the relevant pathways involved in the thermal rearrangement of 2-pinanol is translated into reaction families. Four reaction families are defined: homolytic C-C scission reactions, biradical β-scission reactions, sigmatropic [1,5]-H-shift reactions and ene-cyclization reactions. The fragmentation of the four-membered ring by homolytic C-C scission reactions is the initial step in the thermal isomerization of bicyclic monoterpenes consisting of a bicyclo[3.1.1]heptane system. This scission reaction leads to four 1,4-disubstituted cyclohexanol biradicals. The 1,4-disubstituted biradicals quickly rearrange into linalool enantiomers through C-C β-scission reactions involving the bond in β-position of the two carbon atoms bearing the unpaired electron. An alternative, minor pathway leads β-terpineol, via sigmatropic [1,5]-H-shift reactions. Linalool further isomerizes into plinol products via pericyclic ene-cyclization reactions. Note that the application range of a reaction family needs to be constrained as much as possible so only molecules are considered that can react based on this reaction type.

The calculation of thermochemical properties and symmetry numbers of the species in the model follows the same methodology as described in previous work (Vandewiele et al., 2012; Vandewiele et al., 2014). The assignment of the rate coefficients to each elementary reaction occurs on the level of reaction families in Genesys. Ideally, ΔGAV°s obtained from *ab initio* calculations are used to calculate values for the rate coefficients of the generated reactions. Unfortunately, values for ΔGAV°s for the reaction families required for the thermal decomposition of 2-pinanol are unavailable. Instead, fixed Arrhenius parameters were used for the generated reactions of each reaction family, derived from experiments. The provided pre-exponential factor for a reaction family is a single-event pre-exponential factor that is multiplied by the number of single-events of that particular reaction. The number of single-events is calculated following the algorithm described by Vandewiele et al. (Vandewiele et al., 2014).

The activation energy of the rate coefficient of the homolytic C-C scission reactions of 2-pinanol is 5 kJ mol^{-1} lower when the methyl group of the chiral carbon atom is in *cis* position relative to the dimethyl bridge of the bicyclo[3.1.1]heptane system, compared to *trans*-structure in which the methyl group is in *trans* position relative to the dimethyl bridge. For the homolytic C-C scission reactions, the influence of the arrangement of the methyl group next to the chiral carbon with respect to the dimethyl bridge on the rate coefficients is added as a further specification to the reaction family. The steric relation detection algorithm enables the automatic assignment of the appropriate rate coefficient for the corresponding stereoisomer.

Similarly, rate coefficients were also derived for the β-scission reactions, the H-shift reactions, and the ene-cyclization reactions (Vandewiele et al., 2011).

Results and discussion

The generated kinetic model for the thermal rearrangement of 2-pinanol consists of only 20 elementary reactions between 20 species. The presented kinetic model is an improvement of the kinetic model of Vandewiele et al. (Vandewiele et al., 2011) on 3 levels. First, this new kinetic model is automatically constructed, as opposed to the previous model that was constructed by hand. Second, the new kinetic model distinguishes between enantiomers while the hand-built model lumps them together. Last but not least, the model presented in this chapter accounts for the stereoselectivity of the ene-cyclization of linalool, by incorporating distinct reactions to each of the stereoisomers of plinol. The previously build model only provides a lumped reaction in which a global reaction rate is provided towards the lumped plinol isomers.

The predictions of the generated kinetic model are compared against experimental data by Leiner et al. (Leiner et al., 2013). The reactor in the experiments consists of a 200 mm long quartz tube with a diameter of 15 10-3 m. 5000 ppm of the reactant was fed to the setup in a N_2 flow at temperatures between 623 and 873K and a pressure of 1bar. Conversions ranging from 0 to 100% were obtained with these conditions. Species at the reactor outlet are identified and quantified using GC-FID and GC-MS. Simulations were carried using the plug flow reactor model of the Chemkin 4.1 package (Kee et al., 2007), using the reported reactor dimensions and conditions.
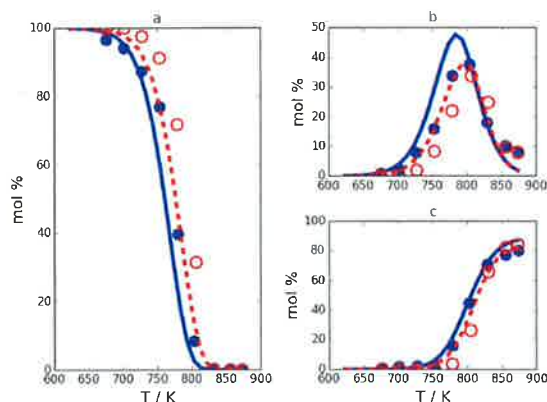

Figure 3: Mole fraction as a function of reactor tempearture of a) 2-pinanol, b) linalool, c) sum of isomers of plinol. Model predictions: cis-2-pinanol as feed (blue, full lines), trans-2-pinanol as feed (red, dashed lines). Experiments: cis-2-pinanol as feed (blue, full symbols), trans-2-pinanol as feed (red, hollow symbols).

Figure 1 shows the mole fractions as a function of the reactor temperature for 2-pinanol, linalool and plinol for experiments with cis- and trans-1 as the feed. Since the feed was not enantiomerically pure, the mole fractions of enantiomeric species in the model were lumped together and the sum of the mole fractions of both enantiomers for the experiments and the model predictions was compared instead. Also,

measured concentrations of minor products were not reported, and thus could not be used for comparison. The conversion of both diastereomers of 2-pinanol is slightly overestimated, resulting in a small overestimation of the mole fractions of linalool. Overall, good agreement was found between model and experiment, especially given that none of the parameters in the model were fitted to the experimental data.

Conclusion

This work presents the extension of Genesys for the automatic construction of kinetic models for molecules and reactions that account for stereochemistry. It uses a 2.5D representation of molecules, i.e. a graph representation augmented with so-called stereoparities for the stereocenters of the molecules. Genesys keeps track of the absolute configurations of existing stereocenters, the creation or destruction of stereocenters, and the generation of stereoisomers.

The functionality of the tool was illustrated by the automated construction of a kinetic model for the thermal rearrangement of 2-pinanol. Using a stereoisomer generation algorithm, the four existing stereoisomers of 2-pinanol were generated from a structure in which the absolute configuration of the stereocenters was not specified. The trans-linked pinanol isomers were identified using the methodology for the detection of the relative arrangements of substituents of stereocenters. Four reaction families were defined and used to generate the kinetic model for 2-pinanol. The model was validated using experimental data from literature and showed the good agreement between the model predictions and the measured concentrations of the reactant and products without adjustment of any of the reaction rate coefficients.

References

Battin-Leclerc, F., Blurock, E., Bounaceur, R., Fournet, R., Glaude, P.A., Herbinet, O., Sirjean, B., and Warth, V. (2011). Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models. *Chem. Soc. Rev.* 40, 4762-4782.

Corma, A., Iborra, S., and Velty, A. (2007). Chemical Routes for the Transformation of Biomass into Chemicals. *Chemical Reviews* 107, 2411-2502.

Eliel, E.L. (1962). Stereochemistry of Carbon Compounds. New York.

Fritsch, H.T., and Schieberle, P. (2005). Identification Based on Quantitative Measurements and Aroma Recombination of the Character Impact Odorants in a Bavarian Pilsner-type Beer. *Journal of Agricultural and Food Chemistry* 53, 7544-7551.

Green, W.H., Allen, J.W., Ashcraft, R.W., Beran, G.J., Class, C.A., Gao, C., Goldsmith, C.F., Harper, M.R., Jalan, A., Magoon, G.R., Matheu, D.M., Merchant, S.S., Mo, J.D., Petway, S., Raman, S., Sharma, S., Song, J., Van Geem, K.M., Wen, J., West, R.H., Wong, A., Wong, H.S., Yelvington, P.E., and Yu, J. (2013). RMG - Reaction Mechanism Generator v4.0.

Kee, R.J., Rupley, F.M., Miller, J.A., Coltrin, M.E., Grcar, J.F., Meeks, E., Moffat, H.K., Lutz, G., Dixon-Lewis, A.E., Smooke, M.D., Warnatz, J., Evans, G.H., Larson, R.S., Mitchell, R.E., Petzhold, L.R., Reynolds, W.C., Caracotsios, M., Stewart, W.E., Glarborg, P., Wang, C., Adigun, O., Houf, W.G., Chou, C.P., Miller, S.F., Ho, P., and Young, D.J. (2007). CHEMKIN Release 4.1.1

Leiner, J., Stolle, A., Ondruschka, B., Netscher, T., and Bonrath, W. (2013). Thermal behavior of pinan-2-ol and linalool. *Molecules* 18, 8358-8375.

McNaught, A.D., and Wilkinson, A. (1997). Compendium of Chemical Terminology. (Oxford: Blackwell Science).

Mercier, C., and Chabardes, P. (1995). Organometallic Chemistry in Industrial Vitamin-A and Vitamin-E Synthesis. *Pure Appl. Chem.* 66, 1509-1518.

Nowicki, J. (2000). Claisen, Cope and related rearrangements in the synthesis of flavour and fragrance compounds. *Molecules* 5, 1033-1050.

Noyori, R. (2002). Asymmetric catalysis: science and opportunities (Nobel lecture). *Angewandte Chemie* 41, 2008-2022.

Ohloff, G., and Klein, E. (1962). Die absolute konfiguration des linalools durch verknüpfung mit dem pinansystem. *Tetrahedron* 18, 37-42.

Pierucci, S., and Ranzi, E. (2008). A review of features in current automatic generation software for hydrocarbon oxidation mechanisms. *Computers & Chemical Engineering* 32, 805-826.

Ratkiewicz, A., and Truong, T.N. (2006). Automated mechanism generation: From symbolic calculation to complex chemistry. *International Journal of Quantum Chemistry* 106, 244-255.

Swift, K.A.D. (2004). Catalytic transformations of the major terpene feedstocks. *Topics in Catalysis* 27, 143-155.

Vandewiele, N.M., Van Geem, K.M., Reyniers, M.F., and Marin, G.B. (2011). Kinetic study of the thermal rearrangement of cis- and trans-2-pinanol. *J. Anal. Appl. Pyrolysis* 90, 187-196.

Vandewiele, N.M., Van Geem, K.M., Reyniers, M.-F., and Marin, G.B. (2012). Genesys: Kinetic model construction using chemo-informatics. *Chem Eng J* 207, 526-538.

Vandewiele, N.M., Van de Vijver, R., Van Geem, K.M., Reyniers, M.-F., and Marin, G.B. (2014). Symmetry calculation for molecules and transition states. In *Journal of Computational Chemistry*, 36, 181-192.

Warth, V., Battin-Leclerc, F., Fournet, R., Glaude, P.A., Côme, G.M., and Scacchi, G. (2000). Computer based generation of reaction mechanisms for gas-phase oxidation. *Comput Chem* 24, 541-560.