

**Automatische ab-initio berekeningen
voor het genereren van kinetische modellen van gasfaseprocessen**

**Automatic Ab Initio Calculations
for Kinetic Model Generation of Gas-Phase Processes**

Ruben Van de Vijver

Promotoren: prof. dr. ir. K. Van Geem, prof. dr. ir. G. B. Marin
Proefschrift ingediend tot het behalen van de graad van
Doctor in de ingenieurswetenschappen: chemische technologie



Vakgroep Materialen, Textiel en Chemische Proceskunde
Voorzitter: prof. dr. P. Kiekens
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2016 - 2017

ISBN 978-90-8578-998-7
NUR 952
Wettelijk depot: D/2017/10.500/33

Promotors

Prof. Dr. Ir. Guy B. Marin

Laboratorium voor Chemische Technologie

Vakgroep Materialen, Textiel en Chemische Proceskunde

Universiteit Gent

Prof. Dr. Ir. Kevin M. Van Geem

Laboratorium voor Chemische Technologie

Vakgroep Materialen, Textiel en Chemische Proceskunde

Universiteit Gent



UNIVERSITEIT
GENT

 FACULTEIT INGENIEURSWETENSCHAPPEN
EN ARCHITECTUUR

The author acknowledges financial support from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT Vlaanderen) and the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 290793.

Examencommissie

Prof. Dr. Marie-Françoise Reyniers
Laboratorium voor Chemische Technologie
Vakgroep Materialen, Textiel en Chemische Proceskunde
Universiteit Gent

Prof. Dr. Ir. Maarten K. Sabbe [secretaris]
Laboratorium voor Chemische Technologie
Vakgroep Materialen, Textiel en Chemische Proceskunde
Universiteit Gent

Prof. Dr. John Simmie
Combustion Chemistry Centre
National University of Ireland

Prof. Dr. Ir. Toon Verstraelen
Centrum voor Moleculaire Modelling
Vakgroep Fysica en Sterrenkunde
Universiteit Gent

Dr. Judit Zádor
Combustion Research Facility
Sandia National Laboratories

Prof. Dr. Ir. Gert de Cooman [voorzitter]

SYSTeMS research group

Vakgroep voor Elektronica en Informatiesystemen

Universiteit Gent

Prof. Dr. Ir. Guy B. Marin [promotor]

Laboratorium voor Chemische Technologie

Vakgroep Materialen, Textiel en Chemische Proceskunde

Universiteit Gent

Prof. Dr. Ir. Kevin M. Van Geem [promotor]

Laboratorium voor Chemische Technologie

Vakgroep Materialen, Textiel en Chemische Proceskunde

Universiteit Gent

Acknowledgments

Eerst en vooral wil ik mijn richten tot Kevin. Bedankt om mij het vertrouwen te geven een doctoraat te starten in jouw groep, en voor de jarenlange begeleiding zowel op wetenschappelijk vlak, maar ook voor de hulp in de praktische kant van de academische wereld, het helpen bij het schrijven van abstracts, papers en dit werk en de goede samenwerking en verstandhouding tijdens mijn doctoraat. Ik hoop dat we in de toekomst op hetzelfde elan verder kunnen samenwerken. Daarnaast wil ik ook Guy bedanken om mij de mogelijkheid te geven aan het LCT mijn doctoraat af te leggen en om mijn werk kritisch en gedetailleerd te bekijken.

Alhoewel mijn twee promotoren een prominente rol hebben gespeeld tijdens de voorbije vier jaar, heb ik ook veel hulp gekregen van andere collega's. Maarten zou ik willen bedanken voor de wegwijs in de statistische thermodynamica, de lunchpauzes die we hebben gespendeerd terwijl tunnelingscoëfficiënten werden besproken, de zoektocht naar traagheidsmomenten, etc.

Also many thanks to Hans-Heinrich, who was always prepared to discuss my work in great detail and help me with the *ab initio* calculations and the revision of several of my texts.

Daarnaast zijn er nog vele mensen aan het LCT die ik zou willen bedanken. Het zijn er te veel om op te noemen, maar in het bijzonder wil ik Pieter, Natalia en Florence bedanken, ik kan me geen beter bureau inbeelden. Ook Ruben, Andres, Carl, David, Pieter, Ismaël, Nick, en Thomas verdienen hier een vermelding.

Ik heb tijdens mijn doctoraat het genoeg gehad drie uitstekende master thesis studenten te begeleiden. Jessica, zonder jou was Genesys slechts rule-based, bedankt om dit ook rate-based te maken. Lennert, een fundament van mijn doctoraat, namelijk de geometrie van een transitietoestand schatten, heb ik aan jou te danken. En Cato, het resultaat van jouw thesis valt nog af te wachten, maar ik ben er zeker van dat het modeleren van de pyrolyse van stikstofcomponenten binnenkort geen toekomstmuziek meer is.

Besides my work at the LCT, I also enjoyed a stay at the CRF in Sandia National Laboratories in Livermore, USA. Judit, thank you for the opportunity to visit you and for the very nice collaboration. I gained a great deal of knowledge and expertise thanks to you, and I hope we can continue this collaboration and develop KinBot so that someday we can say “KinBot automatically found a new reaction type”. I also want to thank Craig for the opportunity I got at the CRF. Finally, Jasmine, thank you for everything.

Mijn doctoraat werd gefinancierd door het toenmalige Instituut voor Innovatie door Wetenschap en Technologie (IWT), dat ik hiervoor zou willen bedanken. Het Fonds voor Wetenschappelijk Onderzoek (FWO) wil ik bedanken voor de financiering van mijn verblijf in Sandia National Laboratories.

Naast het leven in het labo heb ik ook van veel steun kunnen genieten buitenaf. In het bijzonder zou ik Papa, Mama, Sarah en Myrjam willen bedanken voor alle steun en alle begrip dat jullie voor mij hadden in deze chaotische jaren.

Finalement, Bette, il n’y a pas assez de mots pour expliquer ce que tu as voulu dire dans ma vie, pas juste pendant mes études, mais toutes les dernières 23 années. Ça n’a pas toujours était facile, mais je n’y changerai rien. Alors pour tout ça, un très grand merci!

Ruben Van de Vijver

Gent, 2017

Contents

Contents	i
List of symbols	v
Summary	ix
Samenvatting	xiii
Glossary	xvii
Chapter 1: Introduction	1
1.1 Scope and objective	1
1.2 Automatic kinetic model generation.....	2
1.2.1 Building a reaction mechanism	9
1.2.2 From reaction mechanism to kinetic model: data calculation.....	16
1.2.2.1 Thermodynamics.....	17
1.2.2.2 Rate coefficients.....	24
1.3 Genesys.....	44
1.3.1 Genesys input	45
1.3.2 Kinetic model generation algorithm.....	47
1.3.3 Post-processing.....	49
1.3.4 Genesys features.....	50
1.4 Outline	51
1.5 References	53
Chapter 2: Decomposition and isomerization of 1-pentanol radicals and the pyrolysis of 1-pentanol	69
2.1 Abstract.....	69
2.2 Introduction	70
2.3 Methods	73
2.3.1 Exploration of the PES by KinBot	73
2.3.2 OH + 1-pentene channel.....	74
2.3.3 Master Equation calculations	74
2.3.4 Experimental data.....	75

2.3.5	Kinetic model construction	76
2.4	Results	77
2.4.1	PES	77
2.4.2	OH + 1-pentene capture rate coefficient	78
2.4.3	Rate coefficients	79
2.4.4	Modelling	83
2.5	Conclusions	87
2.6	References	88
Chapter 3: On-the-fly <i>ab initio</i> calculations.....		93
3.1	Abstract.....	93
3.2	Introduction	94
3.3	Methodology.....	99
3.3.1	Stable species	99
3.3.1.1	3D coordinates	99
3.3.1.2	<i>Ab initio</i> calculations.....	105
3.3.2	Transition states.....	109
3.3.3	Thermodynamics	111
3.3.4	Kinetics.....	118
3.4	Results	121
3.4.1	Stable species	121
3.4.1.1	Initial coordinates generation.....	122
3.4.1.2	Conformational analysis	122
3.4.1.3	Stereoconfiguration.....	124
3.4.1.4	Thermodynamics.....	126
3.4.2	Reactions	128
3.4.2.1	3D coordinates	128
3.4.2.2	Kinetics	132
3.5	Conclusions	140
3.6	References	141
Chapter 4: Group additivity model for intramolecular hydrogen abstractions		147
4.1	Abstract.....	147

4.2	Introduction	148
4.3	Methodology.....	153
4.3.1	<i>Ab initio</i> calculations.....	153
4.3.2	Group additivity model	154
4.4	Results	159
4.4.1	Transition state geometries.....	159
4.4.2	<i>Ab initio</i> calculations.....	160
4.4.2.1	Rate coefficients.....	160
4.4.2.2	Number of single events	164
4.4.2.3	Tunneling coefficients	166
4.4.2.4	Comparison to experimental data	168
4.4.2.5	Comparison to theoretical data	169
4.4.3	Group additivity model	173
4.4.3.1	Initial model.....	173
4.4.3.2	Final model	174
4.4.3.3	Modelling tunneling coefficients	178
4.4.3.4	Dependence of group additive values on the type of H-shift.....	181
4.4.3.5	Temperature dependence of group additive values.....	181
4.4.3.6	Thermodynamic consistency.....	182
4.4.4	Application of the group additivity model	184
4.4.4.1	1-octyl radical	185
4.4.4.2	Applicability of the model	187
4.5	Conclusions	188
4.6	References	190
Chapter 5: First principles based automated kinetic modeling of heptane pyrolysis		195
5.1	Abstract.....	195
5.2	Introduction	196
5.3	Kinetic model generation.....	199
5.3.1	Network generation	199
5.3.2	Thermochemistry and reaction rate coefficients	202
5.3.3	Model reduction	204

5.4	Results	205
5.4.1	Model validation	205
5.4.2	Rate of production analysis	217
5.4.3	Role of H-shift reactions	222
5.5	Conclusions	226
5.6	References	228
Chapter 6: Conclusions and future outlook.....		231
6.1	Conclusions	231
6.2	Future outlook.....	234
Appendices		237
Appendix A: Hindered rotor treatment.....		238
Appendix B: Correction factors for the OH + 1-pentene addition kinetics.....		239
B.1	SO coupling	239
B.2	Conformer-specific 1-D basis-set correction potentials	240
B.3	Presence of multiple conformers.....	241
Appendix C: Coordinates embedding		243
Appendix D: Wilson B matrix		246
Appendix E: Test reactions		248
Appendix F: Initial group additive values.....		257
Appendix G: Uncertainty on ΔGAV° 's.....		259
Appendix H: Publications.....		262
H.1	A1 publications	262
H.2	C1 publications	262
References		264

List of symbols

Roman symbols

A	pre-exponential factor	s^{-1} or $\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$
\tilde{A}	single-event pre-exponential factor	s^{-1} or $\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$
C_p	heat capacity	$\text{J mol}^{-1} \text{K}^{-1}$
D	binary diffusion coefficient	$\text{m}^2 \text{s}^{-1}$
d_{ab}	distance between a and b	pm
$E(x)$	expected value of x	
E_a	activation energy	J mol^{-1}
F	significance of regression	
g	degeneracy	
G	Gibbs free energy	J mol^{-1}
h	Planck constant	$6.62 \cdot 10^{-34} \text{ J s}$
\hbar	reduced Planck constant	$1.05 \cdot 10^{-34} \text{ J s}$
H	enthalpy	J mol^{-1}
I	moment of inertia	kg m^2
k	rate coefficient	s^{-1} or $\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$
\tilde{k}	single-event rate coefficient	s^{-1} or $\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$
k_B	Boltzmann constant	$1.38 \cdot 10^{-23} \text{ J K}^{-1}$
K_{eq}	equilibrium coefficient	units depend on reaction
k_h	thermal conductivity	$\text{J s}^{-1} \text{m}^{-1} \text{K}^{-1}$
L	reactor length	m
\ln	natural logarithm, i.e. with base e	
\log	common logarithm, i.e. with base 10	
l_{ab}	lower bound between atoms a and b	pm
m	mass	kg
MM	molar mass	g mol^{-1}
n_e	number of single events	
n_{opt}	number of optical isomers	

N_A	Avogadro constant	$6.02 \cdot 10^{23} \text{ mol}^{-1}$
p	pressure	Pa
$P(x)$	probability distribution of x	
q	molecular partition function	
r	random number	
R	universal gas constant	$8.314 \text{ J mol}^{-1} \text{ K}^{-1}$
R^2	multiple correlation coefficient	
S	entropy	$\text{J mol}^{-1} \text{ K}^{-1}$
\tilde{S}	symmetry-independent entropy	$\text{J mol}^{-1} \text{ K}^{-1}$
T	temperature	K
u_{ab}	upper bound between atoms a and b	pm
U	internal energy	J mol^{-1}
v	gas velocity	m s^{-1}
V	volume	m^3
$Var(x)$	variance of x	
V_m^0	standard molar volume	m^3
$V(\varphi)$	potential energy	J
$w_{i,j}$	weight	

Greek symbols

$\Delta_a H$	standard enthalpy of atomization	J mol^{-1}
$\Delta_f H$	standard enthalpy of formation	J mol^{-1}
$\Delta^\ddagger H$	standard enthalpy of activation	J mol^{-1}
$\Delta^\ddagger S$	standard entropy of activation	$\text{J mol}^{-1} \text{ K}^{-1}$
ϵ	energy level	J
κ	tunneling coefficient	
λ	eigenvalue	
μ	dynamic viscosity	$\text{kg s}^{-1} \text{ m}^{-1}$
ν	frequency	s^{-1}
ρ	density	kg m^{-3}
σ	rotational symmetry number	

φ	torsional angle
Ψ	wave function

Sub- and super scripts

‡	transition state
AI	<i>ab initio</i>
ARRH	Arrhenius
elec	electronic
eq	equilibrium
exo	exothermic
for	forward
rot	rotational
trans	translational
red	reduced
ref	reference
rev	reverse
top	rotating top
vib	vibrational

Acronyms

1D	one dimensional
2D	two dimensional
3D	three dimensional
AITSTME	<i>ab initio</i> transition state theory master equation
BAC	bond additive correction
BE	bond electron
CBS	complete basis set
CDK	chemistry development kit
DFT	density functional theory
GAV	group additivity value
IRC	intrinsic reaction coordinate
MMFF	Merck molecular force field
NNI	non-nearest neighbor interaction

PES	potential energy surface
RC-TST	reaction class transition state theory
RES	resonance correction
RMS	root mean square
RSC	ring strain correction
SCR	small curvature tunneling
TS	transition state

Summary

Scope

The chemical industry is facing several important challenges in the coming decade, mainly as a result of the tremendous impact of society on the environment, but also to maintain current living standards taking in to account the economic viability of chemical processes. These challenges entail the use of alternative feedstocks compared to the conventional fossil resources, and the design and optimization of chemical processes with a focus on a minimal use of energy and materials next to avoiding pollutant formation. An important methodology to solve these challenges are experimental measurements to search for optimal process parameters such as reactor configurations, reactor conditions and inlet concentrations. These experiments, however, are time consuming, costly and can only cover a limited range of the aforementioned parameters. Modeling of chemical processes can be a great aid to accelerate the understanding of these parameters on the process outcome and minimize the number of necessary experiments. These modeling methodologies rely on a very detailed knowledge of the underlying chemistry to the overall process. This requires the inclusion of all important elementary chemical reactions and their reactants and products. Even more importantly, the reaction rate coefficient of each reaction and the properties of each species are needed, constructing the so called kinetic model. When dealing with common chemical processes, the number of chemical reactions often exceeds a thousand, with hundreds of molecules and intermediates. The amount of data to construct these kinetic models is thus very large. This thesis focusses on the latter, i.e. on how a large number of accurate data can be retrieved automatically without the need for human interaction, which is part of a larger framework of automatic kinetic model generation. For the latter, many parameters, e.g. reaction rate coefficients or thermodynamic properties of molecules, are calculated using approximation methods such as group additivity or other quantitative structure-property relationships. Several of these methods have been proven to yield results of sufficient accuracy, given that they are not used outside their application range. The latter depends on the data used by the approximation model. These data originate from either experimental measurements or from high-level quantum chemical calculations. Today data from these high-level quantum chemical calculations are scarce due to the necessity of user involvement and expertise.

Therefore, in this work, methods are developed to allow automatic calculations of a large number of reaction rate coefficients and thermodynamic parameters using *ab initio* calculations. The results can then be used to develop new approximation schemes, to improve existing one or to extend their applicability range.

State of the art

Automatically performing *ab initio* calculations has already been implemented in several computer codes, with a different application in mind than the current work. These methods prove to be of great value and were used to investigate the state of the art. One of these computer codes is “KinBot”, developed at Sandia National Laboratories, which automatically searches for reactions on a potential energy surface. KinBot has been used to identify reactions during n-pentanol pyrolysis and to calculate accurate rate coefficients for these reactions. They were implemented in a kinetic model which shows good agreement to experimental data.

Methodology

To automatically perform *ab initio* calculations, the framework of the kinetic model generator tool “Genesys” is used, providing a computational representation of molecules, intermediates and reactions. The representation entails information on the connectivity of atoms within a species or transition state, in the form of a mathematical graph or connectivity matrix. This information is insufficient to start *ab initio* calculations, for which three dimensional coordinates of all the atoms are needed. Using “distance geometry” and force field calculations, the connectivity of a species is translated to three dimensional coordinates. Although these methods have already been applied to molecules and radicals, in this work the application range has been extended to transition states. In several steps, the coordinates are optimized and the lowest energy conformer can be selected. For the latter, high-level *ab initio* calculations are done, e.g. using the CBS-QB3 level of theory.

From the CBS-QB3 calculations, species properties are calculated using statistical thermodynamics, for which the partition functions are calculated using the molecular frequencies. Internal modes resembling rotations around a chemical bond are treated as one dimensional hindered rotors, and the frequencies corresponding to these rotations need to be removed from the partition function calculations. This is automatically done by verifying if the normal mode of a

frequency resembles an internal rotation. From the thermodynamic properties calculated for reactants, products and the transition state of a reaction, kinetic parameters can be easily calculated. These parameters are further improved by calculating tunneling corrections.

Thermodynamics and kinetics

The newly developed methodology has been tested by calculating the thermodynamic properties and rate coefficients of several species and reactions respectively, exhibiting a wide variety of functional groups and molecular structures. These values have been compared to literature data, both experimental as well as theoretic, and an excellent agreement is found, similar to the accuracy of the CBS-QB3 calculations themselves. Automatically performing the calculations has thus no impact on accuracy compared to the “manual” calculations, and can be done much faster and on a larger scale.

Group additivity

Using the newly developed methods to automatically perform *ab initio* calculations, reaction rate coefficients have been calculated for a large set of intramolecular hydrogen abstraction reactions in hydrocarbons, including abstractions by a carbon radical from a neighboring carbon atom up to abstractions by a carbon radical to another carbon atom with a carbon chain of five atoms in between. The influence of the ligands of the attacking and attacked carbon atom has been studied, as well as the influence of substituents on the carbon chain between both reactive atoms. This has led to a new group additivity model for intramolecular hydrogen abstractions, which is able to reproduce the *ab initio* calculations with sufficient accuracy.

Heptane pyrolysis

With the new group additivity model for intramolecular hydrogen abstractions, a kinetic model for the pyrolysis of n-heptane has been automatically constructed using “Genesys”. By accounting for the chemistry of small molecules and for the formation of aromatics through literature data, the new model was able to reproduce two datasets in a very wide range of pressures ($400 - 1 \cdot 10^5$ Pa) and temperatures (800 – 1800 K) without any adjustments to the kinetic model using the experiment data itself. A rate of production analysis has unraveled the

major reaction pathways, and has shown the necessity of including intramolecular hydrogen abstraction reactions for the correct prediction of several pyrolysis products.

Samenvatting

Toepassingsgebied

De chemische industrie staat op korte termijn voor verschillende belangrijke uitdagingen, vooral als gevolg van de enorme impact van de maatschappij op het milieu, maar evenals om de huidige levensstandaarden te behouden en om de economische rendabiliteit van chemische processen te waarborgen. Deze uitdagingen bevatten onder meer het gebruik van alternatieve grondstoffen in de plaats van de traditionele fossiele bronnen en de ontwikkeling en optimalisatie van chemische processen met als doel het gebruik van materialen en energie te minimaliseren en de productie van schadelijke producten te vermijden. Een belangrijk hulpmiddel bij het zoeken naar oplossingen voor deze uitdagingen zijn experimentele metingen om optimale procesparameters te verkrijgen, zoals reactor configuratie, condities en inlaatsamenstellingen. Deze experimentele technieken vergen echter veel tijd, zijn duur, en kunnen slechts een klein gedeelte van de parameter ruimte bestuderen. Daarom wordt het modeleren van chemische processen steeds meer als alternatief gebruikt om een beter inzicht te verkrijgen in de invloed van de verschillende parameters, en om het aantal nodige experimenten sterk te verminderen. Modeleren van een chemisch proces steunt op een gedetailleerde beschrijving van de chemische reacties die erin plaatsvinden. Elke belangrijke reactie en zijn reactanten en producten moet in het model worden opgenomen en daarenboven moeten de snelheidscoëfficiënten van alle reacties en de thermodynamische eigenschappen van alle moleculen en intermediären gekend zijn, wat vervat zit in een zogenaamd kinetisch model. In veel chemische processen zijn er meer dan duizend belangrijke reacties tussen enkele honderden species. Er is dus veel chemische data nodig om een kinetisch model op te stellen. Dit werk is gefocust op het laatste, namelijk op hoe een groot aantal accurate data automatisch kan worden berekend zonder dat hier manuele tussenkomst voor nodig is. Dit kadert in het automatisch genereren van kinetische modellen, waarvoor snelheidscoëfficiënten van chemische reacties en thermodynamische parameters van moleculen en intermediären worden berekend aan de hand van benaderingsmethodes zoals groep additiviteit of andere kwantitatieve methodes die de eigenschappen van een molecule bepalen aan de hand van de moleculaire structuur. Het is aangetoond dat deze methodes accurate resultaten opleveren als ze in het juiste toepassingsdomein worden gebruikt, wat vooral afhangt van de

achterliggende data van de berekeningsmethodes. Deze data, die worden opgemeten uit experimentele technieken of worden bepaald door middel van kwantumchemische berekeningen, zijn schaars, onder andere door de nodige betrokkenheid en deskundigheid van de gebruiker. Daarom worden in dit werk methodes ontwikkeld om hier een oplossing voor te bieden door automatische *ab initio* berekeningen van een groot aantal reactie-snelheidscoëfficiënten en thermodynamische parameters van species toe te laten. De resultaten kunnen vervolgens gebruikt worden om benaderingsschema's op te stellen, te verbeteren en hun toepassingsdomein uit te breiden.

Stand van zaken

Het automatisch uitvoeren van kwantumchemische berekeningen werd al in verschillende computer codes geïmplementeerd, telkens met een verschillend toepassingsgebied dan dit werk. De ontwikkelde methodes kunnen echter van grote waarde zijn en werden gebruikt om de huidige stand van zaken te onderzoeken. Eén van deze codes is "KinBot", die automatisch naar reacties zoekt op een potentieel energie oppervlak. KinBot werd gebruikt om reacties te vinden die van belang zijn voor de pyrolyse van n-pentanol en om accurate snelheidscoëfficiënten te berekenen. Deze resultaten werden in een kinetisch model geïmplementeerd, dat goede overeenkomst vertoont met experimentele data.

Methodes

Om automatisch *ab initio* berekeningen uit te voeren werd "Genesys" gebruikt, dit is een computer code die automatisch kinetische modellen genereert, en die een computationele voorstelling van moleculen, intermediären en reacties omvat. Deze voorstelling is de connectiviteit van de atomen in een species of transitietoestand opgesomd in een wiskundige graaf of een connectiviteitsmatrix, wat onvoldoende is om *ab initio* berekeningen te starten, waarvoor driedimensionale coördinaten van alle atomen nodig zijn. Gebruik makende van "*distance geometry*" en krachtveld berekeningen kan de connectiviteit van een species worden vertaald in driedimensionale coördinaten. Deze methodes worden al toegepast op moleculen en radicalen, en werden in dit werk uitgebreid naar transitietoestanden. In verschillende stappen worden de coördinaten geoptimaliseerd en wordt de laagste energie conformer geselecteerd, waarvoor CBS-QB3 berekeningen worden uitgevoerd.

Uit deze resultaten worden de thermodynamische eigenschappen van de component berekend aan de hand van statistische thermodynamica, waarvoor de partitiefuncties worden behaald met de moleculaire frequenties. Interne modes die rotaties rond chemische bindingen voorstellen worden behandeld als eendimensionale hinderde rotoren, en de moleculair frequenties moeten worden weggelaten om de partitiefuncties te berekenen. Dit wordt automatisch gedaan door te controleren of de normale mode van een frequentie overeenkomt met een interne rotatie. Met de thermodynamische parameters van de reactanten, producten en transitietoestand van een reactie kunnen de snelheidscoëfficiënten berekend worden, die verder worden verbeterd met tunnelingsberekeningen.

Thermodynamica en kinetiek

De ontwikkelde methodes werden getest door het berekenen van thermodynamische parameters en snelheidscoëfficiënten van verschillen species en reacties respectievelijk, met een grote variatie aan functionele groepen en moleculaire structuren. De berekende waarden werden vergeleken met data uit de literatuur, zowel uit experimentele metingen als uit theoretische methodes, en een uitstekende overeenkomst werd verkregen, die vergelijkbaar is met de accuraatheid van CBS-QB3 berekeningen. Het automatisch uitvoeren van die berekeningen heeft dus geen invloed op de accuraatheid vergeleken met het “manueel” uitvoeren ervan, en meer berekeningen kunnen worden uitgevoerd in een kortere tijdspanne.

Groep additiviteit

Om het nut van automatische *ab initio* berekeningen aan te tonen werden snelheidscoëfficiënten van een grote set aan intramoleculaire waterstofabstractiereacties in koolwaterstoffen berekend, waarvan de abstractie plaatsvindt tussen twee naburige koolstofatomen tot en met abstracties tussen twee koolstofatomen waartussen zich een keten van 5 andere koolstofatomen bevindt. The invloed van de liganden op het koolstofradicaal en op het koolstofatoom waarvan een waterstof wordt geabstraheerd werd onderzocht, evenals de invloed van substituenten op de koolstofketen tussen beide reactieve koolstofatomen. Dit heeft geleid tot een nieuw groep additief model for intramoleculaire waterstofabstractiereacties, die met voldoende accuraatheid de *ab initio* data kan reproduceren.

Heptaan pyrolyse

Met het nieuw groep additief model voor intramoleculaire waterstofabstracties werd een kinetisch model automatisch opgesteld voor de pyrolyse van n-heptaan gebruik makende van “Genesys”. De reacties die de chemie van kleine componenten beschrijft en de reacties voor de vorming van aromaten werden uit de literatuur gehaald. Het uiteindelijke model was in staat om verschillende experimentele datasets goed te reproduceren in een grote druk ($400 - 1 \cdot 10^5$ Pa) en temperatuur (680 – 1800 K) bereiken, zonder enige parameter te fitten aan de experimenten. Met behulp van reactiepadanalyse werden de belangrijkste reacties geïdentificeerd, en werd het nut van de intramoleculaire waterstofabstractiereacties voor de correcte beschrijving van de vorming van verschillende pyrolyseproducten aangetoond.

Glossary

Automatic kinetic model generation	Automatic kinetic model generation is an engineering methodology to build a kinetic model by making use of computers with only minor human interaction.
Branching fraction	The branching fraction of a reaction is the relative rate coefficient of that reaction to the sum of the rate coefficients of the other reactions with the same reactant(s).
Chemically activated reaction	A chemically activated reaction is a reaction that proceeds after the collision of the reactant(s) with a bath molecule, leading to the formation of an active complex which lies above the reaction barrier. This complex decomposes to the products of the reaction.
Chemical significance	A structure of a species is chemical significant if the interatomic distances of two non-bonded atoms is smaller than 1.2 times the sum of the covalent radii of both atoms.
Conformer	A conformer of a species is a structure belonging to that species, i.e. in which all the covalent bonds are identical as the ones from the topology of the species.
Diastereomer	Stereoisomers that are not enantiomers are called diastereomers. They are not mirror images of each other and are not superimposable.
Distance bounds matrix	Matrix containing the lower and upper distance limits between each pair of atoms in a species.
Distance geometry	Distance geometry is a numerical method that generates 3D coordinates of all the atoms in a species based on the connectivity of that species.
Distance matrix	Matrix containing the distances between each pair of atoms in a species.
Enantiomer	An enantiomer is one of two stereoisomers that are mirror

	images of each other without being superimposable.
Frobenius norm	The Frobenius norm of a matrix is a scalar obtained by the square root of the sum of the absolute squares of the elements of the matrix.
Group additivity method	A group additivity method allows to calculate properties of molecules or reaction by summing contributions for each group the molecule or reaction exists of. A group is a submolecular pattern existing of a small number of atoms.
Intrinsic reaction coordinate	Intrinsic reaction coordinate calculations are calculations in which starting from a saddle point, the two wells on each side of the saddle point are searched for, corresponding to the reactants and products of a reaction.
Isodesmic reaction	An isodesmic reaction is a reaction of which the bonds that are broken in the reactant are of the same type as the bonds that are formed in the product.
Level of theory	The level of theory of quantum chemical calculations describe the treatment of the electron correlations and defines the basis set.
Potential energy surface	A potential energy surface describes the electronic energy of chemical species and intermediates as a function of the geometrical arrangement of the atoms.
Rate rule	Rate coefficients belonging to a specific combination of reactants.
Reactive atom	An atom that changes in connectivity through an elementary reaction.
Reactive center	Collection of reactive atoms for one elementary reaction
Residence time	The residence time in a tubular flow reactor is calculated by the volume of the reactor divided by the volumetric flow rate of the reactant at the reactor inlet.
Rotamers	A rotamers are conformers of a species that are formed by rotating a part of the molecule along a single bond.

Structure	A structure is a set of 3D coordinates describing the geometrical arrangement of all the atoms of a chemical species.
Template	3D coordinates of the reactive center supplied by the user to Genesys.
Thermodynamic consistency	A reaction is thermodynamic consistent if the ratio of the forward and reverse rate coefficients equals the equilibrium coefficient of the reaction.
Topology	Connectivity information of a chemical species, optionally augmented with the configuration of stereocenters.
Well-skipping reaction	A well-skipping reaction is a reaction that proceeds between reactants and products in between which the energy of the lowest energy path crosses two or more local maxima, and at least one well. The actual energy of the activated complex lies above the highest energy along that path and the activated complex is not relaxed into one of the intermediate wells but directly in the final products.

Chapter 1: Introduction

This chapter is based on the following paper:

Van de Vijver R, Vandewiele NM, Bhoorasingh PL, Slakman BL, Seyedzadeh Khanshan F, Carstensen H-H, Reyniers M-F, Marin GB, West RH, Van Geem KM. Automatic Mechanism and Kinetic Model Generation for Gas- and Solution-Phase Processes: A Perspective on Best Practices, Recent Advances, and Future Challenges. *International Journal of Chemical Kinetics*. 2015;47(4):199-231.

1.1 Scope and objective

Fully automated mechanism generation of detailed kinetic models is within reach in the coming decade. This is necessary to extend our fundamental understanding of chemical processes and resolve many of today's society problems such as energy conversion, emission reduction, greener chemical production processes, etc. In this thesis, the focus is on the core of the automated mechanism generation for gas-phase processes i.e. on how the reaction kinetics and thermodynamic properties of species are calculated starting from the fundamental theory. One of the main challenges that still needs to be overcome deals with data scarcity. The combination of kinetic model generation with affordable computational chemistry calculations to provide new, reliable data seems the logical step forward. The ability to automatically perform quantum

chemical calculations for a large set of species or reactions would allow the fast generation of new calculation methods to obtain the necessary data fast and reliable. However, user knowledge and expertise are nowadays often required to obtain the necessary results. This thesis aims at minimizing the human interaction with computational chemistry and elaborates on automation procedures to calculate thermodynamic properties of species and reaction rate coefficients. The objective is fully automatic *ab initio* calculations that can be done in the framework of kinetic model generation to account for the data gaps and data uncertainties, with the applications of thermal decompositions, pyrolysis and combustion in mind. Therefore, the methods are restricted to gas-phase reactions.

1.2 Automatic kinetic model generation

Chemical kinetic models are extremely powerful and valuable. Many significant public policy and business decisions are made on the basis of kinetic model predictions. For example, the Montreal Protocol, which imposed a worldwide ban on certain halocarbons, was based on a fundamental knowledge of the ozone layer problem established by kinetic modeling.¹ In the chemical industry kinetic models are widely applied, e.g. to simulate steam cracking²⁻⁵, refining⁶ or vinyl chloride production⁷. However, for the majority of technologically important chemical processes, including combustion, pyrolysis, and oxidation of hetero-atomic mixtures, detailed kinetic models are not available. This is because constructing a reliable model remains very difficult and time consuming. One reason for this is that such models typically contain thousands of reactions, involving hundreds of intermediates, while only a small fraction of the reaction rate coefficients have been determined experimentally. Moreover, it is usually impossible to measure the concentrations of all the kinetically significant chemical species.⁸ Numerically solving these large systems of differential equations in a reasonable time also remains a challenge, in particular when these models need to be implemented in Computational Fluid Dynamics codes.⁹

Ideally, kinetic models should capture the essential chemistry while maintaining a manageable size to enable relatively fast simulations. Figure 1 displays the size of selected kinetic models for thermal decomposition, oxidation, and combustion processes, with the outliers listed in Table 1.

Table 1 shows that, among the models generated recently, those that contain several thousands of reactions are generated automatically, such as the model of Herbinet et al. (n. 6 in Table 1), while the small models are either lumped or were constructed by hand. The figure also shows a remarkable trend that was first pointed out by Lu and Law¹⁰: The number of reactions increases linearly with the number of species in the model and not, as could be expected, exponentially. However, the model size increases exponentially as a function of the number of heavy (non-hydrogen) atoms in the reactant molecule(s).

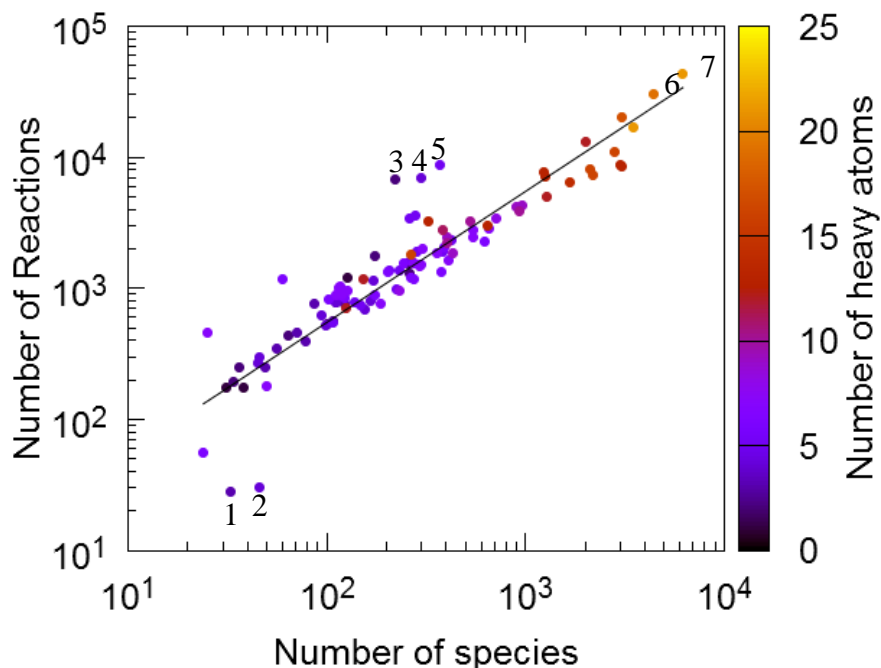


Figure 1: Model sizes: number of reactions as a function of the number of species for gas phase kinetic models of pyrolysis, oxidation and combustion processes. The secondary axis is the number of heavy (non-hydrogen) atoms in the reactant molecule(s). The line through the graph has a slope of 1.

Table 1: Outliers of Figure 1.

Number on graph	Year	Number of species	Number of reactions	Compound	Process	Comments	Ref.
1	2005	33	28	Dimethyl ether	Combustion	Manual/ Lumped	Zheng et al. ¹¹
2	2007	46	30	1,3-butadiene	Combustion	Manual/ Lumped	Zheng et al. ¹²
3	2013	220	6800	Ethylene/methane	Combustion	Automatic	Cuoci et al. ¹³
4	2010	300	7000	Propanol	Combustion	Automatic	Frassoldati et al. ¹⁴
5	2013	372	8732	Iso-butanol	Pyrolysis and oxidation	Automatic	Merchant et al. ¹⁵
6	2011	4442	30425	Methyl palmitate	Combustion	Automatic	Herbinet et al. ¹⁶
7	2011	6203	43444	Methyl stearate	Combustion	Automatic	Herbinet et al. ¹⁶

When considering the number of species and reactions in detailed kinetic models, it becomes obvious that the manual construction of such models is tedious and error-prone. This motivated the construction of automatic kinetic model generation codes in recent decades: NETGEN¹⁷ at the University of Delaware, RMG¹⁸⁻²¹ at MIT based on the earlier codes XMG and NETGEN, EXGAS^{22,23} at the *Université de Nancy and CNRS*, Genesys²⁴ and PRIM^{25,26} at Ghent University, RING^{27,28} at the University of Minnesota, MAMOX^{29,30} at *Politecnico di Milano*, REACTION³¹⁻³³ at the Johannes Kepler University, COMGEN³⁴ at the University of Utah, MECHGEN³⁵ at the Hungarian Academy of Sciences, KING³⁶ at the University of Calabria, RAIN^{37,38} at the *Technische Universität München*, CASB³⁹ at the Mari State University, GRACE^{40,41} and KUCRS⁴² at the University of Tokyo, RNG⁴³ at the Institute of Chemical Technology, Prague, etc.⁴⁴⁻⁴⁸

A reaction mechanism entails the information of a chemical system via the list of molecules, intermediates, and the reactions between these species. A detailed kinetic model is the union of the chemical mechanism with thermodynamic and transport data of the species, and reaction rate coefficients of the reactions, translated into a mathematical model. The model equations describe the variation in species concentrations, temperature, and pressure via conservation laws.

When constructing a kinetic model it is first important to know which species interact through which reactions, i.e. what the reaction mechanism is. Second, accurate values for the thermochemistry, optionally transport properties, and reaction kinetics are indispensable to allow reactor simulations. To solve the continuity equations during reactor simulations, the kinetic data is needed to formulate the corresponding source term. For each species, this source term describes how the species concentration changes due to chemical reactions. The energy equations also have a source term in which the generated thermal power is described as a function of the reaction rates and the reaction enthalpies.

The need for thermochemistry is two-fold. First it is used in the construction of the energy balances in non-isothermal reactor models. To solve these equations, the enthalpy of formation and the heat capacity of all species are required. Secondly, if a reaction is implemented as a reversible reaction, the reverse rate coefficient is calculated based on the forward rate coefficient and the equilibrium coefficient, i.e. via so-called thermodynamic consistency, Eq. 1.1.

$$\frac{k_{\text{for}}(T)}{k_{\text{rev}}(T)} = K_{\text{eq}}(T) = \exp\left(-\frac{\Delta_r G^0}{R \cdot T}\right) \cdot \left(\frac{R \cdot T}{p}\right)^{-\Delta n} \quad \text{Eq. 1.1}$$

Finally, transport data is needed to account for heat and mass transfer throughout the reactor; these quantities are diffusivity, viscosity and thermal conductivity of the species.

Several reviews on automatic mechanism generation have been published in the past, all with their own focus and application in mind. Green et al.¹⁹ summarized the main methods needed to allow computers to build complex kinetic models for real processes. Besides the discussion of all the necessary reaction steps, algorithms to calculate molecular properties and reaction rate coefficients, including pressure dependence, were discussed. Specific attention was paid to the implementation of (on-the-fly) reduced complex kinetic models in reactive flow simulations. Green⁴⁹ further reviewed the obstacles to build kinetic models with a focus on the fundamental data models use, the methods for solving kinetic simulations, and the methods to determine if model predictions are consistent with experimental data. Furthermore, the challenge of handling data, both experimental and theoretical data in the kinetic model, was stressed. Ratkiewicz and Truong⁵⁰ reviewed the use of algebra for the generation of chemical reactions in complex systems. Broadbelt and Pfaendtner⁵¹ wrote a dictionary on quantitative kinetic modeling that, besides unifying the chemical principles, served as a starting point for continuous tracking of the advancements, refinements, and refreshments in this field. The authors emphasized the need for collaboration between experimentalists and theorists and further focused on the state-of-the art of kinetic modeling via theoretical methods. Vinu and Broadbelt⁵² focused on several aspects of kinetic modeling of pyrolysis and oxidation processes involving polymers, such as the occurring elementary reactions steps, the assignment of rate coefficients to reactions, and the influence of end-group functionalities and reactive centers in the polymer chain on the reaction mechanism. It is no surprise that researchers working in the field of combustion are among the most active in the field of automatic kinetic model construction. During the last decade tremendous progress has been made in the combustion community thanks to accurate kinetic models in terms of engine design, fuel blend optimization, etc. This resulted in several recent reviews specifically on kinetic modeling of combustion processes.⁵³⁻⁵⁵ An interested reader is also directed to a recent book on the subject.⁵⁶ However, all these reviews agree that for detailed kinetic models, which typically contain thousands of reactions c.f. Figure 1, it is simply not feasible to measure each reaction rate individually. Moreover, even with the power of today's fastest supercomputers, the calculation of

all the rate coefficients using high level computational chemistry methods remains too time consuming. This implies that cheaper methods are needed to calculate the large number of unknown reaction rate coefficients of the different elementary reactions before one can even attempt to compare experiment with simulation. This is typically resolved by years of expert user involvement combined with several assumptions that help to retain the mechanism within a manageable size, so that it becomes tractable for experienced researchers.

Although several challenges remain in automatic mechanism generation codes, it is already possible to generate large mechanisms by means of computers with only minor human interventions. The interested reader is referred to several papers in the literature where a kinetic model is automatically generated, optionally manually improved, and subsequently validated against experimental data.

Ranzi et al.⁵⁷ studied the capabilities of automatically generated kinetic models. The use of MAMOX+ was illustrated and the obtained models were reliable for the pyrolysis, partial oxidation, and combustion of branched alkanes.

Warth et al.⁵⁸ demonstrated the use of EXGAS to construct an oxidation model for n-butane. EXGAS itself is only a mechanism generation code and no model parameters are provided. The latter is handled by two other codes: THERGAS to calculate thermodynamic data and KINGAS to calculate reaction rate coefficients. The main response variables, such as product yields, conversion, and induction period, are satisfactorily predicted by the model.

Hakka et al.⁵⁹ used EXGAS to build an oxidation kinetic model for methyl and ethyl butanoates and validated it against experimental data published in the same article. The main reaction pathways in the oxidation of butanoates were searched via reaction flux and sensitivity analyses.

Grenda et al.⁶⁰ modeled the pyrolysis of methane at high temperature (1038K) via a reaction mechanism constructed by XMG that showed excellent agreement with experimental data. The authors emphasized the importance of pressure dependence of chemically activated reactions to accurately model such systems.

Van Geem et al.⁶¹ generated a n-hexane steam cracking model using RMG which accurately predicted the conversion and the yields of the major products without any parameter adjustment

against the experimental data. Furthermore, it was shown that the quasi-steady state approximation for large radicals holds and that pressure dependence can be neglected for the studied conditions. Van Geem et al.⁶¹ subsequently reduced the mechanism substantially while still retaining good agreement with the experimental data. This agreement indicates that the generation of very large kinetic models is certainly not always the best option and accurate mechanism generations and reaction simulations can be done at lower computational costs.

Harper et al.⁶² developed a n-butanol pyrolysis and combustion model using RMG and compared it successfully to experimental data. The most important reaction paths were uncovered, by means of sensitivity analyses, in which several rate coefficients needed to be more accurately determined. This was done using quantum chemistry and transition state theory (TST) calculations. These TST calculations are not yet automated as a part of the mechanism generation and this intervention can be seen as a recommended manual improvement of the model.

Moréac et al.⁶³ showed the capabilities of the REACTION code by modeling the oxidation and combustion of n-decane and n-heptane. The ignition delay time was accurately predicted and besides the validation of the entire model, the rate coefficients were also successfully compared. The model was built by combining two sub-models: a manually constructed C4- model and an automatically generated C5+ model.

Mersin et al.⁶⁴ compared a hand-generated kinetic model of hexadecane combustion and oxidation to one generated automatically by REACTION to show that they are similar.

Several other publications on this topic are by Glaude et al.⁶⁵, Buda et al.⁶⁶, Prickett and Mavrovouniotis⁶⁷, Muharam and Warnatz⁶⁸, and Khan and Broadbelt⁶⁹.

Note that the ability to predict chemical processes does not only provide an alternative to time-consuming and costly experiments, it also accelerates process development and innovation. Recently, via the generation of kinetic models, it has already become possible to describe complex gas-phase chemistry occurring in processes such as steam cracking, pyrolysis and combustion. Simulation software based on detailed kinetic models has already found its way into industrial production plants, both to optimize the day-to-day production as well as to further develop and improve the plant.²⁻⁷ Future applications of detailed gas-phase and solvation-phase kinetic models could encompass the detailed development of complex reactor systems and

combustion engines, reactants or fuels with specific requirements, and reactor materials. Furthermore, although this work is focused on gas-phase chemistry, the extensions to catalysis, polymerization and biomass related processes open up a wide array of possible applications of detailed kinetic models. This ranges from catalyst design^{70,71} and optimization of catalytic processes⁷² to the creation of new polymer materials with very specific characteristics^{73,74} or the development of sustainable processes based on biomass feeds⁷⁵. A lot of effort has been made in these fields, but the description of these applications lie outside of the scope of this thesis.

During the last five years clear progress has been made in this field reducing expert user involvement by automating many of the tasks that used to require human intervention, such as sensitivity and rate of production analysis. With this progress, the first accurate and completely in-silico generated kinetic model should become available in the not-so-far future. Starting from the state-of-the-art in automatic kinetic model generation all the necessary ingredients for a good code are outlined. The main focus of the next sections are (1) on the automatic construction of detailed kinetic models for the description of gas-phase processes such as oxidation, pyrolysis and combustion, and (2) on the assignment of kinetic and thermodynamic data for species and reactions. Both fast, scalable methodologies and computationally expensive methods are discussed. Particular attention is paid to recent advancements which are especially of interest for this work, such as the use of on-the-fly quantum mechanical calculations to extract the required data. Most of the examples given in this work are applied to kinetic models incorporating only small elements such as hydrogen, oxygen, carbon and nitrogen but the discussed approaches can be extended to other elements and chemistries.

1.2.1 Building a reaction mechanism

Three elements are key to generate a mechanism automatically, which are illustrated in Figure 2. These will be elaborated in the following paragraphs.

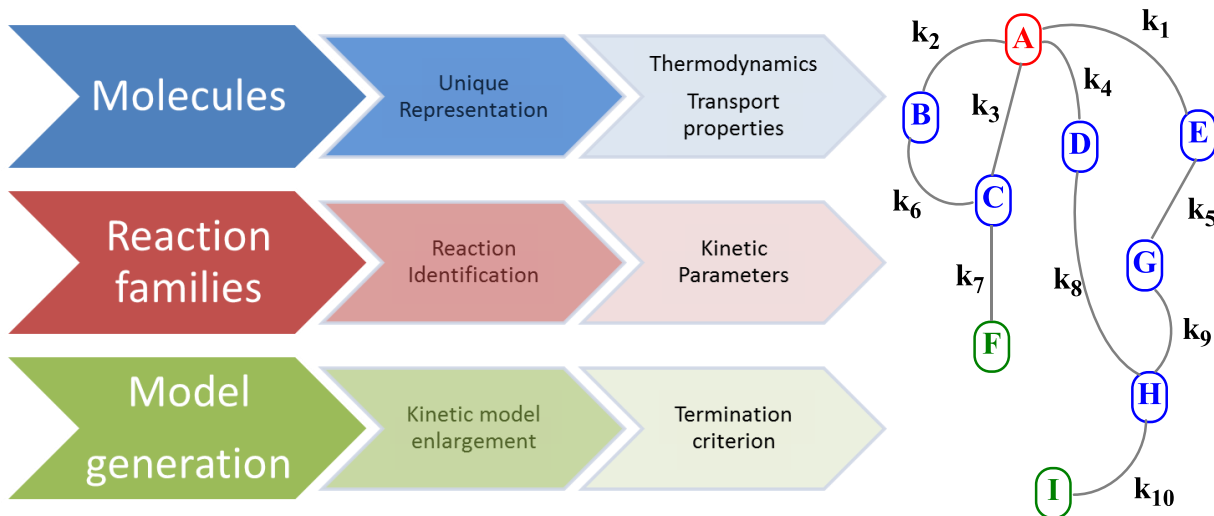


Figure 2: Schematic overview of the essential elements of an automatic kinetic model generation code. Resulting kinetic model of reactant A up to products I and F via intermediates B, C, D, E, G and H.

In computer codes, a species is regarded as a collection of atoms each bonded to one or more other atoms, and knowledge of the atom connectivity, bond orders, lone electron pairs, radical sites, and charges suffices for a correct treatment of the species. It is thus possible for computers to use the topological representation of the species, e.g. bond matrices^{25,36,76,77} or graph objects^{20,24}, or to employ a collection of functional groups to describe the chemical structures.^{78,79} The former only holds information about the elements, the bond orders, and the connectivity of the molecule. The latter is useful when the complete atom connectivity of the molecule is unnecessary for the mechanism generation. To construct a detailed predictive kinetic model it is key to represent the species unambiguously, i.e. to represent each species in a unique fashion. Furthermore, it is necessary to allow unambiguous reading of input species. Several methods have been developed to read, write and generate species names such as SMILES⁸⁰, InChI⁸¹, adjacency lists or in-house developed names.^{82,83}

One possible way to assure each species' uniqueness in a kinetic model is the use of algorithms developed by so-called chem(o)-informatics. This is a cross-disciplinary scientific field involving

computer sciences, mathematics and chemistry. The most important species representation in chemo-informatics projects is the use of graph objects with a set of nodes or vertices representing the atoms, connected by edges that represent the chemical bonds. Furthermore, a graph can be weighted to introduce additional information about the chemical species.⁸⁴ Vertices can point to more than “an atom” by bearing information about the atom type, its electronic configuration, etc. The edges can refer to single, double, triple, aromatic, transition-state, etc. bonds. This species representation can be used to describe stable species, i.e. minima on the potential energy surface (PES) as well as transition state structures, i.e. approximated by saddle points on the PES.

To ensure species are represented in a unique and unambiguous manner, (chemo-) informaticians have developed algorithms for canonicalization, the unique numbering of atoms in a molecule.⁵⁰ The canonical representation of molecules, i.e. the unique set of numbers assigned to the atoms, can be used to compare two molecules, to search in molecular databases for information about the species, or to generate a unique name, e.g. an InChI⁸¹. Several algorithms have been developed to canonicalize a graph⁸⁵⁻⁸⁸, but the most widely used one was introduced by Morgan.⁸⁹ This algorithm starts by partitioning all the atoms based on graph invariants, i.e. a sequence of numbers computed, starting from the graph topology. In each step, the graph invariant of each atom is recursively updated by summing the invariants of all its neighboring atoms from the previous step. The procedure continues until the number of distinct invariants reaches a maximum. This is illustrated in Figure 3: first, an initial invariant of one is chosen for each atom. In the second step, the methyl substituents retain an invariant of one since they only have one neighbor. The secondary carbon atoms get an invariant of two because of their two neighbors, and tertiary carbon atoms obtain an invariant of three, leading to a total of 3 distinct invariants. The third step continues with the same procedure resulting in 6 distinct invariants while the next step results in only 5 distinct variables. This means that the maximum number of invariants is 6, and that all the atoms with the same invariant in the third step are equivalent.

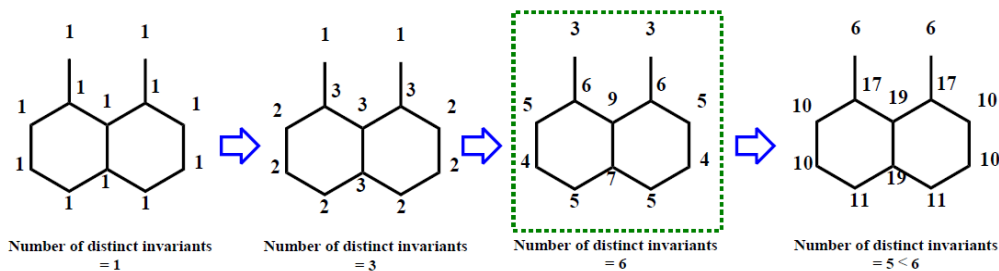


Figure 3: Recursive calculation of graph invariants for the atoms in dimethyldecalin. The initial invariants are chosen to be one.⁹⁰

The first possibility to verify whether two chemical species are identical is to generate a unique name, a so-called string, for both graphs and subsequently compare the strings. Another option is to use graph isomorphism. Two graphs are isomorphic if a bijection exists from the vertices of the first graph to the ones of the second graph in such a way that the connection between the vertices remains. Besides the use of isomorphism algorithms in determining the uniqueness of two graphs, these algorithms can be used for sub-graph recognition as well, i.e. whether graph G contains a sub-graph that is identical to graph H. This results in a mapping between the vertices of G that match vertices of H.

Sub-graph recognition is a valuable tool in automatic mechanism generation for several applications. First, a molecule can be searched for functional groups or reactive centers capable of undergoing a specific reaction. Secondly, if constraints are used to limit the size of the mechanism or only add kinetically significant reactions and species, they can be defined in terms of sub-molecular structures. This way, the presence of a specific atom arrangement with specific bond types can be searched for via isomorphism algorithms to decide whether or not a given species can undergo certain reactions. Finally, to calculate species properties or reaction rate coefficients via group additivity, the groups in a molecule need to be identified and linked to databases. The presence of specific groups can be detected via sub-graph searches.

The use of such sub-graph definitions, to store sub-molecular patterns in libraries or to define reactive centers and reaction constraints, demands an unambiguous and human-readable sub-graph language, the latter is needed to construct data libraries and input files or to analyze output files. An example of a sub-molecular pattern language is the SMARTS language (SMILES ARbitrary Target Specification).⁹¹ The use of atom or bond specifications, alongside logical

operators, allows the definition of both very general as well as very specific sub-molecular patterns.

All the possible reactions a species can undergo, either unimolecular or in the presence of another species, can be described in terms of elementary reaction families.³¹ A reaction family specifies the atom rearrangements occurring when proceeding from the reactants to the products. Abstraction is made of the electronic structure and the atomic nature, only the connectivity pattern, the bond orders, and their changes are important to define a reaction family. These families each represent a particular type of elementary chemical reactions, e.g. a homolytic scission of a single bond or a radical addition to a multiple bond. A popular approach taken by kinetic model generation programs is to complete the reaction mechanism by successive application of the reaction families to large molecules with a limited series of reactions involving small molecules, e.g. reactions involving molecules ranging from hydrogen up to molecules containing three or four carbon atoms. This approach is justified in a number of ways. First of all, the availability of rate coefficients for reactions involving small molecules, either determined from kinetic experiments or from affordable high-level *ab initio* calculations avoids the need for more approximate methods that are designed to provide rate coefficients for each member of the reaction family. Second, group additivity does not yield the same accuracy for small molecules, because the molecule in its entirety influences the rate coefficients rather than the immediate surrounding of the reactive atoms. Furthermore, small molecules lack many reaction pathways that are available for larger molecules.⁵⁴ Therefore, researchers have preferred the implementation of so called seed-mechanisms, which entail the small-molecule chemistry, in their kinetic models.⁹² The molecules in these seed-mechanisms are limited in size and most kinetic data originates directly from experimental or theoretical work. The part of the kinetic model that contains reactions of larger molecules are automatically generated either starting from the seed-mechanism or starting from scratch. In the latter case, the seed-mechanism is merged after the kinetic model generation. In many processes where automatic kinetic model generation is applied, larger molecules break down into smaller ones, and eventually to molecules contained in the seed-mechanisms. A correct description of the small-molecule chemistry is thus of utmost importance to enable the validation of the complete kinetic model. Therefore, extensive validation and optimization of small-molecule chemistry has been done in the past. Metcalfe et al.⁹² published a comparative study of small kinetic models which serve, amongst others, as seed-

mechanisms for the description of the combustion of commonly used and potentially beneficial fuels. Several kinetic models describing small-molecule combustion have been developed in the past.⁹³⁻⁹⁷

To automatically generate all the possible reactions that a species can undergo reaction rules are required to execute the reaction specifying the characteristic transformations corresponding to the considered reaction families. After the execution of these reaction rules, product species are formed, which can subsequently be subjected to the reaction rules themselves giving birth to second generation, third generation, ..., products. The mechanism generation is iteratively continued until the mechanism is considered complete according to user-defined tolerances and constraints.

Each species not yet subjected to the reaction rules first needs to be searched for structural features needed for the execution of a certain reaction family. For example, when a β -scission family is to be executed, the molecule has to possess the required structural feature, i.e. an atom with an unpaired electron bonded to an atom that is connected to a third atom via a single bond, c.f. Figure 4. If a molecule possesses this submolecular structural feature, the reaction rules defining the characteristic structural rearrangement of the reaction family can be executed. For the β -scission, the rules define the increase of the bond order between the first and second atom, the breaking of a bond between the second and third atom, the loss of the unpaired electron of the first atom and the gain of an unpaired electron on the third atom. The formed products are subsequently added to the mechanism if not yet present.

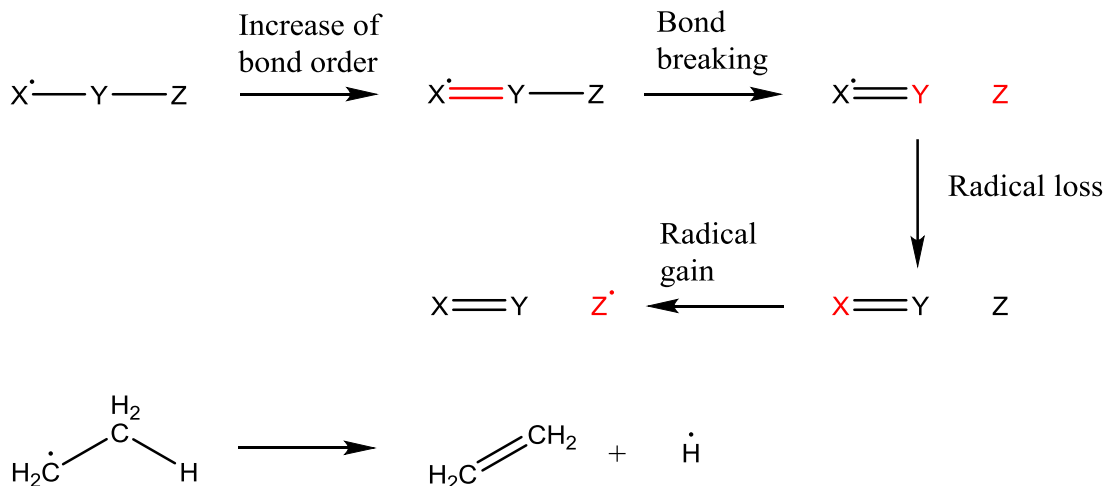


Figure 4: Reaction rules defining the characteristic structural rearrangement of a β -scission reaction exemplified for the reaction of the ethyl radical to ethylene and the hydrogen radical.

When reaction families allow the size of molecules to increase, the aforementioned procedure to construct a mechanism could keep generating reactions infinitely. Furthermore, this procedure does not entail any chemical knowledge and a significant number of kinetically unimportant or improbable reactions would be added to the mechanism. To limit the size and only add important reactions, termination criteria are required. These criteria can be constructed in several ways, of which rule-based, rate-based, and rank-based are the most commonly described in literature. Rate-based criteria²¹, as used in RMG¹⁸ for example, employ reactor simulations to assess which species have high rates of formation and should be added to the kinetic model. To ensure that all the important reactions are included, these criteria thus need relatively accurate values for rate coefficients and thermochemistry. The algorithm is illustrated in Figure 5. The initially defined reactants constitute the initial core of the mechanism that is subjected to all possible reactions specified by the reaction rules of all the defined reaction families. This yields a set of product species that are added to the edge of the mechanism. The model (core plus edge) is then used in a reactor model to identify the edge species with the highest rate of formation. If this highest rate exceeds a threshold value, c.f. Eq. 1.2, the species and its corresponding reactions are added to the core of the kinetic model. In Eq. 1.2, R is the rate of formation, t is the time and ϵ is a user-defined tolerance. The subscript i is the species under consideration to be added to the core, \min is the minimum requirement for a species to be added to the core and j corresponds to the species in the core. This iterative procedure is continued until no edge species have a rate of production that meets the requirement of Eq. 1.2.

$$R_i(t) \geq R_{min}(t) = \varepsilon \cdot \left(\sum_j R_j^2(t) \right)^{1/2} \quad \text{Eq. 1.2}$$

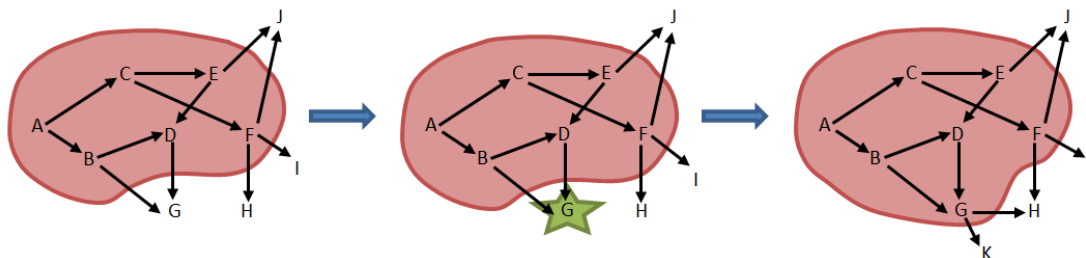
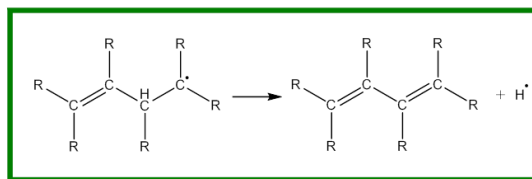


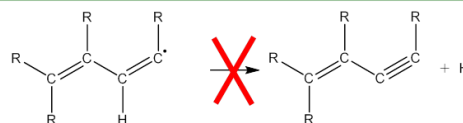
Figure 5: Rate-based model enlargement. The pink area is the core of the mechanism, containing initially six species (A-F) and six reactions. These six species are subjected to all possible reactions specified by the reaction rules of the defined reaction families which lead to the formation of four new species (G-J) that are put in the edge of the mechanism. The reactions of the core and edge are then used in a reactor model to identify the species with the highest rate of formation (G) that is added to the core of the mechanism. Next, all possible reactions of this expanded core are again generated leading to a new species in the edge (K).

Rule-based criteria, on the other hand, incorporate chemical knowledge of the user to evaluate whether or not the structural rearrangement defined by the reaction family is to be executed. For example, this knowledge can be translated into constraints on the required characteristic structural features as is done in Genesys²⁴ and illustrated in Figure 6. These constraints are defined for each reaction family and reactions are only added if all the constraints are met. Constraints can be defined on the atomic level of the reactive atoms, e.g. hybridization, number of neighbors, etc., or on the molecule in its entirety such as the molecule size, the presence of rings or aromaticity, etc.

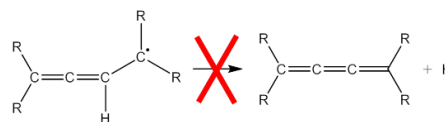
Reaction family: homolytic C-H β scission in presence of double bond in γ position



Constraint 1: single bond between radical C and H-bearing C



Constraint 2: sp^2 C's in double bond, not sp



Constraint 3: radical C not part of ring

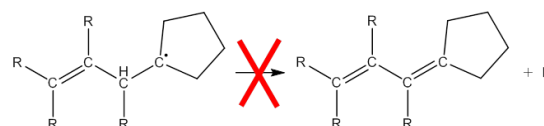


Figure 6: Reaction family constraints. Per reaction family (i.e. a β -scission), constraints can be defined limiting the application domain of this family for a given reaction system. The first constraint limits this reaction family to the formation of double bonds only, triple bonds are not allowed. The second constraint dictates that no allenic structures can be formed. Finally, the third constraint limits the reaction family to open-chain structures, the double bond cannot be formed in a ring or with only one atom in the ring.

A third termination method is based on the maximum number of reaction steps, i.e. rank-based model enlargement, to obtain all products.^{98,99}

Other mechanism-limiting methods are employed *a posteriori* in contrast to the *a priori* techniques described above. These methods are mechanism reduction algorithms that grasp the essential chemistry in a large model and remove the redundant species and reactions. Mechanism reduction to skeletal mechanisms¹⁰⁰⁻¹⁰², lumping^{29,78,79,103-107}, quasi-steady state approximation^{26,108}, and sensitivity analyses¹⁰⁴ are well known and widely used methods.^{10,109} Path flux analysis can also be used to reduce the mechanism.¹¹⁰ Model reduction is in particular important in the context of Computational Fluid Dynamic simulations where only small kinetic models are allowed due to the computational cost of the calculations.^{111,112} Kinetic models with a few hundred of reactions can already be handled.¹¹³

1.2.2 From reaction mechanism to kinetic model: data calculation

To allow reactor simulations, properties of all the species and rate coefficients of all the reactions in a mechanism must be provided. Species properties refer to thermodynamic parameters and optionally transport properties of the molecules. The latter are not considered in this work.

Measuring all these parameters experimentally is unfeasible and accurate quantum mechanical calculations remain computationally extremely demanding. Therefore, several methods have been developed which require only a fraction of the computational resources needed for *ab initio* calculations.

1.2.2.1 Thermodynamics

The thermodynamic parameters are often represented by NASA polynomials¹¹⁴ describing the temperature dependence of the enthalpy of formation, the entropy, and the heat capacity, c.f. Eq. 1.3, Eq. 1.4, and Eq. 1.5.

$$\frac{C_p(T)}{R} = a_1 + a_2 \cdot T + a_3 \cdot T^2 + a_4 \cdot T^3 + a_5 \cdot T^4 \quad \text{Eq. 1.3}$$

$$\frac{\Delta_f H(T)}{R} = a_1 \cdot T + \frac{a_2}{2} \cdot T^2 + \frac{a_3}{3} \cdot T^3 + \frac{a_4}{4} \cdot T^4 + \frac{a_5}{5} \cdot T^5 + a_6 \quad \text{Eq. 1.4}$$

$$\frac{S(T)}{R} = a_1 \cdot \ln(T) + a_2 \cdot T + \frac{a_3 \cdot T^2}{2} + \frac{a_4 \cdot T^3}{3} + \frac{a_5 \cdot T^4}{4} + a_7 \quad \text{Eq. 1.5}$$

Aside from the calculation of enthalpies of reaction, thermodynamic data are needed once reactions are implemented as reversible reactions, i.e. to account for thermodynamic consistency. If no experimental or theoretical values are available for a species, calculation procedures are used to obtain the enthalpy, entropy, and heat capacities.¹¹⁵ These methods are fast and have been proven to be accurate for a wide range of molecules; these two characteristics are extremely valuable for automatic kinetic model generation where thermodynamic properties of many species are required.

1.2.2.1.1 Group additivity

The main method to calculate thermodynamic parameters of ideal gases is group additivity, as introduced by Benson.^{116,117} The basic assumption is that thermodynamic values of an entire molecule can be determined from contributions of all the groups within a species. A group was defined by Benson as a central polyvalent atom with all of its neighbors. Such a group is labeled X-(A)_i(B)_j(C)_k(D)_l with X the central atom having *i* neighbors of atom A, *j* of atom B, and so on. To differentiate between atoms having a double, triple or aromatic bond, a subscript is used. For instance, a carbon atom can be single (C), double (C_d) or triple (C_t) bonded, and can be part of an aromatic ring (unfused: C_b or fused: C_{bf}) or can be allenic (C_a). For example, propylene consists of three groups, as shown in Figure 7. In Figure 7a, the notation for the first atom is trivial,

however, the notation for the second and the third carbon atom (both C_d) does not include all their ligands. When only considering hydrocarbons, a double bonded carbon atom is always double bonded to another carbon atom and, hence, implies another C_d as ligand, which is thus left out of the notation. As a corollary, when heteroatoms are considered, sp^2 carbon atoms bonded to a hetero-atom, such as the oxygen atom in a carbonyl groups, need to be noted differently, which is done by the introduction of a diatomic center instead of a central atom, e.g. (CO), (CS), (CN) corresponding with $-C=O$, $-C=S$, $-C=N-$ etc., c.f. Figure 7b. The same principle is applicable to aromatic carbon atoms, always neighboring two other aromatic carbon atoms in the case of an unfused ring and three in the case of a fused ring, to triple bonded carbon atoms and to allenic carbon atoms.

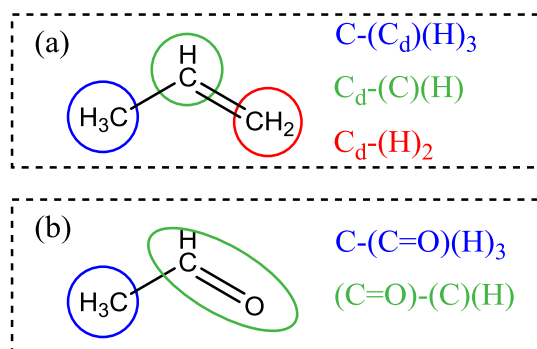


Figure 7: (a) Benson groups in propylene. Propylene has three polyvalent atoms. The first one, depicted in blue, is a sp^3 carbon atom (notation: C) bonded to three hydrogen atoms and an sp^2 carbon atom (notation: C_d). The second atom (in green) has an sp^2 hybridization. It is thus implied that, when only considering hydrocarbons, this atom is double bonded to another C_d atom. Therefore, the latter is not added in the Benson group notation and only the sp^3 carbon atom and the hydrogen atom are mentioned as ligands. (b) Benson groups in acetaldehyde. The carbonyl group is seen as one polyvalent center to differentiate between a C=C bond and a C=O bond.

The thermodynamic properties of molecules and radicals can be calculated by adding a contribution for each group in the species. Group additivity values (GAVs) initially originated from experimental data, but Yamada et al.¹¹⁸ showed that computational chemistry calculations can also be used to derive these GAVs if experimental data are lacking. To account for interactions in the molecule reaching further than the nearest neighbors of each atom, non-next-nearest neighbor interactions (NNI, for example gauche or cis interactions), c.f. Figure 8, resonance corrections (RES) and ring strain corrections (RSC) have been introduced.^{8,119} The thermodynamic properties are accordingly calculated using Eq. 1.6. Several publications

summarize GAVs for species and radicals^{8,119-121}, and stand-alone codes have been written to determine thermodynamic properties based on group additivity.¹²²⁻¹²⁴

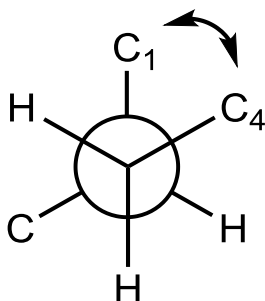


Figure 8: Gauche 1-4 non-next-nearest neighbor interaction (NNI) in 2-methyl butane. Carbon atom number 1 and carbon atom number 4 are not bonded to each other, but their H-ligands do interact which influences the thermodynamics of the molecule. It is thus necessary to account for this contribution when using group additivity for the determination of thermochemical properties.

$$\begin{cases} \Delta_f H^\circ \\ S^\circ \\ C_p^\circ \end{cases} = \sum_{j=1}^n GAV(X_j) + \sum_{j=1}^m NNI_j + \sum_{j=1}^q RSC_j + \sum_{j=1}^p RES_j \quad \text{Eq. 1.6}$$

The entropy calculated via group additivity still needs to be corrected to account for the molecular symmetry¹²⁰, c.f. Eq. 1.7.

$$S = S_{GAV} - R \cdot \ln\left(\frac{\sigma}{n_{opt}}\right) \quad \text{Eq. 1.7}$$

S_{GAV} is the entropy calculated with Eq. 1.6, and S is the entropy for the molecule. R is the universal gas constant, σ the total symmetry number and n_{opt} is the number of optical isomers of the species. The total symmetry number is the number of identical configurations obtained both by rotating the total molecule around symmetry axes as well as by internal rotations around

bonds. Note that the factor $-R \cdot \ln\left(\frac{1}{n_{opt}}\right)$ in Eq. 1.7 arises when enantiomers are lumped together

and considered indistinguishable. The latter assumption is acceptable since enantiomers have identical thermochemical properties and is often applied when lumping enantiomers does not affect the modeling results. In contrast, diastereomeric relations between molecules cannot be treated in the same way as was done for enantiomers. Kinetic model generation tools that distinguish between chiral or cis-trans stereoisomers not only require a molecule representation

that accounts for stereocenters, but also require modifications that allow the creation and destruction of stereocenters in molecules. Only very recently, a proof of concept of stereochemistry-aware kinetic model generation was presented and applied.¹²⁵

GAVs for radicals are more complex to obtain, hence Lay et al.¹²⁶ developed the hydrogen bond increment (HBI) method in which the thermochemistry of a radical $R\cdot$ is calculated by the sum of the values of the parent molecule RH and a contribution to account for the loss of the hydrogen atom.

The above described methods have been shown to calculate thermochemical data within 4 kJ mol^{-1} , i.e. with so-called chemical accuracy similar to high-level *ab initio* calculations.^{8,119,121} The accuracy and applicability of group additivity depends on the availability and accuracy of the GAVs. GAVs for hydrocarbons have been determined and are regularly updated when new experimental or theoretical data became available. Furthermore, GAVs can easily be implemented in computer software to calculate thermochemistry very quickly. Therefore, group additivity offers an excellent way to calculate the thermodynamic parameters in large reaction mechanisms.

Several schemes for additivity have been developed next to the group additivity scheme proposed by Benson. Platt^{127,128} introduced a method to predict the enthalpies of formation, heats of vaporization, boiling points, etc. of paraffins based on the number of C-C and C-H bonds and extra contributions for neighboring bonds. Greenshields and Rossini¹²⁹ calculated the isomeric variation of the same properties as Platt based on bonds, pairs of bonds and other structural features. Thinh et al.^{130,131} introduced a methodology to calculate the heat capacities, standard heats of formation, standard entropies of formation, standard free energies of formation, and absolute entropies using central polyvalent atoms or polyatomic central moieties to defined the group contributions. Smith^{132,133} introduced an approach to calculate standard enthalpies of formation of alkanes and radicals from an additive bond-energy scheme, including additional contributions for interactions between neighboring hydrogen atoms.

1.2.2.1.2 Computational chemistry

Ab initio methodologies, which calculate the electronic energy of a species by approximately solving the corresponding Schrödinger equation, are used to locate local minima, resembling

molecules or radicals, or first-order saddle points, which approximately resemble the transition state of a reaction with pronounced barrier, on the potential energy surface. This equation is very difficult to solve and can only be approximated, using electronic structure models, when the species has more than two atoms. Two major electronic structure methods have been developed: wavefunction-based and density-based both including semi-empirical as well as *ab initio* methods.

Wavefunction-based *ab initio* methods, based on the Hartree-Fock (HF) model, solve the Schrödinger equation by approximating the many-electron system as a set of one-electron systems. This implies that the electron interactions are only calculated as mean fields. The results are the eigenvalues of the Hamiltonian corresponding to the wavefunctions describing the electronic states. The lowest energy found corresponds to the ground state. Post-HF methods, such as Coupled Cluster¹³⁴ or Møller–Plesset perturbation theory^{135,136} methods, include electron correlation effects in the electronic structure calculations to obtain more accurate results. Several semi-empirical models, e.g. the Parameterized Model series^{137,138}, have been developed that can be categorized as wavefunction-based methods.

In the density-based methodology, such as density functional theory¹³⁹ (DFT) functionals are used that express the energy based on the electron density. To obtain an approximation for the electronic kinetic energy, Slater determinant wavefunctions are often used, as introduced by Kohn and Sham¹⁴⁰, which represent a fictitious system of non-interacting electrons. Examples of this methodology are B-LYP¹⁴¹ and Minnesota functionals¹⁴². They employ different approximations of the electron-correlation energy. Here, semi-empirical methods have also been developed in which the parameters to obtain the exchange-correlation energy are fitted to experimental data.

Wavefunction-based and density-based methods can also be combined, as it is done in some composite methods. Here, DFT methods are generally used to optimize geometries and calculate frequencies, while a series of high-level single point quantum chemical calculations is used to calculate the energy. Examples are the complete basis set (CBS)^{143,144}, Gaussian (G)^{145,146} or Weizmann (W)^{147,148} methods in which DFT is used to calculate the geometry and the post-HF methods calculate the energies.⁵³

Starting from a reasonable initial guess, the geometry can be optimized using several levels of theory by making use of the Hessian matrix, which is the second derivative of the energy to the coordinates of the atoms in the molecule, and is thus a force constant matrix. The accuracy of the geometry optimization typically increases with the computational cost of the calculations. Semi-empirical methods are relatively fast, hence large molecules can be handled in a reasonable calculation time, but the results are less accurate compared to the computationally more intensive *ab initio* methods. It is not the purpose of this chapter to give an extensive review on the capacities, applications, and limitations of the available *ab initio* methods because other reviews cover this subject, e.g. Klippenstein et al.¹⁴⁹.

From quantum chemical calculations, thermodynamic properties are obtained by making use of molecular partition functions q . To calculate these partition functions the molecular frequencies, calculated from the Hessian, among other properties are needed, which originate from *ab initio* calculations. By assuming that all the modes (vibrations, rotations, translation and electronic modes) in a molecule are separable, the partition functions can be calculated by a multiplication of the electronic, translational, vibrational, and rotational partition functions, as shown in Eq. 1.8. The calculation of each contribution is explained in Chapter 3.

$$q_{tot} = q_{elec}q_{trans}q_{rot}q_{vib} \quad \text{Eq. 1.8}$$

From the partition functions, which is a microscopic property, macroscopic properties can be calculated, such as the internal energy U , the enthalpy H , the entropy S , and the heat capacity C_p .

$$U(T) = RT^2 \left(\frac{\partial \ln(q(T))}{\partial T} \right) \quad \text{Eq. 1.9}$$

$$H(T) = U(T) + RT \quad \text{Eq. 1.10}$$

$$S(T) = R \left(1 + \ln \left(\frac{q(T)}{N_A} \right) + T \frac{\partial \ln(q(T))}{\partial T} \right) \quad \text{Eq. 1.11}$$

$$C_p(T) = R \left(1 + 2T \frac{\partial \ln(q(T))}{\partial T} + T \frac{\partial^2 \ln(q(T))}{\partial T^2} \right) \quad \text{Eq. 1.12}$$

Finally, the standard enthalpy of formation can be calculated using the atomization method, in this the subscript $\Delta_a H$ stands for the atomization enthalpy.

$$\Delta_f H^\circ(X_m Y_n) = m\Delta_a H^\circ(X) + n\Delta_a H^\circ(Y) - (mH(X) + nH(Y) - H(X_m Y_n)) \quad \text{Eq. 1.13}$$

It is important to note in the context of computational chemistry methods that a large number of conformers exist for most species and transition states. For example, n-butane has three local minima on its PES, 2 unique configurations, n-pentane already has nine with 4 unique configurations, and n-hexane has 27 out of which 14 are unique. The problem becomes even more complex in cases where ring structure or polar moieties are present in the molecule. Several computer codes and algorithms have been developed in order to search for a low-energy conformer and the lowest-energy one. Two main categories of algorithms can be distinguished: deterministic and stochastic algorithms.^{150,151} Deterministic methods scan the torsional space of the molecule on a grid. If the number of rotamers is high enough, the lowest-energy conformer can be found. It is however, not possible to know how many rotamers are necessary to assure a convergence to the lowest-energy conformer, and the torsional scans are computationally intensive. Therefore, the number of rotamers is often reduced to a small number, losing the guarantee of finding the lowest-energy conformer. Stochastic methods only search for subsets of the full conformational space, for example by only scanning the rotation of one bond.¹⁵⁰ Monte Carlo methods^{152,153}, “distance geometry” models¹⁵⁰, genetic algorithms¹⁵⁴ and random perturbations¹⁵⁵ are well-known stochastic methods, which locate local minima but not always the global minimum. A detailed description of each method can be found elsewhere.¹⁵⁶⁻¹⁶⁰

1.2.2.1.3 *On-the-fly computational chemistry*

Although work has already been done for cyclic compounds, e.g. Ince et al.¹⁶¹, group additivity is not always accurate for ring structures such as polycyclic species where almost every molecule needs a unique ring strain correction to obtain adequate results. Also for many molecules, GAVs are simply not available (yet). Therefore, even though it is more time consuming, calculating the thermodynamics on-the-fly using computational chemistry calculations offers alternatives. These calculations have been incorporated in kinetic model generation.¹⁶²⁻¹⁶⁴ The incorporation of on-the-fly computational chemistry calculations into automated kinetic model generation applications requires a reasonable initial guess for the three-dimensional coordinates of the atoms in the molecule. Automated kinetic model generation applications employ a connectivity-based molecular representation throughout the various steps of the kinetic model generation, and hence

require a conversion of the graph-based molecular structure to a format in which explicit three dimensional (3D) atomic coordinates are stored. 3D coordinates generation tools have been developed for decades, and applied in various chemo- and bioinformatics frameworks.¹⁶⁵ In the context of kinetic model generation applications, three key parameters of the 3D coordinates generation tool are important: 1) the robustness to handle a wide variety of molecular structures that arise upon the execution of reaction families, 2) the accuracy of the produced atomic coordinates to ensure an adequate initial guess for the subsequent geometry optimization, and 3) the execution speed of the program, and hence the potential to process hundreds of molecular structures within a reasonable timeframe. Recently, Ebejer et al.¹⁵¹ performed a benchmark study in which these parameters were evaluated. Illustrations of the integration of on-the-fly computational chemistry calculations into automated kinetic model generation applications are scarce. Broadbelt et al.¹⁶⁴ proposed a rule-based 3D coordinates generation algorithm tailored to molecules encountered during the generation of hydrocarbon pyrolysis models and integrated MOPAC¹³⁷ in the kinetic model construction procedure.¹⁶⁶ Magoon and Green¹⁶³ integrated semi-empirical calculations by MOPAC2009 and the Gaussian 03 suite of programs¹⁶⁷ and force-field calculations from MM4¹⁶⁸ through a 3D coordinates generation interface using a “distance geometry” approach by Blaney et al.¹⁶⁹ as implemented in RDKit, initially developed at Novartis.¹⁷⁰ By using this algorithm for a set of 43 polycyclic hydrocarbons (propellanes) and comparing the thermodynamic properties to experimentally measured ones¹⁷¹, a mean absolute deviation (MAD) of 29 kJ mol⁻¹ was observed for the PM3 method compared to a MAD of 167 kJ mol⁻¹ when using group additivity and 124 kJ mol⁻¹ for the MM4 method. No GAVs were calculated for this system, and no ring strain corrections are available. It is thus suspected that the GAV method does not perform well. An update of ring strain corrections could increase the performance, but as stated above, for most molecules a separate single ring strain correction is necessary, which defeats the original purpose of GAVs being a fast and scalable calculation method.

1.2.2.2 Rate coefficients

The calculation methods for reaction rate coefficients are based on the idea that the reactive center of reactions within a reaction family is the same, or at least highly resembling, for all the reactions in that family. For each reaction belonging to the family, the rate coefficient can be calculated based on a contribution for the reactive center alongside reaction-specific kinetic

contributions. In most cases, rate coefficients are written in the Arrhenius or modified Arrhenius form, describing the temperature dependence of the reaction (Eq. 1.14 and Eq. 1.15 respectively).

$$k(T) = A \cdot \exp\left(-\frac{E_a}{R \cdot T}\right) \quad \text{Eq. 1.14}$$

$$k(T) = A \cdot \left(\frac{T}{1 \text{ K}}\right)^n \cdot \exp\left(-\frac{E}{R \cdot T}\right) \quad \text{Eq. 1.15}$$

Here, A is the pre-exponential factor, E_a is the activation energy of the reaction, n is sometimes introduced to account for temperature dependence of A and E_a ¹¹⁶, R is the universal gas constant and T the absolute temperature.

1.2.2.2.1 Semi-empirical methods to calculate rate coefficients

A general class of methods to quickly calculate the rate coefficient or the activation energy of a reaction uses a linear free-energy relationship:

$$\ln(k_i(T)) = \ln(k_0(T)) + m \cdot (x_i - x_0) \quad \text{Eq. 1.16}$$

Where $k_0(T)$ is the rate of a known reaction belonging to the same reaction family, m is determined per reaction family, and x_i is a property of reaction i or the species in reaction i (the 0 subscript refers to the known reaction). The first set of linear free-energy methods compares the reaction enthalpy of a reference reaction (with known kinetics) to the reaction enthalpy of reaction i to calculate $k_i(T)$. Evans and Polanyi¹⁷² first introduced this approach in which the pre-exponential factor is considered constant for all reactions within a reaction family and the activation energy depends linearly on the reaction enthalpy (Eq. 1.17).

$$E_a^i = E_a^0 + \gamma \cdot \Delta_r H^\circ_i \quad \text{Eq. 1.17}$$

Here, E_a^i is the activation energy of reaction i and $\Delta_r H^\circ_i$ is the standard reaction enthalpy of reaction i , the ‘°’ symbol refers to standard conditions, i.e. at a pressure of 1 bar and a chosen temperature. γ is the transfer coefficient for the reaction family and E_a^0 is the intrinsic barrier, which is the activation energy for a thermo-neutral reaction, hence noted with the ‘0’ superscript. For highly exothermic reactions, the Evans-Polanyi relationship leads to negative activation energies, but this is avoided by truncating the relationship to zero. Furthermore, if an Evans-Polanyi model is constructed, the validity range should be reported and one should be careful when extrapolating outside this range.

However, for many reactions the activation energy shows nonlinear dependence of the activation energy on the reaction enthalpy.¹⁷³ Alternative relationships were proposed in the literature, such as the Marcus equation (Eq. 1.18)¹⁷⁴, which predicts the activation energies well for mildly endothermic or exothermic reactions.

$$E_a^i = E_a^0 \cdot \left(1 + \frac{\Delta_r H^{\circ}_i}{8 \cdot E_a^0} \right)^2 \quad \text{Eq. 1.18}$$

Roberts and Steele¹⁷⁵ extended the Evans-Polanyi relation by introducing the role of polar effects. This improves the accuracy of the activation energies, as demonstrated by comparison to literature values, but suffers from a limited applicability.

Blowers and Mase^{173,176} extended the approach of Evans and Polanyi by accounting for the nonlinear behavior. Three regimes were proposed to calculate the activation energy based on the reaction enthalpy. In the first regime, an activation energy of zero is assigned if the reaction enthalpy is less than $-4 \cdot E_a^0$. In the second regime, the activation energy equals the reaction enthalpy if the latter is larger than $4 \cdot E_a^0$. The last regime, i.e. if the reaction enthalpy lies in between $-4 \cdot E_a^0$ and $4 \cdot E_a^0$, the activation energy is calculated according to Eq. 1.19, where V_p is a parameter related to the intrinsic barrier, depending on w_B , w_F and E_a^0 .

$$E_a = \frac{\left(\frac{w_B + w_F + \Delta_r H^{\circ}}{2} \right) (V_p - w_B - w_F + \Delta_r H^{\circ})^2}{V_p^2 - (w_B + w_F)^2 + \Delta_r H^{\circ 2}} \quad \text{Eq. 1.19}$$

The difference between this approach and the one of Evans and Polanyi is illustrated in Figure 9 for a set of hydrogen abstraction reactions by the hydrogen radical on a variety of hydrocarbons. If the full data set is regressed to either an Evans-Polanyi or a Blowers-Masel relationship, the Blowers-Masel model predicts the trends better than the Evans-Polanyi model. However, when the dataset is divided in three parts based on the nature of the carbon atom from which the hydrogen is abstracted, Evans-Polanyi models can predict the activation energies well. In Figure 9, the fraction of the reactions with the highest exothermicity are hydrogen abstractions leading to the formation of resonantly stabilized radicals. The second part are abstractions from an sp^3 carbon atom and the third fraction are endothermic reactions that lead to the formation of vinylic or phenylic radicals.

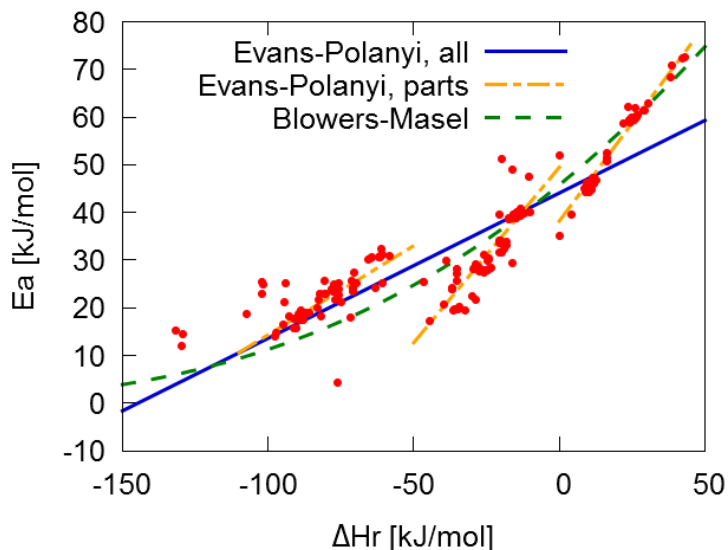


Figure 9: Performance comparison of the Evans-Polanyi (blue line) and Blowers-Masel (green line) methods in predicting the activation energies of hydrogen abstraction reactions by the hydrogen radical. The red points represent calculated activation energies / enthalpies of reaction pairs at the CBS-QB3 *ab initio* level of theory.¹⁷⁷ Also shown is the performance of the Evans-Polanyi model for the three distinct subsets of the data, which reflect different radical types.

1.2.2.2.2 Group contribution methods for Arrhenius parameters

As described in section 1.2.2.1, group contribution methods were initially developed to calculate the thermochemistry of molecules and radicals in a scalable manner, by adding a contribution for each constituting group of the species. This approach can also be applied to the calculation of rate coefficients.^{120,178-188} An expression for rate coefficient calculations based on transition state theory is given in Eq. 1.20 in which k_B is the Boltzmann constant, T the absolute temperature, V_m° the standard molar volume and h is the Planck constant. $\Delta^\ddagger S$ and $\Delta^\ddagger H$ are the standard entropy of activation and the standard enthalpy of activation. The number of single events n_e of the reaction is obtained via Eq. 1.21, in which σ is the total rotational symmetry number and n_{opt} is the number of optical isomers. Eq. 1.21 not only provides a convenient way to determine the number of single events, but also translates it into the more tangible problem of the determination of symmetry numbers. Nevertheless, the determination of the number of single events and symmetry number remains a challenging problem to automate, with many attempts only providing appropriate values in a limited subspace of molecules or reactions.¹⁸⁹⁻¹⁹¹ Moreover, more generally applicable solutions to the algorithmic evaluation of the symmetry of molecules and reactions will inevitably be hampered by *ad hoc* complications originating from stereoisomerism and conformational degrees of freedom present in reactants and/or transition

states.^{192,193} Finally, the contribution of quantum mechanical tunneling to the rate coefficient is incorporated via the tunneling coefficient κ .

Since group additivity was shown to calculate thermodynamic data of molecules and radicals with good accuracy, the same method was successfully extended to rate coefficient calculations. As Eq. 1.20 shows, calculating the thermodynamic properties of transition states and reactants suffices to calculate the rate coefficients of hydrogen abstraction reactions and the high-pressure limits of unimolecular reactions. Hence, the idea to use of group additivity to predict rate coefficients is based on the assumption that thermodynamic properties of transition states can be calculated with similar success compared to species.

$$k(T) = n_e \kappa(T) \frac{k_B T V_m^\circ}{h} \exp\left(\frac{\Delta S^\ddagger}{k_B}\right) \exp\left(-\frac{\Delta H^\ddagger}{k_B T}\right) \quad \text{Eq. 1.20}$$

$$n_e = \frac{n_{opt,\ddagger}}{\prod_r n_{opt,r}} \cdot \frac{\prod_r \sigma_r}{\sigma_\ddagger} \quad \text{Eq. 1.21}$$

Group additivity has several advantages. First, compared to theoretical calculations, it can be used independent of the type of reactions or species, provided that data is available. It provides a fast, easy, and scalable calculation of rate coefficients for automatic kinetic model generation codes. Theoretical calculations need user-expertise once a new reaction family is encountered or once other types of atoms are present in the species. Secondly, group additivity can be extended with non-next-nearest neighbor interactions, resonance corrections, ring strain corrections, and so forth, to improve the accuracy. Empirical methods, on the other hand, only use a small number of parameters for one reaction family and no extra contribution can be implemented.

Several methods have been developed employing an additive approach. One such method uses experimentally determined rate coefficients of a few reactions to calculate the rate coefficient of other reactions. This has been used to describe complex processes such as the oxidation of volatile organic compounds in the atmosphere. Aumont et al.⁷⁶ reported a set of structure-activity relationships for several reaction types such as hydroxyl radical reactions, NO₃ hydrogen abstractions, ozone based reactions, etc. Kwok and Atkinson¹⁹⁴ developed a scheme to calculate rate coefficients of hydroxyl radical reactions using a group contribution method, giving the total reaction rate of a hydroxyl radical with an organic compound, which can be seen as a lumped reaction rate. The addition of hydroxyl radicals to multiple bonds was studied by Peeters et al.¹⁹⁵,

who developed a structure-activity relationship calculated the rate coefficients by a sum of partial site-specific rate coefficients. The latter depends on the stability of the resulting β -hydroxyl radical.

A second method is based on a group additive scheme for Arrhenius parameters, where the rate coefficients of reactions are calculated based on the rate coefficient of a reference reaction belonging to the same reaction family and a contribution for the structural differences between the reaction of interest and the reference reaction.^{120,178-181,185,186,196-200} For every group present in the reactive center, a contribution to the different parameters is taken into account and is called the primary contribution. This is illustrated in Figure 10, the reactive center consists of the three carbon atoms, each bearing several ligands that can influence the reaction rate coefficients.

Secondary contributions can also be obtained if the reaction rate coefficients depend on sub-molecular groups which are not part of the reactive center. If a reactive center is constructed by n atoms, the pre-exponential factor and activation energy can be calculated by Eq. 1.22 and Eq. 1.23 respectively.

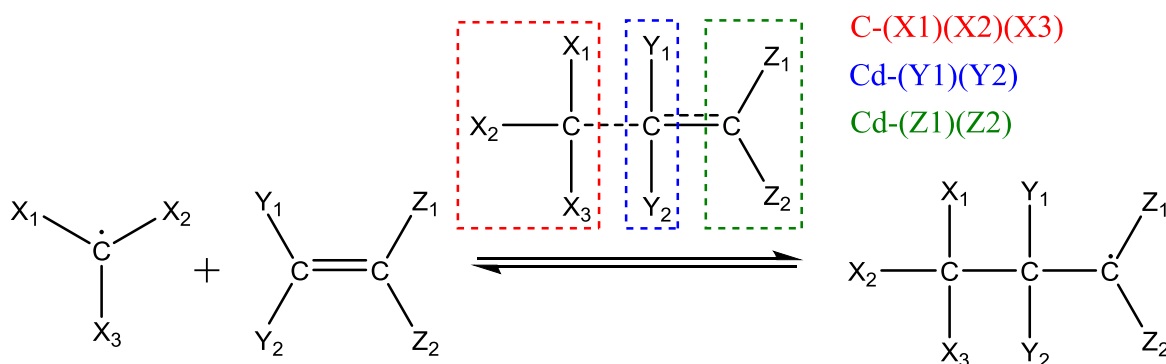


Figure 10: Groups to calculate the reaction rate of a radical addition to a double bond.

$$\log \tilde{A} = \log \tilde{A}_{ref} + \sum_{i=1}^n \Delta GAV_{\log \tilde{A}}^0(X_i) \quad \text{Eq. 1.22}$$

$$E_a = E_{a,ref} + \sum_{i=1}^n \Delta GAV_{E_a}^0(X_i) \quad \text{Eq. 1.23}$$

Here $E_{a,ref}$, and \tilde{A}_{ref} are the Arrhenius parameters of the reference reaction. ΔGAV^0 is the group additivity value of a group within the reactive center, relative to the reference reaction. An important advantage of this method is that only the differences between the reactants and

transition state are needed, the thermodynamic properties of the transition states are not necessary, in contrast to the “supergroup” approach of Green et al.¹⁸²⁻¹⁸⁴ The outcome of these equations can subsequently be employed to obtain the reaction rate coefficient.^{180,181}

$$k(T) = \kappa n_e \tilde{A} \exp\left(-\frac{E_a}{RT}\right) \quad \text{Eq. 1.24}$$

While the Arrhenius parameters A and E_a can be modeled by the summation of groups of localized sub-molecular fragments in the transition state, the rate coefficient k does contain two contributions that are non-local in nature and hence should be evaluated for every reaction separately, cf. Eq. 1.24. The contribution κ to the rate coefficient represents tunneling of atoms through the reaction barrier. Marin and coworkers^{120,181,185,186,196-198,200} used an empirical power law, correlating the tunneling coefficient κ with temperature and the activation energy in the exothermic direction. This proved to be a fast and accurate method to obtain adequate values for the tunneling coefficient without the need for the imaginary frequency of the motion along the reaction coordinate or the potential energy surface of the reaction. The contribution n_e represents the number of single events. Note that non-local contributions are excluded from the rate coefficient of the reference reaction as well. As a result, when evaluating the rate coefficient of a specific reaction, a separate contribution from tunneling and a separate contribution from the number of single events should be accounted for in order to calculate the rate coefficient of the reference reaction and to determine the rate coefficient of the studied reaction.

To determine the group additivity values for a certain reaction family, Marin and coworkers^{120,180,181} started by defining reference reactions for which the kinetics were calculated theoretically. Next, a list of reactions is defined in which all structural features that reactants can possess and are anticipated to influence the kinetics significantly are at least once considered. While group additivity values corresponding to each structural feature should in principle be determined statistically independent, it is not always possible as certain structures are always present together, i.e. group additivity values can be linearly dependent.⁸ In these cases, one of the groups is assigned a predefined ΔGAV^0 and the ΔGAV^0 s of the other structures will depend on this choice. However, similarly to thermodynamics, this arbitrary choice of the reference value does not influence the final Arrhenius parameters.

A third group additive method is based on the thermodynamic properties of the reactants and transition state of a reaction.¹⁸²⁻¹⁸⁴ New group additivity values can be calculated for transition state specific centers based on *ab initio* calculations. Green et al.¹⁸²⁻¹⁸⁴ first used ‘supergroups’ instead of the conventional Benson groups. Subsequently, these supergroups, which contained more than one polyvalent atom with its ligands, were subdivided into groups in line with the Benson definition. For this, distinction was made between methyl, primary, secondary and tertiary carbon atoms around the abstracted hydrogen atom, for hydrogen abstraction reactions. The groups defined around this abstracted hydrogen atom bearing two distinct ligands, H-(X₁)(X₂) in which X can be a hydrogen atom, a methyl group or a primary, secondary or tertiary carbon atom, result in 11 sets of GAVs while the GAVs of the groups with two equal ligands were defined as zero.

The rate coefficients can be calculated using Eq. 1.20. If the enthalpy of formation and entropy of both the reactants and transition state are calculated at the same level of theory one can expect the accuracy to improve as systematic errors partially cancel out, however, inducing that the thermodynamic properties of the transition state have to be known. As stated previously the method of Marin and coworkers does not require the thermodynamic properties of the transition state.

An advantage is that the supergroup approach does not require an Arrhenius form of the rate coefficient as the latter can be calculated at any temperature. To account for this, the group additivity schemes calculating Arrhenius parameters have to report ΔGAV^0 s at several temperatures or include ‘n’ as an extra parameter, Eq. 1.15. However, for reaction families such as beta-scissions and radical additions, Sabbe et al.¹²⁰ showed that ΔGAV^0_{Ea} varies by less than 4.5 kJ/mol in 90% of the cases for a temperature range between 300 and 1300K. In the same study, $\Delta GAV^0_{\log A}$ varies less than 0.4.

Green et al.¹⁸²⁻¹⁸⁴ deduced GAVs alongside non-nearest-neighbor effects for hydrogen abstraction reactions based on the thermodynamic properties of the reactants and transition states. The deviations of the use of group additivity versus the direct rate calculations via transition state theory were reported: for the full temperature range 300-500K, the deviations of the rate coefficients for hydrogen abstractions by hydrogen radicals do not exceed 30%.

1.2.2.2.3 *Tree structure averaging*

The rate coefficient of a reaction is largely determined by the reactive atoms in the specific reaction, i.e. adjacent to where bonds are being formed or broken. Identifying the reactive center allows one to calculate the reaction rate coefficient, and hence, methods have been developed that exploit this characteristic feature. The reactive center can be decomposed into its constituting reactive atoms. For example, in H-abstraction reactions $XH + Y\cdot \rightarrow X\cdot + YH$ the reactive center would be the abstracting atom (Y) and the atom from which a hydrogen is abstracted (X).

In the “rate rules” implementation in RMG¹⁸⁻²¹, the groups X and Y are used to locate the transition state supergroup XHY in the database of rate rules, and the appropriate rule is used to as reaction rate coefficient. This rule may be deduced from an Evans-Polanyi expression, but is often a reaction rate coefficient for a specific reaction containing that supergroup, either from a quantum chemistry and transition state theory calculation or taken from literature. When a rate rule is not available for XHY, the rate coefficients of supergroups close to it in the database are averaged. For example if there is no rule available for X_xHY_y , there may be data for $X_{i\neq x}HY_y$ and/or $X_xHY_{j\neq y}$, or, if necessary, for $X_{i\neq x}HY_{j\neq y}$. The database is arranged in a hierarchical tree structure, so that the most specific group definition with data is used, and the closest possible, and hence most chemically similar, rules are averaged.^{20,201} Details are provided in the PhD theses by Jing Song and Joshua Allen and a publication by Sumathy and Carstensen.^{20,202,203} The scheme provides approximations when only few data are available, using the philosophy that even a highly uncertain rate coefficient is better than no rate coefficient at all. Disadvantages of this averaging scheme include that it can obfuscate the source(s) of the final reaction rate coefficient.

1.2.2.2.4 *Reaction class transition state theory*

Truong²⁰⁴ proposed a prediction scheme for a large set of reactions – the reaction class transition state theory (RC-TST) – based on the idea that reactions belonging to the same reaction class have the same reactive center and thus possess similarities on their PES. Several characteristics of the reactions of the same class are equal, which makes it possible to calculate reaction rate coefficients based on known rate coefficients. Given a known rate coefficient k_1 of a reaction, called the principal reaction, rate coefficients of other reactions belonging to the same reaction class can be calculated as shown in Eq. 1.25. The relative rate depends on four coefficients: the

transmission coefficient, the reaction symmetry number coefficient, the partition function coefficient, and the potential energy coefficient.

$$\frac{k_2}{k_1} = f_\kappa \cdot f_\sigma \cdot f_Q \cdot f_V \quad \text{Eq. 1.25}$$

The ratio of transmission coefficients of both reactions as a function of temperature is calculated using Eq. 1.26 in which κ_1 is the tunneling coefficient of the principal reaction and κ_2 the one of the reaction of interest.

$$f_\kappa = \frac{\kappa_2(T)}{\kappa_1(T)} \quad \text{Eq. 1.26}$$

If the tunneling contributions for both reactions are calculated using the same model, e.g. the Eckart model²⁰⁵, according to Truong²⁰⁴ errors cancel out which increases the accuracy. By analysis of the transition states of several reactions belonging to the same reaction class, Truong found that the imaginary frequency, i.e. the one corresponding to the reaction coordinate, is more or less conserved throughout the reaction class. This way, the assumption was made of a constant imaginary frequency and zero-point energy corrections to the barrier and reaction energy in one reaction class. For each reaction, only the differences in barrier height and reaction energy of the principal reaction need to be known to determine the ratio of transmission coefficients.

The reaction symmetry number coefficient f_σ in Eq. 1.25 is calculated as the ratio of the symmetry numbers of both reactions. The partition function coefficient is given in Eq. 1.27 in which all partition functions are products of the translational, rotational and vibrational partition functions. ‡ is the mark for a transition state and r marks the reactants.

$$f_Q = \frac{\left(\frac{q_{\ddagger}}{\prod_r q_r}\right)_2}{\left(\frac{q_{\ddagger}}{\prod_r q_r}\right)_1} = \frac{\frac{(q_{\ddagger})_2}{(q_{\ddagger})_1}}{\frac{(\prod_r q_r)_2}{(\prod_r q_r)_1}} \quad \text{Eq. 1.27}$$

The translational and rotational contributions result in a constant multiplication factor in the coefficient. Vibrational contributions, however, depend on temperature. By rearranging the equation to a ratio of the transition state partition functions divided by a ratio of reactant partition functions, Truong showed that the substituent effects in the transition state and reactants yield a constant contribution, which cancels out if there is no coupling between the substituents and the

reactive center. It can thus be expected that the partition function coefficients do not strongly depend on temperature. This was confirmed by analyzing several hydrogen abstraction reactions in which the partition function coefficient only varied with temperature for temperatures below 300K. A constant high-temperature coefficient was subsequently introduced. Furthermore, it was observed that a constant f_Q can be defined for one reaction family. This value is calculated by comparing the values obtained for two reactions, one of which is the principal reaction, by high-level vibrational frequencies calculations.

The potential energy coefficient (Eq. 1.28) accounts for the difference in classical barrier height and thus shows the effects of the substituents on the potential energy of the reactants and transition state. The differential height is determined by calculating the differences in potential energy of the transition state and of the reactants of the reaction compared to the principal reaction as shown in Eq. 1.29.

$$f_V = \exp\left(-\frac{(\Delta V_{\ddagger,2} - \Delta V_{\ddagger,1})}{k_b T}\right) = \exp\left(-\frac{\Delta\Delta V_{\ddagger}}{k_b T}\right) \quad \text{Eq. 1.28}$$

$$\Delta\Delta V_{\ddagger} = (E_{\ddagger,2} - E_{\ddagger,1}) - (E_{r,2} - E_{r,1}) \quad \text{Eq. 1.29}$$

Truong validated the approach for a set of hydrogen abstractions by hydrogen atoms. The principal reaction was the abstraction from methane and the one from ethane was used to calculate f_Q . Transition state theory with Eckart tunneling corrections was used to calculate the rate coefficient of these two reactions. By comparing the obtained rate for these reactions both to experimental and theoretical data, it was shown that by only using the barrier height and the reaction energy, it is possible to predict rate coefficients of reactions belonging to one reaction class reasonably well.

In a subsequent publication, Zhang and Truong²⁰⁶ showed that the reaction class transition state theory together with a linear energy relationship (RC-TST/LER) provides an accurate means for the prediction of rate coefficients. The approach was validated against a set of 46 hydrogen abstractions from a sp^3 hybridized carbon atom by a hydrogen radical. Zhang and Truong derived linear energy relationships - using the 46 reactions - both for hydrogen abstractions from alkanes as well as from alkenes. With the barrier heights obtained from the linear energy relationship, the potential energy coefficient could easily be determined. The symmetry coefficients were directly calculated using the symmetry numbers of the reaction and the ones from the principal reaction.

To calculate the transmission coefficients, using the reaction enthalpy and the barrier height obtained from the linear energy relationship, a constant imaginary frequency for all the reactions was assumed. The partition function factor was taken constant for all the reactions. Comparison of the reaction rate coefficients calculated using the RC-TST/LER method against both experimental data and theoretical rate coefficients (from transition state theory with Eckart tunneling) showed a good agreement. The maximum deviation when comparing the rate coefficients to experimental ones amounted to 137% with an average deviation of 51%, while the deviations between the calculated rate coefficients and the theoretical rate coefficients remained below 100%.

Several articles describing the application of the RC-TST/LER method were published: Kungwan and Truong²⁰⁷ applied the method to hydrogen abstractions from alkanes by a methyl radical, Huynh and Truong²⁰⁸ calculated rate coefficients for hydrogen abstractions by a hydroxyl radical from alkanes, Huynh et al.²⁰⁹ applied the method to abstraction reactions ground state oxygen atoms from alkanes, Huynh and Truong²⁰⁹ investigated the method for abstraction reactions by formyl radicals from alkanes, Ratkiewicz et al.²¹⁰ looked at abstraction reactions by hydrogen radicals from alcohols, Ratkiewicz et al.²¹¹ calculated the rate coefficients of hydrogen abstractions by ethyl radicals from alkanes, Piansawan et al.²¹² applied the method for hydrogen abstractions from alkanes by chlorine radicals, and Wang et al.²¹³ used the method for hydrogen abstractions by hydrogen radicals from methyl esters.

Wang et al.²¹⁴ gave a new interpretation of the reaction class transition state theory using an extension of the isodesmic reaction method to the calculation of reaction barriers. They showed that the results of RC-TST are very sensitive to the level of the *ab initio* method used but the new method does not suffer from this sensitivity.

1.2.2.5 Computational chemistry methods

Analogously to thermodynamic data, rate coefficients have been accurately calculated based on high-level computational chemistry in the last decade(s). Most calculations are based on transition state theory, c.f. Eq. 1.30, in which q is the partition function and $\Delta^\ddagger E$ the energy of activation, i.e. the energy difference between the reactants and the transition state at 0K.

$$k(T) = n_e \kappa(T) \frac{k_B T V_m^0}{h} \frac{q_{\ddagger}}{\prod_r q_r} \exp\left(-\frac{\Delta E^{\ddagger}}{k_B T}\right) \quad \text{Eq. 1.30}$$

Transition state theory can be applied using the partition functions directly, or using the enthalpies and entropies of activation as shown in Eq. 1.20. Both approaches yield exactly the same results.¹⁸⁴

Optimizing stable species is the search for a minimum on the PES, as opposed to transition states, where saddle points on the PES are located. These saddle points are characterized by one imaginary frequency, which corresponds to the pathway along the reaction coordinate, c.f. Figure 11. This saddle point is used to calculate reaction rate coefficients of a reaction as approximation of the transition state, being a hypersurface which separates the products from the reactants. Computational chemistry methods can search for transition states and calculate the electronic energy and molecular frequencies. $\Delta^{\ddagger}E$ and q_{\ddagger} can be calculated to obtain the rate coefficient from Eq. 1.20 or Eq. 1.30, provided that these properties are available for the reactants and products of the reaction.

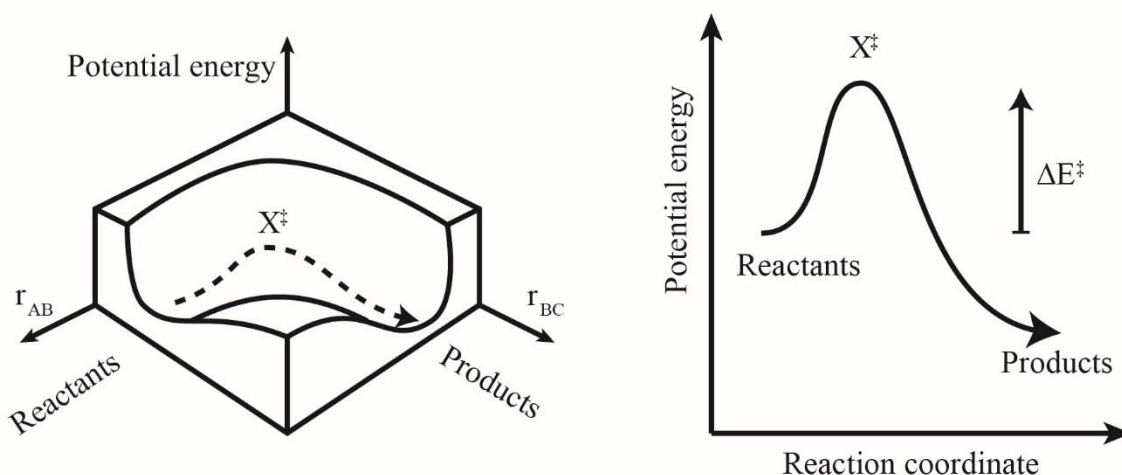


Figure 11: Illustration of a transition state.

1.2.2.2.6 On-the-fly computational chemistry

During automatic kinetic model generation, both the reactants and products are known as a result of the execution of reaction rules. The graph structure of the transition state is thus known and an initial guess of its three dimensional structure can be made. Preliminary work on this methodology has already been done.^{215,216} Bhoorasingh and West²¹⁶ constructed a code to

automatically locate transition states of intermolecular hydrogen abstraction reactions. The geometry of the transition states are constructed based on a “distance geometry” algorithm. Such an algorithm builds the 3D coordinates based on distance limits between each pair of atoms in the species of transition state. For reactive atoms, these distances are not comparable to stable species, and a group contribution method has been developed to calculate the bond lengths of bonds that are formed or broken during the reaction. After optimization of the transition state structure, intrinsic reaction coordinate (IRC) calculations are used to validate the transition state.

Zádor and Najm²¹⁵ introduced a computer code called “KinBot” at Sandia National Laboratories, which is able to automatically explore chemical pathways in an efficient manner for reactions that are relevant in gas-phase systems. Several reaction families are already covered, such as single bond scission, isomerization via hydrogen transfer, cyclic ether formation and direct HO₂ elimination. One of the strengths of this approach is that, although no new reaction classes can be found, new reaction pathways can be discovered. At this point, however, it is computationally too expensive to build an entire mechanism by searching for all possible reactions on the PES. Starting from initial species, reactions can be found by distorting the reactant geometry in many ways until a saddle point on the PES is found. Distorting the equilibrium reactant geometry in this context means perturbations of bond lengths, bond angles and dihedral angles according to previously acquired chemical knowledge. The result of these perturbations is that the code slowly moves on the PES to locate saddle points and subsequently the product structures.

1.2.2.2.7 Accounting for pressure dependence

Until now, only the assignment of high-pressure limit reaction rate coefficients to elementary reactions was described in this work; no attention was given to pressure dependent rate coefficients. However, many steps in gas phase processes such as atmospheric reactions, oxidation, combustion, or pyrolysis can depend on pressure and a full description of the pressure dependence can be necessary to allow numerical simulations under diverse conditions. An example of a rate coefficient as a function of the pressure is given in Figure 12. Both unimolecular and chemically activated bimolecular reactions can be strongly pressure dependent, but most current kinetic model generation codes do not calculate the pressure dependent rate coefficients. Furthermore, experimental data are often measured at relatively low pressure whilst

industrial processes seldom use these low pressures. An accurate extrapolation is necessary when a kinetic model is validated against the low pressure data that are used to predict industrial data.¹⁹

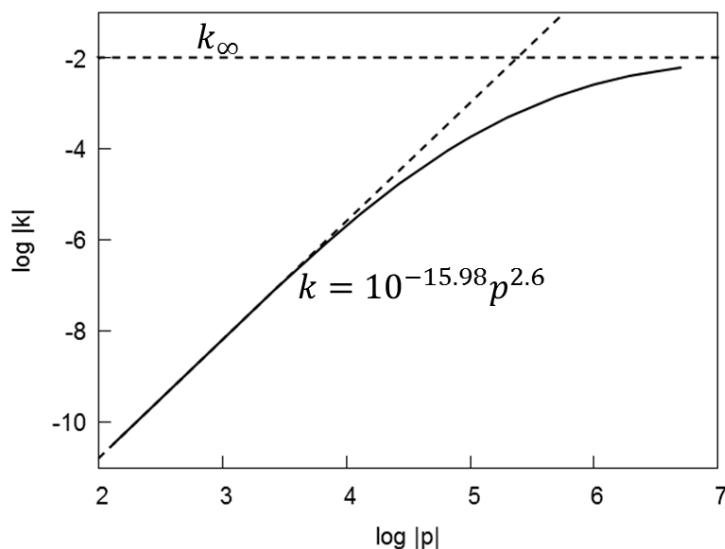


Figure 12: Pressure dependence of the rate coefficient k_{uni} from the reaction of the 1-pentano-3-yl radical to water and the 2-penten-5-yl radical at 900K. The rate coefficients are expressed in s^{-1} and the pressure in Pa.

Especially when dealing with high-temperature low-pressure conditions common in experimental combustion studies ($T \gtrsim 1000 \text{ K}$, $P \lesssim 1 \text{ bar}$) of small molecules, as many as half of the reactions may be in the fall-off regime, i.e. the rate coefficients are lower than the high-pressure limit rate coefficients, and depend on the total pressure. Calculating the density of states from approximated thermochemical data remains a challenge, and a poorly calculated potential energy surface (due to inaccurate thermochemical or kinetic data) can make the resulting Master Equation difficult to solve. There thus remains a trade-off between accuracy, speed, and robustness.²¹⁷ Although this is not straightforward for automatic kinetic model generation programs, it is even harder to do by hand, and most models simply neglect the vast majority of chemically activated reactions, which are reactions that proceed after the collision of the reactant with another molecule. This collision creates an activated complex which decomposes to the products of the reaction.

A potential energy surface of a pressure dependent reaction contains several wells, bimolecular products and saddle points, but only a few of them contribute significantly to the chemistry described by the envisioned kinetic model. Often, a full analysis of the PES is necessary and sample simulations for extreme conditions are required before the important reaction pathways

can be determined. Furthermore, rate coefficient predictions for pressure dependent reactions is a time-consuming and tedious task. Rice-Ramsperger-Kassel-Marcus (RRKM) master equation methods can be used to determine the pressure-dependent rate coefficient, but this requires user-intensive and computationally expensive *ab initio* calculations on the reactants, transition state, and products, and an analysis of the contributions of the densities of states.

Rate coefficients resulting from chemically activated reactions are usually complicated functions of temperature and pressure, and can often not be represented by modified Arrhenius expressions (Eq. 1.15). Relatively simple fall-off reactions can be represented by a high-pressure expression and a low-pressure expression, interpolated either by a parameter-free Lindemann-Hinshelwood form, or an expression with parameters to describe the curve between them such as the Troe form.^{218,219}

In single well systems, the commonly used Troe or SRI formalisms²²⁰ work very well, but some apparent reaction rate coefficients cannot be described accurately with these methods, even if additional parameters are introduced. A purely mathematical approach to handle these ‘problem cases’ is the use of Chebyshev polynomials. Alternative ways to represent more complex $k(T,p)$ dependencies are the sum of two or more modified Arrhenius expressions to capture a complex temperature-dependency, a set of modified Arrhenius expressions at different pressures that are interpolated in $\log(p)$ space, a combination of the previous two approaches (*i.e.* the sum of two or more sets of modified Arrhenius expressions, each at several pressures), or a set of Chebyshev polynomials in (T,p) .²²¹

An advantage – due to explicit definitions of the temperature and pressure limits – of Chebyshev polynomials is to prohibit rate coefficients from being used outside the valid range and thus to prevent unjustified extrapolative predictions. The time spent evaluating these complex expressions, especially while performing sensitivity analyses or fluid dynamics simulations, can be a considerable problem and is a further barrier to their widespread use. Although other forms have been proposed^{222,223}, they tend not to see widespread adoption until they are incorporated into downstream simulation software such as Chemkin²²⁴ or Computational Fluid Dynamics codes.

Matheu et al.²²⁵ developed an algorithm for the exploration and screening of pressure-dependent mechanisms. Only the important pressure-dependent reactions are added to the mechanism for the given temperature and pressure. The prediction of the pressure-dependent rate coefficient via on-the-fly calculations was presented in a subsequent publication.²²⁶ This was achieved by using the high-pressure limit rate coefficient calculated by the calculation methods available in the mechanism generation code alongside group additive thermochemistry of the reactants. An improved Quantum-Rice-Ramsperger-Kassel/Modified Strong Collision (QRRK/MSK) code is used to calculate the rate coefficient as a function of pressure.²²⁷ An estimated error of a factor of 10 is expected, given that the high-pressure limit may have an uncertainty of up to a factor of 10, which is a reasonable accuracy for a rate coefficient.

Green et al.¹⁹ developed a method to assess for which reactions pressure-dependence should be calculated by comparing the high-pressure limit pre-exponential factor to the hard-sphere collision rate constant and the concentration of the bath gas and by controlling whether the temperature is high enough to merit the use of pressure dependent calculations.

1.2.2.2.8 Challenges faced with rate coefficient calculations

One of the biggest challenges facing automatic mechanism generation is the lack of kinetic and to a lesser extent thermodynamic data on which to base the models. The methods used to calculate the thousands up to millions of reaction rate coefficients generally use known reaction rate coefficients from reactions similar to the ones being calculated (i.e. with resembling electronic changes) or data derived from these coefficients. If such data are not available, then analogies are made with less similar reactions, i.e. a broader generalization is made. When the data are arranged in a tree structure with the most general nodes at the top and each level down the tree getting more specific, the calculation method can repeatedly be based on a parent node until data are found. The principle here is that “uncertain rate coefficients are better than rate coefficients”, as the alternative (pretending that the rate coefficient is zero) is usually worse.

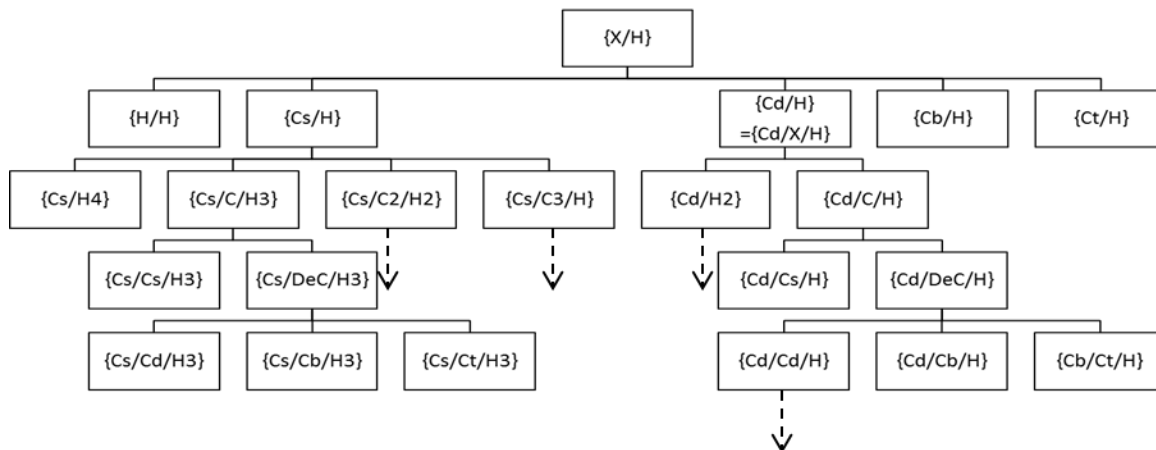


Figure 13: Example of tree structure to assign reaction rate coefficients to hydrogen abstraction reactions. Structure taken from Sumathy et al.²⁰³.

Problems can occur when an inappropriate generalization is made. For example, assuming the database does not contain an entry for the rate coefficient for the insertion reaction of CO_2 into a C–C bond, it may generalize to CO_2 insertion into any single bond R–R, and then use the rate expression for insertion into a C–H bond, which is much too fast.²²⁸ A challenge of this approach is therefore structuring the tree so that chemically similar functional groups are close to each other, resulting in a sensible analogy if the desired data are not found. Another challenge is populating the tree with sufficient data distributed across a wide chemical space. When models are found to depend critically on rate coefficients coming from broad generalizations, the user should seek to fill in the missing areas of the database and build the model again. Such data either come from searching the literature, performing experiments, or performing quantum chemistry and transition state theory calculations of a few characteristic reactions. Once these data are in the database they can be used to improve all future models.

Another important challenge when determining reaction rate coefficients is the existence of barrierless reactions. For these reactions, the reaction rate can decrease as a function of temperature. Regressing Arrhenius or modified Arrhenius parameters can lead to negative values for the activation energy. The term “activation energy” has to be used carefully in this context because the energy of the transition state is not higher than the energy of the reactants (if the reaction is considered in the exothermic direction) since the reaction is barrierless. Evans-Polanyi, Blower-Masel, group additivity, ... are not applied to barrierless reactions, since the activation energies are generally constant with values around zero – independent of the enthalpy

of reaction. Semi-empirical methods and group additivity models introduced above cannot be used since most reaction families with barrierless reactions, such as radical recombination reactions, do not follow these trends. Quantum mechanical methods can lead to rather accurate reaction rate coefficients, but the level of expertise required to obtain reliable results is at present too high to allow inclusion of these reaction families in on-the-fly kinetic model generation. Furthermore, because expert-user involvement is needed and only a limited number of researchers are working on this topic, data is lacking for many of these reactions. It is clear that still a lot of work needs to be done before barrierless reactions can be considered in on-the-fly automatic kinetic model generation codes. A fixed pre-exponential factor is often employed and the activation energy is equated with reaction enthalpy in endothermic direction, again with the philosophy that any rate coefficient is better than no rate coefficient at all.

One category of database entries for rate coefficient calculations that needs particular attention is those involving aromatic groups. Atoms and bonds that belong to aromatic moieties inside molecules feature a reactivity that is distinctly different from non-aromatic species²²⁹ due to the resonance stabilization of the aromatic ring, and hence require specific database entries. Unfortunately only a limited number of elementary rate coefficients are available for reactions involving aromatic species. For example, the pathways to aromatics such as indene, naphthalene, benzene, styrene and other polyaromatic hydrocarbons starting from five-membered rings such as cyclopentadiene and cyclopentadienyl are not completely understood on an elementary basis.^{230,231} An often chosen pragmatic solution that addresses the lack of fundamental understanding of these systems is the introduction of lumped reactions that model the underlying elementary aromatic chemistry, as in the work of Cavallotti and Polino.²³² A consequence of this particular example of data scarcity is the limited capability of current mechanism generators to model aromatic chemistry. In addition to the issue of data scarcity and a lack of fundamental understanding, the algorithmic identification of aromatic atoms and bonds inside molecules during automated kinetic model generation is also non-trivial, not in the least due to the absence of a straightforward definition of aromaticity.²³³

Group contributions for thermochemical properties of species cannot account for ring strain, since this is based on a structural arrangement. Therefore extra corrections are required. This works for simple structures, for example when a 5- or 6-membered ring is present, but as

complexity increases, e.g. when the ring contains heteroatoms (other than carbon), or when there are polycyclic species with fused rings, the number of ad-hoc ring corrections required becomes a major challenge. To overcome this problem, Magoon and coworkers added a feature to RMG to perform on-the-fly semiempirical quantum chemistry calculations (as described in section 1.2.2.1), allowing them to model TCD (exo-Tricyclo[5.2.1.0^{2,6}] decane), the main component of JP-10 jet fuel²³⁴, which contains fused cyclopentadiene rings.

Relatively accurate thermochemistry of ring-containing species is a huge help, since many reaction rate coefficients depend largely on the reaction enthalpy, but rate coefficients can have their own problems with rings: in some reaction families the transition state may have ring structures unlike either the reactant or product. Current efforts to perform on-the-fly calculations of transition states may help in these cases.^{235,236}

The final challenge regarding rings is once again data scarcity: there are currently few rate rules available, even for simple species like cyclopentane. Numerous studies have been carried out on cyclohexane, methylcyclohexane and decalin pyrolysis and combustion.²³⁷⁻²³⁹ However, very little attention is paid to the chemistry of 5-membered rings, although the ring strain is very different. The difficulty to experimentally differentiate and accurately identify components containing five- and six-membered rings with standard techniques such as gas-chromatography and mass spectrometry further complicate kinetic modeling efforts involving naphthenic components.^{231,240,241}

1.3 Genesys

Genesys is a kinetic model generation code recently developed at the Laboratory for Chemical Technology (LCT) at Ghent University²⁴. The main features of the code as it exists today are highlighted in this section. As mentioned above, many kinetic model generators are built with an application domain in mind. The species representation, the definition of reaction rules and the accompanying databases are tailored to these specific applications, hampering the use of the code outside of it. To overcome this limitation, Genesys has been developed by integrating open-source chemo-informatics codes. These codes allow a general representation of species, independent of the elements and reaction families. For Genesys, the chemo-informatics library called the “Chemistry Development Kit” (CDK)^{242,243}, developed at the University of Notre Dame, is used as shown in Figure 14. Besides supporting nearly all the chemical elements and providing a robust molecular representation, CDK provides different algorithms such as the conversion of species identifiers such as InChI⁸¹ and SMILES^{80,86,91} to the internal representation, graph isomorphism and automorphism and property identification algorithms such as aromaticity detection, resonance detection, subgraph recognition, among others by using SMARTS⁹¹, etc.

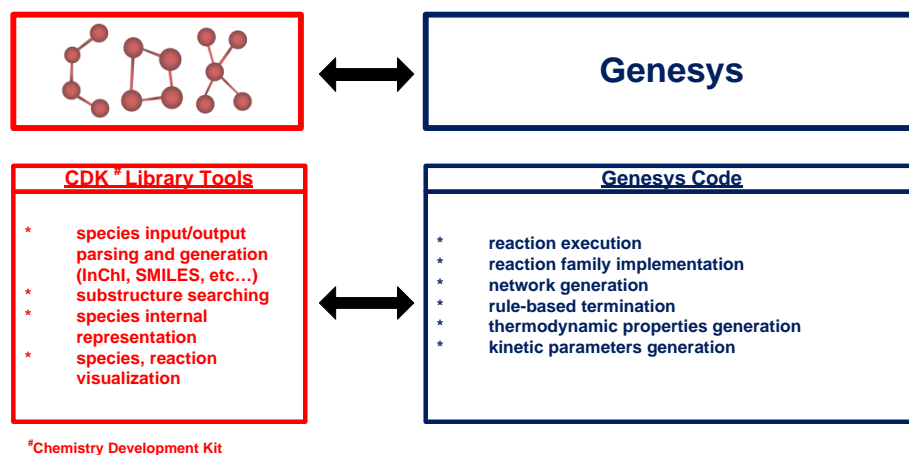


Figure 14: Interaction between Genesys and CDK

Genesys is written in the programming language JAVA, as is CDK, allowing an easy integration of both codes and is distributed via a closed repository on <https://github.ugent.be/LCT/Genesys>.

Genesys solely generates elementary reactions, i.e. reactions characterized by a single transition state, and the generation of these reactions is terminated using either the rule-based method, the rate-based method or a combination of both. Figure 15 illustrates how a kinetic model is

generated. The user input, i.e. the reactants and the reactions families, is read and used to iteratively build the reaction mechanism. This is followed by the assignment of the necessary data to species and reactions. When Genesys employs the rule-based termination criterion, the assignment of data can be done after the network generation, and the final reaction network is independent of these data. For the rate-based algorithm, after each iteration step, the thermodynamic and kinetic data are calculated and a reactor simulation is performed to assess which species and reactions should be added to the model.

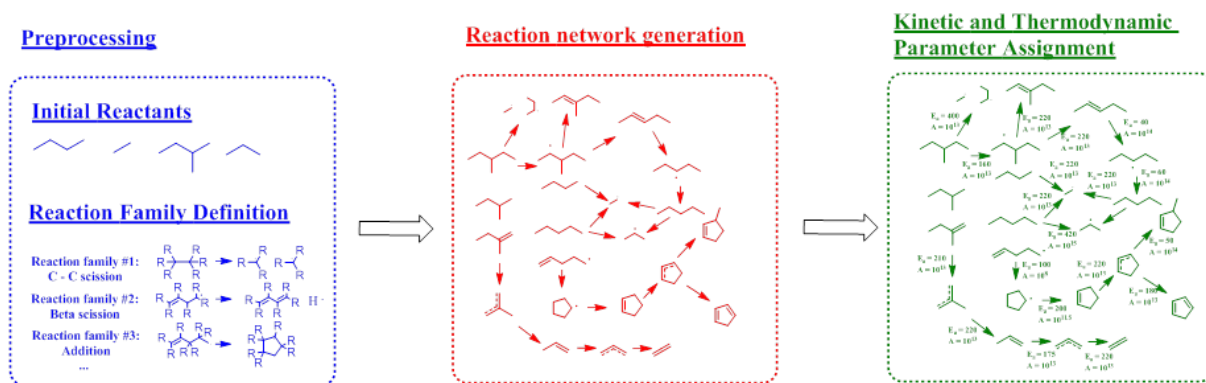


Figure 15: Illustration of the generation of a kinetic model using Genesys.

The four following sections elaborate on (1) user input to Genesys, (2) the kinetic model generation algorithm, (3) post-processing, and (4) Genesys features.

1.3.1 Genesys input

The input needed to generate a kinetic model consists of a set of input species and a set of reaction families. Input species, which are often limited to the reactants to a chemical process, can be declared in InChI or SMILES format. Genesys reads these chemical identifiers and CDK translates them to the internal graph representation of species. A significant difference of Genesys compared to most other kinetic model generation codes is the externalization of reaction families, allowing the user to input a set of reaction families appropriate to his or her chemical process, without having knowledge nor access to the source code of Genesys. As consequence the user is responsible to build and select the set or reaction families needed for the process. The reaction families are written in the “extensible markup language” XML. A reaction family contains several elements. Each family is assigned a name. Next is defined whether a reaction family is unimolecular or bimolecular. This is followed by the reaction recipe. This recipe contains information of the breaking and forming of bonds, the increase and decrease of bond

orders and the gain or loss of charges and single electrons. Each of these elements is defined on the level of the reactive atoms. Which atoms can undergo this recipe is defined in the next section. Each reactive center of each reactant needs a separate definition in which the entire reactive centers is defined using the SMARTS language. Besides this definition, each reactive atom separately is also defined using a SMARTS string. The definition of each reactant is concluded by adding a list of constraints the reactant needs to comply with to allow reaction through this reaction family. These constraints can be defined on the atomic level, e.g. hybridization, elements, number of neighbors, etc. or on the entire molecule, such as molecule size, presence of rings, etc. The final part of the reaction family definition contains the procedure to calculate the rate coefficients of the reactions belonging to this reaction family. Several procedures are enabled in Genesys, such as Evans-Polanyi¹⁷², Blowers-Masel^{173,176}, group additivity^{185,198,244}, etc. An example of a reaction family definition is given in Figure 16.

Node	Content
xml	version="1.0" encoding="UTF-8"
config	
inp-temperature	1000
inp-reaction-family	
name	homolytic substitution reaction of a hydrogen atom on a sulfur atom - vandeputte H8 Ph.D Thesis (unpublished)
bimolecular	true
inp-recipe	
inp-transformation	
type	FORM_BOND
centers	A,B
inp-transformation	
inp-transformation	
inp-transformation	
inp-reactant	
inp-reactant	
value	2
smarts	[C,v4,v3]-[S;X2v2]-*
inp-reactive-center	
value	B
smarts	!\$([S;X2v2](-*)-[C,v4,v3])
inp-reactive-center	
inp-reactive-center	
inp-molecule-constraint	
type	SINGLEELECTRONCOUNT
limit	MAX
value	0
inp-molecule-constraint	
inp-molecule-constraint	
inp-kinetics	
type	GROUP_ADDITIVITY
path	homolytic substitution reaction of a hydrogen atom on a sulfur atom - vandeputte H8 Ph.D Thesis (unpublished).xml
inp-reaction-family	
inp-reaction-family	
inp-reaction-family	
inp-reaction-family	
inp-reaction-family	

Figure 16: Screenshot of the input for a reaction family in Genesys read by the “Rinzo XML Editor”. The reaction family is the homolytic substitution of a hydrogen atom on a sulfur atom, which is part of the reaction families which describe the pyrolysis of alkyl sulfides.

1.3.2 Kinetic model generation algorithm

Once the input species and reaction families are defined, the construction of the model can start, including the generation of new product species, the determination of the reactions between the species and the assignment of kinetic and thermodynamic parameters. The generation commences with a pool of initial species. These species are added to a set of species called the “source”, which contains all species that have not been subjected to the reaction families yet. In each iteration step, one species is taken out of the “source” and moved to the set of species called the “core” species. For this species, all the possible reactions are generated, according to the user-defined constraints per reaction family. For monomolecular reaction families, only that species is used. Bimolecular reactions are generated by using the species at hand with each species from the “core” species set. The reactions lead to the formation of species. If these product species are new, i.e. they are not present in the “source” nor in the “core” species set, they are added to the “source”. The resulting reactions are added to the model, if they are new. This procedure is continued until the “source” is empty. This scheme for network generation is shown in Figure 17.

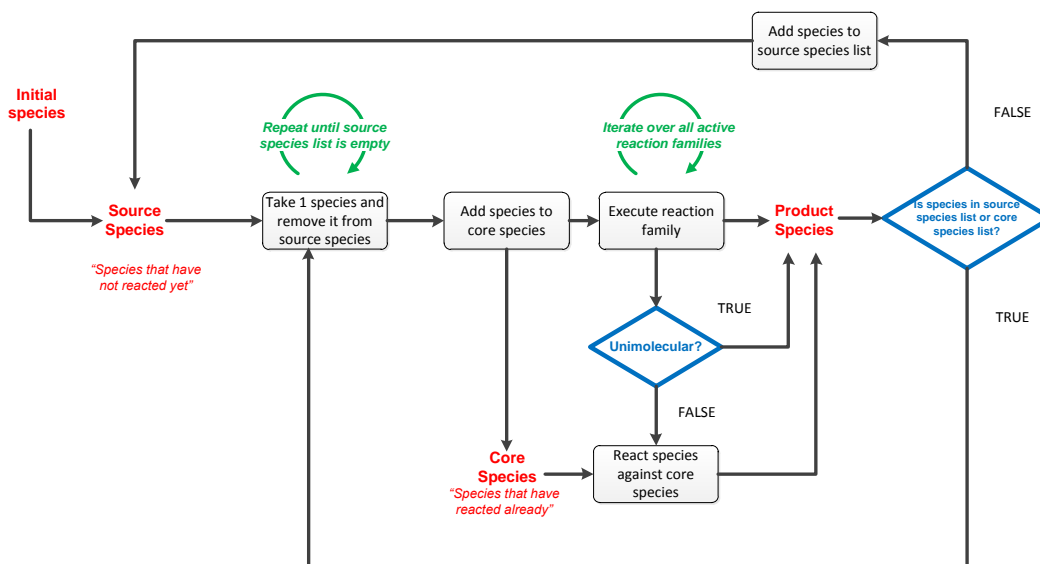


Figure 17: Network generation scheme of Genesys.

In order to distinguish whether or not a reaction has already been encountered in the network, a method to compare chemical structures needs to be developed. Within Genesys, this method comprises two steps. First, both the reactants and the products InChI strings are generated, this enables the comparison of simple string objects. Each new reaction is then compared to the reactions in the network, both in the forward and the reverse direction, by the use of those InChI

strings. When a reaction is discovered with the same reactants and products, an isomorphism check of the transition state is performed. This way, reactions with the same reactants and products but a different transition state can be added to the network. This is for example necessary to distinguish between the two intramolecular hydrogen abstractions from the 1-pentyl to the 2-pentyl radical, of which the transition states can possess a 3 membered ring structure or a 5 membered ring structure.

Once the network generation is done, i.e. once the “source” is empty, thermodynamic and kinetic parameters need to be calculated and assigned to the different reactions and species. For species, the search for appropriate thermodynamic parameters is two-fold. First, a library containing thermodynamic data for a large amount of species is visited. If one of the species in the network is present in this library, the thermodynamic data is pulled from the library and used as such. Second, if a species from the network is not present in the first library, its thermodynamic properties are calculated using group additivity. All group additivity values originate from CBS-QB3 calculations.^{8,120,121,161,245}

The procedure to calculate reaction rate coefficients is supplied by the user as part of the definition of reaction families. Genesys has a database containing many group additivity values calculated by Marin and coworkers^{120,181,185,186,196-198,200}. The procedure to calculate rate coefficients from group additivity is shown in Figure 18. If the kinetic parameters are assigned by use of group additivity models, an external library is needed, which contains the values for a given reaction family. Those values are the single-event pre-exponential factor and activation energy of the reference reaction, plus the different group additivity values. For each surrounding of each reactive atom, corresponding group additivity values are required. In order to connect specific sub-molecular patterns to those values, SMARTS string are used. Group additivity only calculates single-event pre-exponential factors, which is multiplied by the number of single events and optionally a tunneling coefficient to obtain the final pre-exponential factor, which entails the need for accurate symmetry calculations. Empirical tunneling models are typically used, which depend on the temperature and the activation energy in the exothermic direction.

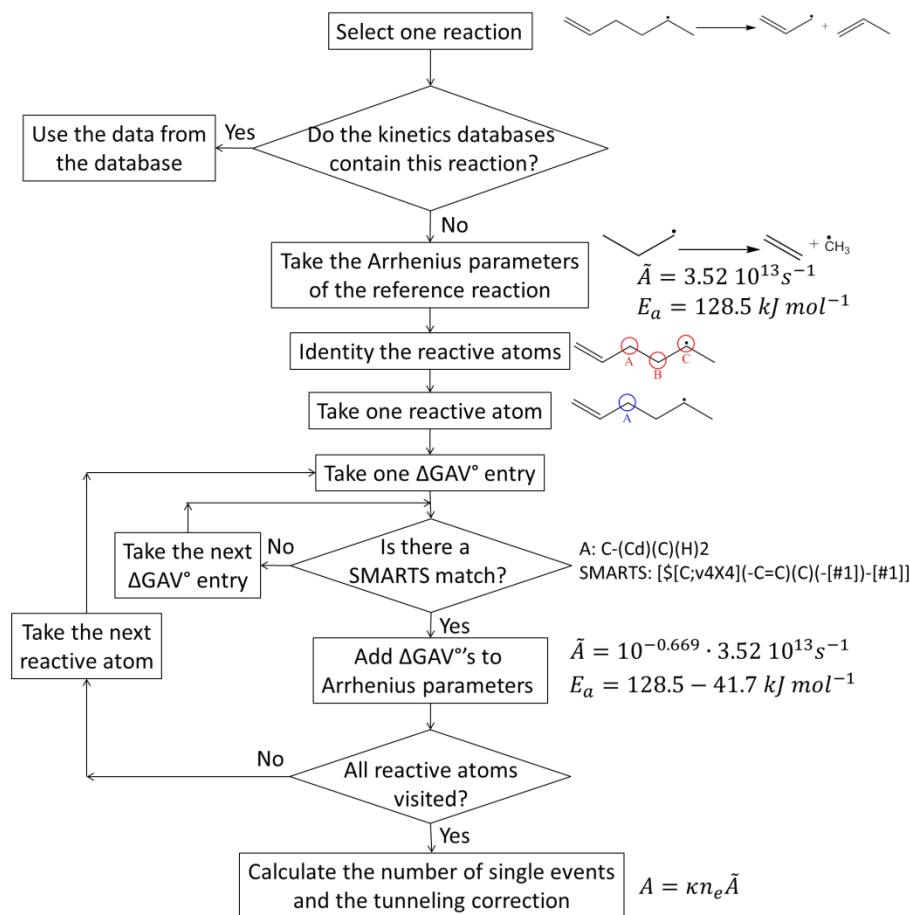


Figure 18: Calculation scheme for kinetics using group additivity. The example is the β -scission of the 1-hexen-5-yl radical to the allyl radical and propene.

The externalization of the libraries with kinetic group additive values has several advantages, one being the possible use of the values in two reaction families with different constraints without the need of storing the same values in each reaction family. Furthermore, this externalization facilitates simple and straightforward extensions and updates of the libraries, which are easily decipherable and well structured.

1.3.3 Post-processing

A final step of the generation of the kinetic model holds the post-processing, this results in human readable information of the model and input files that allow the use of the generated model in third-party tools.

Since Chemkin²²⁴ is widely used software for kinetic modelling, and their input formats have been standardized and are generally accepted by the kinetic modelling community, Genesys automatically writes the final kinetic model in Chemkin format. Furthermore, for each species

and each reaction in the model, the source of the thermodynamic and kinetic data respectively, is summarized, allowing the user to verify where all the data comes from. Visualization of all the species in a model is also enabled.

1.3.4 Genesys features

Genesys has several features facilitating a fully automatic kinetic model generation with, besides the building of proper input, minimal user interaction.

First, both to calculate the total entropy of molecules, as well as to calculate the number of single events of reactions, the total rotational symmetry number of molecules and transition states is necessary. This number is automatically calculated using the automorphism group order of the corresponding graph, corrected for the presence of so called “label stereocenters”, which are stereocenters that appear in a species by labelling indistinguishable atoms making them distinguishable. More information on the algorithms to obtain the symmetry numbers can be found elsewhere.¹⁹¹

Secondly, Genesys has been developed to explicitly account for stereoisomers.¹²⁵ This entails algorithms to detect stereocenters and generate all possible stereoisomers of a species or transition state, algorithms to construct and destruct stereocenters through the course of chemical reactions, algorithms to assign stereospecific and stereoselective rate coefficients and algorithms to assign separate thermodynamic parameters to stereoisomers. The latter two sets of algorithms are also driven by databases build to distinguish stereoisomers.

Finally, Genesys also allows to automatically merge the final kinetic model with existing kinetic models, if they are provided in Chemkin format. This is useful when the small molecule chemistry cannot be described by the reaction families nor by the available databases, or when considerable work on similar molecules has already been done, and the current study is an extension of existing kinetic models.

1.4 Outline

The search for and the development of reliable kinetic models is hampered by large data scarcity. The data gap needs to be accounted for, which cannot be done by experimental studies alone. They need to be complemented with high level quantum chemical calculations, which although they are time intensive, they are in most cases still faster than experimental work. However, current state-of-the-art *ab initio* calculations require user knowledge and expertise, slowing down the use of these data in industry. This work takes a first step towards fully automated *ab initio* calculations in the framework of kinetic model generation.

The state-of-the-art searches for reaction on a potential energy surface is examined in **Chapter 2**, where the pyrolysis of 1-pentanol is studied. One of the most important radicals in this process is the 1-pentanol radical formed after hydrogen abstractions from 1-pentanol. The potential energy surface of this radical is studied using the KinBot²¹⁵ software. By means of Master Equation calculations, pressure- and temperature-dependent rate coefficients are obtained, which are subsequently used to build a 1-pentanol pyrolysis model using Genesys. The final model is compared to experimental data and a good agreement is found.

The kinetic model generator Genesys has then been extended to allow on-the-fly *ab initio* calculations, both to obtain thermodynamic as well as kinetic data. **Chapter 3** elaborates on the methodology used to minimize human interaction and to ensure a reliable, robust and widely applicable code. Calculated data on thermodynamics of molecules and radicals, and on reaction rate coefficients are promising. The comparison to literature data is satisfactory.

Chapter 4 uses the on-the-fly *ab initio* calculation methodology of Chapter 3 to build a new group additivity model for intramolecular hydrogen abstraction reactions. No such model has been reported in literature so far. Many reaction rate coefficients have been calculated automatically and are very similar to experimental and theoretic literature results. Their regression into a group additivity model is validated; independently calculated rate coefficients are well within the uncertainty boundaries of the model.

The use and necessity of a group additive model for intramolecular hydrogen abstractions is illustrated by building a kinetic model for n-heptane pyrolysis, reported in **Chapter 5**. Genesys has been used as kinetic model generator, including the new group additive model. Several sets of

experimental data have been simulated showing a good accuracy of the model. Rate of production analysis shows the most important pathways and can highlight where intramolecular hydrogen abstractions are important.

This thesis is concluded in **Chapter 6** where the main findings are summarized, and an outlook is given for the near future.

1.5 References

1. Crutzen PJ, Arnold F. Nitric acid cloud formation in the cold Antarctic stratosphere: a major cause for the springtime 'ozone hole'. *Nature*. 1986;324(6098):651-655.
2. Dente M, Ranzi E, Goossens AG. Detailed prediction of olefin yields from hydrocarbon pyrolysis through a fundamental simulation model (SPYRO). *Computers & Chemical Engineering*. 1979;3(1-4):61-75.
3. van Goethem MWM, Kleinendorst FI, van Leeuwen C, van Velzen N. Equation-based SPYRO® model and solver for the simulation of the steam cracking process. *Computers & Chemical Engineering*. 2001;25(4-6):905-911.
4. Van Geem KM, Reyniers M-F, Marin GB. Challenges of modeling steam cracking of heavy feedstocks. *Oil & Gas Science and Technology-Revue D Ifp Energies Nouvelles*. 2008;63(1):79-94.
5. Dijkmans T, Pyl SP, Reyniers M-F, Abhari R, Van Geem KM, Marin GB. Production of bio-ethene and propene: alternatives for bulk chemicals and polymers. *Green Chemistry*. 2013;15(11):3064-3076.
6. Klein MT, Hou G, Bertolacini R, Broadbelt LJ, Kumar A. *Molecular Modeling in Heavy Hydrocarbon Conversions*: Taylor & Francis; 2005.
7. Incavo JA. A Detailed Quantitative Study of 1,2-Dichloroethane Cracking to Vinyl Chloride by a Gas Chromatographic Pyrolysis Device. *Industrial & Engineering Chemistry Research*. 1996;35(3):931-937.
8. Sabbe MK, Saeys M, Reyniers M-F, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *Journal of Physical Chemistry A*. 2005;109(33):7466-7480.
9. He K, Ierapetritou MG, Androulakis IP. Integration of on-the-fly kinetic reduction with multidimensional CFD. *Aiche Journal*. 2010;56(5):1305-1314.
10. Lu TF, Law CK. Toward accommodating realistic fuel chemistry in large-scale computations. *Progress in Energy and Combustion Science*. 2009;35(2):192-215.
11. Zheng XL, Lu TF, Law CK, Westbrook CK, Curran HJ. Experimental and computational study of nonpremixed ignition of dimethyl ether in counterflow. *Proceedings of the Combustion Institute*. 2005;30(1):1101-1109.
12. Zheng XL, Lu TF, Law CK. Experimental counterflow ignition temperatures and reaction mechanisms of 1,3-butadiene. *Proceedings of the Combustion Institute*. 2007;31:367-375.
13. Cuoci A, Frassoldati A, Faravelli T, Ranzi E. A computational tool for the detailed kinetic modeling of laminar flames: Application to C₂H₄/CH₄ coflow flames. *Combustion and Flame*. 2013;160(5):870-886.
14. Frassoldati A, Cuoci A, Faravelli T, Niemann U, Ranzi E, Seiser R, Seshadri K. An experimental and kinetic modeling study of n-propanol and iso-propanol combustion. *Combustion and Flame*. 2010;157(1):2-16.
15. Merchant SS, Zanoelo EF, Speth RL, Harper MR, Van Geem KM, Green WH. Combustion and pyrolysis of iso-butanol: Experimental and chemical kinetic modeling study. *Combustion and Flame*. 2013;160(10):1907-1929.
16. Herbinet O, Biet J, Hakka MH, Warth V, Glaude P-A, Nicolle A, Battin-Leclerc F. Modeling study of the low-temperature oxidation of large methyl esters from C11 to C19. *Proceedings of the Combustion Institute*. 2011;33(1):391-398.

17. Broadbelt LJ, Stark SM, Klein MT. Computer-Generated Pyrolysis Modeling - on-the-Fly Generation of Species, Reactions, and Rates. *Industrial & Engineering Chemistry Research*. 1994;33(4):790-799.
18. *RMG - Reaction Mechanism Generator v4.0* [computer program] 2013.
19. Green WH, Barton PI, Bhattacharjee B, Matheu DM, Schwer DA, Song J, Sumathi R, Carstensen HH, Dean AM, Grenda JM. Computer Construction of Detailed Chemical Kinetic Models for Gas-Phase Reactors. *Industrial & Engineering Chemistry Research*. 2001;40(23):5362-5370.
20. Song J. *Building Robust Chemical Reaction Mechanisms: Next Generation of Automatic Model Construction Software*: Chemical engineering, Massachusetts Institute of Technology; 2004, Ph. D. Thesis.
21. Susnow RG, Dean AM, Green WH, Peczak P, Broadbelt LJ. Rate-based construction of kinetic models for complex systems. *Journal of Physical Chemistry A*. 1997;101(20):3731-3740.
22. Warth V, Battin-Leclerc F, Fournet R, Glaude PA, Come GM, Scacchi G. Computer based generation of reaction mechanisms for gas-phase oxidation. *Computers & Chemistry*. 2000;24(5):541-560.
23. Battin-Leclerc F, Glaude PA, Warth V, Fournet R, Scacchi G, Côme GM. Computer tools for modelling the chemical phenomena related to combustion. *Chemical Engineering Science*. 2000;55(15):2883-2893.
24. Vandewiele NM, Van Geem KM, Reyniers M-F, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chemical Engineering Journal*. 2012;207:526-538.
25. Clymans PJ, Froment GF. Computer-Generation of Reaction Paths and Rate-Equations in the Thermal-Cracking of Normal and Branched Paraffins. *Computers & Chemical Engineering*. 1984;8(2):137-142.
26. Hillewaert LP, Dierickx JL, Froment GF. Computer generation of reaction schemes and rate equations for thermal cracking. *Aiche Journal*. 1988;34(1):17-24.
27. Rangarajan S, Bhan A, Daoutidis P. Language-oriented rule-based reaction network generation and analysis: Description of RING. *Computers & Chemical Engineering*. 2012;45:114-123.
28. Rangarajan S, Bhan A, Daoutidis P. Language-oriented rule-based reaction network generation and analysis: Applications of RING. *Computers & Chemical Engineering*. 2012;46:141-152.
29. Ranzi E, Dente M, Goldaniga A, Bozzano G, Faravelli T. Lumping procedures in detailed kinetic modeling of gasification, pyrolysis, partial oxidation and combustion of hydrocarbon mixtures. *Progress in Energy and Combustion Science*. 2001;27(1):99-139.
30. Ranzi E, Sogaro A, Gaffuri P, Pennati G, Westbrook CK, Pitz WJ. A new comprehensive reaction mechanism for combustion of hydrocarbon fuels. *Combustion and Flame*. 1994;99(2):201-211.
31. Blurock ES. Reaction: System for Modeling Chemical Reactions. *Journal of Chemical Information and Computer Sciences*. 1995;35(3):607-616.
32. Blurock ES. Detailed Mechanism Generation. 1. Generalized Reactive Properties as Reaction Class Substructures. *Journal of Chemical Information and Computer Sciences*. 2004;44(4):1336-1347.
33. Blurock ES. Detailed Mechanism Generation. 2. Aldehydes, Ketones, and Olefins. *Journal of Chemical Information and Computer Sciences*. 2004;44(4):1348-1357.

34. Ratkiewicz A, Truong TN. Application of Chemical Graph Theory for Automated Mechanism Generation. *Journal of Chemical Information and Computer Sciences*. 2003;43(1):36-44.
35. Németh A, Vidóczy T, Héberger K, Kúti Z, Wágner J. MECHGEN: Computer Aided Generation and Reduction of Reaction Mechanisms. *Journal of Chemical Information and Computer Sciences*. 2002;42(2):208-214.
36. Di Maio FP, Lignola PG. KING, a KInetic Network Generator. *Chemical Engineering Science*. 1992;47(9-11):2713-2718.
37. Fontain E, Bauer J, Ugi I. Computer Assisted Bilateral Generation of Reaction Networks from Educts and Products. *Chemistry Letters*. 1987;16(1):37-40.
38. Fontain E, Reitsam K. The generation of reaction networks with RAIN. 1. The reaction generator. *Journal of Chemical Information and Computer Sciences*. 1991;31(1):96-101.
39. Porollo AA, Lushnikov DE, Pivina TS, Ivshin VP. Computer representation and generation of possible pathways for thermal decomposition reactions of organic compounds. *Journal of Molecular Structure: THEOCHEM*. 1997;391(1-2):117-124.
40. Yoneda Y. Chemogram, a Computer-Program Package for Chemical Logic .3. Estimation of the Thermodynamic Properties of Organic-Compounds in the Ideal-Gas State .1. Acyclic Compounds and Cyclic-Compounds with a Ring of Cyclopentane, Cyclohexane, Benzene, or Naphthalene. *Bulletin of the Chemical Society of Japan*. 1979;52(5):1297-1314.
41. 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 - GRACE(I). *Bulletin of the Chemical Society of Japan*. 1979;52(1):8-14.
42. *KUCRS software library* [computer program] 2012.
43. Karaba A, Zamostny P, Lederer J, Belohlav Z. Generalized Model of Hydrocarbons Pyrolysis Using Automated Reactions Network Generation. *Industrial & Engineering Chemistry Research*. 2013;52(44):15407-15416.
44. Chinnick SJ, Baulch DL, Ayscough PB. An expert system for hydrocarbon pyrolysis reactions. *Chemometrics and Intelligent Laboratory Systems*. 1988;5(1):39-52.
45. Prickett SE, Mavrovouniotis ML. Construction of complex reaction systems—I. Reaction description language. *Computers & Chemical Engineering*. 1997;21(11):1219-1235.
46. Prickett SE, Mavrovouniotis ML. Construction of complex reaction systems—II. Molecule manipulation and reaction application algorithms. *Computers & Chemical Engineering*. 1997;21(11):1237-1254.
47. Khan SS, Zhang Q, Broadbelt LJ. Automated mechanism generation. Part 1: mechanism development and rate constant estimation for VOC chemistry in the atmosphere. *Journal of Atmospheric Chemistry*. 2009;63(2):125-156.
48. Chevalier C, Warnatz J, Melenk H. Automatic Generation of Reaction Mechanisms for Description of Oxidation of Higher Hydrocarbons. *Berichte der Bunsengesellschaft für physikalische Chemie*. 1990;94(11):1362-1367.
49. Green WH. Predictive Kinetics: A New Approach for the 21st Century. In: Guy BM, ed. *Advances in Chemical Engineering*. Vol Volume 32: Academic Press; 2007:1-313.
50. Ratkiewicz A, Truong TN. Automated mechanism generation: From symbolic calculation to complex chemistry. *International Journal of Quantum Chemistry*. 2006;106(1):244-255.

51. Broadbelt LJ, Pfaendtner J. Lexicography of kinetic modeling of complex reaction networks. *Aiche Journal*. 2005;51(8):2112-2121.
52. Vinu R, Broadbelt LJ. Unraveling Reaction Pathways and Specifying Reaction Kinetics for Complex Systems. *Annual Review of Chemical and Biomolecular Engineering, Vol 3* 2012:29-54.
53. Zádor J, Taatjes CA, Fernandes RX. Kinetics of elementary reactions in low-temperature autoignition chemistry. *Progress in Energy and Combustion Science*. 2011;37(4):371-421.
54. Battin-Leclerc F, Blurock E, Bounaceur R, Fournet R, Glaude PA, Herbinet O, Sirjean B, Warth V. Towards cleaner combustion engines through groundbreaking detailed chemical kinetic models. *Chemical Society Reviews*. 2011;40(9):4762-4782.
55. Simmie JM. Detailed chemical kinetic models for the combustion of hydrocarbon fuels. *Progress in Energy and Combustion Science*. 2003;29(6):599-634.
56. Battin-Leclerc F, Simmie JM, Blurock E. *Cleaner Combustion: Developing Detailed Chemical Kinetic Models*: Springer; 2013.
57. Ranzi E, Faravelli T, Gaffuri P, Garavaglia E, Goldaniga A. Primary Pyrolysis and Oxidation Reactions of Linear and Branched Alkanes. *Industrial & Engineering Chemistry Research*. 1997;36(8):3336-3344.
58. Warth V, Stef N, Glaude PA, Battin-Leclerc F, Scacchi G, Côme GM. Computer-Aided Derivation of Gas-Phase Oxidation Mechanisms: Application to the Modeling of the Oxidation of n-Butane. *Combustion and Flame*. 1998;114(1-2):81-102.
59. Hakka MH, Bennadji H, Biet J, Yahyaoui M, Sirjean B, Warth V, Coniglio L, Herbinet O, Glaude PA, Billaud F, Battin-Leclerc F. Oxidation of methyl and ethyl butanoates. *International Journal of Chemical Kinetics*. 2010;42(4):226-252.
60. Grenda JM, Androulakis IP, Dean AM, Green WH. Application of Computational Kinetic Mechanism Generation to Model the Autocatalytic Pyrolysis of Methane. *Industrial & Engineering Chemistry Research*. 2003;42(5):1000-1010.
61. Van Geem KM, Reyniers M-F, Marin GB, Song J, Green WH, Matheu DM. Automatic reaction network generation using RMG for steam cracking of n-hexane. *Aiche Journal*. 2006;52(2):718-730.
62. Harper MR, Van Geem KM, Pyl SP, Marin GB, Green WH. Comprehensive reaction mechanism for n-butanol pyrolysis and combustion. *Combustion and Flame*. 2011;158(1):16-41.
63. Moréac G, Blurock ES, Mauss F. AUTOMATIC GENERATION OF A DETAILED MECHANISM FOR THE OXIDATION OF n-DECANE. *Combustion Science and Technology*. 2006;178(10-11):2025-2038.
64. Mersin IE, Blurock ES, Soyhan HS, Konnov AA. Hexadecane mechanisms: Comparison of hand-generated and automatically generated with pathways. *Fuel*. 2014;115(0):132-144.
65. Glaude PA, Warth V, Fournet R, Battin-Leclerc F, Scacchi G, Côme GM. Modeling of the oxidation of n-octane and n-decane using an automatic generation of mechanisms. *International Journal of Chemical Kinetics*. 1998;30(12):949-959.
66. Buda F, Bounaceur R, Warth V, Glaude PA, Fournet R, Battin-Leclerc F. Progress toward a unified detailed kinetic model for the autoignition of alkanes from C4 to C10 between 600 and 1200 K. *Combustion and Flame*. 2005;142(1-2):170-186.

67. Prickett SE, Mavrovouniotis ML. Construction of complex reaction systems—III. An example: alkylation of olefins. *Computers & Chemical Engineering*. 1997;21(12):1325-1337.
68. Muharam Y, Warnatz J. Kinetic modelling of the oxidation of large aliphatic hydrocarbons using an automatic mechanism generation. *Physical Chemistry Chemical Physics*. 2007;9(31):4218-4229.
69. Khan SS, Broadbelt LJ. Automated mechanism generation. Part 2: application to atmospheric chemistry of alkanes and oxygenates. *Journal of Atmospheric Chemistry*. 2009;63(2):157-186.
70. Thybaut JW, Marin GB. Single-Event MicroKinetics: Catalyst design for complex reaction networks. *Journal of Catalysis*. 2013;308:352-362.
71. Singh S, Li S, Carrasquillo-Flores R, Alba-Rubio AC, Dumesic JA, Mavrikakis M. Formic acid decomposition on Au catalysts: DFT, microkinetic modeling, and reaction kinetics experiments. *Aiche Journal*. 2014;60(4):1303-1319.
72. Vandegehuchte BD, Choudhury IR, Thybaut JW, Martens JA, Marin GB. Integrated Stefan–Maxwell, Mean Field, and Single-Event Microkinetic Methodology for Simultaneous Diffusion and Reaction inside Microporous Materials. *The Journal of Physical Chemistry C*. 2014;118(38):22053-22068.
73. Vinu R, Levine SE, Wang L, Broadbelt LJ. Detailed mechanistic modeling of poly(styrene peroxide) pyrolysis using kinetic Monte Carlo simulation. *Chemical Engineering Science*. 2012;69(1):456-471.
74. Derboven P, D’hooge DR, Stamenovic MM, Espeel P, Marin GB, Du Prez FE, Reyniers M-F. Kinetic Modeling of Radical Thiol–Ene Chemistry for Macromolecular Design: Importance of Side Reactions and Diffusional Limitations. *Macromolecules*. 2013;46(5):1732-1742.
75. Hou Z, Bennett CA, Klein MT, Virk PS. Approaches and Software Tools for Modeling Lignin Pyrolysis. *Energy & Fuels*. 2010;24(1):58-67.
76. Aumont B, Szopa S, Madronich S. Modelling the evolution of organic carbon during its gas-phase tropospheric oxidation: development of an explicit model based on a self-generating approach. *Atmospheric Chemistry and Physics*. 2005;5(9):2497-2517.
77. Broadbelt LJ, Stark SM, Klein MT. Computer generated reaction modelling: Decomposition and encoding algorithms for determining species uniqueness. *Computers & Chemical Engineering*. 1996;20(2):113-129.
78. Quann RJ, Jaffe SB. Structure-Oriented Lumping - Describing the Chemistry of Complex Hydrocarbon Mixtures. *Industrial & Engineering Chemistry Research*. 1992;31(11):2483-2497.
79. Quann RJ, Jaffe SB. Building useful models of complex reaction systems in petroleum refining. *Chemical Engineering Science*. 1996;51(10):1615-1635.
80. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences*. 1988;28(1):31-36.
81. Heller SR, Stein SE, Tchekhovskoi DV. InChI: Open access/open source and the IUPAC international chemical identifier. *Abstracts of Papers of the American Chemical Society*. 2005;230:U1025-U1026.
82. Côme GM, Muller C, Cunin PY, Griffiths M. A linear chemical notation. *Computers & Chemistry*. 1984;8(4):233-237.

83. Muller C, Scacchi G, Côme GM. A compiler for a linear chemical notation. *Computers & Chemistry*. 1991;15(4):337-342.
84. Marin G, Yablonsky GS. *Kinetics of Chemical Reactions*: Wiley; 2011.
85. McKay BD. Practical Graph Isomorphism. *Congressus Numerantium*. 1981;30:45-87.
86. Weininger D, Weininger A, Weininger JL. SMILES. 2. Algorithm for generation of unique SMILES notation. *Journal of Chemical Information and Computer Sciences*. 1989;29(2):97-101.
87. Faulon J-L, Collins MJ, Carr RD. The Signature Molecular Descriptor. 4. Canonizing Molecules Using Extended Valence Sequences. *Journal of Chemical Information and Computer Sciences*. 2004;44(2):427-436.
88. Braun J, Gugisch R, Kerber A, Laue R, Meringer M, Rücker C. MOLGEN-CID A Canonizer for Molecules and Graphs Accessible through the Internet. *Journal of Chemical Information and Computer Sciences*. 2004;44(2):542-548.
89. Morgan HL. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *Journal of Chemical Documentation*. 1965;5(2):107-&.
90. Faulon JL, Bender A. *Handbook of Chemoinformatics Algorithms*: Taylor & Francis; 2010.
91. Daylight Chemical Information Systems. Daylight Theory Manual2008.
92. Metcalfe WK, Burke SM, Ahmed SS, Curran HJ. A Hierarchical and Comparative Kinetic Modeling Study of C1 – C2 Hydrocarbon and Oxygenated Fuels. *International Journal of Chemical Kinetics*. 2013;45(10):638-675.
93. Westbrook CK, Dryer FL. Simplified Reaction Mechanisms for the Oxidation of Hydrocarbon Fuels in Flames. *Combustion Science and Technology*. 1981;27(1-2):31-43.
94. Miller JA, Bowman CT. Mechanism and modeling of nitrogen chemistry in combustion. *Progress in Energy and Combustion Science*. 1989;15(4):287-338.
95. Hughes KJ, Turányi T, Clague AR, Pilling MJ. Development and testing of a comprehensive chemical mechanism for the oxidation of methane. *International Journal of Chemical Kinetics*. 2001;33(9):513-538.
96. Bowman CT, Hanson RK, Davidson DF, Gardiner WC, Jr., Lissianski V, Smith GP, D.M. G, Frenklach M, Goldenberg M. GRI-Mech. <http://combustion.berkeley.edu/gri-mech/>.
97. Chevalier C. *Entwicklung eines detaillierten Reaktionsmechanismus zur Modellierung der Verbrennungsprozesse von Kohlenwasserstoffen bei Hoch- und Niedertemperaturbedingungen*: Institut für Technische Verbrennung, Universität Stuttgart; 1993, Ph. D. Thesis.
98. Broadbelt LJ, Stark SM, Klein MT. Termination of Computer-Generated Reaction Mechanisms: Species Rank-Based Convergence Criterion. *Industrial & Engineering Chemistry Research*. 1995;34(8):2566-2573.
99. De Witt MJ, Dooling DJ, Broadbelt LJ. Computer Generation of Reaction Mechanisms Using Quantitative Rate Information: Application to Long-Chain Hydrocarbon Pyrolysis. *Industrial & Engineering Chemistry Research*. 2000;39(7):2228-2237.
100. Tomlin AS, Pilling MJ, Turányi T, Merkin JH, Brindley J. Mechanism reduction for the oscillatory oxidation of hydrogen: Sensitivity and quasi-steady-state analyses. *Combustion and Flame*. 1992;91(2):107-130.
101. Tomlin AS, Pilling MJ, Merkin JH, Brindley J, Burgess N, Gough A. Reduced Mechanisms for Propane Pyrolysis. *Industrial & Engineering Chemistry Research*. 1995;34(11):3749-3760.

102. Whitehouse LE, Tomlin AS, Pilling MJ. Systematic reduction of complex tropospheric chemical mechanisms, Part I: sensitivity and time-scale analyses. *Atmospheric Chemistry and Physics*. 2004;4:2025-2056.
103. Pierucci S, Ranzi E. A review of features in current automatic generation software for hydrocarbon oxidation mechanisms. *Computers & Chemical Engineering*. 2008;32(4-5):805-826.
104. Klinker DJ, Broadbelt LJ. Mechanism reduction during computer generation of compact reaction models. *Aiche Journal*. 1997;43(7):1828-1837.
105. Ranzi E, Faravelli T, Gaffuri P, Sogaro A. Low-temperature combustion: Automatic generation of primary oxidation reactions and lumping procedures. *Combustion and Flame*. 1995;102(1-2):179-192.
106. Li G, Tomlin AS, Rabitz H, Tóth J. A general analysis of approximate nonlinear lumping in chemical kinetics. I. Unconstrained lumping. *The Journal of Chemical Physics*. 1994;101(2):1172-1187.
107. Tomlin AS, Li G, Rabitz H, Tóth J. A general analysis of approximate nonlinear lumping in chemical kinetics. II. Constrained lumping. *The Journal of Chemical Physics*. 1994;101(2):1188-1201.
108. Laidler KJ. *Chemical Kinetics*: Pearson Education; 1987.
109. Ho TC. Kinetic Modeling of Large-Scale Reaction Systems. *Catalysis Reviews*. 2008;50(3):287-378.
110. Sun W, Chen Z, Gou X, Ju Y. A path flux analysis method for the reduction of detailed chemical kinetic mechanisms. *Combustion and Flame*. 2010;157(7):1298-1307.
111. Liang L, Stevens JG, Raman S, Farrell JT. The use of dynamic adaptive chemistry in combustion simulation of gasoline surrogate fuels. *Combustion and Flame*. 2009;156(7):1493-1502.
112. Tosatto L, Bennett BAV, Smooke MD. A transport-flux-based directed relation graph method for the spatially inhomogeneous instantaneous reduction of chemical kinetic mechanisms. *Combustion and Flame*. 2011;158(5):820-835.
113. Schietekat CM, Van Cauwenberge DJ, Van Geem KM, Marin GB. Computational fluid dynamics-based design of finned steam cracking reactors. *Aiche Journal*. 2014;60(2):794-808.
114. Burcat A, Gardiner WC. Ideal Gas Thermochemical Data for Combustion and Air Pollution Use. In: Gardiner WC, ed. *Gas-Phase Combustion Chemistry*. New York, NY: Springer New York; 2000:489-538.
115. van Speybroeck V, Gani R, Meier RJ. The calculation of thermodynamic properties of molecules. *Chemical Society Reviews*. 2010;39(5):1764-1779.
116. Benson SW. *Thermochemical kinetics: methods for the estimation of thermochemical data and rate parameters*: Wiley; 1976.
117. Cohen N, Benson SW. Estimation of heats of formation of organic compounds by additivity methods. *Chemical Reviews*. 1993;93(7):2419-2438.
118. Yamada T, Lay TH, Bozzelli JW. Ab initio calculations and internal rotor: Contribution for thermodynamic properties S-298(o) and C-p(T)'s (300 < T/K < 1500): Group additivity for fluoroethanes. *Journal of Physical Chemistry A*. 1998;102(37):7286-7293.
119. Vandeputte AG, Sabbe MK, Reyniers M-F, Marin GB. Modeling the Gas-Phase Thermochemistry of Organosulfur Compounds. *Chemistry – A European Journal*. 2011;17(27):7656-7673.

120. Sabbe MK, De Vleeschouwer F, Reyniers M-F, Waroquier M, Marin GB. First Principles Based Group Additive Values for the Gas Phase Standard Entropy and Heat Capacity of Hydrocarbons and Hydrocarbon Radicals. *Journal of Physical Chemistry A*. 2008;112(47):12235-12251.
121. Paraskevas PD, Sabbe MK, Reyniers M-F, Papayannakos N, Marin GB. Group Additive Values for the Gas-Phase Standard Enthalpy of Formation, Entropy and Heat Capacity of Oxygenates. *Chemistry – A European Journal*. 2013;19(48):16431-16452.
122. Ritter ER, Bozzelli JW. THERM: Thermodynamic property estimation for gas phase radicals and molecules. *International Journal of Chemical Kinetics*. 1991;23(9):767-778.
123. Muller C, Michel V, Scacchi G, Come GM. THERGAS: a computer program for the evaluation of thermochemical data of molecules and free radicals in the gas phase. *Journal De Chimie Physique Et De Physico-Chimie Biologique*. 1995;92(5):1154-1178.
124. Blurock ES, Warth V, Grandmougin X, Bounaceur R, Glaude PA, Battin-Leclerc F. JTHERGAS: Thermodynamic estimation from 2D graphical representations of molecules. *Energy*. 2012;43(1):161-171.
125. Vandewiele NM, Van De Vijver R, Carstensen H-H, Van Geem KM, Reyniers M-F, Marin GB. Implementation of Stereochemistry in Automatic Kinetic Model Generation. *International Journal of Chemical Kinetics*. 2016;48(12):755-769.
126. Lay TH, Bozzelli JW, Dean AM, Ritter ER. Hydrogen Atom Bond Increments for Calculation of Thermodynamic Properties of Hydrocarbon Radical Species. *Journal of Physical Chemistry*. 1995;99(39):14514-14527.
127. Platt JR. Influence of Neighbor Bonds on Additive Bond Properties in Paraffins. *The Journal of Chemical Physics*. 1947;15(6):419-420.
128. Platt JR. Prediction of Isomeric Differences in Paraffin Properties. *The Journal of Physical Chemistry*. 1952;56(3):328-336.
129. Greenshields JB, Rossini FD. Molecular Structure and Properties of Hydrocarbons and Related Compounds. *The Journal of Physical Chemistry*. 1958;62(3):271-280.
130. Thinh TP, Duran JL, Ramalho RS. Estimation of Ideal Gas Heat Capacities of Hydrocarbons from Group Contribution Techniques. New and Accurate Approach. *Industrial & Engineering Chemistry Process Design and Development*. 1971;10(4):576-582.
131. Thinh TP, Trong TK. Estimation of standard heats of formation $\Delta H^{\circ}fT$, standard entropies of formation $\Delta S^{\circ}fT$, standard free energies of formation $\Delta F^{\circ}fT$ and absolute entropies $S^{\circ}fT$ of hydrocarbons from group contributions: An accurate approach. *The Canadian Journal of Chemical Engineering*. 1976;54(4):344-357.
132. Smith DW. Additive bond-energy scheme for the calculation of enthalpies of formation of hydrocarbons including geminal H-H terms. *Journal of the Chemical Society, Faraday Transactions*. 1996;92(7):1141-1147.
133. Smith DW. Additive bond energy scheme for the calculation of enthalpies of formation and bond dissociation energies for alkyl radicals. *Journal of the Chemical Society, Faraday Transactions*. 1996;92(22):4415-4417.
134. Paldus J. Correlation problems in atomic and molecular systems. V. Spin-adapted coupled cluster many-electron theory. *The Journal of Chemical Physics*. 1977;67(1):303-318.
135. Cremer D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 2011;1(4):509-530.

136. Raghavachari K, Trucks GW, Pople JA, Head-Gordon M. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters*. 1989;157(6):479-483.
137. Stewart JJ. Optimization of parameters for semiempirical methods I. Method. *Journal of Computational Chemistry*. 1989;10(2):209-220.
138. Stewart JJP. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling*. 2007;13(12):1173-1213.
139. Gross E, Dreizler R. Density functional theory: Plenum Press ASI series B; 1995.
140. Kohn W, Sham LJ. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*. 1965;140(4A):A1133-A1138.
141. Gill PMW, Johnson BG, Pople JA, Frisch MJ. The performance of the Becke-Lee-Yang-Parr (B-LYP) density functional theory with various basis sets. *Chemical Physics Letters*. 1992;197(4):499-505.
142. Zhao Y, Schultz NE, Truhlar DG. Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions. *The Journal of Chemical Physics*. 2005;123(16):161103.
143. Montgomery Jr JA, Frisch MJ, Ochterski JW, Petersson GA. A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. *Journal of Chemical Physics*. 1999;110(6):2822.
144. Ochterski JW, Petersson GA, Montgomery Jr JA. A complete basis set model chemistry. V. Extensions to six or more heavy atoms. *Journal of Chemical Physics*. 1996;104(7):2598.
145. Pople JA, Head-Gordon M, Fox DJ, Raghavachari K, Curtiss LA. Gaussian-1 theory: A general procedure for prediction of molecular energies. *Journal of Chemical Physics*. 1989;90(10):5622.
146. Curtiss LA, Redfern PC, Raghavachari K. Gaussian-4 theory. *Journal of Chemical Physics*. 2007;126(8):084108.
147. Boese AD, Oren M, Atasoylu O, Martin JML, Kállay M, Gauss J. W3 theory: Robust computational thermochemistry in the kJ/mol accuracy range. *The Journal of Chemical Physics*. 2004;120(9):4129-4141.
148. Karton A, Rabinovich E, Martin JML, Ruscic B. W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions. *Journal of Chemical Physics*. 2006;125(14):144108.
149. Klippenstein SJ, Pande VS, Truhlar DG. Chemical Kinetics and Mechanisms of Complex Systems: A Perspective on Recent Theoretical Advances. *Journal of the American Chemical Society*. 2013;136(2):528-546.
150. Spellmeyer DC, Wong AK, Bower MJ, Blaney JM. Conformational analysis using distance geometry methods. *Journal of Molecular Graphics & Modelling*. 1997;15(1):18-36.
151. Ebejer J-P, Morris GM, Deane CM. Freely Available Conformer Generation Methods: How Good Are They? *Journal of Chemical Information and Modeling*. 2012;52(5):1146-1158.
152. Chang G, Guida WC, Still WC. An internal-coordinate Monte Carlo method for searching conformational space. *Journal of the American Chemical Society*. 1989;111(12):4379-4386.

153. Senderowitz H, Guarnieri F, Still WC. A Smart Monte Carlo Technique for Free Energy Simulations of Multiconformational Molecules. Direct Calculations of the Conformational Populations of Organic Molecules. *Journal of the American Chemical Society*. 1995;117(31):8211-8219.
154. Judson RS, Colvin ME, Meza JC, Huffer A, Gutierrez D. Do intelligent configuration search techniques outperform random search for large molecules? *International Journal of Quantum Chemistry*. 1992;44(2):277-290.
155. Ferguson DM, Raber DJ. A new approach to probing conformational space with molecular mechanics: random incremental pulse search. *Journal of the American Chemical Society*. 1989;111(12):4371-4378.
156. Saunders M, Houk KN, Wu YD, Still WC, Lipton M, Chang G, Guida WC. Conformations of cycloheptadecane. A comparison of methods for conformational searching. *Journal of the American Chemical Society*. 1990;112(4):1419-1427.
157. Leach AR. A Survey of Methods for Searching the Conformational Space of Small and Medium-Sized Molecules. *Reviews in Computational Chemistry*: John Wiley & Sons, Inc.; 2007:1-55.
158. Foloppe N, Chen IJ. Conformational Sampling and Energetics of Drug-Like Molecules. *Current Medicinal Chemistry*. 2009;16(26):3381-3413.
159. Beusen DD, Shands EFB. Systematic search strategies in conformational analysis. *Drug Discovery Today*. 1996;1(10):429-437.
160. Chen IJ, Foloppe N. Is conformational sampling of drug-like molecules a solved problem? *Drug Development Research*. 2011;72(1):85-94.
161. Ince A, Carstensen HH, Reyniers MF, Marin GB. First-principles based group additivity values for thermochemical properties of substituted aromatic compounds. *Aiche Journal*. 2015;61(11):3858-3870.
162. Magoon GR. *Automated Reaction Mechanism Generation: Improving Accuracy and Broadening Scope*: Chemical Engineering, Massachusetts Institute of Technology; 2012, Ph. D. Thesis.
163. Magoon GR, Green WH. Design and implementation of a next-generation software interface for on-the-fly quantum and force field calculations in automated reaction mechanism generation. *Computers & Chemical Engineering*. 2013;52:35-45.
164. Broadbelt LJ, Stark SM, Klein MT. Computer generated reaction networks: on-the-fly calculation of species properties using computational quantum chemistry. *Chemical Engineering Science*. 1994;49(24, Part 2):4991-5010.
165. Sadowski J, Gasteiger J. From atoms and bonds to three-dimensional atomic coordinates: automatic model builders. *Chemical Reviews*. 1993;93(7):2567-2581.
166. Broadbelt LJ, Stark SM, Klein MT. Computer Generated Pyrolysis Modeling: On-the-Fly Generation of Species, Reactions, and Rates. *Industrial & Engineering Chemistry Research*. 1994;33(4):790-799.
167. *Gaussian 03* [computer program]. Wallingford, CT, USA: Gaussian, Inc.; 2004.
168. Allinger NL, Chen K, Lii J-H. An improved force field (MM4) for saturated hydrocarbons. *Journal of Computational Chemistry*. 1996;17(5-6):642-668.
169. Blaney JM, Dixon JS. Distance Geometry in Molecular Modeling. *Reviews in Computational Chemistry*. 1994;5:299-335.
170. *RDKit: Open-source cheminformatics*; <http://rdkit.sourceforge.net> [computer program] 2006.

171. Osmont A, Catoire L, Goekalp I. Physicochemical properties and thermochemistry of propellanes. *Energy & Fuels*. 2008;22(4):2241-2257.
172. Evans MG, Polanyi M. Further considerations on the thermodynamics of chemical equilibria and reaction rates. *Transactions of the Faraday Society*. 1936;32(0):1333-1360.
173. Blowers P, Masel R. Engineering Approximations for Activation Energies in Hydrogen Transfer Reactions. *Aiche Journal*. 2000;46(10):2041-2052.
174. Denisov E. Models for Abstraction and Addition Reactions of Free Radicals. *ChemInform*. 1999;30(50).
175. Roberts BP, Steel AJ. An extended form of the Evans-Polanyi equation: a simple empirical relationship for the prediction of activation energies for hydrogen-atom transfer reactions. *Journal of the Chemical Society, Perkin Transactions 2*. 1994(10):2155-2162.
176. Blowers P, Masel RI. Extensions of the Marcus equation for the prediction of approximate transition state geometries in hydrogen transfer and methyl transfer reactions. *Theoretical Chemistry Accounts*. 2000;105(1):46-54.
177. Carstensen HH. Personal communication: Ab initio hydrogen abstraction rate coefficients and reaction enthalpies 2014.
178. Willems PA, Froment GF. Kinetic modeling of the thermal cracking of hydrocarbons. 1. Calculation of frequency factors. *Industrial & Engineering Chemistry Research*. 1988;27(11):1959-1966.
179. Willems PA, Froment GF. Kinetic modeling of the thermal cracking of hydrocarbons. 2. Calculation of activation energies. *Industrial & Engineering Chemistry Research*. 1988;27(11):1966-1971.
180. Saeyns M, Reyniers M-F, Marin GB, Van Speybroeck V, Waroquier M. Ab initio group contribution method for activation energies for radical additions. *Aiche Journal*. 2004;50(2):426-444.
181. Sabbe MK, Vandeputte AG, Reyniers M-F, Waroquier M, Marin GB. Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions. *Physical Chemistry Chemical Physics*. 2010;12(6):1278-1298.
182. Sumathi R, Carstensen HH, Green WH. Reaction rate predictions via group additivity. Part 3: Effect of substituents with CH₂ as the mediator. *Journal of Physical Chemistry A*. 2002;106(22):5474-5489.
183. Sumathi R, Carstensen HH, Green WH. Reaction rate prediction via group additivity, part 2: H-abstraction from alkenes, alkynes, alcohols, aldehydes, and acids by H atoms. *Journal of Physical Chemistry A*. 2001;105(39):8969-8984.
184. Sumathi R, Carstensen HH, Green WH. Reaction rate prediction via group additivity Part 1: H abstraction from alkanes by H and CH₃. *Journal of Physical Chemistry A*. 2001;105(28):6910-6925.
185. Vandeputte AG, Sabbe MK, Reyniers M-F, Marin GB. Kinetics of alpha hydrogen abstractions from thiols, sulfides and thiocarbonyl compounds. *Physical Chemistry Chemical Physics*. 2012;14(37):12773-12793.
186. Vandeputte AG, Reyniers M-F, Marin GB. Kinetics of Homolytic Substitutions by Hydrogen Atoms at Thiols and Sulfides. *Chemphyschem*. 2013;14(8):1703-1722.
187. Adamczyk AJ, Reyniers M-F, Marin GB, Broadbelt LJ. Exploring 1,2-Hydrogen Shift in Silicon Nanoparticles: Reaction Kinetics from Quantum Chemical Calculations and Derivation of Transition State Group Additivity Database. *The Journal of Physical Chemistry A*. 2009;113(41):10933-10946.

188. Adamczyk AJ, Reyniers M-F, Marin GB, Broadbelt LJ. Kinetic correlations for H₂ addition and elimination reaction mechanisms during silicon hydride pyrolysis. *Physical Chemistry Chemical Physics*. 2010;12(39):12676-12696.
189. Muller C, Scacchi G, Come GM. A topological method for determining the external symmetry number of molecules. *Computers & Chemistry*. 1991;15(1):17-27.
190. Yu J. *Estimation method for the thermochemical properties of polycyclic aromatic molecules*: Department of Chemical Engineering, Massachusetts Institute of Technology; 2005, Ph. D. Thesis.
191. Vandewiele NM, Van de Vijver R, Van Geem KM, Reyniers M-F, Marin GB. Symmetry calculation for molecules and transition states. *Journal of Computational Chemistry*. 2015;36(3):181-192.
192. Pollak E, Pechukas P. Symmetry Numbers, Not Statistical Factors, Should be Used in Absolute Rate Theory and in Bronsted Relations. *Journal of the American Chemical Society*. 1978;100(10):2984-2991.
193. Fernández-Ramos A, Ellingson BA, Meana-Pañeda R, Marques JMC, Truhlar DG. Symmetry numbers and chemical reaction rates. *Theoretical Chemistry Accounts*. 2007;118:813-826.
194. Kwok ESC, Atkinson R. Estimation of hydroxyl radical reaction rate constants for gas-phase organic compounds using a structure-reactivity relationship: An update. *Atmospheric Environment*. 1995;29(14):1685-1695.
195. Peeters J, Vertommen J, Langhans I, Boullart W, Hoeymissen J, Pultau V. Kinetic Studies of Reactions of Alkylperoxy and Haloalkylperoxy Radicals with NO. A Structure-Activity Relationship for Reactions of OH with Alkenes and Polyalkenes. In: Bras G, ed. *Chemical Processes in Atmospheric Oxidation*. Vol 3: Springer Berlin Heidelberg; 1997:179-185.
196. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Kinetic Modeling of α -Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Hydrogen Atoms. *Journal of Physical Chemistry A*. 2014;118(40):9296-9309.
197. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos N, Marin GB. Kinetic Modeling of α -Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Carbon-Centered Radicals. *Chemphyschem*. 2014;15(9):1849-1866.
198. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Group Additive Kinetics for Hydrogen Transfer Between Oxygenates. *Journal of Physical Chemistry A*. 2015;119(27):6961-6980.
199. Paraskevas PD, Sabbe MK, Reyniers MF, Marin GB, Papayannakos NG. Group additive kinetic modeling for carbon-centered radical addition to oxygenates and -scission of oxygenates. *Aiche Journal*. 2016;62(3):802-814.
200. Sabbe MK, Reyniers M-F, Waroquier M, Marin GB. Hydrogen Radical Additions to Unsaturated Hydrocarbons and the Reverse β -Scission Reactions: Modeling of Activation Energies and Pre-Exponential Factors. *Chemphyschem*. 2010;11(1):195-210.
201. Song J, Raman S, Yu J, Wijaya CD, Stephanopoulos G, Green WH. RMG: the next generation of automatic chemical reaction mechanism generator. Paper presented at: AIChE Annual Meeting 2003; CA, USA.
202. Allen JW. *Predictive chemical kinetics : enabling automatic mechanism generation and evaluation*: Chemical engineering, Massachusetts Institute of Technology; 2013, Ph. D. Thesis.

203. Sumathy R, Carstensen H-H. Tree structure for intermolecular hydrogen abstraction from hydrocarbons (C/H) and generic rate constant rules for abstraction by vinyl radical. *International Journal of Chemical Kinetics*. 2012;44(5):327-349.
204. Truong TN. Reaction class transition state theory: Hydrogen abstraction reactions by hydrogen atoms as test cases. *Journal of Chemical Physics*. 2000;113(12):4957-4964.
205. Eckart C. The Penetration of a Potential Barrier by Electrons. *Physical Review*. 1930;35(11):1303-1309.
206. Zhang S, Truong TN. Kinetics of Hydrogen Abstraction Reaction Class H + H-C(sp³): First-Principles Predictions Using the Reaction Class Transition State Theory. *The Journal of Physical Chemistry A*. 2003;107(8):1138-1147.
207. Kungwan N, Truong TN. Kinetics of the Hydrogen Abstraction •CH₃ + Alkane → CH₄ + Alkyl Reaction Class: An Application of the Reaction Class Transition State Theory. *The Journal of Physical Chemistry A*. 2005;109(34):7742-7750.
208. Huynh LK, Ratkiewicz A, Truong TN. Kinetics of the Hydrogen Abstraction OH + Alkane → H₂O + Alkyl Reaction Class: An Application of the Reaction Class Transition State Theory. *The Journal of Physical Chemistry A*. 2006;110(2):473-484.
209. Huynh LK, Zhang S, Truong TN. Kinetics of hydrogen abstraction O(³P) + alkane → OH + alkyl reaction class: An application of the reaction class transition state theory. *Combustion and Flame*. 2008;152(1-2):177-185.
210. Ratkiewicz A, Bieniewska J, Truong TN. Kinetics of the hydrogen abstraction R-OH + H → R•-OH + H₂ reaction class: An application of the reaction class transition state theory. *International Journal of Chemical Kinetics*. 2011;43(2):78-98.
211. Ratkiewicz A, Huynh LK, Pham QB, Truong TN. Kinetics of the hydrogen abstraction •C₂H₅ + alkane → C₂H₆ + alkyl reaction class: an application of the reaction class transition state theory. *Theoretical Chemistry Accounts*. 2013;132(3):1-17.
212. Piansawan T, Kungwan N, Jungsuttiwong S. Application of the reaction class transition state theory to the kinetics of hydrogen abstraction reactions of alkanes by atomic chlorine. *Computational and Theoretical Chemistry*. 2013;1011(0):65-74.
213. Wang Q-D, Wang X-J, Kang G-J. An application of the reaction class transition state theory to the kinetics of hydrogen abstraction reactions of hydrogen with methyl esters at the methoxy group. *Computational and Theoretical Chemistry*. 2014;1027(0):103-111.
214. Wang B-Y, Li Z-R, Tan N-X, Yao Q, Li X-Y. Interpretation and Application of Reaction Class Transition State Theory for Accurate Calculation of Thermokinetic Parameters Using Isodesmic Reaction Method. *The Journal of Physical Chemistry A*. 2013;117(16):3279-3291.
215. Zádor J, Najm HN. KinBot 1.0: A code for automatic PES exploration. *8th U. S. National Combustion Meeting* 2013.
216. Bhoorasingh PL, West RH. Transition state geometry prediction using molecular group contributions. *Physical Chemistry Chemical Physics*. 2015;17(48):32173-32182.
217. Allen JW, Goldsmith CF, Green WH. Automatic estimation of pressure-dependent rate coefficients. *Physical Chemistry Chemical Physics*. 2012;14(3):1131-1155.
218. Troe J. Predictive possibilities of unimolecular rate theory. *The Journal of Physical Chemistry*. 1979;83(1):114-126.
219. Troe J. Theory of Thermal Unimolecular Reactions in the Fall-off Range. I. Strong Collision Rate Constants. *Berichte der Bunsengesellschaft für physikalische Chemie*. 1983;87(2):161-169.

220. Stewart PH, Larson CW, Golden DM. Pressure and temperature dependence of reactions proceeding via a bound complex. 2. Application to $2\text{CH}_3 \rightarrow \text{C}_2\text{H}_5 + \text{H}$. *Combustion and Flame*. 1989;75(1):25-31.
221. Venkatesh PK, Chang AY, Dean AM, Cohen MH, Carr RW. Parameterization of pressure- and temperature-dependent kinetics in multiple well reactions. *Aiche Journal*. 1997;43(5):1331-1340.
222. Wang H, Frenklach M. Modification of Troe's fall-off broadening. *Chemical Physics Letters*. 1993;205(2-3):271-276.
223. Zhang P, Law CK. A fitting formula for the falloff curves of unimolecular reactions. *International Journal of Chemical Kinetics*. 2009;41(11):727-734.
224. *CHEMKIN Release 4.1.1* [computer program]. San Diego, CA, USA: Reaction Design, Inc.; 2007.
225. Matheu DM, Lada Ii TA, Green WH, Dean AM, Grenda JM. Rate-based screening of pressure-dependent reaction networks. *Computer Physics Communications*. 2001;138(3):237-249.
226. Matheu DM, Green WH, Grenda JM. Capturing pressure-dependence in automated mechanism generation: Reactions through cycloalkyl intermediates. *International Journal of Chemical Kinetics*. 2003;35(3):95-119.
227. Dean AM. Predictions of pressure and temperature effects upon radical addition and recombination reactions. *The Journal of Physical Chemistry*. 1985;89(21):4600-4608.
228. Seyedzadeh Khanshan F, West R. Using Reaction Mechanism Generator (RMG) to Build Detailed Kinetic Model of Biofuels. *AIChE Annual Meeting*. San Fransisco, CA2013.
229. Schleyer PvR, Jiao H. What is aromaticity? *Pure and Applied Chemistry*. 1996;68(2):209-218.
230. Djokic M, Carstensen H-H, Van Geem KM, Marin GB. The thermal decomposition of 2,5-dimethylfuran. *Proceedings of the Combustion Institute*. 2013;34(1):251-258.
231. Vandewiele NM, Magoon GR, Van Geem KM, Reyniers M-F, Green WH, Marin GB. Experimental and Modeling Study on the Thermal Decomposition of Jet Propellant-10. *Energy & Fuels*. 2014;28(8):4976-4985.
232. Cavallotti C, Polino D. On the kinetics of the $\text{C}_5\text{H}_5 + \text{C}_5\text{H}_5$ reaction. *Proceedings of the Combustion Institute*. 2013;34(1):557-564.
233. Cyrański MK, Krygowski TM, Katritzky AR, Schleyer PvR. To What Extent Can Aromaticity Be Defined Uniquely? *The Journal of Organic Chemistry*. 2002;67(4):1333-1338.
234. Magoon GR, Green Jr WH, Oluwole O, Wong HS, Albo S, Lewis D. Updating Our Understanding of JP-10 Decomposition Chemistry: A Detailed JP-10 Combustion Mechanism Constructed Using RMG, an Automatic Reaction Mechanism Generator. *46th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*: American Institute of Aeronautics and Astronautics; 2010.
235. Bhoorasingh PL, West RH. Automatic Transition State Searches for On-The-Fly Kinetic Calculations. Paper presented at: AIChE Annual Meeting2013; San Francisco, CA.
236. Bhoorasingh PL, West RH. Automatically Locating Hydrogen Abstraction Transition States via Double-Ended Searches. *International Journal of Chemical Kinetics*. submitted.
237. Comandini A, Dubois T, Abid S, Chaumeix N. Comparative Study on Cyclohexane and Decalin Oxidation. *Energy & Fuels*. 2014;28(1):714-724.

238. Chen L, Zhang T, Li C, Wang W, Lü J, Wang W. Theoretical kinetic investigation of thermal decomposition of methylcyclohexane. *Computational and Theoretical Chemistry*. 2013;1026(0):38-45.
239. Wang Z, Cheng Z, Yuan W, Cai J, Zhang L, Zhang F, Qi F, Wang J. An experimental and kinetic modeling study of cyclohexane pyrolysis at low pressure. *Combustion and Flame*. 2012;159(7):2243-2253.
240. Serinyel Z, Herbinet O, Frottier O, Dirrenberger P, Warth V, Glaude PA, Battin-Leclerc F. An experimental and modeling study of the low- and high-temperature oxidation of cyclohexane. *Combustion and Flame*. 2013;160(11):2319-2332.
241. De Bruycker R, Carstensen H-H, Simmie JM, Van Geem KM, Marin GB. Experimental and computational study of the initial decomposition of gamma-valerolactone. *Proceedings of the Combustion Institute*.
242. Steinbeck C, Han Y, Kuhn S, Horlacher O, Luttmann E, Willighagen E. The Chemistry Development Kit (CDK): An Open-Source Java Library for Chemo- and Bioinformatics. *Journal of Chemical Information and Computer Sciences*. 2003;43(2):493-500.
243. Steinbeck C, Hoppe C, Kuhn S, Floris M, Guha R, Willighagen EL. Recent Developments of the Chemistry Development Kit (CDK) - An Open-Source Java Library for Chemo- and Bioinformatics. *Current Pharmaceutical Design*. 2006;12(17):2111-2120.
244. Sabbe MK, Vandeputte AG, Reyniers M-F, Van Speybroeck V, Waroquier M, Marin GB. Ab Initio Thermochemistry and Kinetics for Carbon-Centered Radical Addition and β -Scission Reactions. *The Journal of Physical Chemistry A*. 2007;111(34):8416-8428.
245. Vandeputte AG, Sabbe MK, Reyniers M-F, Marin GB. Modeling the Gas-Phase Thermochemistry of Organosulfur Compounds. *Chemistry-a European Journal*. 2011;17(27):7656-7673.

Chapter 2: Decomposition and isomerization of 1-pentanol radicals and the pyrolysis of 1-pentanol

2.1 Abstract

Stable species and saddle points on the potential energy surface of the terminal addition of hydroxyl to 1-pentene have been determined, which are of interest for n-pentanol pyrolysis and combustion. All the stationary points have been determined automatically with the KinBot software and have been calculated using UCCSD(T)-F12/cc-pVTZ-F12//M062X/6-311++G(d,p) quantum chemistry calculations. The master equation has been solved to obtain $k(T,p)$ for all the reactions on the potential energy surface in a temperature range of 50-3000K and a pressure range of 0.001-10 MPa. The entrance channel is characterized by a barrierless step into a Van der Waals-well followed by a transition state, for which an effective two transition states model has been employed. The newly obtained rate coefficients have been implemented in a kinetic model for the thermal decomposition of 1-pentanol diluted in a nitrogen stream. The results prove that the reactions automatically found by KinBot are of significant importance to correctly describe conversion and product selectivities.

2.2 Introduction

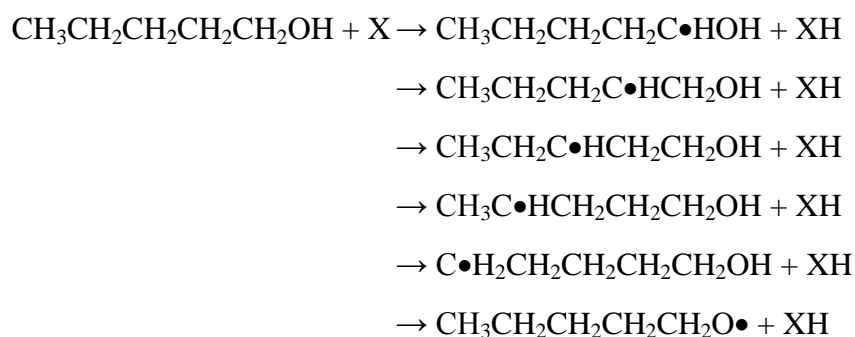
The increasing worldwide energy demand and the major environmental constraints on greenhouse gas emissions are strongly driving the search for alternative fuel and base chemical resources. Many studies have reported industrial, experimental, and theoretical results on feedstocks derived from non-fossil processes, among which urban waste and biological feedstocks are the most promising. Primary alcohols have already been considered for decades as sustainable fuel; ethanol, belonging to the first generation biofuels, is a widespread substitute for conventional diesel.¹ Other alcohols such as propanol and n- and iso-butanol gained significant interest in the past for their higher energy density, higher cetane numbers and lower hygroscopic properties. To increase the energy content even more, the second-generation biofuel n-pentanol is now extensively studied. Its performance in engines has been studied to minimize the CO, hydrocarbon, NO_x, and smoke emission,^{2,3} particulate emissions,⁴ mixing characteristics,⁵ and optimize the cetane number⁶ of a pentanol/diesel blend. Experimental and modeling work on the combustion properties such as laminar flame speeds^{7,8} and ignition characteristics⁹ have been reported.

Optimizing a well-established and well-understood chemical process such as combustion or pyrolysis for a new, alternative feedstock requires the knowledge of the underlying chemistry. Alternative feedstocks contain a fundamentally different chemical structure compared to fossil resources, and their chemistry is thus not comparable. The growing experimental database on these feedstocks is a first step towards optimal chemical processes. However, experimental work is expensive, time-consuming, and has several constraints on operating conditions and reactor configurations. Therefore, chemists and chemical engineers have developed engineering approximations and computer simulations, which complement the experimental results and allow directed experimental design, lowering the cost and time of the process development.

A fundamental part of the computer simulations for chemical processes is the availability of a chemical kinetic model, i.e. a set of rate expressions of the occurring reactions. Kinetic models for thermal decomposition, steam cracking, and combustion processes can contain up to thousands of reactions, for which most of the rate coefficients have never been measured experimentally. Typically many of the reaction rate coefficients are not accurately known and calculation methods are applied such as group additivity and linear free energy correlations, but

their accuracy is questionable if they are applied beyond the training set they were developed for. This limitation is nowadays overcome by calculating many rate coefficients ab initio by taking advantage of the growing computational power and availability of high-performance computing solutions. However, calculating all the thermodynamic properties of species and reaction rate coefficients of the selected reactions through high level quantum chemistry calculations is still too time consuming to be implemented on-the-fly since (1) many manual interventions are necessary, (2) systematic conformational searches become practically impossible for molecules with several torsional modes, (3) even for small molecules, finding all the reactions is tedious and pathways can easily be overlooked, and (4) locating transition states sometimes needs to be done on a trial-and-error basis. To overcome these challenges and ensure a full exploration of the potential energy surface (PES), a computer code called KinBot¹⁰ has been developed, which will be discussed in more detail in section 2.3.1.

In this chapter the capabilities of KinBot¹⁰ are illustrated for the thermal decomposition of n-pentanol. This is a challenging problem because of the intrinsic complexity of the decomposition chemistry. During the pyrolysis of 1-pentanol the first radicals are formed when the C–C, C–O, C–H, or O–H bonds break. At the same time the unimolecular dissociation of n-pentanol continues to be an important process, abstraction reactions mostly by X = H, CH₃, OH radicals are driving the secondary chemistry. These abstraction reactions yield six different pentanol radicals:



In the remaining of the chapter, the radical site will not be marked in the structures, instead, the radicals will be denoted α -R, β -R, γ -R, δ -R, ε -R, and o-R, alluding to the location of the radical site.

While parameters for the unimolecular reactions of smaller alcoholic radicals are available in the literature, the information on the decomposition reactions of these radicals only consist of approximations or values derived from indirect or relative experiments. The unimolecular decomposition of β -R is known to primarily lead to OH + 1-pentene, and the reverse reaction has been investigated more directly. Nip and Paraskevopoulos,¹¹ Biermann et al.,¹² and McGillen et al.¹³ measured the absolute rate coefficient of the OH + 1-pentene reaction around room temperature, and the latter¹³ also determined its temperature dependence in the 262 – 312 K range. These values are only of limited use, for three reasons. First, the addition reactions can lead to terminal and non-terminal adducts, while for 1-pentanol only the terminal adduct is relevant. Second, a fraction of this reaction proceeds via abstraction, even at room temperature, according to Biermann et al.¹² this fraction amounts around 13%. And finally, the temperature range at which these experiments were carried out is outside the range of interest for pyrolysis studies. Nevertheless, these calculations can be used to compare with the theoretical results obtained in this work. There are two other studies that determined the unimolecular dissociation and isomerization rate coefficients for ε -R. The work of Heimann et al.¹⁴ is an indirect measurement, while Hein et al.¹⁵ used a direct measurement. However, both measurements were only able to put a lower limit on the isomerization rate coefficient. In summary, there is very little information in the literature about these and the related reactions.

In this work, the Ab Initio Transition State Theory Master Equation (AITSTME) approach has been used, which has been shown to be successful in the past to study the analogous reactions of smaller alcohol radicals up to C₄,¹⁶⁻²¹. AITSTME-based pressure- and temperature-dependent rate coefficients for 1-pentanol radical decomposition and isomerization have been calculated at the UCCSD(T)-F12/cc-pVTZ-F12//M06-2X/6-311++G(d,p) level of theory. However, because the potential energy surface is even more complicated than in the case of 1-butanol,²⁰ KinBot was employed to explore and characterize the relevant stationary points. The trends for radical decompositions in this class of reactions is summarized and compared the calculations to state-of-the-art group additivity rate coefficients of Genesys. Finally, the new rate coefficients have been used to update the Genesys-generated model for 1-pentanol pyrolysis, which is then compared against flow reactor experiments.

2.3 Methods

2.3.1 Exploration of the PES by KinBot

A large number of the known reaction types for hydrocarbon and oxygenate molecules and radicals have been translated to automatic procedures in the KinBot code.¹⁰ KinBot builds on the idea that most reactions can be classified in a reaction family, and that the transition states of all reactions belonging to the same family have important similarities. These similarities are translated in an estimation of the geometrical structure of the transition state, i.e. bond lengths, bond angles and torsional angles. By collecting the available chemical knowledge on possible reactions for hydrocarbons and oxygenate molecules and implementing this in the code, KinBot can automatically explore reaction pathways. The input comprises solely of the geometry of the starting structure and KinBot finds most of the reactions that are important for that structure. With this computational approach of searching for reactions, conformational searches can also be automated with systematic calculations ensuring the lowest energy conformer is found.

Starting from the geometry of a reactant, KinBot crawls on the potential energy surface in search for transition states. From the structure of the transition state, the code searches for products by following the reaction coordinate, i.e. intrinsic reaction coordinate calculations, into local minima on the PES. One of the local minima is the reactant in KinBot, on the other side of the energy barrier, a product can be found. This product is either a new well on the same PES or a bimolecular product. In case of a well, KinBot continues to search for reactions with this well as reactant. In case of bimolecular products, KinBot halts its search and considers this an endpoint of the current pathway. By iteratively applying this method, the full PES is searched for, according to user-defined rules. For example, the search past an energy barrier will only continue if the barrier height is lower than a user-defined energy, which amounted 418 kJ mol^{-1} in this work. Other input variables are the level of theory each calculation is done at, the computational aspects such as memory and number of cores, etc.

First, all stationary points were located at the B3LYP/6-31+G level of theory, which was also used for the conformational sampling, intrinsic reaction coordinates calculations, and the hindered rotor potentials. The final geometry of each stationary point was obtained using the M06-2X/6-311++G(d,p) level of theory, at which level the molecular frequencies were

calculated. All the aforementioned calculations were done using the Gaussian09 suite of programs²². Accurate single-point electronic energies were calculated using UCCSD(T)-F12/cc-pVTZ-F12 level of theory as implemented in Molpro 2009.²³

To limit the calculation time, the hindered rotor potentials were obtained at the B3LYP/6-31+G level of theory. Several rotors were repeated at M06-2X/6-311++G(d,p) to evaluate the difference in level of theory. Only small deviations were seen, as reported in Appendix A, suggesting that the B3LYP potentials can be used to calculate the rate coefficients.

2.3.2 OH + 1-pentene channel

One crucial part of the PES is the OH + 1-pentene exit channel. This is a dynamically complex pathway that consists of a barrierless outer transition state and a submerged inner barrier. The reaction is controlled by the inner TS at the temperatures relevant for pyrolysis. However, in order to compare it to the available room temperature data an effective two-transition-state model^{24,25} was used. In this, the sum of states for the transitional modes of the barrierless channel is calculated using variable-reaction coordinate transition state theory (VRC-TST).^{26,27} State averaged CASPT2(5e,4o)/cc-pVDZ²³ calculations was used to evaluate the long-range potential on-the-fly between OH and 1-pentene with an active space consisting π and π^* orbitals of 1-pentene (2 electrons), and the p orbital (2 electrons) and the radical orbital (1 electron) of OH. Geometry relaxation is negligible in this region,^{17,28} but a 1-D correction to achieve aug-cc-pVDZ accuracy is used. The calculated conserved mode number of states was convolved with the internal mode state counts including the hindered rotor sum of states. Finally, the inner and outer transition states were combined in a two-transition-state model. The effect of the 139.7 cm^{-1} spin-orbit coupling of the OH radical on the canonical level in the barrierless calculations is included, and a systematic study on how the calculations are affected by the various conformers of the pentene molecule is also conducted. The resulting corrections were less than 10%. The details of the calculations for the corrections are described in Appendix B.

2.3.3 Master Equation calculations

To calculate pressure- and temperature-dependent rate coefficients, the MESS code was used.^{29,30} In these calculations a 1-D hindered rotor corrections for the state counts is included, in which the rotational frequencies were manually removed from the set of molecular frequencies, and Eckart barriers were used to approximate the tunneling corrections at the microcanonical level. The

symmetry numbers for each species were determined using the principles described by Pollak and Pechukas.³¹

2.3.4 Experimental data

The pyrolysis of n-pentanol has been studied experimentally on a Bench Scale set-up, which has been described previously by Djokic et al.³² and Harper et al.³³.

Nitrogen and n-pentanol have been purchased from Sigma Aldrich and Air Liquide respectively, with a purity of 99.999+% and 99% respectively. The reactor is an Incoloy 800HT tube of 1.475m long with an internal diameter of 6mm, placed in a vertical furnace consisting of four heating sections controlled by four thermocouples inside the reactor tube. The pressure is measured at the in- and outlet of the reactor, with an inlet set point of 0.17 MPa. The pressure drop along the reactor can be neglected. Two sets of experiments were done, one with at a pentanol dilution of 1:4 in nitrogen gas, another one with a dilution of 1:1. For the first set, the pentanol inlet flow rate amounts 0.013 g s^{-1} , for the second one 0.033 g s^{-1} . The temperature inside the reactor was varied between 913 and 1073 K with increments of 20 K.

The reactor effluent is directed to a sample valve maintained at 573K, after which a gaseous part is injected on-line on a gas chromatograph to quantify C4- hydrocarbon species using a flame ionization detector. Permanent gasses are quantified with two thermal conductivity detectors. The concentrations of all the species are calculated based on the flow rate of nitrogen gas. Response factors were deducted using a calibration mixture provided by Air Liquide, Belgium. Another fraction of the reactor effluent is injected on a gas chromatograph used to identify small oxygenated molecules such as formaldehyde, methanol and water. Propylene, which is quantified on the first detector, is used as secondary internal standard here. A last part of the effluent is sent to an on-line GC×GC with is used to identify and quantify the species, using a time of flight mass spectrometer and a flame ionization detector respectively. Ethene is used as secondary internal standard. The effective carbon number approach³⁴ is used to obtain response factors. The transfer lines between the reactor and the analysis section are heated to prevent condensation. The analysis procedure allows on-line analysis of the complete product spectrum. Carbon balances close with 5% percent. The uncertainty on the mole fractions is estimated to be 5%, according to earlier work.^{32,33}

The Bench Scale set-up can be modelled as an ideal plug-flow reactor.³³ Chemkin³⁵ is used as kinetic model simulation tool. The reactor conditions and inlet flow rates are taken from the experimental measurements. The temperature profile and pressure are imposed on the reactor.

2.3.5 Kinetic model construction

To test the rate coefficients, a pentanol pyrolysis model has been built. Genesys³⁶ is a recently developed automatic kinetic model generator integrated with advanced open-source cheminformatics libraries, allowing the generation of kinetic models independent of the reactants, elements, or chemistry. For this work, the rule-based termination criterion of Genesys was employed. Reaction families relevant for hydrocarbon and oxygenates pyrolysis have been constructed, consisting of hydrogen abstraction reactions, β -scission reactions and the reverse radical addition reactions, homolytic bond scissions and the reverse recombination reactions, and CO formation reactions. Each reaction family is constrained to only allow relevant reactions to be generated. These constraints are defined based on the immediate surroundings of reactive centers, e.g. valency or hybridization, as well as on the molecule in its entirety, such as molecule size, presence of rings or aromaticity. For most reaction families, group additivity is employed to calculate the rate coefficients.³⁷⁻⁴³ Species thermochemistry is also calculated from group additivity.⁴⁴⁻⁴⁶

In this work the n-pentanol pyrolysis sub-model generated by Genesys is merged with the Aramco Mech,⁴⁷ which describes the pyrolysis and oxidation chemistry of small compounds, with up to 4 carbon atoms. The kinetics for the homolytic scission reactions of C-H, C-C, C-O and O-H bonds are taken from the literature, as well as the H₂O elimination in n-pentanol.⁴⁸ Finally, the reactions and kinetics generated by KinBot for n-pentanol radicals are added to the model. The merging of kinetic models is done semi-automatically in several steps. First, to be able to compare species, their names are mapped with the corresponding InChI's. Then, the models are compared to identify large differences in thermodynamics and kinetics. If these differences exist, it is the responsibility of the user to select the proper set of data. In a next step the models are merged using a master/slave approach. Here, the master kinetic model is taken as is, and the reactions and species of the slave model that are not present in the master model are added. The naming of the species is adapted to correspond to the naming of the master model. A

final manual check is necessary to identify inconsistencies by calculating reverse reaction rate coefficients and comparing them to the reverse reaction rate coefficients in the original models.

2.4 Results

2.4.1 PES

For the current study, β -R was used as starting structure in KinBot. From this structure, the other 5 radicals were found via H-transfer transition states, alongside 17 bimolecular products. All the reactions found by KinBot can be found in Figure 1, in which all energies are relative to the pentene + OH energy. All the reactions can be categorized into three reaction families. Firstly, 15 reactions are intramolecular hydrogen abstractions among 6 radicals. These reactions comprise of 1-2, 1-3, 1-4, 1-5 and one 1-6 internal hydrogen shifts. 1-2 and 1-3 hydrogen shift reactions have a high barrier, between 159 and 201 kJ mol⁻¹, compared to the reactions with a larger ring structure in the transition state, where the barrier amounts between 84 and 134 kJ mol⁻¹. Secondly, 13 more reactions are β -scission reactions yielding a bimolecular product pair, i.e., a radical species and an unsaturated molecule. The entrance channel is the OH addition to 1-pentene on the first carbon atom, and is characterized by a barrierless step into a van der Waals-well followed by a barrier to form the initial well on the PES, as shown in Figure 1a. The other reactions are H- β -scissions forming C₅H₁₀O unsaturated compounds and C- β -scissions. The highest barriers correspond to the H- β -scission reactions and amount up to 197 kJ mol⁻¹ compared to its well, and the lowest energy pathway to a bimolecular product is the formation of formaldehyde and the n-propyl radical (123 kJ mol⁻¹ above the well). Finally, 4 direct water elimination reactions were found, which have considerably higher reaction barriers compared to all previous reactions: 285-331 kJ mol⁻¹. From Figure 1, the lowest energy well can be identified as the α -R, which lies 44 kJ mol⁻¹ below the highest one, the o-R.

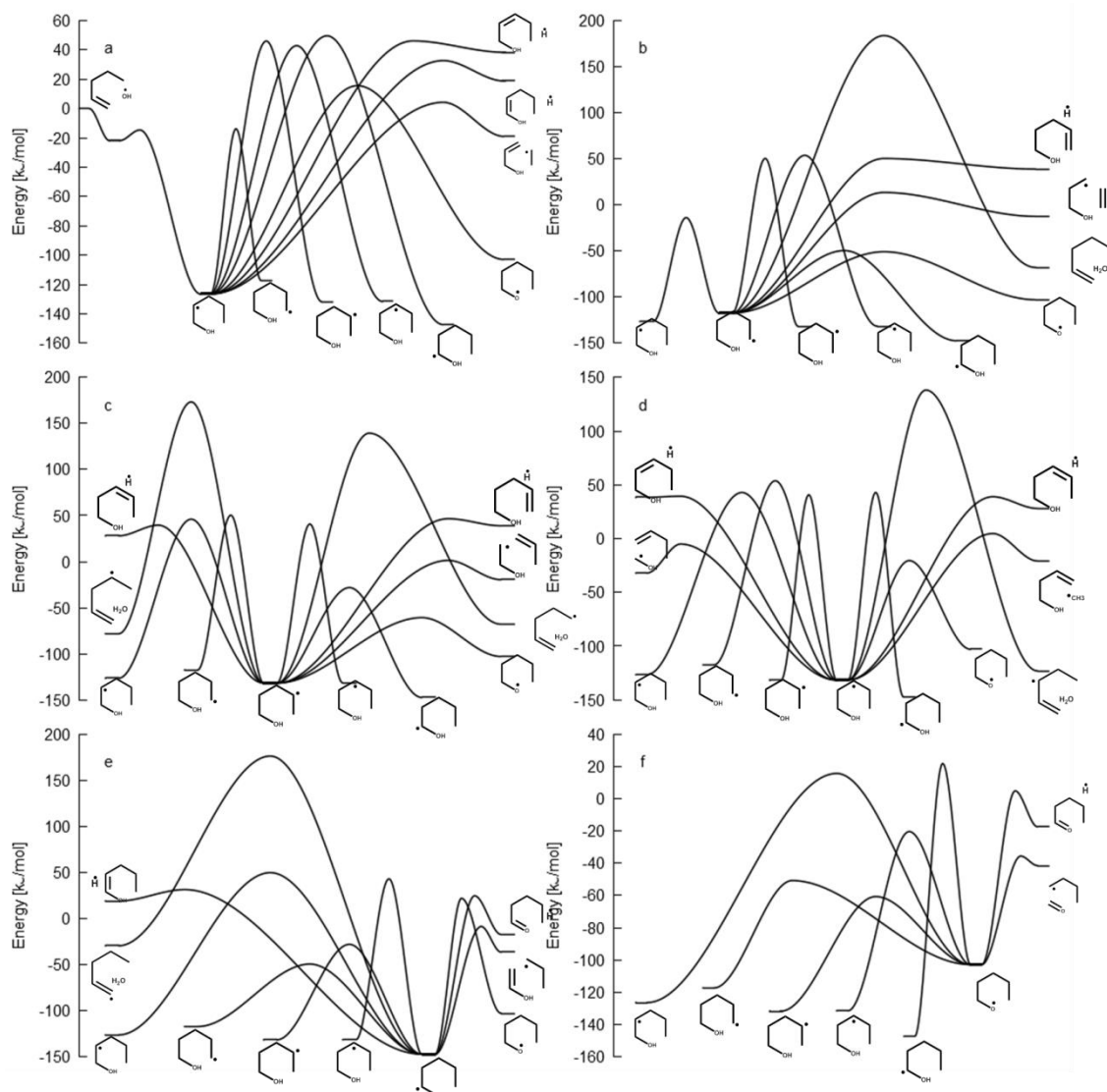


Figure 1: Potential energy surface of the 1-pentanol radicals. The transition states and products are separated per well: (a) β -R, (b) ϵ -R, (c) δ -R, (d) γ -R, (e) α -R, and (f) o-R.

2.4.2 OH + 1-pentene capture rate coefficient

The calculated OH + 1-pentene capture rate coefficients at low temperatures can be compared to experimental data available in the literature.¹¹⁻¹³ However, as mentioned in the introduction, the experiments measure addition both to the terminal and the non-terminal C-atom of the double bond, and potentially include contributions at room temperature from abstraction. The latter contribution was measured to be 13 ± 5 % by Biermann et al.¹² at 298 K, but there is no information on the branching ratio for the two adducts. Based on previous calculations for propene¹⁷, the branching into the two wells is estimated to be 1:1, and the contribution from the abstraction is $\sim 30\%$ lower at the low, and $\sim 10\%$ higher at the high end of the 262 – 312 K

experimental temperature range of McGillen et al.¹³ Applying these estimated corrections, the calculations are compared to the experimental values as shown in Figure 2.

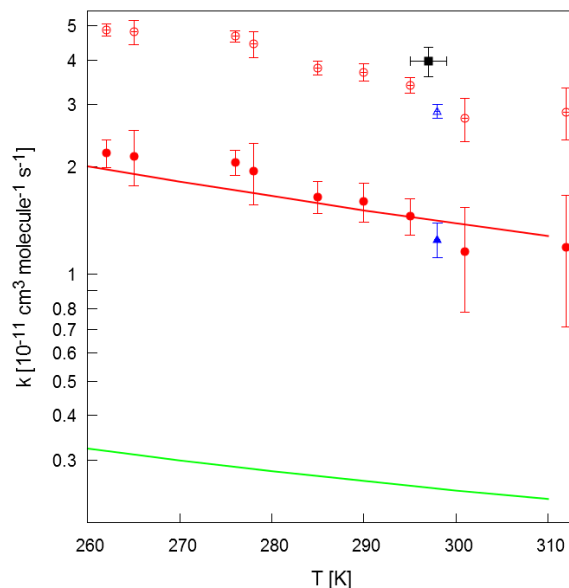


Figure 2: Calculated rate coefficients for the terminal addition of OH on 1-pentene compared to literature data of Nip and Paraskevopoulos¹¹ (black), Biermann et al.¹² (blue) and McGillen et al.¹³ (red). The empty symbols are the total rate coefficients as reported in literature, the full symbols are our estimate for the terminal addition reaction. The green line is the calculate terminal addition and the red line is the calculated terminal addition for which the barrier was decreased by 4.18 kJ mol⁻¹.

Similarly to other studies, the addition rate coefficient is underpredicted, largely because the inner barrier for addition is too high. Decreasing it by 4.18 kJ mol⁻¹, a change similar to what was used in previous studies,^{16,17} brings theory and experiment in almost perfect accord.

2.4.3 Rate coefficients

The OH addition to a multiple bond and its two-transition-state characteristics have been determined for several hydrocarbon molecules. Greenwald et al.²⁵ studied the OH addition to ethene in the 10–600 K range. Their rate coefficients $k(T,P)$ compare well with experimental data. The full C₂H₅O PES of the OH addition to ethene has been further explored by Senosiain et al.,¹⁶ consisting of three wells and four bimolecular products. The two addition reactions of OH to propene have been studied by Zádor et al.¹⁷ The C₃H₅O PES has also been studied by Zádor and Miller⁴⁹ in which OH adds to either propyne or propadiene, and there are further high-level kinetics calculations in the literature on the hydroxyl additions to 1-butene,²⁰ 2-butene,²¹ and isoprene.⁵⁰

The above theoretically calculated rate coefficients have all been obtained at a very similar level of theory, including high level of theory for the electronic structure calculations. Furthermore, they all solved the master equation to obtain temperature and pressure dependent rate coefficients. This makes a comparison of these studies appropriate. The rate coefficients for hydroxyl addition to alkenes are shown in Figure 3 for ethene, propene, cis- and trans-2-butene, isoprene and 1-pentene. All the rate coefficients correspond to the addition of OH to one carbon atom. The propene study only includes the total addition rate coefficient, i.e. both to the terminal as well as central carbon atom.¹⁷ The authors report a 50:50 branching fractions for both additions, the total rate coefficient has thus been divided by two to only include the addition to one well.

Figure 4 shows the total dissociation rate coefficient of the initial adduct, i.e. an alcohol with a radical site in β position of the hydroxyl group. This dissociation contains C-H, C-C and C-O β -scission reactions. The branching fractions of the latter, which results back in the initial OH and alkene, are reported in Figure 5. As already reported in the literature, the dissociation to the OH and alkene is largely favored at low temperature, with branching fractions of 1 at 300K and below.

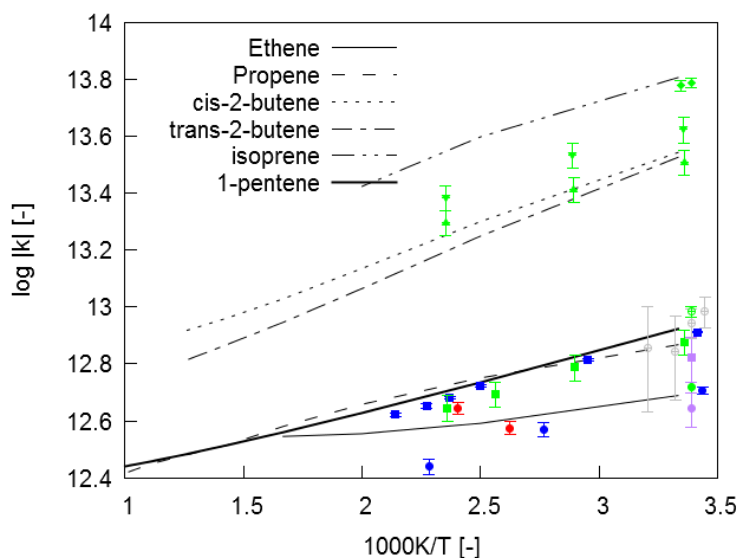


Figure 3: Theoretical OH addition rate coefficients to several alkenes in their high pressure limit calculated by Greenwald et al.²⁵ for ethene, Zádor et al.¹⁷ for propene, Antonov et al.²¹ for 2-butene, Greenwald et al.⁵⁰ for isoprene, and this study for 1-pentene. The symbols correspond to experimental data of ethene (○), propene (●), pentene (○), trans-2-butene(▲), cis-2-butene (▼), and isoprene (◆). The data originates from Tully et al.^{51,52} (blue), Gordon and Mulac⁵³ (red), Schmidt et al.⁵⁴ (purple), Atkinson and coworkers^{55,56} (green), and McGillen et al.¹³ (grey). The rate coefficients k are expressed in $\text{cm}^3 \text{mol}^{-1} \text{s}^{-1}$.

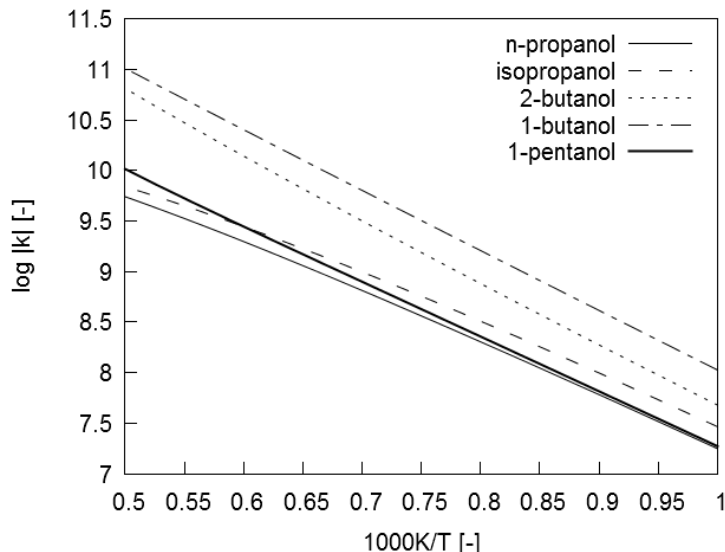


Figure 4: Total theoretical dissociation rate coefficient of alcohols with a radical site in β position relative to the hydroxyl group in their high pressure limit. In the case of 2-butanol, the secondary radical is considered, i.e. formed by the OH addition to 2-butene. The rate coefficients originate from Zádor et al.¹⁷ for n-propanol and isopropanol, Antonov et al.²¹ for 2-butanol, Zhang et al.²⁰ for 1-butanol, and this study for 1-pentanol. The rate coefficients are expressed in s^{-1} .

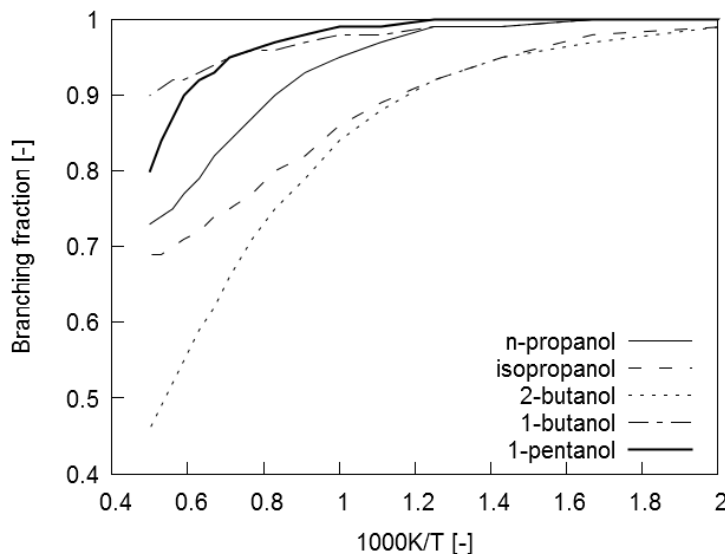


Figure 5: Branching fraction of the dissociation to OH and alkene to the total dissociation rates (all in their high pressure limit). In the case of 2-butanol, the secondary radical is considered, i.e. formed by the OH addition to 2-butene. The rate coefficients originate from Zádor et al.¹⁷ for n-propanol and isopropanol, Antonov et al.²¹ for 2-butanol, Zhang et al.²⁰ for 1-butanol, and this study for 1-pentanol.

Besides the comparison to similar reactions on other PESs, the rate coefficients can also be compared to the current group additivity rate coefficients calculated by Genesys. As long as the new calculations are not incorporated in Genesys, the group additivity values will be used instead. These rate coefficients, although they are valuable to obtain a large amount of data very fast, show relatively high uncertainties compared to *ab initio* data. The values used for this comparison are either in-house or can be found in literature.^{37,40} All the group additive values

have been obtained from CBSQB3 calculations, with an overall uncertainty of 2. Group additivity itself introduces an additional uncertainty factor of 3^{38} , yielding a total uncertainty factor of 3.66. To calculate this, the differences between the logarithms of the rate coefficients is used instead of their ratio. It can be assumed that the probability distribution of the logarithm of the *ab initio* rate coefficient is a normal distribution with the logarithm of the true rate coefficient as expected value and a standard deviation of $\log 2$, c.f. Eq. 2.1. Similarly, the probability distribution of the logarithm of the group additivity rate coefficient has the logarithm of the *ab initio* rate coefficient as expected value and a standard deviation of $\log 3$ as shown in Eq. 2.2. Assuming that the uncertainty of the group additivity model is independent of the uncertainty of the *ab initio* rate coefficients, the probability distribution of the logarithm of the group additivity rate coefficient compared to the logarithm of the true rate coefficient is also a normal distribution with a standard deviation of 3.66, calculated in Eq. 2.3 to Eq. 2.6. In the equations Eq. 2.1 to Eq. 2.6, the subscript “AI” stands for *ab initio*, “GA” stands for group additivity, the rate coefficient without subscript is the true rate coefficient, $P(x)$ is the probability distribution of x , $E(x)$ is the expected value of x , and $\text{Var}(x)$ is the variance.

$$P(\log k_{AI}) = \frac{1}{\sqrt{2\pi(\log 2)^2}} e^{-\frac{(\log k_{AI} - \log k)^2}{2(\log 2)^2}} \quad \text{Eq. 2.1}$$

$$P(\log k_{GA}) = \frac{1}{\sqrt{2\pi(\log 3)^2}} e^{-\frac{(\log k_{GA} - \log k_{AI})^2}{2(\log 3)^2}} \quad \text{Eq. 2.2}$$

$$\text{Var}(\log k_{AI}) = (\log 2)^2 = E[(\log k_{AI})^2] - (\log k)^2 \quad \text{Eq. 2.3}$$

$$\text{Var}(\log k_{GA})_{AI} = (\log 3)^2 = E[(\log k_{GA})^2] - (\log k_{AI})^2 \quad \text{Eq. 2.4}$$

$$\text{Var}(\log k_{GA})_{tot} = E[(\log k_{GA})^2] - (\log k)^2 = \sigma_{AI}^2 + \sigma_{GA}^2 \quad \text{Eq. 2.5}$$

$$\text{Var}(\log k_{GA})_{tot} = (\log 2)^2 + (\log 3)^2 = (\log 3.66)^2 \quad \text{Eq. 2.6}$$

Table 1 summarizes 8 reactions for which the rate coefficients were both calculated in this work as well as with Genesys. The second column gives the high pressure limit rate coefficients at 1000 K calculated in this study. The third column shows the ratio of the rate coefficients in this work compared to the Genesys calculated rate coefficients, both evaluate at 1000 K in the high pressure limit. In general, the accuracy of the Genesys group additivity rate coefficients is good, only three rate coefficients show a deviation factor higher than 2, and only one has a deviation

factor higher than 4. The Genesys rate coefficients are thus well within the expected uncertainty boundaries, but their values can be corrected when using high level *ab initio* calculations, which should result in a more accurate final kinetic model.

Table 1: High pressure limit rate coefficients of 8 reactions calculated in this work at 1000 K and the ratio of the rate coefficients of this work compared to the Genesys group additivity rate coefficients.

Reaction	k_{theory} ($cm^3 mol^{-1} s^{-1}$)	$\frac{k_{theory}}{k_{estimate}}$ (-)
$CH_3CH_2CH=CH_2 + CH_2\bullet OH \rightarrow \gamma\text{-R}$	6.0×10^9	1.07
$CH_3CH_2CH_2\bullet + CH_2=CHOH \rightarrow \alpha\text{-R}$	3.1×10^9	0.52
$CH_2=CHCH_2CH_2CH_2OH + H\bullet \rightarrow \varepsilon\text{-R}$	4.3×10^{12}	0.99
$CH_3\bullet + CH_2=CHCH_2CH_2OH \rightarrow \gamma\text{-R}$	1.1×10^{10}	1.14
$CH_2=CH_2 + CH_2\bullet CH_2CH_2OH \rightarrow \varepsilon\text{-R}$	1.5×10^{10}	0.37
$CH_3CH=CH_2 + CH_2\bullet CH_2OH \rightarrow \delta\text{-R}$	8.8×10^9	0.19
$CH_2=CHCH_2CH_2CH_2OH + H\bullet \rightarrow \delta\text{-R}$	5.8×10^{12}	1.39
$CH_3CH_2\bullet + CH_2=CHCH_2OH \rightarrow \beta\text{-R}$	3.9×10^9	0.42

2.4.4 Modelling

The experimental and simulation results are given in Figure 6 and Figure 7 for the major and minor compounds respectively. Simulations have also been done with three models found in the literature by Heufer et al.⁴⁸, Togbé et al.⁵⁷ and Wang et al.⁵⁸. Compared to the other models, the current work gives the closest agreement of the conversion to the experimental data. The predictions of the three main products, ethene, methane and CO, as a function of conversion is comparable to the experiments for the four models. Propene and ethane both exhibit a maximum in their mass fraction profile at a conversion of 90%. The simulations are able to capture the maximum, but it lies at a higher conversion. The model of Heufer et al. leads to the best propene simulations, although the differences are small between the models. For ethane, Togbé et al. shows a better agreement to the experimental data, the models of Heufer et al. and Wang et al. are not able to capture the trend of the selectivity towards ethane as a function of the conversion. The water yields are similar for the four models, and the current work gives the best butadiene predictions, although secondary reaction pathways of butadiene are probably missing, which explains the overprediction only at high temperatures. Although the relative deviations of the

simulations on these compounds are higher compared to the major products, the absolute deviations are of a similar order of magnitude.

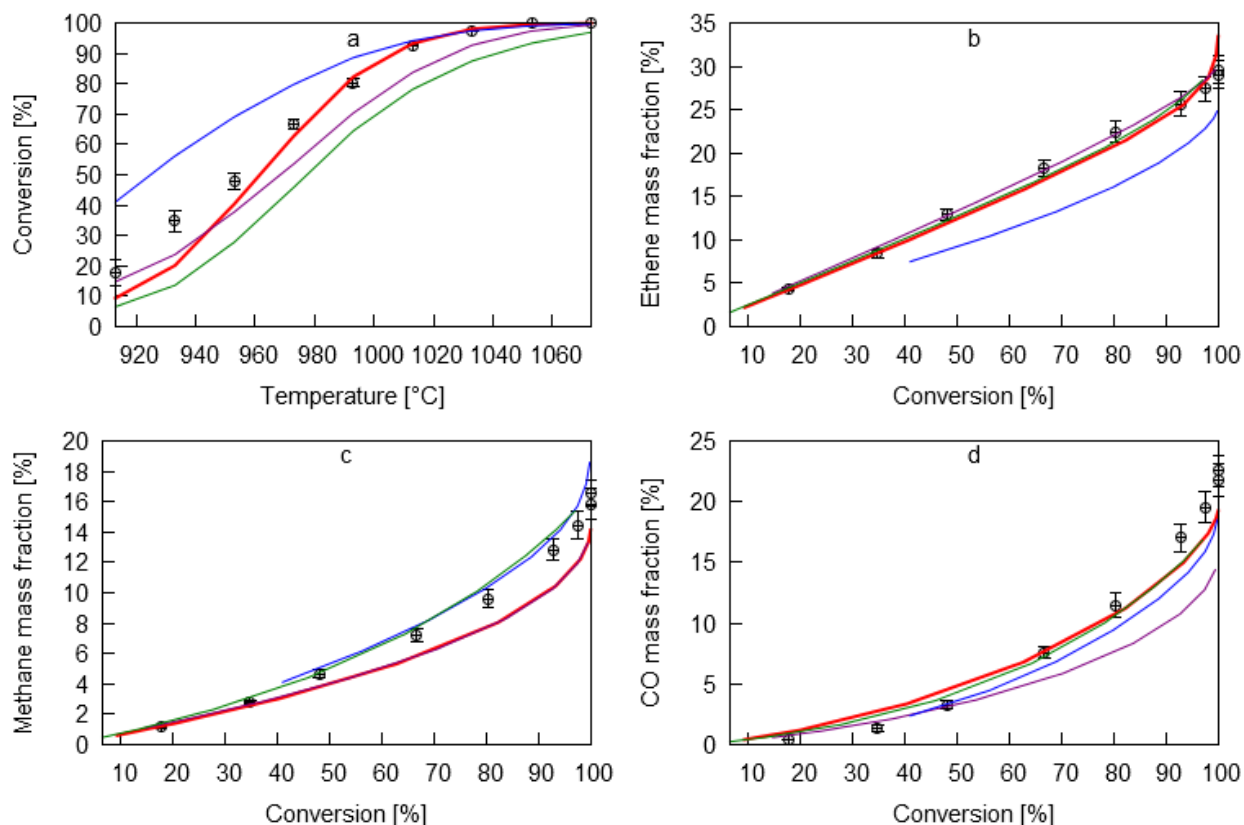


Figure 6: Experimental results (dots) and simulations (lines) for (a) the conversion of pentanol as a function of temperature, (b) the mass fraction of ethene as function of the conversion, (c) the mass fraction of methane as a function of conversion and (c) the mass fraction of CO as a function of conversion. The model generated by Genesys (red) is compared to the models of Heufer et al.⁴⁸ (purple), Togbé et al.⁵⁷ (blue), and Wang et al.⁵⁸ (green).

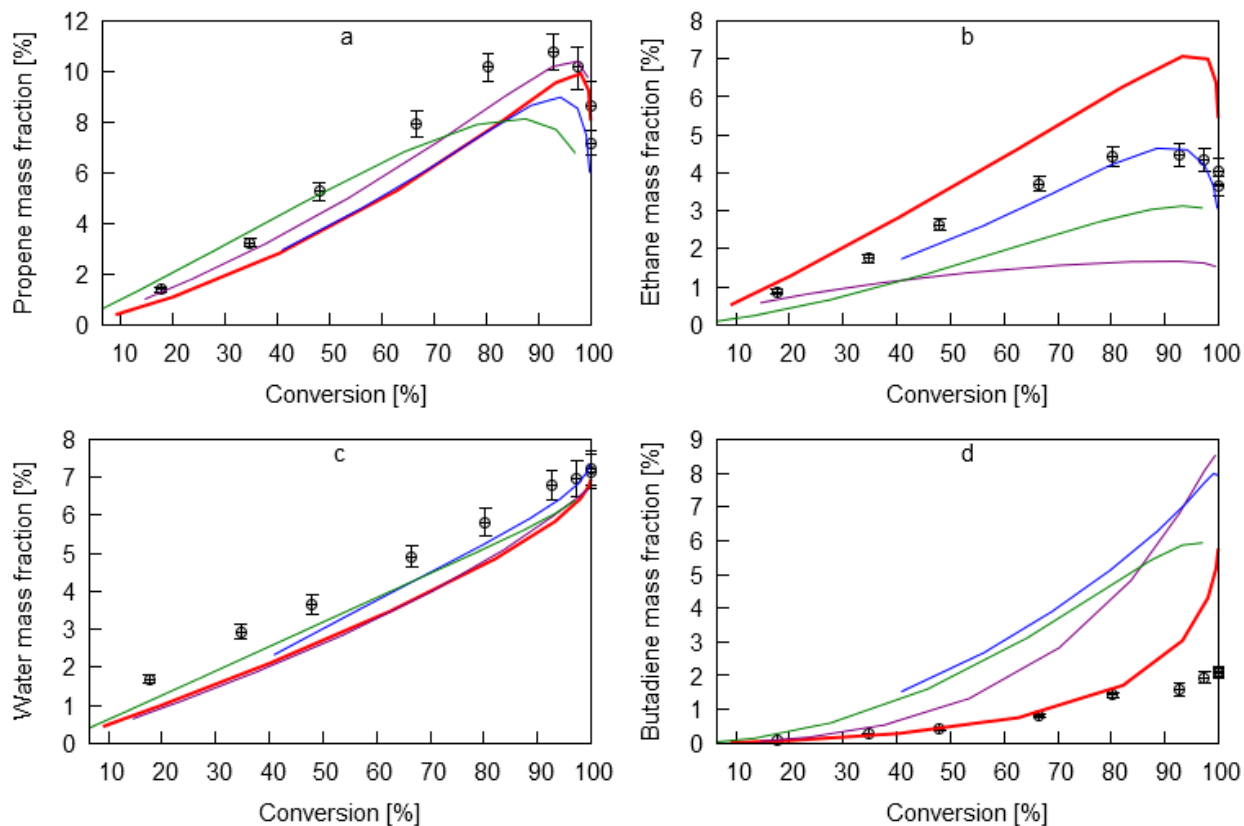


Figure 7: Experimental results (dots) and simulations (lines) for (a) the mass fraction of propylene, (b) the mass fraction of ethane, (c) the mass fraction of water and (c) the mass fraction of butadiene as a function of conversion. The model generated by Genesys (red) is compared to the models of Heufer et al.⁴⁸ (purple), Togbé et al.⁵⁷ (blue), and Wang et al.⁵⁸ (green).

From the reactor simulations, the main decomposition pathways can be identified and their relative importance can be quantified. The initial decomposition of the n-pentanol radical happens through either a C-C bond cleavage or the direct elimination of a water molecule, forming pentene. Homolytic scission of C-O, C-H and O-H bonds do not play an important role. The weakest bond is, in agreement to studies of pentanol^{48,59} or other alcohols^{60,61}, the bond in β position of the hydroxyl group, followed by the terminal C-C bond. Once a radical pool is created, these scission reactions decrease very rapidly in importance, and the main decomposition route is hydrogen abstractions, shown in Figure 8. The reaction scheme is recorded at the highest n-pentanol decomposition rate, which is at a distance of 37 cm into the reactor for the simulation at 973 K. At 37 cm, the set-point temperature is already reached. The percentages show the relative rate of production to the total decomposition rate of n-pentanol. For the sake of clarity, no intramolecular hydrogen abstractions and well-skipping reactions have been added to the figure. Two intramolecular hydrogen abstraction reactions have a contribution of more than 1% compared to the total conversion rate of n-pentanol: the net conversion of δ -R to o-R amounts

2%, and the total net conversion of ϵ -R to α -R is around 1%. Also, one well-skipping reaction has a noteworthy contribution: o-R reacts to propene and the $\text{CH}_2\bullet\text{CH}_2\text{OH}$ with a production rate of 3%.

More or less half of the o-R first undergoes a C-C- β -scission to formaldehyde and n-butyl. The remainder of the o-R decomposes through the well-skipping reaction mentioned above. The α -R, which is the lowest lying well on the OH+1-pentene PES, see Figure 1, and is also the main radical formed through hydrogen abstractions from n-pentanol, decomposes to n-propyl and vinyl alcohol. The β -R reacts via a C-O- β -scission, forming 1-pentene and a hydroxyl radical. The C-C- β -scission is also somewhat important, leading to ethyl and 1-buten-3-ol. Two C-C- β -scission are possible in γ -R, forming methyl and 1-buten-4-ol, and 1-butene and $\text{C}\bullet\text{H}_2\text{OH}$, respectively. As already mentioned, δ -R is consumed by an intramolecular hydrogen abstraction reaction to o-R. Furthermore, it decomposes through two pathways: the formation of propene and $\text{C}\bullet\text{H}_2\text{CH}_2\text{OH}$ and a well-skipping reaction to formaldehyde and n-butyl, accountable for almost 10% of the total decomposition of n-propanol.

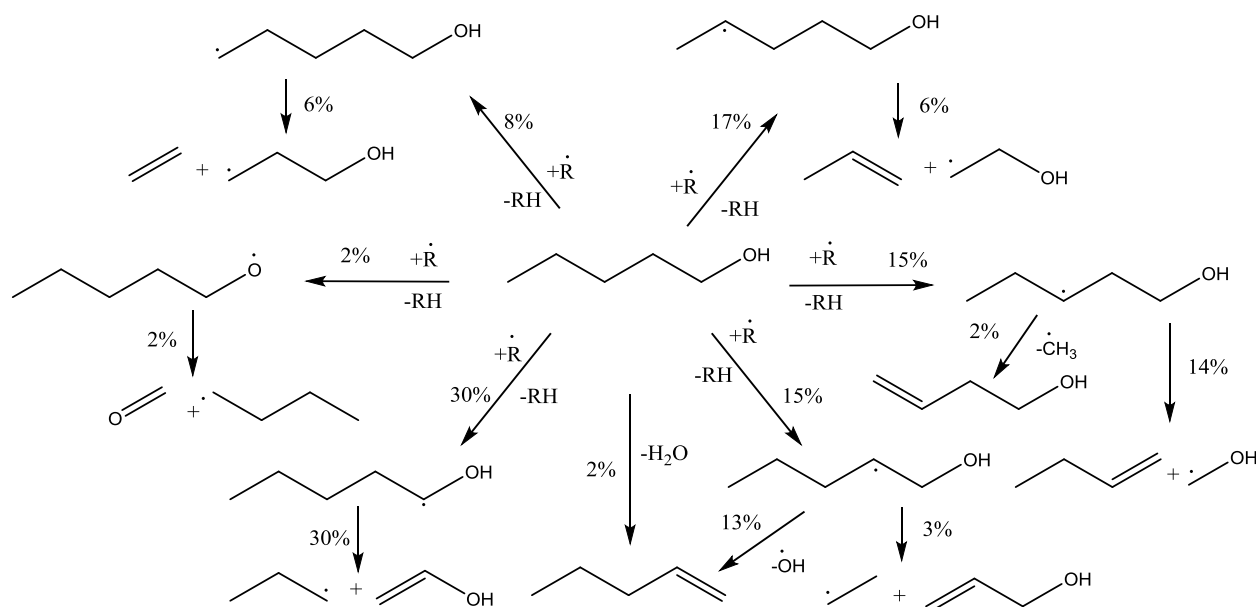


Figure 8: Initial decomposition of n-pentanol at a position of 37 cm in the Bench Scale set-up at 973 K, which corresponds to a conversion of 38%. The percentages are the relative rates to the total rate of consumption of n-pentanol.

2.5 Conclusions

The decomposition of n-pentanol has been studied in several steps. First, all the stationary points on the n-pentanol radical potential energy surface have been automatically found using the KinBot software. The wells and saddle points were optimized at the M062X/6-311++G(d,p) level of theory with additional single point energy calculations at the UCCSD(T)-F12/cc-pVTZ-F12 level of theory. The OH + 1-pentene channel was treated with a two-transition state model to account for both the inner saddle points as well as the barrierless outer transition state. Master equation calculations were used to get the pressure and temperature dependent rate coefficients.

The resulting potential energy surface is comparable to other alcohol pyrolysis potential energy surfaces, which typically exhibit the van der Waals well when adding OH to and olefin giving the β -radical. Intramolecular hydrogen abstractions are the only isomerization reactions, and exit channels exist of β -scissions and water eliminations. The rate coefficients from this potential energy surface compare well to experimental data and theoretical data of similar systems.

Second, the decomposition of n-pentanol has been studied in a flow reactor at various temperatures covering a wide conversion range. A kinetic model developed by Genesys and augmented with the newly calculated rate coefficients is able to describe the conversion and selectivities of the main products well.

2.6 References

1. Boundy B, Diegel SW, Wright L, Davis SC. *Biomass energy data book*. 4th ed. ed. Oak Ridge National Laboratory 2011.
2. Rajesh Kumar B, Saravanan S, Rana D, Nagendran A. Combined effect of injection timing and exhaust gas recirculation (EGR) on performance and emissions of a DI diesel engine fuelled with next-generation advanced biofuel – diesel blends using response surface methodology. *Energy Conversion and Management*. 2016;123:470-486.
3. Wei L, Cheung CS, Huang Z. Effect of n-pentanol addition on the combustion, performance and emission characteristics of a direct-injection diesel engine. *Energy*. 2014;70:172-180.
4. Zhang Z-H, Chua S-M, Balasubramanian R. Comparative evaluation of the effect of butanol–diesel and pentanol–diesel blends on carbonaceous particulate composition and particle number emissions from a diesel engine. *Fuel*. 2016;176:40-47.
5. Li L, Wang J, Wang Z, Xiao J. Combustion and emission characteristics of diesel engine fueled with diesel/biodiesel/pentanol fuel blends. *Fuel*. 2015;156:211-218.
6. Imdadul HK, Masjuki HH, Kalam MA, Zulkifli NWM, Alabdulkarem A, Rashed MM, Ashraful AM. Influences of ignition improver additive on ternary (diesel-biodiesel-higher alcohol) blends thermal stability and diesel engine performance. *Energy Conversion and Management*. 2016;123:252-264.
7. Nativel D, Pelucchi M, Frassoldati A, Comandini A, Cuoci A, Ranzi E, Chaumeix N, Faravelli T. Laminar flame speeds of pentanol isomers: An experimental and modeling study. *Combustion and Flame*. 2016;166:1-18.
8. Köhler M, Kathrotia T, Oßwald P, Fischer-Tammer ML, Moshhammer K, Riedel U. 1-, 2- and 3-Pentanol combustion in laminar hydrogen flames – A comparative experimental and modeling study. *Combustion and Flame*. 2015;162(9):3197-3209.
9. Heufer KA, Bugler J, Curran HJ. A comparison of longer alkane and alcohol ignition including new experimental results for n-pentanol and n-hexanol. *Proceedings of the Combustion Institute*. 2013;34(1):511-518.
10. Zádor J, Najm HN. KinBot 1.0: A code for automatic PES exploration. Paper presented at: 8th U. S. National Combustion Meeting 2013.
11. Nip WS, Paraskevopoulos G. Rates of OH radical reactions. VI. Reactions with C₃H₆, 1-C₄H₈ and 1-C₅H₁₀ at 297 K. *Journal of Chemical Physics*. 1979;71(5):2170-2174.
12. Biermann HW, Harris GW, Pitts JN. Photoionization mass spectrometer studies of the collisionally stabilized product distribution in the reaction of hydroxyl radicals with selected alkenes at 298 K. *Journal of Physical Chemistry*. 1982;86(15):2958-2964.
13. McGillen MR, Percival CJ, Shallcross DE, Harvey JN. Is hydrogen abstraction an important pathway in the reaction of alkenes with the OH radical? *Physical Chemistry Chemical Physics*. 2007;9(31):4349-4356.
14. Heimann G, Benkelberg H-J, Böge O, Warneck P. Photodecomposition of iodopentanes in air: Product distributions from the self-reactions of n-pentyl peroxy radicals. *International Journal of Chemical Kinetics*. 2002;34(2):126-138.
15. Hein H, Hoffmann A, Zellner R. Direct investigations of reactions of 1-butoxy and 1-pentoxy radicals using laser pulse initiated oxidation: reaction with O₂ and isomerisation at 293 K and 50 mbar. *Physical Chemistry Chemical Physics*. 1999;1(16):3743-3752.

16. Senosiain JP, Klippenstein SJ, Miller JA. Reaction of ethylene with hydroxyl radicals: A theoretical study. *Journal of Physical Chemistry A*. 2006;110(21):6960-6970.
17. Zádor J, Jasper AW, Miller JA. The reaction between propene and hydroxyl. *Physical Chemistry Chemical Physics*. 2009;11(46):11040-11053.
18. Kappler C, Zádor J, Welz O, Fernandes RX, Olzmann M, Taatjes CA. Competing channels in the propene + OH reaction: Experiment and validated modeling over a broad temperature and pressure range. *Zeitschrift für Physikalische Chemie*. 2011;225:1271-1293.
19. Zádor J, Miller JA. Unimolecular dissociation of hydroxypropyl and propoxy radicals. *Proceedings of the Combustion Institute*. 2013;34:519-526.
20. Zhang P, Klippenstein SJ, Law CK. Ab Initio Kinetics for the Decomposition of Hydroxybutyl and Butoxy Radicals of n-Butanol. *Journal of Physical Chemistry A*. 2013;117(9):1890-1906.
21. Antonov IO, Kwok J, Zador J, Sheps L. A Combined Experimental and Theoretical Study of the Reaction OH+2-Butene in the 400-800 K Temperature Range. *Journal of Physical Chemistry A*. 2015;119(28):7742-7752.
22. *Gaussian 09, Revision A.02* [computer program]. Wallingford CT: Gaussian, Inc.; 2009.
23. *MOLPRO, version 2012.1, a package of ab initio programs* [computer program] 2012.
24. Miller WH. Unified statistical model for "complex" and "direct" reaction mechanisms. *The Journal of Chemical Physics*. 1976;65(6):2216-2223.
25. Greenwald EE, North SW, Georgievskii Y, Klippenstein SJ. A Two Transition State Model for Radical-Molecule Reactions: A Case Study of the Addition of OH to C₂H₄. *The Journal of Physical Chemistry A*. 2005;109(27):6031-6044.
26. Georgievskii Y, Klippenstein SJ. Transition State Theory for Multichannel Addition Reactions: Multifaceted Dividing Surfaces *J. Phys. Chem. A*. 2003;107(46):9776-9781.
27. Georgievskii Y, Klippenstein SJ. Variable reaction coordinate transition state theory: Analytic results and application to the C₂H₃ + H → C₂H₄ reaction. *Journal of Chemical Physics*. 2003;118(12):5442-5455.
28. Antonov IO, Kwok J, Zádor J, Sheps L. OH + 2-butene: A combined experimental and theoretical study in the 300-800 K temperature range. *Journal of Physical Chemistry A*. 2015;119:7742-7752.
29. *MESS.2016.3.23*. [computer program].
30. Georgievskii Y, Miller JA, Burke MP, Klippenstein SJ. Reformulation and solution of the master equation for multiple-well chemical reactions. *Journal of Physical Chemistry A*. 2013;117(46):12146-12154.
31. Pollak E, Pechukas P. Symmetry numbers, not statistical factors, should be used in absolute rate theory and in bronsted relations. *Journal of the American Chemical Society*. 1978;100(10):2984-2991.
32. Djokic M, Carstensen HH, Van Geem KM, Marin GB. The thermal decomposition of 2,5-dimethylfuran. *Proceedings of the Combustion Institute*. 2013;34:251-258.
33. Harper MR, Van Geem KM, Pyl SP, Marin GB, Green WH. Comprehensive reaction mechanism for n-butanol pyrolysis and combustion. *Combustion and Flame*. 2011;158(1):16-41.
34. Beens J, Boelens H, Tijssen R, Blomberg J. Quantitative aspects of comprehensive two-dimensional gas chromatography (GC x GC). *Hrc-Journal of High Resolution Chromatography*. 1998;21(1):47-54.

35. *CHEMKIN Release 4.1.1* [computer program]. San Diego, CA, USA: Reaction Design, Inc.; 2007.
36. Vandewiele NM, Van Geem KM, Reyniers MF, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chemical Engineering Journal*. 2012;207:526-538.
37. Sabbe MK, Reyniers MF, Van Speybroeck V, Waroquier M, Marin GB. Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors. *Chemphyschem*. 2008;9(1):124-140.
38. Sabbe MK, Vandeputte AG, Reyniers MF, Waroquier M, Marin GB. Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions. *Physical Chemistry Chemical Physics*. 2010;12(6):1278-1298.
39. Vandeputte AG, Sabbe MK, Reyniers M-F, Van Speybroeck V, Waroquier M, Marin GB. Theoretical study of the thermodynamics and kinetics of hydrogen abstractions from hydrocarbons. *Journal of Physical Chemistry A*. 2007;111(46):11771-11786.
40. Paraskevas PD, Sabbe MK, Reyniers MF, Marin GB, Papayannakos NG. Group additive kinetic modeling for carbon-centered radical addition to oxygenates and -scission of oxygenates. *Aiche Journal*. 2016;62(3):802-814.
41. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos N, Marin GB. Kinetic Modeling of alpha-Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Carbon-Centered Radicals. *Chemphyschem*. 2014;15(9):1849-1866.
42. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Group Additive Kinetics for Hydrogen Transfer Between Oxygenates. *Journal of Physical Chemistry A*. 2015;119(27):6961-6980.
43. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Kinetic Modeling of a-Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Hydrogen Atoms. *Journal of Physical Chemistry A*. 2014;118(40):9296-9309.
44. Paraskevas PD, Sabbe MK, Reyniers M-F, Papayannakos N, Marin GB. Group Additive Values for the Gas-Phase Standard Enthalpy of Formation, Entropy and Heat Capacity of Oxygenates. *Chemistry – A European Journal*. 2013;19(48):16431-16452.
45. Sabbe MK, De Vleeschouwer F, Reyniers MF, Waroquier M, Marin GB. First Principles Based Group Additive Values for the Gas Phase Standard Entropy and Heat Capacity of Hydrocarbons and Hydrocarbon Radicals. *Journal of Physical Chemistry A*. 2008;112(47):12235-12251.
46. Sabbe MK, Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *Journal of Physical Chemistry A*. 2005;109(33):7466-7480.
47. Metcalfe WK, Burke SM, Ahmed SS, Curran HJ. A Hierarchical and Comparative Kinetic Modeling Study of C-1 - C-2 Hydrocarbon and Oxygenated Fuels. *International Journal of Chemical Kinetics*. 2013;45(10):638-675.
48. Heufer KA, Sarathy SM, Curran HJ, Davis AC, Westbrook CK, Pitz WJ. Detailed Kinetic Modeling Study of n-Pentanol Oxidation. *Energy & Fuels*. 2012;26(11):6678-6685.
49. Zádor J, Miller JA. Adventures on the C₃H₅O potential energy surface: OH + propyne, OH + allene and related reactions. *Proceedings of the Combustion Institute*. 2015;35(1):181-188.
50. Greenwald EE, North SW, Georgievskii Y, Klippenstein SJ. A Two Transition State Model for Radical-Molecule Reactions: Applications to Isomeric Branching in the OH-Isoprene Reaction. *The Journal of Physical Chemistry A*. 2007;111(25):5582-5592.

51. Tully FP. Laser photolysis/laser-induced fluorescence study of the reaction of hydroxyl radical with ethylene. *Chemical Physics Letters*. 1983;96(2):148-153.
52. Tully FP, Goldsmith JEM. Kinetic study of the hydroxyl radical-propene reaction. *Chemical Physics Letters*. 1985;116(4):345-352.
53. Gordon S, Mulac WA. Reaction of the OH(X²-PI) radical produced by the pulse radiolysis of water vapor. *International Journal of Chemical Kinetics*. 1975;1:289-299.
54. Schmidt V, Zhu GY, Becker KH, Fink EH. Study of OH Reactions at High Pressures by Excimer Laser Photolysis - Dye Laser Fluorescence. *Berichte Der Bunsen-Gesellschaft-Physical Chemistry Chemical Physics*. 1985;89(3):321-322.
55. Atkinson R, Pitts JN. Rate Constants for the Reaction of OH Radicals with Propylene and the Butenes over the Temperature Range 297-425°K. *Journal of Chemical Physics*. 1975;63(8):3591-3595.
56. Atkinson R, Aschmann SM. Rate constants for the reaction of OH radicals with a series of alkenes and dialkenes at 295 ± 1 K. *International Journal of Chemical Kinetics*. 1984;16(10):1175-1186.
57. Togbé C, Halter F, Foucher F, Mounaim-Rousselle C, Dagaut P. Experimental and detailed kinetic modeling study of 1-pentanol oxidation in a JSR and combustion in a bomb. *Proceedings of the Combustion Institute*. 2011;33(1):367-374.
58. Wang G, Yuan W, Li Y, Zhao L, Qi F. Experimental and kinetic modeling study of n-pentanol pyrolysis and combustion. *Combustion and Flame*. 2015;162(9):3277-3287.
59. Zhao L, Ye L, Zhang F, Zhang L. Thermal Decomposition of 1-Pentanol and Its Isomers: A Theoretical Study. *The Journal of Physical Chemistry A*. 2012;116(37):9238-9244.
60. Cai J, Zhang L, Zhang F, Wang Z, Cheng Z, Yuan W, Qi F. Experimental and Kinetic Modeling Study of n-Butanol Pyrolysis and Combustion. *Energy & Fuels*. 2012;26(9):5550-5568.
61. Frassoldati A, Cuoci A, Faravelli T, Niemann U, Ranzi E, Seiser R, Seshadri K. An experimental and kinetic modeling study of n-propanol and iso-propanol combustion. *Combustion and Flame*. 2010;157(1):2-16.

Chapter 3: On-the-fly *ab initio* calculations

3.1 Abstract

During automatic kinetic model generation a large number of species and reactions are generated. All these species and reactions need to have thermodynamic and kinetic data assigned to them. Currently, databases and approximation methods are the main alternative to experimental and high level theoretical methods. In many cases, however, the approximation methods give poor results, leading to inaccurate kinetic models. The ever-growing computational power and availability of high performance computing solutions allows to envisage on-the-fly *ab initio* calculations for a large number of species and reactions. Currently, the need for extensive user knowledge and involvement is the main hurdle to develop a fully *ab initio* kinetic model. The present chapter introduces automation procedures to minimize the user involvement and to allow fully automated quantum chemical calculations for a large number of species and reactions for gas-phase processes, with a wide variety in chemical structure.

3.2 Introduction

Reliable procedures to calculate thermodynamic and kinetic data that cover a wide range of species and reactions are indispensable to automatically build a kinetic model. Although a kinetic model can contain thousands of reactions, most of the reactions can be classified in a limited set of reaction families and the calculation of rate coefficients can be done per reaction family. When data is lacking or when the calculation procedures do not yield satisfactory results, quantum chemical calculations can complement this data gap. In the framework of automatic kinetic model generation, these quantum chemical calculations include several reactions belonging to the same reaction family. Based on the *ab initio* results, calculation procedures such as linear free-energy relationships^{1,2} or group additivity³⁻⁸ methods can be developed. These methods are fast and scalable, and do not require any additional quantum chemical calculations. If developed properly these methods are more than adequate for current state-of-the-art in kinetic modeling. Since *ab initio* calculations are computationally intensive, require several manual interventions, user knowledge and last but not least expertise, the development of a reaction rate coefficients calculation methodology remains time consuming. Automation algorithms which can perform *ab initio* calculations and subsequently extract accurate thermochemical data would thus be of great interest for kinetic model generation. As mentioned in Chapter 1, species and reactions are represented as mathematical graphs containing the connectivity information between the atoms. This topology information, also called two dimensional (2D) representation of species, is insufficient to start *ab initio* calculations. Therefore, to automatically perform *ab initio* calculations, three dimensional (3D) coordinates of species and transition states need to be generated. This is then followed by calculations using quantum chemical software packages. Finally, thermodynamic data and reaction rate coefficients can be calculated.

Several approaches to incorporate quantum mechanical calculations in automatic kinetic model generation are possible. For example, as was discussed in the previous chapter, reactions on a potential energy surface (PES) can be found using the KinBot⁹ software. Although the results are useful for automatic kinetic model generation, i.e. the rate coefficients can be stored in the kinetic model generators databases for later use, the approach has several disadvantages when aiming at automated construction of calculation procedures for reaction rate coefficients. First, only a single PES can be calculated per simulation. A kinetic model typically contains hundreds of

species and up to thousands of reactions. The stationary points of these species and reactions lie on numerous potential energy surfaces, which have to be calculated separately. Second, reaction families are typically hard coded into the current tools, which means that extension of the chemical space requires the availability and knowledge of the source code. Third, the exploration of potential energy surfaces is very useful when the chemistry of a given molecule or radical is not well known, however, many reaction families and many rate coefficients are already available in databases. Potential energy surface searches identify reactions independent of their reaction family, i.e. they are not limited to a certain reaction families. When, for a chemical process, data is already available for a few reaction families, and data is necessary for the other families, the potential energy searches will find all the reaction families and thus re-calculate a lot of data. This increases the computational demands, calculation times and user involvement significantly.

When the kinetic model generation program does not include all possible reaction families for a molecule or radical, quantum chemistry can be used to search for new reaction types and new reactions which are not well documented yet. These calculations are computationally very expensive and cannot yet be coupled with automatic kinetic model generation in a straightforward way. Nevertheless, the methods used to discover new reaction types can be of interest for this work. They share similar challenges and they could be implemented in automatic kinetic model generation codes in the future. Suleimanov and Green¹⁰ published the discovery of chemical reaction steps using the double-ended freezing string method and the single-ended Berny optimization methods. Starting from the Bond Electron (BE) matrix, in which each row i contains the electronic configuration of the i^{th} atom, reactions are identified by multiplying the BE matrix with a reaction matrix A yielding another BE matrix, corresponding to the product structure. Via force field calculations, the reactant and product connectivity are converted to 3D coordinates. To locate the transition state, single-ended and double-ended searches can be used.¹¹ Single-ended searches start from an initial guess of the 3D coordinates of the atoms of the transition state and explore the PES using local gradients. An example is the Berny saddle point optimization method.¹²⁻¹⁴ Double-ended algorithms start from the reactant and product wells on the PES to locate the saddle point in between. Several algorithms have been developed for this purpose.¹⁵⁻¹⁷ Suleimanov and Green employ a double-ended search to locate a good initial guess for an optimization with a single-ended method. Zimmerman^{18,19} developed a similar approach as the BE matrices of Suleimanov and Green. Reactions are defined by the changes in coordination

number of atoms and in connectivity between the atoms. Force field calculations are employed prior to DFT optimizations. The latter allows screening for kinetically insignificant intermediates, which are too high in energy compared to the starting species. Double-ended searches are subsequently used to locate the transition state.

Instead of starting from one species and exhaustively identifying saddle points and wells on the PES, it is also possible to start from a known reaction or species and automatically perform *ab initio* calculations aiming at improving the thermodynamic and kinetic databases. This approach was already mentioned in literature by several authors. Broadbelt et al.²⁰ introduced algorithms to build 3D coordinates of species from the connectivity information. The geometrical structure was then used as input in a semi-empirical quantum chemistry program to obtain heats of formation of the species in a kinetic model. Magoon and Green²¹ extended this methodology by introducing more flexibility in terms of application range and also allowed the search for the lowest energy conformer. Semi-empirical methods are used to extract thermodynamic parameters. Broadbelt et al. and Magoon and Green only enabled the calculations of species, i.e. wells on the potential energy surface. It is, however, also possible to extend this approach during the calculation of reaction rate coefficients by optimizing transition state structures. Bhoorasingh and West²² developed a group additive scheme to estimate the structure of transition states of intermolecular hydrogen abstraction reactions which were then used as input for *ab initio* calculations. Intrinsic reaction coordinates calculations validate a successful transition state search. Finally, rate coefficients were extracted using transition state theory.

To solve the above mentioned shortcomings to automatically perform *ab initio* calculations in the framework of kinetic model generation, a new tool has been developed which is described in this chapter. The tool starts from a reaction network, i.e. a set of known species and reactions, and automatically performs quantum chemical calculations to obtain electronic energies and molecular frequencies. These results are used to generate thermodynamic and kinetic data, which are directly incorporated in kinetic models. The new tool is, in contrast to the work of Bhoorasingh and West²², not limited to one specific chemistry, reaction family or calculation method. With the applications of thermal decomposition, pyrolysis, combustion, oxidation, and atmospheric chemistry in mind, radical and molecular reactions in the gas-phase are covered, independent of the elements present in the reactions. Semi-Empirical, (post-)Hartree-Fock,

Density Functional Theory, or composite methods can be used, and new methods can easily be added, with only small changes in the source code. An overview of the methodology used in this work is given in Figure 1.

The present chapter introduces this tool, which is an extension to the existing kinetic model generator Genesys²³. Genesys is a recent kinetic model generator, based on open-source chemoinformatics, which allows an unrestricted model generation in terms of application range. Genesys is written in the programming language JAVA, which was also chosen as programming language for this work. This chapter first describes the methodology in which, starting from molecular identifiers, initial 3D structures are generated. These structures are refined in several steps and are used in extensive conformational scans to identify the lowest energy conformer. The final optimizations are preferably done at high level of theory. Torsional modes are treated with a one dimensional (1D) hindered rotor approach. Secondly, the calculation methods to obtain thermodynamic and kinetic data are described. Thirdly, results are presented for several species and reactions exhibiting a wide variety of features, such as hetero-atoms, ring structure, reaction families, radical centers, resonance, etc.

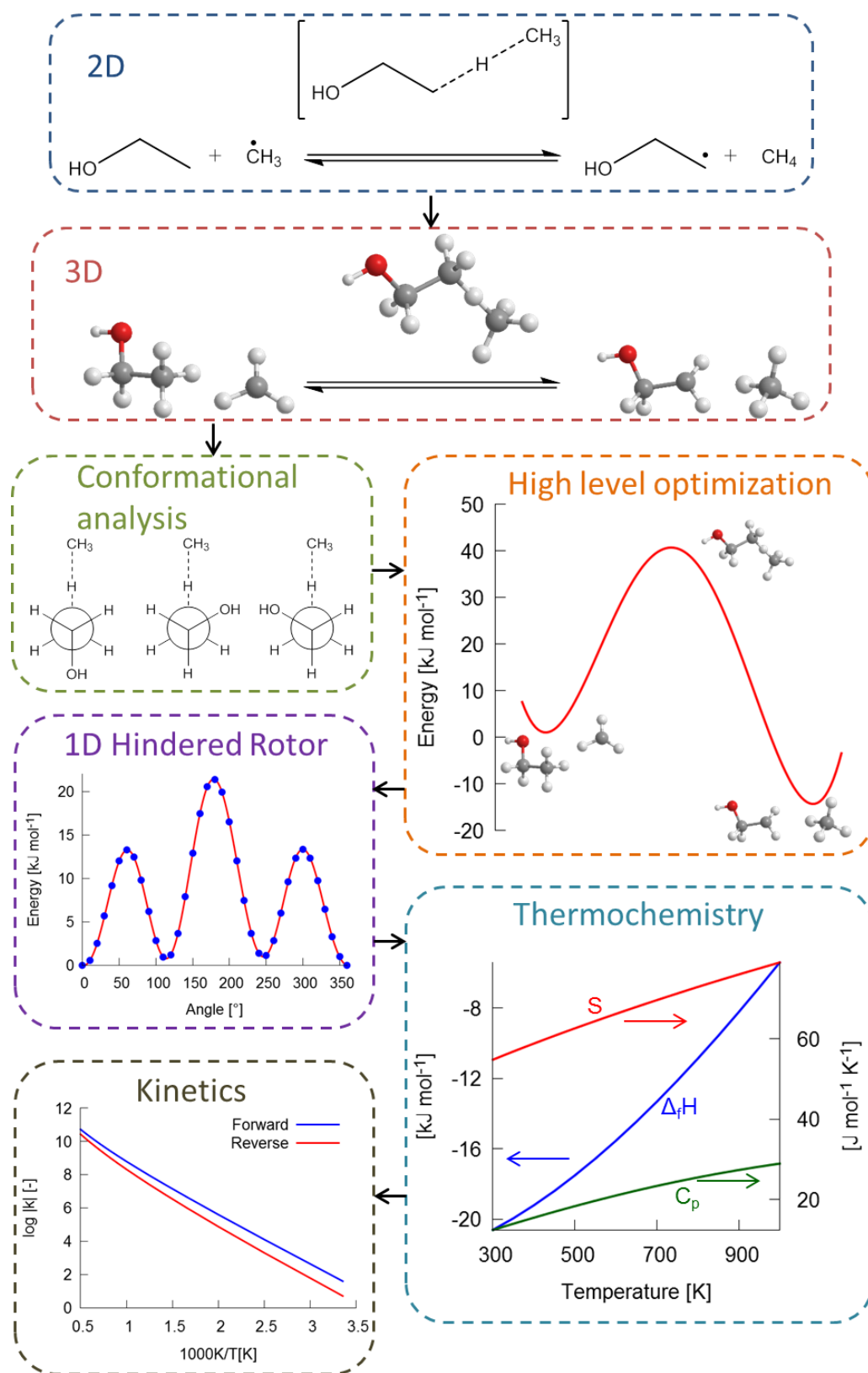


Figure 1: Illustration of the methodology of on-the-fly *ab initio* calculations.

3.3 Methodology

3.3.1 Stable species

3.3.1.1 3D coordinates

During automatic kinetic model generation, chemical species are typically represented using a matrix or mathematical graph. This 2D representation is very powerful for the generation of reactions, molecular comparison, recognizing sub-molecular patterns, user input and output, and database searches.²⁴ However, the main disadvantage is the loss of information on the spatial arrangements of the atoms within a molecule or radical. 3D coordinates are not available, which is an intrinsic limitation to start *ab initio* calculations. To allow automatic quantum chemical calculations, the 2D representation needs to be extended to a 3D one. The problem of generating a 3D structure for a molecule or radical has already been widely studied in the past for the purpose of drug discovery, retrosynthetic analysis, molecular depiction, etc. In general, algorithms to generate 3D coordinates can be categorized in two groups: rule-based and numerical methods.²⁰ Rule-based methods employ heuristics to obtain a molecular structure. The heuristics originate from the chemical knowledge of the software developer. Therefore, it is always tailored to one or a limited set of application domains. This restricts its use significantly. Numerical techniques²⁵⁻²⁹ such as “distance geometry” offer a broader application domain and robustness and typically consist of two steps. First, a “distance bounds matrix” is built, which is used to embed all the atoms in the three dimensional space. This structure is further refined using an energy minimization algorithm.

In order to create a general tool for 3D coordinates generation, the “distance geometry” method has been implemented in Genesys, which is illustrated in Figure 2. With “distance geometry” the author implies the generation of 3D structures based on a distance bounds matrix, which describes the lower and upper distance limits between two atoms. From these distance limits, an initial structure can directly be generated without any optimization of the coordinates. The distance bounds matrix can be built based on the topology of the molecule or radical, i.e. its 2D representation. However, the initial structure can be far away from the equilibrium, and optimization procedures to refine the initial guess are recommended.

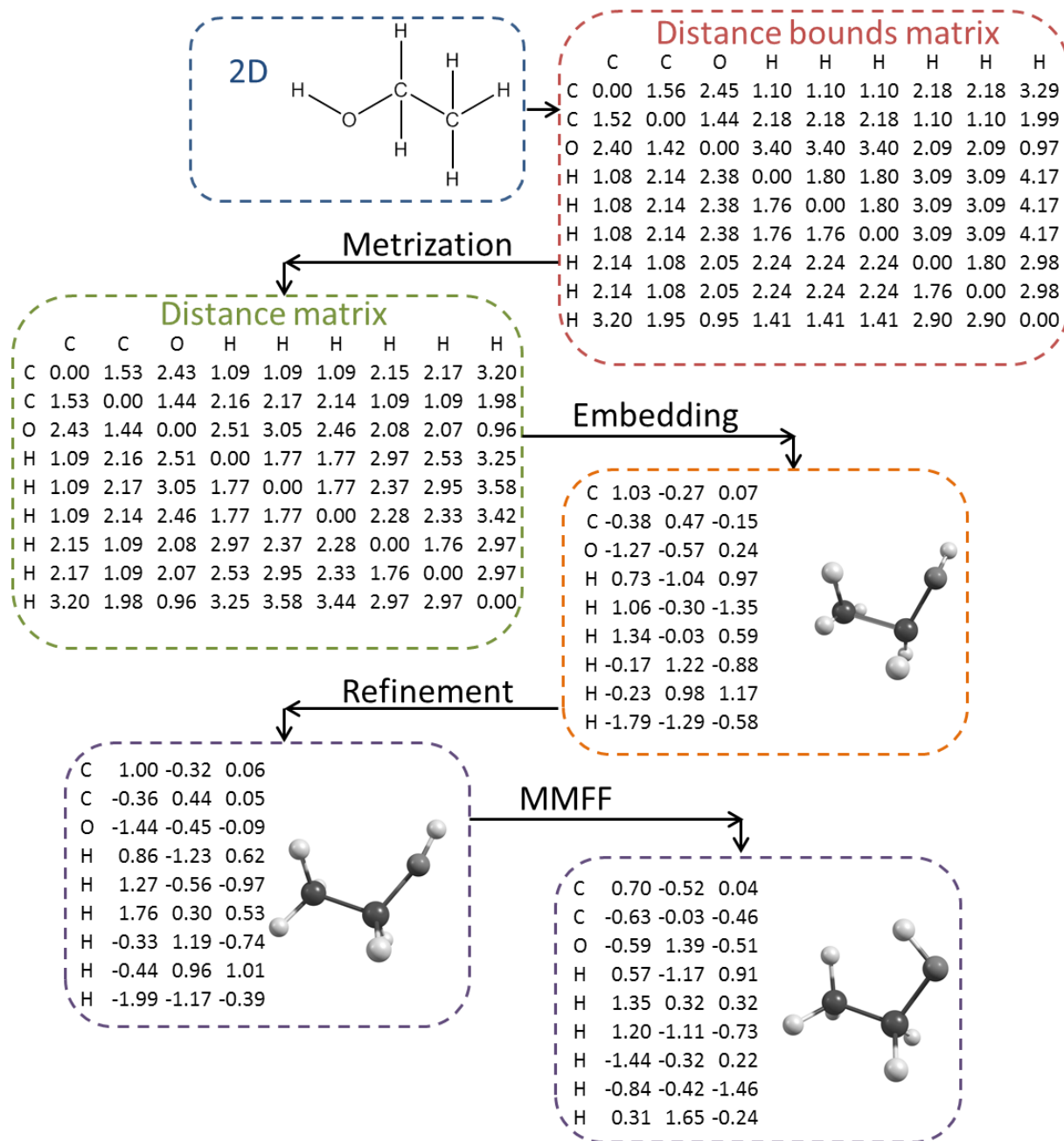


Figure 2: Illustration of the coordinates generation procedure.

The following paragraphs on distance geometry are based on the work of Havel³⁰. In order to allow distance geometry, the lower l_{ab} and upper u_{ab} distance limits between each pair of atoms a and b in a molecule or radical needs to be generated. This information is put in a matrix with the upper limits in the upper triangle and the lower limits in the lower triangle. A well-constructed distance bounds matrix contains the full conformational space of a molecule or radical, i.e. the

interatomic distances in each conformer all lie within the lower and upper bound defined in the distance bounds matrix.

The generation of the lower and upper bounds between two atoms depends on the atomic topology. Let A, B, C, and D be four atoms subsequently bonded to each other as shown in Figure 3. The distances between A and B, B and C, and C and D are 1-2 distances, i.e. a 1-2 distance is the distance between two atoms that are bonded to each other. Similarly, a 1-3 distance is the distance between two atoms that are both bonded to the same atoms, in Figure 3 the distances between A and C, and B and D are 1-3 distances. Finally, a 1-4 distance is defined by the distance between two atoms that are bonded to two other atoms which are bonded to each other, such as the distance between A and D in Figure 3. The distance limits of atoms that are bonded to each other, are defined by a small range around the average bond length between them. Average bond lengths are tabulated in databases. 1-3 distances are calculated based on the two bond lengths and the average bond angle, which is also tabulated. The 1-4 distance must lie between the minimum distance defined by the syn-conformation and the maximum distance defined by the anti-conformation. A small range below and above the minimum and maximum distances respectively is used to allow flexibility of the algorithm. The ranges chosen around tabulated equilibrium values are important for the accuracy of the algorithm. Both too loose and too strict limits will lead to unrealistic molecular structures.

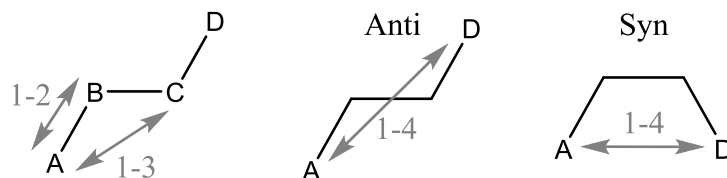


Figure 3: Illustration of 1,2, 1,3 and 1,4 distances.

Before the matrix can be used for the generation of 3D coordinates, inconsistencies need to be corrected. An algorithm called “triangle inequality bound smoothing” is chosen as adequate method to locate contradictions in the bounds. In this algorithm, the upper and lower limits between two atoms a and b need to be compared to the path through a third atom c, as illustrated in Figure 4. The path is a sequence of three atoms (a,c,b) with the length calculated based on the two upper and two lower bounds (u_{ac} , u_{bc} , l_{ac} , l_{bc}). The upper bound u_{ab} cannot exceed the sum of the two lower bounds l_{ac} and l_{bc} . Similarly, the lower bound l_{ab} should be larger than $l_{km} - \bar{u}_{ik} -$

\bar{u}_{jm} in which i, j, k and m are not necessarily distinct atom indices. The overbars denote the triangle inequality limits.³¹

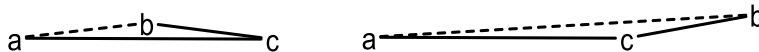


Figure 4: Illustration of the upper and lower limit of the distance between A and C based on the distances of A and B and B and C.

After the triangle inequality bound smoothing, a random distance matrix is chosen that fits within the boundaries imposed by the distance bounds matrix, a procedure which is called “metrization”. For each two atoms A and B, a distance d_{ab} is obtained by Eq. 3.1 in which r is a random number between 0 and 1. By applying this procedure several times, different distance matrices can be build, belonging to a different conformer or a different stereo-isomer.

$$d_{ab} = l_{ab} + r(u_{ab} - l_{ab}) \quad \text{Eq. 3.1}$$

This is followed by the actual fitting of coordinates to the distances, i.e. embedding. More information on the embedding algorithm can be found in Appendix C or in literature^{29,30}. Following paragraphs give the approach used in Genesys in brief. In an embed algorithm, the coordinates $\mathbf{x}_1, \dots, \mathbf{x}_N$ of the N atoms that are a best-fit to the estimated distances can be found fast and reliable by eigenvalue methods. For this work, unweighted embedding is performed, i.e. all the weight factors $w_{i,j}$ are equal to 1. The squared distances d_{ab}^2 are used, which are denoted as D_{ij} . In a first step, the distances of each atom to the center of mass of the species is calculated, in which m_j is the mass of the j^{th} atom and M is the total mass of the species:

$$D_{0i} = M^{-1} \sum_{j=1}^N m_j D_{ij} - M^{-2} \sum_{1=j<k}^{N,N} m_j m_k D_{jk} \quad \text{Eq. 3.2}$$

It can be easily proven that the distances to the center of mass, i.e. center of mass coordinates, is exact if the estimated distances among the points are exact. The embedding procedure searches for a minimum of the function given in Eq. 3.3.

$$\frac{1}{2} \sum_{i,j=1}^{N,N} \left((\mathbf{x}_i \cdot \mathbf{x}_j - a_{ij}) \right)^2 \quad \text{Eq. 3.3}$$

$$a_{ij} := \frac{1}{2}(D_{0i} + D_{0j} - D_{ij}) \quad \text{Eq. 3.4}$$

By introducing the matrices \mathbf{A} and \mathbf{X} , c.f. Eq. 3.5 and Eq. 3.6, the function in Eq. 3.3 can be expressed as a squared Frobenius norm, shown in Eq. 3.7.

$$\mathbf{A} := [a_{ij}] \quad \text{Eq. 3.5}$$

$$\mathbf{X} := [x_1 \dots x_N]^T \quad \text{Eq. 3.6}$$

$$F(\mathbf{X}) = \frac{1}{2} \|\mathbf{X}\mathbf{X}^T - \mathbf{A}\|^2 \quad \text{Eq. 3.7}$$

A necessary condition for the global minimum is that the gradient becomes zero.

$$\left[\frac{\partial F}{\partial x_{ij}} \right] = (\mathbf{X}\mathbf{X}^T - \mathbf{A})\mathbf{X} = \mathbf{0} \quad \text{Eq. 3.8}$$

$$\mathbf{A}\mathbf{X} = \mathbf{X}(\mathbf{X}^T\mathbf{X}) \quad \text{Eq. 3.9}$$

$\mathbf{X}^T\mathbf{X}$ is a 3x3 matrix named the tensor. It can be assumed that the coordinates are rotated in space so that the inertial tensor is diagonal.

$$\mathbf{X}^T\mathbf{X} = \mathbf{diag}(\lambda_1, \lambda_2, \lambda_3) \quad \text{Eq. 3.10}$$

The columns of \mathbf{X} are proportional to eigenvectors of \mathbf{A} and the moments of inertia $\lambda_1, \lambda_2, \lambda_3$ are the corresponding eigenvalues. Combining the previous two equations result in Eq. 3.11.

$$\mathbf{X}^T\mathbf{A}\mathbf{X} = (\mathbf{X}^T\mathbf{X})^2 = (\mathbf{diag}(\lambda_1, \lambda_2, \lambda_3))^2 \quad \text{Eq. 3.11}$$

By making use of the expression resulting from the global minimum equation we can rewrite the Frobenius norm, c.f. Eq. 3.12.

$$F(\mathbf{X}) = \text{tr}(\mathbf{A}^2) - \lambda_1^2 - \lambda_2^2 - \lambda_3^2 \quad \text{Eq. 3.12}$$

The global minimum \mathbf{X} is obtained by taking the three largest nonnegative eigenvalues of \mathbf{A} and scaling the corresponding eigenvectors by their square roots, which contains the initial coordinates of all the atoms.

The initial Cartesian coordinates generated by embedding reflect the overall shape. However, the quality of this estimate is poor and it does not satisfy the original distance bounds matrix as a

result of the compression of the structure during its projection from $N - 1$ dimensions to three dimensions. Therefore an improvement step is introduced in which the coordinates are refined by use of an error function that forces all constraints to lie between the lower and upper bounds. The distance error function takes following form, the first term enforces upper bound constraints while the second term handles the lower bounds.

$$F_{dist} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \max \left[0, \left(\frac{d_{ij}^2}{u_{ij}^2} - 1 \right)^2 \right] + \max \left[0, \left(\frac{2l_{ij}^2}{l_{ij}^2 + d_{ij}^2} - 1 \right)^2 \right] \quad \text{Eq. 3.13}$$

Distance geometry is a computationally fast method to generate conformers, but it has the downside that purely distance-based constraints tend to lead to distorted aromatic rings and sp^2 centers. Optimization with force field calculations can correct this. Therefore, the Merck Molecular Force Field (MMFF)³² has been implemented in Genesys as refinement step. Since several optimization steps are done after the force field calculations, the original version of the MMFF was chosen for this work, without any adjustments.

Stereoisomers can have distinct thermodynamic parameters, e.g. in the case of multiple chiral centers or when cis-trans centers are present in the species. Genesys explicitly accounts for each stereoisomer of a species³³, and thus needs to build the coordinates for each stereoisomer separately. In the case of chiral centers, the distance geometry algorithm is applied several times, as shown in Figure 5, each time with a different distance matrix, i.e. by varying the random numbers, and structures that do not comply with the stereoconfiguration are eliminated. Because distance geometry is very fast, this step does not influence the total calculation time significantly. The same approach is adopted for species in which the cis-trans center reaches further than a 1-4 distance. For cis-trans isomers over a double bond, the lower and upper limit for the 1-4 distances are based on the stereoconfiguration of the double bond.

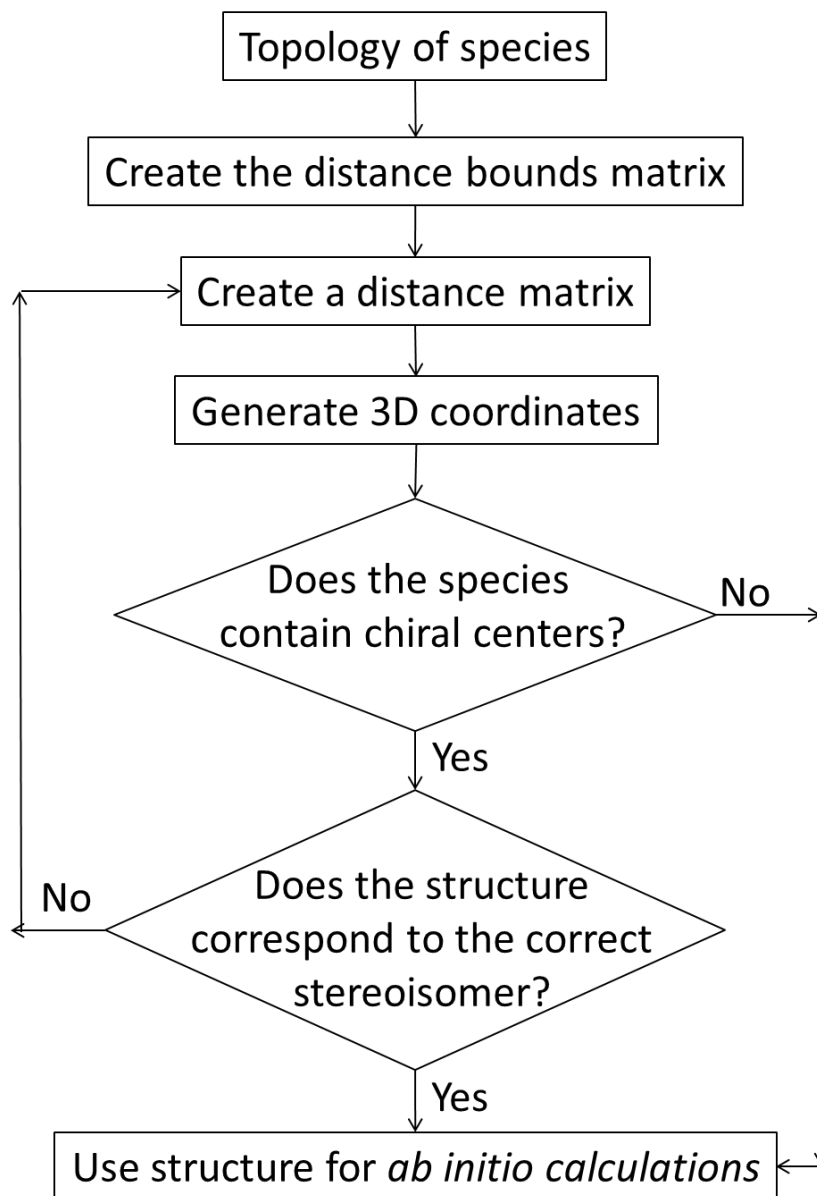


Figure 5: Flowsheet of the generation of 3D coordinates for species with chiral centers in Genesys

3.3.1.2 *Ab initio* calculations

Once an initial estimate for the 3D structure of a molecule or radical is generated, *ab initio* calculations can be started using the Gaussian09 suite of program³⁴. Genesys has been programmed to allow flexibility of the user in terms of the used *ab initio* steps, but a basis of four steps is recommended. Initially, a pre-optimization method is used at low level to refine the 3D structure. Which method to employ is user-defined, common methods are HF/STO-3G or semi-empirical methods such as PM3. For this work, PM3 was chosen for its wide applicability and its speed. Genesys only records the final 3D coordinates of the Gaussian calculation. The

connectivity of the final structure is verified by comparing it to the 2D connectivity of the species or transition state.

The second step is the search for the lowest energy conformers. Again, the level of theory is user-defined, and B3LYP/3-21G* was used in this work, although any level can be specified. The procedure for the conformational search is given in Figure 6. The conformational analysis is exhaustive for open chain species and non-exhaustive for ring structures. Each conformer is generated by Genesys and submitted as a separate calculation to Gaussian. For open-chain species, each internal rotor is first identified. In a second step, starting from the pre-optimized 3D coordinates, all the combinations of k dihedral increments in the n rotors are generated by rotating all the atoms on one side of the rotor bond around the axis of the bond around an angle of $2\pi/k$. This results in a total of k^n structures, as is illustrated for diethyl sulfide in Figure 7, in which two internal rotor exists: the two C-S bonds. For diethyl sulfide n is thus 2. The value of k is user-defined, and a value of 3 is used in this work. A total of $3^2 = 9$ structures are generated.

Although a total of k^n structures are initially generated, the number of Gaussian calculations is lower than k^n because (1) conformational duplicates can be formed in the case of symmetry in molecules and radicals, and (2) some combinations of dihedral angles result in chemically insignificant or different structures, i.e. non-bonded atoms are too close to each other. In Figure 7, several structures are equivalent due to the external symmetry. For example, structure *c* can be superimposed to structure *d* by rotating the structure around a vertical axis going through the sulfur atom. Analogously, structures *b* and *g* are identical, structures *e* and *i* are identical, and structures *f* and *h* are identical. Furthermore, structures *b* and *c* are each other's mirror structure, and have thus the same energy. This results in four energetically different structures as shown on the figure.

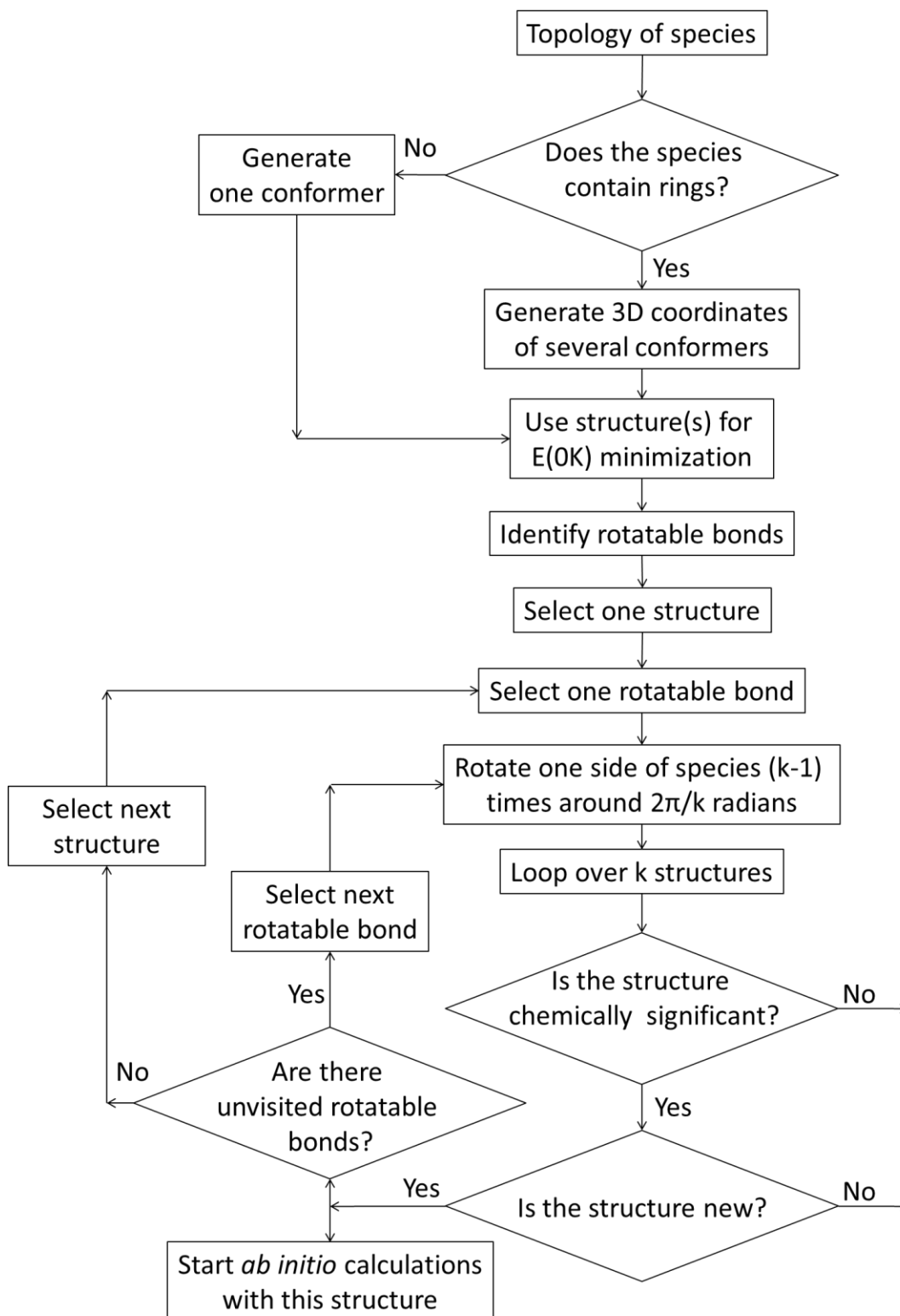


Figure 6: Flowsheet of the conformer generation procedure in Genesys.

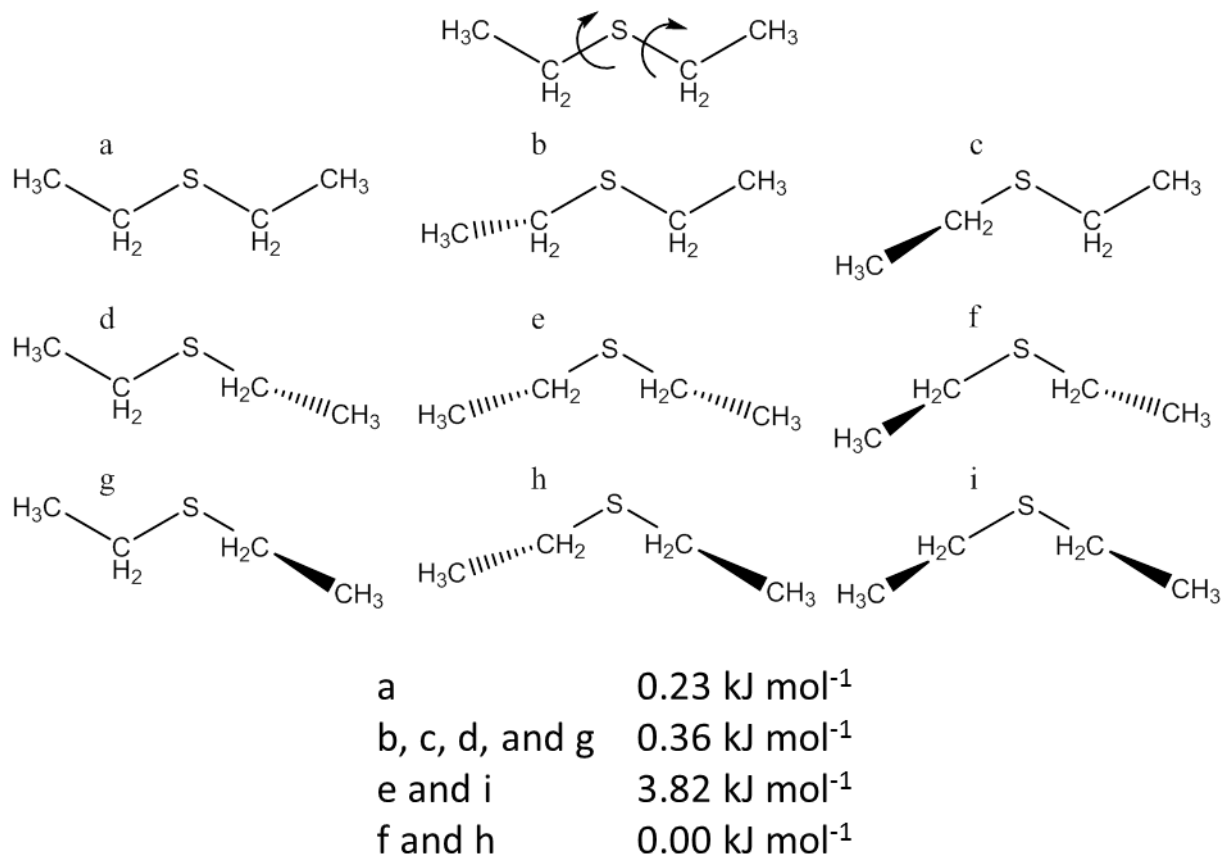


Figure 7: Application of the conformational analysis algorithm to diethyl sulfide. The energies are the electronic CBSQB3 energies relative to the lowest energy conformer.

For ring structures, the distance geometry algorithm is used to generate several conformers. By applying the algorithm a number of times, different sets of dihedral angles in the ring structure are formed. All these structures are first pre-optimized as discussed above. The conformation of side chains to the ring structure is exhaustively analyzed analogously to open chain species. Genesys keeps track of the geometries of all the conformers together with their electronic energy.

A third step in the Gaussian calculations is the optimization at high level of theory. This step is done for all the conformations that are low in energy. For example, in this work, all the conformations within 5 kJ mol⁻¹ of the lowest energy conformer are optimized at high level. Since the lower level of theory of the conformational analysis cannot always identify the lowest energy conformer, it is necessary to also calculate the high level energy of structures that are slightly higher in energy in the conformational scan. In this work, the high level optimization was done at the CBS-QB3 level of theory. The final geometries are recorded, together with the electronic energy, and the frequencies and normal modes of each conformer. The conformer with the lowest energy is selected to continue the calculations. This energy is the electronic energy at

OK, this is thus only an approximation of the lowest energy conformer. A more accurate determination of the conformers would be a Boltzmann average of all the conformers, evaluated at the temperature of interest.

The final step in the Gaussian calculations is the determination of 1D hindered rotor profiles around rotatable bonds. The same rotatable bonds as identified during the conformational scan are used to perform a MODREDUN calculation in Gaussian in which the dihedral angle is gradually increased, with the size of the increment supplied as user option. A default of 10 degrees is set, which was used in this work. MODREDUN calculations in Gaussian can freeze one of the internal coordinates of the molecule or radical while optimizing all other coordinates. The stepwise application of these optimizations while increasing the dihedral angle results in an energy profile as function of that angle. Genesys verifies whether the final energy of the rotational scan is the same energy as for the initial geometry. If this is not the case, the hindered rotor profile is considered to be incorrect.

3.3.2 Transition states

The above methodology for species can also be applied for transition states. However, several additional steps are required.

First, the generation of initial 3D coordinates is more complex compared to molecules or radicals because there are no universal force field parameters available for transition states. To use Genesys, users need to supply reaction families, reaction family constraints, computational details, etc. based on their knowledge of the chemistry at hand. This philosophy is further employed in the generation of transition state structures. It is well-known that all reactions belonging to the same reaction family have strong similarities in terms of geometrical arrangement of the atoms that change in connectivity throughout the reaction, i.e. the reactive atoms. When calculating reaction rate coefficients with the *ab initio* module of Genesys, the 3D structure of the reactive center, i.e. all the reactive atoms relative to each other, needs to be supplied by the user, this structure is called the template. This structure can be filtered from the vast knowledge on transition state structures generated in the past. Furthermore, the structure of each transition state optimized by Genesys is saved in its database. These databases will be used in the future to decrease the user involvement and aim at fully automated generation of 3D coordinates for transition states.

The 3D coordinates of the reactive center is used to build the distance bounds matrix. All the atoms that do not change in connectivity are treated as atoms not part of a transition state. With the distance bounds matrix, an initial structure can be formed, which is also further refined with force field calculations. For this, new parameters need to be defined relative to the reactive atoms. The equilibrium distances and angles are deduced from the user supplied template. Since no force field parameters are available for transition states, force constants do not exist to accompany the equilibrium distances. Instead, the force constants are estimated based on known force constants. For this, the topology of the transition state is used, in which transition state specific bonds are replaced by conventional bonds. The force constants for most of these bonds already exists in the MMFF.³² The force field contributions for which no force constants are defined yet are obtained by assigning force constants from similar bonds.

The initial transition state structure is pre-optimized with *ab initio* calculations using Gaussian software. In these calculations the structure of the reactive atom is fixed and only the remaining part of the transition state is energy-optimized. This allows to relax the non-reactive atoms, but no saddle point needs to be identified at this low level of theory, for this, the MODREDUN option is used. On the other hand, during the conformational analysis, the full transition state structure is optimized. No coordinates are fixed and the energy of the final 3D coordinates are used to select the low energy conformers. The identification of the transition state is done with the TS option in the optimization step. If the user-supplied template is indeed close to the saddle point structure, this step will easily identify the full transition state structure. The final optimization and hindered rotor steps are also calculated with the TS option.

Optimizing a molecular structure to a saddle point on the PES can sometimes result in an unwanted structure, often correlated with a maximum energy along a rotational path. The assessment of a correct transition state calculation can be done in several ways. Often, user involvement is required by visually controlling the structure and animating the normal mode corresponding to the imaginary frequency. However, since Genesys aims at automatically performing *ab initio* calculations without any user involvement, this is not an option. Another possibility is performing intrinsic reaction coordinate (IRC) calculations. These calculations start from the transition state structure and follow the imaginary frequency downwards in the two wells that are connected through this transition state. If the wells correspond to the expected

reactants and products of this reaction, an adequate transition state was found. Although these calculations are reliable and can be fully automated, Genesys does not make any use of them because (1) IRC calculations are slow and computationally demanding and (2) Gaussian checkpoint files are required, which can become very large, often leading to computational memory limitations. Therefore, a new method was developed, to verify the transition states in two steps. First, the 3D coordinates of the reactive atoms are compared with the template of the reaction family. The transition state structure optimization is only considered successful if the bonds lengths do not deviate too much from the user-defined bond lengths. Second, the normal mode corresponding to the imaginary frequency is analyzed. Analysis of normal modes is explained in more detail in section 3.3.3. The normal mode corresponding to the imaginary frequency in a transition state has large contributions in the length variation of bonds that are formed or broken throughout the reaction. If these contributions are larger than the bond length contribution of other bonds, the transition state optimization is considered successful.

3.3.3 Thermodynamics

To obtain thermodynamic values from *ab initio* calculations, i.e. the enthalpy, entropy and heat capacities of species, ideal gas statistical thermodynamics is used. The approach used in Genesys is similar to the work of Van Speybroeck³⁵ and Sabbe and coworkers^{36,37}, from which the equations below are taken.

The molecular partition functions are calculated based on the results from Gaussian. The translational, rotational, vibrational, and electronic contributions are assumed to be uncoupled and can be separated, c.f. Eq. 3.14.

$$q = q_{trans} \cdot q_{rot} \cdot q_{vib} \cdot q_{elec} \quad \text{Eq. 3.14}$$

For the electronic partition function, it is assumed that only the ground state contributes, which has a value of 1 for closed shell and 2 for open shell species. Another contribution that is straightforward to calculate is the translational contribution, which depends on the mass of the species and the temperature only as shown in Eq. 3.15.

$$q_{trans} = V \left(\frac{2\pi m k_B T}{h^2} \right)^{\frac{3}{2}} \quad \text{Eq. 3.15}$$

The vibrational contributions are obtained from the harmonic frequencies of the species. When using CBS-QB3, a scaling factor of 0.99 on the frequencies is used, and the total vibrational partition function is calculated as a product of the separate ones, Eq. 3.16. A non-linear species with N atoms has $3N-6$ degrees of freedom, i.e. $3N-6$ molecular frequencies.

$$q_{vib} = \prod_{i=1}^{3N-6} q_{vib,i} = \prod_{i=1}^{3N-6} \frac{e^{-hv_i/2k_B T}}{1 - e^{-hv_i/k_B T}} \quad \text{Eq. 3.16}$$

Finally, the rotational contribution is calculated as follows, c.f. Eq. 3.17 (for non-linear species). The principal moments of inertia are calculated based on the optimized 3D coordinates of the atoms.

$$q_{rot} = \frac{1}{\sigma_{rot}} 8\pi \left(\frac{2\pi k_B T}{h^2} \right)^{\frac{3}{2}} \sqrt{I_1 I_2 I_3} \quad \text{Eq. 3.17}$$

The vibrational contributions as described above assume harmonic oscillations around the optimized 3D coordinates, which is a reasonable assumption for most of the molecular frequencies. However, the frequencies that resemble an internal rotation around a bond cannot be calculated based on the harmonic oscillator approximation. For these modes, hindered rotor calculations are necessary. In this work, it is assumed that each rotor is independent of the other rotors and that 1D relaxed scans can be used. All rotational scans with a barrier lower than 50 kJ mol⁻¹ are treated as hindered rotors, others are considered harmonic oscillators. To calculate the contribution to the partition function $q_{vib,i}$ of the i^{th} vibrational mode resembling an internal rotation, the potential energy of the relaxed scan is approximated by a function as shown in Eq. 3.18, in which φ is the torsional angle. The coefficients A_k and B_k are calculated per hindered rotor scan using a linear least-square regression method. This function is chosen as opposed to Fourier series because it fits hindered rotor profiles better. Since the rotational scan starts at the lowest energy configuration of the species, the potential energy is zero when φ is zero, which is always the case in Eq. 3.18. The fit of the potential energy is finally also compared to the *ab initio* potential energy, which allows to identify discontinuities in the *ab initio* profile. These are communicated to the end-user.

$$V(\varphi) = \sum_{k=1}^n \frac{1}{2} A_k (1 - \cos(k\varphi)) + \sum_{k=1}^n B_k (\sin(k\varphi)) \quad \text{Eq. 3.18}$$

Before the Schrödinger equation can be solved for an internal rotation, the reduced moment of inertia of the rotation needs to be calculated. In this work, the $I^{(2,3)}$ estimator of the moment of inertia, according to East and Radom³⁸ is used. The first number (2) corresponds to the approximation of the coupling of one rotor with the rotation of the entire species. The second number (3) corresponds to the use of the axis through the centers of mass of each top as rotational axis.

On each side of the bond around which the hindered rotor profile was obtained, the moment of the rotating top is calculated.

$$I_{top,1} = \sum_{k=1}^{N_1} m_k d_{A,k}^2 \quad \text{Eq. 3.19}$$

N_1 is the number of atoms on one side of the rotating top and $d_{A,k}$ is the distance from the k^{th} atom to the axis of rotations, i.e. the axis through the centers of mass of the tops. Similarly, $I_{top,2}$ is constructed with the atoms on the other side of the rotating top. The $I^{(3,2)}$ estimator can then be calculated by Eq. 3.20.

$$\frac{1}{I_{red}} = \frac{1}{I_{top,1}} + \frac{1}{I_{top,2}} \quad \text{Eq. 3.20}$$

With the reduced moment of inertia and $V(\varphi)$ representing the energy profile of the 1D hindered rotation, the 1D Schrödinger equation, Eq. 3.21, can be numerically solved, resulting in the energy eigenvalues ϵ_{ki} of the i^{th} rotation.

$$\left[-\frac{\hbar^2}{2I_{red}} \frac{\partial^2}{\partial \phi^2} + V_i(\phi_i) \right] \Psi_{ki}(\phi_i) = \epsilon_{ki} \Psi_{ki}(\phi_i) \quad \text{Eq. 3.21}$$

Finally, the partition function corresponding to the rotation can be calculated, c.f. Eq. 3.22. σ_i is the internal symmetry number of the rotation, g_k is the degeneracy of the energy level ϵ_{ki} , and m is the number of energy levels included when solving the Schrödinger equation.

$$q_i = \frac{1}{\sigma_i} \sum_{k=1}^m g_k e^{-\epsilon_{ki}/k_B T} \quad \text{Eq. 3.22}$$

The above treatment of rotors only applies for hindered rotors, free rotors also need to be taken into account. When the barrier for rotations is low, i.e. smaller than 1 kJ mol⁻¹, a rotor is treated as free. The partition function corresponding to a free rotor is given in Eq. 3.23.

$$q_i = \frac{(8\pi^3 I_{red} k_B T)^{\frac{1}{2}}}{h\sigma_i} \quad \text{Eq. 3.23}$$

The vibrational contribution to the partition function q_{vib} is multiplied by q_i to account for the hindered or free rotation. However, in order for q_{vib} to be correct, the molecular frequency corresponding to the torsional mode should not be accounted for in Eq. 3.16. Genesys thus needs to automatically select which frequencies ν_i to include in Eq. 3.16 and which ones to leave out. The selection procedure is based on the methodology developed by Ayala and Schlegel³⁹. This methodology identifies frequencies that correspond to rotations based on the normal mode of the frequency. This normal mode is given in Cartesian coordinates by Gaussian, normalized such that the sum of the lengths of all vectors is one. Obviously, this definition of the normal mode is only valid for one set of 3D coordinates, i.e. the final structure in its standard orientation. In order to compare frequencies in terms of their contributions to a rotational movement around a bond, the normal modes are translated into internal coordinates, describing the changes in bond lengths, bond angles, out-of-plane motions, and dihedral angles in a species. This translation is obtained by multiplying the vector with all the Cartesian coordinates of the normal mode with the so-called Wilson B matrix. The calculation of this matrix is elaborated in Appendix C. Once a normal mode is expressed in internal coordinates, the ratio of the contributions to changes in dihedral angles around one bond divided by the sum of all the other contributions can be calculated. In the case of a rotational movement, this ratio is high. For each bond around which rotation is possible, a ratio corresponding to each frequency is calculated and the frequency with the highest ratio is considered as the one corresponding to the torsional mode. For the next bonds, this frequency is no longer considered. This iterative procedure ensures that one frequency is selected per bond around which rotation is possible.

Once the molecular partition function $q(T)$ is known, macroscopic thermodynamic values can be calculated according to the equations Eq. 3.24-Eq. 3.26.

$$S(T) = R \left(1 + \ln \left(\frac{q(T)}{N_A} \right) + T \frac{\partial \ln(q(T))}{\partial T} \right) \quad \text{Eq. 3.24}$$

$$C_P(T) = R \left(1 + 2T \frac{\partial \ln(q(T))}{\partial T} + T \frac{\partial^2 \ln(q(T))}{\partial T^2} \right) \quad \text{Eq. 3.25}$$

$$H(T) = RT \left(1 + T \left(\frac{\partial \ln(q(T))}{\partial T} \right) \right) \quad \text{Eq. 3.26}$$

From the above calculated enthalpy $H(T)$, the standard enthalpy of formation at 298K is obtained from the atomization enthalpy method, i.e. Eq. 3.27 for a species with structural formula $X_m Y_n$. $\Delta_f H^\circ$ is the standard enthalpy of formation of the species, H is the enthalpy as calculated by Eq. 3.26 and $\Delta_a H^\circ$ is the atomization enthalpy, which is based on experimental data. All the values of $\Delta_f H^\circ$, $\Delta_a H^\circ$ and H are evaluated at 298K. The atomization enthalpy method is illustrated in Figure 8 for ethane. The calculated atomization enthalpy consist of the calculated standard enthalpies of the elements constituting the species minus the calculated standard enthalpy of the compound under study. The atomization enthalpies of the elements constituting the species from their standard originate from experimental data, c.f. Table 1.

$$\begin{aligned} \Delta_f H^\circ(X_m Y_n) = & m\Delta_a H^\circ(X) + n\Delta_a H^\circ(Y) - (mH(X) + nH(Y) \\ & - H(X_m Y_n)) \end{aligned} \quad \text{Eq. 3.27}$$

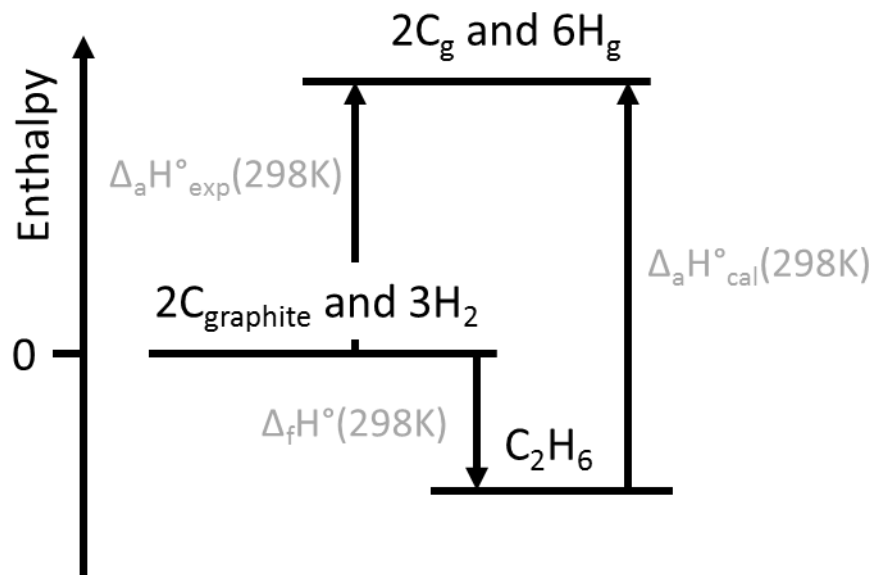


Figure 8: Schematic representation of the atomization enthalpy method to obtain the standard enthalpy of formation for the calculated enthalpy and experimental atomization enthalpies, with ethane as example. The atomization enthalpy $\Delta_a H^\circ_{exp}$ is experimentally determined, $\Delta_a H^\circ_{calc}$ is the calculated atomization of ethane, which is obtained from the *ab initio* enthalpies of ethane and its atoms.

Table 1: Experimental atomization enthalpies used in the atomization enthalpy method taken from the NIST Chemistry WebBook⁴⁰.

Element	$\Delta_a H^\circ(298K)$ (kJ mol ⁻¹)
H	217.998
C	716.68
N	472.68
O	249.18
Si	450.00
P	316.39
S	276.98

Finally, the standard enthalpy of formation can be corrected by making use of Bond Additive Corrections (BAC's) and spin-orbit corrections (SO). For this work, BAC's from literature have been taken for hydrocarbon and oxygenate species.⁴¹ No consistent set of BAC's was found for sulfur and nitrogen for CBS-QB3 results. Their standard enthalpies of formation were thus not corrected with BAC's. The enthalpy of formation was finally increased by applying spin-orbit corrections for carbon (0.35 kJ mol⁻¹), oxygen (0.93 kJ mol⁻¹), and sulfur (2.35 kJ mol⁻¹) atoms.

Table 2: Bond additive corrections taken from the literature.⁴¹ The value for the triple bonded carbon atoms (C≡C) comes from the work of Sabbe et al.³⁶.

Bond	BAC (kJ mol ⁻¹)
C-C	-2.07
C-H	-0.19
C-O	1.58
C=O	3.11
O-H	-1.77
C=C	-3.45
C≡C	-3.97

The final values for the standard enthalpy of formation, the standard entropy and the heat capacities can be used as such for kinetic model generation and reactor simulations. However, most reactor simulation codes require the thermodynamics to be in Chemkin⁴² format. This format does not contain the thermodynamic values directly but describes the temperature dependence of the thermodynamic properties in so-called NASA polynomials. These polynomials are constructed from two sets of 7 coefficients, one set for low and one for high temperatures. The 7 coefficients describe the thermodynamic properties as shown in Eq. 3.28-Eq. 3.30.

$$\frac{C_p(T)}{R} = a_1 + a_2 \cdot T + a_3 \cdot T^2 + a_4 \cdot T^3 + a_5 \cdot T^4 \quad \text{Eq. 3.28}$$

$$\frac{\Delta_f H(T)}{R} = a_1 \cdot T + \frac{a_2}{2} \cdot T^2 + \frac{a_3}{3} \cdot T^3 + \frac{a_4}{4} \cdot T^4 + \frac{a_5}{5} \cdot T^5 + a_6 \quad \text{Eq. 3.29}$$

$$\frac{S(T)}{R} = a_1 \cdot \ln(T) + a_2 \cdot T + \frac{a_3 \cdot T^2}{2} + \frac{a_4 \cdot T^3}{3} + \frac{a_5 \cdot T^4}{4} + a_7 \quad \text{Eq. 3.30}$$

To obtain the NASA polynomials of a chemical species, a linear least-square regression method is used based on the *ab initio* derived heat capacities. The unknown coefficients a_1 through a_5 constitute the \mathbf{b} vector, each row of the matrix \mathbf{X} contains the temperature exponents, as shown in Eq. 3.31, and each column corresponds to a temperature at which the head capacity is evaluated. Finally, the \mathbf{y} vector contains the heat capacities.

$$\mathbf{X}_i = [\mathbf{1} \quad T_i \quad T_i^2 \quad T_i^3 \quad T_i^4] \quad \text{Eq. 3.31}$$

The unknown coefficients in \mathbf{b} are given by:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad \text{Eq. 3.32}$$

Once a_1 through a_5 are known, a_6 and a_7 follow from one data point of the standard enthalpy of formation and the standard entropy respectively.

3.3.4 Kinetics

The above description of thermodynamics calculations is valid for stable species as well as transition state structures. The resulting thermodynamic properties can be employed to deduce reaction rate coefficients according to Eq. 3.33. In this expression, n_e is the number of single events, κ is the tunneling coefficient, k_B is the Boltzmann constant, V_m° is the standard molar volume, and h is the Planck constant. For bimolecular reactions, the difference in number of species between the transition state and the reactants $-\Delta n^\ddagger$ is -1, for monomolecular reactions, this value is 0. The single event entropy difference $\Delta \tilde{S}^\ddagger$ and the enthalpy difference ΔH^\ddagger are calculated from the standard entropy and standard enthalpy of formation of the reactants and the transition state of a reaction.

$$k(T) = n_e \kappa(T) \frac{k_B T (V_m^\circ)^{-\Delta n^\ddagger}}{h} \exp\left(\frac{\Delta \tilde{S}^\ddagger}{k_B}\right) \exp\left(-\frac{\Delta H^\ddagger}{k_B T}\right) \quad \text{Eq. 3.33}$$

Eq. 3.33 is often referred to as the macroscopic expression because of the use of the macroscopic thermodynamic properties of the reactants and products, i.e. the standard enthalpy of formation and the standard entropy. A rate coefficient can however also be constructed from partition functions, which are microscopic properties, leading to the microscopic expression Eq. 3.34. Here, ΔE^\ddagger the energy difference between the reactants and the transition state at 0K.

$$k(T) = n_e \kappa(T) \frac{k_B T (V_m^\circ)^{-\Delta n^\ddagger}}{h} \frac{q_\ddagger}{\prod_r q_r} \exp\left(-\frac{\Delta E^\ddagger}{k_B T}\right) \quad \text{Eq. 3.34}$$

The number of single events, also referred to as the reaction path degeneracy⁴³, is calculated from the rotational symmetry numbers and numbers of optical isomers of the reactants and products:

$$n_e = \frac{n_{opt,\ddagger}}{\prod_r n_{opt,r}} \cdot \frac{\prod_r \sigma_r}{\sigma_\ddagger} \quad \text{Eq. 3.35}$$

Tunneling is calculated using the Eckart⁴⁴ method. This method fits an Eckart potential through the zero-point-corrected energies of the stationary points, i.e. the reactants, transition state and products, c.f. Eq. 3.36.

$$V(s) = \frac{ae^{\alpha(s-s_0)}}{1 + e^{\alpha(s-s_0)}} + \frac{be^{\alpha(s-s_0)}}{(1 + e^{\alpha(s-s_0)})^2} + c \quad \text{Eq. 3.36}$$

With this potential, the Schrödinger equation can be solved exactly. This way, an analytic form of the transmission probability can be constructed. The tunneling coefficient is calculated from the formulas of Coote et al.⁴⁵ which implement the Eckart⁴⁴ method.

Analogously to thermodynamic parameters, rate coefficients $k(T)$ can often not be used as such in reactor modeling or kinetic model generation, but they need to be converted to another format, typically Arrhenius or modified Arrhenius expressions, as shown in Eq. 3.37 and Eq. 3.38 respectively.

$$k(T) = A \exp\left(-\frac{E_a}{RT}\right) \quad \text{Eq. 3.37}$$

$$k(T) = AT^n \exp\left(-\frac{E}{RT}\right) \quad \text{Eq. 3.38}$$

The Arrhenius coefficients A , n and E_a can be estimated by regression of a set of values of $k(T)$ in the temperature range of interest. A linear least-square regression is possible after transforming the Arrhenius expression into a linear form:

$$\ln(k(T)) = \ln(A) + n \ln(T) - \frac{E}{RT} \quad \text{Eq. 3.39}$$

For the linear least-square regression, c.f. Eq. 3.41, the \mathbf{b} vector contains $\ln(A)$, n and E_a , each row of the matrix \mathbf{X} is shown in Eq. 3.40 and each column corresponds to a temperature at which the rate coefficient is evaluated. The \mathbf{y} vector contains the rate coefficients.

$$\mathbf{X}_i = [1 \quad \ln(T_i) \quad -(RT_i)^{-1}] \quad \text{Eq. 3.40}$$

The Arrhenius coefficients are given by:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad \text{Eq. 3.41}$$

The significance of the regression can be verified by calculating the F-value, c.f. Eq. 3.42, in which n is the number of rate coefficients and p is the number of parameters, p is 2 for Arrhenius expressions and 3 for modified Arrhenius expressions.

$$F = \frac{\frac{(Xb)^T Xb}{p}}{\frac{y^T y - (Xb)^T y}{n - p}} \quad \text{Eq. 3.42}$$

The quality of the regression can further be analyzed using R^2 values, i.e. multiple correlation coefficients. R^2 values lie between 0 and 1, and the higher the value, the better the Arrhenius parameters are able to describe the rate coefficients.

$$R^2 = \frac{(Xb)^T Xb}{y^T y} \quad \text{Eq. 3.43}$$

In calculation procedures such as group additivity^{5-8,46-48}, the pre-exponential factor A is calculated based on several contributions. First, the group additivity model calculates a so-called single event pre-exponential factor \tilde{A} . The final pre-exponential A is obtained by multiplying \tilde{A} with the tunneling coefficient and the number of single events:

$$A = n_e \kappa \tilde{A} \quad \text{Eq. 3.44}$$

The tunneling coefficient and number of single events are calculated separately. The number of single events can be obtained from the topology of the reactants and transition state.⁴⁹ Many calculation procedures exist for tunneling coefficients, which depend on the reaction family. For example, Sabbe et al.⁶ introduced a correlation to calculate the tunneling coefficient in hydrogen abstraction reactions between hydrocarbons, c.f. Eq. 3.45. Similarly Eq. 3.46 shows a correlation for the tunneling coefficients in homolytic substitutions by hydrogen radicals at thiols and sulfides developed by Vandeputte et al.⁵⁰.

$$\kappa(T) = 1 + \left(\frac{162}{T}\right)^3 E_{a,exo} + 2.71 \cdot 10^{-6} \cdot \exp\left(-\frac{T - 300}{26}\right) E_{a,exo}^4 \quad \text{Eq. 3.45}$$

$$\kappa(T) = 1 + \left(\frac{83}{T}\right)^3 E_{a,exo} \quad \text{Eq. 3.46}$$

3.4 Results

Genesys has been applied to a range of species and reactions to illustrate its capabilities and the approach in automatic *ab initio* calculations during automatic kinetic model generation.

3.4.1 Stable species

Thermodynamic properties have been calculated for 21 species with variety of structural features: The set contains both small (C3-) and big (C10) species, hetero-atomic molecule (with oxygen, sulfur and nitrogen), ring structures, aromatic structures, radicals and resonantly stabilized species, optical and cis/trans isomers, and a molecule with H-bridge bonding. They are shown in Figure 9.

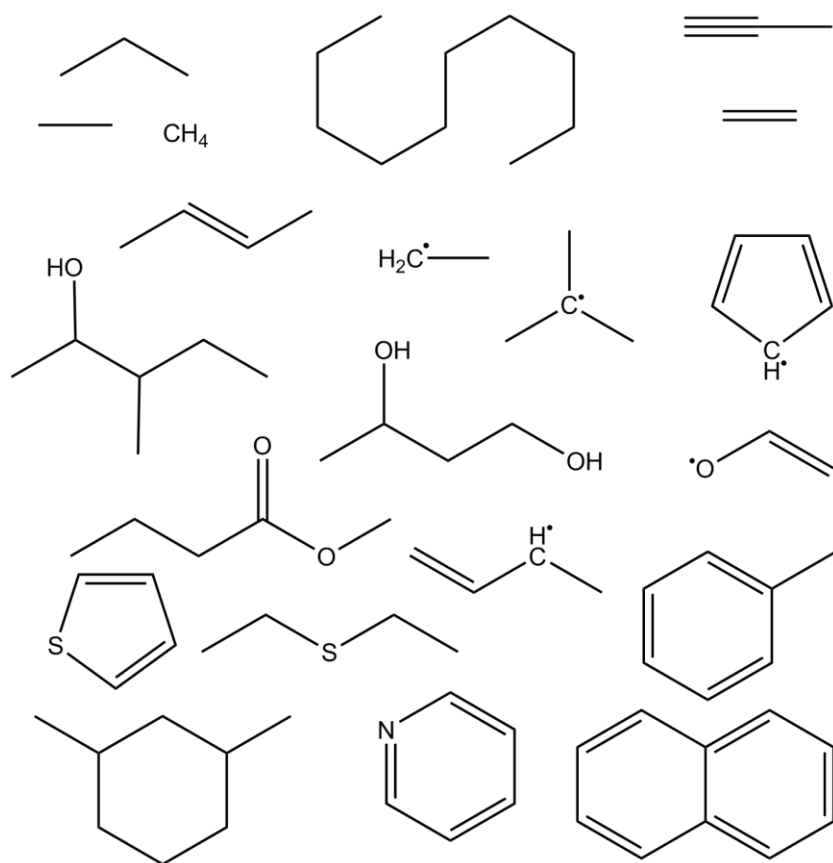


Figure 9: 21 species used to illustrate the on-the-fly *ab initio* thermodynamics calculations in Genesys.

A total of 2162 calculations were done with Gaussian, including 215 pre-optimizations, 1393 conformational scans, 502 final optimizations and the calculation of 52 1D hindered rotor potentials.

3.4.1.1 Initial coordinates generation

A crucial step in automatically performing *ab initio* calculations is the generation of initial 3D coordinates, which are close enough to a local minimum on the potential energy surface. The Merck Molecular Force Field³² has been shown to be very robust, all the necessary force field parameters needed to estimate the 3D coordinates of the atoms of the species in Figure 9 are present. The initial distance geometry algorithm results in structures in which several coordinates are far from the force field equilibrium values. The force field implemented in Genesys can however correct these coordinates, while keeping the connectivity of the species. In ethane for example, as shown in Figure 10, the distance geometry algorithm creates a 3D structure in which the bond lengths and bond angles are close to the force field equilibrium values. However, the dihedral angle can be further away from the equilibrium, and the ethane structure is close to the eclipsed configuration. The force field optimization corrects this by rotating around the C-C bond finding a minimum in forces for a dihedral angle close to 60°, which corresponds to the staggered configuration.

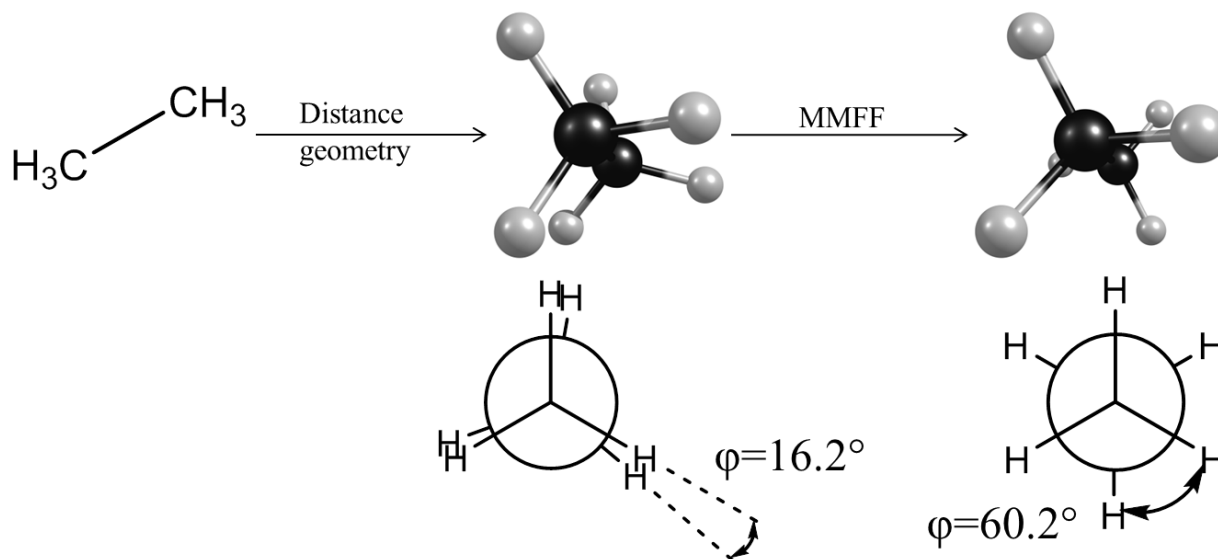


Figure 10: Illustration of the generation of 3D coordinates via “distance geometry” followed by force field calculations, ϕ is the H-C-C-H dihedral angle.

3.4.1.2 Conformational analysis

The conformational analysis is performed in two steps, depending on whether the species or transition state contains a cyclic structure. For ring structures, the distance geometry is used as conformer generator. The distance geometry algorithm is used several times to generate a number of initial structures, by varying the random numbers to select the distance matrix from the

distance bounds matrix. These structures are compared to each other via the Root Mean Square (RMS) distances between the atoms and identical structures are filtered out. For 1,3-dimethylcyclohexane, for example, this leads to 11 different sets of coordinates. For such a species, it is necessary to evaluate the structure of the cyclohexane skeleton and to evaluate the position of the methyl groups on the ring. Genesys automatically generated chair and twist-boat ring structures. For each conformer, the methyl groups can be each in equatorial, each in axial or one in equatorial and one in axial position. In the twist-boat structure, the carbon atoms are not equivalent, i.e. the symmetry is lower, this leads to more possibilities to substitute the ring with two methyl groups.

Conformational analysis is also important in open chain structures. After the pre-optimization, each conformer is generated by rotating only one side of a rotatable bond. This is exhaustively done for the whole species. In decane, there are 9 rotatable bonds. The methyl groups are not considered in the conformational scan, leading to 7 rotatable bonds that contribute in the conformational analysis. 3 scan were done around 120° ; a total of $3^7=2187$ structures were generated. These structures are filtered to eliminate chemically insignificant ones, i.e. when non-bonded atoms are too close to each other. Furthermore, due to symmetry, several conformers are present twice in the set of structures. These are identified based on RMS distances, and are eliminated from the calculations. For decane, 581 Gaussian calculations were performed during the conformational analysis. Out of the 581 calculations, 14 are within 5 kJ mol^{-1} from the lowest energy conformer, another 190 structures are within 10 kJ mol^{-1} from the minimum. The highest energy conformer is 35.7 kJ mol^{-1} above the lowest one. The structure in which every C-C-C-C dihedral angle is approximately 180° , was identified as global energy minimum for decane, as expected.

Hydrogen-bridge bonds are not included in the topology information of a species in Genesys. However, they play an important role when calculating thermodynamic parameters. During the conformational analysis of 1,3-butanediol, several conformers were created in which the hydrogen of one alcohol group approaches the oxygen atom of the other alcohol group. Figure 11 shows the difference in electronic energy between all the conformers of 1,3-butanediol, relative to the lowest energy conformer, calculated at the B3LYP/3-21G* level of theory. Each point on the graph corresponds to a different conformer. Flat areas of the graph mean that several

conformers have similar energies, whereas a steep slope represent a big gap in energy between the conformers. Such a gap is seen around 14 kJ mol^{-1} , where the next conformer has an energy of 25 kJ mol^{-1} . All conformers with an energy up to 14 kJ mol^{-1} contain a hydrogen-bridge bond, the others do not. Therefore, these structures are used in the high level optimization and downstream calculations.

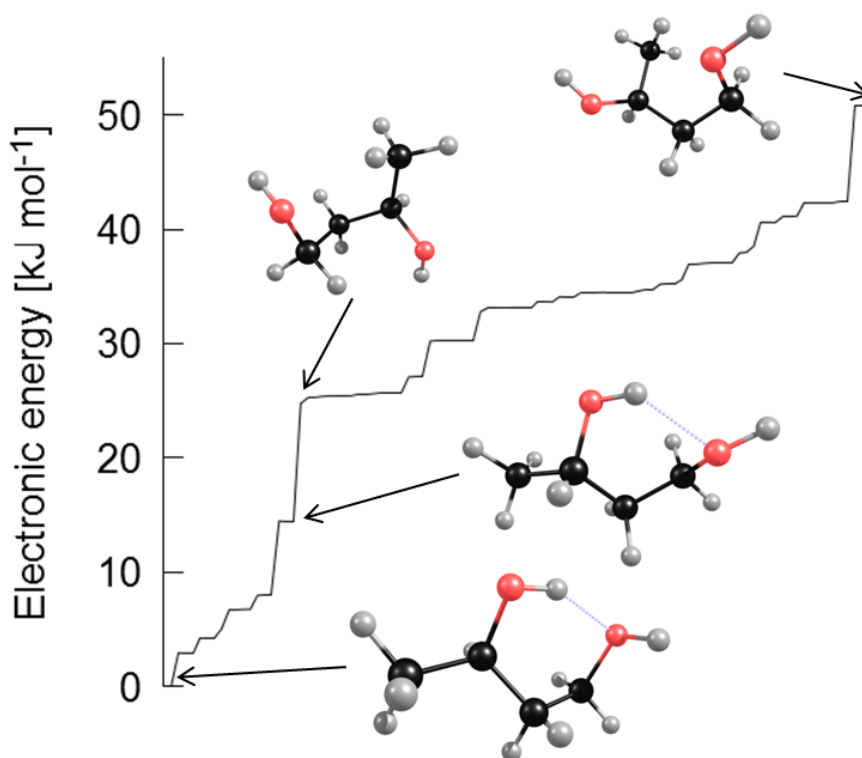


Figure 11: Electronic energy of all the conformers of 1,3-butanediol relative to the lowest energy conformer.

As mentioned in section 3.3.1.2, a cutoff energy of 5 kJ mol^{-1} is employed to start CBS-QB3 calculations. This cutoff is defined by making a tradeoff between the certainty to identify the lowest energy conformer and the number of high level of theory calculations that need to be done. In order to verify whether this value is not too low, energy calculations at high level of theory were performed for all the conformers. The maximum deviation in DFT energy between the conformer with the lowest CBS-QB3 energy and the conformer with the lowest DFT energy amounts less than 3 kJ mol^{-1} making the value of 5 kJ mol^{-1} appropriate for the molecules in this work.

3.4.1.3 Stereoconfiguration

As mentioned for 1,3-dimethylcyclohexane, Genesys explicitly takes the stereoconfiguration of species into account. When a species with a stereocenter is encountered, each stereo-isomer is

generated and treated as a separate species. In the case of 3-methyl-2-pentanol, there are two chiral centers. This leads to 4 possible stereoisomers, two set of two enantiomers, who relate as diastereomers to each other. Since enantiomers have the same electronic energy and frequencies, it is not necessary to treat them as separate species in Gaussian. Diastereomers, however, need to each have their own thermodynamic parameters assigned to. Figure 12 shows the pairs of enantiomers, together with their standard enthalpy of formation and entropy. Although the differences are small, 1.5 kJ mol^{-1} for the standard enthalpy of formation and $1.3 \text{ J mol}^{-1} \text{ K}^{-1}$ for the entropy, Genesys is able to account for these differences. Furthermore, it has been shown that these differences can play an important role during kinetic model simulations.³³

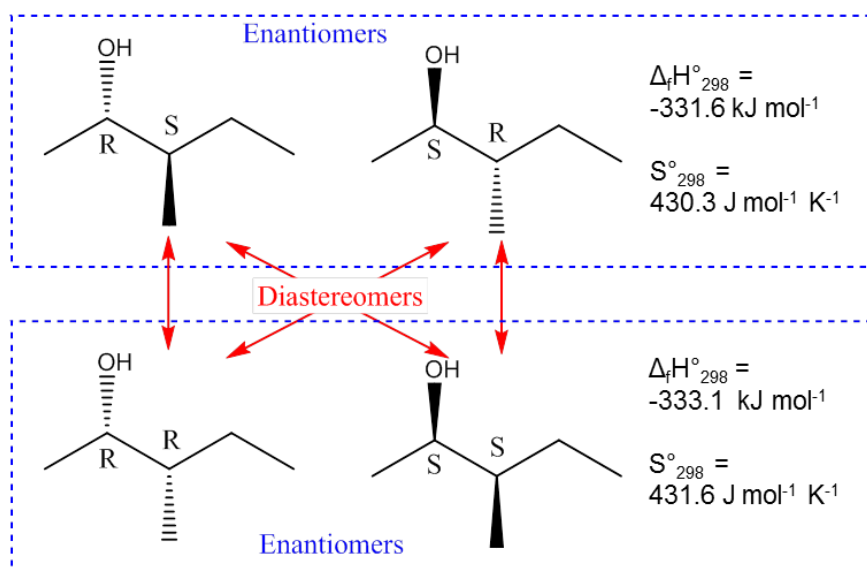


Figure 12: The stereoisomers of 3-methyl-2-pentanol consist of two pairs of enantiomers. Two elements of different pairs are diastereomers of each other. Diastereomers have different thermodynamic properties.

Besides multiple chiral centers, the stereoconfiguration of cis-trans centers also influences the thermodynamics. As for diastereomers, Genesys explicitly accounts for the cis and the trans configuration of a species. Both configurations are treated as separate species and separate coordinates are assigned to them. In the case of cis-trans stereoisomers over a double bond, the initial distance bounds matrix is different for each isomer, in contrast to chiral centers where the distances are identical for each isomer, since they are only based on the topology. When defining the 1-4 distances, a dihedral angle of 0° is chosen for cis-isomers, whereas the dihedral angle in the trans-isomer is 180° .

The equilibrium bond lengths in 2-butene are 154 pm for the single C-C bonds *a* and 134 pm for the double C-C bond *b*. The angle α is assumed to be 120° . With these geometrical constraints, the distance *d*, c.f. Figure 13, between the two terminal carbon atoms is 393 pm in the case of the trans structure, and 288 pm for the cis isomer. These values are used to build the distance bounds matrix, yielding different initial 3D coordinates.

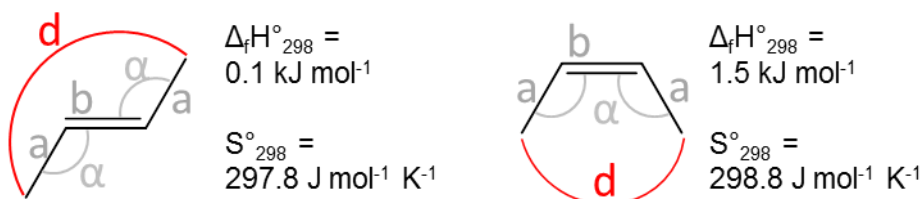


Figure 13: Distances between the two terminal carbon atoms in 2-butene depends on the stereoconfiguration of the double bond.

In species where cis-trans isomers exist, but where the locus of stereoisomerism is larger than two atoms, e.g. as a result of substituents on a ring structure, the distance bounds matrix cannot account for the stereocenters. Two distance bounds matrices of these isomers are thus identical. In this case, the proper isomer is selected after the generation of the 3D structure. As mentioned above, 1,3-dimethylcyclohexane has a cis and a trans isomer. Each isomer is tracked in Genesys, and separate thermodynamics are calculated.

3.4.1.4 Thermodynamics

Finally, the obtained thermodynamic properties, i.e. the standard enthalpy of formation and the standard entropy are compared to literature data, both experimental as well as theoretical data. Table 3 shows the comparison between the on-the-fly *ab initio* calculated standard enthalpies of formation and standard entropies and the literature data. Open-chain hydrocarbons show very good agreements in standard enthalpy of formation, definitely when corrected with BAC's. The standard entropy is also comparable to the experimental data. Cyclic structures show higher deviations, the BAC's decrease the standard enthalpy of formation too much. Standard entropy values are close to the experimental ones. For hetero-atomic species, there is only a limited amount of data available in the literature. The calculated standard enthalpy of formation of methyl butanoate is within chemical accuracy, i.e. an agreement within 4 kJ mol⁻¹, of the experimental data. Although the hydrogen-bridge was successfully found in 1,3-butanediol, a large deviation in standard enthalpy of formation is obtained.

Table 3: Comparison between Genesys and literature data of the standard enthalpy of formation and the standard entropy. Enthalpies are expressed in kJ mol⁻¹ and entropies in J mol⁻¹ K⁻¹.

	$\Delta_f H^\circ_{298}$					S°_{298}		
	Genesys	BAC +SO	Literature	Δ^a	Δ_{corr}^a	Genesys	Literature	Δ^a
Methane	-74.2	-0.4	-74.9 ^b	-0.7	-0.3	186.1	188.7 ^b	2.6
Ethane	-81.8	-2.5	-84.0 ^b	-2.2	0.3	229.2	229.1 ^d	-0.1
Ethylene	55.9	-3.5	52.5 ^b	-3.4	0.1	218.9	219.3 ^b	0.4
Propane	-100.7	-4.6	-104.7 ^b	-4.0	0.6	270.8	270.2 ^d	-0.6
Trans-2-butene	-3.5	-7.7	-10.8 ^b	7.2	0.4	298.8	296.5 ^d	-2.3
Cis-2-butene	1.5	-7.7	-7.7 ^b	-9.2	-1.5	297.8	300.8 ^d	3.0
Propyne	192.0	-5.8	185.4 ^b	-6.5	-0.8	247.4	248.1 ^d	0.7
Decane	-228.5	-19.3	-249.7 ^b	-21.2	-1.9	541.4	545.8 ^b	4.4
Trans-1,3-dimethylcyclohexane	-173.3	-16.8	-176.5 ^d	-3.2	13.6	378.7	376.2 ^d	-2.5
Toluene	58.1	-17.7	50.1 ^b	-8.0	9.8	320.1	320.7 ^d	0.6
Naphthalene	159.4	-27.7	150.0 ^b	-9.4	18.3	332.6	333.1 ^d	0.5
Methyl butanoate	-455.3	1.8	-454.4 ^c	0.9	-0.8	400.2		
3-methylpentane-2-ol	-331.7	-10.0				423.7		
1,3-butanediol	-459.7	-4.9	-433.0 ^b	26.7	31.5	359.7		
Thiophene	114.4	3.8	116.4 ^b	2.0	-1.8	278.6	278.9 ^d	0.3
Diethyl sulfide	-83.0	3.8	-83.5 ^b	-0.5	-4.3	369.9		
Pyridine	144.1	1.8	140.2 ^b	-3.9	-5.6	281.9	282.8 ^d	0.9
Ethyl radical	124.4	-2.3	119.0 ^b	-5.4	-3.1	247.8	248.0 ^f	0.3
t-butyl radical	59.7	-6.5	48.0 ^b	-11.7	-5.2	320.0	319.1 ^h	-0.9
1-buten-3-yl radical	141.3	-7.5	135.2 ^f	-6.1	1.4	301.1	301.1 ^h	0.0
Cyclopentadienyl radical	273.0	-12.3	271.0 ^e	-2.0	10.3	278.0	264.9 ⁱ	-13.1
Vinyloxyl radical	13.8	-1.7	14.6 ^g	0.8	2.6	258.9	258.9 ^g	0.0

^a Δ corresponds to $\Delta_f H^\circ_{298, \text{lit}} - \Delta_f H^\circ_{298, \text{genesys}}$ for the standard enthalpy of formation and $S^\circ_{298, \text{lit}} - S^\circ_{298, \text{genesys}}$. The subscript _{corr} refers to the calculated standard enthalpy of formation corrected with bond additive and spin-orbit corrections.

^b Nist Chemistry WebBook (<http://webbook.nist.gov/chemistry/>)

^c El-Nahas et al.⁵¹

^d Lange's Handbook of Chemistry, 14th Edition⁵²

^e Nunes et al.⁵³

^f Sabbe et al.³⁶

^g Paraskevas et al.⁴¹

^h Sabbe et al.³⁷

ⁱ Kiefer et al.⁵⁴

No BAC's were used to calculate the standard enthalpy of formation of nitrogen and sulfur containing compounds. Nevertheless, good agreement with the standard enthalpy of formation is found for pyridine, thiophene and diethyl sulfide. Their standard entropies also compare well with the literature data.

Five radical species were calculated and compared to other theoretical results. Except for the cyclopentadienyl radical, chemical accuracy is obtained for the standard enthalpy of formation. The entropies are very close to the literature data.

In general, good agreement has been obtained between automatically calculated *ab initio* thermodynamic values of several molecule and radicals and literature data, within chemical accuracy. Only minor manual interventions are necessary, i.e. the correction of the rotational symmetry, giving promising results to use Genesys for the calculations of thermodynamic parameters where data is lacking.

3.4.2 Reactions

To validate the automated *ab initio* reaction rate coefficient calculations, 8 reactions belonging to 8 different reaction families have been calculated automatically. These reaction families are inter- and an intramolecular radical addition, inter- and intramolecular hydrogen abstraction, substitution, dehydration, dehydrogenation, and retro-ene reaction. All reactions can be found in Table 4.

3.4.2.1 3D coordinates

For each reaction, a reaction family is constructed in Genesys. A part of the reaction family definition contains the coordinates of the reactive atoms. This information describes the 3D arrangement of the atoms that change in connectivity throughout the reaction at the saddle point, and has been constructed from literature and manually calculated transition states. The initial bond lengths and bond angles of the reactive atoms in the transition state are listed in Table 4, together with the final geometries. The letters refer to the bonds and angles depicted in the transition state (TS) figures.

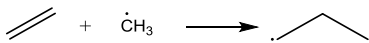
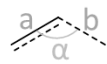
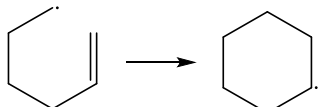
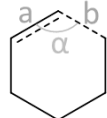
For radical addition reactions, both inter- and intramolecular, it has been assumed that the bond length between the attacking and the attacked carbon atom amounts to 231pm, according to literature data. Because this transition state is rather early, the bond length of the bond that changes from a double to a single bond is close to the average bond length of a double bond, i.e. 135pm compared to 133pm. This value is further away from the average singly bonded carbon-carbon bond length of 154pm. For the intramolecular addition, the bond angle between the three reactive atoms is restricted because of the ring structure in the transition state, whereas for the intermolecular case, the bond angle is higher due to steric hindrance.

The bond lengths in transition states of hydrogen abstractions can often be easily estimated by multiplying the stable bond lengths by a factor around 1.2. For a C-H bond and an O-H bond, this leads to a bond length of respectively 135pm and 239pm in the transition state. For intermolecular hydrogen abstractions, the bond angle between the three reactive atoms is assumed to be as high as possible, i.e. 180°. Similarly to intramolecular radical addition reactions, if the reactive atoms are part of a ring structure in the transition state, the bond angle is lower. In this case it is approximately 100°.

Bond lengths in substitution reactions such as **R5**, where the bond being broken and the bond being formed are similar, i.e. they are both single C-S bonds, can be assumed to be equal. The bond angle should be high because of the steric hindrance, but the sulfur atom is also bonded to a hydrogen atom, a tradeoff between both leads to a bond angle of around 153° as starting point.

Transition states for **R6-R8** are more complex to find since these reactions have more than three reactive atoms. Furthermore, the transition state is cyclic. The dehydration reaction proceeds through a four-membered ring structure as shown in Table 4. The transition states of **R7** and **R8** both contain a six-membered ring of which all atoms are part of the reactive atom. The transition state for the dehydrogenation reaction **R7** belongs to the S_1 point group, a reflection plane through the middle of the forming H-H bond, the C-C bond that decreases in order and the carbon atom that is not a reactive atom, reflects the species onto itself. Therefore, only 7 parameters (4 bond lengths and 3 bond angles) are required instead of 12. Finally, the retro-ene reaction **R8** depends on 6 bond lengths and 6 bond angles.

Table 4: Reactions used for the validation of the on-the-fly *ab initio* reaction rate coefficient calculations with the initial and final geometrical parameters of the reactive atoms. The distances are expressed in pm, the angles in degrees.

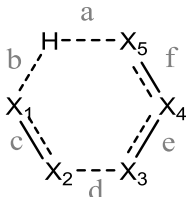
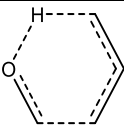
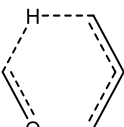
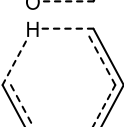
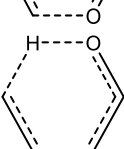
	Reactions	TS		Initial	Final
R1			a	135	135
			b	231	233
			α	120°	110°
R2			a	135	136
			b	231	231
			α	98°	100°

(Table 4 continued)

	Reactions	TS		Initial	Final
R3			a	135	132
			b	135	138
			α	180°	179°
R4			a	123	118
			b	239	237
			α	156°	153°
R5			a	188	211
			b	188	206
			α	153°	165°
R6			a	141	142
			b	160	190
			c	158	141
			d	115	128
			α	101°	94°
			β	73°	61°
			γ	116°	130°
			δ	68°	75°
R7			a	97	97
			b	156	156
			c	141	141
			d	139	139
			α	115°	115°
			β	96°	96°
			γ	108°	108°
R8			a	133	135
			b	137	140
			c	129	132
			d	184	189
			e	139	139
			f	141	142
			α	139°	147°
			β	105°	105°
			γ	123°	114°
			δ	102°	100°
			ϵ	122°	117°
			ζ	104°	99°

Each of the geometrical values used for the generation of initial coordinates for the atoms of transition states is valid for one specific reaction family because a reaction family is by definition a set of reactions with similar electronic structures in the transition state. The bonds that are formed, broken or change in order should be of identical nature, i.e. same elements with the same hybridization.

Table 5: Bond lengths (in pm) in the transition state of several retro-ene reactions with an oxygen atom in different positions.

	Transition state	a	b	c	d	e	f
		129	135	132	198	140	142
		135	140	132	189	139	142
		138	144	141	188	138	138
		118	144	139	229	137	133

Retro-ene reactions, for example, are important reactions in hydrocarbon and oxygenates pyrolysis and combustion. The parameters as listed in Table 4 can only be used in the case an oxygen atom is present in the β position of the hydrogen reactive atom. The other four reactive atoms need to be carbon atoms. In the case another retro-ene reaction is considered, with an oxygen atom in a different position, different geometrical parameters are required. This is illustrated in Table 5: Four transition states are depicted belonging to four different retro-ene reactions. The position of the oxygen atom is varied. The bond lengths between the reactive atoms in the transition state are strongly dependent on the position of the oxygen atom. Although the absolute deviations are limited to 40pm, the total deviation on the transition state structure is rather high and led to failed saddle point searches by Gaussian without the proper initial geometries.

The initial 3D coordinates generation leads to transition states in which all the bond lengths and angles are close to the ones defined in the reaction families. These starting geometries are used as input for Gaussian as explained above. The final geometrical parameters in Table 4 correspond to the CBS-QB3 optimized structures. Of course, the initial structure of the transition states needs to be close enough to the saddle point for a successful calculation, but from the differences between the initial and final geometries it can be seen that a reasonable estimate is acceptable, with a maximum deviation of 30pm for bond lengths and 14° for bond angles.

3.4.2.2 Kinetics

The rate coefficients are shown in Figure 14 and Figure 15 as a function of temperature, the modified Arrhenius parameters are listed in Table 6 together with the rate coefficients at 300 and 1000K, and the simple Arrhenius parameters are listed in Table 7. Tunneling coefficients and numbers of single events can be found in Table 8. Figure 14 and Figure 15 also report rate coefficients from the literature to compare the results. The modified Arrhenius parameters in Table 6 have been obtained via linear regression of the modified expression, after transforming it to a linear form. The uncertainty intervals in Table 6 correspond to a confidence of 95%. Regression of the conventional Arrhenius equation, i.e. Eq. 3.37, has also been done, the results can be found in Table 7. For most reactions, the regression has been done in a temperature range of 300 to 2000K. However, for two reactions, **R6** and **R7**, tunneling coefficients are very high at low temperature, which can be seen in Table 8. The uncertainty on the Arrhenius parameters is too high when considering the same temperature range and the Arrhenius parameters have thus been regressed in a temperature range of 600 to 2000K.

The rate coefficients in Table 6 are reported calculated from *ab initio* (AI) and from the modified Arrhenius parameters (ARRH). The ratio of both is also given, which shows a good agreement. An average deviation of 5% is observed with a maximum deviation of 17%. Conventional Arrhenius parameters do not reach this accuracy, as shown in Table 7, especially for the reactions for which tunneling coefficients are high. The average deviation amounts 47% with a maximum deviation of 202%. For each regression, the significance of the regression is verified by the F values, which all lie between $6.2 \cdot 10^4$ and $3.9 \cdot 10^7$, which are very high compared to the $F_{0.05}$ values of 3.68 and 3.89 for the regression of 300 to 2000K and 600 to 2000K respectively. The R^2 values are always larger than 0.999. The three parameters show very high binary correlation coefficients: -0.9994 for $\ln A$ and n , 0.9736 for $\ln A$ and E , and -0.9669 for E and n . For the simple Arrhenius parameters, the R^2 values are higher than 0.99, the binary correlation coefficient for $\ln A$ and E_a amounts 0.8242. Although the binary correlation coefficient is lower for simple Arrhenius parameters, their accuracy in a large temperature domain is too low to be used for automatic kinetic model generation. Simple Arrhenius parameters can be used in a more narrow temperature range, if the application of the kinetic model allows this.

Table 6: Rate coefficients for the validation set. The first row of each reaction corresponds to the forward rate coefficient, the second row to the reverse one. The units of E are kJ mol^{-1} , A and k are expressed in s^{-1} for monomolecular reactions and $\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$ for bimolecular reactions. Modified Arrhenius parameters have been obtained from linear regression of the modified Arrhenius equation. The uncertainty intervals of the regression correspond to a confidence of 95%. The pre-exponential factors contain the contributions for tunneling and the numbers of single events. The rate coefficients have been evaluated using the modified Arrhenius parameters (ARRH) and using the *ab initio* calculations (AI), their ratio is also reported.

	T range (K)	A	n	E	k_{ARRH}	k_{ARRH}	k_{AI}	k_{AI}	$\frac{k_{AI}}{k_{ARRH}}$	$\frac{k_{AI}}{k_{ARRH}}$
					300K	1000K	300K	1000K	300K	1000K
R1	300-2000	$6.6 \pm 1.5 \cdot 10^{-2}$	2.37 ± 0.02	23.0 ± 0.2	$4.8 \cdot 10^0$	$5.3 \cdot 10^4$	$4.9 \cdot 10^0$	$5.4 \cdot 10^4$	1.01	1.01
		$5.7 \pm 4.4 \cdot 10^{10}$	0.97 ± 0.07	123.3 ± 0.4	$4.9 \cdot 10^{-9}$	$1.7 \cdot 10^7$	$5.0 \cdot 10^{-9}$	$1.7 \cdot 10^7$	1.02	1.01
R2	300-2000	$4.0 \pm 2.6 \cdot 10^7$	0.70 ± 0.13	29.1 ± 0.8	$1.9 \cdot 10^4$	$1.5 \cdot 10^8$	$1.8 \cdot 10^4$	$1.4 \cdot 10^8$	0.97	0.92
		$8.0 \pm 4.3 \cdot 10^{11}$	0.77 ± 0.10	124.3 ± 0.6	$1.5 \cdot 10^{-8}$	$5.2 \cdot 10^7$	$1.5 \cdot 10^{-8}$	$5.2 \cdot 10^7$	1.02	0.99
R3	300-2000	$9.1 \pm 8.5 \cdot 10^{-9}$	4.48 ± 0.35	32.0 ± 2.0	$3.1 \cdot 10^{-3}$	$5.3 \cdot 10^3$	$3.4 \cdot 10^{-3}$	$5.9 \cdot 10^3$	1.11	1.11
		$3.1 \pm 2.9 \cdot 10^{-10}$	4.83 ± 0.39	48.3 ± 2.2	$1.1 \cdot 10^{-6}$	$2.9 \cdot 10^2$	$1.3 \cdot 10^{-6}$	$3.2 \cdot 10^2$	1.17	1.11
R4	300-2000	$2.6 \pm 2.0 \cdot 10^2$	2.56 ± 0.18	64.4 ± 1.0	$3.5 \cdot 10^{-3}$	$5.4 \cdot 10^6$	$3.7 \cdot 10^{-3}$	$5.7 \cdot 10^6$	1.06	1.06
		$6.0 \pm 3.9 \cdot 10^1$	2.40 ± 0.14	11.7 ± 0.8	$4.9 \cdot 10^5$	$2.3 \cdot 10^8$	$5.2 \cdot 10^5$	$2.5 \cdot 10^8$	1.07	1.07
R5	300-2000	$5.6 \pm 2.5 \cdot 10^{-1}$	1.90 ± 0.08	50.6 ± 0.4	$4.4 \cdot 10^{-5}$	$6.4 \cdot 10^2$	$4.6 \cdot 10^{-5}$	$6.1 \cdot 10^2$	1.04	0.96
		$1.1 \pm 0.1 \cdot 10^{-3}$	2.64 ± 0.01	50.5 ± 0.0	$6.1 \cdot 10^{-6}$	$2.1 \cdot 10^2$	$5.9 \cdot 10^{-6}$	$2.1 \cdot 10^2$	0.96	1.00
R6	600-2000	$8.2 \pm 6.9 \cdot 10^4$	2.50 ± 0.23	268.8 ± 2.0		$2.4 \cdot 10^{-2}$	$1.2 \cdot 10^{-31}$	$2.3 \cdot 10^{-2}$		0.98
		$1.5 \pm 1.2 \cdot 10^{-10}$	4.52 ± 0.19	224.1 ± 1.7		$1.1 \cdot 10^{-8}$	$8.2 \cdot 10^{-34}$	$1.1 \cdot 10^{-8}$		1.03
R7	600-2000	$8.7 \pm 5.4 \cdot 10^7$	1.46 ± 0.12	228.5 ± 1.1		$2.4 \cdot 10^0$	$1.1 \cdot 10^{-27}$	$2.3 \cdot 10^0$		0.95
		$2.3 \pm 0.3 \cdot 10^{-4}$	2.70 ± 0.02	125.0 ± 0.2		$8.6 \cdot 10^{-3}$	$1.2 \cdot 10^{-18}$	$8.3 \cdot 10^{-3}$		0.97
R8	300-2000	$6.5 \pm 4.3 \cdot 10^3$	2.29 ± 0.14	170.4 ± 0.8	$6.5 \cdot 10^{-21}$	$6.1 \cdot 10^1$	$7.0 \cdot 10^{-21}$	$6.5 \cdot 10^1$	1.07	1.07
		$1.8 \pm 1.0 \cdot 10^{-11}$	4.31 ± 0.11	148.8 ± 0.6	$1.1 \cdot 10^{-26}$	$2.6 \cdot 10^{-6}$	$1.1 \cdot 10^{-26}$	$2.6 \cdot 10^{-6}$	1.04	1.01

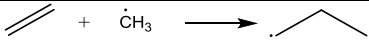
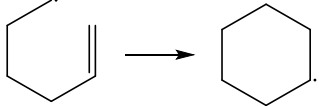
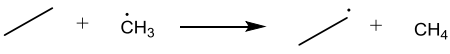
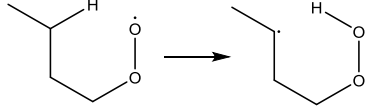
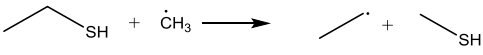
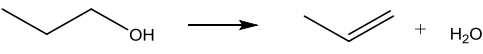
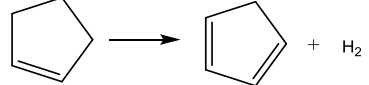
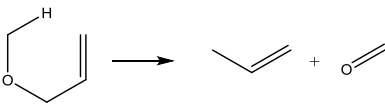
Table 7: Simple Arrhenius parameters for the validation set. The first row of each reaction corresponds to the forward rate coefficient, the second row to the reverse one. The units of E_a are kJ mol^{-1} , A is expressed in s^{-1} for monomolecular reactions and $\text{m}^3 \text{mol}^{-1} \text{s}^{-1}$ for bimolecular reactions. The Arrhenius parameters have been obtained from linear regression of the Arrhenius equation. The uncertainty intervals of the regression correspond to a confidence of 95%. The rate coefficients have been evaluated using the Arrhenius parameters (ARRH). The *ab initio* (AI) rate coefficients can be found in Table 6. The ratios of the Arrhenius and *ab initio* rate coefficients are also reported.

	T range (K)	A	E_a	k_{ARRH}		$\frac{k_{AI}}{k_{ARRH}}$	
				300K	1000K	300K	1000K
R1	300-2000	$5.7 \pm 1.5 \cdot 10^6$	36.0 ± 1.7	$3.1 \cdot 10^0$	$7.5 \cdot 10^4$	1.59	0.72
		$9.9 \pm 1.2 \cdot 10^{13}$	128.6 ± 0.7	$4.0 \cdot 10^{-9}$	$1.9 \cdot 10^7$	1.25	0.90
R2	300-2000	$8.7 \pm 0.8 \cdot 10^9$	32.9 ± 0.5	$1.6 \cdot 10^4$	$1.7 \cdot 10^8$	1.11	0.84
		$2.9 \pm 0.3 \cdot 10^{14}$	128.4 ± 0.6	$1.3 \cdot 10^{-8}$	$5.7 \cdot 10^7$	1.18	0.91
R3	300-2000	$8.4 \pm 3.7 \cdot 10^6$	56.4 ± 3.2	$1.3 \cdot 10^{-3}$	$9.5 \cdot 10^3$	2.68	0.62
		$4.2 \pm 2.0 \cdot 10^6$	74.6 ± 3.4	$4.3 \cdot 10^{-7}$	$5.3 \cdot 10^2$	3.02	0.60
R4	300-2000	$9.4 \pm 2.7 \cdot 10^{10}$	78.3 ± 1.8	$2.2 \cdot 10^{-3}$	$7.6 \cdot 10^6$	1.69	0.75
		$6.5 \pm 1.7 \cdot 10^9$	24.8 ± 1.7	$3.1 \cdot 10^5$	$3.3 \cdot 10^8$	1.66	0.76
R5	300-2000	$1.2 \pm 0.3 \cdot 10^6$	60.8 ± 1.4	$3.1 \cdot 10^{-5}$	$8.0 \cdot 10^2$	1.45	0.76
		$7.2 \pm 2.1 \cdot 10^5$	64.9 ± 1.8	$3.6 \cdot 10^{-6}$	$2.9 \cdot 10^2$	1.64	0.72
R6	600-2000	$4.2 \pm 0.9 \cdot 10^{13}$	290.6 ± 2.1		$2.8 \cdot 10^{-2}$		0.83
		$8.6 \pm 2.9 \cdot 10^5$	263.5 ± 3.7		$1.5 \cdot 10^{-8}$		0.74
R7	600-2000	$1.1 \pm 0.1 \cdot 10^{13}$	241.4 ± 1.2		$2.7 \cdot 10^0$		0.85
		$5.8 \pm 1.3 \cdot 10^5$	148.5 ± 2.2		$1.0 \cdot 10^{-2}$		0.82
R8	300-2000	$3.0 \pm 0.8 \cdot 10^{11}$	182.9 ± 1.6	$4.3 \cdot 10^{-21}$	$8.4 \cdot 10^1$	1.64	0.78
		$4.6 \pm 2.0 \cdot 10^3$	172.3 ± 3.0	$4.6 \cdot 10^{-27}$	$4.6 \cdot 10^{-6}$	2.40	0.57

Radical addition reactions to unsaturated bonds often contain an early transition state, with a low energy barrier, which is seen in the simple Arrhenius parameters of the reactions **R1** and **R2**, i.e. 36.0 and 32.9 kJ mol^{-1} respectively as shown in Table 7. The reactions are exothermic, and the reverse reactions have thus a higher barrier. Tunneling coefficients for these reactions are typically low, as seen in Table 8. The pre-exponential factor of the reverse reaction of **R2** is several orders of magnitude higher than the forward one, which reflects the strong loss in entropy when closing an open carbon chain to a cyclic structure. The kinetics of **R1**, shown in Figure 14a

in the reverse direction, show a satisfactory agreement with the experimental data over a wide temperature range. The rate coefficients are within a factor of 1.75 from the experimental data of Pappic and Laidler⁵⁵ and Mintz and Le Roy⁵⁶, the deviations to the data of Tsang et al.⁵⁷ is higher, with an average deviation of a factor of 4, and a maximum of 4.91 at 800K. The rate coefficients are also comparable to other *ab initio* studies such as the one from Sabbe et al.⁵

Table 8: Tunneling coefficients at 300 and 1000K, and numbers of single events of the 8 reactions.

	Reactions	κ (300K)	κ (1000K)	n_e forward	n_e reverse
R1		1.17	1.02	8	2
R2		1.22	1.02	4	4
R3		37.14	1.29	12	8
R4		25.08	1.31	2	2
R5		1.15	1.01	2	2
R6		2542444.93	1.48	2	4
R7		112.78	1.24	1	4
R8		5.49	1.12	3	6

Intramolecular additions, and their reverse β -scission reactions, are important in the pyrolysis and combustion of cyclic species, and also for the production of aromatic compounds, and PAH formation. An accurate understanding of these reactions is thus crucial for modeling and optimization of fuel blends and combustion engines. One of the simplest reactions belonging to this family is the exo-addition of the 1-hexen-6-yl radical to the cyclohexyl radical (**R2**), of which the rate coefficients are shown in Figure 14b. Wang et al.⁵⁸ also used CBS-QB3 calculations with 1D hindered rotor potential to calculate the rate coefficients. In this reaction, where 4 internal rotations are frozen in the transition state and product, the use of the 1D hindered rotor approach has a major impact on the calculated rate coefficient. This maximum

ratio between Wang et al. and this work is 2.6, which is obtained at low temperature. Sirjean et al.⁵⁹ reports rate coefficients that are more than an order of magnitude higher. The experimental data of Handford-Styring and Walker⁶⁰ deviate a factor of 0.8 to 1.0, which is in very good agreement of the *ab initio* rate coefficients.

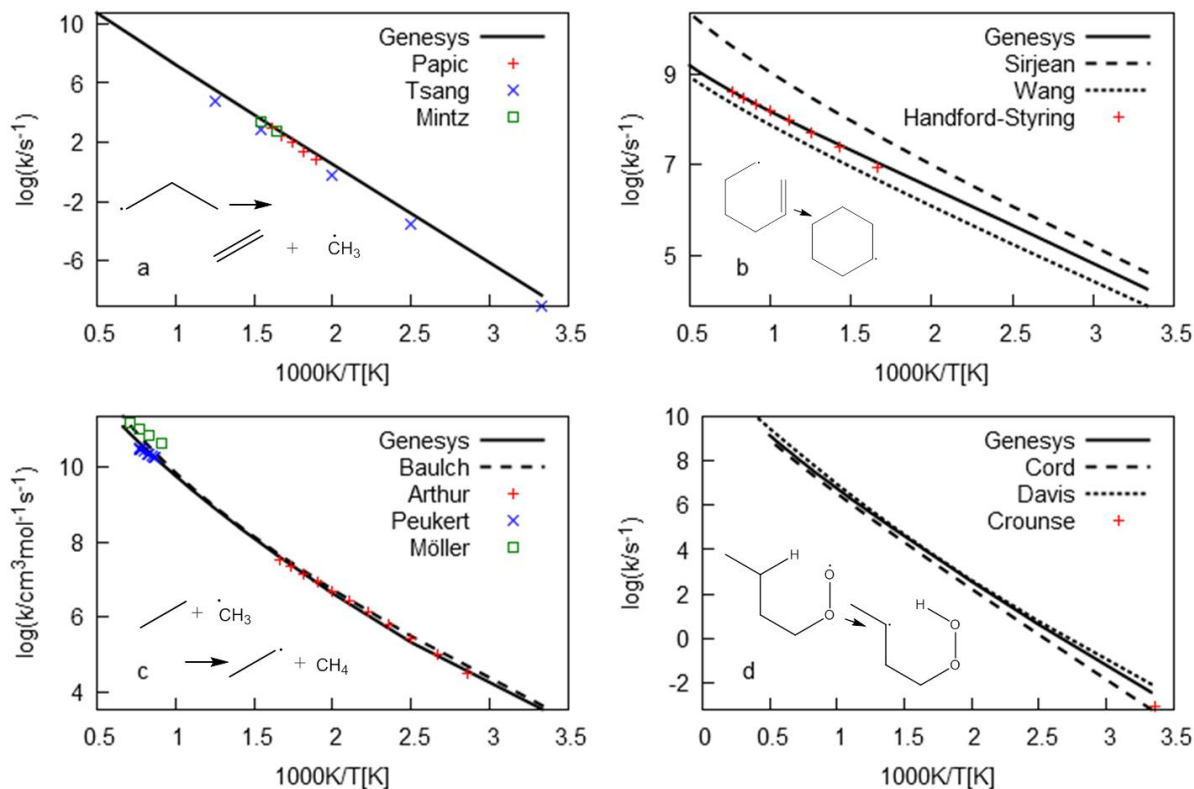


Figure 14: Rate coefficients calculated by Genesys for (a) the β -scission of the *n*-propyl radical to a methyl radical and ethylene (reverse of R1) compared to experimental data from Papis and Laidler⁵⁵, Tsang et al.⁵⁷, Mintz and Le Roy⁵⁶, (b) the intramolecular addition of the 1-hex-5-enyl radical to the cyclohexyl radical (R2) compared to experimental data from Sirjean et al.⁵⁹, Wang et al.⁵⁸, and Handford-Styring and Walker⁶⁰, (c) the hydrogen abstraction by a methyl radical from ethane (R3) to literature data from Baulch et al.⁶¹, Arthur and Bell⁶², Peukert⁶³, and Möller et al.⁶⁴, and (d) the ROO to QOOH reaction (R4) compared to literature data from Cord et al.⁶⁵, Davis and Francisco⁶⁶, and Crouse et al.⁶⁷.

Hydrogen abstraction reactions, such as **R3** and **R4**, have high tunneling coefficients at low temperatures, as shown in Table 8. Regression of a modified Arrhenius expression for these reactions thus leads to high n values, and an accompanied higher uncertainty on n and on the pre-exponential factor. **R3** is compared to experimental data, where Genesys deviates up to a factor of 2 compared to the work of Peukert et al.⁶⁸, a factor of 1.3 compared to the work of Arthur and Bell⁶², and a factor of 3.3 compared to the work of Möller et al.⁶⁴. Furthermore, the rate coefficients reported by Baulch et al.⁶¹, who evaluated a large number of reactions and

summarized their Arrhenius parameters, are within a factor of 1.9 of the results obtained with Genesys, the average deviation between 300 and 1500K only amounts to 1.4.

The ROO to QOOH reaction, which is an intramolecular hydrogen abstraction from an alkylperoxy radical to a hydroperoxyalkyl radical, has a large impact on the combustion and oxidation of hydrocarbons.⁶⁹ **R4** belongs to this reaction family. Its rate coefficients are shown in Figure 14d. Since it is very difficult to isolate and measure the reactants and products of this reaction family, direct measurements of the rate coefficients are lacking. The rate coefficients of **R4** are thus compared to other *ab initio* obtained rate coefficients, and an excellent agreement is achieved.

Homolytic substitutions, such as the reaction of ethane thiol with a methyl radical yielding methane thiol and an ethyl radical (**R5**), are encountered during the thermal decomposition of sulfur containing compounds. The process is not well understood yet, and many experimental and theoretical studies are still required to establish a set of rate coefficients for this reaction family. Since no comparison is possible, Figure 15a only shows the forward and reverse rate coefficient calculated with Genesys. These reactions occur at a rate of similar order of magnitude as other bimolecular reactions such as intermolecular hydrogen abstractions (**R3**), c.f. Table 6, and can thus be expected to indeed play an important role in thermal decomposition of thiols, sulfides, or other sulfur-containing compounds. This has already been shown for substitutions by hydrogen radicals.⁵⁰ A tool such as Genesys is of great value to identify important pathways and subsequently generate a set of rate coefficients for reaction families of which the kinetics are not well known.

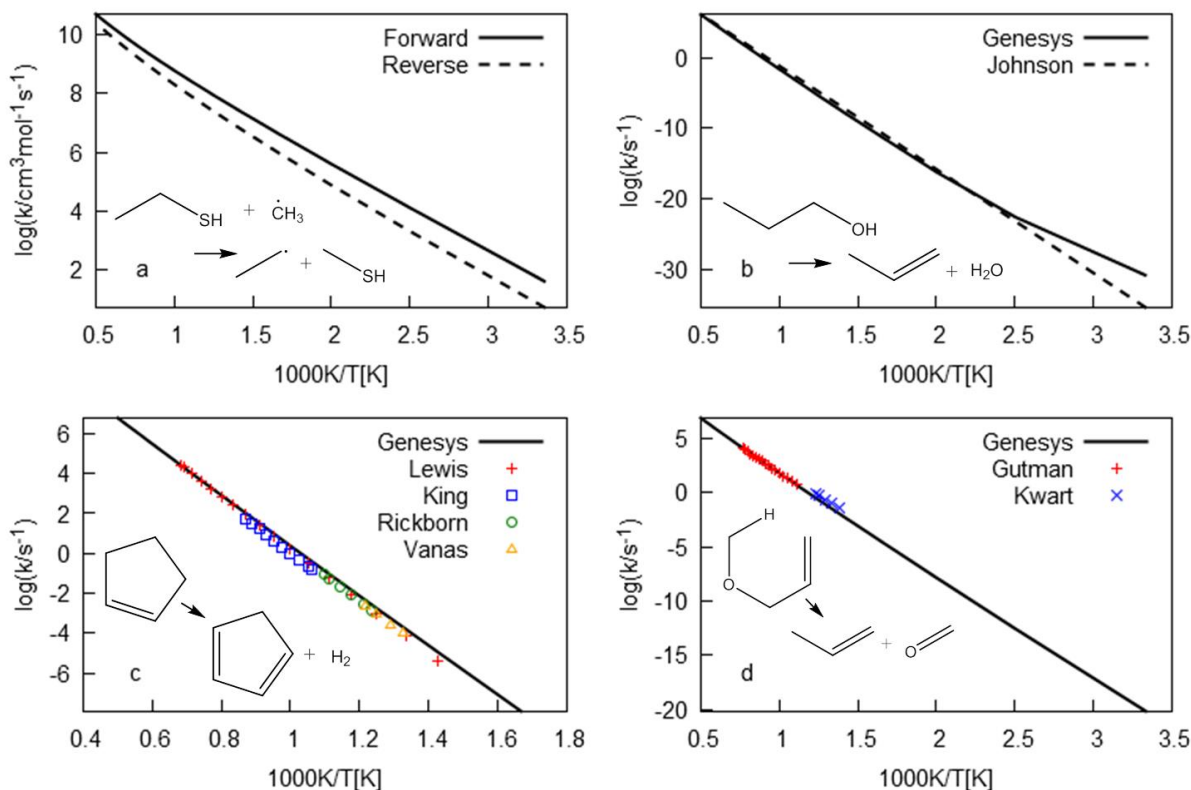


Figure 15: Rate coefficients calculated by Genesys for (a) the homolytic substitution from ethane thiol by a methyl radical to methane thiol and an ethyl radical (R5), (b) the dehydration of propanol (R6) compared to literature data from Johnson et al.⁷⁰, (c) the H_2 elimination of cyclopentene to cyclopentadiene (R7) compared to literature data from Lewis et al.⁷¹, King⁷², Rickborn et al.⁷³, and Vanas and Walters⁷⁴, and (d) the retro-ene reaction from 3-methoxy-1-propene to formaldehyde and propylene (R8) compared to literature data from Gutman et al.⁷⁵ and Kwart et al.⁷⁶.

The last three reactions are more complex in the determination of their transition state as they involve reaction centers containing 4 to 6 polyvalent atoms and exhibit a cyclic structure in the transition state. **R6** and **R7** have very high tunneling coefficients at low temperature, as mentioned above. Their rate coefficients can thus only be regressed into an (modified) Arrhenius equation at higher temperatures, starting at 600K. Although both reactions have low rate coefficients at 1000K, order of magnitudes 1 s^{-1} or lower, these reactions do not involve radicals and are often important for the initial decomposition of molecules before the radical chemistry becomes the main decomposition route. The same applies for **R8**. The rate coefficients of **R6**, the direct elimination of water from a primary alcohol, is shown in Figure 15b. The rate coefficients compare well to the data of Johnson et al.⁷⁰, in which the deviation factor amounts up to 2, except at low temperatures, where tunneling is important.

The direct elimination of hydrogen gas from an unsaturated 5 membered ring structure, as shown in Figure 15c, has already been experimentally observed in the 1940's⁷⁴. To the best of the authors knowledge, no *ab initio* rate coefficients are reported in literature. Genesys calculated the rate coefficient, which is in very good agreement to several experimental studies, the deviation factor does not exceed 2.

Finally, a retro-ene reaction (**R8**) has been calculated by Genesys. The reaction rate coefficients reported by Gutman et al.⁷⁵ are within a factor of 1.1, the ones of Kwart et al.⁷⁶ are up to a factor of 2.4 higher.

Similarly to thermodynamic values, the rate coefficients obtained by Genesys have a promising accuracy when compared to literature data. Currently, minor manual interventions are necessary, which can and will be eliminated in the future. These calculations can be used to automatically fill databases with *ab initio* reaction rate coefficients or to re-calculate rate coefficients when crucial pathways in a kinetic model have been identified, which are often based on less accurate calculation procedures. Furthermore, this opens the door to rate-based kinetic model generation, which is sensitive to the accuracy of the reaction rate coefficients.

3.5 Conclusions

This chapter introduces automation procedures for on-the-fly *ab initio* calculations in the framework of the automatic kinetic model generator Genesys. Genesys represents species and transition states as graph structures, and provides 3D coordinates of the constituting atoms to start *ab initio* calculations. A distance geometry algorithm has been implemented in Genesys to allow the generation of an initial 3D structure. For transition states, the user knowledge of the reaction family at hand, translated in the 3D coordinates of the reactive atoms near the saddle point, is used to estimate the coordinates. The 3D structures are then used in Gaussian in several steps to calculate the necessary thermochemical data for automatic kinetic model generation. An extensive conformational search is employed to ensure that the lowest energy conformer is used. Stereochemistry is explicitly accounted for in the *ab initio* calculations and thermochemical and kinetic data because the thermodynamic properties and rate coefficient can differ significantly depending on the isomer. Rotational modes are treated with a 1D hindered rotor approach.

The methodology is used to calculate the thermodynamic parameters of a set of species and the rate coefficients of 8 reactions. Comparison to manually obtained *ab initio* data and experimental data shows an excellent agreement, illustrating the capabilities of Genesys to automatically generate an accurate set of data readily usable in automatic kinetic model generation. Only minor manual interventions were needed, which will be eliminated in the future.

On the longer term, to generate kinetic models at larger scale with Genesys first the thermodynamic calculations need to be improved. For example, the generation of new BAC values from a large experimental dataset will increase the accuracy of the standard enthalpies of formation.

3.6 References

1. Evans MG, Polanyi M. Further considerations on the thermodynamics of chemical equilibria and reaction rates. *Transactions of the Faraday Society*. 1936;32(0):1333-1360.
2. Blowers P, Masel RI. Engineering approximations for activation energies in hydrogen transfer reactions. *Aiche Journal*. 2000;46(10):2041-2052.
3. Saeys M, Reyniers M-F, Marin GB, Van Speybroeck V, Waroquier M. Ab Initio Calculations for Hydrocarbons: Enthalpy of Formation, Transition State Geometry, and Activation Energy for Radical Reactions. *The Journal of Physical Chemistry A*. 2003;107(43):9147-9159.
4. Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Ab initio group contribution method for activation energies for radical additions. *Aiche Journal*. 2004;50(2):426-444.
5. Sabbe MK, Reyniers MF, Van Speybroeck V, Waroquier M, Marin GB. Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors. *Chemphyschem*. 2008;9(1):124-140.
6. Sabbe MK, Vandeputte AG, Reyniers MF, Waroquier M, Marin GB. Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions. *Physical Chemistry Chemical Physics*. 2010;12(6):1278-1298.
7. Vandeputte AG, Sabbe MK, Reyniers M-F, Marin GB. Kinetics of alpha hydrogen abstractions from thiols, sulfides and thiocarbonyl compounds. *Physical Chemistry Chemical Physics*. 2012;14(37):12773-12793.
8. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos N, Marin GB. Kinetic Modeling of alpha-Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Carbon-Centered Radicals. *Chemphyschem*. 2014;15(9):1849-1866.
9. Zádor J, Najm HN. KinBot 1.0: A code for automatic PES exploration. Paper presented at: 8th U. S. National Combustion Meeting2013.
10. Suleimanov YV, Green WH. Automated Discovery of Elementary Chemical Reaction Steps Using Freezing String and Berny Optimization Methods. *Journal of Chemical Theory and Computation*. 2015;11(9):4248-4259.
11. Wales DJ. Energy landscapes: calculating pathways and rates. *International Reviews in Physical Chemistry*. 2006;25(1-2):237-282.
12. Schlegel HB. Optimization of equilibrium geometries and transition structures. *Journal of Computational Chemistry*. 1982;3(2):214-218.
13. Schlegel HB. Estimating the hessian for gradient-type geometry optimizations. *Theoretica chimica acta*. 1984;66(5):333-340.
14. Peng CY, Ayala PY, Schlegel HB, Frisch MJ. Using redundant internal coordinates to optimize equilibrium geometries and transition states. *Journal of Computational Chemistry*. 1996;17(1):49-56.
15. Henkelman G, Uberuaga BP, Jonsson H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *Journal of Chemical Physics*. 2000;113(22):9901-9904.
16. Henkelman G, Jónsson H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *The Journal of Chemical Physics*. 2000;113(22):9978-9985.

17. Weinan E, Ren WQ, Vanden-Eijnden E. String method for the study of rare events. *Physical Review B*. 2002;66(5).
18. Zimmerman PM. Automated discovery of chemically reasonable elementary reaction steps. *Journal of Computational Chemistry*. 2013;34(16):1385-1392.
19. Zimmerman PM. Navigating molecular space for reaction mechanisms: an efficient, automated procedure. *Molecular Simulation*. 2015;41(1-3):43-54.
20. Broadbelt LJ, Stark SM, Klein MT. Computer generated reaction networks: on-the-fly calculation of species properties using computational quantum chemistry. *Chemical Engineering Science*. 1994;49(24):4991-5010.
21. Magoon GR, Green WH. Design and implementation of a next-generation software interface for on-the-fly quantum and force field calculations in automated reaction mechanism generation. *Computers & Chemical Engineering*. 2013;52:35-45.
22. Bhoorasingh PL, West RH. Transition state geometry prediction using molecular group contributions. *Physical Chemistry Chemical Physics*. 2015;17(48):32173-32182.
23. Vandewiele NM, Van Geem KM, Reyniers MF, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chemical Engineering Journal*. 2012;207:526-538.
24. Van de Vijver R, Vandewiele NM, Bhoorasingh PL, Slakman BL, Seyedzadeh Khanshan F, Carstensen H-H, Reyniers M-F, Marin GB, West RH, Van Geem KM. Automatic Mechanism and Kinetic Model Generation for Gas- and Solution-Phase Processes: A Perspective on Best Practices, Recent Advances, and Future Challenges. *International Journal of Chemical Kinetics*. 2015;47(4):199-231.
25. Mizutani MY, Nakamura K, Ichinose T, Itai A. Starting point to molecular design: Efficient automated 3D model builder Key3D. *Chemical & Pharmaceutical Bulletin*. 2006;54(12):1680-1685.
26. Vainio MJ, Johnson MS. Generating conformer ensembles using a multiobjective genetic algorithm. *Journal of Chemical Information and Modeling*. 2007;47(6):2462-2474.
27. Crippen GM. Chemical distance geometry: Current realization and future projection. *Journal of Mathematical Chemistry*. 1991;6(1):307-324.
28. Wertz DH. Distance geometry and conformational calculations, G. M. Crippen. Research Studies Press, Chichester, New York, 1981, 58 pp. *Journal of Computational Chemistry*. 1982;3(4):603-603.
29. Havel TF, Kuntz ID, Crippen GM. The theory and practice of distance geometry. *Bulletin of Mathematical Biology*. 1983;45(5):665-720.
30. Havel TF. Distance Geometry: Theory, Algorithms, and Chemical Applications. *Encyclopedia of Computational Chemistry*: John Wiley & Sons, Ltd; 2002.
31. Dress AWM, Havel TF. Shortest-path problems and molecular conformation. *Discrete Applied Mathematics*. 1988;19(1):129-144.
32. Halgren TA. Merck molecular force field .1. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*. 1996;17(5-6):490-519.
33. Vandewiele NM, Van De Vijver R, Carstensen H-H, Van Geem KM, Reyniers M-F, Marin GB. Implementation of Stereochemistry in Automatic Kinetic Model Generation. *International Journal of Chemical Kinetics*. 2016;48(12):755-769.
34. *Gaussian 03* [computer program]. Wallingford, CT, USA: Gaussian, Inc.; 2004.
35. Van Speybroeck V. *Ab initio and dynamic methods : A useful tool in the study of chemical reactions*: Center for Molecular Modeling, Ghent University; 2001, Ph. D. Thesis.

36. Sabbe MK, Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *Journal of Physical Chemistry A*. 2005;109(33):7466-7480.
37. Sabbe MK, De Vleeschouwer F, Reyniers MF, Waroquier M, Marin GB. First Principles Based Group Additive Values for the Gas Phase Standard Entropy and Heat Capacity of Hydrocarbons and Hydrocarbon Radicals. *Journal of Physical Chemistry A*. 2008;112(47):12235-12251.
38. East ALL, Radom L. Ab initio statistical thermodynamical models for the computation of third-law entropies. *Journal of Chemical Physics*. 1997;106(16):6655-6674.
39. Ayala PY, Schlegel HB. Identification and treatment of internal rotation in normal mode vibrational analysis. *Journal of Chemical Physics*. 1998;108(6):2314-2325.
40. Linstrom PJ, Mallard WG, eds. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. Gaithersburg MD: National Institute of Standards and Technology; 2016.
41. Paraskevas PD, Sabbe MK, Reyniers M-F, Papayannakos N, Marin GB. Group Additive Values for the Gas-Phase Standard Enthalpy of Formation, Entropy and Heat Capacity of Oxygenates. *Chemistry – A European Journal*. 2013;19(48):16431-16452.
42. *CHEMKIN Release 4.1.1* [computer program]. San Diego, CA, USA: Reaction Design, Inc.; 2007.
43. Pollak E, Pechukas P. Symmetry Numbers, Not Statistical Factors, Should be Used in Absolute Rate Theory and in Bronsted Relations. *Journal of the American Chemical Society*. 1978;100(10):2984-2991.
44. Eckart C. The Penetration of a Potential Barrier by Electrons. *Physical Review*. 1930;35(11):1303-1309.
45. Coote ML, Collins MA, Radom L. Calculation of accurate imaginary frequencies and tunnelling coefficients for hydrogen abstraction reactions using IRCmax. *Molecular Physics*. 2003;101(9):1329-1338.
46. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Kinetic Modeling of α -Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Hydrogen Atoms. *Journal of Physical Chemistry A*. 2014;118(40):9296-9309.
47. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Group Additive Kinetics for Hydrogen Transfer Between Oxygenates. *Journal of Physical Chemistry A*. 2015;119(27):6961-6980.
48. Paraskevas PD, Sabbe MK, Reyniers MF, Marin GB, Papayannakos NG. Group additive kinetic modeling for carbon-centered radical addition to oxygenates and -scission of oxygenates. *Aiche Journal*. 2016;62(3):802-814.
49. Vandewiele NM, Van de Vijver R, Van Geem KM, Reyniers M-F, Marin GB. Symmetry calculation for molecules and transition states. *Journal of Computational Chemistry*. 2015;36(3):181-192.
50. Vandeputte AG, Reyniers M-F, Marin GB. Kinetics of Homolytic Substitutions by Hydrogen Atoms at Thiols and Sulfides. *Chemphyschem*. 2013;14(8):1703-1722.
51. El-Nahas AM, Navarro MV, Simmie JM, Bozzelli JW, Curran HJ, Dooley S, Metcalfe W. Enthalpies of Formation, Bond Dissociation Energies and Reaction Paths for the Decomposition of Model Biofuels: Ethyl Propanoate and Methyl Butanoate. *The Journal of Physical Chemistry A*. 2007;111(19):3727-3739.
52. Dean J. *Lange's Handbook of Chemistry, 14th Edition* 1992.

53. Nunes PM, Agapito F, Costa Cabral BJ, Borges dos Santos RM, Martinho Simões JA. Enthalpy of Formation of the Cyclopentadienyl Radical: Photoacoustic Calorimetry and *ab Initio* Studies. *The Journal of Physical Chemistry A*. 2006;110(15):5130-5134.
54. Kiefer JH, Tranter RS, Wang H, Wagner AF. Thermodynamic functions for the cyclopentadienyl radical: The effect of Jahn–Teller distortion. *International Journal of Chemical Kinetics*. 2001;33(12):834-845.
55. Papic MM, Laidler KJ. Kinetics of the Mercury-Photosensitized Decomposition of Propane. Part II. Reactions of the Propyl Radicals. *Canadian Journal of Chemistry*. 1971;49(4):549-554.
56. Mintz KJ, Le Roy DJ. Kinetics of radical reactions in sodium diffusion flames. *Canadian Journal of Chemistry*. 1978;56(7):941-949.
57. Tsang W. The stability of alkyl radicals. *Journal of the American Chemical Society*. 1985;107(10):2872-2880.
58. Wang K, Villano SM, Dean AM. Reactivity–Structure-Based Rate Estimation Rules for Alkyl Radical H Atom Shift and Alkenyl Radical Cycloaddition Reactions. *The Journal of Physical Chemistry A*. 2015;119(28):7205-7221.
59. Sirjean B, Glaude PA, Ruiz-Lopez MF, Fournet R. Theoretical Kinetic Study of Thermal Unimolecular Decomposition of Cyclic Alkyl Radicals. *The Journal of Physical Chemistry A*. 2008;112(46):11598-11610.
60. Handford-Styring SM, Walker RW. Addition of cyclopentane to slowly reacting mixtures of H₂+ O₂ between 673 and 783 K: reactions of H and OH with cyclopentane and of cyclopentyl radicals. *Journal of the Chemical Society, Faraday Transactions*. 1995;91(10):1431-1438.
61. Baulch DL, Cobos CJ, Cox RA, Esser C, Frank P, Just T, Kerr JA, Pilling MJ, Troe J, Walker RW, Warnatz J. Evaluated Kinetic Data for Combustion Modelling. *Journal of Physical and Chemical Reference Data*. 1992;21(3):411-734.
62. Arthur NL, Bell TN. An evaluation of the kinetic data for hydrogen abstraction from silanes in the gas phase. *Reviews of Chemical Intermediates*. 1978;2(1):37-74.
63. Peukert SL, Labbe NJ, Sivaramakrishnan R, Michael JV. Direct Measurements of Rate Constants for the Reactions of CH₃ Radicals with C₂H₆, C₂H₄, and C₂H₂ at High Temperatures. *The Journal of Physical Chemistry A*. 2013;117(40):10228-10238.
64. Möller W, Mozzhukhin E, Gg. Wagner H. High Temperature Reactions of CH₃. 2. H-Abstraction from Alkanes. *Berichte der Bunsengesellschaft für physikalische Chemie*. 1987;91(6):660-666.
65. Cord M, Sirjean B, Fournet R, Tomlin A, Ruiz-Lopez M, Battin-Leclerc F. Improvement of the Modeling of the Low-Temperature Oxidation of n-Butane: Study of the Primary Reactions. *The Journal of Physical Chemistry A*. 2012;116(24):6142-6158.
66. Davis AC, Francisco JS. *Ab Initio* Study of Hydrogen Migration in 1-Alkylperoxy Radicals. *The Journal of Physical Chemistry A*. 2010;114(43):11492-11505.
67. Crouse JD, Nielsen LB, Jørgensen S, Kjaergaard HG, Wennberg PO. Autoxidation of Organic Compounds in the Atmosphere. *The Journal of Physical Chemistry Letters*. 2013;4(20):3513-3520.
68. Peukert S, Naumann C, Braun-Unkhoff M, Riedel U. Formation of H-atoms in the pyrolysis of cyclohexane and 1-hexene: A shock tube and modeling study. *International Journal of Chemical Kinetics*. 2011;43(3):107-119.

69. Savee JD, Papajak E, Rotavera B, Huang H, Eskola AJ, Welz O, Sheps L, Taatjes CA, Zádor J, Osborn DL. Direct observation and kinetics of a hydroperoxyalkyl radical (QOOH). *Science*. 2015;347(6222):643-646.
70. Johnson MV, Goldsborough SS, Serinyel Z, O'Toole P, Larkin E, O'Malley G, Curran HJ. A Shock Tube Study of n- and iso-Propanol Ignition. *Energy & Fuels*. 2009;23(12):5886-5898.
71. Lewis DK, Bergmann J, Manjoney R, Paddock R, Kalra BL. Rates of reactions of cyclopropane, cyclobutane, cyclopentene, and cyclohexene in the presence of boron trichloride. *The Journal of Physical Chemistry*. 1984;88(18):4112-4116.
72. King KD. Very low-pressure pyrolysis (VLPP) of cyclopentene. *International Journal of Chemical Kinetics*. 1978;10(1):117-123.
73. Rickborn SF, Rogers DS, Ring MA, O'Neal HE. Kinetic and product studies of the thermal decomposition of dimethylsilane in a single-pulse shock tube and in a stirred flow reactor. *The Journal of Physical Chemistry*. 1986;90(3):408-414.
74. Vanas DW, Walters WD. The Thermal Decomposition of Cyclopentene. *Journal of the American Chemical Society*. 1948;70(12):4035-4039.
75. Gutman D, Braun W, Tsang W. Comparison of the thermal and infrared laser induced unimolecular decompositions of allylmethylether, ethylacetate, and isopropylbromide. *Journal of Chemical Physics*. 1977;67(9):4291-4296.
76. Kwart H, Sarner SF, Slutsky J. Mechanisms of thermolytic fragmentation of allyl ethers. I. *Journal of the American Chemical Society*. 1973;95(16):5234-5242.

Chapter 4: Group additivity model for intramolecular hydrogen abstractions

4.1 Abstract

Group additivity is a powerful concept for calculating, amongst others, thermodynamic data and reaction rate coefficients. A novel methodology is introduced to determine a set of ΔGAV° values regressed from data obtained by quantum chemical calculations. The method is demonstrated for intramolecular hydrogen abstraction reactions of alkanes, alkenes and alkynes, for which no group additivity values are currently available. Calculating 448 reaction rate coefficients at the CBS-QB3 level of theory for 1-2 up to 1-7 hydrogen shift reactions allowed the estimation ΔGAV° values for 270 groups. The influence of substituents on (1) the attacking radical, (2) the attacked carbon atom, and (3) the carbon chain between the attacking and attacked reactive atom has been systematically studied. Substituents have been varied between hydrogen atoms and sp^3 , sp^2 and sp hybridized carbon atoms. Hence, the primary contributions to the Arrhenius parameters, i.e. the influence of the ligands of the attacked and attacking carbon atom, as well as the secondary contributions to the Arrhenius parameters, i.e. the influence of the ligands of the carbon chain between the attacking and attacked carbon atom, are accounted for. This group additivity model is applicable to a wide variety of reactions in the 300-1800K temperature range. Correlations for tunneling coefficients have been generated which are complementary to the ΔGAV° to obtain accurate rate coefficients without the need for imaginary

frequencies nor electronic energies of activation. These correlations depend on the temperature and activation energy of the exothermic step. The group additivity model has been successfully applied to a test set of reactions also calculated at the CBS-QB3 level of theory.

4.2 Introduction

On-the-fly *ab initio* calculations for each reaction during automated kinetic model generation is computationally infeasible. One option is that the previously calculated *ab initio* data are retrieved from databases and used in future model generations. The straightforward way to implement this is the extension of the databases with a single entry per reaction and species calculated previously. This database is visited during kinetic model generation and the data is pulled from the database in case of an exact hit between the target molecule (or reaction) and the database molecule (or reaction). Another, more general, solution is the generations of procedures to calculate new thermodynamics or rate coefficients methods based on the *ab initio* data. In this way, the calculations can be constantly improved by performing new *ab initio* calculations. Thermodynamics are mainly based on group additivity procedures, introduced by Benson¹, and extended by many researchers. For kinetic parameters, a variety of methods is available, as described in Chapter 1, each with different accuracies, theoretical bases, data requirements, etc. One of the methods that exhibits a similar accuracy compared to the direct use of *ab initio* calculations, is group additivity²⁻¹¹, i.e. ΔGAV° values. This method has been successfully applied to many reactions such as hydrogen abstraction reactions, radical addition reactions to multiple bonds, and homolytic substitution reactions in hydrocarbon, oxygenates and sulfur compounds.

One of the important reaction families during combustion, steam cracking or pyrolysis of hydrocarbons is the intramolecular hydrogen abstraction family, also called hydrogen migrations or hydrogen shifts (H-shifts). This isomerization reaction happens through a cyclic transition state in which a hydrogen atom migrates from a carbon atom to a radical site. This is thus a monomolecular reaction with a single product species, and the reverse reaction belongs to the same reaction family. The distance between the attacking radical and the attacked carbon atom is crucial to obtain kinetic parameters. In this work, the following reactions are considered: reactions between neighboring carbon atoms and reactions with one up to 5 carbon atoms

between the attacking and attacked carbon atom. These reactions are noted 1-2, 1-3, 1-4, 1-5, 1-6 and 1-7 intramolecular hydrogen abstractions. The number 1 is the position of the attacking radical center and the numbers 2 to 7 are the positions of the attacked carbon atom. A 1-2 H-shift proceeds through a three-membered cyclic transition state, a 1-3 H-shift through a four-membered cyclic transition state, and so on. This is illustrated in Figure 1.

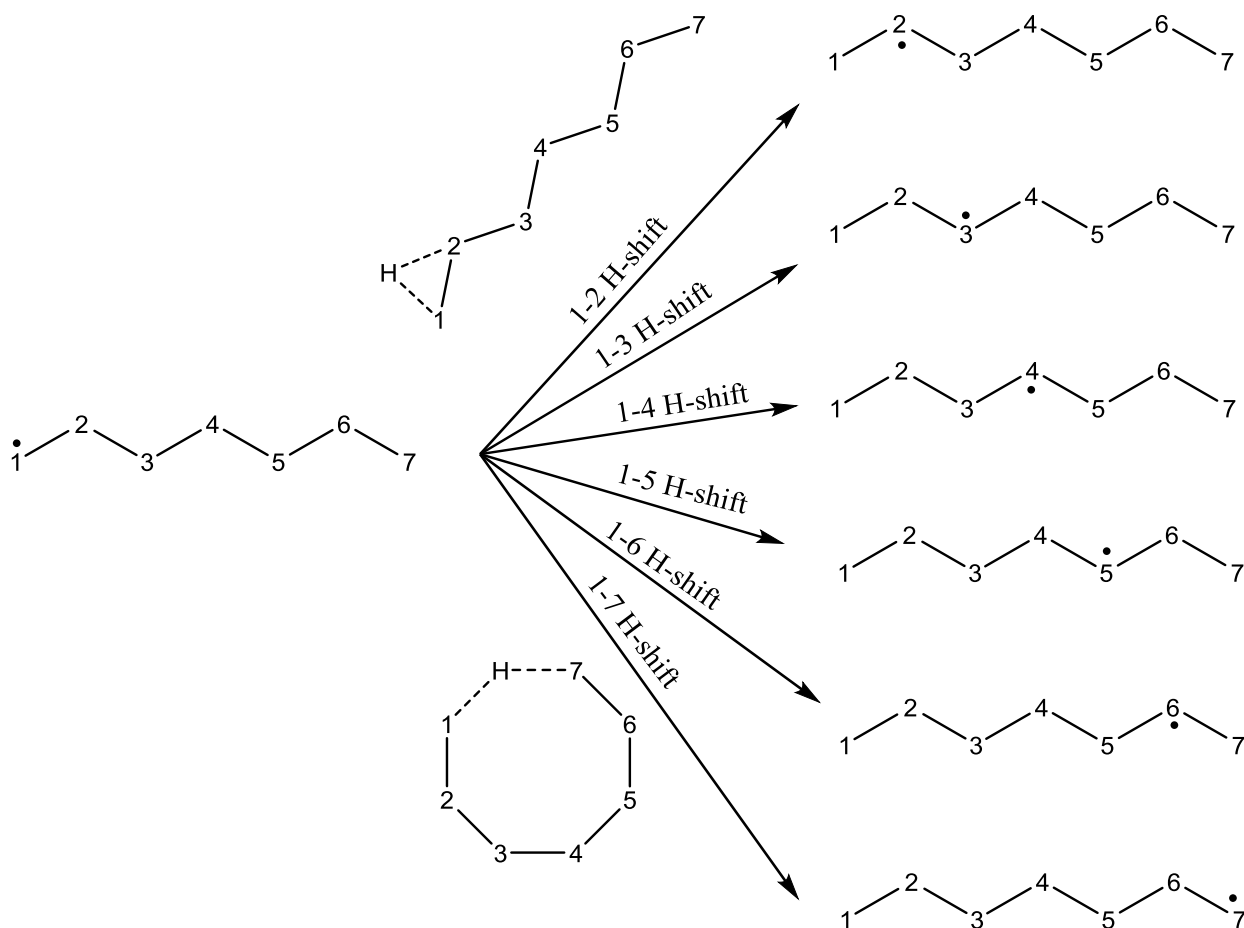


Figure 1: Illustration of 1-2 to 1-7 intramolecular hydrogen abstraction reactions (H-shifts). The transition state of a 1-2 H-shift proceeds through a three-membered cyclic structure, a 1-7 H-shift has an eight-membered cyclic transition state. The numbers correspond to the position of the carbon atoms, the hydrogen atoms have not been drawn for the sake of clarity, except in the transition state structures: the migrating hydrogen atom is visualized.

Considerable work has already been done to develop procedures to calculate the rate coefficients of intramolecular hydrogen abstractions. Benson¹² introduced many calculation methods both for thermodynamic as well as kinetic parameters. Activation energies of intermolecular abstraction reactions were modelled using linear-free energy relationships.¹² This implies that to obtain the activation energies of H-shifts, the activation energy of an intermolecular hydrogen abstraction

reaction is calculated and augmented with a contribution for the ring strain introduced by the cyclic transition state.

During the development of a n-heptane oxidation kinetic model, Curran et al.¹³ included intramolecular hydrogen abstraction reactions. The method for kinetics calculations is similar to Benson's approach. The kinetics were calculated using a linear free-energy method, using the activation energy of a similar bimolecular abstraction, together with a ring strain correction and the reaction enthalpy. The pre-exponential factors were obtained from transition state theory by calculating the difference in entropy between the radical and the transition state. This method was also employed for other kinetic models.^{14,15}

Zheng and Truhlar¹⁶ calculated the 1-4 hydrogen transfer in the 1-pentyl radical and the 1-4 and 1-5 hydrogen transfers in the 1-hexyl radical using variational transition state theory with multidimensional tunneling. The resulting rate coefficients agree well with experimental data and the tunneling coefficients are able to accurately describe the rate coefficients at low temperatures.

The barriers for hydrogen shift reactions in alkyl, alkenyl and oxoallylic radicals have been investigated by Hayes and Burgess¹⁷ using the composite *ab initio* methods G3MP2B3 and G3B3. A reduction in reaction barrier of 10 and 20 kJ mol⁻¹ has been found for secondary and tertiary radical formation, respectively, compared to the formation of a primary radical, which is not influenced by the transition state ring size. The formation of alkenyl and oxoalkyl radicals also exhibits lower barriers compared to their alkyl radical analogues.

Ratkiewicz and coworkers^{18,19} employed the Reaction Class Transition State Theory (RC-TST) to obtain calculation procedures for 1-4 and 1-5-migration reactions in alkyl radicals. The calculation procedures allowed to approximate the rate coefficients by distinguishing between primary, secondary and tertiary carbon atoms of which a hydrogen atom is abstracted. Furthermore, Ratkiewicz et al.²⁰ calculated the rate coefficients of thermo-neutral 1-3 up to 1-6 hydrogen migrations and used those results to successfully calculate the rate coefficients of hydrogen migration reactions of non-thermo-neutral reactions.

Quantum tunneling plays a prominent role in hydrogen migration reactions at lower temperature. Using n-heptyl as model molecule, Sirjean et al.²¹ computed tunneling coefficients using a multi-dimensional treatment (small curvature tunneling (SCT) approximations^{22,23}) and compared it to

three one-dimensional methods: Wigner²⁴, Skodje and Truhlar²⁵, and Eckart²⁶ tunneling. In general, Sirjean and coworkers found that Eckart tunneling yields the most accurate results compared to SCT approximations. Neglecting corner cutting, compensated by a narrower potential energy curve results in this apparent accuracy of Eckart tunneling coefficients. This conclusion should thus be used with care; other reaction families do not necessarily show this behavior.

Davis and Francisco²⁷ studied all the possible intramolecular hydrogen abstractions in C2 to C7 n-alkyl radicals and assessed the differences in rate coefficient when varying the location of the abstraction site relative to the terminal carbon atom. The influence of equatorial or axial transition states was also investigated. In another article, Davis and Francisco²⁸ calculated thermodynamic properties and rate coefficients for H-shifts in methylalkyl radicals to show the influence of methyl groups either on the transition state ring structure or immediately outside of it.

Resonance stabilization can play an important role in intramolecular hydrogen abstraction reactions as shown by Wang et al.²⁹. Reactions between an alkyl radical and an allyl radical, or between two allyl radicals were studied as the CBS-QB3 level of theory, including 1-2 up to 1-7 shifts. Wang et al.³⁰ also reported reactivity-structure based on calculation rules for 1-2 up to 1-7 hydrogen shifts in saturated alkyl radicals.

Bian et al.³¹ studied 3 types of unimolecular reactions for the 2-octenyl and 3-octenyl radicals: intramolecular additions, internal hydrogen migrations and bond dissociations. The presence of a double bond strongly affects the isomerization rate through internal hydrogen migration reactions. The formation of resonantly stabilized allylic radicals is favored, both when the double bond is a part of the transition state ring structure, as well as with a saturated transition state ring structures. Other hydrogen migration reactions that bypass the double bond can be neglected. For long-chain alkenyl radicals, 1-4 and 1-5-migrations are significantly faster, up to two orders of magnitude, compared to 1-7 and 1-8 migrations. At temperatures above 1600K, 1-2 and 1-3-migrations become more important due to entropy effects.

The concept of group additivity has, since its introduction by Benson^{1,12,32} in the 1950's, widely been used in the chemical community to calculate properties of single compounds or mixtures.

This concept originates from the idea that a property of a molecule can be broken down into separate contributions from each group in the molecule. A group has originally been defined as a central polyvalent atom and its immediate ligands. This led to the notation $X-(A)_k(B)_l(C)_n(D)_m$ for the central atom X with k ligands of type A, l of type B, and so forth. By summing the contributions of all groups in a molecule, the property of that molecule can be calculated. Of course, several contributions in a molecule cannot be comprised in central atoms and their neighbors only. Interactions reaching further in a molecule need to be accounted for as well. In the framework of kinetic modelling, where thermodynamic parameters of molecules are indispensable, these interactions are for example non-nearest neighbor interactions (NNIs), ring strain corrections (RSCs), or symmetry corrections to the entropy. For kinetics, contributions such as the number of single events and tunneling coefficients also need a separate treatment as complement to the group additivity scheme. In the past two decades, alongside the development of large kinetic models, calculation of kinetic parameters gained significant interest. Group additivity has become a prominent procedure for kinetic parameters. Several definitions of groups can be used for kinetic data. Sumathi et al.³³⁻³⁵ introduced so-called supergroups, which consist of a central reactive moiety instead of a single central atom. The groups are defined on the level of the transition state topology. Marin and coworkers²⁻¹¹ extended the concept of group additivity for kinetic data by defining the groups as the reactive atoms in the reactants. In this way, the groups are still defined as one central atom with its ligands.

This chapter introduces a group additive model for intramolecular hydrogen abstractions in hydrocarbon radicals. The influence of ligands of the attacking and attacked carbon atom (primary contributions) on the Arrhenius parameters are studied, together with the contributions of substituents on the carbon chain between the attacking and attacked carbon atom (secondary contributions) to the pre-exponential factor and activation energy. This chain is of significant importance in this reaction family, because it composes the cyclic structure in the transition state. A set of 60 reactions to generate the primary contributions, together with 81 reactions for the secondary contributions, have been automatically calculated at the CBS-QB3 level of theory, with a systematic conformational search and one-dimensional (1D) hindered rotor potentials. A test set of 108 reactions has been calculated at the same level of theory to test the approach.

4.3 Methodology

4.3.1 *Ab initio* calculations

An extensive description of the theoretical methods to automatically obtain *ab initio* rate coefficient can be found in Chapter 3, section 3. In brief, each reaction is first generated based on the reactant topology, building the connectivity of the transition state and product structure. From the topology, initial three-dimensional (3D) coordinates of the atoms in the reactant, transition state and product are generated using an algorithm called “distance geometry”. The coordinates are refined using the Gaussian09 suite of programs³⁶. This refinement is done in several steps, including a fast, semi-empirical pre-optimization, a conformational search, an optimization at high level of theory and the calculation of 1D hindered rotor potentials. All these steps are automatically generated and post-processed with the Genesys software.³⁷ The results are used to extract high-pressure limit rate coefficients for temperatures between 300 and 1800K.

Many of the reactants, products and transition states calculated have multiple equivalent hydrogen atoms, and can exhibit external rotational symmetry or optical isomers. To account for this, the number of single events is calculated for each reaction, which is done automatically in Genesys. The symmetry is calculated using the automorphism group of the connectivity graph of the molecule.³⁸ Optical isomers are perceived as reported by Vandewiele et al.³⁹. Tunneling plays an important role in hydrogen abstractions reactions. Similarly to previous work^{18-21,31}, the Eckart method²⁶ is chosen to calculate tunneling coefficients.

For the generation of initial 3D coordinates of the atoms in the transition state, Genesys requires the user to supply the coordinates of the reactive atoms, see Chapter 3, section 3.2. For intramolecular hydrogen abstraction reactions, three coordinates are sufficient. First, two bond lengths are necessary, i.e. the two C-H bond lengths between the hydrogen that is abstracted and the attacking and attacked carbon atom. Secondly, the C-H-C angle between the same three reactive atoms is needed. Because the reverse reaction to an intramolecular abstraction reaction belongs to the same reaction family, it is assumed that the two C-H bond lengths are equal. It is obvious that the geometry of a transition state depends on its ring size, i.e. whether the reaction is a 1-2, 1-3, etc. shift. The initial estimates for the C-H bond lengths are all set to 135pm. However, six different angles were defined, as shown in Table 1.

Table 1: Coordinates input for the initial generation of 3D coordinates for the transition states and the minimum and maximum final values.

	C-H bond length (pm)			C-H-C angle (°)		
	Initial value	Minimum final value	Maximum final value	Initial value	Minimum final value	Maximum final value
1-2 shift	135	123	143	68	67	71
1-3 shift	135	133	153	103	103	106
1-4 shift	135	127	154	132	132	135
1-5 shift	135	124	153	152	152	157
1-6 shift	135	122	151	165	162	171
1-7 shift	135	126	152	168	165	177

4.3.2 Group additivity model

Marin and coworkers²⁻¹¹ introduced a group additive method to calculate the pre-exponential factor and activation energies of reactions belonging to a reaction family. Rate coefficients for monomolecular reactions such as intramolecular hydrogen abstractions can be calculated from Eq. 4.1, with $\Delta^\ddagger H$ the standard enthalpy of activation and $\Delta^\ddagger S$ the standard activation entropy.

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta^\ddagger H - T\Delta^\ddagger S}{RT}\right) \quad \text{Eq. 4.1}$$

The activation entropy can be calculated by summing two contributions: a symmetry-independent entropy contribution and a term depending on the symmetry numbers σ and the numbers of optical isomers n_{opt} of the reactant and transition state, c.f. Eq. 4.2 and Eq. 4.3. The number of single events n_e has been defined as the reaction path degeneracy by Pollak and Pechukas⁴⁰.

$$\Delta^\ddagger S = \Delta^\ddagger \tilde{S} + R \ln(n_e) \quad \text{Eq. 4.2}$$

$$n_e = \frac{n_{\text{opt},\ddagger} \cdot \sigma_r}{n_{\text{opt},r} \cdot \sigma_\ddagger} \quad \text{Eq. 4.3}$$

This formulation of the activation entropy can be substituted in Eq. 4.1. By then dividing both sides by the number of single events, the so-called single-event rate coefficient is obtained.

$$\frac{k}{n_e} = \tilde{k} = \frac{k_B T}{h} \exp\left(-\frac{\Delta^\ddagger H - T\Delta^\ddagger \tilde{S}}{RT}\right) \quad \text{Eq. 4.4}$$

The Arrhenius activation energy can be calculated from the rate coefficient:

$$E_a = RT^2 \frac{\partial}{\partial T} \ln k = \Delta^\ddagger H + RT \quad \text{Eq. 4.5}$$

The rate coefficient can thus be rewritten into:

$$\tilde{k} = \frac{k_B T}{h} \exp\left(\frac{\Delta^\ddagger \tilde{S}}{R}\right) \exp(1) \exp\left(-\frac{\Delta^\ddagger H + RT}{RT}\right) \quad \text{Eq. 4.6}$$

The last exponential factor contains the exponential part of the Arrhenius equation, the other factors can be identified as the single-event pre-exponential factor.

Obtaining the Arrhenius parameters of a reaction thus only depends on the activation enthalpy $\Delta^\ddagger H$ and activation entropy $\Delta^\ddagger \tilde{S}$. The enthalpy of formation and the standard entropy of a stable molecule can be approximated using group additivity.^{1,12,32} Marin and coworkers²⁻¹¹ and Green and coworkers³³⁻³⁵ extended this concept to transition states, which allows to calculate the standard enthalpy of formation and the entropy by summing group additive values and optionally non-nearest neighbor interactions, as shown in Eq. 4.7 and Eq. 4.8.

$$\Delta_f H_T^o(TS) = \sum_i GAV_H(C_i^{TS}) + \sum_j NNI_{H,j} \quad \text{Eq. 4.7}$$

$$\tilde{S}_T^o(TS) = \sum_i GAV_{\tilde{S}}(C_i^{TS}) + \sum_j NNI_{\tilde{S},j} \quad \text{Eq. 4.8}$$

By defining ΔGAV as the difference between a group additive value in the transition state and the group additive value in the reactant, corresponding to the same polyvalent atom, the activation energy and the activation entropy can be modelled as:

$$E_a(T) = \sum_i \Delta GAV_{E_a}(C_i) + \sum_j \Delta NNI_{E_a,j} + RT \quad \text{Eq. 4.9}$$

$$\Delta^\ddagger \tilde{S}^o(T) = \sum_i \Delta GAV_{\Delta \tilde{S}}(C_i) + \sum_j \Delta NNI_{\Delta \tilde{S},j} \quad \text{Eq. 4.10}$$

Each group in the reactant that is not part of the reactive moiety, i.e. the polyvalent atoms that do not change in connectivity throughout the reaction, and of which the immediate neighbors also do not change in connectivity, has the same group additive values in the transition state. For these groups, ΔGAV is zero.

Since only transition state specific groups need to be accounted for, only the direct surroundings of the transition states need to have group additive values assigned to. It is thus more practical to

introduce a reference reaction. The activation energy can be calculated using Eq. 4.11, in which $\Delta GAV_{E_a}^o = \Delta GAV_{E_a} - \Delta GAV_{E_a,ref}$. The index i iterates over the groups that are transition state specific, which depends on the reaction family. For intramolecular abstraction reactions, the attacking and attacked carbon atoms and the carbon chain connecting them are the groups that are considered to influence the kinetics. It is initially assumed that other non-nearest neighbor interactions can be neglected.

$$E_a(T) = E_{a,ref}(T) + \sum_i \Delta GAV_{E_a}^o(C_i) \quad \text{Eq. 4.11}$$

In a similar fashion, the single-event pre-exponential factor can be calculated according to Eq. 4.12. The log notation is used for the common logarithm, i.e. the logarithm with base 10, throughout this chapter.

$$\log \tilde{A}(T) = \log \tilde{A}(T)_{ref} + \sum_i \Delta GAV_{\log \tilde{A}}^o(C_i) \quad \text{Eq. 4.12}$$

With

$$\Delta GAV_{\log \tilde{A}}^o = \frac{\log e}{R} (\Delta GAV_{\log \tilde{A}} - \Delta GAV_{\log \tilde{A},ref}) \quad \text{Eq. 4.13}$$

In general, *ab initio* methods yield more accurate relative than absolute values, which makes these methods well suited to calculate ΔGAV^o 's, rather than ΔGAV 's. Also, the temperature dependence of the kinetic parameters is mostly accounted for in $E_{a,ref}(T)$ and $\log \tilde{A}(T)$, making the ΔGAV^o values less temperature dependent.

The group additive values for the pre-exponential factor, and the reference reaction pre-exponential factor are evaluated without taking the number of single events into account. The final pre-exponential factor of a reaction thus needs to be multiplied by the number of single events, as shown in Eq. 4.14.

$$\log A(T) = \log \tilde{A}(T)_{ref} + \sum_i \Delta GAV_{\log \tilde{A}}^o(C_i) + \log n_e \quad \text{Eq. 4.14}$$

The reference reactions for the 1-2 up to 1-7 hydrogen shifts are shown in Figure 2. The most simple reaction per hydrogen shift is chosen, i.e. a thermo-neutral reaction with identical reactants and products and with hydrogen atoms as substituents for each group.

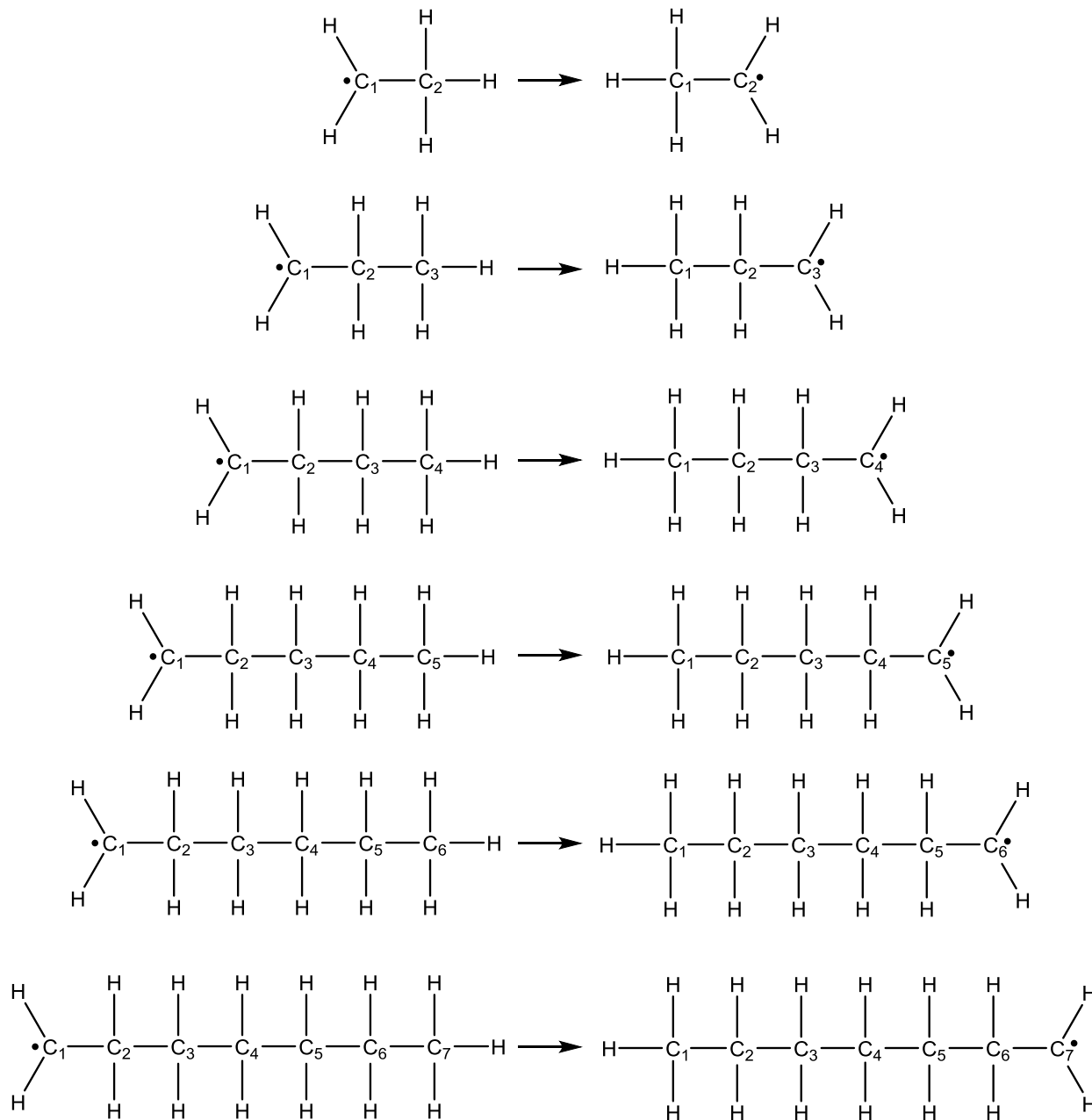


Figure 2: Reference reactions for the 1-2 up to 1-7 intramolecular hydrogen abstraction reactions. Each reaction is thermo-neutral, the products and reactants are identical. Both the attacking carbon atom, attacked carbon atom as well as the carbon chain in between have the simplest surroundings, i.e. hydrogen atoms.

To obtain group additive values, several reactions are calculated in which the hydrogen atoms are substituted with three possible groups: methyl groups noted as C, vinyl groups noted as C_d, and C_t as a triply bonded carbon atom. This leads to 10 possible combinations around each central carbon atom: C-(H)₂, which corresponds to the reference reactions, and which has group additive values of zero, C-(C)(H), C-(C_d)(H), C-(C_t)(H), C-(C)₂, C-(C_d)(C), C-(C_t)(C), C-(C_d)₂, C-(C_d)(C_t), and C-(C_t)₂. In this notation, the ligands that are necessary for the reactions are not included. These are an extra C-ligand for the attacking radical, extra C- and H-ligands for the attacked carbon atom and two extra C-ligands for the atoms along the carbon chain between the reactive atoms. The substitutions are done on C₁ and C₂ (Figure 2) to study the primary contributions and C₃ to C₇ for secondary contributions. The nature of the attacking and attacked carbon atoms and the carbon chain in between both is not changed, only sp³ hybridized atoms are considered in this work.

So far, the pre-exponential factors have been described without considering tunneling corrections. From the *ab initio* calculations, tunneling coefficients have been calculated using the Eckart procedure. These tunneling coefficients are subsequently used to develop correlations which can model the tunneling coefficients of new reactions belonging to the same reaction family. These correlations depend on the temperature and the exothermic activation energy. The latter can be obtained from group additivity and no new *ab initio* calculations need to be done.

In a first step, one reaction is used to determine one group additive value by substituting one group on the reference reaction with one of the nine groups listed above, constituting the training set. This leads to a straightforward determination of the ΔGAV^o 's by taking the difference between the Arrhenius parameters of the reaction and the Arrhenius parameters of the reference reaction. In a next step, the group additivity methodology is validated against a test set of reactions that contain more than one group. If satisfactory agreements are obtained between the *ab initio* and group additively calculated rate coefficients, group additivity can be used to model this reaction family. Finally, to increase the accuracy of the group additive values, all the reactions, i.e. the training set and the test set, are used to determine final ΔGAV^o 's which can deviate slightly from the previously calculated ones, but increase the overall accuracy of the group additivity model. For this step, two least-square regressions are used, one for the pre-exponential factors and one for the activation energies. The ΔGAV^o 's constitute the **b** vector. The

i^{th} element of the \mathbf{y} vector contains the difference in pre-exponential factor or activation energy of reaction i compared to the reference reaction. Each element x_{ij} of the matrix \mathbf{X} is 1 if reaction i contains group j , and 0 otherwise. The ΔGAV^o 's are calculated by:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad \text{Eq. 4.15}$$

The significance of the regression can be verified by calculating the F value, c.f. Eq. 4.16, in which n is the number of reactions and p is the number of group additivity values.

$$F = \frac{\frac{(\mathbf{Xb})^T \mathbf{Xb}}{p}}{\frac{\mathbf{y}^T \mathbf{y} - (\mathbf{Xb})^T \mathbf{y}}{n - p}} \quad \text{Eq. 4.16}$$

The quality of the regression can further be analyzed using R^2 values, i.e. multiple correlation coefficients. R^2 values lie between 0 and 1, and the higher the value, the better the group additive values are able to describe the Arrhenius parameters.

$$R^2 = \frac{(\mathbf{Xb})^T \mathbf{Xb}}{\mathbf{y}^T \mathbf{y}} \quad \text{Eq. 4.17}$$

4.4 Results

4.4.1 Transition state geometries

Table 1 gives the extrema of the two geometrical parameters compared to the user input values. The bond length provided to Genesys, which applies to both the forming and breaking bond of the abstraction reactions, amounts 135pm. The final values in the full set of the 249 transition states lie in between 122pm and 154pm. The extrema in bond lengths do not show much deviation as a function of the type of H-shift. However, bond angles do depend strongly on the type of H-shift. 1-2 H-shifts, proceeding through a three-membered cyclic transition state, have a C-H-C angle between 67° and 71° . The angles in 1-7 H-shifts, of which the transition state contains an eight-membered ring structure, are much higher and can amount up to 177° , which approaches the maximum angle of 180° .

4.4.2 *Ab initio* calculations

This section reports the *ab initio* rate coefficients which were automatically calculated using the Genesys software. The rate coefficients include tunneling corrections and the number of single events and are thus the effective rate coefficients of the reactions. They have been calculated automatically from 300 up to 1800K with intervals of 100K. Each set of 5 consecutive rate coefficients is regressed to obtain Arrhenius parameters.

4.4.2.1 Rate coefficients

The Arrhenius parameters for the training set, regressed between 800-1200K, are shown in Table 4 and Table 5. The tables in Appendix E give the Arrhenius parameters for the test set, also regressed in the temperature interval 800-1200K. Since the regression was done in a small temperature range, the 95% confidence intervals on the pre-exponential factors and activation energies are low, c.f. Table 2. For all the reactions, the uncertainty on $\log A$ is lower than 0.17, which corresponds to a factor of 1.17 on the pre-exponential factor. The minimum values show negligible uncertainties on the pre-exponential factors. A noticeable trend is the increase of the uncertainty as a function of the type of H-shift. The average uncertainty increases from a value of 0.07 for 1-2 H-shifts to a value of 0.11 for 1-7 H-shifts. The maximum uncertainties show a similar trend. For the activation energies, all uncertainties of the regression are within 1.33 kJ mol⁻¹. A similar trend as for the pre-exponential factors is seen when comparing the uncertainties of the different families. Both the average and the maximum values increase from 1-2 H-shifts to 1-7 H-shifts.

Table 2 : Minimum (Min), maximum (Max), and average (Av) uncertainties on the pre-exponential factors and activation energies. The uncertainties on $\log A$ and E_a correspond to the 95% confidence intervals of the linear regression of the Arrhenius equation.

	Uncertainty on $\log A$ (-)			Uncertainty on E_a (kJ mol ⁻¹)		
	Min	Max	Av	Min	Max	Av
1-2 shift	0.05	0.12	0.07	0.37	1.00	0.55
1-3 shift	0.01	0.13	0.09	0.06	1.07	0.72
1-4 shift	0.002	0.13	0.09	0.01	1.02	0.71
1-5 shift	0.00	0.14	0.09	0.04	1.15	0.75
1-6 shift	0.01	0.15	0.10	0.06	1.23	0.80
1-7 shift	0.02	0.17	0.11	0.18	1.33	0.91
All	0.002	0.17	0.09	0.01	1.33	0.76

A total of 249 reactions were considered yielding 448 distinct rate coefficients of single reactions. For several reactions, the reactant and product are identical. These are identity reactions and their

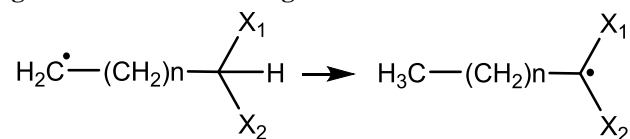
forward and reverse rate coefficients are the same. For the other reactions, the forward and reverse rate coefficients are different and can both be used to generate and validate the group additive approach. Table 3 tabulates the number of reactions and distinct rate coefficients calculated by Genesys for the 1-2 up to 1-7 H-shifts considering (1) the primary contributions, (2) the secondary contributions and (3) the test set reactions. It was observed, as shown in section 4.4, that 1-7 H-shifts do not play an important role over a wide temperature range. Therefore, the test set is limited.

Table 3: Number of reactions and distinct rate coefficients automatically obtained with Genesys.

	Reactions			Distinct rate coefficients		
	Primary	Secondary	Test set	Primary	Secondary	Test set
1-2 shift	10	0	17	19	0	33
1-3 shift	10	9	20	19	9	33
1-4 shift	10	9	19	19	18	34
1-5 shift	10	18	22	19	27	42
1-6 shift	10	18	23	19	36	44
1-7 shift	10	27	7	19	45	13
Total	60	81	108	114	135	199

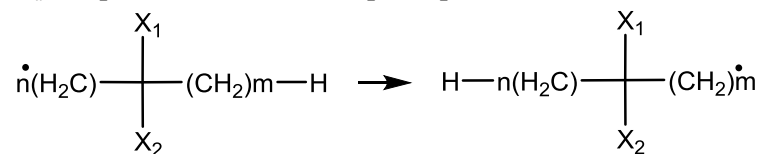
For primary contributions, 10 reactions were calculated per type of H-shift, in which the substituents of the attacked carbon atom were varied between hydrogen atoms, methyl groups, vinyl groups and ethynyl groups. The resulting Arrhenius parameters are reported in Table 4, n is the number of CH_2 groups between the radical site and the attacked carbon atom. For a 1-2 shift, n is 0, for a 1-3 shift, n is 1, and so on. Secondary contributions were calculated by substituting the hydrogen atoms on the carbon chain between the radical site and the attacked carbon atom by methyl groups, vinyl groups and ethynyl groups. The Arrhenius parameters are shown in Table 5. The two substituents X_1 and X_2 are bonded to a carbon atom located at a distance of n CH_2 groups of the radical site and m CH_2 groups of the attacked carbon atom. For 1-2 shifts, no secondary contributions exist. Finally, all the other reactions belong to the test set, tabulated in Appendix E. The test set was constructed by two possible changes from the reference reactions. First, the hydrogen atoms of the reference reaction can be substituted by larger groups, e.g. ethyl groups instead of methyl groups. Second, hydrogen atoms on two carbon atoms were substituted. The full test set of 1-2 up to 1-6 H-shifts contain each group on each position at least once. The 1-7 H-shift test set is smaller due to the relative unimportance of 1-7 H-shifts compared to the other intramolecular hydrogen abstractions. In this test set, not all the groups are present.

Table 4: Arrhenius parameters for the reactions used to deduce primary ΔGAV° 's at 1000K regressed from the rate coefficients between 800 and 1200K. Per reaction, the first line corresponds to the forward direction, the second line to the reverse. The pre-exponential factors are expressed in s^{-1} and E_a is expressed in $kJ\ mol^{-1}$. All pre-exponential factors include tunneling and the number of single events.



X_1	X_2	1-2 (n=0)		1-3 (n=1)		1-4 (n=2)		1-5 (n=3)		1-6 (n=4)		1-7 (n=5)	
		logA	Ea	logA	Ea	logA	Ea	logA	Ea	logA	Ea	logA	Ea
H	H	13.4	177.3	12.7	176.2	11.4	106.0	10.5	73.1	10.0	69.3	8.8	76.0
		13.4	177.3	12.7	176.2	11.4	106.0	10.5	73.1	10.0	69.3	8.8	76.0
H	CH ₃	13.6	167.4	13.0	166.4	11.7	93.7	10.8	62.8	9.8	57.1	8.6	63.5
		13.9	183.6	13.1	181.8	11.8	110.7	10.8	78.8	10.1	74.1	9.0	80.3
H	C ₂ H ₃	13.4	135.6	12.9	144.0	11.1	74.9	10.3	46.6	9.7	44.4	8.3	49.4
		14.3	211.0	13.9	217.4	12.4	148.5	11.5	120.1	10.8	117.8	9.2	123.0
H	C ₂ H	13.5	144.7	13.0	150.9	11.6	77.2	10.7	49.0	9.3	44.5	8.6	51.8
		13.8	206.6	13.2	207.0	11.9	134.0	11.0	104.7	9.5	102.5	8.9	107.0
CH ₃	CH ₃	13.3	157.2	12.7	157.5	11.5	84.7	10.4	52.1	9.2	47.8	8.3	56.3
		14.1	183.8	13.0	182.5	11.8	111.1	10.6	78.5	9.7	75.0	8.4	81.9
CH ₃	C ₂ H ₃	12.9	127.6	12.6	137.8	11.1	66.8	10.0	38.7	8.7	37.2	7.3	44.6
		14.2	215.2	13.8	221.2	12.4	152.1	11.3	123.3	10.5	127.2	9.1	130.6
CH ₃	C ₂ H	13.2	136.0	12.7	143.2	11.3	68.8	10.4	40.5	9.2	34.2	7.9	43.9
		14.0	211.7	12.9	213.2	11.6	140.6	10.8	111.2	9.7	104.9	8.9	117.2
C ₂ H ₃	C ₂ H ₃	12.7	115.0	12.5	126.2	11.1	56.7	10.2	35.1	9.2	33.3	7.5	38.6
		13.9	233.7	13.6	241.3	12.1	175.4	10.8	151.2	10.7	154.6	9.0	155.2
C ₂ H ₃	C ₂ H	13.0	114.6	12.6	129.0	11.2	56.0	10.2	29.5	9.0	25.3	7.7	32.1
		14.3	234.8	13.8	239.3	12.4	168.5	11.5	140.0	10.4	136.6	9.3	143.4
C ₂ H	C ₂ H	13.1	115.3	12.7	130.1	11.3	55.3	10.3	30.8	9.0	26.3	8.2	32.9
		13.8	233.7	12.8	235.3	11.7	161.9	10.7	133.6	9.9	129.9	8.6	136.4

Table 5: Arrhenius parameters for the reactions used to deduce secondary ΔGAV° 's at 1000K regressed from the rate coefficients between 800 and 1200K. The pre-exponential factors are expressed in s^{-1} and E_a is expressed in $kJ\ mol^{-1}$. All pre-exponential factors include tunneling and the number of single events.



		1-3		1-4				1-5					
		n=1, m=1		n=1, m=2		n=2, m=1		n=1, m=3		n=2, m=2		n=1, m=3	
X_1	X_2	logA	E_a	logA	E_a	logA	E_a	logA	E_a	logA	E_a	logA	E_a
H	CH ₃	13.1	174.5	11.7	100.5	12.1	103.9	10.7	67.7	11.1	71.0	10.6	67.3
H	C ₂ H ₃	12.7	172.9	11.7	99.3	11.7	101.7	10.9	68.0	10.9	70.4	10.7	68.9
H	C ₂ H	12.6	170.5	11.6	97.9	11.8	103.4	10.8	67.2	11.0	73.8	10.6	71.3
CH ₃	CH ₃	13.3	171.5	11.7	98.0	12.2	102.7	11.1	66.0	11.6	70.1	10.7	66.7
CH ₃	C ₂ H ₃	13.0	169.9	11.7	95.3	11.9	100.6	11.0	63.9	11.3	67.6	10.7	65.3
CH ₃	C ₂ H	13.0	169.5	11.7	96.9	12.1	102.7	10.9	63.0	11.3	68.3	10.6	64.2
C ₂ H ₃	C ₂ H ₃	12.9	167.4	11.7	95.2	11.7	104.0	11.0	61.6	11.0	68.2	11.2	62.4
C ₂ H ₃	C ₂ H	12.9	168.0	11.8	93.9	11.8	101.3	11.0	62.1	11.0	69.4	10.6	68.3
C ₂ H	C ₂ H	12.7	167.3	11.7	94.4	11.7	103.2	10.9	64.4	10.8	72.6	10.8	69.2

		1-6								1-7									
		n=1, m=4		n=2, m=3		n=3, m=2		n=4, m=1		n=1, m=5		n=2, m=4		n=3, m=3		n=4, m=2		n=5, m=1	
X_1	X_2	logA	E_a	logA	E_a	logA	E_a	logA	E_a	logA	E_a	logA	E_a	logA	E_a	logA	E_a	logA	E_a
H	CH ₃	10.0	65.3	10.0	66.3	10.0	65.9	10.3	68.0	8.8	72.1	8.9	73.7	9.1	72.3	8.7	74.4	8.9	74.1
H	C ₂ H ₃	9.8	64.8	10.3	68.5	10.1	68.2	9.9	67.4	8.8	71.3	8.8	74.7	9.3	74.0	8.9	74.3	8.4	73.6
H	C ₂ H	9.7	62.8	10.1	67.5	9.9	67.6	9.9	69.0	8.7	71.5	8.9	72.6	9.2	79.2	8.9	74.0	8.8	78.1
CH ₃	CH ₃	9.8	63.2	10.1	66.7	10.1	65.5	10.4	68.5	9.0	69.2	9.5	74.7	9.3	80.1	8.9	70.4	9.1	76.3
CH ₃	C ₂ H ₃	9.9	61.2	10.3	66.5	10.4	65.5	10.3	66.2	9.1	72.5	9.2	69.6	10.0	75.7	9.3	69.1	9.1	76.8
CH ₃	C ₂ H	9.9	61.9	10.2	62.8	10.2	63.8	10.3	67.7	8.9	73.3	9.1	70.8	9.6	77.4	9.0	72.1	9.5	79.9
C ₂ H ₃	C ₂ H ₃	9.9	58.7	10.4	63.3	10.6	64.5	9.8	65.2	9.0	69.9	9.2	65.8	9.9	71.4	9.2	66.3	9.5	74.9
C ₂ H ₃	C ₂ H	9.9	57.7	10.3	67.0	10.3	65.9	10.0	65.9	8.9	70.3	9.0	75.0	9.8	78.1	9.1	75.1	8.5	78.4
C ₂ H	C ₂ H	9.8	60.4	10.1	65.3	10.2	68.3	10.0	72.0	9.1	73.0	9.0	69.7	9.7	77.6	9.1	74.4	9.2	83.5

From all the *ab initio* calculated rate coefficients, the activation energy can be plotted as a function of the reaction enthalpy, c.f. Figure 3. Although a clear trend is visible, the use of linear free-energy relationships such as Evans-Polanyi⁴¹ or analogous methods such as Blowers-Masel⁴² leads to a maximum deviation of 11 kJ mol⁻¹ or more, which is high compared to the group additivity results below.

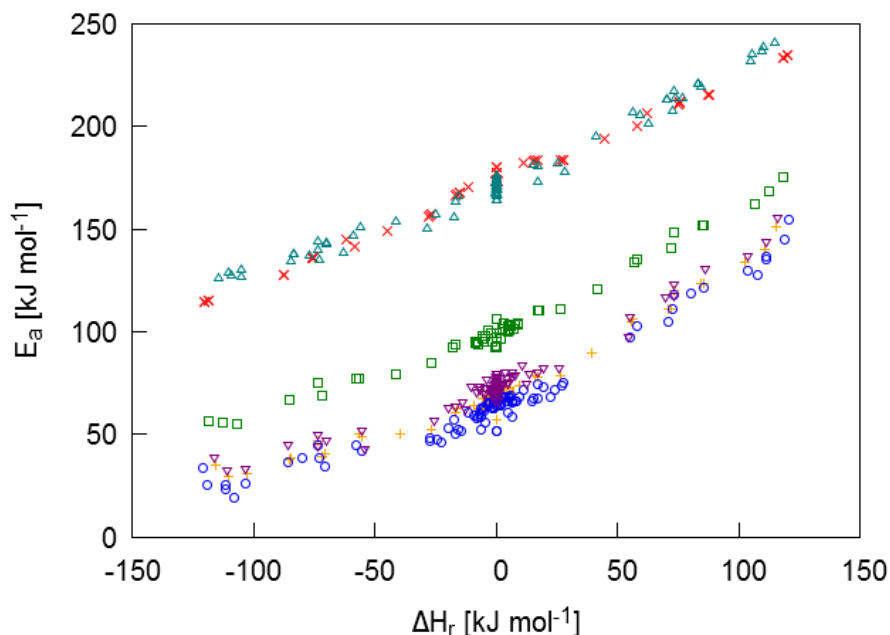


Figure 3: *Ab initio* activation energies of all the training and test reactions as a function of the *ab initio* reaction enthalpy, calculated at 1000K. The points correspond to 1-2 (×), 1-3 (△), 1-4 (□), 1-5 (+), 1-6 (○), and 1-7 (▽) H-shifts.

4.4.2.2 Number of single events

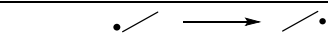
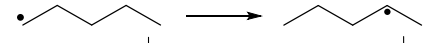
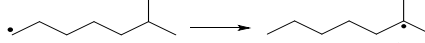
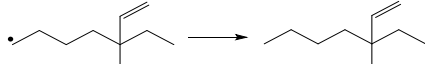
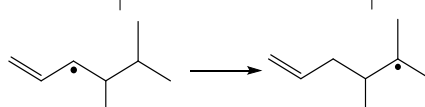
Genesys automatically calculates the total rotational symmetry number and number of optical isomers to obtain the number of single events. A few typical symmetrical contributions have been tabulated in Table 6. The six reference reactions all have 6 as number of single events. The reactants have a two-fold and three fold internal symmetry contribution, without any external symmetry, and the transition state has an external symmetry number of 2. The latter assumes that flipping ring structures into another conformer has a very low energy barrier. The reactants and transition states do not exhibit any optical isomerism. However, the reference reactions are all identity reactions, which implies that the number of single events needs to be increased by a factor 2, from the indistinguishability of the forward versus the reverse reaction. This leads to a number of single events of $2 \cdot \frac{6/1}{2/1} = 6$. This is illustrated by the 1-2 H-shift of ethyl (R1) in Table

6. The value of 6 can be intuitively explained by the three indistinguishable hydrogen atoms of the methyl group and the 2 sides of the planar radical structure that can approach the abstracted hydrogen atom.

When one of the four hydrogen atoms neighboring the reactive atoms is substituted by a larger group, e.g. a methyl group, the two-fold external symmetry of the transition state disappears. Instead, a three-fold internal symmetry arises as a result of the methyl rotation. Furthermore, in the transition state, the reactive atom bonded to the methyl group is chiral. Its four neighbors are different in nature since the hydrogen atom participating in the reaction is further away from the carbon atom compared to the other hydrogen atom. This chiral center leads to a number of optical isomers of 2. R2 in Table 6 illustrates this for the 1-4 H-shift in 1-pentyl. The value of four comes from two indistinguishable hydrogen atoms that can be abstracted by 2 possible sides of the CH₂ radical group.

If another methyl group is added to a reactive atom, the chirality of the transition state disappears again. The transition state still does not contain any external rotational symmetry and the methyl rotor is the only contribution of the transition state to the number of single events. However, because these methyl rotors are also present in the reactant, this contribution cancels out. Again, the resulting value of 2 can be explained by the number of hydrogen atoms, i.e. one, and the number of ways the planar CH₂ radical group can face the abstracted hydrogen atom, i.e. two.

Table 6 : Numbers of single events of selected reactions.

	Reaction	Reactant		TS		
		σ	n_{opt}	σ	n_{opt}	n_e
R1		6	1	2	1	6
R2		6	1	3	2	4
R3		18	1	9	1	2
R4		18	2	3	2	6
R5		18	2	18	4	2

Substituents on the carbon chain between the two reactive atoms have also an influence on the symmetry numbers and numbers of optical isomers, such as R4 in Table 6. The optical isomers in the reactant are preserved throughout the reaction leading to a cancelation of its contribution to the number of single events. Since the internal symmetry due to the methyl substituent also

cancel out, the only final contribution comes from the terminal methyl group and the radical center, leading to number of single events of 6. This value originates from the three indistinguishable hydrogen atoms that can be abstracted, and the two ways the planar CH_2 radical group can face the abstracted hydrogen atom.

R5 shows another example in which the reactant has one chiral center, leading to a number of optical isomers of 2. The transition state, however, contains two chiral centers, and 4 isomers of this transition state exist: two pairs of diastereomers. In principle, the energy of these two diastereomers is different and the rate coefficient depends thus on the specific diastereomer. However, this difference is neglected in this work, and the lowest energy diastereomer is always taken. In this case, the number of optical isomers amounts to 4, yielding a total number of single events of two. This number represents the two ways the planar radical group can face the abstracted hydrogen atom. If the diastereomers would be made distinguishable, two rate coefficients would be calculated, each with a number of single events of 1. In this case, the dihedral angle around the bond between the two chiral centers in the transition state is fixed per diastereomer, and the planar radical structure can only approach the abstracted hydrogen atom with one direction. The total rate coefficient used in this work will slightly overestimate the real total rate coefficient because the highest rate coefficient is multiplied by 2 instead of using the sum of the highest and lowest value.

The above examples give an overview of the different types of contributions. However, many other combinations are encountered in this study, but a complete analysis of all the number of single events falls outside the scope of this chapter.

4.4.2.3 Tunneling coefficients

Tunneling plays a crucial role in hydrogen abstractions, whether they are intramolecular or intermolecular. At low temperatures, the tunneling coefficient can amount up to $2 \cdot 10^5$ for 1-3 H-shifts. However, at these temperatures, the total rate coefficient is insignificant for practical purposes such as combustion, oxidation or pyrolysis reactions. At higher temperatures, tunneling is still important, where the coefficients amount up to 1.52 for 1-3 H-shifts at 1000K. The average and extreme values of 1-2 to 1-7 H-shifts are given in Table 7. In general, 1-3 H-shifts have the highest tunneling coefficients, followed by 1-2 and 1-4 H-shifts. 1-5, 1-6 and 1-7 H-shifts have comparable values for tunneling coefficients.

Figure 4 shows the tunneling coefficients at 1000K as a function of the imaginary frequency. The imaginary frequency is, besides the electronic energy of activation, the second parameter to calculate Eckart tunneling corrections. From Figure 4, the range in imaginary frequencies can also be observed. 1-3 H-shifts have the highest imaginary frequencies, which range from 1945 to 2100 cm^{-1} , this is also reflected in the high tunneling coefficients for these reactions. In 1-2 H-shifts, although the reaction proceeds through a three-membered ring, there are no small C-C-C bond angles. In 1-3 H-shifts, however, the four-membered ring has a small C-C-C angle, leading to a higher ring strain for these reactions. The imaginary frequencies in 1-2 H-shifts vary between 1723 and 1982 cm^{-1} . From 1-4 to 1-6 H-shifts, there is a downward trend in the imaginary frequencies with averages of 1780 cm^{-1} for 1-4, 1611 cm^{-1} for 1-5, and 1545 cm^{-1} for 1-6 H-shifts. The imaginary frequencies for 1-7 H-shifts increase again and lie in a range of 1393 to 1693 cm^{-1} . For each H-shift, most of the reactions show an imaginary frequency that is relatively close to the maximum value for the family. Only a few reactions show imaginary frequencies that are closer to the minimum value. These reactions all include the formation or destruction of a strong resonantly stabilized radical, e.g. a hydrogen is abstracted between two vinyl groups.

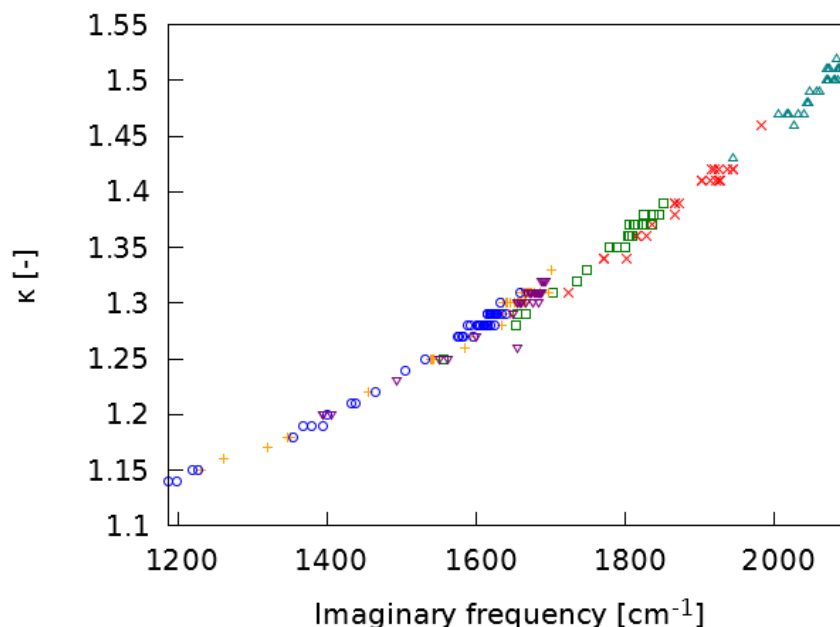


Figure 4: Tunneling coefficients, calculated by the Eckart²⁶ procedure, at 1000K as a function of the imaginary frequency of the saddle point on the PES. The points correspond to 1-2 (×), 1-3 (△), 1-4 (□), 1-5 (+), 1-6 (○), and 1-7 (▽) H-shifts.

Table 7: Average and extreme values for tunneling coefficients of 1-2 to 1-7 H-shifts at temperatures ranging from 300 to 1800K.

T/K	1-2			1-3			1-4		
	Average	Maximum	Minimum	Average	Maximum	Minimum	Average	Maximum	Minimum
300	10860.65	59138.61	290.02	98656.95	213897.87	8151.39	270.59	585.33	26.95
400	22.97	46.40	9.09	63.85	85.07	28.43	10.35	14.10	4.84
500	4.88	6.52	3.39	7.59	8.59	5.58	3.74	4.39	2.55
600	2.72	3.20	2.22	3.49	3.74	2.96	2.38	2.63	1.88
700	2.02	2.25	1.76	2.38	2.49	2.14	1.86	2.00	1.58
800	1.69	1.83	1.53	1.91	1.97	1.76	1.60	1.69	1.42
900	1.51	1.60	1.40	1.65	1.69	1.56	1.45	1.51	1.32
1000	1.39	1.46	1.31	1.49	1.52	1.43	1.35	1.39	1.25
1100	1.31	1.36	1.25	1.39	1.41	1.34	1.28	1.32	1.20
1200	1.26	1.30	1.21	1.32	1.34	1.28	1.23	1.26	1.17
1300	1.21	1.25	1.17	1.27	1.28	1.23	1.20	1.22	1.14
1400	1.18	1.21	1.15	1.23	1.24	1.20	1.17	1.19	1.12
1500	1.16	1.18	1.13	1.19	1.20	1.17	1.15	1.16	1.11
1600	1.14	1.16	1.11	1.17	1.18	1.15	1.13	1.14	1.09
1700	1.12	1.14	1.10	1.15	1.16	1.13	1.11	1.13	1.08
1800	1.11	1.12	1.09	1.13	1.14	1.12	1.10	1.11	1.07
T/K	1-5			1-6			1-7		
	Average	Maximum	Minimum	Average	Maximum	Minimum	Average	Maximum	Minimum
300	40.64	66.41	4.95	30.04	49.12	4.37	64.99	102.65	9.48
400	5.60	6.91	2.36	4.88	6.00	2.23	6.39	7.56	3.17
500	2.79	3.13	1.72	2.58	2.94	1.66	2.96	3.22	2.04
600	2.00	2.16	1.46	1.90	2.10	1.42	2.06	2.18	1.63
700	1.65	1.76	1.32	1.59	1.72	1.30	1.69	1.75	1.43
800	1.47	1.54	1.24	1.43	1.52	1.22	1.49	1.53	1.32
900	1.36	1.41	1.19	1.33	1.40	1.17	1.37	1.40	1.25
1000	1.28	1.33	1.15	1.26	1.31	1.14	1.29	1.32	1.20
1100	1.23	1.26	1.12	1.21	1.26	1.12	1.24	1.26	1.16
1200	1.19	1.22	1.10	1.18	1.22	1.10	1.20	1.21	1.13
1300	1.16	1.19	1.09	1.15	1.18	1.08	1.17	1.18	1.11
1400	1.14	1.16	1.08	1.13	1.16	1.07	1.14	1.15	1.10
1500	1.12	1.14	1.07	1.11	1.14	1.06	1.12	1.13	1.09
1600	1.11	1.13	1.06	1.10	1.12	1.06	1.11	1.12	1.08
1700	1.10	1.11	1.05	1.09	1.11	1.05	1.10	1.11	1.07
1800	1.09	1.10	1.05	1.08	1.10	1.04	1.09	1.09	1.06

4.4.2.4 Comparison to experimental data

The rate coefficients automatically calculated by Genesys were compared to experimental data from literature. The 1-4 H-shift in 1-pentyl has been extensively measured in several independent experiments and its rate coefficients can serve as a good data to compare the *ab initio* data to.

Figure 5a shows the 1-4 shift of 1-pentyl to 2-pentyl in the temperature range 300-1800K in the high pressure limit, compared to 6 experimental studies. A good agreement is seen between the theoretical and experimental rate coefficients. The temperature dependence at low temperatures is well-captured by the tunneling coefficients, which amounts to 468 at 300K and has thus a major contribution to the rate coefficient at these low temperatures.

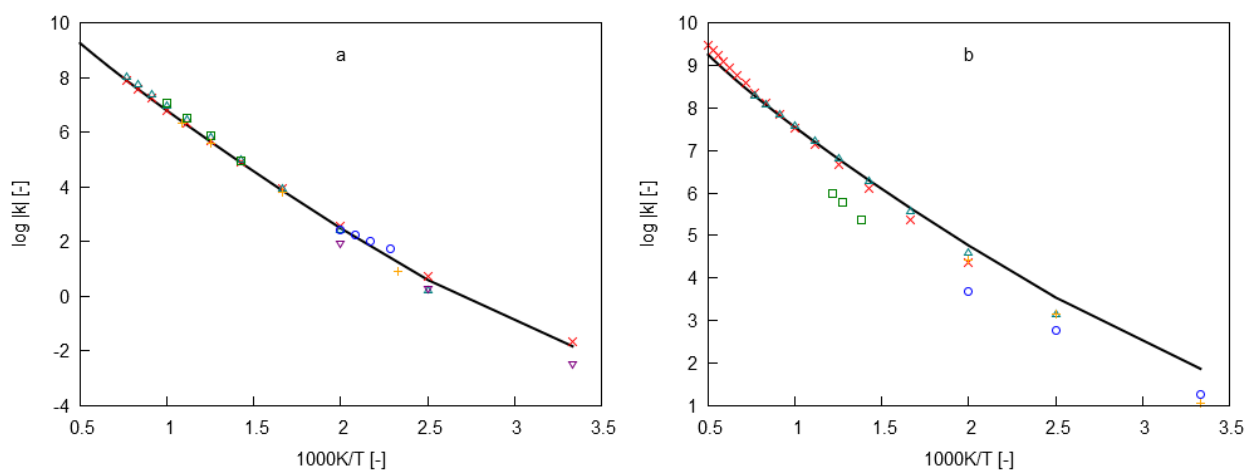


Figure 5: (a) Arrhenius plot of the 1-4 H-shift of 1-pentyl to 2-pentyl compared to experimental data from the literature of Miyoshi et al.⁴³ (×), Yamauchi et al.⁴⁴ (△), Tsang et al.⁴⁵ (□), Marshall⁴⁶ (+), Endrenyi and Le Roy⁴⁷ (○), and Watkins⁴⁸ (▽). (b) Arrhenius plot of the 1-5 H-shift of 1-hexyl to 2-hexyl compared to experimental data from the literature of Tsang et al.⁴⁹ (×), Yamauchi et al.⁴⁴ (△), Imbert and Marshall⁵⁰ (□), Dóbé et al.⁵¹ (+), and Watkins and Ostreko⁵² (○). The lines correspond to the *ab initio* rate coefficients calculated by Genesys. Log(k) is reported in which the rate coefficients are expressed in s^{-1} .

Another reaction for which many experimental data are available is the 1-5 H-shift in the 1-hexyl radical yielding the 2-hexyl radical, of which the rate coefficients are given in Figure 5b. Although much effort has been done to obtain experimental rate coefficients, the results are scattered over more than one order of magnitude, especially at temperatures between 500 and 1000K. The rate coefficients calculated by Genesys show a good agreement at high temperatures, i.e. a maximum deviation factor below 1.4 above 500K compared to the data of Yamauchi et al.⁴⁴ and a deviation factor up to 2 above 500K compared to the data of Tsang et al.⁴⁹. However, at lower temperatures, the rate coefficients are higher than the experimental ones.

4.4.2.5 Comparison to theoretical data

As mentioned in the introduction of this chapter, many different groups have been working on intramolecular hydrogen abstraction reactions, using a variety of theoretical methods. The rate coefficients of 46 reactions calculated by Genesys at 1000K were compared to literature data from several sources, as tabulated in Table 8. Reactions from 1-2 up to 1-7 H-shift families were

compared, mainly consisting of alkyl radicals and a few alkenyl radicals, since these are the only literature data available for which comparable rate coefficients were calculated with Genesys. Of the 70 rate coefficients compared, 40 of the literature rate coefficients are within a factor of 2 of the Genesys rate coefficients. 17 more rate coefficients are within a factor of 3, giving only 17 rate coefficients that deviate more than a factor 3. A few of these outliers are the rate coefficients of Ratkiewicz et al.²⁰ in which the 1-3 H-shift in 1-propyl, the 1-4 H-shift in 1-butyl, the 1-5 H-shift in 1-pentyl, the 1-6 H-shift in 1-hexyl are 5, 3, 55 and 717 times higher than the Genesys rate coefficients, respectively. These very high deviations are also seen for the Ratkiewicz data compared to other literature data.

In conclusion, the rate coefficients of Genesys are in general well within the uncertainty limits of the literature data. These results also indicate that the generation of a set of group additive values seems feasible with the current dataset of rate coefficients.

Table 8 : Rate coefficients of several reactions calculated by Genesys (Gen) and the ratio of the Genesys rate coefficients to literature rate coefficients evaluated at 1000K.

Family	Reaction	k_{Gen} (s^{-1})	Wang et al. ^{29,30}	Davis et al. ²⁷	$\frac{k_{Gen}}{k_{lit}}$	Ratkiewicz et al. ²⁰	Zheng and Truhlar ⁶
1-2		$2.8 \cdot 10^4$	0.95	2.69			
1-2		$6.8 \cdot 10^4$	0.98	2.00			
1-2		$6.6 \cdot 10^4$	1.05	2.77			
1-2		$1.8 \cdot 10^4$	0.54				
1-2		$1.2 \cdot 10^5$	0.96				
1-2		$6.5 \cdot 10^4$	0.97	3.23			
1-2		$1.5 \cdot 10^4$	0.40				
1-2		$2.0 \cdot 10^4$	0.81				
1-2		$1.9 \cdot 10^6$	1.02				
1-3		$6.2 \cdot 10^3$	0.95	0.79		0.22	
1-3		$1.9 \cdot 10^4$	0.91	1.01			
1-3		$2.0 \cdot 10^4$	1.01				
1-3		$4.8 \cdot 10^4$	1.08				
1-3		$6.3 \cdot 10^3$	0.42				
1-3		$2.9 \cdot 10^4$	0.92				
1-3		$4.1 \cdot 10^4$	2.93				
1-3		$1.8 \cdot 10^4$		1.46			
1-3		$2.5 \cdot 10^5$	2.47				
1-4		$1.6 \cdot 10^6$	0.92	0.46		0.32	
1-4		$4.3 \cdot 10^6$	2.38				
1-4		$6.1 \cdot 10^6$	0.91	0.48	1.30		0.62
1-4		$3.4 \cdot 10^6$	1.74				
1-4		$2.3 \cdot 10^6$	0.36				

(Table 8 continued)

Family	Reaction	k_{Gen} (s^{-1})	Wang et al. ^{29,30}	Davis et al. ²⁷	$\frac{k_{Gen}}{k_{lit}}$	Sirjean et al. ²¹	Ratkiewicz et al. ²⁰	Zheng and Truhlar ¹⁶
1-4		$1.4 \cdot 10^6$	0.32					
1-4		$1.1 \cdot 10^7$	0.89					
1-4		$4.2 \cdot 10^6$	0.64	0.70	1.30			
1-4		$1.6 \cdot 10^7$	0.55					
1-5		$1.0 \cdot 10^7$	0.90	0.23			0.02	
1-5		$2.5 \cdot 10^7$	0.72					
1-5		$1.6 \cdot 10^7$	0.45					
1-5		$3.5 \cdot 10^7$	0.95	0.53	1.73			0.53
1-5		$1.7 \cdot 10^7$	0.94					
1-5		$3.1 \cdot 10^7$	1.48					
1-5		$3.0 \cdot 10^7$	0.76	0.41	1.54			
1-5		$4.5 \cdot 10^7$	0.77					
1-5		$7.6 \cdot 10^7$	0.48					
1-6		$2.4 \cdot 10^6$	0.90	0.22			0.001	
1-6		$6.6 \cdot 10^6$	0.31	0.19	1.64			
1-6		$3.4 \cdot 10^6$	0.57					
1-6		$4.7 \cdot 10^6$	0.61					
1-6		$2.6 \cdot 10^7$	0.93					
1-7		$6.6 \cdot 10^4$	0.63	0.09				
1-7		$1.1 \cdot 10^5$	0.74					
1-7		$3.9 \cdot 10^5$	1.49					
1-7		$1.9 \cdot 10^5$	0.52	0.07	1.72			
1-7		$2.3 \cdot 10^5$	0.48					

4.4.3 Group additivity model

As explained previously, to obtain group additive values, the Arrhenius parameters are used, excluding the number of single events and tunneling corrections. The resulting pre-exponential factor is referred to as the single-event pre-exponential factor. Tunneling corrections and the number of single event are calculated separately from the group additive values and are added to the group additive rate coefficient to again obtain the total rate coefficient.

4.4.3.1 Initial model

The reactions from the training set, i.e. Table 4 and Table 5, are used to deduce an initial set of ΔGAV° 's, in which each ΔGAV° is calculated based on the rate coefficient of a single reaction. The reactions only deviate from the reference reactions by one change to a reactive atom or to a carbon atom along the chain in between the two reactive carbon atoms, all other groups remain equal to those of the reference reaction. The tables in Appendix F tabulate the initial group additive values and the reference Arrhenius parameters.

The validity of the initial model has been assessed by comparing the *ab initio* rate coefficients by group additivity rate coefficients. Appendix E list the test set reactions for 1-2 to 1-7 H-shifts respectively. The last column shows the ratio of the group additivity calculated rate coefficient to the *ab initio* calculated rate coefficient. The parity plots given in Figure 6 summarize these data. In Table 9, the mean absolute deviations and maximum deviations between the group additive rate coefficient and the *ab initio* rate coefficients are given. All the deviations were calculated by dividing the highest rate coefficient by the lowest one, to only get values higher than 1. Except for 1-7 H-shift, the mean deviations are all well within a factor of two, which is a typical accuracy for group additivity⁴, and which is of the same order of magnitude as the uncertainty on CBS-QB3 *ab initio* rate coefficients. A few outliers show higher deviations, which can be attributed to either not considering non-nearest neighbor effects, which will be discussed later, or to the *ab initio* uncertainties.

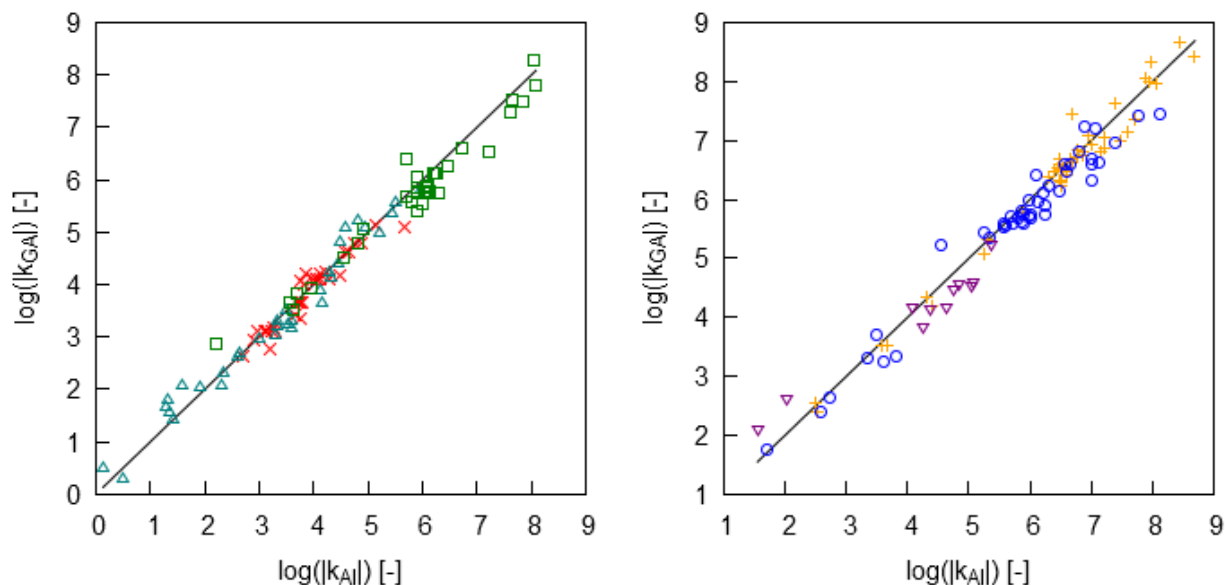


Figure 6: Parity plots of the group additivity predicted single-event rate coefficients k_{GA} versus the *ab initio* rate coefficients k_{AI} at 1000K for 1-2 (\times), 1-3 (\triangle) and 1-4 (\square) H-shifts (left), and for 1-5 ($+$), 1-6 (\circ) and 1-7 (∇) H-shifts (right). The rate coefficients exclude tunneling corrections in this figure. $\text{Log}(|k|)$ is reported for all reactions in which the rate coefficient is expressed in s^{-1} .

Table 9: Mean absolute deviation (MAD) and maximum (MAX) deviation between the group additivity rate coefficients and the *ab initio* rate coefficients evaluated at 1000K of the training and test reactions using both the initial set as well as the final set of group additivity values. All the deviations are calculated by dividing the highest rate coefficient by the lowest one. For each type of H-shift, the first row contains the deviations obtained from the initial set of group additive values. The training set is not included because all deviation factors amount 1 exactly. The bottom two rows contain the deviations of the training set and the ones of the test set respectively obtained from the final set of group additive values.

		1-2		1-3		1-4	
		MAD	MAX	MAD	MAX	MAD	MAX
Initial set of ΔGAV° 's	Test set	1.45	3.58	1.75	3.31	1.93	4.82
Final set of ΔGAV° 's	Training set	1.28	1.78	1.23	1.72	1.27	1.85
	Test set	1.25	1.81	1.35	2.60	1.44	2.91
		1-5		1-6		1-7	
		MAD	MAX	MAD	MAX	MAD	MAX
Initial set of ΔGAV° 's	Test set	1.60	5.82	1.90	4.96	2.41	3.81
Final set of ΔGAV° 's	Training set	1.20	1.82	1.25	1.78	1.07	1.90
	Test set	1.18	1.81	1.31	2.45	1.75	3.96

4.4.3.2 Final model

By using a training and test set, it has been shown that the group additivity approach is valid for intramolecular hydrogen abstraction reactions. Although the calculations of the rate coefficients are within the expected limits compared to the *ab initio* data, it is possible to increase the accuracy of the overall model by including the kinetic data of the test set in the calculation of

group additivity values. Since for all H-shifts but the 1-7 H-shifts, all groups are present at least twice in the combined set, a linear regression approach is necessary to obtain new group additivity values. For 1-7 H-shifts, the test set is small due to the relative unimportance of these reactions. The group additive values are summarized in Table 11 and Table 12. Although no independently calculated reactions are available to validate this set of group additive values, the validity of group additivity for these reactions has already been shown, and it can be expected that the new set of group additive values is more accurate because it was retrieved from a larger set of reactions. The gain in accuracy is shown in Table 9.

Table 10 shows F and R^2 values of the least-square regressions of ΔGAV° 's from the Arrhenius parameters of the *ab initio* calculations. All the significances are higher than the $F_{0.05}$ value, except for the pre-exponential factors of the 1-7 H-shifts. More reactions are thus necessary to conclude this reaction family, but their relative importance is negligible, as shown in a next section. Therefore, it was chosen not to extend the dataset for 1-7 H-shifts. The multiple correlation coefficients between the ΔGAV° 's are high, especially for the activation energies. The Arrhenius parameters are thus well represented by the group additivity model. The uncertainties on the group additive values, defined by the 95% confidence intervals, are given in Appendix G.

Table 10 : Significances of the regressions (F) and multiple correlation coefficients (R^2) of regression of ΔGAV° 's for the 1-2 to 1-7 H-shifts.

	log A		E_a	
	F	R^2	F	R^2
1-2	13.55	0.878	254.1	0.993
1-3	8.48	0.871	456.4	0.997
1-4	7.72	0.888	298.1	0.997
1-5	8.48	0.899	364.6	0.997
1-6	7.12	0.895	258.6	0.997
1-7	1.62	0.879	108.0	0.998

Table 11: Primary group additive values ΔGAV° for 1-2 to 1-7 intramolecular hydrogen abstractions deduced from the combined training set and test set. The reference reactions are given in Figure 2. The index i depends on the H-shift. For 1-2 H-shifts, i equals 2, for 1-3 H-shifts 3, etc. The units for ΔGAV°_{Ea} values are kJ mol^{-1} . For the reference reaction, the single-event pre-exponential factors are expressed in s^{-1} and Ea is expressed in kJ mol^{-1} . The reference reaction pre-exponential factor does not include tunneling nor the number of single events.

	1-2		1-3		1-4		1-5		1-6		1-7	
	$\log \tilde{A}$	Ea	$\log \tilde{A}$	Ea	$\log \tilde{A}$	Ea	$\log \tilde{A}$	Ea	$\log \tilde{A}$	Ea	$\log \tilde{A}$	Ea
Reference	12.783	177.30	12.043	176.19	10.823	106.03	9.943	73.11	9.122	69.30	7.892	75.99
$C_1-(C)(H)_2$	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00
$C_1-(C)_2(H)$	-0.032	7.01	0.086	5.34	0.034	4.48	0.013	4.76	0.087	4.26	0.340	4.29
$C_1-(C_d)(C)(H)$	0.402	33.33	0.761	41.99	0.688	42.31	0.775	46.08	0.764	48.24	0.450	47.04
$C_1-(C_t)(C)(H)$	0.120	25.32	0.203	31.59	0.183	27.82	0.065	29.97	-0.332	33.83	0.140	31.04
$C_1-(C)_3$	0.021	10.31	0.147	8.76	-0.005	4.56	-0.357	5.26	-0.360	6.74	-0.350	5.88
$C_1-(C_d)(C)_2$	0.506	35.53	0.794	47.49	0.654	46.83	0.515	49.99	0.540	56.35	0.070	53.90
$C_1-(C_t)(C)$	-0.020	31.18	-0.061	38.02	-0.002	36.18	0.059	38.31	-0.140	38.11	-0.310	37.32
$C_1-(C_d)_2(C)$	0.480	56.23	0.818	64.07	0.665	66.51	0.428	77.93	0.508	82.53	0.240	79.16
$C_1-(C_d)(C_t)(C)$	0.617	55.01	0.695	63.30	0.640	64.29	0.745	66.77	0.474	68.44	0.590	67.39
$C_1-(C_t)_2(C)$	0.197	56.04	0.200	59.90	0.250	52.32	0.415	61.84	-0.031	60.59	0.110	60.39
$C_i-(C)(H)_3$	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00
$C_i-(C)_2(H)_2$	0.069	-10.63	0.223	-11.26	0.135	-13.12	0.079	-11.69	0.150	-11.99	-0.010	-12.31
$C_i-(C_d)(C)(H)_2$	-0.210	-42.29	0.008	-32.99	-0.387	-32.25	-0.196	-27.19	-0.141	-25.64	-0.300	-26.59
$C_i-(C_t)(C)(H)_2$	-0.064	-35.02	0.202	-27.40	0.018	-30.08	-0.046	-25.84	-0.201	-24.03	-0.030	-24.19
$C_i-(C)_3(H)$	0.160	-16.47	0.455	-16.65	-0.070	-22.78	-0.057	-21.20	-0.374	-21.70	0.010	-19.69
$C_i-(C_d)(C)_2(H)$	-0.296	-45.88	-0.051	-36.46	-0.137	-40.57	-0.360	-36.15	-0.613	-32.06	-0.835	-32.32
$C_i-(C_t)(C)(H)$	-0.039	-39.06	0.194	-33.30	0.144	-36.48	-0.030	-32.30	-0.240	-33.19	-0.510	-34.84
$C_i-(C_d)_2(C)(H)$	-0.473	-62.70	-0.348	-53.33	-0.028	-52.32	-0.182	-39.01	-0.498	-36.53	-0.800	-37.35
$C_i-(C_d)(C_t)(C)(H)$	-0.140	-63.71	0.165	-46.56	-0.003	-49.57	-0.032	-45.34	-0.366	-43.60	-0.530	-43.85
$C_i-(C_t)_2(C)(H)$	-0.080	-62.13	0.220	-45.13	0.198	-53.33	-0.075	-42.45	-0.341	-44.54	-0.040	-43.12

Table 12: Secondary group additive values ΔGAV° for 1-2 to 1-7 intramolecular hydrogen abstractions deduced from the combined training set and test set. The reference reactions and the meaning of the index i are given in Figure 2. The units for $\Delta\text{GAV}^\circ_{\text{Ea}}$ values are kJ mol^{-1} . For the reference reaction, the single-event pre-exponential factors are expressed in s^{-1} and E_a is expressed in kJ mol^{-1} .

	1-3		1-4			
	i=2		i=2		i=3	
	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log A$	E_a
$\text{C}_i\text{-(C)}_2\text{(H)}_2$	0.000	0.00	0.000	0.00	0.000	0.00
$\text{C}_i\text{-(C)}_3\text{(H)}$	-0.057	-3.32	0.073	-6.62	0.164	-4.52
$\text{C}_i\text{-(C}_d\text{)}\text{(C)}_2\text{(H)}$	-0.080	-2.50	0.064	-9.81	0.067	-6.31
$\text{C}_i\text{-(C}_t\text{)}\text{(C)}_2\text{(H)}$	-0.224	-7.22	-0.112	-7.41	0.069	-2.68
$\text{C}_i\text{-(C)}_4$	0.300	-5.07	-0.043	-7.66	-0.110	-1.57
$\text{C}_i\text{-(C}_d\text{)}\text{(C)}_3$	-0.018	-6.82	0.007	-11.72	-0.020	-5.74
$\text{C}_i\text{-(C}_t\text{)}\text{(C)}_3$	-0.030	-8.57	0.073	-10.11	0.110	-3.79
$\text{C}_i\text{-(C}_d\text{)}_2\text{(C)}_2$	0.199	-8.51	0.037	-10.97	0.087	-2.22
$\text{C}_i\text{-(C}_d\text{)}\text{(C}_t\text{)}\text{(C)}_2$	0.263	-7.52	0.100	-12.50	0.067	-5.75
$\text{C}_i\text{-(C}_t\text{)}_2\text{(C)}_2$	0.001	-6.38	0.060	-12.14	0.053	-3.79

	1-5					
	i=2		i=3		i=4	
	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a
$\text{C}_i\text{-(C)}_2\text{(H)}_2$	0.000	0.00	0.000	0.00	0.000	0.00
$\text{C}_i\text{-(C)}_3\text{(H)}$	-0.011	-4.83	0.100	-3.81	0.043	-2.94
$\text{C}_i\text{-(C}_d\text{)}\text{(C)}_2\text{(H)}$	-0.027	-5.60	0.206	-3.81	-0.025	-3.37
$\text{C}_i\text{-(C}_t\text{)}\text{(C)}_2\text{(H)}$	-0.072	-6.19	-0.018	-1.40	0.030	-0.68
$\text{C}_i\text{-(C)}_4$	-0.134	-6.63	-0.013	-5.95	-0.024	-2.55
$\text{C}_i\text{-(C}_d\text{)}\text{(C)}_3$	-0.003	-11.81	0.238	-9.74	0.098	-1.53
$\text{C}_i\text{-(C}_t\text{)}\text{(C)}_3$	0.061	-9.92	0.350	-7.47	0.176	-4.59
$\text{C}_i\text{-(C}_d\text{)}_2\text{(C)}_2$	0.191	-9.95	0.175	-9.24	0.106	-7.52
$\text{C}_i\text{-(C}_d\text{)}\text{(C}_t\text{)}\text{(C)}_2$	0.270	-9.76	0.150	-3.53	0.300	-3.52
$\text{C}_i\text{-(C}_t\text{)}_2\text{(C)}_2$	0.096	-9.90	-0.083	-5.00	0.211	-0.07

	1-6							
	i=2		i=3		i=4		i=5	
	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a
$\text{C}_i\text{-(C)}_2\text{(H)}_2$	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00
$\text{C}_i\text{-(C)}_3\text{(H)}$	-0.205	-6.87	0.123	-2.57	0.193	-3.84	-0.165	-4.08
$\text{C}_i\text{-(C}_d\text{)}\text{(C)}_2\text{(H)}$	-0.067	-6.53	0.280	-2.37	0.257	-1.10	0.032	-3.79
$\text{C}_i\text{-(C}_t\text{)}\text{(C)}_2\text{(H)}$	-0.108	-4.65	0.130	0.12	0.010	0.85	-0.119	1.35
$\text{C}_i\text{-(C)}_4$	-0.073	-5.46	0.136	-2.62	0.152	-4.57	0.077	-0.21
$\text{C}_i\text{-(C}_d\text{)}\text{(C)}_3$	-0.157	-5.56	0.426	-4.41	0.415	-4.43	-0.039	-2.93
$\text{C}_i\text{-(C}_t\text{)}\text{(C)}_3$	-0.116	-7.77	0.278	-6.53	0.243	-7.81	0.067	-1.00
$\text{C}_i\text{-(C}_d\text{)}_2\text{(C)}_2$	-0.190	-11.05	0.341	-6.10	0.670	-5.83	-0.239	-4.68
$\text{C}_i\text{-(C}_d\text{)}\text{(C}_t\text{)}\text{(C)}_2$	-0.120	-10.68	0.256	-4.79	0.445	-4.44	-0.129	-2.65
$\text{C}_i\text{-(C}_t\text{)}_2\text{(C)}_2$	-0.174	-8.00	0.125	-4.08	0.166	-0.84	-0.020	2.59

(Table 12 continued)

	i=2		i=3		1-7 i=4		i=5		i=6	
	log \tilde{A}	E _a	log \tilde{A}	E _a	log \tilde{A}	E _a	log \tilde{A}	E _a	log \tilde{A}	E _a
C ₁ -(C) ₂ (H) ₂	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00	0.000	0.00
C ₁ -(C) ₃ (H)	0.070	-5.04	0.195	-1.52	0.350	-3.73	0.240	0.23	0.020	-2.37
C ₁ -(C _d)(C) ₂ (H)	0.030	-4.69	0.030	-1.34	0.530	-2.01	0.100	-1.68	-0.400	-2.42
C ₁ -(C _i)(C) ₂ (H)	-0.070	-4.54	0.120	-3.38	0.460	3.20	0.080	-1.99	0.060	2.08
C ₁ -(C) ₄	0.190	-6.84	0.730	-1.26	0.060	4.09	0.150	-5.59	0.310	0.29
C ₁ -(C _d)(C) ₃	0.360	-3.53	0.440	-6.37	1.250	-0.25	0.550	-6.85	0.050	0.84
C ₁ -(C _i)(C) ₃	0.150	-2.74	0.300	-5.23	0.840	1.45	0.220	-3.85	0.430	3.91
C ₁ -(C _d) ₂ (C) ₂	0.240	-6.10	0.430	-10.19	1.080	-4.60	0.430	-9.68	0.670	-1.13
C ₁ -(C _d)(C _i)(C) ₂	0.130	-5.65	0.220	-0.99	1.010	2.12	0.300	-0.89	-0.260	2.43
C ₁ -(C _i) ₂ (C) ₂	0.280	-3.02	0.240	-6.32	0.960	1.62	0.340	-1.60	0.410	7.53

4.4.3.3 Modelling tunneling coefficients

The direct calculation of tunneling coefficients cannot be done from group additivity. Eckart tunneling coefficients are calculated using the electronic barrier and the imaginary frequency. However, these properties are not available without *ab initio* calculations. Instead, correlations are typically constructed to model the tunneling coefficients from properties that can be obtained from group additivity. In this work, the activation energy is used to model the tunneling coefficients, which is a good measure for the electronic barrier. Activation energies can be obtained from group additivity and do not require additional quantum chemical calculations. For each reaction, only the activation energy in the exothermic direction, E_{a,exo}, is considered, i.e. the smallest activation energy when the forward and reverse Arrhenius parameters are compared, since tunneling only pertains to the net electronic barrier

Tunneling coefficients at 300 and 1000K for all the 1-2 up to 1-6 *ab initio* calculated training and test set reactions are depicted in Figure 7. The 1-7 H-shifts have not been included in the figure for the sake of clarity; their values are comparable to 1-6 H-shifts. The tunneling coefficients can be modelled with an exponential function given in Eq. 4.18.

$$\kappa(T) = 1 + a E_{a,exo}^b \quad \text{Eq. 4.18}$$

The temperature dependence of *a* and *b* is difficult to model without major loss in accuracy. Therefore, *a* and *b* are tabulated in Table 13 and linear or second order interpolation is suggested to obtain tunneling coefficients at intermediate temperatures. The linear regression of *a* and *b* led to F values ranging from 104.8 to 198626.6, which is much larger than the tabulated F_{0.05} values;

with the degrees of freedom of each H-shift, a maximum $F_{0.05}$ value of 3.385 is reached. The binary correlation coefficient of $\log a$ and b is very close to -1, the maximum value is -0.996. The uncertainties in Table 13 correspond to the 95% confidence intervals. Most of the uncertainty intervals on the parameters are limited to 20% of the parameter itself, except for 1-3 H-shifts.

Table 13: Coefficients a and b for Eq. 4.18 of the tunneling corrections of 1-2 to 1-7 hydrogen shifts at temperatures from 300 to 1800K. The uncertainties correspond to the 95% confidence intervals.

T (K)	1-2		1-3		1-4	
	$\log a$	b	$\log a$	b	$\log a$	b
300	-13.39 ± 1.41	7.96 ± 0.65	-8.16 ± 1.89	6.02 ± 0.87	-5.03 ± 0.58	3.85 ± 0.30
400	-3.98 ± 0.55	2.45 ± 0.25	-2.57 ± 0.74	2.01 ± 0.34	-1.76 ± 0.26	1.42 ± 0.14
500	-1.96 ± 0.28	1.18 ± 0.13	-1.03 ± 0.35	0.85 ± 0.16	-1.21 ± 0.17	0.86 ± 0.09
600	-1.57 ± 0.21	0.83 ± 0.10	-0.79 ± 0.24	0.55 ± 0.11	-1.13 ± 0.14	0.66 ± 0.07
700	-1.51 ± 0.19	0.70 ± 0.09	-0.81 ± 0.19	0.43 ± 0.09	-1.16 ± 0.12	0.57 ± 0.06
800	-1.52 ± 0.18	0.63 ± 0.08	-0.87 ± 0.17	0.38 ± 0.08	-1.24 ± 0.11	0.53 ± 0.06
900	-1.57 ± 0.17	0.59 ± 0.08	-0.95 ± 0.16	0.35 ± 0.07	-1.32 ± 0.11	0.50 ± 0.06
1000	-1.64 ± 0.17	0.57 ± 0.08	-1.02 ± 0.15	0.33 ± 0.07	-1.40 ± 0.11	0.49 ± 0.06
1100	-1.70 ± 0.17	0.55 ± 0.08	-1.10 ± 0.15	0.31 ± 0.07	-1.50 ± 0.11	0.49 ± 0.06
1200	-1.77 ± 0.17	0.54 ± 0.08	-1.17 ± 0.14	0.31 ± 0.07	-1.57 ± 0.11	0.48 ± 0.06
1300	-1.84 ± 0.18	0.54 ± 0.08	-1.24 ± 0.14	0.30 ± 0.06	-1.65 ± 0.10	0.48 ± 0.05
1400	-1.90 ± 0.18	0.53 ± 0.08	-1.30 ± 0.14	0.30 ± 0.06	-1.72 ± 0.10	0.49 ± 0.05
1500	-1.96 ± 0.18	0.53 ± 0.08	-1.36 ± 0.14	0.30 ± 0.06	-1.78 ± 0.10	0.48 ± 0.05
1600	-2.02 ± 0.18	0.53 ± 0.08	-1.42 ± 0.14	0.30 ± 0.06	-1.84 ± 0.11	0.48 ± 0.06
1700	-2.09 ± 0.19	0.54 ± 0.09	-1.48 ± 0.14	0.29 ± 0.06	-1.89 ± 0.11	0.48 ± 0.06
1800	-2.16 ± 0.19	0.55 ± 0.09	-1.54 ± 0.14	0.30 ± 0.06	-1.94 ± 0.12	0.47 ± 0.06
T (K)	1-5		1-6		1-7	
	$\log a$	b	$\log a$	b	$\log a$	b
300	-2.91 ± 0.31	2.55 ± 0.18	-2.47 ± 0.27	2.26 ± 0.16	-2.50 ± 0.41	2.35 ± 0.23
400	-1.55 ± 0.19	1.27 ± 0.11	-1.37 ± 0.16	1.14 ± 0.09	-1.07 ± 0.23	1.00 ± 0.13
500	-1.35 ± 0.15	0.92 ± 0.09	-1.25 ± 0.13	0.85 ± 0.08	-0.92 ± 0.18	0.67 ± 0.10
600	-1.36 ± 0.13	0.78 ± 0.08	-1.27 ± 0.11	0.72 ± 0.07	-0.97 ± 0.15	0.56 ± 0.09
700	-1.43 ± 0.12	0.72 ± 0.07	-1.35 ± 0.11	0.66 ± 0.06	-1.07 ± 0.14	0.51 ± 0.08
800	-1.53 ± 0.12	0.69 ± 0.07	-1.45 ± 0.11	0.64 ± 0.06	-1.17 ± 0.14	0.48 ± 0.08
900	-1.62 ± 0.12	0.67 ± 0.07	-1.57 ± 0.11	0.63 ± 0.06	-1.29 ± 0.14	0.48 ± 0.08
1000	-1.73 ± 0.12	0.67 ± 0.07	-1.68 ± 0.11	0.64 ± 0.07	-1.40 ± 0.14	0.48 ± 0.08
1100	-1.83 ± 0.13	0.67 ± 0.07	-1.79 ± 0.12	0.65 ± 0.07	-1.50 ± 0.14	0.48 ± 0.08
1200	-1.92 ± 0.13	0.68 ± 0.07	-1.91 ± 0.13	0.66 ± 0.08	-1.59 ± 0.15	0.49 ± 0.08
1300	-2.01 ± 0.14	0.68 ± 0.08	-2.01 ± 0.14	0.68 ± 0.08	-1.69 ± 0.16	0.50 ± 0.08
1400	-2.10 ± 0.15	0.70 ± 0.08	-2.13 ± 0.16	0.70 ± 0.09	-1.77 ± 0.16	0.50 ± 0.09
1500	-2.22 ± 0.16	0.72 ± 0.09	-2.26 ± 0.17	0.74 ± 0.10	-1.86 ± 0.18	0.51 ± 0.09
1600	-2.33 ± 0.16	0.75 ± 0.09	-2.39 ± 0.18	0.78 ± 0.10	-1.93 ± 0.19	0.52 ± 0.10
1700	-2.43 ± 0.17	0.78 ± 0.09	-2.51 ± 0.20	0.81 ± 0.11	-2.05 ± 0.21	0.55 ± 0.11
1800	-2.53 ± 0.18	0.80 ± 0.10	-2.62 ± 0.22	0.84 ± 0.12	-2.16 ± 0.22	0.58 ± 0.11

Eq. 4.18 can only be used to model tunneling between 300 and 1800K, extrapolation outside this temperature range leads to inaccurate results. The correlations are also depicted in Figure 7 and show a good agreement of the tunneling coefficients as a function of the exothermic activation energy, both at 300 and 1000K.

Eq. 4.18 assures that tunneling coefficients are always greater than one, a values are always positive. When the activation energy of the exothermic step approaches zero, the tunneling coefficients will approach 1 at each temperature. Furthermore, a and b are always positive, which leads to a monotonic increase of tunneling coefficients as a function of the activation energy in the exothermic direction. Tunneling coefficients typically decrease with increasing temperature. This is not straightforward to see from the tabulated values of a and b . Table 13 shows that a always exhibits a maximum between 500K and 700K whereas a minimum for b is found between 1500 and 1700K. However, the interplay of both does lead to a monotonous decrease of κ as a function of temperature.

In Figure 7, the necessity to divide the intramolecular hydrogen abstraction reactions as different reaction families is highlighted. A single tunneling correlation will not be able to capture the large differences between 1-2, 1-3 and the other reaction families, which can surpass an order of magnitude at low temperatures. Tunneling coefficients of 1-4 up to 1-7 hydrogen shifts could be modelled together without much loss in accuracy, but the categorization for 1-2 and 1-3 hydrogen shifts is kept for other reaction families as well for reasons of consistency.

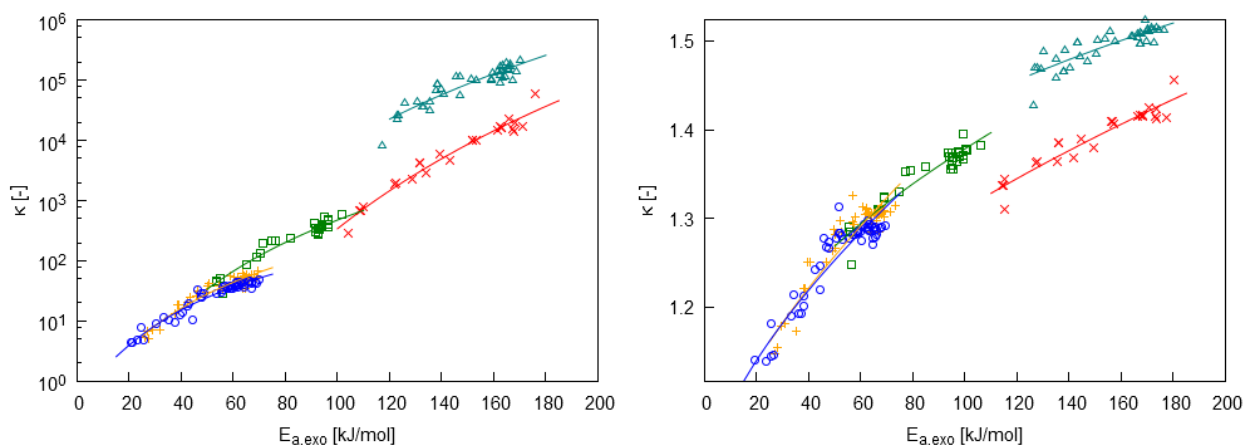


Figure 7: Tunneling coefficients of all the 1-2 (\times), 1-3 (\triangle), 1-4 (\square), 1-5 ($+$), and 1-6 (\circ) hydrogen abstraction reactions (points) at 300K (left) and 1000K (right) compared to the correlations (lines). 1-7 shifts have been left out for the sake of clarity; the values are of the same order of magnitude as 1-6 shifts.

4.4.3.4 Dependence of group additive values on the type of H-shift

In the past¹²⁻¹⁵, intramolecular hydrogen abstraction rate coefficients have often been calculated by using the rate coefficient of a similar intermolecular hydrogen abstraction and adding a single correction depending on the ring size in the transition state, i.e. the type of H-shift. The reactive atoms of the intermolecular hydrogen abstraction should have a similar surrounding compared to the intramolecular reaction. Although this approach might be accurate for a limited number of reactions, it cannot be generalized to the whole intramolecular hydrogen abstraction reactions family. The effect of certain groups on the rate coefficient is dependent on the type of H-shift, and for each type, a separate contribution was thus calculated in this work.

Secondary contributions, which are due to substituents on the carbon chain in between the two reactive atoms strongly depend on the length of that carbon chain. These contributions can thus not be generalized for all H-shifts. For example, if a hydrogen atom bonded to the adjacent carbon atom of the radical is substituted by a vinyl group, the influence on a 1-4 H-shift activation energy amounts $-10.97 \text{ kJ mol}^{-1}$. For a 1-7 H-shift, the activation energy is only decreased by 1.13 kJ mol^{-1} . The difference between both is more than 9 kJ mol^{-1} and should thus not be neglected.

Furthermore, primary contributions also need to be treated separately. Table 11 shows, for example, that two vinyl substituents on the attacked carbon atom have a contribution of $-62.70 \text{ kJ mol}^{-1}$ in 1-2 H-shifts, but only $-37.35 \text{ kJ mol}^{-1}$ in 1-7 H-shifts. At 1000K, this leads to a deviation of a factor of 20 on the rate coefficient.

4.4.3.5 Temperature dependence of group additive values

Previously generated group additivity models for activation energies and pre-exponential factors²⁻¹¹ observed that ΔGAV^0 only show a minor temperature dependence because (1) Arrhenius parameters do not vary much when regressed over several temperature domains and (2) the temperature dependence is mostly covered by the Arrhenius parameters of the reference reactions.

The temperature dependence is verified by calculating the group additive values at 300 to 1800K and relating the values to the value at 300K as show in Figure 8. The figure only shows the group additive values with the highest temperature dependence per H-shift. Most of the group additive

values with high temperature dependence correspond to the $C_1-(C_d)_1(C)$ group. 1-6 H-shifts exhibit the highest temperature dependence, followed by 1-7 and 1-5 H-shifts. The temperature dependence in 1-2 and 1-3 shifts remains limited. The difference between 300K and 1800K reaches to 18 kJ mol^{-1} for the $C_1-(C_d)_1(C)$ group in 1-6 H-shifts. The temperature dependence can thus become significant depending on the substituents of the reactive atoms and the carbon chain in between. Group additive values should thus only be used in the relevant temperature range where they were regressed from, which is 800 to 1200K for the values in Table 11 and Table 12.

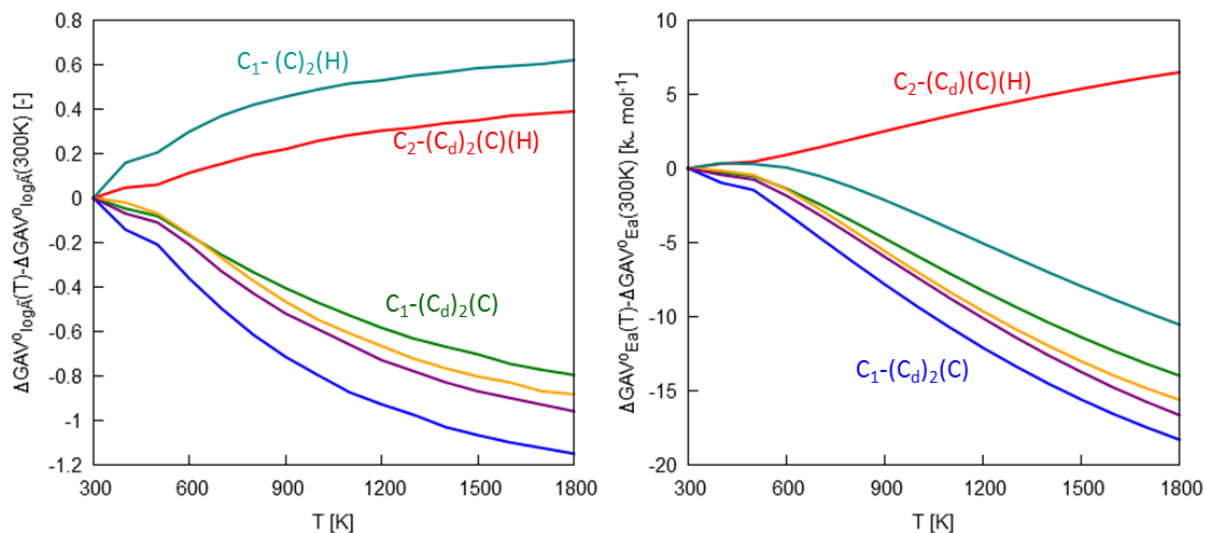


Figure 8: Temperature dependence of the ΔGAV° 's with the highest temperature dependence for the 1-2 (—), 1-3 (—), 1-4 (—), 1-5 (—), 1-6 (—), and 1-7 (—) H-shifts. The temperature dependence is obtained by subtracting the group additive value at 300K from the group additive values at a given temperature. The lines to which no label are added all correspond to the $C_1-(C_d)_1(C)$ group.

4.4.3.6 Thermodynamic consistency

One important note to be made when using a group additive model is that thermodynamic consistency is not necessarily met. Thermodynamic consistency is met if the ratio of the forward and reverse rate coefficients of a reaction equals the equilibrium coefficient of that reaction. When calculating the forward and reverse rate coefficients of a reaction using group additivity, their ratio is not equal to the equilibrium coefficient obtained from the *ab initio* thermodynamic parameters of the reactant and product, as shown in Table 14. The highest deviation for the 1-2 H-shifts occurs in a reaction of which both the reactive atoms have non-hydrogen ligands. These reactions also show a higher deviation in the group additivity rate coefficients compared to the *ab initio* one. Furthermore, this deviation is in opposite direction for the forward versus the reverse rate coefficient, leading to the high deviation in equilibrium coefficients. Another higher

mismatch between the group additivity and the *ab initio* equilibrium coefficients is seen for a 1-7 H-shift in which the group additivity reverse rate coefficient deviates with a factor of three compared to the *ab initio* one, as shown in Appendix E. Besides these two deviations, most group additivity equilibrium coefficients agree well with the *ab initio* coefficients.

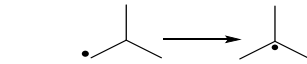
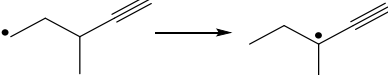
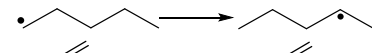
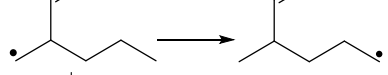
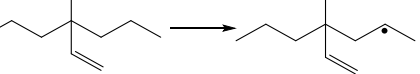
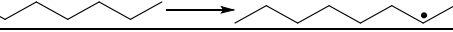
Table 14: Mean absolute deviation (MAD) and maximum (MAX) deviation between the group additivity equilibrium coefficients and the *ab initio* equilibrium coefficients evaluated at 1000K of the training and test reactions. All the deviations are calculated by dividing the highest equilibrium coefficient by the lowest one.

	1-2		1-3		1-4	
	MAD	MAX	MAD	MAX	MAD	MAX
Training set	1.40	2.01	1.24	1.46	1.22	1.53
Test set	1.40	3.02	1.23	1.64	1.22	1.65
	1-5		1-6		1-7	
	MAD	MAX	MAD	MAX	MAD	MAX
Training set	1.22	1.87	1.21	1.67	1.16	2.76
Test set	1.22	1.84	1.22	1.65	2.37	4.15

During the generation of a kinetic model, often no thermodynamic data of species is available directly from *ab initio* calculations. Instead, calculation procedures such as group additivity are used to obtain thermodynamic data of all the species. These procedures are only applicable to a limited group of molecules, and have an accuracy lower than direct *ab initio* calculations. It is thus key to know if these data are sufficient to calculate the reverse reaction rate coefficients from thermodynamic consistency. First, the applicability domain of the calculation procedure needs to be verified. In the framework of Genesys, thermodynamic data is calculated based on the group additivity values of Marin and coworkers^{53,54}. These values allow the calculation of standard enthalpies of formations, entropies and heat capacities for open chain hydrocarbon molecules and radicals, including double bonds, triple bonds and allenic bonds. Furthermore, the calculations also contain monoaromatic species and a limited set of aliphatic cyclic molecules and radicals. The application domain of the current work, i.e. reactions in open chain hydrocarbon radicals, including double and triple bonds, is thus covered by the thermochemistry calculations. Second, the accuracy of the group additivity thermochemistry data needs to be verified. Sabbe et al.⁵³ reported a mean absolute deviation of the group additivity standard enthalpy of formation versus the *ab initio* calculated standard enthalpy of formation better than 2kJ mol⁻¹, which is good compared to CBS-QB3 accuracies of 4kJ mol⁻¹. Furthermore, entropies and heat capacities are within 5J mol⁻¹ according to Sabbe et al.⁵⁴ Besides these reported uncertainties, the accuracy can be controlled by calculated equilibrium coefficients from group

additivity and comparing them to the *ab initio* equilibrium coefficients. Table 15 shows 6 reactions for which the equilibrium coefficient has been calculated from the group additivity thermochemistry of the reactants and products $K_{eq,GA}$ compared to the ratio of the forward and reverse rate coefficients. The ratios have been calculated by *ab initio* and by group additivity. In general, calculating the reverse rate coefficient using $K_{eq,GA}$ gives an addition uncertainty of a factor of 2 compared to the *ab initio* reverse rate coefficient.

Table 15 : Comparison of the group additivity equilibrium coefficients calculated from the thermodynamic data of Sabbe et al.^{53,54} to the ratio of the forward and reverse rate coefficients calculated from *ab initio* (AI) and from group additivity (GA).

Reaction	$K_1 = \frac{k_{for,AI}}{k_{rev,AI}}$	$K_2 = \frac{k_{for,GA}}{k_{rev,GA}}$	$K_{eq,GA}$	$\frac{K_1}{K_2}$	$\frac{K_1}{K_{eq,GA}}$
	$1.4 \cdot 10^3$	$9.2 \cdot 10^2$	$3.3 \cdot 10^3$	1.45	0.41
	$8.8 \cdot 10^3$	$9.6 \cdot 10^3$	$1.7 \cdot 10^4$	0.92	0.53
	$9.6 \cdot 10^0$	$1.1 \cdot 10^1$	4.410^0	0.92	2.19
	$7.8 \cdot 10^{-1}$	$7.7 \cdot 10^{-1}$	$1.0 \cdot 10^0$	1.03	0.79
	$6.0 \cdot 10^0$	$8.0 \cdot 10^0$	$4.4 \cdot 10^0$	0.75	1.36
	$4.5 \cdot 10^0$	$3.3 \cdot 10^0$	$4.4 \cdot 10^0$	1.38	1.04

4.4.4 Application of the group additivity model

The calculation of Arrhenius parameters of an intramolecular hydrogen abstraction reaction using the new group additivity model requires several steps. This is illustrated based on the reaction in Figure 9 for which the group additivity Arrhenius parameters will be calculated at 1000K. The reference Arrhenius parameters first need to be known. The example reaction belongs to the 1-4 H-shift family for which the logarithm of the reference single event pre-exponential factor is 10.823 and reference activation energy is 106.03 kJ mol⁻¹. In a 1-4 H-shift, four contributions ΔGAV° 's are necessary: two primary and two secondary contributions. The primary contributions in the example reaction are C₁-(C)₂(H) and C₄-(C)₃(H), with 0.034 and -0.070 as $\Delta GAV^\circ_{\log \tilde{A}}$ and 4.48 and -22.78 as ΔGAV°_{Ea} , respectively. The two secondary groups are C₂-(C)₂(C_d)(H) and C₃-(C)₂(H)₂, the latter corresponds to the reference reaction and its ΔGAV° 's are thus zero both for the pre-exponential factor as well as for the activation energy. The former however does deviate from the reference reaction, $\Delta GAV^\circ_{\log \tilde{A}}$ is 0.064 and ΔGAV°_{Ea} is -9.81. The resulting activation energy and single-event pre-exponential factor are:

$$\begin{aligned}
 E_a(1000K) &= E_{a,ref}(1000K) + \sum_i \Delta GAV_{E_a}^o(C_i, 1000K) \\
 &= 106.03 + 4.48 + (-22.78) + (-9.81) + 0.0 = 77.92
 \end{aligned}
 \tag{Eq. 4.19}$$

$$\begin{aligned}
 \log \tilde{A}(1000K) &= \log \tilde{A}_{ref}(1000K) + \sum_i \Delta GAV_{\log \tilde{A}}^o(C_i, 1000K) \\
 &= 10.823 + 0.034 + (-0.070) + 0.064 + 0.0 = 10.851
 \end{aligned}
 \tag{Eq. 4.20}$$

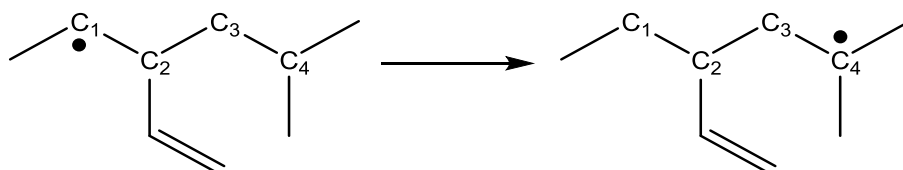


Figure 9: Illustration of the calculation of Arrhenius parameters from group additivity for intramolecular hydrogen abstraction reactions.

Finally, the single event pre-exponential factor needs to be multiplied by two factors: the tunneling coefficient and the number of single events. The tunneling coefficient at 1000K for 1-4 H-shifts is calculated according to $1 + 10^{-1.40} E_{a,exo}^{0.49}$, see Table 13. The activation energy in exothermic direction needs to be known. The activation energy in the forward direction has already been calculated and amounts $77.91 \text{ kJ mol}^{-1}$. In the reverse direction, the activation can be calculated analogously and amounts $91.16 \text{ kJ mol}^{-1}$. The lowest activation energy belongs to the exothermic direction, and the resulting tunneling coefficient is 1.34.

To calculate the number of single events, the total rotational symmetry numbers and the numbers of optical isomers of the reactant and transition state are needed. Neither contain any external rotational symmetry, but have three methyl groups leading to an internal symmetry number of 27. The number of optical isomers of the reactant is 2, C_2 is a chiral center. The number of optical isomers in the transition state is 4, both C_1 and C_2 are chiral centers. The transition state thus exists as two pairs of diastereomers. In this work, the difference in energy between these diastereomers is neglected, instead the number of optical isomers is set to 4. The number of single events is thus 2. The total pre-exponential factor can now be calculated:

$$\begin{aligned}
 \log A(1000K) &= \log \tilde{A} + \log(\kappa \cdot n_e) \\
 &= 10.851 + \log(1.34 \cdot 2) = 11.279
 \end{aligned}
 \tag{Eq. 4.21}$$

4.4.4.1 1-octyl radical

The 1-octyl radical can isomerize via 6 H-shifts yielding 2-, 3-, and 4-octyl radicals, as shown in Figure 10. The calculation of rate coefficients for these 6 reactions can be done using the newly

developed group additive model. The methodology is similar for all reactions, and is only illustrated for the 1-7 H-shift. The primary groups in this reaction are $C_1-(C)(H)_2$ and $C_7-(C)_2(H)_2$. The Arrhenius parameters of 1-7 H-shifts also depend on 5 secondary contributions which are all $C_1-(C)_2(H)_2$ in this example. These are identical to the secondary groups in the reference reaction, and their group additive value are thus zero. The pre-exponential factor and activation energy at 1000K are calculated according to Eq. 4.22 and Eq. 4.23. The tunneling correction is calculated using the activation energy of $63.68 \text{ kJ mol}^{-1}$, which is in the exothermic direction, the reverse activation energy is $80.28 \text{ kJ mol}^{-1}$. The numbers of single event for all the reactions in Figure 10 are equal to 4.

$$E_a(1000K) = 75.99 + (-12.31) = 63.68 \quad \text{Eq. 4.22}$$

$$\log A(1000K) = 7.882 + (-0.010) + \log((1 + 10^{-1.40} \cdot 63.68^{0.48}) \cdot 4) = 8.595 \quad \text{Eq. 4.23}$$

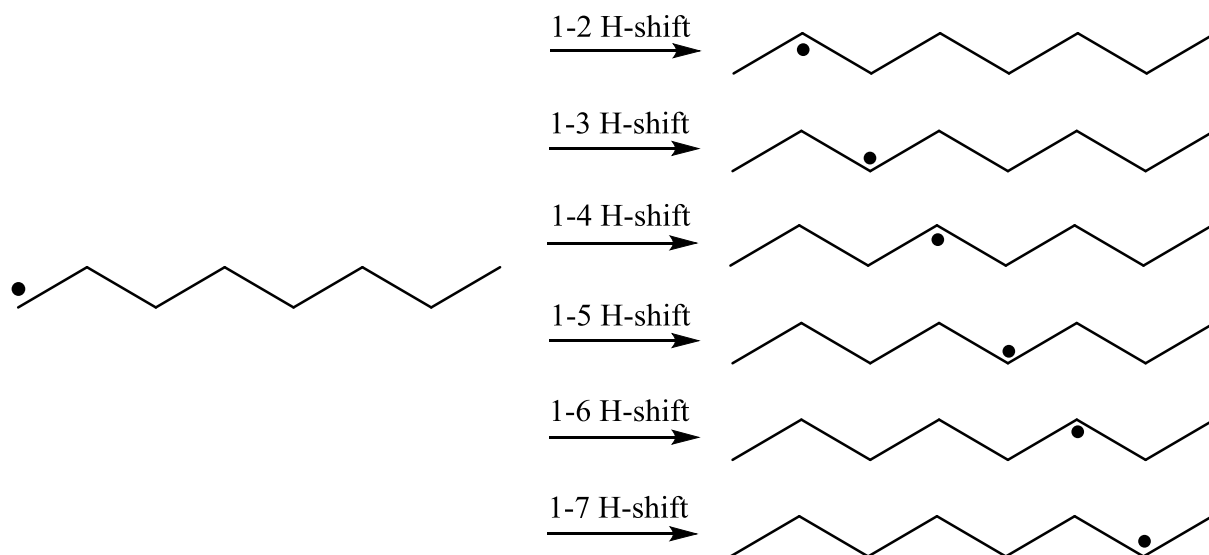


Figure 10: H-shifts in the 1-octyl radical

The other 5 H-shifts in the 1-octyl radical can be calculated in an analogous manner. Each of the 6 reactions has the same two primary group contributions: $C_1-(C)(H)_2$ and $C_2-(C)_2(H)_2$, all the secondary contributions are zero. The branching fractions, i.e. the rate coefficients of each H-shift reaction compared to the sum of the 6 rate coefficients, are given in Figure 11a. The rate coefficients themselves are depicted in Figure 11b. As is already known for this reaction family, 1-7 H-shifts only play a very small role compared to the other 5 H-shifts. The branching fraction of this 1-7 H-shift does not exceed 0.004 in the studied temperature domain. At low temperatures,

1-5 and 1-6 H-shifts are the most important, due to enthalpy effects. When the temperature rises, 1-4 H-shifts become more important, followed by 1-2 and 1-3 H-shifts. The transition states of these three H-shifts have higher ring strain, and thus a less advantageous enthalpy of activation. However, because less rotors are frozen when the open chain reactants change to a cyclic structure in the transition state, their entropy of activation is increased. The same is visible when looking at the rate coefficients themselves. The slope of the 1-5 to 1-7 H-shifts is relatively flat, but their extrapolated intercept with the y-axis is lower compared to 1-2 and 1-3 H-shifts. 1-4 H-shift exhibits an intermediate interaction of the slope and extrapolated intercept with the y-axis.

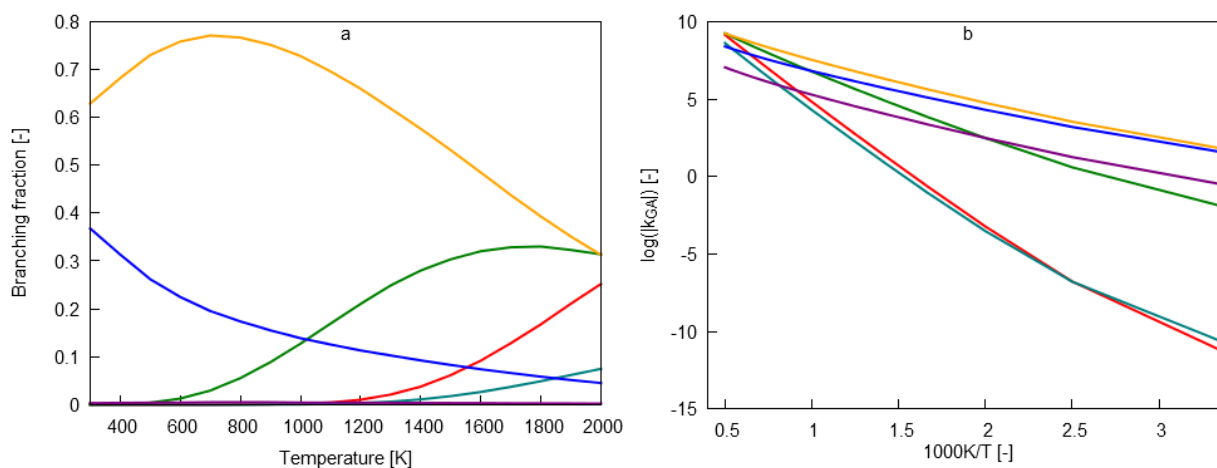


Figure 11: 1-2 (—), 1-3 (—), 1-4 (—), 1-5 (—), 1-6 (—), and 1-7 (—) H-shifts in the 1-octyl radical: (a) branching fractions of each H-shift compared to the total H-shift rate coefficient and (b) rate coefficients of the 6 reactions in which $\log(k/k_A)$ is reported. The rate coefficients k are expressed in s^{-1} .

It should also be noted that, although the 1-2 and 1-3 H-shift seem to gain importance at higher temperature, many radical species will decompose to bimolecular products instead at these temperatures.

4.4.4.2 Applicability of the model

The model only includes group additive values, no non-nearest neighbor effects have been included so far. However, a few reactions do show larger deviations. These deviations arise when two neighboring carbon atom have groups that come in each other's vicinity. One of the most important of these deviations is seen when both the attacking and attacked carbon atoms are tertiary atoms instead of primary atoms. Although the total group additive rate coefficients agrees well with the *ab initio* values, both the activation energy and the pre-exponential factor show a too high deviation. The activation energy in a 1-2 shift is 14.4 kJ mol^{-1} too low via group

additivity, 5.3 kJ mol^{-1} in a 1-3 H-shift and 5.0 kJ mol^{-1} in a 1-4 H-shift. The methyl groups on each side of the transition state thus influence each other, and their influence is higher when the transition state ring is smaller.

In a similar fashion, the need for non-nearest neighbor interactions was verified by adding three methyl groups on the three carbon atoms between the reactive atoms in a 1-5 H-shift. In this case, the predicted activation energy is 4.5 kJ mol^{-1} too high.

All the other predicted activation energies were within 4 kJ mol^{-1} compared to the *ab initio* values. Nevertheless, future calculations to come up with a consistent set of non-nearest neighbor interactions is required to increase the accuracy of a limited number of reactions.

4.5 Conclusions

One of the new features of Genesys, i.e. automatically performing *ab initio* calculations, has been illustrated in this chapter by building a group additive model for intramolecular hydrogen abstractions in hydrocarbons. The geometries of all species and reactions have been generated by Genesys and submitted to Gaussian to obtain electronic energies and frequencies of the lowest energy conformers. Furthermore, 1D hindered rotor potentials have been calculated for all molecular frequencies resembling torsional modes. Rate coefficients have subsequently been calculated for all reactions, leading to a large set of kinetic data. These results have been compared to previous experimental and theoretic studies of intramolecular hydrogen abstractions and show a good agreement to literature data.

The rate coefficients have been used to initially calculate a set of group additive values from a limited training set. These values have been used to calculate the rate coefficients of the reactions of the test set and have been compared to additional *ab initio* rate coefficients. Good agreement was found between both, i.e. most group additivity rate coefficients are within a factor of 2 of the *ab initio* values, validating the group additive approach for intramolecular hydrogen abstraction reactions. In a next step, both the training set and test set have been used to regress group additive values in order to increase the accuracy of the group additive model.

Since tunneling corrections cannot be modelled by group additivity, correlations have been obtained to calculate the transmission probability of intramolecular hydrogen abstractions. These correlations follow an exponential trend as function of the exothermic activation energies of the reaction, which is a property that can be modelled with group additivity. The parameters of the correlation are strongly temperature dependent, and their temperature dependence is too complex to be accurately modelled. Instead, the parameters have been tabulated and interpolation between these values is required to obtain tunneling coefficients at intermediate temperatures.

Compared to alternative methods for calculating rate coefficients of intramolecular hydrogen abstractions, the current approach is more elaborate by taking both the immediate surroundings of the attacking and attacked carbon atom into account, as well as substituents on the carbon chain that connect these carbon atoms. The surroundings considered have also been expanded to multiple sp^2 and sp hybridized carbon atoms.

4.6 References

1. Benson SW, Buss JH. Additivity Rules for the Estimation of Molecular Properties. Thermodynamic Properties. *The Journal of Chemical Physics*. 1958;29(3):546-572.
2. Saeys M, Reyniers M-F, Marin GB, Van Speybroeck V, Waroquier M. Ab Initio Calculations for Hydrocarbons: Enthalpy of Formation, Transition State Geometry, and Activation Energy for Radical Reactions. *The Journal of Physical Chemistry A*. 2003;107(43):9147-9159.
3. Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Ab initio group contribution method for activation energies for radical additions. *Aiche Journal*. 2004;50(2):426-444.
4. Sabbe MK, Reyniers MF, Van Speybroeck V, Waroquier M, Marin GB. Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors. *Chemphyschem*. 2008;9(1):124-140.
5. Sabbe MK, Vandeputte AG, Reyniers MF, Waroquier M, Marin GB. Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions. *Physical Chemistry Chemical Physics*. 2010;12(6):1278-1298.
6. Vandeputte AG, Sabbe MK, Reyniers M-F, Marin GB. Kinetics of alpha hydrogen abstractions from thiols, sulfides and thiocarbonyl compounds. *Physical Chemistry Chemical Physics*. 2012;14(37):12773-12793.
7. Vandeputte AG, Reyniers M-F, Marin GB. Kinetics of Homolytic Substitutions by Hydrogen Atoms at Thiols and Sulfides. *Chemphyschem*. 2013;14(8):1703-1722.
8. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Kinetic Modeling of α -Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Hydrogen Atoms. *Journal of Physical Chemistry A*. 2014;118(40):9296-9309.
9. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos N, Marin GB. Kinetic Modeling of alpha-Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Carbon-Centered Radicals. *Chemphyschem*. 2014;15(9):1849-1866.
10. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Group Additive Kinetics for Hydrogen Transfer Between Oxygenates. *Journal of Physical Chemistry A*. 2015;119(27):6961-6980.
11. Paraskevas PD, Sabbe MK, Reyniers MF, Marin GB, Papayannakos NG. Group additive kinetic modeling for carbon-centered radical addition to oxygenates and -scission of oxygenates. *Aiche Journal*. 2016;62(3):802-814.
12. Benson SW. *Thermochemical kinetics: methods for the estimation of thermochemical data and rate parameters*: Wiley; 1976.
13. Curran HJ, Gaffuri P, Pitz WJ, Westbrook CK. A Comprehensive Modeling Study of n-Heptane Oxidation. *Combustion and Flame*. 1998;114(1-2):149-177.
14. Curran HJ, Gaffuri P, Pitz WJ, Westbrook CK. A comprehensive modeling study of iso-octane oxidation. *Combustion and Flame*. 2002;129(3):253-280.
15. Curran HJ, Pitz WJ, Westbrook CK, Dagaut P, Boettner JC, Cathonnet M. A wide range modeling study of dimethyl ether oxidation. *International Journal of Chemical Kinetics*. 1998;30(3):229-241.

16. Zheng J, Truhlar DG. Direct Dynamics Study of Hydrogen-Transfer Isomerization of 1-Pentyl and 1-Hexyl Radicals. *The Journal of Physical Chemistry A*. 2009;113(43):11919-11925.
17. Hayes CJ, Burgess DR. Kinetic Barriers of H-Atom Transfer Reactions in Alkyl, Allylic, and Oxoallylic Radicals as Calculated by Composite Ab Initio Methods. *The Journal of Physical Chemistry A*. 2009;113(11):2473-2482.
18. Ratkiewicz A, Bankiewicz B. Kinetics of 1,5-Hydrogen Migration in Alkyl Radical Reaction Class. *The Journal of Physical Chemistry A*. 2012;116(1):242-254.
19. Bankiewicz B, Huynh LK, Ratkiewicz A, Truong TN. Kinetics of 1,4-Hydrogen Migration in the Alkyl Radical Reaction Class. *The Journal of Physical Chemistry A*. 2009;113(8):1564-1573.
20. Ratkiewicz A, Bankiewicz B, Truong TN. Kinetics of thermoneutral intermolecular hydrogen migration in alkyl radicals. *Physical Chemistry Chemical Physics*. 2010;12(36):10988-10995.
21. Sirjean B, Dames E, Wang H, Tsang W. Tunneling in Hydrogen-Transfer Isomerization of n-Alkyl Radicals. *The Journal of Physical Chemistry A*. 2012;116(1):319-332.
22. Lu D-h, Truong TN, Melissas VS, Lynch GC, Liu Y-P, Garrett BC, Steckler R, Isaacson AD, Rai SN, Hancock GC, Lauderdale JG, Joseph T, Truhlar DG. POLYRATE 4: A new version of a computer program for the calculation of chemical reaction rates for polyatomics. *Computer Physics Communications*. 1992;71(3):235-262.
23. Liu YP, Lynch GC, Truong TN, Lu DH, Truhlar DG, Garrett BC. Molecular modeling of the kinetic isotope effect for the [1,5]-sigmatropic rearrangement of cis-1,3-pentadiene. *Journal of the American Chemical Society*. 1993;115(6):2408-2415.
24. Wigner E. Crossing of Potential Thresholds in Chemical Reactions. *Zeitschrift für Physikalische Chemie*. 1932;B19:203-216.
25. Skodje RT, Truhlar DG. Parabolic tunneling calculations. *The Journal of Physical Chemistry*. 1981;85(6):624-628.
26. Eckart C. The Penetration of a Potential Barrier by Electrons. *Physical Review*. 1930;35(11):1303-1309.
27. Davis AC, Francisco JS. Ab Initio Study of Hydrogen Migration across n-Alkyl Radicals. *The Journal of Physical Chemistry A*. 2011;115(14):2966-2977.
28. Davis AC, Francisco JS. Ab Initio Study of Key Branching Reactions in Biodiesel and Fischer-Tropsch Fuels. *Journal of the American Chemical Society*. 2011;133(47):19110-19124.
29. Wang K, Villano SM, Dean AM. The Impact of Resonance Stabilization on the Intramolecular Hydrogen-Atom Shift Reactions of Hydrocarbon Radicals. *Chemphyschem*. 2015;16(12):2635-2645.
30. Wang K, Villano SM, Dean AM. Reactivity-Structure-Based Rate Estimation Rules for Alkyl Radical H Atom Shift and Alkenyl Radical Cycloaddition Reactions. *The Journal of Physical Chemistry A*. 2015;119(28):7205-7221.
31. Bian H, Wang Z, Zhang F, Wang Z, Zhu J. Unimolecular Reaction Properties for the Long-Chain Alkenyl Radicals. *International Journal of Chemical Kinetics*. 2015;47(11):685-694.
32. Cohen N, Benson SW. Estimation of heats of formation of organic compounds by additivity methods. *Chemical Reviews*. 1993;93(7):2419-2438.

33. Sumathi R, Carstensen HH, Green WH. Reaction rate prediction via group additivity, part 2: H-abstraction from alkenes, alkynes, alcohols, aldehydes, and acids by H atoms. *Journal of Physical Chemistry A*. 2001;105(39):8969-8984.
34. Sumathi R, Carstensen HH, Green WH. Reaction rate prediction via group additivity Part 1: H abstraction from alkanes by H and CH₃. *Journal of Physical Chemistry A*. 2001;105(28):6910-6925.
35. Sumathi R, Carstensen HH, Green WH. Reaction rate predictions via group additivity. Part 3: Effect of substituents with CH₂ as the mediator. *Journal of Physical Chemistry A*. 2002;106(22):5474-5489.
36. *Gaussian 03* [computer program]. Wallingford, CT, USA: Gaussian, Inc.; 2004.
37. Vandewiele NM, Van Geem KM, Reyniers MF, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chemical Engineering Journal*. 2012;207:526-538.
38. Vandewiele NM, Van de Vijver R, Van Geem KM, Reyniers M-F, Marin GB. Symmetry calculation for molecules and transition states. *Journal of Computational Chemistry*. 2015;36(3):181-192.
39. Vandewiele NM, Van De Vijver R, Carstensen H-H, Van Geem KM, Reyniers M-F, Marin GB. Implementation of Stereochemistry in Automatic Kinetic Model Generation. *International Journal of Chemical Kinetics*. 2016;48(12):755-769.
40. Pollak E, Pechukas P. Symmetry Numbers, Not Statistical Factors, Should be Used in Absolute Rate Theory and in Bronsted Relations. *Journal of the American Chemical Society*. 1978;100(10):2984-2991.
41. Evans MG, Polanyi M. Further considerations on the thermodynamics of chemical equilibria and reaction rates. *Transactions of the Faraday Society*. 1936;32(0):1333-1360.
42. Blowers P, Masel RI. Engineering approximations for activation energies in hydrogen transfer reactions. *Aiche Journal*. 2000;46(10):2041-2052.
43. Miyoshi A, Widjaja J, Yamauchi N, Koshi M, Matsui H. Direct investigations on the thermal unimolecular isomerization reaction of 1-pentyl radicals. *Proceedings of the Combustion Institute*. 2002;29(1):1285-1293.
44. Yamauchi N, Miyoshi A, Kosaka K, Koshi M, Matsui H. Thermal Decomposition and Isomerization Processes of Alkyl Radicals. *The Journal of Physical Chemistry A*. 1999;103(15):2723-2733.
45. Tsang W, Walker JA, Manion JA. Single-pulse shock-tube study on the decomposition of 1-pentyl radicals. *Symposium (International) on Combustion*. 1998;27(1):135-142.
46. Marshall RM. The rate constant for the intramolecular isomerization of pentyl radicals. *International Journal of Chemical Kinetics*. 1990;22(9):935-950.
47. Endrenyi L, Roy DJL. The Isomerization of n-Pentyl and 4-Oxo-1-pentyl Radicals in the Gas Phase. *The Journal of Physical Chemistry*. 1966;70(12):4081-4084.
48. Watkins KW. Photolysis of n-pentylazomethane vapor. Reactions of the n-pentyl radical. *Journal of the American Chemical Society*. 1971;93(24):6355-6359.
49. Tsang W, Walker JA, Manion JA. The decomposition of normal hexyl radicals. *Proceedings of the Combustion Institute*. 2007;31(1):141-148.
50. Imbert FE, Marshall RM. The mechanism and rate parameters for the pyrolysis of n-hexane in the range 723–823 K. *International Journal of Chemical Kinetics*. 1987;19(2):81-103.
51. Dóbbé S, Bérces T, Réti F, Márta F. Isomerization of n-hexyl and s-octyl radicals by 1,5 and 1,4 intramolecular hydrogen atom transfer reactions. *International Journal of Chemical Kinetics*. 1987;19(10):895-921.

52. Watkins KW, Ostreko LA. Isomerization of n-hexyl radicals in the gas phase. *The Journal of Physical Chemistry*. 1969;73(6):2080-2083.
53. Sabbe MK, Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *Journal of Physical Chemistry A*. 2005;109(33):7466-7480.
54. Sabbe MK, De Vleeschouwer F, Reyniers MF, Waroquier M, Marin GB. First Principles Based Group Additive Values for the Gas Phase Standard Entropy and Heat Capacity of Hydrocarbons and Hydrocarbon Radicals. *Journal of Physical Chemistry A*. 2008;112(47):12235-12251.

Chapter 5: First principles based automated kinetic modeling of heptane pyrolysis

5.1 Abstract

This chapter demonstrates the automated generation of a kinetic model for n-heptane pyrolysis using Genesys. n-Heptane has been chosen because it is an important surrogate naphtha and a large experimental dataset is available to validate the pyrolysis behavior. The considered reaction families are: hydrogen abstraction reactions, radical addition to multiple bonds and the reverse β -scission reactions. For additions and abstractions, both intermolecular and intramolecular reactions have been considered. The thermochemistry and the reaction rate coefficients of all the steps in the primary decomposition of n-heptane have been calculated by group additivity. Besides the primary decomposition chemistry, the chemistry of small compounds and the formation of aromatics have been incorporated in the model via literature reactions and kinetics. The final kinetic model contains 4221 reactions between 691 species. However, several species and reactions do not contribute to the pyrolysis of n-heptane, and model reduction steps using the DREGP method led to a more concise kinetic model, i.e. 1043 reactions and 154 species, significantly improving simulation times without any change in predicted reactant conversion or products yields. The kinetic model has been compared to two independently reported experimental datasets, showing a good agreement for most of the response variables without any

tuning of the first principles derived reaction rate coefficients, and performing better than previously reported mechanisms. A rate of production analysis was employed to find the main reaction pathways. Finally, considering intramolecular hydrogen abstractions (H-shift reactions), for which group additive values were calculated in previous chapter, proves to be important, especially to understand the formation of C5 and C6 olefins, which lead to soot precursors such as benzene and 1,3-cyclopentadiene.

5.2 Introduction

n-Heptane is one of the main linear alkanes in crude oil, and can be found in industrial processes as an important feed component for the petrochemical industry, as well as energy carrier both for gasoline or as surrogate for diesel. Its pyrolysis and combustion have been widely investigated to better understand in the decomposition and ignition chemistry relevant for these processes. Especially for pyrolysis, both experimental and modeling studies have been carried out.¹⁻¹¹

Murata et al.¹ developed a free-radical reaction mechanism for heavy paraffinic hydrocarbons able to predict the product distribution under pyrolysis conditions. The authors stressed the need for radical isomerization reactions, which they tackled by a heuristic process that decreases the concentrations of primary radicals and increases the concentrations of secondary radicals. Their results showed good agreement with experimental pyrolysis data for n-butane to n-hexadecane. Experiments for steam cracking of n-heptane were done by Bajus et al.² at atmospheric pressure and temperatures of 953 to 1033K. The reactor effluent was analyzed using gas chromatography. The conversion of n-heptane was approximated by an Arrhenius equation with a pre-exponential factor of $1.34 \cdot 10^{11} \text{ s}^{-1}$ and an activation energy of $195.5 \text{ kJ mol}^{-1}$. Aribike and Susu³ investigated steam cracking of n-heptane in a temperature range of 933 to 1053K at atmospheric pressure. The authors also reported a first order reaction describing the conversion of n-heptane, with a pre-exponential factor of $5.88 \cdot 10^{10} \text{ s}^{-1}$ and an activation energy of $206.1 \text{ kJ mol}^{-1}$. The formation of products has been modelled using secondary reactions including hydrogen abstraction reactions and addition and β -scission reactions. The same authors⁴ proposed a free radical mechanism for steam cracking of n-heptane. This mechanism also includes the formation of benzene via two pathways: Dehydrocyclization of alkenyl radicals and a Diels-Alder reaction. Both pathways are continued via the loss of hydrogen gas molecules. The kinetics of all the reactions in the

mechanism were taken from literature or estimated via the experimental data. Pant and Kunzru⁵ studied the kinetics and the product distribution during steam cracking of n-heptane in a temperature range of 953-1023K under atmospheric pressure in a tubular reactor. A kinetic model consisting of one global reaction describing the formation of the main products from n-heptane, together with 24 non-elementary reactions between the primary products was constructed, with the rate coefficients calculated using the experimental data. Again, a single Arrhenius equation was used to model the conversion of n-heptane, with a pre-exponential factor of $6.02 \cdot 10^{13} \text{ s}^{-1}$ and an activation energy of $250.7 \text{ kJ mol}^{-1}$. Held et al.⁶ developed a semi-empirical mechanism for n-heptane oxidation and pyrolysis and validated the model using independent datasets. The final model consist of 266 reactions between 41 species. The pyrolysis of n-heptane was also studied by Chakraborty and Kunzru⁷ in a temperature range of 793-953K and a pressure range of 0.1-2.93 MPa. It was found that the selectivities to several species depends on the pressure. At higher pressures, less hydrogen, methane, ethene, and propene are formed whereas more propane, butane, and 1-butene was measured. An overall activation energy for the decomposition of n-heptane was found to be between 209 and 219 kJ mol^{-1} . Garner et al.⁸ reported pyrolysis experiments of n-heptane using a high-pressure shock tube setup operated at pressures ranging from 2.5 to 5.0 MPa and temperatures ranging from 1000 to 1350K. Using these data, a kinetic model for pyrolysis of C7 hydrocarbons was built. Yuan et al.⁹ studied the pyrolysis of n-heptane diluted in argon in a tubular flow reactor at temperatures of 780 to 1780K. The reactor effluent was analyzed using synchrotron vacuum ultraviolet photoionization mass spectrometry, which enables the quantification of radical species in the reactor. Olahová et al.¹⁰ carried out steam cracking experiments of heptane at temperatures between 953 and 1033K and observed a conversion of n-heptane that can be modeled with an Arrhenius equation with a pre-exponential factor of $3.13 \cdot 10^{13} \text{ s}^{-1}$ and an activation energy of $249.1 \text{ kJ mol}^{-1}$. By using 24 secondary reactions, simulations were able to predict the conversion and the yields of ethene and hydrogen gas. Other response variables showed high deviations. Zámotný et al.¹¹ developed a model for n-heptane pyrolysis using automatic kinetic model generation. The kinetics were obtained using a group contribution approach, and were further refined by fitting them through experimental data. The simulated yields of the main products are in good agreement to the experimental data, but the yields of other products, e.g. hydrogen gas, ethane, and propane, were not well described by the model.

Many articles have already reported the chemistry or kinetic models for the pyrolysis of alkanes. Their thermal decomposition can be described by radical chain reactions, of which the typical reaction steps are listed in Table 1. The homolytic breaking of C-C (**R1**), and to a lesser of extent C-H, bonds, forms radical species. These radicals will abstract hydrogen atoms from the reactant (**R2**). The formed radicals will subsequently decompose via β -scission pathways (**R3**), forming smaller radicals and alkenes. This can be preceded by isomerization by H-shift reactions. The decomposition continues without much change in the total radical concentration. Finally, the recombination of radicals (reverse of **R1**) terminates the chain mechanism. This primary chemistry needs to be complemented by secondary chemistry, in which radical species can also add to olefins, again forming larger radicals (reverse of **R3**). Furthermore, large radicals can undergo cyclization through intramolecular radical addition steps. Subsequent hydrogen abstraction reactions and hydrogen β -scission reactions of these cyclic structures leads to the formation of aromatics. Besides the radical chemistry mechanism, other monomolecular reactions can be important for the aromatics formation.

Table 1: Typical reaction steps in a alkane thermal decomposition chain mechanism.

R1	Homolytic scission and reverse recombination	$\begin{array}{c} \text{R}_2 \\ \diagup \quad \diagdown \\ \text{R}_1 \end{array} \rightleftharpoons \cdot\text{R}_1 + \cdot\text{R}_2$
R2	Hydrogen abstraction	$\cdot\text{R}_1 + \text{H}-\begin{array}{c} \text{R}_2 \\ \diagup \quad \diagdown \\ \text{H} \end{array} \rightleftharpoons \text{H}-\begin{array}{c} \text{R}_1 \\ \diagup \quad \diagdown \\ \text{H} \end{array} + \cdot\text{R}_2$
R3	β -scission and reverse radical addition	$\begin{array}{c} \cdot \\ \diagup \quad \diagdown \\ \text{R}_1 \quad \text{R}_2 \quad \text{R}_3 \end{array} \rightleftharpoons \cdot\text{R}_1 + \text{R}_2=\text{R}_3$

Although the above described chemistry is well-known and accepted, the individual steps and their rate coefficients strongly depend on the reactant. The design or optimization of a process can be accelerated by accurate kinetic modelling listing all the significant reaction steps and their kinetic data. In the current society, largely driven by a fossil feedstock industry, the search for cleaner and more sustainable options is inevitable to maintain living standards. To obtain these requirements, chemical processes in general, and steam cracking and combustion processes in particular, need to be optimized wherever possible to maximize the energy efficiency and minimize pollutant emissions. An accurate description of the chemistry of even relatively simple molecules such as n-heptane can serve as starting point for these optimizations.

The pyrolysis studies described above are a first step in this direction, but a complete and accurate kinetic model for n-heptane thermal decomposition is to the best of the authors

knowledge not available. Moreover, the currently available kinetic models contain lumped reactions and species, which could raise some questions related to the true fundamental nature of the proposed models. Therefore, this chapter introduces the generation of a new n-heptane pyrolysis microkinetic model by merging an automatically generated model with well-established reaction steps taken from literature. The obtained kinetic model is first compared to experimental data without any tuning, i.e. to test its predictive capabilities. A rate of production analysis is then used to uncover the important pathways and assess their importance depending on the reactor conditions. The necessity of H-shift reactions is evaluated by performing simulations with and without this reaction family and looking at the differences in radical concentrations and in product yields.

5.3 Kinetic model generation

The generation of the microkinetic model has been done in two parts. First, reaction families of which a large number of reactions can be important for the pyrolysis of n-heptane have been included in an automatic kinetic model generation scheme. The kinetics of these reactions can be accurately calculated using group additivity approaches. Second, reactions that cannot be generalized in reaction families or whose kinetics are more difficult to calculate have been added manually to the model.

For the first part, the automatic kinetic model generator Genesys¹² has been employed to generate a large number of reactions which can be categorized in a limited number of reaction families. The network generation, i.e. the search for the reactions themselves, followed by the calculation of the necessary data is described in the following sections.

5.3.1 Network generation

Genesys can be used as rule-based kinetic model generator, which means that the termination of the kinetic model does not depend on the reaction rates, as opposed to rate-based kinetic model generators. In a rule-based scheme, a reaction is included to the model if it complies with a number of constraints defined by the end-user. These constraints can be defined for the whole kinetic model or on a per reaction family basis. Both the inclusion of molecular constraints, such as the molecular size, the number and size of ring structures, aromaticity, etc., as well as atomic constraints on the reactive center are possible. The latter include the immediate surrounding or

atoms that change in connectivity throughout the reactions. These constraints can include atomic properties such as valence, number of neighbors, hybridization, etc. and can also be defined by the role of the atom in the molecule, i.e. if the atom can or cannot be part of a ring or aromatic structure, what functional group can be part of its surroundings, etc. A good definition of the constraints is key to generate a model that can grasp the relevant chemistry without including too many unimportant reaction steps.

Genesys has been used to automatically find hydrogen abstractions and β -scission reactions during the decomposition of n-heptane. Although these are only two reaction families, they have been subdivided in several reaction families in Genesys for two reasons:

- (1) Constraints can depend on the nature of the reaction itself and can be different for reactions belonging to the same family, and
- (2) kinetics are calculated per reaction family in Genesys and separate calculation schemes are often needed to assure the necessary accuracy.

For β -scission reactions, the reverse radical addition reactions also need to be accounted for in the final model. Genesys considers all reactions reversible, and the kinetics of the reverse reaction step are calculated using thermodynamic consistency. Nevertheless, for technical reasons the reverse reaction families, or to be more specific the reaction recipe that needs to be executed, also need to be defined for a proper reaction network model generation. This can be explained as follows: because some reactions will first proceed through an addition, their products will not be present in the model before this reaction is generated by Genesys, and will thus not be found by only including the β -scission families. The resulting radical addition reaction families are not written to the final kinetic model, and should thus also be generated in the β -scission direction for their inclusion in the kinetic model. The reaction families used in Genesys and their subdivisions are summarized in Table 2.

Table 2: Reaction families and their subdivisions used in Genesys.

Reaction family	Subdivision 1	Subdivision 2	Example	Kinetics
Hydrogen abstractions	Intermolecular	By C- radical		Sabbe et al. ¹³
		By H- radical		Paraskevas et al. ¹⁴
	Intermolecular	1-2 H-shift		Chapter 4
		1-3 H-shift		Chapter 4
		1-4 H-shift		Chapter 4
		1-5 H-shift		Chapter 4
		1-6 H-shift		Chapter 4
		1-7 H-shift		Chapter 4
Radical additions and β -scissions	Intermolecular	Hydrogen-Centered β -Scission		In house data
		Hydrogen radical addition		Reverse from 2.1.1.
		Carbon-Centered β -Scission		Sabbe et al. ¹⁵
		Carbon radical addition		Reverse from 2.1.3.
	Intramolecular	C6 ring opening		In house data
		C6 cyclization		Reverse from 2.2.1.
		C5 ring opening		In house data
C5 cyclization		Reverse from 2.2.3		

Due to the radical addition reactions defined in the reaction families, the growth of molecules is allowed, which would lead to an indefinite model generation without any constraints. Furthermore, many of these addition, leading to large molecules, are insignificant under pyrolysis conditions. The constraints on additions were set such that all species contain 7 carbon atoms or less because species with more than 7 carbon atoms were not reported in the validation data. Note that, if all the isomers are considered, and an unlimited amount of double or triple bonds are allowed, the network becomes too elaborate. The maximum number of double and triple bonds in a molecule is limited to 3, and triple bonds are not allowed in C₄+ molecules because they have not been observed experimentally. Finally, to limit the number of isomers, only methyl branching is allowed, with a maximum of two branches per molecule.

Next to the reaction families, a species input needs to be given to Genesys prior to the network generation. The reactant, n-heptane, is inputted by its SMILES notation “CCCCCCC”. Since rate coefficients of initiation reactions cannot be calculated using group additivity or other simple models, no initiation reactions are present in the reaction families. Therefore, a radical species needs to be added to the reactants list to start the generation of reactions. The methyl radical “[CH3]” has been provided next to n-heptane.

The two initial species are put in the so-called “source” species, which is a set of species in Genesys containing all the species for which no reactions were searched yet. In each iteration step in Genesys, one species is taken from the source and added to the “reacted” species set. The latter contains all the species for which reactions have been searched. The species is checked against all the reaction families and all the possible monomolecular reactions are generated. Next, bimolecular reactions are searched for by using this species and each of the molecules in the reacted species. Newly generated species, that are not present in the source or reacted set, are added to the source. This procedure continues until the source is empty. A more elaborate description of the algorithm can be found elsewhere.¹²

5.3.2 Thermochemistry and reaction rate coefficients

Genesys contains two databases to obtain thermodynamic parameters of all the species in a kinetic model. Firstly, a database containing the thermochemistry of single species is present. After the network generation, all the species from the network are checked against the species in this database and if an exact hit is found between the target molecule and the database entry, the

accompanying database thermochemistry is used in the kinetic model, this database was used for 103 species in the kinetic model. Secondly, if a molecule is not found in the databases, its thermodynamic data is calculated based on group additivity, used for 283 species. An elaborate explanation of the concept and use of group additivity can be found in Chapter 1. In brief, each polyvalent atom in a molecule is assigned a group additive value, based on its nature and surroundings. These group additive values are tabulated in the databases of Genesys. The sum of all these group additive values gives an approximation of the property of the molecule. Besides the group additive values, a few other non-local contributions need to be accounted for such as non-nearest neighbor interactions or symmetry contributions. Entropy, standard enthalpy of formation and heat capacity can be calculated using this scheme. All the group additive values used in this work originate from CBS-QB3 *ab initio* calculations.^{16,17}

Kinetic data can, analogously to thermodynamics, be pulled from two separate databases: One containing single rate coefficients per reaction and one containing group additive values. The former database has not been used in this work, all the kinetic data is calculated from group additivity. Details on the group additivity scheme for kinetic data can be found in literature^{13-15,18-21} as well as in Chapter 4 of this work. In short, the pre-exponential factor and activation energy of a reaction are calculated by taking the pre-exponential factor and activation energy of a reference reaction belonging to the same reaction family, and adding contributions for the ligands of the reactive atoms. This is shown in Eq. 5.1 and Eq. 5.2 in which C_i are the reactive atoms. For the pre-exponential factor, additional contributions for the number of single events n_e and for tunneling corrections κ are required. The ΔGAV^o 's have all been obtained using *ab initio* rate coefficients calculated at the CBS-QB3 level of theory.

$$E_a(T) = E_{a,ref}(T) + \sum_i \Delta GAV_{E_a}^o(C_i) \quad \text{Eq. 5.1}$$

$$\log A(T) = \log \tilde{A}(T)_{ref} + \sum_i \Delta GAV_{\log \tilde{A}}^o(C_i) + \log \kappa \cdot n_e \quad \text{Eq. 5.2}$$

The above description of automatic reaction generation and kinetic parameter calculation are very powerful for their wide applicability, scalability, accuracy, and speed. However, several reaction steps cannot be described by these procedures and need a separate treatment. First, several reactions only happen once or a limited number of times and are difficult to be generalized in a reaction family. Among these reactions, the formation of the aromatics benzene, toluene and

naphthalene are of particular interest for this work. Second, for several reactions, although they can be described by general reaction recipes, the kinetics cannot be calculated by group additivity or other simple calculation procedures. Homolytic scission reactions and their reverse radical recombination reactions belong to this group of reactions. Calculations schemes typically result in large uncertainties, and most of the process responses are very sensitive to their rate coefficients. Therefore, the values are taken from literature when available. To account for the chemistry of small molecules, the model generated by Genesys was merged with a literature model for hydrocarbons and oxygenates, of which only the hydrocarbon species and reactions were kept. This literature model is the AramcoMech2.0 model²², of which the hydrocarbon part contains 114 species 498 reactions. The formation of aromatic species was also implemented in the kinetic model through literature reactions. A model generated by RMG^{23,24}, including 743 reactions between 360 species, was merged to the kinetic model. Finally, the 3 initial homolytic C-C scissions of n-heptane were taken from Mohamed et al.²⁵ The specifications on the different parts of the model are given in Table 3. The number of species and reactions from the complete model is less than the sum of the different parts because several species and reactions are present in two or more models. Model merging is done semi-automatically using a master/slave approach as is explained in section 2.3.5. The complete model is archived at the LCT and is available on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1\.

Table 3: Specifications on the automatic part of the kinetic model (Genesys) and on the parts that were added from literature.

Model	Number of species	Number of reactions
Genesys	387	3398
AramcoMech2.0 ²²	114	498
RMG ^{23,24}	360	743
Initiations ²⁵	9	3
Complete model	691	4221
Reduced model	154	1043

5.3.3 Model reduction

The initial kinetic model contains 4221 reactions among 691 species. Many of the species and elementary steps are however insignificant for the conversion, the intermediates formation and the products formation during pyrolysis of n-heptane. Therefore, the model was reduced using the Chemkin Workbench²⁶ which implements the “Directed Relation Graph incorporating Error

Propagation²⁷ (DRGEP) reduction method. This method is a skeletal model reduction method, which removes redundant species and the reactions they participate in. The chemical scheme and kinetic structure of the important species is not altered by the reduction, in contrast to lumping techniques. The DRGEP method²⁷ is graph-based, meaning that the reduction is based on identifying groups of species that may be internally coupled, but are not strongly coupled to selected important species. The coupling between two species is defined by the influence the removal of the first species has on the error to the production of the second species. Without error propagation, the importance of a species A for the production or consumption of species B depends on the reaction with the lowest relative rate of production along the pathway from A to B. With error propagation, errors are damped as they propagate along the network, and the importance of species A depends on the product of all the relative rates of production along the pathway from A to B. Since the relative rates of production lie between 0 and 1, the final importance of species A is lower compared to the DRG method without error propagation, and the difference between the two methods increases with an increasing number of reaction steps between A and B. In the case of several pathways from A to B, the pathway resulting in the highest importance of A is used.

An important note to make is that this method is a local method, i.e. it depends on rates of production and thus on the conditions at which these are calculated. To allow the final model to maintain its accuracy under divergent conditions, two sets of conditions were used, taken from Yuan et al.⁹ and Zámostný et al.¹¹. The resulting model contains 154 species reacting in 1043 elementary steps. It can be found in the archive S:\vakgroep\ea12archieff\SoftLib\r\ruvdv\1\.

5.4 Results

5.4.1 Model validation

As mentioned in the introduction, Yuan et al.⁹ performed n-heptane pyrolysis experiments in a tubular reactor. The reactor configuration and process conditions are summarized in Table 4. Heptane is highly diluted in an argon stream, the n-heptane inlet mole fraction is 0.02, and low pressures are applied.

Table 4: Reactor configuration and conditions for the n-heptane pyrolysis experiments of Yuan et al.⁹.

Reactor length (m)	0.3
Reactor length in furnace (m)	0.1
Diameter (m)	$6 \cdot 10^{-3}$
Temperature (K)	780-1780
Total pressure, absolute (Pa)	400
Inlet Argon mass flow rate (g s^{-1})	$1.5 \cdot 10^{-2}$
Inlet n-heptane mass flow rate (g s^{-1})	$7.5 \cdot 10^{-4}$

Five kinetic models have been used to simulate the reactor outcome: The model of Genesys, and four literature models for combustion and pyrolysis of n-heptane, i.e. the models of Herbinet et al.²⁸, Zhang et al.²⁹, Cai et al.³⁰, and Garner et al.⁸. The four literature models and the model generated in this work are compared in Table 5 in terms of size and the generation method. All the models were automatically generated, and three of the models only contain elementary steps, whereas two models contain lumped reactions and species.

Table 5: Comparison of the five kinetic models used for the simulations.

Model	Number of species	Number of reactions	Automatic generation	Elementary reactions
Genesys	154	1043	yes	yes
Herbinet et al. ²⁸	1268	5336	yes	no, lumped
Zhang et al. ²⁹	303	1990	yes	no, lumped
Cai et al. ³⁰	1692	5804	yes	yes
Garner et al. ⁸	135	717	yes	yes

The Chemkin²⁶ software has been used to perform the reactor simulations. The reactor has been modelled as an ideal plug flow reactor. The reactor has a small diameter and the gas velocities are high, above 100 m s^{-1} . Furthermore, the Péclet numbers amount between 615 and 1093 for mass, c.f. Eq. 5.3, and between 542 and 981 for heat transport, c.f. Eq. 5.4. Hence, although the Reynolds numbers, calculated by Eq. 5.5, are low (between 688 and 1195), the use of a plug flow reactor model is justified. In the equations Eq. 5.3 to Eq. 5.5, L is the length of the reactor, v is the gas velocity, D is the binary diffusion coefficient, k_h is the thermal conductivity, MM is the molar mass, ρ is the density, C_p is the heat capacity, and μ is the dynamic viscosity. These equations have been evaluated using the inlet conditions.

$$Pé_{mass} = \frac{Lv}{D} \quad \text{Eq. 5.3}$$

$$Pe_{heat} = \frac{Lv}{k_h MM / \rho C_p} \quad \text{Eq. 5.4}$$

$$Re = \frac{\rho Lv}{\mu} \quad \text{Eq. 5.5}$$

The continuity equation to calculate concentration of n-heptane using a plug flow reactor model is given in Eq. 5.6 and Eq. 5.7, in which F_{A_0} is the inlet flow rate of n-heptane, X_A is the conversion of n-heptane, $R_{v,A}$ is volumetric production rate and V is the reactor volume. The continuity equation is integrated as shown in Eq. 5.8. From this equation, the residence time t can be calculated according to Eq. 5.9 in which R is the ideal gas constant, T is the temperature and p_{A_0} is the inlet partial pressure of n-heptane. The residence time is thus calculated based on inlet conditions.

$$F_{A_0} dX_A = -R_{v,A} dV \quad \text{Eq. 5.6}$$

$$\frac{dX_A}{d(V/F_{A_0})} = -R_{v,A} \quad \text{Eq. 5.7}$$

$$\frac{V}{F_{A_0}} = - \int_0^{X_A} \frac{dX}{R_{v,A}} \quad \text{Eq. 5.8}$$

$$t = \frac{V}{F_{A_0} RT / p_{A_0}} \quad \text{Eq. 5.9}$$

The n-heptane mole fraction measured by Yuan et al.⁹ is shown Figure 1 together with the results from the five simulations. Using this reactor configuration, the residence time amounts between $2 \cdot 10^{-4}$ and $4 \cdot 10^{-4}$ s, and n-heptane only starts to decompose at temperatures above 1100K. This is well predicted by four of the five kinetic models, among which the one developed in this work. All kinetic models show a too high decrease of the mole fraction of n-heptane as a function of temperature. The new model shows the lowest decrease of the n-heptane mole fraction and is thus slightly closer to the experimental measurements. Above 1600K, almost all the n-heptane is decomposed.

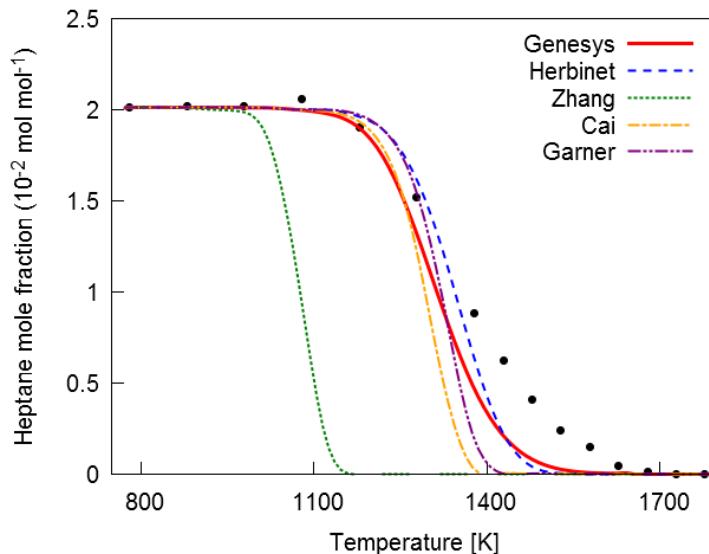


Figure 1: Simulated n-heptane conversion compared to the experimental data of Yuan et al.⁹ using the kinetic models of Genesys (red, full line), Herbinet et al.²⁸ (blue, dashed line), Zhang et al.²⁹ (green, dotted line), Cai et al.³⁰ (orange, dashed – dotted line), and Garner et al.⁸ (purple, dashed – double dotted line). The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1.

The formation of H₂, methane, ethyne, ethene, ethane, propene, propyne, and allene as a function of temperature is depicted in Figure 2. All the simulations exhibit a too low hydrogen gas formation, with the model of Genesys being the closest to the experimental data, this without any fitting of model parameters to the experiments. For methane, the new model is capable of reproducing the experimental yields, together with the model of Cai et al.³⁰. The other models not only show too low methane yields, the trend of the methane yield as a function of temperature is also not well predicted. Ethyne, which is mainly formed at high temperatures has a monotonously increasing yield as function of the temperature, which is well predicted by all kinetic models. The model of Genesys slightly overpredicts its formation, while the other models have a slightly too low ethyne prediction. One of the main products of n-heptane pyrolysis is ethene, of which the mole fraction amounts twice the n-heptane inlet mole fraction at high conversions, i.e. close to 4 atoms of the 7 carbon atoms in n-heptane end up in ethene. The trends of ethene yields are well predicted by the four of the five kinetic models, with the Genesys model again being the closest to the experimental data.

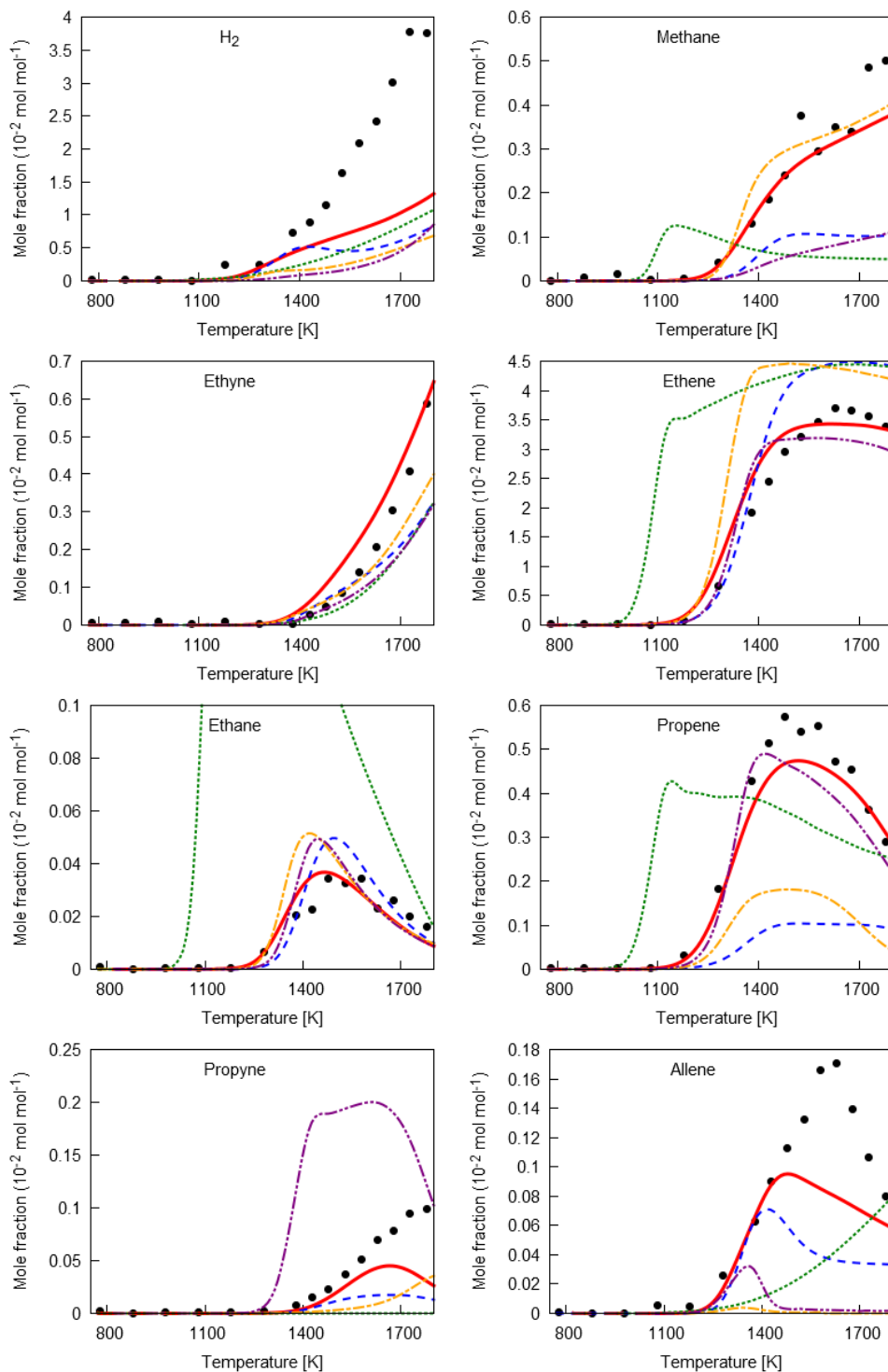


Figure 2: Simulated H₂, methane, ethyne, ethene, ethane, propene, propyne, and allene product yields compared to the experimental data of Yuan et al.⁹ using the kinetic models of Genesys (red, full line), Herbinet et al.²⁸ (blue, dashed line), Zhang et al.²⁹ (green, dotted line), Cai et al.³⁰ (orange, dashed – dotted line), and Garner et al.⁸ (purple, dashed – double dotted line). The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1.

The mole fraction of ethane is largely overpredicted by the model of Zhang et al.²⁹. The other four models are closer to the experimental data. The peak of the ethane yield as a function of temperature shown by the experiments is also visible in all the simulations. Three important C3 products are propyne, allene and propene. The models Cai et al.³⁰, Garner et al.⁸ and Genesys are able to grasp the trend of the temperature dependence of propene. The two latter also predict the mole fractions themselves well. The trends of propyne and allene are not well predicted by Genesys, and for both, a too low mole fractions is simulated. These compounds are important intermediates towards the propargyl radical, which can further react to benzene and other cyclic species.

The product yields of 1-buten-3-yne (C_4H_4), 1,3-butadiene, 1-butene, 1,4-pentadiene (C_5H_8), 1-pentene, benzene, and 1-hexene are given in Figure 3. The models of Genesys and Garner et al.⁸ show a good C_4H_4 mole fractions while the other three models largely underpredict the formation of this product. The kinetic model of Genesys is also capable of predicting the C_4H_6 mole fraction until 1500K, above which the experimental data reports higher C_4H_6 yields. Three 1-alkenes, C_4H_8 , C_5H_{10} , and C_6H_{12} , show a similar trend of their mole fraction: A narrow peak of their yield as a function of temperature is observed around 1400K, after which their mole fractions rapidly decrease. These trends are well captured by the new kinetic model, the other three models show higher deviations. The simulated trends of the C_5H_8 yields show higher deviations compared to the experimental data. Finally, the trend of the benzene formation can be described by the Genesys model, with an underprediction of the yields themselves.

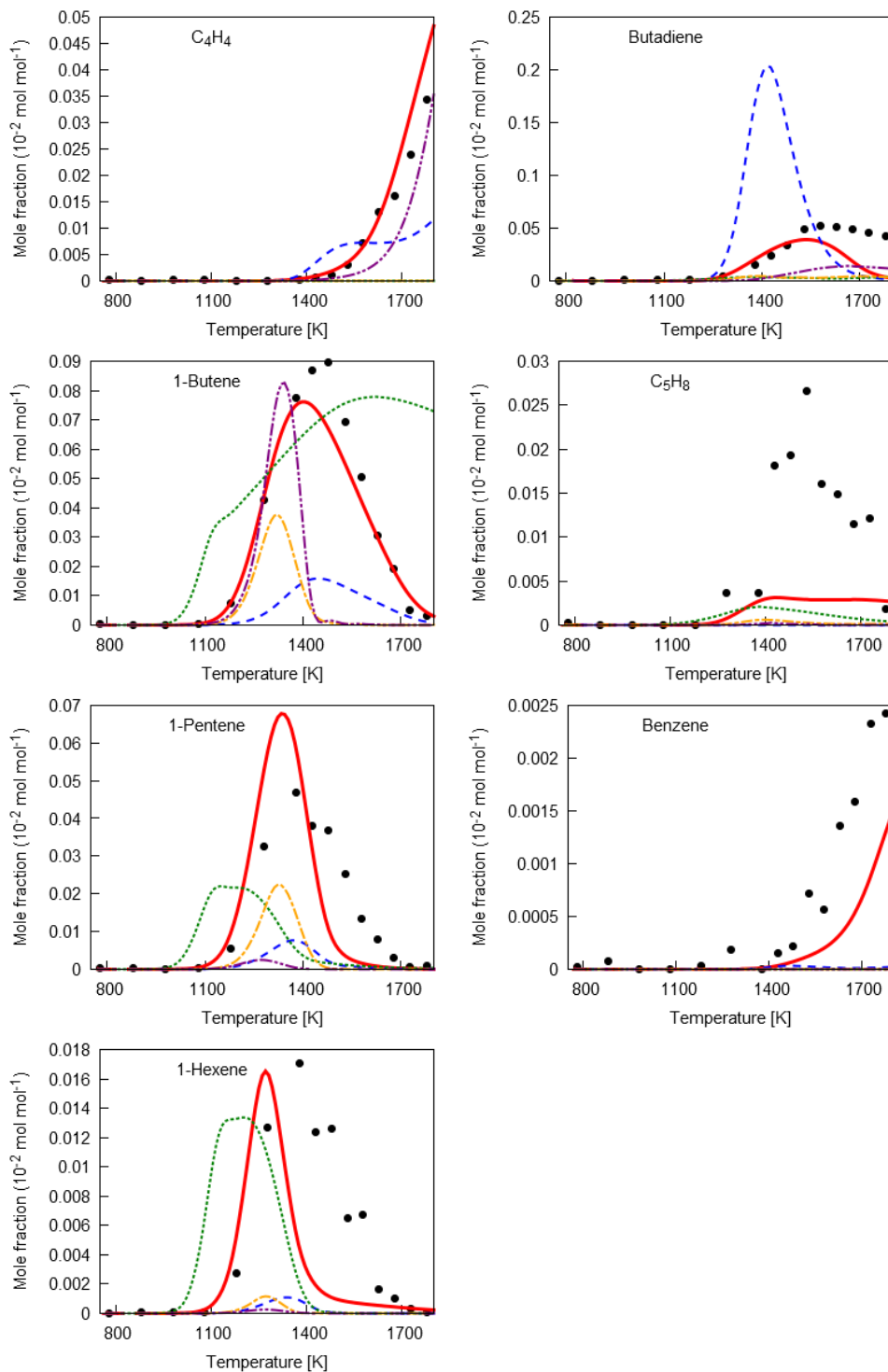


Figure 3: Simulated 1-buten-3-yne (C_4H_4), 1,3-butadiene, 1-butene, 1,4-pentadiene (C_5H_8), 1-pentene, benzene, and 1-hexene product yields compared to the experimental data of Yuan et al.⁹ using the kinetic models of Genesys (red, full line), Herbinet et al.²⁸ (blue, dashed line), Zhang et al.²⁹ (green, dotted line), Cai et al.³⁰ (orange, dashed – dotted line), and Garner et al.⁸ (purple, dashed – double dotted line). The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1\.

Four radicals have been measured by Yuan et al.⁹: methyl, ethyl, propargyl and allyl, of which their experimental and simulated mole fractions are shown in Figure 4. The experimental methyl and ethyl trends are also seen in the simulations, but the absolute values show higher deviations. The kinetic model of Genesys is the closest to the experimental data. The increase of the propargyl radical yield at high temperatures is well predicted by the model of Garner et al.⁸, the model of Genesys shows an underprediction. Finally, the allyl radical is not well predicted by any of the kinetic models.

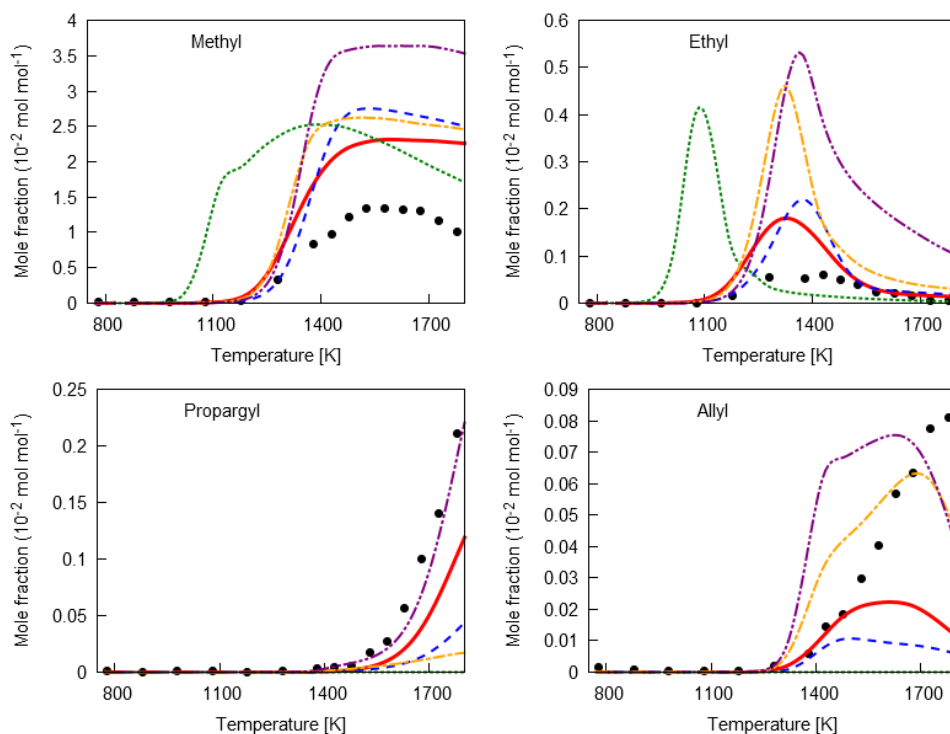


Figure 4: Simulated methyl, ethyl, propargyl, and allyl radical yields compared to the experimental data of Yuan et al.⁹ using the kinetic models of Genesys (red, full line), Herbinet et al.²⁸ (blue, dashed line), Zhang et al.²⁹ (green, dotted line), Cai et al.³⁰ (orange, dashed – dotted line), and Garner et al.⁸ (purple, dashed – double dotted line). The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1\.

Besides the high temperature and low pressure experiments of Yuan et al.⁹, the new kinetic model has been compared to the experimental data of Zámostný et al.¹¹, who reported n-heptane steam cracking experiments of n-heptane in a tubular reactor at lower temperatures and higher pressure. The reactor configuration and experimental conditions are shown in Table 6.

Table 6: Reactor configuration and conditions for the n-heptane steam cracking experiments of Zámostný et al.¹¹.

Reactor length (m)	0.75
Reactor length in furnace (m)	0.3
Inner diameter (m)	$6 \cdot 10^{-3}$
Outer diameter (m)	$12 \cdot 10^{-3}$
Temperature (K)	953-1033
Absolute pressure (kPa)	101
Inlet n-heptane mass flow rate (g s^{-1})	$4.8\text{-}9.7 \cdot 10^{-3}$
Inlet water mass flow rate (g s^{-1})	$1.3\text{-}2.9 \cdot 10^{-2}$
Steam dilution ($\text{kg}_{\text{water}}/\text{kg}_{\text{n-heptane}}$)	3

Two kinetic models have been used to simulate the reactor outcome: The model of Genesys, and the model of Garner et al.⁸. The Chemkin²⁶ software has been used to perform the reactor simulations. The reactor has been modelled as an ideal plug flow reactor. Similarly to the experiments of Yuan et al.⁹, the reactor has a small diameter and high gas velocities. Furthermore, the Péclet numbers range from 294 to 734 for heat and 2310 to 5737 for mass transport. The Reynolds numbers amount between 1592 and 40000. These dimensionless numbers have been calculating using equations Eq. 5.3 to Eq. 5.5. Again, the ideal plug flow approximation of the reactor is justified. Zámostný et al.¹¹ also reported simulated product yields which can be used for the comparison of the kinetic model performances. The kinetic parameters in the model of Zámostný et al.¹¹ originate from fitting the model performance through six experiments and the other three experiments were used for validation.

The experimental and modeled conversions are given in Figure 5. The model of Zámostný et al.¹¹ is able to predict the conversion of n-heptane. For the model of Genesys, the conversion was initially overpredicted. However, it was seen that a considerable amount of 1-, 2-, and 3-heptene are modeled by Genesys, i.e. up to 5wt%. The yields of these species are not measured by Zámostný et al.¹¹. Very likely they could not be measured with the used analytical equipment, but the mass balances of the experimental data close. Therefore, the heptenes have been added to

unreacted heptane to calculate the conversion, which brings the experimental and modeled conversion close to each other. The model of Garner et al.⁸ largely underpredicts the conversion.

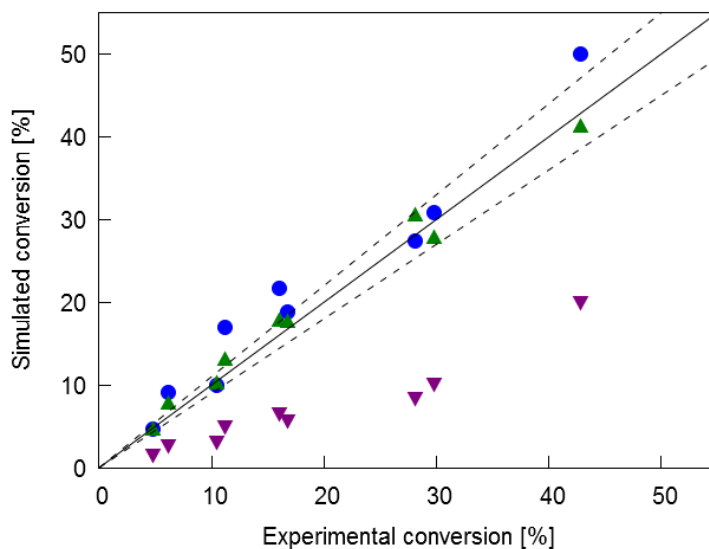


Figure 5: Simulated n-heptane conversion compared to the experimental data of Zámotný et al.¹¹ using the kinetic models of Genesys (blue circles), Zámotný et al.¹¹ (green upwards-pointing triangles), and Garner et al.⁸ (purple downwards-pointing triangles). The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdv\jv\1\.

The yields of the main products as a function of the conversion are given in Figure 6. Hydrogen gas is overpredicted by all three models, the simulations show a linear trend of the hydrogen yield as a function of the n-heptane conversion, whereas the experimental data exhibits an increase of the hydrogen yield that is less than linear. The methane yield, showing a linear increase as a function of the conversion, is very well predicted by all three models. For ethene and propene, the model of Genesys slightly underpredicts their yields, while the model of Zámotný et al.¹¹ is closer to the experimental data. The model of Garner et al.⁸ shows a good ethene yield but also underpredicts the propene yield. 1-butene is well described by the model of Genesys and the model of Zámotný et al.¹¹ while the model of Garner et al.⁸ overpredicts the yields. Figure 6 also shows the yield of C5 and C6 species, which is overpredicted by the new model while the other two models are close to the experimental data. The overprediction is mainly attributed to a high 1-pentene yield.

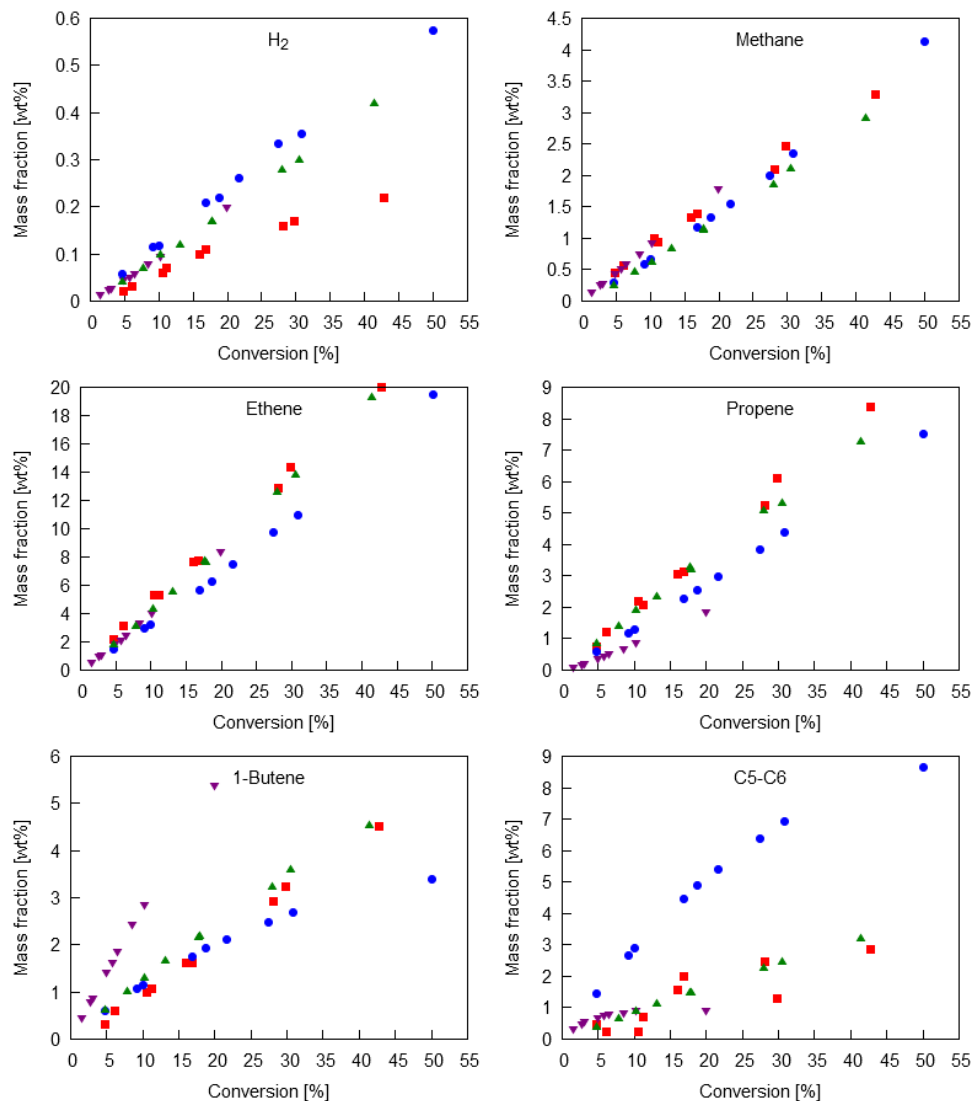


Figure 6: Simulated H₂, methane, ethene, propene, 1-butene and C5-C6 product yields compared to the experimental data of Zámotný et al.¹¹ (red squares) using the kinetic models of Genesys (blue circles), Zámotný et al.¹¹ (green upwards-pointing triangles), and Garner et al.⁸ (purple downwards-pointing triangles). The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1.

Figure 7 shows the experimental and simulated yields of minor products. Five products, i.e. ethane, propane, ethyne, propyne, and n-butane, are not calculated by the model of Zámotný et al.¹¹. Ethane is well predicted by the model of Genesys, and propane is slightly overpredicted. The kinetic model of Garner et al.⁸ underpredicts the ethane yield and shows a good agreement for the propane yield. Ethyne and propyne are largely overpredicted by the model of Garner et al.⁸ whereas the kinetic model of Genesys shows a much better agreement.

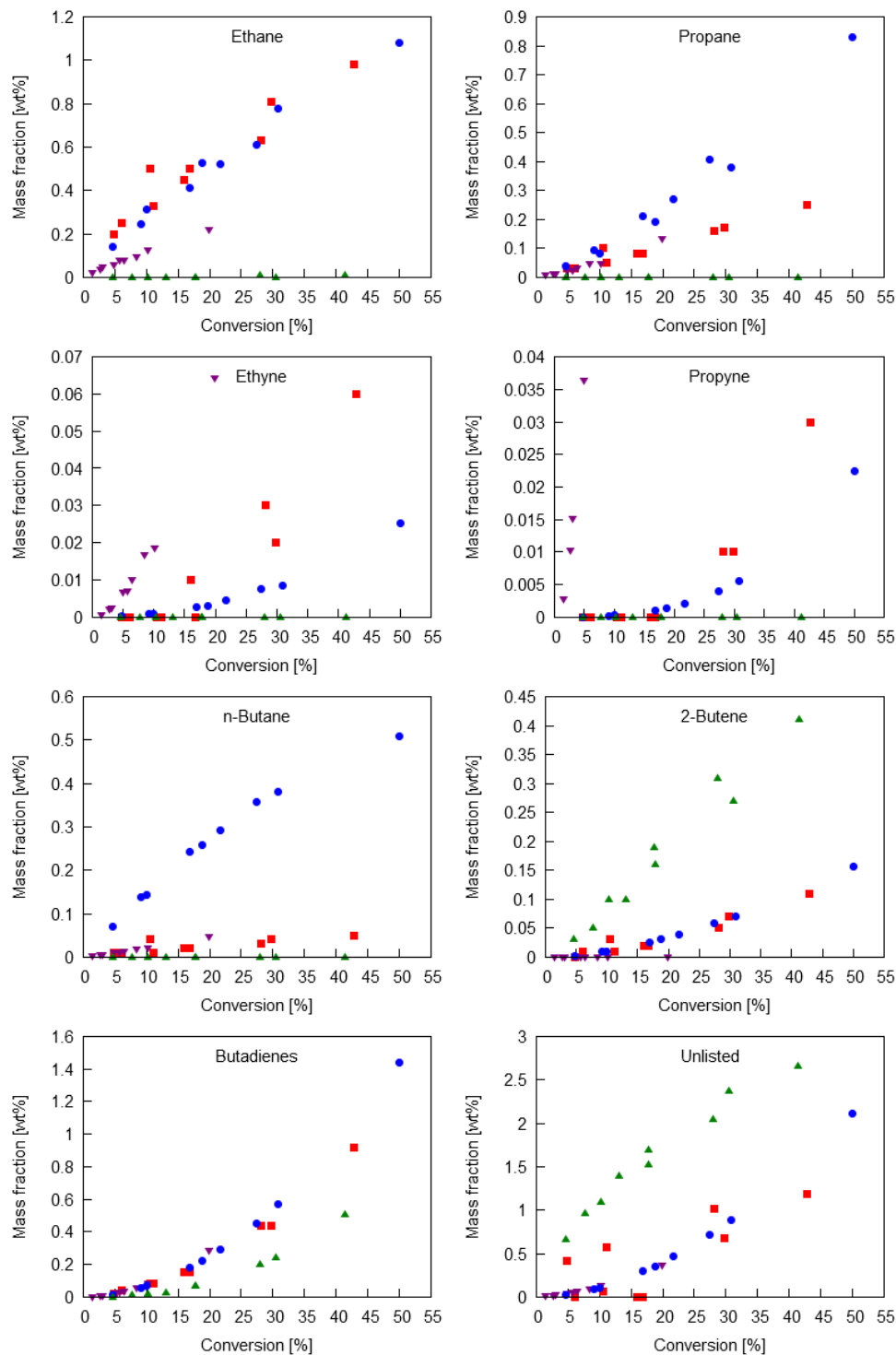


Figure 7: Simulated ethane, propane, ethyne, propyne, n-butane, 2-butene, butadiene and unlisted product yields compared to the experimental data of Zámstný et al.¹¹ (red squares) using the kinetic models of Genesys (blue circles), Zámstný et al.¹¹ (green upwards-pointing triangles), and Garner et al.⁸ (purple downwards-pointing triangles). The unlisted product yields are obtained by closing the mass balance compared to all the listed product yields. The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\lea12archief\SoftLib\r\ruvdv\1.

The model of Genesys calculates too much n-butane in contrast to the model of Garner et al.⁸. 2-butene is well predicted by the model of Genesys. The other two models show higher deviations. Butadiene is well predicted by the model of Genesys and the model of Garner et al.⁸. The model of Zámostný et al.¹¹ exhibits a small underprediction.

The final graph of Figure 7 shows the amount of remaining species, which was calculated by adding all the mole fractions up of all the species that were not reported in the experimental study. The models of Genesys and Garner et al.⁸ have a similar amount of remaining species compared to the experimental data. The model of Zámostný et al.¹¹ predicts too many other species.

In conclusion, the model of Genesys has been tested in a wide range of temperatures, i.e. 680-1800K, both at low (400 Pa) and high ($1 \cdot 10^5$ Pa) absolute pressures. A wide variety of product yields have been compared to experimental data, and considering the fundamental nature of the model clear improvements can be seen.

5.4.2 Rate of production analysis

To gain insights in the main decomposition pathways of n-heptane, a rate of production analysis has been done based on the reactor configurations and conditions of the experiments of Yuan et al.⁹ and Zámostný et al.¹¹. The reaction scheme showing the initial reactions during n-heptane pyrolysis are depicted in Figure 8. The decomposition of n-heptane is characterized by a free radical chain mechanism. The origin of the radicals are the three homolytic C-C scissions in n-heptane leading to methyl and 1-hexyl, ethyl and 1-pentyl, and n-propyl and 2-butyl radicals respectively. As expected, homolytic C-H scissions have been found to be unimportant. Another important family for the initial decomposition of n-heptane are hydrogen abstractions yielding the four n-heptyl radicals. The formed radicals further decompose via mainly C-H- β -scission reactions and C-C- β -scission reactions. This results in a wide variety of 1-alkene species such as ethene, propene, 1-butene, 1-pentene, and 1-hexene and in the three heptene species. The radicals with a carbon chain above four carbon atoms also isomerize via H-shift reactions to form secondary radicals. These can in turn also undergo β -scission reactions forming other olefin species. The relative importance of the competing reaction pathways strongly depends on the reactor conditions.

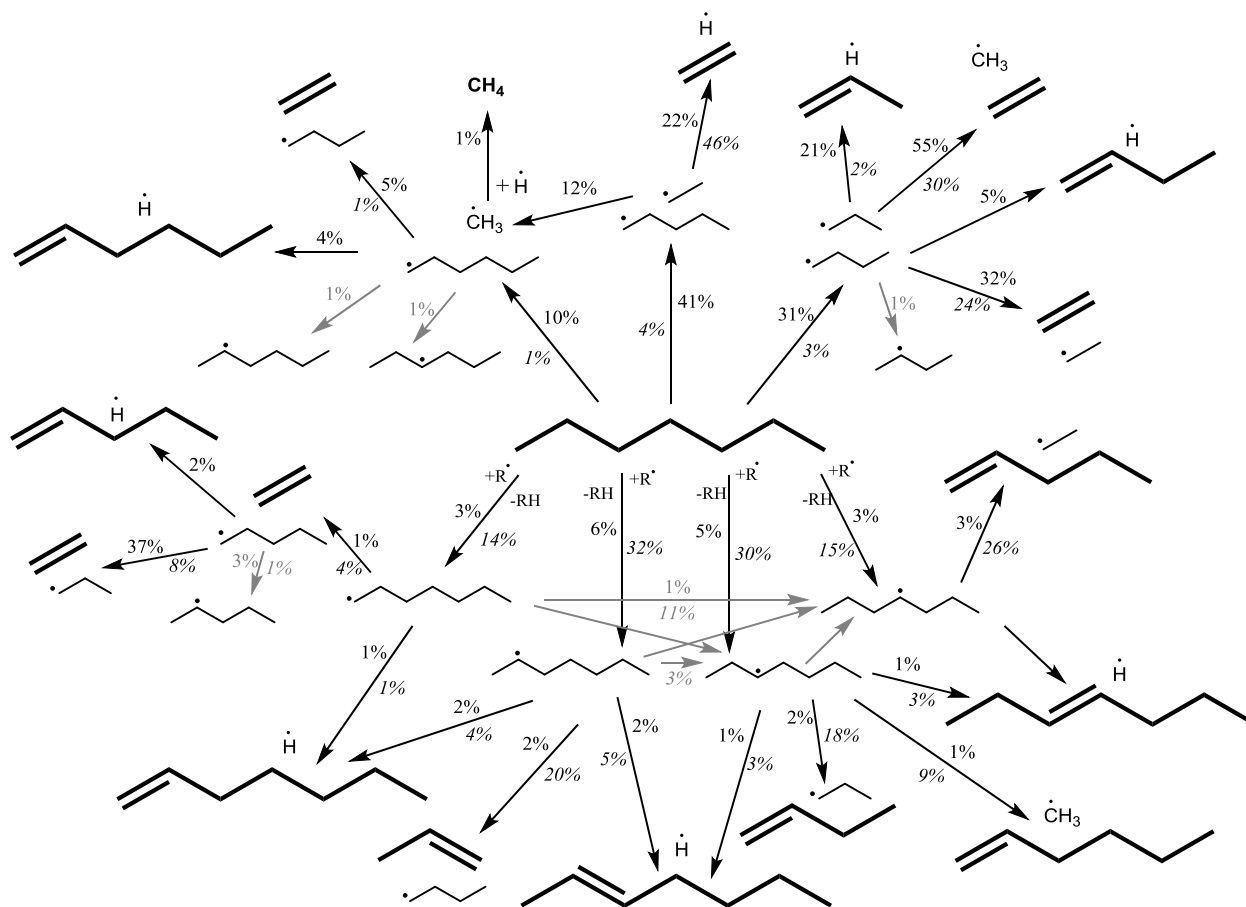


Figure 8: Reaction scheme of the decomposition of n-heptane. n-Heptane and the products are depicted in bold. The percentages correspond to the rate of production of the reaction relative to the total rate of consumption of n-heptane, percentages are only added if they are 1% or higher. The grey arrows and percentages correspond to H-shift reactions. The non-italic percentages correspond to the experimental conditions of Yuan et al.⁹ at 1473K and a conversion of 20%. The italic percentages correspond to the experimental conditions of Zámotný et al.¹¹ at the highest temperatures, i.e. with a peak temperature of 1033K, and a conversion of 20%. The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archiefl(SoftLib\r\ruvdvijv\1).

In the following sections, the rate of production of several important species during the decomposition of n-heptane are elaborated.

When simulating the experimental data of Yuan et al.⁹, which take place at low pressure, high temperatures and low residence time, the main n-heptane decomposition pathways are the homolytic C-C scission reactions. Figure 9 shows the rate of consumption of n-heptane at four different temperatures: 1223K, 1473K and 1773K, with conversions of 12.8%, 95.0% and 99.9999% respectively. The graph on the left shows the total rate of consumption of n-heptane at each temperature. On the right, the relative importance of three reaction families are depicted, i.e. the homolytic C-C scission reactions, hydrogen abstractions by hydrogen radicals and hydrogen abstractions by methyl radicals. Other reaction families account for less than 1% of the

conversion of n-heptane. The homolytic C-C scission reactions are important at each temperature, and the importance of hydrogen abstraction increases with decreasing temperature, of which abstraction reactions by hydrogen radicals are the most important ones. At the highest temperature, n-heptane decomposes early in the reactor and its rate of consumption thus rapidly decreases as a function of the distance in the reactor. It is important to note that these reactions are in the falloff region under the process conditions of Yuan et al.⁹.

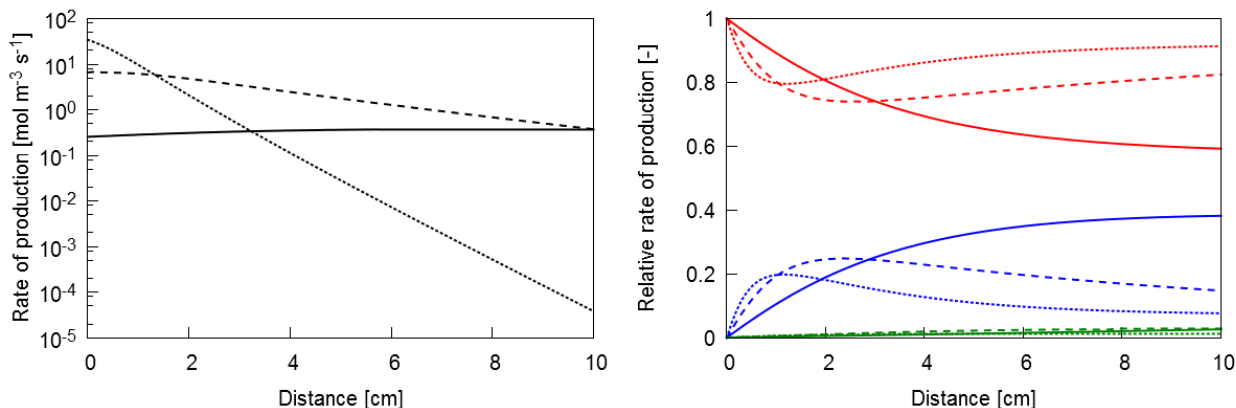


Figure 9: Total rate of consumption of n-heptane (left) and relative rate of consumption of n-heptane (right) of homolytic C-C scission reactions (red), hydrogen abstractions by hydrogen radicals (blue) and hydrogen abstractions by methyl radicals (green) using the experimental conditions of Yuan et al.⁹ at 1223K with a conversion of 0 to 12.8% (full lines), 1473K with a conversion of 0 to 95% (dashed lines) and 1773 with a conversion of 0 to 99.9999% (dotted lines). The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archief\SoftLib\r\ruvdv\jv\1\.

When simulating the reactor of Zámostný et al.¹¹, of which the pyrolysis takes place at higher pressures, lower temperatures and higher residence times (0.1 to 0.4 s according to Eq. 5.9), the relative importance of homolytic C-C scissions is much less, as can be seen in Figure 10. Hydrogen abstractions lead to a much higher conversion rate of n-heptane, with abstractions by hydrogen radicals being mainly important early in the reactor and abstractions by methyl radicals increasing in importance halfway through the reactor. This is mainly due to the concentration differences of hydrogen and methyl radicals. Hydrogen radicals are consumed much faster compared to methyl radicals. For the experimental conditions of Zámostný et al.¹¹, the initiation kinetics are closer to the high-pressure limits.

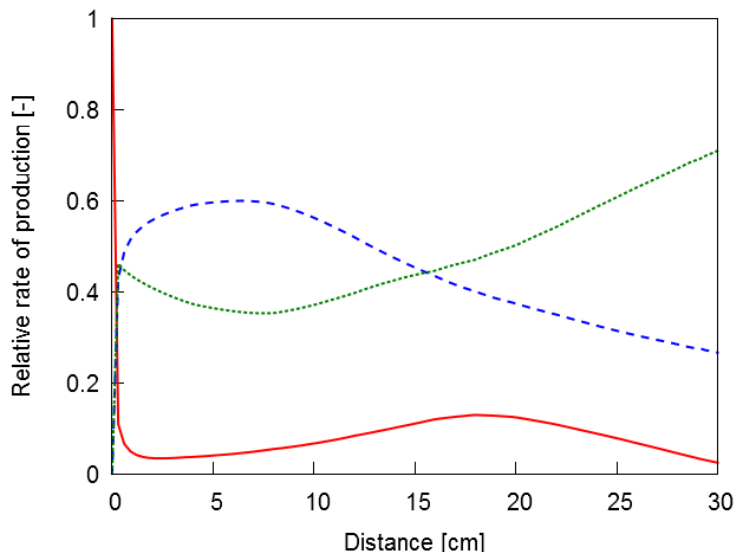


Figure 10: Relative rate of consumption of n-heptane of homolytic C-C scission reactions (red full lines), hydrogen abstractions by hydrogen radicals (blue dashed lines) and hydrogen abstractions by methyl radicals (green dotted lines) using the experimental conditions of Zámostný et al.¹¹ at 1033K with a conversion of 0 to 54.3%. The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1.

The formation of methane is mainly driven by hydrogen abstractions by methyl radicals from n-heptane and by the radical recombination of a methyl and a hydrogen radical. The reaction of methyl with an ethyl radical yielding ethene and methane is also important. Methyl radicals are created through C-C- β -scissions in various larger radicals. The reaction of a hydrogen radical with an ethyl radical giving two methyl radicals also contributes significantly, although it turns in a consumption route of methyl at high temperatures and high conversion. Further consumption pathways of methyl radicals are, besides hydrogen abstractions to methane, radical recombinations to methane, ethane and propane.

The pathways to hydrogen gas all include hydrogen abstractions by hydrogen radicals. Abstractions from n-heptane have the highest rate of production, followed by abstractions from olefins yielding resonantly stabilized radicals and abstractions from small hydrocarbons such as ethane, ethene or methane. The hydrogen radicals themselves are formed through C-H- β -scission reactions.

Ethene is almost exclusively formed by β -scission reactions of primary radicals. The C-H- β -scission in the ethyl radical is the most important route to ethene, followed by the C-C- β -scissions in 1-propyl, 1-butyl and 1-pentyl radicals. At low conversion and low temperature,

using the experimental condition of Yuan et al.⁹, the C-C- β -scissions in 1-propyl radical dominates the C-H- β -scission in the ethyl radical, but its importance decreases along the reactor, whereas the rate of production of ethene via the latter increases. At higher temperatures, the C-H- β -scission in the ethyl radical are more important, which is also the case when using the experiments of Zámotný et al.¹¹.

Using the experimental conditions of Zámotný et al.¹¹, propene is mainly formed by a C-C- β -scission of the 2-heptyl radical. Early in the reactor, the C-H- β -scission of the propyl radical also contributes to the formation of propene. However, further in the reactor, the reverse reaction becomes a consumption pathway of propene to the 1-propyl radical. Other reactions that contribute to the formation of propene are the C-C- β -scissions of the 2-butyl, 2-pentyl and 2-hexyl reactions. Also, the reaction of 1-butene with a hydrogen radical yielding propene and a methyl radical is important. Using the experimental conditions of Yuan et al.⁹, the same reactions are important, but their relative importance changes depending on the temperature.

Ethyne formation proceeds through C-C- β -scission and C-H- β -scission reactions of vinylic radicals. The vinyl radical is an important source of ethyne, but the 1-propen-1-yl, 1-buten-1-yl, and 1-penten-1-yl radicals also contribute to the formation of ethyne. The reaction of a hydrogen radical or methyl radical with a vinyl radical yielding ethyne and hydrogen gas or methane is also important.

Larger product species are mostly formed by consecutive β -scissions and hydrogen abstractions starting from the radicals formed from n-heptane, i.e. 1-butyl, 1-pentyl, 1-hexyl and the four heptyl radicals. These are precursors to, among others, heptenes, 1-hexene, 1-pentene, butenes, and butadiene.

Although benzene is not formed in high amounts, its rate of production shows several interesting pathways, which are depicted in Figure 11. At a temperature of 1473K, the main pathway to form benzene is the H₂ elimination of 1,4-cyclohexadiene. The latter is formed by C-H- β -scission of the 4-cyclohexenyl radical. This species is created after intramolecular radical addition reaction of the 1,4-hexadien-6-yl radical. At higher temperatures, the importance of benzene formation by a C-H- β -scission of the 1,3-cyclohexadien-5-yl radical increases. This radical is formed from 1,3-cyclohexadiene, either by hydrogen abstractions or by a homolytic C-H scission. At high

temperatures and higher conversions, the formation of benzene via fulvene and via propargyl radicals also gains in importance. The formation of these species is also shown on Figure 11.

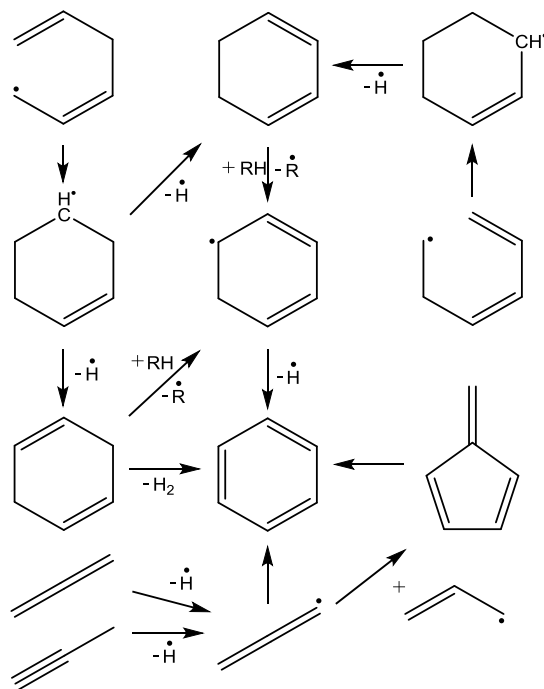


Figure 11: Main reaction pathways towards benzene. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1\.

5.4.3 Role of H-shift reactions

The previous chapter introduced a set of group additivity values to calculate the rate coefficients of H-shift reactions. This reaction family has been added to the Genesys input, and a large number of H-shift reactions have been included in the kinetic model, with their rate coefficients being calculated from group additivity. It can be assumed that these reactions will influence the relative concentrations of the radicals of one parent molecule, i.e. a molecule that is obtained by substituting the radical site with a hydrogen atom. A difference in radical concentration will lead to different decomposition pathways and thus another product outcome. In order to quantify these differences, a new kinetic model was created by removing all the H-shift reactions from the model of Genesys. Reactor simulations with this new model can be compared to the original Genesys kinetic model to assess the differences in radical and product concentrations.

As shown in section 5.4.2, n-heptane has two main decomposition pathways. First, homolytic C-C scissions lead to the net formation of radicals in the reactor and result in 1-alkyl radicals. Second, radicals can abstract a hydrogen atom from n-heptane leading to the four different heptyl

radicals. Figure 12 shows the concentrations of the latter, i.e. the concentrations of the 1-, 2-, 3- and 4-heptyl radicals using the models with and without H-shift reactions. The rate coefficients of the hydrogen abstractions on n-heptane depend on the position of the hydrogen that is abstracted, and the both models show that 1-heptyl is present in a lesser amount compared to 2-, and 3-heptyl, while 4-heptyl is the most important radical. However, when H-shift reactions are included in the model, the relative concentrations change significantly, mainly for the 1-, and 4-heptyl radicals, which was also concluded from the rate of production analysis. At 1473K, the consumption of the 1-heptyl radical to the 4-heptyl radical via a 1-4 H-shift reaction amounts for more than 25% of the total consumption of the 1-heptyl radical. 2% of the 1-heptyl radical reacts to the 3-heptyl radical via a 1-5 H-shift reaction. The 2-heptyl radical also isomerizes to the 3-heptyl radical (via a 1-4 H-shift reaction) and to the 4-heptyl radical (via a 1-3 H-shift reactions). These two reactions account for maximum 13% of the total consumption of the 2-heptyl radical. Furthermore, a small fraction of the 3-heptyl radical reacts to the 4-heptyl radical via a 1-2 H-shift reaction.

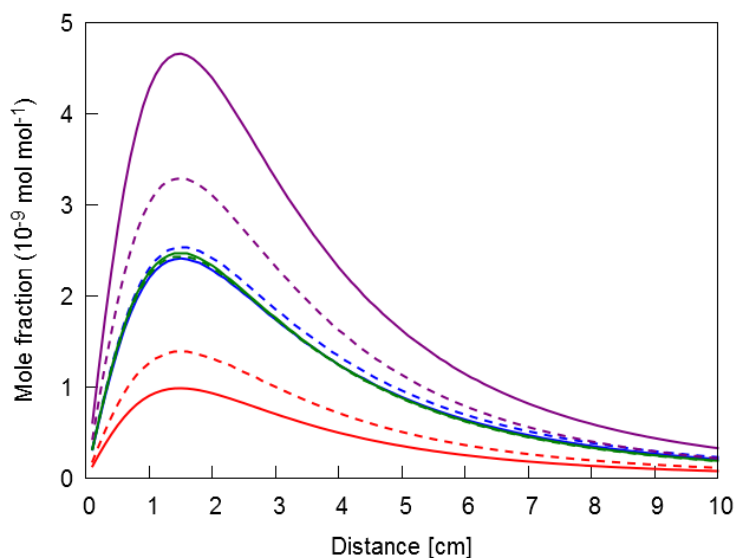


Figure 12: Mole fractions of the 1-heptyl (red), 2-heptyl (blue), 3-heptyl (green), and 4-heptyl (purple) radicals as a function of the distance to the inlet of the reactor for the model of Genesys including the H-shift reactions (full lines) and excluding the H-shift reactions (dashed lines). The reactor configuration and conditions are taken from the experimental data of Yuan et al.⁹ with a temperature of 1473K with a conversion of 0 to 12.8%. The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1\.

Besides hydrogen abstractions, the initial homolytic C-C scissions also lead to several large radicals such as 1-hexyl and 1-pentyl radicals. Under the studied conditions, 2-hexyl, 3-hexyl, 2-pentyl and 3-pentyl radicals are not formed in significant amounts without H-shift reactions. The

ratio of the concentrations of the three hexyl radicals using the model with H-shift reactions compared to the model without H-shift reactions is shown in Figure 13. The ratio of the concentrations of the 2-hexyl radical exceeds an order of magnitude and the ratio for the 3-hexyl radical is even larger, up to more than 5 orders of magnitude at the beginning of the reactor. These radicals are mainly formed from the 1-hexyl radical via a 1-4 and a 1-5 H-shift reaction respectively, accounting for more than 10% of the consumption of the latter, of which the ratio is thus slightly below 1. The absolute concentrations of the 2-, and 3-hexyl are an order of magnitude lower than the 1-hexyl radicals. Pentyl and butyl radicals show a similar trend compared to hexyl radicals.

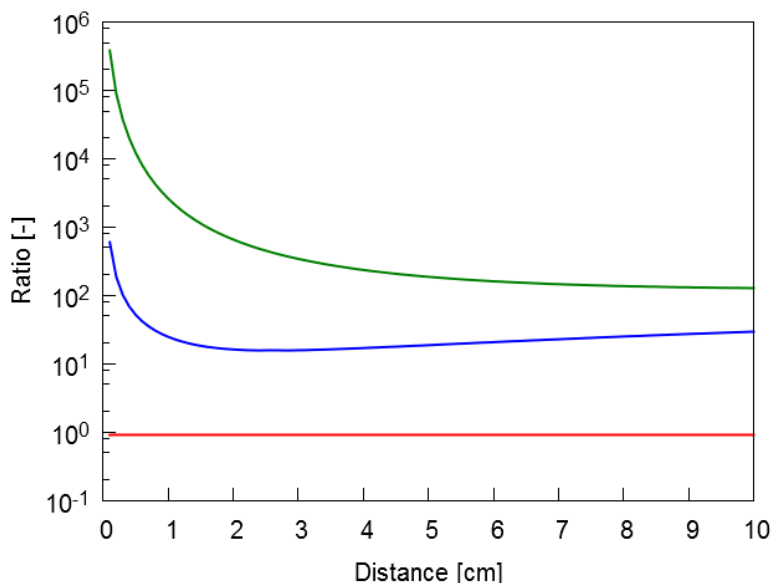


Figure 13: Ratio of the mole fractions of 1-hexyl (red), 2-hexyl (blue) and 3-hexyl (green) using the model including H-shift reactions compared to the model excluding H-shift reactions as a function of the distance to the inlet of the reactor. The reactor configuration and conditions are taken from the experimental data of Yuan et al.⁹ with a temperature of 1473K with a conversion of 0 to 12.8%. The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archieff\SoftLib\r\ruvdvijv\1\.

The concentrations of propyl radicals using both kinetic models are given in Figure 14, which shows that their concentrations are mostly independent of hydrogen shift reactions. At 1473K, the *i*-propyl radical is formed via two reactions. First, the reaction of an ethyl radical with a methyl radical yielding an *i*-propyl and a hydrogen radical is important in the beginning of the reactor. Second, the abstraction from a hydrogen atom from propane gains in importance throughout the reactor. The *n*-propyl radical is formed via homolytic C-C scission reactions and β -scission reactions. The 1-2 H-shift reaction for *n*-propyl to *i*-propyl or vice versa is not important.

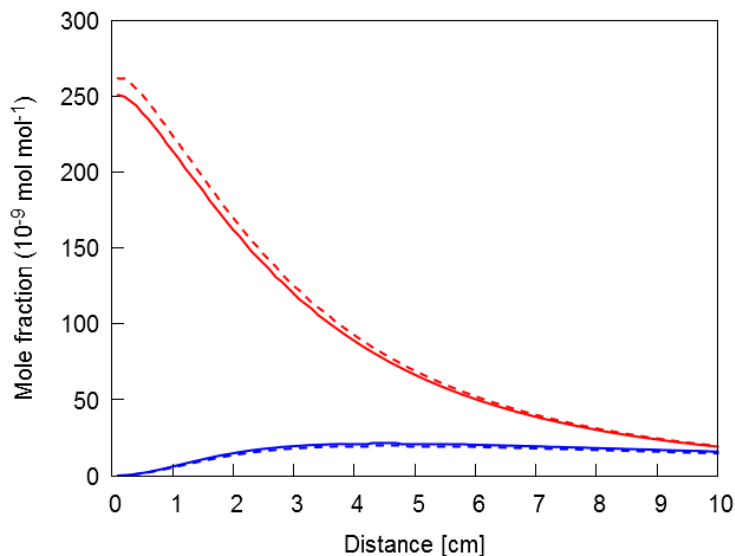


Figure 14: Mole fractions of the n-propyl (red) and iso-propyl (blue) radicals as a function of the distance to the inlet of the reactor for the model of Genesys including the H-shift reactions (full lines) and excluding the H-shift reactions (dashed lines). The reactor configuration and conditions are taken from the experimental data of Yuan et al.⁹ with a temperature of 1473K with a conversion of 0 to 12.8%. The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archie\SoftLib\r\ruvdv\1\.

Not only does the inclusion of H-shift reactions influence the radical concentrations, but the yields of the stable products are also dependent on these reactions, as can be seen in Figure 15 for ethene, propene, 1-butene, 2-butene, 1-pentene, 2-pentene, 1-hexene, 2-hexene, and 3-hexene. When including H-shift reactions, the concentrations of 1-alkyl radicals decreases, leading to a decrease in ethene formation, because 1-alkyl radicals are the main intermediates to form ethene. In contrary, propene can be formed from secondary radicals, from which the concentration increases as a result of H-shift reactions. The propene concentration is thus higher when including these reactions. 1-butene and 2-butene are both more formed in the case H-shift reactions are included in the kinetic model. Similarly, the concentration of the two pentene species is higher with taking H-shift reactions into account. The 1-hexyl radical concentration decreases due to the H-shift reactions, as can also be seen in Figure 13. This leads to a decrease in the 1-hexene concentrations while 2-hexene and 3-hexene are preferably formed compared to the kinetic model without H-shift reactions. The relative concentration of these compounds is also important for the formation of cyclic species. An important fraction of C6 ring structures originate from an intramolecular addition of an unsaturated 1-hexyl radical. As show in the rate of production analysis, the formation of 1,3- and 1,4-cyclohexadiene depends on the open chain

radical. Similarly, the formation of fulvene and cyclopentadiene depends strongly on these radicals.

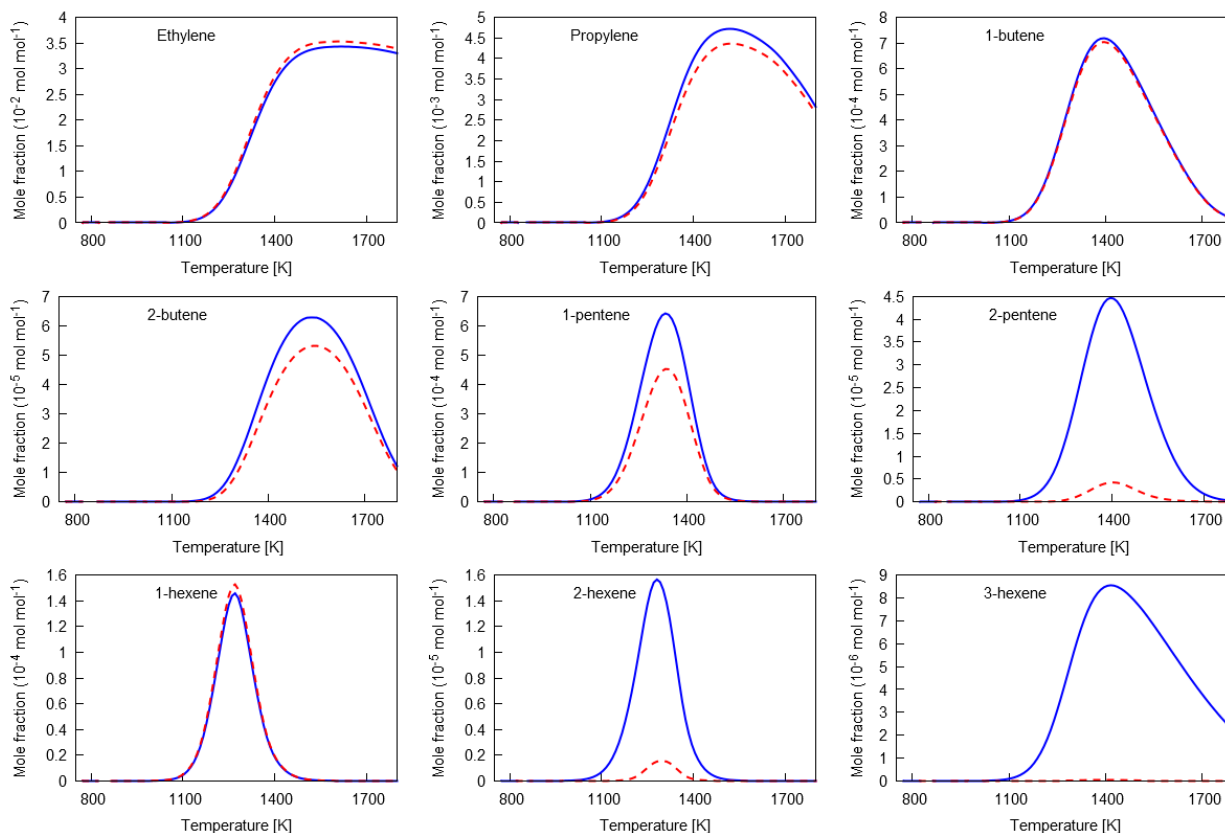


Figure 15: Mole fractions of ethene, propene, 1-butene, 2-butene, 1-pentene, 2-pentene, 1-hexene, 2-hexene, and 3-hexene as a function of the reactor temperature for the model of Genesys including the H-shift reactions (blue full lines) and excluding the H-shift reactions (red dashed lines). The reactor configuration and conditions are taken from the experimental data of Yuan et al.⁹. The simulations are performed using an ideal plug flow reactor model, see Eq. 5.8. The kinetics can be found on S:\vakgroep\ea12archie\SoftLib\r\ruvdvijv\1\.

5.5 Conclusions

A microkinetic model for the thermal decomposition and steam cracking of n-heptane has been automatically generated using the Genesys software. Most of the kinetic data originate from group additivity, of which the group additive values are based on high level quantum chemical calculations. The model has been extended to account for the chemistry of small compounds and for aromatics formation using literature data. The resulting model, containing 4221 reactions between 691 species has been reduced to remove redundant species and reactions, yielding a kinetic model with 154 species reacting in 1043 elementary steps.

By comparing the kinetic model to two independently measured experimental datasets, a wide applicability of the model can be observed over a broad range of pressure and temperature. The conversion and the formation of the most of the experimentally determined products were well described by the model, and only a few compounds showed a higher deviation between experiments and simulations, such as propyne, allene, n-butane, and H₂. Furthermore, a comparison was made with literature models for n-heptane combustion and pyrolysis. In general, the model of Genesys performs better than other models.

Via a rate of production analysis, the important pathways to consume n-heptane and to form the major products were outlined. Most pathways are very strongly dependent on the applied conditions. For example, for short residence times, low pressures and high temperatures, n-heptane consumption is driven by homolytic scissions whereas under milder conditions, hydrogen abstractions are more important.

Finally, the necessity of considering H-shift reactions was evaluated by comparing the original kinetic model with a kinetic model excluding this reaction family. Several important differences were observed, such as the lower 1-heptyl radical concentration and higher 4-heptyl radical concentration in the case H-shift reactions are included. Also, 1-hexyl and 1-pentyl radicals can react to other hexyl and pentyl radicals giving very large concentration differences. Differences in radical concentrations are also projected onto the product yields. Small changes can be seen for ethene and propene, but large deviations have been observed for C₅ and C₆ olefins. The inclusion of H-shift reactions is thus definitely necessary.

A few product yields were not well predicted by the kinetic model, such as H₂ and propyne. Improvements to the model can thus still be made, in which both the search for new reactions as well as the updating of the kinetic data can be suggested. Also, more detailed experimental data for a more thorough comparison of simulations to these measurements could lead to additional insights on shortcomings of the kinetic model.

The above conclusions show that, although reactions were taken from literature to account for the chemistry of small molecules and aromatics formation, Genesys is able to describe the thermal decomposition of n-heptane accurately.

5.6 References

1. Murata M, Saito S, Amano A, Maeda S. Prediction of Initial Product Distributions from Pyrolysis of Normal Paraffinic Hydrocarbons. *Journal of Chemical Engineering of Japan*. 1973;6(3):252-258.
2. Bajus M, Veselý V, Leclercq PA, Rijks JA. Steam Cracking of Hydrocarbons. 1. Pyrolysis of Heptane. *Industrial & Engineering Chemistry Product Research and Development*. 1979;18(1):30-37.
3. Aribike DS, Susu AA. Kinetics and mechanism of the thermal cracking of n-heptane. *Thermochimica Acta*. 1988;127:247-258.
4. Aribike DS, Susu AA. Mechanistic modeling of the pyrolysis of n-heptane. *Thermochimica Acta*. 1988;127:259-273.
5. Pant KK, Kunzru D. Pyrolysis of n-heptane: kinetics and modeling. *Journal of Analytical and Applied Pyrolysis*. 1996;36(2):103-120.
6. Held TJ, Marchese AJ, Dryer FL. A Semi-Empirical Reaction Mechanism for n-Heptane Oxidation and Pyrolysis. *Combustion Science and Technology*. 1997;123(1-6):107-146.
7. Chakraborty JP, Kunzru D. High pressure pyrolysis of n-heptane. *Journal of Analytical and Applied Pyrolysis*. 2009;86(1):44-52.
8. Garner S, Sivaramakrishnan R, Brezinsky K. The high-pressure pyrolysis of saturated and unsaturated C7 hydrocarbons. *Proceedings of the Combustion Institute*. 2009;32(1):461-467.
9. Yuan T, Zhang L, Zhou Z, Xie M, Ye L, Qi F. Pyrolysis of n-Heptane: Experimental and Theoretical Study. *Journal of Physical Chemistry A*. 2011;115(9):1593-1601.
10. Olahová N, Bajus M, Hájeková E, Šugár L, Markoš J. Kinetics and modelling of heptane steam-cracking. *Chemical Papers*. 2014;68(12):1678-1689.
11. Zámotný P, Karaba A, Olahová N, Petrů J, Patera J, Hájeková E, Bajus M, Bělohav Z. Generalized model of n-heptane pyrolysis and steam cracking kinetics based on automated reaction network generation. *Journal of Analytical and Applied Pyrolysis*. 2014;109:159-167.
12. Vandewiele NM, Van Geem KM, Reyniers MF, Marin GB. Genesys: Kinetic model construction using chemo-informatics. *Chemical Engineering Journal*. 2012;207:526-538.
13. Sabbe MK, Vandeputte AG, Reyniers MF, Waroquier M, Marin GB. Modeling the influence of resonance stabilization on the kinetics of hydrogen abstractions. *Physical Chemistry Chemical Physics*. 2010;12(6):1278-1298.
14. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Kinetic Modeling of α -Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Hydrogen Atoms. *Journal of Physical Chemistry A*. 2014;118(40):9296-9309.
15. Sabbe MK, Reyniers MF, Van Speybroeck V, Waroquier M, Marin GB. Carbon-centered radical addition and beta-scission reactions: Modeling of activation energies and pre-exponential factors. *Chemphyschem*. 2008;9(1):124-140.
16. Sabbe MK, Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M. Group additive values for the gas phase standard enthalpy of formation of hydrocarbons and hydrocarbon radicals. *Journal of Physical Chemistry A*. 2005;109(33):7466-7480.
17. Sabbe MK, De Vleeschouwer F, Reyniers MF, Waroquier M, Marin GB. First Principles Based Group Additive Values for the Gas Phase Standard Entropy and Heat Capacity of

- Hydrocarbons and Hydrocarbon Radicals. *Journal of Physical Chemistry A*. 2008;112(47):12235-12251.
18. Vandeputte AG, Sabbe MK, Reyniers M-F, Marin GB. Kinetics of alpha hydrogen abstractions from thiols, sulfides and thiocarbonyl compounds. *Physical Chemistry Chemical Physics*. 2012;14(37):12773-12793.
 19. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos N, Marin GB. Kinetic Modeling of alpha-Hydrogen Abstractions from Unsaturated and Saturated Oxygenate Compounds by Carbon-Centered Radicals. *Chemphyschem*. 2014;15(9):1849-1866.
 20. Paraskevas PD, Sabbe MK, Reyniers MF, Papayannakos NG, Marin GB. Group Additive Kinetics for Hydrogen Transfer Between Oxygenates. *Journal of Physical Chemistry A*. 2015;119(27):6961-6980.
 21. Paraskevas PD, Sabbe MK, Reyniers MF, Marin GB, Papayannakos NG. Group additive kinetic modeling for carbon-centered radical addition to oxygenates and -scission of oxygenates. *Aiche Journal*. 2016;62(3):802-814.
 22. Li Y, Zhou C-W, Somers KP, Zhang K, Curran HJ. The oxidation of 2-butene: A high pressure ignition delay, kinetic modeling study and reactivity comparison with isobutene and 1-butene. *Proceedings of the Combustion Institute*. 2017;36(1):403-411.
 23. Gao CW, Allen JW, Green WH, West RH. Reaction Mechanism Generator: Automatic construction of chemical kinetic mechanisms. *Computer Physics Communications*. 2016;203:212-225.
 24. Merchant S. *Molecules to Engines: Combustion chemistry of Alcohols and their application to advanced engines*, MIT; 2015, Ph. D. Thesis.
 25. Mohamed SY, Cai L, Khaled F, Banyon C, Wang Z, Al Rashidi MJ, Pitsch H, Curran HJ, Farooq A, Sarathy SM. Modeling Ignition of a Heptane Isomer: Improved Thermodynamics, Reaction Pathways, Kinetics, and Rate Rule Optimizations for 2-Methylhexane. *Journal of Physical Chemistry A*. 2016;120(14):2201-2217.
 26. *CHEMKIN Release 4.1.1* [computer program]. San Diego, CA, USA: Reaction Design, Inc.; 2007.
 27. Pepiot-Desjardins P, Pitsch H. An efficient error-propagation-based reduction method for large chemical kinetic mechanisms. *Combustion and Flame*. 2008;154(1-2):67-81.
 28. Herbinet O, Husson B, Serinyel Z, Cord M, Warth V, Fournet R, Glaude P-A, Sirjean B, Battin-Leclerc F, Wang Z, Xie M, Cheng Z, Qi F. Experimental and modeling investigation of the low-temperature oxidation of n-heptane. *Combustion and Flame*. 2012;159(12):3455-3471.
 29. Zhang K, Banyon C, Bugler J, Curran HJ, Rodriguez A, Herbinet O, Battin-Leclerc F, B'Chir C, Heufer KA. An updated experimental and kinetic modeling study of n-heptane oxidation. *Combustion and Flame*. 2016;172:116-135.
 30. Cai L, Pitsch H, Mohamed SY, Raman V, Bugler J, Curran H, Sarathy SM. Optimized reaction mechanism rate rules for ignition of normal alkanes. *Combustion and Flame*. 2016;173:468-482.

Chapter 6: Conclusions and future outlook

6.1 Conclusions

The current gradual shift from fossil to more renewable feedstocks in the (petro)chemical industry, in combination with more stringent environmental constraints and the push to higher energy efficiency are the key drivers for innovation in the 21st century. The chemical industry faces many challenges and proper estimations of the best scenarios need advanced, reliable, fast, predictive modelling tools and methods. Although in-silico process development is becoming important, still the combinations of accurate experimental measurements and theoretical calculations is considered the most reliable approach to come up with well-balanced answers for the future. This thesis aims at contributing to this sustainable development initiative by allowing an easier generation of accurate kinetic models, which are crucial to understand the underlying chemistry in a reactor, and in the end to do more with less. The approach developed in this work can also be used for innovation, design new solutions for many of today's chemical processes.

One of the main hurdles that needs to be overcome is the large data gap to build kinetic models. Data entails among others thermodynamic and kinetic parameters as well as experimental data used for validation. One of the key objectives of this work is identifying and resolving the lack of thermodynamic and kinetic data. This is primarily motivated by the fact that the current calculation procedures in kinetic model generation tools can suffer from a low applicability

domain or low accuracy, and many new quantum chemical calculations are necessary for the generation of predictive kinetic models applicable under divergent conditions. Two different tools to calculate new data have been employed in this work.

First, the KinBot software has been used to find reactions on the potential energy surface of the 1-pentanol radicals. KinBot is a computer code that automatically identifies stationary points on a potential energy surface using *ab initio* methods. Such potential energy searches are important to identify the main pathways of molecules and radicals, and can serve as high level calculations for a limited number of rate coefficients. The interest in 1-pentanol originates from its advantageous characteristics as substitute for diesel in combustion engines. 1-pentanol can be formed via biological processes and is thus a sustainable and clean solution compared to the fossil resources for diesel. To illustrate capabilities of KinBot, the pyrolysis of n-pentanol has been investigated, because the pyrolysis chemistry needs to be understood to model its combustion behavior, which allows in its turn optimizing fuel blends and combustion engines. Using KinBot, the six different radicals of n-pentanol were automatically found, including all their isomerization reactions, β -scission reactions and water elimination reactions. The β -scission of the β -radical, i.e. with the radical located in β position to the oxygen atom, leading to the hydroxyl radical and 1-pentene proved to be of special interest. When examining this reaction in the reverse direction, as the addition of the hydroxyl radical to 1-pentene, a barrierless step into a van der Waals-well is found. This is followed by an inner transition state leading to the β -radical of n-pentanol. An effective two-transition state model has been used to model this complex reaction. The resulting rate coefficients have been implemented in a kinetic model for the pyrolysis of n-pentanol built by Genesys, which showed good agreement to experimental data on n-pentanol pyrolysis.

Second, when building a kinetic model, most of the important pathways are already known and potential energy surfaces are thus of lesser importance. Instead, methods to automatically calculate a set of rate coefficients is of great interest for kinetic models, where hundreds of reactions belonging to the same reaction family can be encountered. In this work, it was chosen to rely on high level quantum chemical calculations to serve this purpose. However, these calculations still required extensive user knowledge and expertise. Therefore, several automation procedures have been developed to minimize human interaction and allow fast generation of thermodynamic properties and reaction rate coefficients. Although these calculations are still

several orders of magnitude slower than for example group additivity, they can serve as training data to build new, more accurate, and more complete group additivity methods or other calculation methods such as Evans-Polanyi or Blowers-Masel relationships. Starting from solely connectivity information of the atoms in a molecule, 3D coordinates of the species and transition states can be build, for which this work relies on “distance geometry”. The initial 3D coordinates are used for several energy minimizations using commercial quantum chemical software. The initial energy minimization is a rough pre-optimization using semi-empirical methods. Next, conformers can be exhaustively searched for. Algorithms to scan the conformation space and ensure finding a global energy minimum, or at least a very low minimum have been created. The lowest energy conformers can subsequently be used for high level calculations, optionally followed by 1D hindered rotor calculations. All these steps are automatically performed, without the need for any manual intervention. The results of the *ab initio* calculations are used to obtain thermodynamic parameters and kinetic data using statistical thermodynamics. The above algorithms have been used for a set of species and a set of reactions and expected results are found, i.e. the final data is comparable to *ab initio* data “manually” calculated by kineticists.

The automated calculation of a large set of intramolecular hydrogen abstraction reaction rate coefficients in hydrocarbon radicals proves the potential of the approach. A wide span of reactions is considered, including 1-2 up to 1-7 H-shifts. The *ab initio* calculations show good agreement to experimental data as well as to theoretical data calculated using a similar level of theory. From these calculations, it proved to be relatively straightforward to develop a group additivity scheme, allowing the calculations of rate coefficients for many new intramolecular hydrogen abstractions. The latter is very beneficial for automatic kinetic model generation. The group additivity scheme includes the influence of ligands on the attacking and attacked carbon atoms as well as the influence of the carbon chain bonded in between both reactive carbon atoms. The ligands vary from hydrogen atoms to sp^3 , sp^2 and sp carbon atoms. Each combination of these four ligands was considered. The group additivity scheme is able to reproduce most of the rate coefficients of reactions within a factor of 2 compared to the *ab initio* values.

To verify the necessity and accuracy of the group additivity scheme, a kinetic model for the pyrolysis of n-heptane has been built. All reactions significant for the thermal decomposition of n-heptane have been automatically generated using four reaction families: Intramolecular

hydrogen abstractions, intermolecular hydrogen abstractions, and inter- and intramolecular radical additions to double bonds and their reverse β -scission reactions. The initiation chemistry, the behavior of small molecules and the formation of larger secondary products such as aromatics were taken from literature. The kinetic model has been compared to two independently obtained experimental datasets for comparison. The model was able to describe most of the response variables accurately, under very divergent pressure, temperature and conversion ranges. Rate of production analysis revealed the important pathways and highlighted the large dependence of the importance of a reaction pathway to the reactor conditions such as pressure and temperature. This also allowed to verify the role of intramolecular hydrogen abstractions, which was found to be significant for the radical concentrations inside the reactor as well as for the products yields.

6.2 Future outlook

The methodologies developed in this work can be the starting point to develop automation procedures to calculate more accurate data in several ways. Firstly, although the use of the level of theory is hard-coded in the software, the extension to higher levels of theory only requires small changes. G4 calculations, for example, have proven to yield better results compared to CBS-QB3. Although these G4 calculations are slower, an automatic framework to calculate them could be beneficial since no human interactions would be necessary anymore, and the calculations can run in the background. Other improvements of the level of theory could include coupled cluster calculations, which would further increase the calculation time. Although full automation is not possible yet due to the complexity of the calculations, it can become possible within the next decade(s).

Secondly, improvements can be made on the approximations to obtain the thermodynamic and kinetic data. For example, 1D hindered rotor approximations are used to this point to treat internal rotors. Since many conformers are already at hand, these calculations could be replaced by multidimensional hindered rotor approximations. In the case where several rotors show strong coupling, multidimensional approximations are necessary to increase the accuracy of the results.

Thirdly, if many conformers have been calculated, using Boltzmann averaging for the calculation of thermodynamic parameters can further improve the results. Finally, the inclusion of master equation calculations can become important when comparing the results to low-pressure and high

temperature experiments. Up to now, the code can only calculate high-pressure limit rate coefficients, and the fall-off region of reactions is completely ignored. Calculating rate coefficients as a function of pressure and temperature could enlarge the application range of the results.

On the application side, although 1-pentanol pyrolysis has been studied and compared to experimental data, understanding its role in diesel engines will require additional simulations and experiments to (1) account for the combustion behavior of 1-pentanol and (2) include reactions between the n-pentanol combustion intermediates, and the ones from diesel combustion. New kinetic models are necessary to find the main pathways and search for optimal fuel blends and engine conditions.

On the group additive level, the developed model constructed for intramolecular hydrogen abstraction reactions only considers open-chain hydrocarbon radicals. However, these reactions also play a role in several other molecules. In cyclic structures, for example, the formation of aromatics depends strongly on several hydrogen shift pathways. The kinetics of these reactions will be strongly influenced by the cyclic structure in the reactant, and a bicyclic transition state will be formed. Extending the group additivity scheme to include cyclic structures could thus be valuable for the pyrolysis of cyclic reactants, or for aromatics formation.

For the modelling aspect, improvements are also still possible even for such a simple system as n-heptane pyrolysis. The n-heptane models found in literature, as well as the newly developed one, show both good agreements with some of the response variables, but also lesser agreements for others. The reaction pathways of all the models need to be verified and compared to one another, and their kinetics need further study to allow the generation of a widely applicable model able to describe the formation of all important products and intermediates, independently of the conditions. Additional experiments with varying conversion and dilution can be helpful to reach this goal and correctly validate the kinetic model.

Last but not least, one of the biggest challenges is related to the fact that most chemical processes are influenced by hetero-atomic species involving mainly oxygen, sulfur and nitrogen. These species can be present as impurities, e.g. originating from the fossil resources, or can be one of the reactants, which is often the case for biomass feedstocks or when using additives.

Intramolecular hydrogen abstraction rate coefficients will be strongly depending on the presence and the position of these hetero-atoms and extending the group additive model is required to account for this. For example, in the case of combustion, where oxygen is present in most of the reactive intermediates, an accurate description of the hydrogen shifts could help to understand the chemical behavior. For steam cracking, the presence of substantial amounts of hetero-atomic compounds, mainly containing oxygen, nitrogen and sulfur, poses many challenges to the safety, the operability and the quality of the product streams and an accurate understanding of the chemistry would be of great aid to overcome these challenges.

Appendices

Appendix A: Hindered rotor treatment

All the internal rotors have been treated as uncoupled rotors and relaxed scans have been done at the B3LYP/6-31+G level of theory. For a few rotors, the M06-2X/6-311++G(d,p) level of theory has also been used to assess the difference between the two calculations methods, which is visualized in Figure A-1. The rotors corresponding to the energy profiles are drawn above or in the graph; they include both rotors in transition states as well as radical species and an ethyl rotor, a methyl rotor and two hydroxyl rotors have been studied. The difference between the B3LYP energy profile and the M06 one is small, certainly around the equilibrium geometry of the species. This suggests that the influence of the level of theory for the hindered rotor profiles on the final rate coefficients is small.

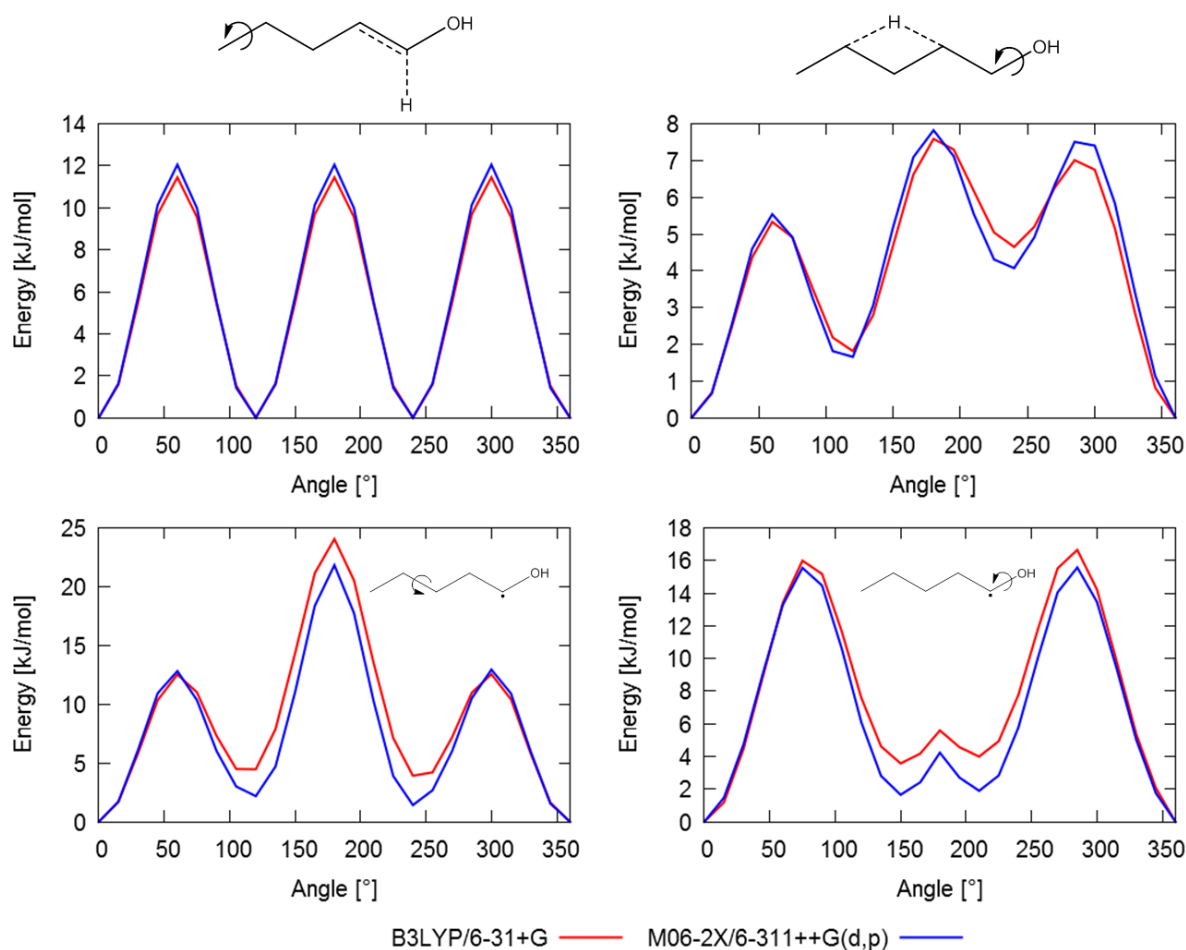


Figure A-1: Hindered rotor potentials for four rotational modes.

Appendix B: Correction factors for the OH + 1-pentene addition kinetics

Three corrections for the barrierless part of the OH + 1-pentene channel were included, one for the spin-orbit (SO) coupling of the OH radical on the state counts, one to account for the small basis set during the sampling, and one to estimate the effect of conformers on the capture rate coefficient.

B.1 SO coupling

The SO coupling of the OH radical (139.7 cm^{-1}) should be considered in the state counts of the reactant and of the outer transition state, assuming that the conserved modes are separable from the transitional ones. (At the inner transition state the SO coupling effectively just increases the inner barrier height by 0.8 kJ mol^{-1} , but that barrier needs to be fine-tuned anyway by more than this, see main text). The partition functions of the reactant OH and of the outer transition state directly at the canonical level was corrected.¹ This treatment is free of an ad hoc interpolation scheme,² and automatically preserves the correct value of the equilibrium constant, which in turn influences the dissociation rate coefficient of β -R. Based on Refs.^{2,3} the temperature-dependent correction factor is $c_{\text{SO}} = Q_{e,r}^{\text{coup}} / (Q_e Q_r)$, where $Q_{e,r}^{\text{coup}}$ is the true, coupled partition function of OH, while Q_e and Q_r are the uncoupled electronic and rotational partition functions, respectively. c_{SO} is given in Table B-1.

Table B-1: Canonical correction factor to account for the SO of OH (c_{SO}) and for the conformer dependency of the long-range attractive potential (c_{conf}). The outer TS of the OH + 1-pentene channel was corrected at the canonical level by the net $c_{\text{SO}} \times c_{\text{conf}}$ factor.

T (K)	c_{SO}	c_{conf}	$c_{\text{SO}} \times c_{\text{conf}}$
300	1.17	0.93	1.09
350	1.14	0.94	1.07
400	1.12	0.94	1.05
450	1.11	0.94	1.05
500	1.10	0.94	1.04
550	1.09	0.94	1.03
600	1.08	0.94	1.01
650	1.07	0.94	1.00
700	1.07	0.94	1.00
750	1.06	0.93	0.99
800	1.06	0.93	0.99
850	1.06	0.93	0.99
900	1.05	0.93	0.98
950	1.05	0.93	0.98
1000	1.05	0.93	0.97

B.2 Conformer-specific 1-D basis-set correction potentials

There are two rotors in 1-pentene that generate conformers, rotor A and rotor B as shown in Figure B-1. Five conformers were found, not accounting for the non-identical mirror images. The conformer group 1, 2, and 3 differs from the 4 and 5 group in the dihedral A. Within the two groups dihedral B is different. Conformers 1, 2, 3, and 5 have non-identical mirror images, while conformer 4 has C_s symmetry. The ZPE-inclusive energies of the conformers at the M06-2X/6-311++G** level is also shown in Figure B-1, with conformer 2 being the lowest one.

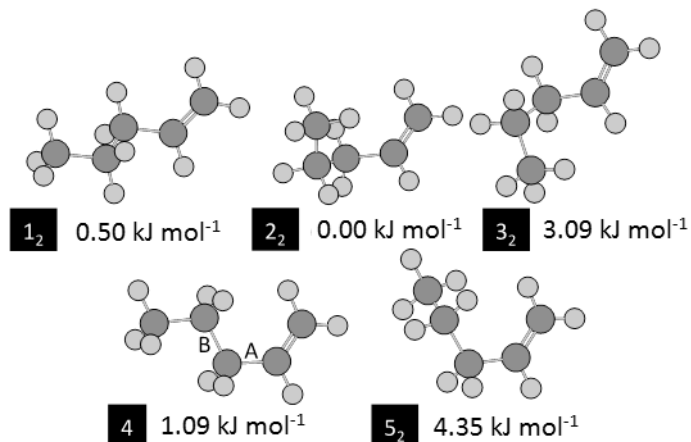


Figure B-1: The five conformers of 1-pentene. A and B in structure 4 label the two bonds that generate them, while the subscript '2' means that there is a non-identical mirror image that is not shown.

Sampling the interfragmental potential at the CASPT2(5e,4o)/aug-cc-pVDZ level was unfeasible for the VRC-TST calculations, and instead, a 1-D correction potential and sample at the CASPT2(5e,4o)/cc-pVDZ level of theory had to be used. To this end, minimum energy paths for all five distinguishable conformers starting from their respective van der Waals wells on both sides of the plane defined by the alkene functional group were calculated. The distance between the O atom and the inner sp^2 carbon was varied. The resulting potentials for the five conformers can be seen in Figure B-2a. The conformers' MEP deviate below ~ 4 Å, after which the lower-energy branches all generally had the same value, i.e., the MEP was independent of the conformation in this coordinate. During the VRC-TST sampling, the correction from the side that had a lower CASPT2(5e,4o)/aug-cc-pVDZ energy for a given conformer was used, because it was not always possible to decide the energetically dominant side. The corrections are, therefore, somewhat noisy, as shown in Figure B-2b.

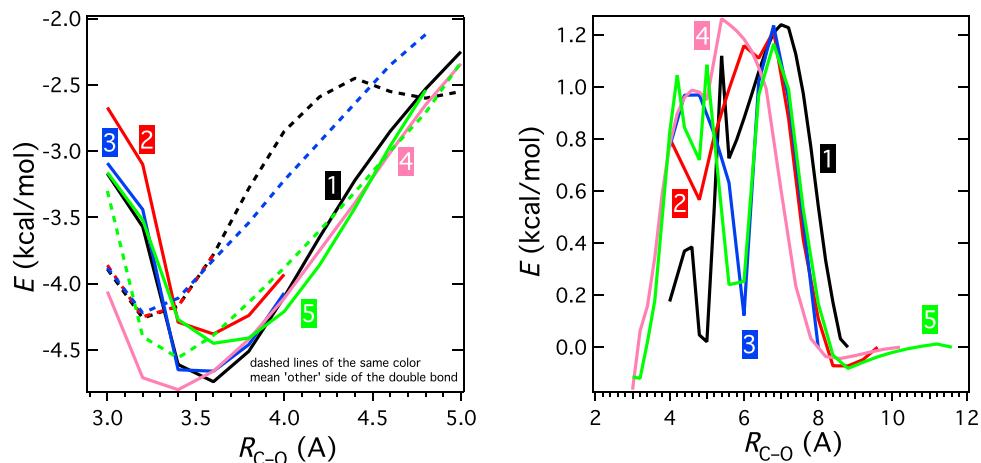


Figure B-2: (a) Interfragmental potential for the five conformers of 1-pentene. (b) Potential correction functions for the five conformers to go from CASPT2(5e,4o)/cc-pVDZ to CASPT2(5e,4o)/aug-cc-pVDZ level of theory.

During the sampling up to 14 bohr, the pivot point was placed on the middle point of the double bond and 0.5 bohr outside the O atom along the O–H bond, while in an overlapping 10–20 bohr separation the center-of-mass pivot points were used. The significant overlap between the two types of sampling ensures variational flexibility in these calculations and removes the arbitrariness of the switching. For each conformer ~4000 geometries were used that converged the results to 5%.

B.3 Presence of multiple conformers

In general, in the VRC-TST calculations the conserved and the transitional modes are treated as separable. The former ones are the internal vibrational modes of the fragments, while the latter ones are the ones that change from rotation to vibrations in the complex. The separability assumption implies that at the critical region of the barrierless potential the conserved internal modes are the same as the internal modes of the fragments at infinite separation. It has been previously shown² that this part of the approximation is valid for alkene + OH reactions. However, typically the transitional modes are only evaluated for the lowest energy conformer of the fragments, and it is usually not investigated how the interfragmental potential, and thus the flux via the barrierless potential changes as one chooses another conformer. One can imagine that having several conformers close in energy (such as in the case of pentene) the differences in the attractive potential (see Figure B-2a) can lead to non-negligible corrections in the capture rate coefficient as a function of temperature.

To investigate this, two test cases were used. In both, the hindered rotors of 1-pentene with the five separate conformers within the HO approximation was used, which is a relatively good approximation at low temperature where the outer TS plays a more significant role. In the first case it was assumed that the attractive potential is the same for all five conformers and it is that of the lowest energy conformer, #2. In the second case, for each channel, its own transitional mode

is used. Comparing the high-pressure rate coefficients, it was found that between 300 and 1000 K the assumed invariability of the interfragmental potential causes an overestimation of the barrierless rate coefficient by 15-5%, with the larger numbers being at the lower temperature. The overestimation is 15% at 100 K. The resulting correction factors c_{conf} are given in Table B-1 along with the combined canonical corrections of $c_{\text{SO}} \times c_{\text{conf}}$, which were in turn used in the calculations.

Appendix C: Coordinates embedding

This section explains the weighted coordinates embedding procedure. This method has proven to be robust. In an embed algorithm, the coordinates that are a best-fit to the estimated distances can be found fast and reliable by eigenvalue methods. The most straight-forward way to fit coordinates to distances is to minimize either the in algebra so-called STRESS shown in Eq. C.1 or the smoother SSTRESS shown in Eq. C.2. The STRESS function is a weighted error function, by defining the error as the difference between the actual distance between two atoms, and the distance from the distance matrix. The SSTRESS uses the squares of the distances instead of the distances themselves. In the equations below, $\mathbf{x}_1, \dots, \mathbf{x}_N$ are the coordinates, d_{ij}^2 the estimated squared distances, which will be renamed D_{ij} , and $w_{i,j}$ are weights larger than one.

$$\sum_{l=i<j}^{N,N} \left(w_{i,j} (\|\mathbf{x}_i - \mathbf{x}_j\| - d_{ij}) \right)^2 \quad \text{Eq. C.1}$$

$$\sum_{l=i<j}^{N,N} \left(w_{i,j} (\|\mathbf{x}_i - \mathbf{x}_j\|^2 - d_{ij}^2) \right)^2 \quad \text{Eq. C.2}$$

The STRESS and SSTRESS functions cannot be minimized fast and reliably. Therefore SSTRESS is expanded resulting in Eq. C.3.

$$4 \sum_{l=i<j}^{N,N} w_{i,j}^2 (\mathbf{x}_i \cdot \mathbf{x}_j) - \frac{1}{2} (\mathbf{x}_i^2 + \mathbf{x}_j^2 - D_{ij})^2 \quad \text{Eq. C.3}$$

Eq. C.4, in which m_j are the masses of the points and M the total mass of the species, is used to calculate the squared distances to the center of mass from the squared distances among the points. The distance between atom i and the center of mass is D_{0i} . This way the dependency of the estimate of the inner products of the coordinates can be removed.

$$D_{0i} = M^{-1} \sum_{j=1}^N m_j D_{ij} - M^{-2} \sum_{1=j<k}^{N,N} m_j m_k D_{jk} \quad \text{Eq. C.4}$$

It can be easily proven that the distances to the center of mass, i.e. center of mass coordinates, is exact if the estimated distances among the points are exact. Based on Eq. C.3 and Eq. C.4 and the restriction to weights of the form $w_i w_j$ the algebraic STRAIN is now introduced, c.f. Eq. C.5. The STRAIN is a new error function in which, as opposed to the STRESS and SSTRESS functions, the distances are related to the center of mass of the molecule instead of the direct distance between two atoms.

$$\frac{1}{2} \sum_{i,j=1}^{N,N} \left(w_i w_j (\mathbf{x}_i \cdot \mathbf{x}_j - a_{ij}) \right)^2 \quad \text{Eq. C.5}$$

$$a_{ij} := \frac{1}{2} (D_{0i} + D_{0j} - D_{ij}) \quad \text{Eq. C.6}$$

The problem is now reduced to minimize the STRAIN. By introducing the matrices \mathbf{A} , \mathbf{W} , \mathbf{X} , and \mathbf{D} , c.f. Eq. C.7 to Eq. C.10, the STRAIN can be expressed as a squared Frobenius norm, shown in Eq. C.11. It can be proven that \mathbf{A} and \mathbf{D} are related to each other by a so called two-sided projection (Eq. C.12). In which \mathbf{I} is the identity matrix. A similar relationship can be found for the matrices.

$$\mathbf{A} := [a_{ij}] \quad \text{Eq. C.7}$$

$$\mathbf{W} := \mathbf{diag}(w_1 \dots w_N) \quad \text{Eq. C.8}$$

$$\mathbf{X} := [\mathbf{x}_1 \dots \mathbf{x}_N]^T \quad \text{Eq. C.9}$$

$$\mathbf{D} := [D_{ij}] \quad \text{Eq. C.10}$$

$$F(\mathbf{X}) = \frac{1}{2} \|\mathbf{W}(\mathbf{X}\mathbf{X}^T - \mathbf{A})\mathbf{W}\|^2 \quad \text{Eq. C.11}$$

$$\mathbf{A} = -\frac{1}{2} (\mathbf{I} - \mathbf{1}\mathbf{m}^T/M) \mathbf{D} (\mathbf{I} - \mathbf{m}\mathbf{1}^T/M) \quad \text{Eq. C.12}$$

$$\mathbf{1} := [1, \dots, 1]^T \quad \text{Eq. C.13}$$

$$\mathbf{m} := [m_1, \dots, m_N]^T \quad \text{Eq. C.14}$$

The STRAIN can thus now be rewritten as Eq. C.15.

$$F(\mathbf{Y}) = \frac{1}{2} \|\mathbf{Y}\mathbf{Y}^T - \mathbf{B}\|^2 \quad \text{Eq. C.15}$$

$$\mathbf{Y} := \mathbf{W}\mathbf{X} \quad \text{Eq. C.16}$$

$$\mathbf{B} := \mathbf{W}\mathbf{A}\mathbf{W} \quad \text{Eq. C.17}$$

A necessary condition for the global minimum is that the gradient becomes zero.

$$\left[\frac{\partial F}{\partial y_{ij}} \right] = (\mathbf{Y}\mathbf{Y}^T - \mathbf{B})\mathbf{Y} = \mathbf{0} \quad \text{Eq. C.18}$$

$$\mathbf{B}\mathbf{Y} = \mathbf{Y}(\mathbf{Y}^T\mathbf{Y}) \quad \text{Eq. C.19}$$

$\mathbf{Y}^T\mathbf{Y}$ is a 3x3 matrix named the tensor. It can be assumed that the coordinates are rotated in space so that the inertial tensor is diagonal.

$$\mathbf{Y}^T\mathbf{Y} = \mathbf{diag}(\lambda_1, \lambda_2, \lambda_3) \quad \text{Eq. C.20}$$

The columns of \mathbf{Y} are proportional to eigenvectors of the scaled estimated Gram matrix \mathbf{B} and the moments of inertia $\lambda_1, \lambda_2, \lambda_3$ are the corresponding eigenvalues. Combining the previous two equations result in Eq. C.21.

$$\mathbf{Y}^T \mathbf{B} \mathbf{Y} = (\mathbf{Y}^T \mathbf{Y})^2 = (\text{diag}(\lambda_1, \lambda_2, \lambda_3))^2 \quad \text{Eq. C.21}$$

The squared Frobenius norm is expanded as Eq. C.22.

$$\begin{aligned} F(\mathbf{Y}) &= \text{tr}((\mathbf{Y}\mathbf{Y}^T - \mathbf{B})^2) \\ F(\mathbf{Y}) &= \text{tr}((\mathbf{B}^2 - 2\mathbf{Y}\mathbf{Y}^T\mathbf{B} + \mathbf{Y}\mathbf{Y}^T)^2) \\ F(\mathbf{Y}) &= \text{tr}(\mathbf{B}^2) - \text{tr}(2\mathbf{Y}^T\mathbf{B}\mathbf{Y} - (\mathbf{Y}^T\mathbf{Y})^2) \end{aligned} \quad \text{Eq. C.22}$$

By making use of the expression resulting from the global minimum equation we can rewrite the Frobenius norm, c.f. Eq. C.23.

$$F(\mathbf{Y}) = \text{tr}(\mathbf{B}^2) - \lambda_1^2 - \lambda_2^2 - \lambda_3^2 \quad \text{Eq. C.23}$$

The global minimum \mathbf{Y} is obtained by taking the three largest nonnegative eigenvalues of \mathbf{B} and scaling these by their square root. These are then transformed back to the original coordinates.

$$\mathbf{X} = \mathbf{W}^{-1}\mathbf{Y} \quad \text{Eq. C.24}$$

A variety of iterative methods can be used to find the three largest eigenvalues in an efficient way without fully diagonalizing the matrix.

Appendix D: Wilson B matrix

Since Gaussian only prints out the normal modes in Cartesian coordinates, and automatic selection of frequencies corresponding to torsional modes requires the normal mode to be expressed in internal coordinates, a transformation is necessary. Internal coordinates are, according to Wilson et al.⁴, coordinates that are unaffected by translation and rotation of the molecule in its entirety. A typical representation, the so-called Z-Matrix coordinates, of a molecule with N atoms only uses the 3N-6 independent coordinates. This includes (N-1) bond lengths, (N-2) bond angle and (N-3) dihedral angles. For analysis of a normal mode, defining more (redundant) coordinates can be helpful. In this case, the number of bond lengths equals the number of chemical bonds in the molecule and all the possible bond angles are considered. Dihedral angles can be described in several ways. For example, one could enumerate all the combinations of four consecutively bonded atoms. This can lead to a large number of unnecessary parameters. Instead, per chemical bond, the atoms on both sides are listed, and one reference atom on one side is chosen to define all the dihedral angles on the other side. Furthermore, because out-of-plane motions can strongly affect dihedral angles, but are not torsional motions, these normal modes can be filtered out by adding the out-of-plane angles of planar structures of four atoms.

The transformation from a displacement in Cartesian coordinates \mathbf{x} to a displacement in internal coordinates \mathbf{q} is straightforward by using Eq. D.1 in which B is the Wilson B matrix.

$$\delta \mathbf{q} = \mathbf{B} \delta \mathbf{x} \quad \text{Eq. D.1}$$

The Wilson B matrix only depends on the geometry of the molecule, which should be in its standard orientation, i.e. the principle axes of inertia need to be aligned with the x, y and z-axes. The vector \mathbf{q} is divided in four parts, bond stretching, angle bending, torsions and out-of-plane coordinates. Let S_t be one of the internal coordinates:

$$S_t = \sum_{i=1}^{3N} B_{ti} x_i \quad \text{Eq. D.2}$$

Although the form of Eq. D.2 is simple to obtain the internal coordinates, its use is less practical since Cartesian coordinates are expressed as a set of N vectors with 3 elements, rather than one 3N dimensional vector. To account for this, Eq. D.2 is rewritten as Eq. D.3 in which an atom is denoted as α and $\mathbf{s}_{t\alpha}$ contains the elements $3\alpha-2$ to 3α of the t^{th} row of the Wilson B matrix. $\mathbf{s}_{t\alpha}$ is the vector that, given all the atoms are in their equilibrium position, will give the highest increase in S_t .

$$S_t = \sum_{\alpha=1}^N \mathbf{s}_{t\alpha} \cdot \mathbf{x}_{\alpha} \quad \text{Eq. D.3}$$

For bond stretching internal coordinates, $\mathbf{s}_{t\alpha}$ can be easily be calculated. When considering the bond between atoms α and β , $\mathbf{s}_{t\alpha}$ equals the unit vector from β to α , i.e. $\mathbf{e}_{\beta\alpha}$, whereas $\mathbf{s}_{t\alpha}$ equals $\mathbf{e}_{\alpha\beta}$.

When three atoms α , β and γ form a bond angle between the bonds α - γ and β - γ , the vectors $\mathbf{s}_{t\alpha}$, $\mathbf{s}_{t\beta}$, $\mathbf{s}_{t\gamma}$ can be computed based on Eq. D.4, Eq. D.5, and Eq. D.6 with ϑ the equilibrium bond angle, $r_{\alpha\beta}$ is the distance between atoms α and β .

$$\mathbf{s}_{t\alpha} = \frac{\cos \vartheta \mathbf{e}_{\gamma\alpha} - \mathbf{e}_{\gamma\beta}}{r_{\gamma\alpha} \sin \vartheta} \quad \text{Eq. D.4}$$

$$\mathbf{s}_{t\beta} = \frac{\cos \vartheta \mathbf{e}_{\gamma\beta} - \mathbf{e}_{\gamma\alpha}}{r_{\gamma\beta} \sin \vartheta} \quad \text{Eq. D.5}$$

$$\mathbf{s}_{t\gamma} = \frac{[(r_{\gamma\alpha} - r_{\gamma\beta} \cos \vartheta) \mathbf{e}_{\gamma\alpha} + (r_{\gamma\beta} - r_{\gamma\alpha} \cos \vartheta) \mathbf{e}_{\gamma\beta}]}{r_{\gamma\alpha} r_{\gamma\beta} \sin \vartheta} \quad \text{Eq. D.6}$$

Out-of-plane internal coordinates can be obtained from the coordinates of four atoms: a central atom δ bonded to three other atoms, α , β and γ . θ is the angle between the vector $\mathbf{e}_{\delta\alpha}$ and the plane through β , γ and δ . ϕ_1 is the angle between the bonds β - δ and γ - δ .

$$\mathbf{s}_{t\alpha} = \frac{1}{r_{\delta\alpha}} \left(\frac{\mathbf{e}_{\delta\beta} \mathbf{e}_{\delta\gamma}}{\cos \theta \sin \phi_1} - \tan \theta \mathbf{e}_{\delta\alpha} \right) \quad \text{Eq. D.7}$$

$$\mathbf{s}_{t\beta} = \frac{1}{r_{\delta\beta}} \left(\frac{\mathbf{e}_{\delta\gamma} \mathbf{e}_{\delta\alpha}}{\cos \theta \sin \phi_1} - \frac{\tan \theta}{\sin^2 \phi_1} (\mathbf{e}_{\delta\beta} - \cos \phi_1 \mathbf{e}_{\delta\gamma}) \right) \quad \text{Eq. D.8}$$

$$\mathbf{s}_{t\gamma} = \frac{1}{r_{\delta\gamma}} \left(\frac{\mathbf{e}_{\delta\alpha} \mathbf{e}_{\delta\beta}}{\cos \theta \sin \phi_1} - \frac{\tan \theta}{\sin^2 \phi_1} (\mathbf{e}_{\delta\gamma} - \cos \phi_1 \mathbf{e}_{\delta\beta}) \right) \quad \text{Eq. D.9}$$

$$\mathbf{s}_{t\delta} = -\mathbf{s}_{t\alpha} - \mathbf{s}_{t\beta} - \mathbf{s}_{t\gamma} \quad \text{Eq. D.10}$$

Finally, four atoms that are subsequently bonded to each other define a dihedral angle, let the order of the atoms be α , β , γ and δ . ϕ_1 and ϕ_2 are the angle between the bonds α - β and β - γ and the bonds β - γ and γ - δ , respectively. τ is the dihedral angle. The expressions $(\alpha\delta)$ and $(\beta\gamma)$ mean that the subscripts α and δ and the subscripts β and γ need to be permuted, respectively.

$$\mathbf{s}_{t\alpha} = \frac{\mathbf{e}_{\alpha\beta} \mathbf{e}_{\beta\gamma}}{r_{\alpha\beta} \sin^2 \phi_1} \quad \text{Eq. D.11}$$

$$\mathbf{s}_{t\beta} = \frac{r_{\beta\gamma} - r_{\alpha\beta} \cos \phi_1}{r_{\beta\gamma} r_{\alpha\beta} \sin \phi_1} \frac{\mathbf{e}_{\alpha\beta} \mathbf{e}_{\beta\gamma}}{\sin \phi_1} + \frac{\cos \phi_2}{r_{\beta\gamma} \sin \phi_2} \frac{\mathbf{e}_{\delta\gamma} \mathbf{e}_{\gamma\beta}}{\sin \phi_2} \quad \text{Eq. D.12}$$

$$\mathbf{s}_{t\gamma} = [(\alpha\delta)(\beta\gamma)] \mathbf{s}_{t\beta} \quad \text{Eq. D.13}$$

$$\mathbf{s}_{t\delta} = [(\alpha\delta)(\beta\gamma)] \mathbf{s}_{t\alpha} \quad \text{Eq. D.14}$$

Appendix E: Test reactions

Table E-1: Arrhenius parameters for the test reactions for 1-2 H-shifts regressed between 800 and 1200 and the comparison of the initial group additive model to the *ab initio* rate coefficients at 1000K. The pre-exponential factors are expressed in s^{-1} and E_a is expressed in $kJ\ mol^{-1}$. For thermo-neutral reactions, i.e. when the reactant and product are identical, the forward and rate coefficients are the same, and is only tabulated once. In the other case, the first line corresponds to the forward rate coefficient, the second line to the reverse rate coefficient. All pre-exponential factors include tunneling and the number of single events.

Reaction	logA	Ea	k_{GA}/k_{AI}
	13.6	183.1	0.95
	13.6	168.0	1.04
	13.0	173.7	0.74
	13.7	181.9	0.62
	13.8	170.5	0.95
	13.1	149.3	0.83
	13.3	194.1	0.82
	13.7	183.1	0.74
	13.5	168.0	1.24
	13.0	141.7	0.94
	13.8	199.9	1.06
	12.9	173.6	0.87
	13.9	183.8	1.08
	13.3	156.5	0.94
	13.4	180.4	2.30
	13.5	166.4	1.06
	13.6	183.4	0.96
	13.3	173.6	0.82
	13.2	173.0	0.93
	13.0	170.3	0.56
	13.6	171.2	1.02
	13.4	185.5	0.77
	13.5	153.0	1.12
	13.9	185.6	1.44
	13.4	160.2	1.12
	13.3	137.9	0.28
	14.1	194.8	0.39

(Table E-1 continued)

	13.4	159.7	0.64
	13.6	175.3	2.02
	13.4	161.8	0.49
	13.5	173.4	0.41
	13.6	173.7	1.22
	12.8	163.2	1.53

Table E-2: Arrhenius parameters for the test reactions for 1-3 H-shifts regressed between 800 and 1200 and the comparison of the initial group additive model to the *ab initio* rate coefficients at 1000K. The pre-exponential factors are expressed in s^{-1} and E_a is expressed in $kJ\ mol^{-1}$. For thermo-neutral reactions, i.e. when the reactant and product are identical, the forward and rate coefficients are the same, and is only tabulated once. In the other case, the first line corresponds to the forward rate coefficient, the second line to the reverse rate coefficient. All pre-exponential factors include tunneling and the number of single events.

Reaction	logA	E_a	k_{GA}/k_{AI}
	12.5	172.6	1.00
	12.6	146.9	0.94
	13.6	206.1	0.94
	12.8	153.8	0.68
	13.0	195.2	0.61
	12.8	163.9	1.11
	13.0	180.8	1.03
	12.6	170.1	0.54
	13.4	173.5	0.32
	13.0	135.0	0.84
	12.7	207.7	0.97
	12.3	140.4	2.26
	13.4	214.2	3.31
	13.2	171.5	0.55
	12.6	150.2	0.87
	12.9	178.4	1.22
	13.0	166.1	0.75
	12.7	155.8	0.61
	13.0	173.1	0.96

(Table E-2 continued)

	12.6	127.8	1.19
	13.2	237.2	2.42
	13.6	169.0	0.53
	13.4	126.9	0.73
	14.3	231.8	0.64
	13.2	138.5	0.64
	13.4	201.4	1.36
	13.1	164.4	0.38
	12.3	134.9	2.57
	13.7	219.3	2.40
	12.9	137.1	1.39
	13.7	214.0	3.28
	11.8	128.3	3.06
	13.4	222.6	1.65

Table E-3: Arrhenius parameters for the test reactions for 1-4 H-shifts regressed between 800 and 1200 and the comparison of the initial group additive model to the *ab initio* rate coefficients at 1000K. The pre-exponential factors are expressed in s^{-1} and E_a is expressed in $kJ\ mol^{-1}$. For thermo-neutral reactions, i.e. when the reactant and product are identical, the forward and rate coefficients are the same, and is only tabulated once. In the other case, the first line corresponds to the forward rate coefficient, the second line to the reverse rate coefficient. All pre-exponential factors include tunneling and the number of single events.

Reaction	logA	Ea	k_{GA}/k_{AI}
	11.4	97.3	0.93
	11.3	99.4	0.86
	11.9	96.7	0.78
	11.8	94.9	0.66
	11.3	79.2	0.79
	11.5	120.6	0.95
	10.9	77.1	0.67
	12.1	135.1	0.93

(Table E-3 continued)

	11.8	92.4	0.57
	11.8	92.1	0.48
	11.4	92.2	1.47
	11.9	110.0	1.48
	10.7	92.8	4.81
	11.0	49.2	1.66
	12.5	170.5	4.75
	11.8	91.2	0.60
	12.1	89.6	0.80
	11.7	87.6	0.84
	11.8	87.5	0.71
	11.8	92.7	0.60
	11.4	54.6	0.52
	11.3	132.0	0.97
	11.1	55.3	0.45
	11.6	146.9	1.31
	11.4	63.7	0.78
	12.1	139.4	0.75
	12.0	98.9	0.32
	11.9	96.4	0.36
	11.1	59.6	0.49
	12.3	147.0	1.40
	11.8	91.2	0.47
	11.9	90.0	0.29
	12.0	76.8	0.21
	11.9	98.3	0.59

Table E-4: Arrhenius parameters for the test reactions for 1-5 H-shifts regressed between 800 and 1200 and the comparison of the initial group additive model to the *ab initio* rate coefficients at 1000K. The pre-exponential factors are expressed in s^{-1} and E_a is expressed in $kJ\ mol^{-1}$. For thermo-neutral reactions, i.e. when the reactant and product are identical, the forward and rate coefficients are the same, and is only tabulated once. In the other case, the first line corresponds to the forward rate coefficient, the second line to the reverse rate coefficient. All pre-exponential factors include tunneling and the number of single events.

Reaction	logA	Ea	k_{GA}/k_{AI}
	10.7	60.7	1.16
	10.8	77.6	0.77
	10.5	68.0	0.96
	10.9	64.3	1.11
	10.8	65.2	1.36
	10.5	63.4	0.68
	10.2	50.3	0.71
	10.4	89.8	1.04
	10.8	66.6	1.25
	10.8	69.5	1.24
	10.9	67.4	0.59
	10.8	63.7	0.76
	11.0	64.0	0.84
	11.1	67.2	0.85
	10.7	64.3	1.19
	10.8	64.1	1.02
	10.2	50.0	0.43
	11.2	106.6	0.66
	10.9	66.9	0.86
	10.2	57.0	1.69
	10.9	57.5	0.79
	11.3	62.8	0.80
	9.8	28.0	2.16
	11.3	149.0	1.19
	11.6	57.1	0.43
	11.6	62.4	0.33

(Table E-4 continued)

	11.2	52.3	0.36
	11.3	62.6	0.44
	10.9	49.5	1.69
	11.4	66.3	5.82
	10.2	27.0	1.68
	11.4	132.8	0.73
	11.0	59.7	1.31
	10.8	58.0	0.75
	10.9	67.9	0.52
	10.9	67.0	0.68
	10.3	23.1	0.56
	11.5	135.0	0.91
	10.2	36.3	1.52
	11.3	108.7	0.62
	10.0	31.4	1.11
	11.4	119.1	1.07

Table E-5: Arrhenius parameters for the test reactions for 1-6 H-shifts regressed between 800 and 1200 and the comparison of the initial group additive model to the *ab initio* rate coefficients at 1000K. The pre-exponential factors are expressed in s^{-1} and E_a is expressed in $kJ\ mol^{-1}$. For thermo-neutral reactions, i.e. when the reactant and product are identical, the forward and rate coefficients are the same, and is only tabulated once. In the other case, the first line corresponds to the forward rate coefficient, the second line to the reverse rate coefficient. All pre-exponential factors include tunneling and the number of single events.

Reaction	logA	Ea	k_{GA}/k_{AI}
	9.9	42.2	0.20
	9.8	97.3	0.33
	9.5	38.3	0.99
	10.5	118.6	1.63
	10.1	66.3	0.77
	9.9	58.7	0.70
	10.1	59.0	0.83
	10.1	68.3	0.87
	10.1	52.3	0.45
	10.5	58.7	0.79
	10.1	68.0	1.01
	10.3	64.6	0.54
	10.4	52.2	0.47
	10.6	67.6	0.33
	10.0	60.3	0.52
	10.0	55.5	0.44
	8.9	51.8	0.78
	10.6	64.3	0.49
	9.5	19.4	0.21
	10.1	127.3	0.64
	10.0	59.8	0.69
	9.9	66.5	1.01

(Table E-5 continued)

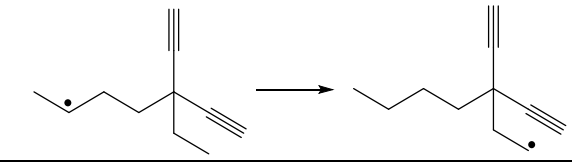
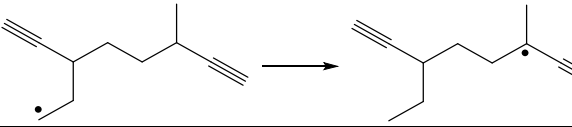
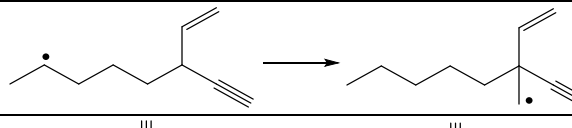
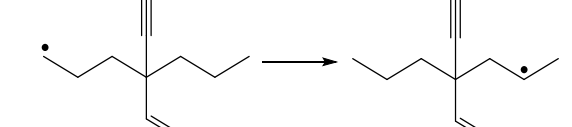
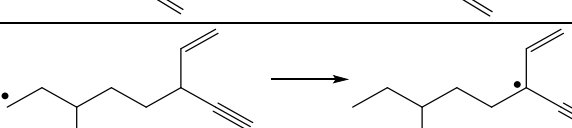
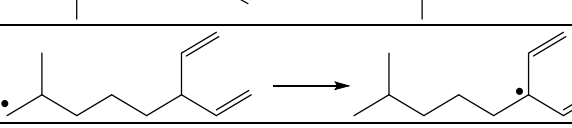
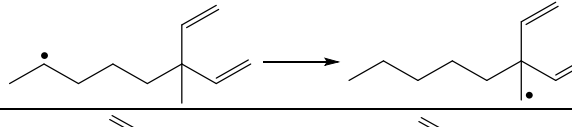
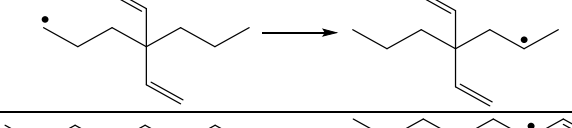
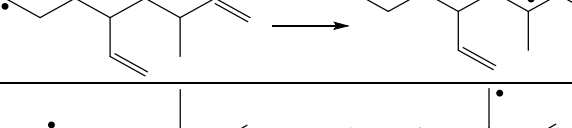
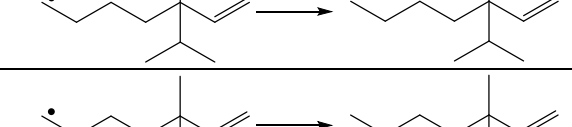
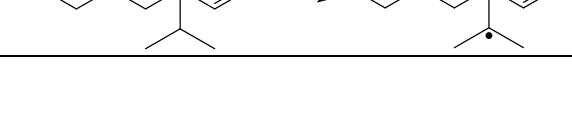




	10.3	72.9	0.92
	10.1	53.2	0.72
	9.4	38.1	1.41
	10.0	110.8	0.95
	9.9	71.7	1.51
	9.9	47.5	0.88
	10.5	51.8	0.34
	10.3	66.3	0.49
	9.4	23.5	0.43
	10.6	135.0	0.82
	8.6	25.3	2.15
	10.4	144.9	1.16
	9.8	68.3	1.04
	9.7	45.8	1.09
	10.7	50.5	0.38
	10.4	67.3	1.01
	9.3	36.2	0.37
	10.8	121.7	0.43
	10.1	64.3	0.59
	10.0	62.7	0.88
	9.0	47.0	1.99
	9.3	74.0	4.90

Table E-6: Arrhenius parameters for the test reactions for 1-7 H-shifts regressed between 800 and 1200 and the comparison of the initial group additive model to the *ab initio* rate coefficients at 1000K. The pre-exponential factors are expressed in s^{-1} and E_a is expressed in $kJ\ mol^{-1}$. For thermo-neutral reactions, i.e. when the reactant and product are identical, the forward and rate coefficients are the same, and is only tabulated once. In the other case, the first line corresponds to the forward rate coefficient, the second line to the reverse rate coefficient. All pre-exponential factors include tunneling and the number of single events.

Reaction	logA	E _a	k _{GA} /k _{AI}
	8.9	72.6	0.37
	9.2	70.6	0.34
	8.6	62.5	1.00
	8.4	78.7	3.00
	8.9	74.5	1.19
	8.7	61.8	0.53
	9.3	71.6	0.56
	8.0	42.7	0.72
	7.8	97.1	3.81
	9.4	79.5	0.58
	9.2	65.7	0.30
	7.9	47.0	0.33
	8.4	116.9	3.43

Appendix F: Initial group additive values

Table F-1: Primary group additive values ΔGAV° for 1-2 to 1-7 intramolecular hydrogen abstractions deduced from the training set.. The index i depends on the H-shift. For 1-2 H-shifts, i equals 2, for 1-3 H-shifts 3, etc. The units for $\Delta\text{GAV}^\circ_{\text{Ea}}$ values are kJ mol^{-1} . For the reference reaction, the single-event pre-exponential factors are expressed in s^{-1} and Ea is expressed in kJ mol^{-1} . The reference reaction pre-exponential factor does not include tunneling nor the number of single events.

	1-2		1-3		1-4		1-5		1-6		1-7	
	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea
Reference	12.78	177.3	12.04	176.2	10.82	106.0	9.94	73.1	9.12	69.3	7.89	76.0
$\text{C}_1\text{-(C)(H)}_2$	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0
$\text{C}_1\text{-(C)}_2\text{(H)}$	-0.09	6.3	0.08	5.6	0.03	4.6	-0.04	5.7	0.06	4.8	0.22	4.3
$\text{C}_1\text{-(C}_d\text{)(C)(H)}$	0.61	33.7	0.95	41.3	0.69	42.5	0.71	47.0	0.85	48.5	0.45	47.0
$\text{C}_1\text{-(C}_l\text{)(C)(H)}$	0.08	29.3	0.17	30.9	0.19	28.0	0.13	31.5	-0.50	33.2	0.14	31.0
$\text{C}_1\text{-(C)}_3$	-0.06	6.5	-0.03	6.3	0.04	5.1	-0.27	5.4	-0.28	5.7	-0.35	5.9
$\text{C}_1\text{-(C}_d\text{)(C)}_2$	0.55	37.9	0.85	45.0	0.68	46.1	0.47	50.2	0.50	57.9	0.36	54.6
$\text{C}_1\text{-(C}_l\text{)(C)}$	-0.04	34.4	-0.11	37.1	-0.10	34.5	-0.03	38.1	-0.28	35.6	0.17	41.2
$\text{C}_1\text{-(C}_d\text{)}_2\text{(C)}$	0.49	56.4	0.96	65.1	0.68	69.4	0.36	78.1	0.69	85.3	0.24	79.2
$\text{C}_1\text{-(C}_d\text{)(C}_l\text{)(C)}$	0.58	57.5	0.81	63.1	0.71	62.5	0.71	66.9	0.47	67.3	0.59	67.4
$\text{C}_1\text{-(C}_l\text{)}_2\text{(C)}$	0.13	56.4	0.13	59.2	0.24	55.9	0.21	60.4	-0.06	60.6	0.11	60.4
$\text{C}_i\text{-(C)(H)}_3$	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0
$\text{C}_i\text{-(C)}_2\text{(H)}_2$	0.05	-9.9	0.18	-9.8	0.13	-12.4	0.17	-10.3	-0.02	-12.3	0.00	-12.5
$\text{C}_i\text{-(C}_d\text{)(C)(H)}_2$	-0.15	-41.7	0.12	-32.2	-0.42	-31.1	-0.31	-26.5	-0.07	-24.9	-0.30	-26.6
$\text{C}_i\text{-(C}_l\text{)(C)(H)}_2$	-0.02	-32.6	0.19	-25.3	0.03	-28.8	0.01	-24.1	-0.48	-24.8	-0.03	-24.2
$\text{C}_i\text{-(C)}_3\text{(H)}$	0.05	-20.1	0.19	-18.7	0.21	-21.3	0.03	-21.0	-0.35	-21.5	0.01	-19.7
$\text{C}_i\text{-(C}_d\text{)(C)}_2\text{(H)}$	-0.28	-49.7	0.05	-38.4	-0.17	-39.2	-0.32	-34.4	-0.76	-32.1	-0.97	-31.4
$\text{C}_i\text{-(C}_l\text{)(C)(H)}$	-0.04	-41.3	0.18	-33.0	0.04	-37.3	0.05	-32.6	-0.29	-35.1	-0.38	-32.1
$\text{C}_i\text{-(C}_d\text{)}_2\text{(C)(H)}$	-0.54	-62.3	-0.01	-50.0	-0.15	-49.4	-0.15	-38.0	-0.30	-36.0	-0.80	-37.4
$\text{C}_i\text{-(C}_d\text{)(C}_l\text{)(C)(H)}$	-0.24	-62.7	0.10	-47.2	-0.02	-50.0	-0.07	-43.6	-0.50	-44.0	-0.53	-43.9
$\text{C}_i\text{-(C}_l\text{)}_2\text{(C)(H)}$	-0.07	-62.0	0.17	-46.1	0.10	-50.7	-0.01	-42.3	-0.51	-43.0	-0.04	-43.1

Table F-2: Secondary group additive values ΔGAV° for 1-2 to 1-7 intramolecular hydrogen abstractions deduced from the training set. The units for $\Delta\text{GAV}^\circ_{\text{Ea}}$ values are kJ mol^{-1} . For the reference reaction, the single-event pre-exponential factors are expressed in s^{-1} and E_a is expressed in kJ mol^{-1} .

	1-3		1-4				1-5					
	i=2		i=2		i=3		i=2		i=3		i=4	
	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log A$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a
$\text{C}_i\text{-(C)}_2\text{(H)}_2$	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0
$\text{C}_i\text{-(C)}_3\text{(H)}$	0.10	-1.7	-0.02	-5.6	0.02	-2.2	-0.09	-5.4	0.08	-5.8	-0.01	-2.2
$\text{C}_i\text{-(C}_d\text{)(C)}_2\text{(H)}$	-0.01	-3.3	-0.02	-6.7	0.01	-4.4	0.05	-5.1	0.18	-4.2	0.07	-2.8
$\text{C}_i\text{-(C}_l\text{)(C)}_2\text{(H)}$	-0.07	-5.7	-0.11	-8.2	0.09	-2.6	-0.07	-5.9	0.07	-1.9	-0.05	-0.5
$\text{C}_i\text{-(C)}_4$	0.17	-4.7	-0.03	-8.1	-0.02	-3.4	-0.01	-7.2	-0.09	-6.4	0.00	-3.0
$\text{C}_i\text{-(C}_d\text{)(C)}_3$	0.00	-6.3	-0.02	-10.7	-0.10	-5.4	0.12	-9.3	0.13	-7.9	0.15	-5.5
$\text{C}_i\text{-(C}_l\text{)(C)}_3$	0.01	-6.7	0.00	-9.1	0.06	-3.3	0.02	-10.1	0.06	-9.0	0.14	-4.8
$\text{C}_i\text{-(C}_d\text{)}_2\text{(C)}_2$	0.20	-8.8	-0.08	-10.8	-0.03	-2.1	0.17	-11.5	0.34	-10.7	0.18	-4.9
$\text{C}_i\text{-(C}_d\text{)(C}_l\text{)(C)}_2$	0.20	-8.2	0.02	-12.2	0.05	-4.8	0.13	-11.0	0.10	-4.8	0.15	-3.8
$\text{C}_i\text{-(C}_l\text{)}_2\text{(C)}_2$	-0.01	-8.9	0.00	-11.6	-0.02	-2.8	0.03	-8.7	-0.03	-4.0	0.17	0.7

	1-6								1-7									
	i=2		i=3		i=4		i=5		i=2		i=3		i=4		i=5		i=6	
	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a
$\text{C}_i\text{-(C)}_2\text{(H)}_2$	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00	0.0
$\text{C}_i\text{-(C)}_3\text{(H)}$	-0.06	-4.0	-0.04	-3.0	0.02	-3.4	-0.03	-1.3	0.04	-3.9	0.10	-2.3	0.35	-3.7	-0.13	-1.6	-0.20	-1.9
$\text{C}_i\text{-(C}_d\text{)(C)}_2\text{(H)}$	-0.17	-4.5	0.24	-0.8	0.11	-1.1	-0.09	-1.9	0.03	-4.7	0.03	-1.3	0.53	-2.0	0.10	-1.7	-0.40	-2.4
$\text{C}_i\text{-(C}_l\text{)(C)}_2\text{(H)}$	-0.28	-6.5	0.08	-1.8	-0.13	-1.7	-0.13	-0.3	-0.07	-4.5	0.12	-3.4	0.46	3.2	0.08	-2.0	0.06	2.1
$\text{C}_i\text{-(C)}_4$	-0.21	-6.1	0.05	-2.6	0.12	-3.8	-0.06	-0.9	0.19	-6.8	0.73	-1.3	0.06	4.1	0.15	-5.6	0.31	0.3
$\text{C}_i\text{-(C}_d\text{)(C)}_3$	-0.12	-8.1	0.30	-2.8	0.36	-3.8	-0.03	-3.1	0.36	-3.5	0.44	-6.4	1.25	-0.3	0.55	-6.9	0.05	0.8
$\text{C}_i\text{-(C}_l\text{)(C)}_3$	-0.15	-7.4	0.23	-6.6	0.19	-5.5	0.02	-1.6	0.15	-2.7	0.30	-5.2	0.84	1.5	0.22	-3.9	0.43	3.9
$\text{C}_i\text{-(C}_d\text{)}_2\text{(C)}_2$	-0.07	-10.6	0.38	-6.0	0.59	-4.8	-0.17	-4.1	0.24	-6.1	0.43	-10.2	1.08	-4.6	0.43	-9.7	0.67	-1.1
$\text{C}_i\text{-(C}_d\text{)(C}_l\text{)(C)}_2$	-0.12	-11.6	0.26	-2.3	0.34	-3.4	-0.05	-3.4	0.13	-5.7	0.22	-1.0	1.01	2.1	0.30	-0.9	-0.26	2.4
$\text{C}_i\text{-(C}_l\text{)}_2\text{(C)}_2$	-0.19	-8.9	0.13	-4.1	0.17	-1.0	-0.02	2.7	0.28	-3.0	0.24	-6.3	0.96	1.6	0.34	-1.6	0.41	7.5

Appendix G: Uncertainty on ΔGAV 's

Table G-1: Uncertainties on the primary group additive values defined by the 95% confidence intervals

	1-2		1-3		1-4		1-5		1-6		1-7	
	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea	$\log\tilde{A}$	Ea
$C_1-(C)_2(H)$	0.077	2.03	0.173	1.92	0.111	2.22	0.133	2.04	0.146	1.96	0.323	2.08
$C_1-(C_d)(C)(H)$	0.171	4.52	0.230	2.56	0.176	3.51	0.194	2.97	0.261	3.50	0.720	4.65
$C_1-(C_t)(C)(H)$	0.121	3.19	0.261	2.90	0.176	3.51	0.194	2.97	0.261	3.50	0.720	4.65
$C_1-(C)_3$	0.106	2.81	0.235	2.62	0.145	2.89	0.211	3.22	0.219	2.94	0.720	4.65
$C_1-(C_d)(C)_2$	0.141	3.71	0.313	3.48	0.175	3.50	0.193	2.95	0.271	3.63	0.534	3.45
$C_1-(C_t)(C)$	0.126	3.31	0.324	3.60	0.175	3.50	0.188	2.88	0.271	3.63	0.534	3.45
$C_1-(C_d)_2(C)$	0.185	4.88	0.313	3.48	0.182	3.63	0.211	3.22	0.262	3.52	0.720	4.65
$C_1-(C_d)(C_t)(C)$	0.171	4.52	0.324	3.60	0.193	3.85	0.188	2.88	0.262	3.52	0.720	4.65
$C_1-(C_t)_2(C)$	0.185	4.88	0.313	3.48	0.182	3.63	0.204	3.12	0.262	3.52	0.720	4.65
$C_1-(C)_2(H)_2$	0.077	2.03	0.173	1.92	0.111	2.22	0.133	2.04	0.146	1.96	0.323	2.08
$C_1-(C_d)(C)(H)_2$	0.171	4.52	0.230	2.56	0.176	3.51	0.194	2.97	0.261	3.50	0.720	4.65
$C_1-(C_t)(C)(H)_2$	0.121	3.19	0.261	2.90	0.176	3.51	0.194	2.97	0.261	3.50	0.720	4.65
$C_1-(C)_3(H)$	0.106	2.81	0.235	2.62	0.145	2.89	0.211	3.22	0.219	2.94	0.720	4.65
$C_1-(C_d)(C)_2(H)$	0.141	3.71	0.313	3.48	0.175	3.50	0.193	2.95	0.271	3.63	0.534	3.45
$C_1-(C_t)(C)(H)$	0.126	3.31	0.324	3.60	0.175	3.50	0.188	2.88	0.271	3.63	0.534	3.45
$C_1-(C_d)_2(C)(H)$	0.185	4.88	0.313	3.48	0.182	3.63	0.211	3.22	0.262	3.52	0.720	4.65
$C_1-(C_d)(C_t)(C)(H)$	0.171	4.52	0.324	3.60	0.193	3.85	0.188	2.88	0.262	3.52	0.720	4.65
$C_1-(C_t)_2(C)(H)$	0.185	4.88	0.313	3.48	0.182	3.63	0.204	3.12	0.262	3.52	0.720	4.65

Table G-2: Uncertainties on the secondary group additive values defined by the 95% confidence intervals

	1-3		1-4			
	i=2		i=2		i=3	
	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log A$	E_a
$C_i-(C)_3(H)$	0.223	2.477	0.105	2.099	0.105	2.099
$C_i-(C_d)(C)_2(H)$	0.290	3.219	0.144	2.872	0.144	2.872
$C_i-(C_l)(C)_2(H)$	0.256	2.848	0.175	3.495	0.175	3.495
$C_i-(C)_4$	0.236	2.628	0.193	3.850	0.193	3.850
$C_i-(C_d)(C)_3$	0.236	2.628	0.193	3.850	0.193	3.850
$C_i-(C_l)(C)_3$	0.211	2.349	0.193	3.850	0.193	3.850
$C_i-(C_d)_2(C)_2$	0.270	2.999	0.193	3.850	0.193	3.850
$C_i-(C_d)(C_l)(C)_2$	0.290	3.219	0.193	3.850	0.193	3.850
$C_i-(C_l)_2(C)_2$	0.264	2.934	0.193	3.850	0.193	3.850

	1-5					
	i=2		i=3		i=4	
	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a
$C_i-(C)_3(H)$	0.094	1.44	0.126	1.93	0.094	1.44
$C_i-(C_d)(C)_2(H)$	0.188	2.88	0.155	2.37	0.188	2.88
$C_i-(C_l)(C)_2(H)$	0.188	2.88	0.155	2.37	0.188	2.88
$C_i-(C)_4$	0.211	3.22	0.182	2.79	0.211	3.22
$C_i-(C_d)(C)_3$	0.204	3.12	0.182	2.79	0.204	3.12
$C_i-(C_l)(C)_3$	0.204	3.12	0.182	2.79	0.204	3.12
$C_i-(C_d)_2(C)_2$	0.204	3.12	0.182	2.79	0.204	3.12
$C_i-(C_d)(C_l)(C)_2$	0.204	3.12	0.182	2.79	0.204	3.12
$C_i-(C_l)_2(C)_2$	0.204	3.12	0.182	2.79	0.204	3.12

	1-6							
	i=2		i=3		i=4		i=5	
	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a	$\log\tilde{A}$	E_a
$C_i-(C)_3(H)$	0.235	3.14	0.235	3.14	0.235	3.14	0.235	3.14
$C_i-(C_d)(C)_2(H)$	0.228	3.06	0.271	3.63	0.271	3.63	0.228	3.06
$C_i-(C_l)(C)_2(H)$	0.228	3.06	0.271	3.63	0.271	3.63	0.228	3.06
$C_i-(C)_4$	0.271	3.63	0.224	3.00	0.224	3.00	0.271	3.63
$C_i-(C_d)(C)_3$	0.205	2.75	0.246	3.29	0.246	3.29	0.205	2.75
$C_i-(C_l)(C)_3$	0.198	2.65	0.262	3.52	0.262	3.52	0.198	2.65
$C_i-(C_d)_2(C)_2$	0.246	3.29	0.246	3.29	0.246	3.29	0.246	3.29
$C_i-(C_d)(C_l)(C)_2$	0.246	3.29	0.246	3.29	0.246	3.29	0.246	3.29
$C_i-(C_l)_2(C)_2$	0.246	3.29	0.246	3.29	0.246	3.29	0.246	3.29

(Table G-2 continued)

	i=2		i=3		1-7 i=4		i=5		i=6	
	$\log \tilde{A}$	E_a	$\log \tilde{A}$	E_a	$\log \tilde{A}$	E_a	$\log \tilde{A}$	E_a	$\log \tilde{A}$	E_a
$C_1-(C)_3(H)$	0.429	2.77	0.534	3.45	0.720	4.65	0.534	3.45	0.429	2.77
$C_1-(C_d)(C)_2(H)$	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65
$C_1-(C_i)(C)_2(H)$	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65
$C_1-(C)_4$	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65
$C_1-(C_d)(C)_3$	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65
$C_1-(C_i)(C)_3$	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65
$C_1-(C_d)_2(C)_2$	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65
$C_1-(C_d)(C_i)(C)_2$	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65
$C_1-(C_i)_2(C)_2$	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65	0.720	4.65

Appendix H: Publications

H.1 A1 publications

1. Van de Vijver, R.; Vandewiele, N. M.; Bhoorasingh, P. L.; Slakman, B. L.; Seyedzadeh Khanshan, F.; Carstensen, H.-H.; Reyniers, M.-F.; Marin, G. B.; West, R. H.; Van Geem, K. M., Automatic Mechanism and Kinetic Model Generation for Gas- and Solution-Phase Processes: A Perspective on Best Practices, Recent Advances, and Future Challenges. *International Journal of Chemical Kinetics* **2015**, 47, (4), 199-231.
2. Van de Vijver, R.; Vandewiele, N. M.; Vandeputte, A. G.; Van Geem, K. M.; Reyniers, M.-F.; Green, W. H.; Marin, G. B., Rule-based ab initio kinetic model for alkyl sulfide pyrolysis. *Chemical Engineering Journal* **2015**, 278, 385-393.
3. Vandewiele, N. M.; Van de Vijver, R.; Van Geem, K. M.; Reyniers, M.-F.; Marin, G. B., Symmetry calculation for molecules and transition states. *Journal of Computational Chemistry* **2015**, 36, (3), 181-192.
4. Van de Vijver, R.; Devocht, B. R.; Van Geem, K. M.; Thybaut, J. W.; Marin, G. B., Challenges and opportunities for molecule-based management of chemical processes. *Current Opinion in Chemical Engineering* **2016**, 13, 142-149.
5. Vandewiele, N. M.; Van de Vijver, R.; Van Geem, K. M.; Reyniers, M.-F.; Marin, G. B., Implementation of Stereochemistry in Automatic Kinetic Model Generation. *International Journal of Chemical Kinetics* **2016**, 48, (12), 755-769.
6. Olahova, N.; Djokic, M. R.; Van de Vijver, R.; Ristic, N. D.; Marin, G. B.; Reyniers, M.-F.; Van Geem, K. M., Thermal Decomposition of Sulfur Compounds and their Role in Coke Formation during Steam Cracking of Heptane. *Chemical Engineering & Technology* **2016**, 39, (11), 2096-2106.

H.2 C1 publications

1. Vermeire, F.; Paraskevas, P.; Van de Vijver, R.; De Bruycker, R.; Sabbe, M. K.; Reyniers, M.-F.; Papayannakos, N.; Marin, G. B.; Van Geem, K. M., First Principles Based Microkinetic Modelling of Methyl Butanoate Pyrolysis. *Chemical Research in Flanders*, Blankenberge, Belgium, **2016**.
2. Vandewiele, N. M.; Van de Vijver, R.; Van Geem, K. M.; Reyniers, M.-F.; Marin, G. B., Implementation of stereochemistry in automatic kinetic model generation. *International Conference on Chemical Kinetics*, Ghent, Belgium, **2015**.
3. Khandavilli, M.; Van Geem, K. M.; Van de Vijver, R.; Marin, G. B., A new group additive kinetic model for pyrolysis of C2-C4 hydrocarbons. *International Conference on Chemical Kinetics*, Ghent, Belgium, **2015**.

4. Olahova, N.; Djokic M. R.; Sarris, S. A.; Van de Vijver, R.; Ristic, N. D.; Marin, G. B.; Reyniers, M.-F.; Van Geem, K. M.; Marin, G.B., Understanding the decomposition of sulfur compounds during steam cracking. *International Conference on Chemical Kinetics*, Ghent, Belgium, **2015**.
5. Van de Vijver, R.; Vandewiele, N. M.; Vandeputte, A. G.; Van Geem, K. M.; Reyniers, M.-F.; Green, W. H.; Marin, G. B., Rule based ab initio kinetic model for alkyl sulfide pyrolysis. *23rd International Symposium on Chemical Reaction Engineering*, Bangkok, Thailand, **2014**.

References

1. Antonov IO, Kwok J, Zádor J, Sheps L. OH + 2-butene: A combined experimental and theoretical study in the 300-800 K temperature range. *Journal of Physical Chemistry A*. 2015;119:7742-7752.
2. Zádor J, Jasper AW, Miller JA. The reaction between propene and hydroxyl. *Physical Chemistry Chemical Physics*. 2009;11(46):11040-11053.
3. Herzberg G. *Molecular Spectra and Molecular Structure. I. Spectra of Diatomic Molecules* New York: Van Nostrand Reinhold; 1950.
4. Wilson EB, Decius JC, Cross PC. *Molecular vibrations : the theory of infrared and Raman vibrational spectra*. New York: McGraw-Hill; 1955.